Introduction to

Many Body physics

Thierry Giamarchi
Anibal Iucci
Christophe Berthod

Lecture notes

University of Geneva, 2008–2013

Contents

Some bare numbers and unsettling questions

0.1 Goal of the game

Condensed matter physics is a remarkable domain where the effects of quantum mechanics combine with the presence of a very large ($\sim 10^{23}$) coupled degrees of freedom. The interplay between these two ingredients leads to the richness of phenomena that we observe in everyday's materials and which have led to things such useful in our daily life as semiconductors (transistors, computers !), liquid crystals (flat panel displays), superconductivity (cables and magnets used in today's MRI equipments) and more recently giant magnetoresistance (hard drive reading heads).

When looking at this type of problems, one important question is how should we model them. In particular one essential question that one has to address is whether the interaction among the particles is an important ingredient to take into account or not. The answer to that question is not innocent. If the answer is no, then all is well (but perhaps a little bit boring) since all we have to do is to solve the problem of one particle, and then we are done. This does not mean that all trivial effects disappear since fermions being indistinguishable have to obey the Pauli principle. But it means at least that the heart of the problem is a single particle problem and excitations. This is what is routinely done in all beginner's solid state physics course, where all the properties of independent electrons are computed.

If the answer to the above questions is no, then we have a formidable problem, where all degrees of freedom in the system are coupled. Solving a Schroedinger equation with 10^{23} variables is completely out of the question, so one should develop tools to be able to tackle such a problem with some chance of success.

What is the appropriate situation in most materials is thus something of importance, and one should address in turn the following points

1. Are the quantum effects important in a solid at the one particle level. Here there is no surprise, the answer for most metals is yes, given the ratio of the typical energy scale in a solid ($\sim 10000K$) due to the Pauli principle, compared to the standard thermal energy scale ($\sim 300K$)

2. From a purely empirical point of view, does the independent electron picture works to explain the properties of many solids.

3. From a more theoretical point of view can one estimate the ratio of the interaction energy (essentially the Coulomb interaction in a solid) to the kinetic energy and work out the consequences of such interactions.

The answer to the first question is without surprise, and can be found in all standard textbooks on solid state physics. The answer to the second question is much more surprising, since in practice the free electron theory works extremely well to describe most of the solids. When one is faced with such a fact the standard reaction is to think that the interactions are indeed negligible in most solid. Unfortunately (or rather quite fortunately), this naive interpretation of the data does not corresponds with a naive estimate of the value of the interactions. One is thus faced with the formidable puzzle to have to treat the interactions, and also to understand why, by some miracle they seem to magically disappear in the physical observables. The miracle is in fact called Fermi liquid theory and was discovered by L. D. Landau, and we will try to understand and explain the main features of this theory in these lectures.

The first part of these lectures will thus be devoted to set up the technology to deal with systems made of a very large number of interacting quantum particles (the so-called many body physics). We will use this technology to understand the theory of Fermi liquids.

In the second part we will see cases where the Fermi liquid theory actually fails, and where interaction effects leads to drastically new physics compared to the non interacting case. This is what goes under the name of non-Fermi liquids or strongly correlated systems.

0.2 Bibliography

The material discussed in these notes can be in part found in several books. Here is a partial bibliographical list:

- Basics of solid state physics [Zim72, Kit88, AM76]
- Many body physics: the techniques of many body are well explained in the book [Mah81] (which is going beyond the techniques we will see in this course) The course will mostly follow the notations of this book.
- Fermi liquid theory: The first chapters of [Noz61]. At a more advanced level [PN66] or [AGD63].
- A good (yet to be) “book” concerning the topics discussed in this course can be found on line at

  `http://www.physics.rutgers.edu/~coleman/mbody/pdf/bk.pdf`

0.3 Disclaimer

Warning !!!!!!

These notes are in progress and still contains several bugs, and should not be treated as some kind of sacro saint text. So don’t buy blindly what is in it, and in case of doubt don’t hesitate to recheck and correct the calculations using your own notes. And of course do not hesitate to ask if needed. This is certainly the best way to really learn the material described in these lectures.

Part I

Electronic properties of metals: Fermi liquids

CHAPTER 1

Basics of basic solid state physics

The goal of this chapter is to review the salient features of noninteracting electrons. This will useful in order to determine whether the interactions lead or not to drastic changes in the physics. We will also estimate the order of magnitude of the interactions in a normal metal, starting from the Coulomb interaction and recall the main differences between Coulomb interactions in the vacuum and in a solid.

1.1 Non interacting electrons

Most of the material in this chapter is classical knowledge of solid state physics [Zim72, Kit88, AM76]. We will however use as soon as possible the proper technology to perform the calculations.

1.1.1 Free electrons

The simple case of free electrons allows us to introduce most of the quantities we will use. Let us consider free electrons described by the Hamiltonian

$$H = \frac{P^2}{2m} \tag{1.1}$$

The eigenstates are the plane waves $|\boldsymbol{k}\rangle$, defined by

$$\langle \boldsymbol{r}|\boldsymbol{k}\rangle = \frac{1}{\sqrt{\Omega}} e^{i\boldsymbol{k}\boldsymbol{r}} \tag{1.2}$$

where Ω is the volume of the system. The corresponding eigenvalue is

$$\varepsilon_{\boldsymbol{k}} = \frac{\hbar^2 k^2}{2m} \tag{1.3}$$

In addition to the above orbital part, the electron possesses a spin 1/2. A complete basis of the spin degrees of freedom is provided by the two eigenstates of one of the spin component. One usually takes the S_z one, and we define the corresponding basis as $|\uparrow\rangle$ and $|\downarrow\rangle$ The ensemble α of quantum numbers needed to fully characterize the electrons is thus its momentum and its spin $\alpha = (\sigma, \boldsymbol{k})$.

For a system of finite size the values of $\boldsymbol{k}$ are quantized by the boundary conditions. In the limit of a very large size the precise boundary condition does not matter so we will take periodic boundary conditions for simplicity. This means that for a system of linear dimensions L (the volume being $\Omega = L^d$ for a system in d dimensions), the wavefunction ψ must satisfy

$\psi(x+L) = \psi(x)$ and similar relations in all directions. This imposes that each component of $\boldsymbol{k}$ is of the form

$$k_l = \frac{2\pi m_l}{L} \tag{1.4}$$

where the m_l are integers for $l = 1, \dots, d$ with d the dimension of the system.

At zero temperature the Pauli principle states that each quantum state is occupied by at most one fermion. One thus starts to fill the lowest energy levels to accommodate the $\mathcal{N}$ electrons of the system. One thus fills the energy level up to the Fermi energy $E_{\rm F}$ and up to a momentum $k_{\rm F}$ such that $\varepsilon_{k_{\rm F}} = E_{\rm F}$. At finite temperature, the states are occupied with a probability that is given by the Fermi-Dirac factor

$$f_{\rm F}(\varepsilon) = \frac{1}{e^{\beta(\varepsilon-\mu)}+1} \tag{1.5}$$

where μ is the chemical potential. The total number of electrons in the system is given by

$$\mathcal{N} = \sum_{\boldsymbol{k}\sigma} f_{\rm F}(\varepsilon_{\boldsymbol{k}}) \tag{1.6}$$

The sum over the integers can be simplified in the large L limit since the values of k_i are nearly continuous. Using (1.4) one gets

$$\sum_{\boldsymbol{k}} \to \frac{\Omega}{(2\pi)^d} \int d\boldsymbol{k} \tag{1.7}$$

One has thus (the sum over the spin degrees of freedom simply giving a factor of two)

$$\mathcal{N} = \frac{\Omega}{(2\pi)^3} \frac{4\pi}{3} k_{\rm F}^3 \tag{1.8}$$

one can thus introduce the density of particles $n = \mathcal{N}/\Omega$ and $n = k_{\rm F}^3/(6\pi^2)$.

The existence of a Fermi level is of prime importance for the properties of solids. Let us put some numbers on the above formulas. Some numbers for the Fermi energy and related quantities will be worked out as an exercise.

A specially important quantity is the density of states $\mathcal{N}(\varepsilon)$ or the density of states per unit volume $n(\varepsilon) = \mathcal{N}(\varepsilon)/\Omega$. $\mathcal{N}(\varepsilon)d\varepsilon$ measures the number of states that have an energy between ε and $\varepsilon + d\varepsilon$. Its expression can be easily obtained by noting that the total number of states with an energy lower than ε is given by

$$L(\varepsilon) = \sum_{\alpha} \theta(\varepsilon - \varepsilon_\alpha) \tag{1.9}$$

where ε_α denotes the energy of the state with quantum numbers α. The density of states is obviously the derivative of this quantity, leading to

$$\mathcal{N}(\varepsilon) = \sum_{\alpha} \delta(\varepsilon - \varepsilon_\alpha) \tag{1.10}$$

As an illustration we will recompute the density of states for free fermions in any dimension.

$$\begin{aligned} \mathcal{N}(\varepsilon) &= \sum_{\sigma,\boldsymbol{k}} \delta(\varepsilon - \frac{\hbar^2 k^2}{2m}) \\ &= \frac{2\Omega}{(2\pi)^d} \int d\boldsymbol{k} \delta(\varepsilon - \frac{\hbar^2 k^2}{2m}) \end{aligned} \tag{1.11}$$

We now use the very convenient formula for δ functions

$$\delta(f(x)) = \sum_i \frac{1}{|f'(x_i)|}\delta(x - x_i) \tag{1.12}$$

where the x_i are the zeros of the function f (i.e. $f(x_i) = 0$).

Since the energy depends only on k^2 it is convenient to use spherical coordinates. One has $\int d\boldsymbol{k} = \int_0^\infty k^{d-1} dk S_d$ where S_d is the surface of the unit sphere [$S_1 = 2$, $S_2 = 2\pi$, $S_3 = 4\pi$ and $1/S_d = 2^{d-1}\pi^{d/2}\Gamma(d/2)$] and thus

$$\begin{aligned}
\mathcal{N}(\varepsilon) &= \frac{2\Omega}{(2\pi)^d} S_d \int_0^\infty dk\, k^{d-1} \frac{m}{\hbar^2 k} \delta(k - \frac{\sqrt{2m\varepsilon}}{\hbar}) \\
&= \frac{2\Omega S_d m}{\hbar^2 (2\pi)^d} \int_0^\infty dk\, k^{d-2} \delta(k - \frac{\sqrt{2m\varepsilon}}{\hbar}) \\
&= \frac{2\Omega S_d m}{\hbar^2 (2\pi)^d} \left(\frac{2m\varepsilon}{\hbar^2}\right)^{\frac{d-2}{2}}
\end{aligned} \tag{1.13}$$

for $\varepsilon > 0$ and zero otherwise. One thus sees that the density of states behaves in three dimensions as $n(\varepsilon) \propto \varepsilon^{1/2}$ while it is a constant in two dimensions and has a $1/\sqrt{\varepsilon}$ singularity at the bottom of the band in one dimension. In three dimensions the density of states per unit volume is (with the factor 2 coming from the spin degrees of freedom included)

$$n(\varepsilon) = \frac{m}{2\pi^2\hbar^2}\left(\frac{2mE_\mathrm{F}}{\hbar^2}\right)^{1/2} = \frac{3}{2}\frac{n}{E_\mathrm{F}} \tag{1.14}$$

Given the relative energies of E_F and, say, the temperature, most of the excitations will simply be blocked by the Pauli principle, and the ones that will play a role will be the ones close to the Fermi level. This simple fact is what gives to most solids their unusual properties, and allow for quantum effects to manifest themselves even at high (by human standards) temperature.

1.1.2 Electrons in periodic potentials: band theory

One of the most important features in solids is the presence of the potential imposed by the crystalline structure of the solids. The ions, charged positively act as a periodic potential on the electron and lead to the formation of energy bands.

There are two ways to view the formation of bands. The first one is to start from the free electrons and add a periodic potential on them. The total Hamiltonian of the system becomes

$$H = \frac{P^2}{2m} + V_0 \cos(QX) \tag{1.15}$$

where for simplicity we have written the periodic Hamiltonian in one dimension only. As explained in the previous section, the solutions of the $\frac{P^2}{2m}$ term are plane waves with a given momentum k. In order to understand the effect of the perturbation V_0 one can use simple perturbation theory. The perturbation is important when it couples states that have degenerate energy, which means that the states $-Q/2$ and $Q/2$ will be strongly coupled.

We will not follow this route here but look at the second way to obtain the main features of bands, namely to start from the opposite limit where the electrons are tightly bound to one site. Around the atom the electron is characterized by a certain atomic wavefunction $\langle r|\phi_i\rangle = \phi(r - r_i)$ that is not very important here. If the wave function is tightly bound around the atom then the overlap between the wavefunctions is essentially zero

$$\langle\phi_i|\phi_j\rangle = \delta_{ij} \tag{1.16}$$

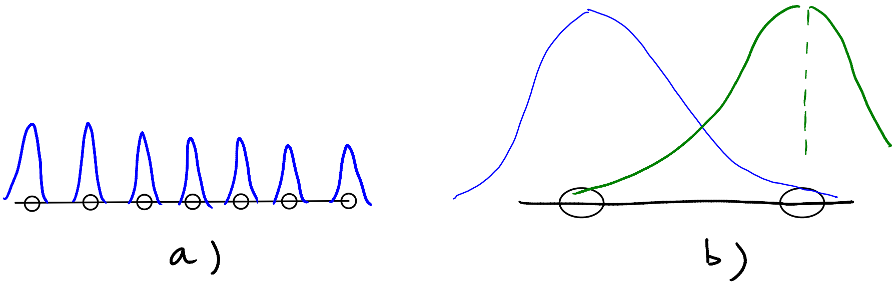


Figure 1.1: a) independent electrons; b) small overlap between the wavefunctions which defines the hopping integral t.

we assume in the following that the energy corresponding to this atomic wavefunction is E_0. This is shown in Fig. 1.1

In the following we will forget the notation $|\phi_i\rangle$ and simply denote the corresponding wavefunction by $|i\rangle$ to denote that this is the wavefunction around the i-th atom (centered on point r_i). The full state of the system is thus described by the basis of all the functions $|i\rangle$ and the energy of the problem would be

$$H = \sum_i E_0 |i\rangle\langle i| \tag{1.17}$$

Of course the wave functions between different sites are not completely orthogonal and there is a small overlap. The dominant one is of course the one between nearest neighbors but this can depend also on the shape of the individual atomic functions that could also favor some directions. This small overlap ensures that $|i\rangle$ is not an eigenstate of the Hamiltonian but that the matrix element $t_{ij} = \langle i|H|j\rangle$ is finite. The tight binding approximation consists in keeping this matrix element while still assuming that the direct overlap between the wavefunctions is zero (1.16). Physically t_{ij} describes the amplitude of tunnelling of a particle from the site r_i to the site r_j. It is important to note that systems such as cold atomic gases in optical lattices are excellent realizations of such a tight binding model. The Hamiltonian becomes

$$H = \sum_i E_0 |i\rangle\langle i| - t \sum_{\langle i,j\rangle} |i\rangle\langle j| \tag{1.18}$$

where we have here for simplicity only retained the overlap between nearest neighbors (denoted by $\langle i,j\rangle$). The first term is the energy of the degenerate atomic levels while the second term t describes the tunnelling between the different sites. The particles will thus delocalize to gain energy from the second term.

In order to solve the Hamiltonian (1.18) one notices that this Hamiltonian is invariant by translation. This means that the momentum is a conserved quantity, and one can simultaneously diagonalize the momentum operator and the Hamiltonian. The eigenstates of the momentum being plane waves, it means that it will be convenient to work in the Fourier space to get a simpler, and hopefully diagonal Hamiltonian. We use

$$\begin{aligned} |k\rangle &= \frac{1}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{ikr_j} |j\rangle \\ |j\rangle &= \frac{1}{\sqrt{N_s}} \sum_k e^{-ikr_j} |k\rangle \end{aligned} \tag{1.19}$$

where N_s is the number of lattice sites. For simplicity we have confined ourselves to one dimension, the generalization being obvious.

Two conditions constraint the allowed values of k. One is the usual quantification condition inside the box $k = \frac{2\pi n}{L}$ where n is a relative integer. As usual in Fourier transform large distances give a condition on the small values of k. Contrarily to the case of the continuum there is here a second condition coming from the fact that the space is discrete and that $r_j = aj$ where j is an integer can only take a set of discrete values. In order to get vectors $|j\rangle$ that are different from the second relation in (1.19) it is necessary for the coefficients in the sum to be different. It is easy to see that translating the value of k by $\frac{2\pi p}{a}$ where p is an integer leaves the exponentials unchanged and thus correspond in fact to identical $|k\rangle$. One must thus restrict the values of k in an interval of size $2\pi/a$. Here it is the small values of r that block the large values of k. One can take any interval. In order to have the symmetry $k \to -k$ obvious it is convenient to choose $[-\pi/a, +\pi/a]$ which is known as the first Brillouin zone. All other values of the k can be deduced by periodicity. The total number of allowed k values is

$$\frac{2\pi}{a}\frac{L}{2\pi} = \frac{L}{a} = N_s \tag{1.20}$$

which is indeed the number of independent states in the original state basis.

Using this new basis we can work out the Hamiltonian. Let us first look at the term

$$H_\mu = -\mu \sum_{j=0}^{N_s-1} |j\rangle\langle j| \tag{1.21}$$

Using (1.19) this becomes

$$H_\mu = -\mu \frac{1}{N_s} \sum_{j=0}^{N_s-1} \sum_{k_1} \sum_{k_2} e^{i(k_1-k_2)r_j} |k_1\rangle\langle k_2| \tag{1.22}$$

The sum over j can now be done

$$\frac{1}{N_s} \sum_{j=0}^{N_s-1} e^{i(k_1-k_2)r_j} \tag{1.23}$$

If $k_1 = k_2$ the sum is obviously 1. If $k_1 \neq k_2$ then one has a geometric series and the sum is

$$\frac{e^{i(k_1-k_2)aN_s} - 1}{e^{i(k_1-k_2)a} - 1} \tag{1.24}$$

which is always zero given the quantization condition on k. One has thus that the sum is δ_{k_1,k_2}. This gives

$$H_\mu = -\mu \sum_k |k\rangle\langle k| \tag{1.25}$$

as could be expected the Hamiltonian is diagonal. This could have been even directly written since this is just a chemical potential term counting the total number of particle which can be expressed in the same way regardless of the base (this is just the closure relation).

Let us now look at

$$H = -t \sum_{j=0}^{N_s-1} (|j\rangle\langle j+1| + \text{h.c.}) \tag{1.26}$$

a similar substitution now leads to

$$H = -t \frac{1}{N_s} \sum_{j=0}^{N_s-1} \sum_{k_1} \sum_{k_2} e^{i(k_1-k_2)r_j} e^{ik_2 a} |k_1\rangle\langle k_2| + \text{h.c.} \tag{1.27}$$

which after the sum over j has been made leads to

$$H = -t \sum_k 2\cos(ka)|k\rangle\langle k| \tag{1.28}$$

The transformed Hamiltonian, known as the tight-binding Hamiltonian thus reads

$$H = -t \sum_k 2\cos(ka)|k\rangle\langle k| + E_0 \sum_k |k\rangle\langle k| \tag{1.29}$$

As could be expected it is diagonal in k. This is because the initial Hamiltonian is invariant by translation and we have here only one state per unit cell. Thus the number of eigenstates in each k sector is only one. If one has had two atoms per unit cell, going to Fourier space would have reduced the $N_s \times N_s$ matrix to a 2×2 to diagonalize and so on. It is thus very important to notice the symmetries of the Hamiltonian and to use them to find the proper base.

The Hamiltonian (1.29) contains the atomic energy E_0. In the absence of hybridization the ground state is N_s times degenerate since the electrons can be put on each site. When there is hybridization t the electrons can gain energy by delocalizing (another expression of the uncertainty principle), which leads to the formation of energy bands. The tight binding is thus a very simple description that encompasses all the properties of the bands: counting the number of states, the proper analytical properties for the energy etc.

The generalization of the above formula to a square or cubic lattice is straightforward and gives

$$\varepsilon(\boldsymbol{k}) = -2 \sum_l t_l \cos(k_l a_l) \tag{1.30}$$

where l denotes each coordinate axis. Close to the bottom of the band one can expand the cosine to get an energy of the form

$$\varepsilon(k) = E_0 - 2t + tk^2 \tag{1.31}$$

this allows to define an effective mass $m^* = 1/(2t)$ by analogy with the energy of free electrons. Here the "mass" has nothing to do with the real mass of the electron but simply describes the facility with which the electron is able to move from one site to the next. The mass can (and in general will) of course be anisotropic since there is no reason why the overlap of atomic orbital in different directions be the same.

It is worth noticing that the filling of the band is crucial for the electronic properties of the system. A system which has one electron per site will fill half of the allowed values of k in the band (because of the spin one value of k can accommodate two electrons of opposite spins). One has thus a half filled band, which usually gives a very good density of states at the Fermi level. One can thus expect, based on independent electrons, in general systems with one electron per site to be good metals. On the contrary a system with two electrons per site will fill all values of k and thus correspond to an insulator, or a semiconductor if the gap to the next band is not too large. It was a tremendous success of band theory to predict based on band filling which elements should conduct or not.

1.1.3 Thermodynamic observables

Let us now examine some of the physical consequences for physical observables of this peculiar features of the electron gas.

A very simple thermodynamic quantity that one can compute is the specific heat of the solid. The specific heat is simply the change in energy (heat) of the system with respect with the

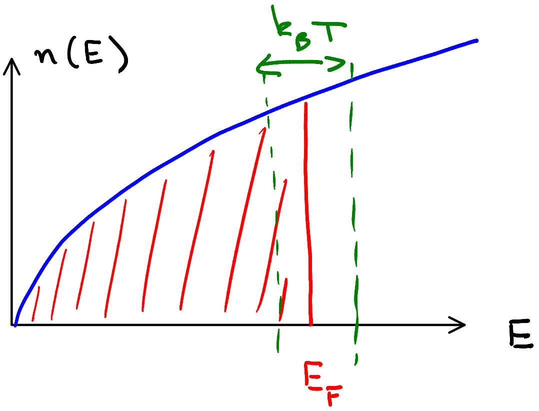


Figure 1.2: The difference in energy between a system at $T = 0$ and T finite is due to the thermal excitations of particles within a slice of $k_{\mathrm{B}}T$ around the Fermi energy E_{F}. All the others are blocked by the Pauli principle.

temperature. The total energy per spin degree of freedom is given by

$$E(T) = \sum_{\boldsymbol{k}} f_{\mathrm{F}}(\varepsilon(\boldsymbol{k}) - \mu(T))\varepsilon(\boldsymbol{k}), \tag{1.32}$$

while the chemical potential is given by the conservation of the total number of particles

$$\mathcal{N} = \sum_{\boldsymbol{k}} f_{\mathrm{F}}(\varepsilon(\boldsymbol{k}) - \mu(T)). \tag{1.33}$$

Notice that in these equations what is fixed is the number of particles, and therefore the chemical potential depends on temperature. Even if one normally uses the grand-canonical ensemble to obtain Eqs. (1.32) and (1.33), they are also valid in the canonical ensemble, by fixing $\mathcal{N}$. Differentiating (1.32) with respect to T gives the specific heat. The full calculation will be done as an exercise. We will here just give a qualitative argument, emphasizing the role of the Fermi surface.

When going from $T = 0$ to the small temperature T, particle in the system will gain an energy of the order of $k_{\mathrm{B}}T$ since they can be thermally excited. However the Pauli principle will block most of such excitations and thus only the particles that are within a slice of $k_{\mathrm{B}}T$ in energy around the Fermi energy can find the empty states in which they can be excited as indicated in Fig. 1.2. The number of such excitations is thus

$$\Delta N(T) = k_{\mathrm{B}}T\mathcal{N}(E_{\mathrm{F}}) \tag{1.34}$$

and the gain in energy is

$$\Delta E(T) = k_{\mathrm{B}}^2 T^2 \mathcal{N}(E_{\mathrm{F}}) \tag{1.35}$$

leading to a specific heat (at constant volume)

$$C_V(T) \propto k_{\mathrm{B}}^2 \mathcal{N}(E_{\mathrm{F}})T \tag{1.36}$$

The Pauli principle and the large Fermi energy compared to the temperature thus directly imply that the specific heat of an independent electron gas is linear in temperature. The proportionality coefficient γ is, up to nonimportant constants directly proportional to the density of states at the Fermi level. This is the first illustration of something that we will encounter often: because of the Pauli principle, most of the states are blocked and thus useless. Only a very small fraction of the electron, close to the Fermi level contributes to the physical observables. This is a very important point, since it means that we can essentially ignore, in most cases, most of the precise details of the band structure and kinetic energy, provided that we know what is the density of states at the Fermi level. In practise, because the energy scale that we are probing (here the temperature) is usually much smaller than the typical energy scale over which the density of state varies we can consider that this quantity is a constant.

The linear dependence of the specific heat of the fermions, is a spectacular manifestation of the Pauli principle. Indeed let us assume instead that our electrons were classical particles. Then we could compute the total energy using the equipartition, and the fact that this is $\frac{1}{2}k_{\mathrm{B}}T$ per degree of freedom. We would have

$$C_V^{\mathrm{cl}}(T) = \frac{1}{2}\mathcal{N}k_{\mathrm{B}} \tag{1.37}$$

which using (1.14) would lead to

$$C_{\mathrm{el}}/C_{\mathrm{cl}} \equiv \frac{\pi^2}{3}\left(\frac{k_{\mathrm{B}}T}{E_{\mathrm{F}}}\right) \tag{1.38}$$

which would lead easily at temperatures of the order of $10K$ but even at ambient temperature to an error of several orders of magnitude.

Let us now move to another thermodynamic quantity namely the compressibility. Normally the compressibility (at constant temperature) of a system is the way the volume varies when one varies the pressure, namely

$$\kappa = -\frac{1}{\Omega}\left(\frac{d\Omega}{dP}\right)_T \tag{1.39}$$

where the $\frac{1}{\Omega}$ normalization is to define an extensive quantity independent of the volume of the system, and the minus sign is a simple convention to get positive numbers since most systems have a diminishing volume when the pressure increases.

This thermodynamic definition of the compressibility is quite inconvenient to work with for the electron gas. However one can relate the compressibility to

$$\kappa = \left(\frac{d\mathcal{N}}{d\mu}\right)_T \tag{1.40}$$

At zero temperature the compressibility can be readily computed by noting that

$$\mathcal{N} = \int_{-\infty}^{\mu} d\varepsilon \mathcal{N}(\varepsilon) \tag{1.41}$$

and thus

$$\kappa = \mathcal{N}(E_{\mathrm{F}}) \tag{1.42}$$

as is obvious from Fig. 1.3.

One notes that again, only the density of states at the Fermi level enters in the value of the compressibility (up to non important factors, that are independent of the physical system considered). This is again a consequence of the Pauli principle. Insulators for the which the density of states is zero at the Fermi level are incompressible. If the chemical potential is varied no additional electron can enter the system. A naive picture of this could be to say that if we have already two electrons per site (a filled band) then there is no “place” where one could squeeze an additional electron. Alternatively a metal, which has a finite density of states at the Fermi level can accommodate additional electrons when the Fermi level is increased. The same image would apply since in that case the band would be partly filled and one would have places with zero or only one electro where one could insert additional particles.

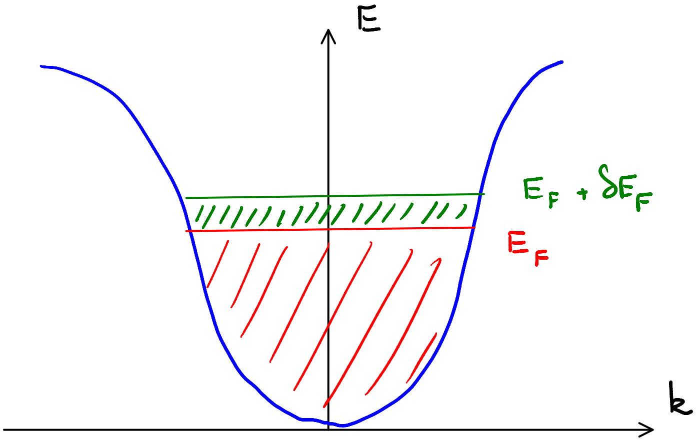


Figure 1.3: Change of number of particles for a change of chemical potential

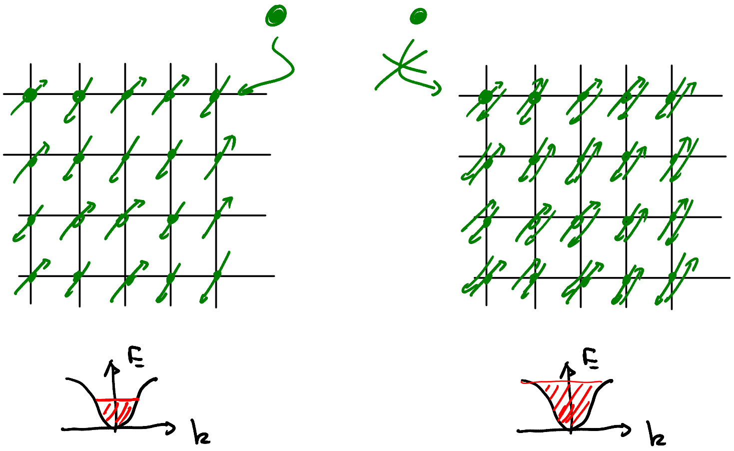


Figure 1.4: Cartoon of incompressibility. A system with a half filled band has many sites where an additional electron could be added and is thus compressible (left). On the other hand a filled band corresponds to two electron per site. No additional electron could be added even if the chemical potential is increased. The system is incompressible.

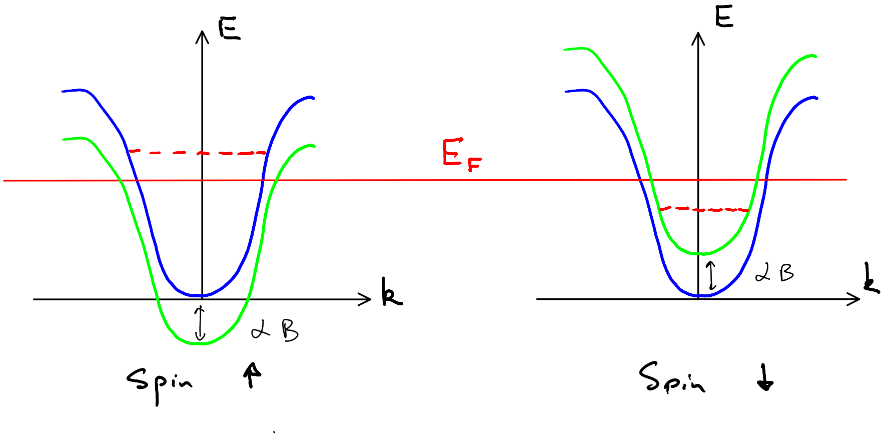


Figure 1.5: Cartoon of susceptibility. The energy levels of the two spin species are shifted up and down (green curve) in the presence of a magnetic field, compared to the zero field case (blue curve). This is equivalent in saying that the two spin species see a different chemical potential (dashed red line) than the true chemical potential (red line). This creates an imbalance of populations and thus a magnetization.

Finally for a solid the last simple useful thermodynamic quantity is the magnetic susceptibility. Quite generally the magnetic susceptibility is the way the magnetization varies when an external magnetic field is applied on the system

$$\chi = \left(\frac{dM}{dH}\right)_T \tag{1.43}$$

The main source of magnetization in the solid is provided by the spins of the electrons (there are also orbital effects but let us ignore those for the moment). The magnetization per spin is given by

$$\boldsymbol{m} = g\mu_{\mathrm{B}}\boldsymbol{\sigma} \tag{1.44}$$

where μ_{B} is the Bohr magneton, a quantity depending on the unit system, allowing the conversion of orbital moments into magnetic moments, and g the Lande factor is a dimensionless number telling for each particle how the orbital moment converts into a magnetic moment ($g \simeq 2$ for the electron in a vacuum). The energy gained by the spins when coupled with an external magnetic field is thus

$$H_B = -\boldsymbol{B} \cdot \sum_i \mathcal{N} g\mu_{\mathrm{B}}\boldsymbol{\sigma}_i \tag{1.45}$$

Applying the field in the z direction and using the fact that for a spin 1/2 only two quantized values of the spin are possible one obtains

$$H_B = -\frac{g\mu_{\mathrm{B}}}{2}B(\mathcal{N}_\uparrow - \mathcal{N}_\downarrow) \tag{1.46}$$

The energies for each up (resp down) spins is thus shifted by $\varepsilon(k) \to \varepsilon(k) \mp (g\mu_{\mathrm{B}})B$. As shown in Fig. 1.5 this implies, since the chemical potential remains unchanged that more spin up and less spin downs will be present in the system. In a total similarity with the compressibility

$$\Delta\mathcal{N}_{\uparrow,\downarrow} = \pm\mathcal{N}(E_{\mathrm{F}})\frac{g\mu_{\mathrm{B}}}{2}B \tag{1.47}$$

leading to a variation of magnetization due to the spins

$$\Delta M_z = \frac{(g\mu_{\mathrm{B}})}{2}(\Delta\mathcal{N}_\uparrow - \Delta\mathcal{N}_\downarrow) = B\frac{(g\mu_{\mathrm{B}})^2}{4}\mathcal{N}(E_{\mathrm{F}}) \tag{1.48}$$

and thus to a spin susceptibility

$$\chi = \frac{(g\mu_{\mathrm{B}})^2}{4}\mathcal{N}(E_{\mathrm{F}}) \tag{1.49}$$

We again see that only the states very close to the Fermi level contribute, which implies that the spin susceptibility is again controlled by the density of states at the Fermi level.

This little survey of noninteracting electrons thus disclosed various important facts that constitute the essence of what a non-interacting electron gas looks like, and that we can summarize below. These properties will of course be crucial to set a frame of comparison with the case of interacting particles.

The ground state of the system is a Fermi sea with a certain number of states occupied, the other are empty at zero temperature. There is a *sharp* separation between these two set of states and in particular a discontinuity in the occupation factor $n(k)$ at the Fermi level k_{F}. For a non interacting electron gas all states below the Fermi energy are occupied with probability one, all states above with probability zero.

The thermodynamics corresponding to this state, dominated by the Pauli principle, leads to

1. A specific heat that is linear in temperature $C_V \propto \gamma T$ for temperatures much smaller than the Fermi energy ($T \ll E_{\mathrm{F}}$)

2. A charge compressibility that goes to a constant κ_0 at zero temperature.

3. A spin susceptibility that goes to a constant χ_0 at zero temperature.

For noninteracting electrons, these three constants γ, κ_0 and χ_0 are up to non system dependent constants simply the density of states at the Fermi level $\mathcal{N}(E_{\mathrm{F}})$.

Finally the excitations above the ground state are easy to identify for the case of independent electrons. They consist is adding an electron in an eigenstate of momentum $\boldsymbol{k}$ and spin σ, or in removing one electron from the occupied states below the Fermi level (in other words creating a hole), again with a well defined momentum and spin.

1.2 Coulomb interaction

Let us now turn to the effects of the interactions. The dominant interaction in a solid is provided by the Coulomb interaction between the charges. There is the interaction between the electrons and the ions (positively charged) of the lattice, and also of course the interaction between the electrons themselves.

The first part is already partly taken into account when one computes the bandstructure of the material, and thus incorporated in the energy and the density of states. Of course this is not the only effects of this interaction and many additional effects are existing, in particular when the lattice vibrates. But the main part of the electron-ion interaction is already taken into account.

The electron-electron interaction is a totally different matter since it directly gives an interaction between all the 10^{23} particles in the what was our band for independent electrons. How much remains of the free electron picture when this interaction is taken into account is what we need to understand.

1.2.1 Coulomb interaction in a solid

Let us first look what the Coulomb interaction does in a solid. One could think naively that this is the same thing than for two charges in the vacuum, but this would be too naive since there are plenty of mobile charges around and thus they can provide screening of the interaction.

In order to study the Coulomb interaction let us compute the potential $V(\boldsymbol{r})$ created by a test charge Q placed at the origin of the solid. The potential obeys the Poisson equation

$$\Delta V(\boldsymbol{r}) + \frac{\rho(r)}{\epsilon_0} = 0 \tag{1.50}$$

where $\rho(\boldsymbol{r})$ is the charge density of *all* the charges in the solid. In the vacuum one has simply

$$\rho(\boldsymbol{r}) = Q\delta(\boldsymbol{r}) \tag{1.51}$$

and the solution of (1.50) is simply

$$V(\boldsymbol{r}) = \frac{Q}{4\pi\epsilon_0 r} \tag{1.52}$$

In a solid, in addition to the test charge there are the charges present of the solid. This includes the electrons and the ions. All these charges will be affected by the presence of the test charge Q and will try to get closer or move away from her. In order to make a simple calculation let us assume for the moment that the ions are massive enough not to move, and besides that they are simply providing a uniform positive potential to ensure the total charge neutrality with the electrons. This model is known under the name of jelium model. The total density of charges is thus

$$\rho(\boldsymbol{r}) = Q\delta(\boldsymbol{r}) + [\rho_e(\boldsymbol{r}) - \rho_0] \tag{1.53}$$

where $\rho_e(\boldsymbol{r})$ is the electronic charge and ρ_0 the uniform background provided by the ions. Calling e the charge of the electron, n_0 the density of electrons and $n(\boldsymbol{r})$ the particle density at point $\boldsymbol{r}$ one has

$$\rho(\boldsymbol{r}) = Q\delta(\boldsymbol{r}) + e[n(\boldsymbol{r}) - n_0] \tag{1.54}$$

Since the electrons are mobile, the density of the electrons at a given point depends on the actual electrostatic potential at that point making (1.50) a rather complicated equation to solve. To get a simple solution, let us make an approximation known as the Thomas-Fermi approximation. Namely we will assume that the external potential $V(\boldsymbol{r})$ is varying slowly enough in space so that one can consider each little volume of electrons as an independent system, subjected to a uniform potential V (dependent on the point). This is sketched in Fig. 1.6. Typically the "important" electrons being the ones at the Fermi level, one can imagine that the relevant set of wavevectors is k_{F} and thus the corresponding lengthscales is k_{F}^{-1}. For typical metals this is a scale of the order of the thenth of nanometers. As long as $V(\boldsymbol{r})$ varies more smoothly than that one could expect the Thomas-Fermi approximation to be a good one. If we admit this approximation, then each little volume of electron has, in addition to its kinetic energy, the electrostatic contribution of the total charge $e\Omega n$ in the potential V

$$H_V = -\Omega e V n \tag{1.55}$$

Thus each energy level of each electron in the little volume is shifted by $\varepsilon(\boldsymbol{k}) \to \varepsilon(\boldsymbol{k}) - eVn$, which by the same reasoning as in the previous chapter leads to a variation of density which is

$$\Delta n = eVn\mathcal{N}(E_{\mathrm{F}}) \tag{1.56}$$

Electrons are attracted to regions of positive electrostatic potential, while they are repelled from regions with negative ones. The essence of the Thomas-Fermi approximation is that we can use this formula for each "point" in space and thus

$$\Delta n(\boldsymbol{r}) = eV(\boldsymbol{r})\mathcal{N}(E_{\mathrm{F}}) \tag{1.57}$$

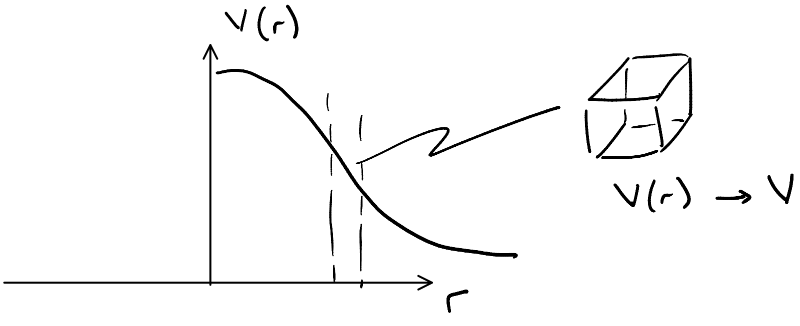


Figure 1.6: In the Thomas-Fermi approximation the potential is supposed to vary sufficiently slowly over the characteristics lengthscales of the electron gas that each little volume surrounding a point in space can be viewed as a whole thermodynamic system, seeing the uniform potential $V(r) \to V$.

which gives us the needed equation to relate the density of charge and the electrostatic potential. The variation of density due to the potential is exactly $n(\boldsymbol{r}) - n_0$ since in the absence of the test charge the electron gas is homogeneous and its uniform density neutralizes exactly the one of the ions. One thus has

$$\Delta V(\boldsymbol{r}) + \frac{Q\delta(r) + e^2\mathcal{N}(E_{\mathrm{F}})V(\boldsymbol{r})}{\epsilon_0} = 0 \tag{1.58}$$

To solve this equation it is important to recognize that this is a linear equation. This should start a Pavlovian reflex that immediately induce the use of Fourier transform. Indeed one the great interest of Fourier transform is to transform differentiation into simple multiplications, and thus allowing to replace a differential equation by a simple algebraic one. In that case one uses

$$V(\boldsymbol{r}) = \frac{1}{\Omega}\sum_{\boldsymbol{k}} V(\boldsymbol{k})e^{i\boldsymbol{k}\boldsymbol{r}} \tag{1.59}$$

A word on the notations. We will always denote the sums over $\boldsymbol{k}$ by a discrete sum, thus implicitly taking into account a quantization in a large box. For the case when the volume goes to the infinity one simply replaces $\sum_{\boldsymbol{k}} \to \frac{\Omega}{(2\pi)^2}\int d\boldsymbol{k}$. The direct Fourier transform is

$$V(\boldsymbol{k}) = \int_{\Omega} d\boldsymbol{r} V(\boldsymbol{r})e^{-i\boldsymbol{k}\boldsymbol{r}} \tag{1.60}$$

One can either substitute (1.59) in (1.58) or perform the Fourier transform of the equation. Let us do the later to detail the calculation. The Fourier transform of the equation becomes

$$\int d\boldsymbol{r}e^{-i\boldsymbol{k}\boldsymbol{r}}[\Delta V(\boldsymbol{r})] + \frac{Q}{\epsilon_0}\int d\boldsymbol{r}e^{-i\boldsymbol{k}\boldsymbol{r}}\delta(\boldsymbol{r}) + \frac{e^2\mathcal{N}(E_{\mathrm{F}})}{\epsilon_0}\int d\boldsymbol{r}e^{-i\boldsymbol{k}\boldsymbol{r}}V(\boldsymbol{r}) = 0 \tag{1.61}$$

The first term corresponds to sums of the form

$$\int d\boldsymbol{r}e^{-i\boldsymbol{k}\boldsymbol{r}}[\partial_x^2 V(\boldsymbol{r})] \tag{1.62}$$

where x denotes here one of the spatial variables $\boldsymbol{r} = (x, y, z, \ldots)$. One can integrate twice by part to obtain

$$(-ik_x)^2 \int d\boldsymbol{r} e^{-i\boldsymbol{k}\boldsymbol{r}} V(\boldsymbol{r}) = (-ik_x)^2 V(\boldsymbol{k}) \tag{1.63}$$

which is of course the great advantage of having used the Fourier representation. The equation thus becomes

$$k^2 V(\boldsymbol{k}) + \frac{e^2 \mathcal{N}(E_{\mathrm{F}})}{\epsilon_0} V(\boldsymbol{k}) = \frac{Q}{\epsilon_0} \tag{1.64}$$

which immediately gives the Fourier transform of the potential created by the charge Q in the solid

$$V(\boldsymbol{k}) = \frac{Q/\epsilon_0}{k^2 + \frac{e^2 \mathcal{N}(E_{\mathrm{F}})}{\epsilon_0}} \tag{1.65}$$

We see immediately that this defines a lengthscale $\lambda^{-2} = \frac{e^2 \mathcal{N}(E_{\mathrm{F}})}{\epsilon_0}$. To understand its meaning let us perform the inverse Fourier transform.

$$V(\boldsymbol{r}) = \frac{1}{\Omega} \int d\boldsymbol{k} \, \frac{Q/\epsilon_0}{k^2 + \lambda^{-2}} \tag{1.66}$$

Let us specialize to $d = 3$ and take the limit of an infinite volume. The integral becomes

$$V(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int d^3k \, \frac{Q/\epsilon_0}{k^2 + \lambda^{-2}} e^{i\boldsymbol{k}\boldsymbol{r}} \tag{1.67}$$

The rotational symmetry of the integrand immediately suggests to use the spherical coordinates. One gets

$$\begin{aligned} V(\boldsymbol{r}) &= \frac{1}{(2\pi)^2} \int_0^\infty k^2 dk \int_{-\pi}^{+\pi} \sin\theta d\theta \frac{Q/\epsilon_0}{k^2 + \lambda^{-2}} e^{ikr\cos\theta} \\ &= \frac{1}{(2\pi)^2} \int_0^\infty k^2 dk \frac{Q/\epsilon_0}{k^2 + \lambda^{-2}} \frac{e^{ikr} - e^{-ikr}}{ikr} \\ &= \frac{1}{(2\pi)^2} \int_{-\infty}^\infty k dk \frac{Q/\epsilon_0}{k^2 + \lambda^{-2}} \frac{e^{ikr}}{ir} \end{aligned} \tag{1.68}$$

There are various ways to finish the calculation, using conventional integration techniques. Let us illustrate however on this simple example the use of integration in the complex plane (see Appendix A for a reminder). Since r is positive, we can replace the integral by an integral over the closed contour of Fig. 1.7 without changing the value of the integral. One has thus

$$V(\boldsymbol{r}) = \frac{Q}{ir\epsilon_0 (2\pi)^2} \oint_C dz \frac{z}{z^2 + \lambda^{-2}} e^{ikr} \tag{1.69}$$

One can rewrite the fraction as

$$\frac{z}{z^2 + \lambda^{-2}} = \frac{1}{2} \left[\frac{1}{z + i\lambda^{-1}} + \frac{1}{z - i\lambda^{-1}} \right] \tag{1.70}$$

which shows directly the two poles $z = \pm i\lambda^{-1}$. Only the upper pole is inside the contour. Using the residue theorem one gets

$$\begin{aligned} V(\boldsymbol{r}) &= \frac{Q}{ir\epsilon_0 (2\pi)^2} (2i\pi) \frac{1}{2} e^{-\lambda r} \\ &= \frac{Q}{4\pi\epsilon_0 r} e^{-\lambda r} \end{aligned} \tag{1.71}$$

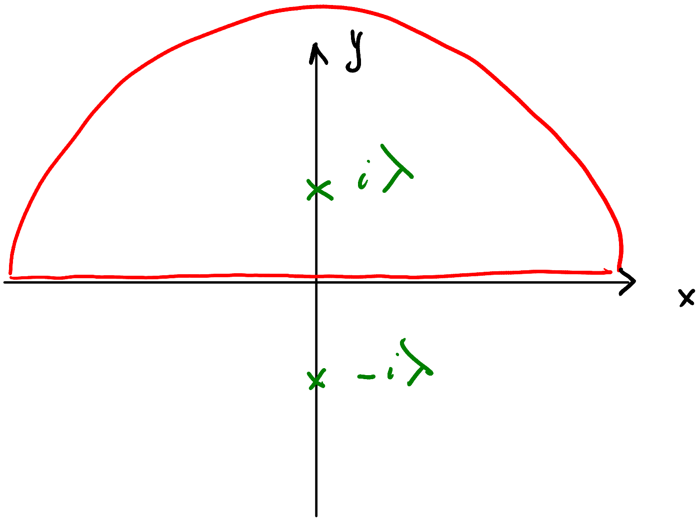


Figure 1.7: Contour for computing the integral. The circle of infinite radius gives a zero contribution to the integral since the integrand decays fast enough with the radius. The integral over the contour is thus equal to the integral on the real axis. Only the poles inside the contour contribute to the residue theorem.

Before we tackle the physics of this solution, let us make some comments on the calculation itself. One sees that the presence of the λ term in the Fourier transform $V(\boldsymbol{k})$ which changes the behavior at small $\boldsymbol{k}$ affects indeed the behavior of $V(\boldsymbol{r})$ at large distance and transforms a powerlaw decay $(1/\boldsymbol{r}^2)$ into an exponential decay. This is the logical correspondence in Fourier transform between the small $\boldsymbol{k}$ and the large $\boldsymbol{r}$. In the absence of such a term the Fourier transform can be evaluated by simple dimensional analysis. Indeed in

$$\int d\boldsymbol{k}\frac{1}{k^2}e^{i\boldsymbol{k}\boldsymbol{r}} \tag{1.72}$$

the singularity in the integrand is coming from small $\boldsymbol{k}$. One can consider roughly that the exponential term is a constant as long as $r < 1/k$ and will start oscillating when $r > 1/k$. In that case the oscillations essentially cancel the integral. One can thus roughly replace the true integral by

$$\int_{k>1/r} d\boldsymbol{k}\frac{1}{k^2} \sim \int_{k>1/r} k^{d-1-2}dk \sim r^{2-d} \tag{1.73}$$

by simple dimensional analysis. This is indeed the correct result [in $d=2$ the power zero gives in fact a $\log(r)$] and one recovers in particular the $1/r$ behavior of the Coulomb potential. Conversely one can see that the slow decay $1/r$ of the Coulomb potential means that the $k=0$ Fourier component cannot be finite since

$$\int^L d\boldsymbol{r}\frac{1}{r} \sim L^{d-1} \tag{1.74}$$

and thus diverge (in $d=2,3$ and even logarithmically in $d=1$). This means by the same arguments that the Fourier transform is a powerlaw of k

$$\int d\boldsymbol{r}\frac{1}{r}e^{-i\boldsymbol{k}\boldsymbol{r}} \sim \int^{1/k} d\boldsymbol{r}\frac{1}{r} \sim k^{1-d} \tag{1.75}$$

leading back to the $1/k^2$ in three dimensions. For the same reasons, if the potential is now exponentially decreasing (or with a high enough power) one sees immediately that the $\boldsymbol{k} = 0$ Fourier component is must now be finite since

$$V(\boldsymbol{k} = 0) = \int d\boldsymbol{r} V(\boldsymbol{r}) \tag{1.76}$$

will now be finite. These simple dimensional arguments, and the identification of the dominant divergence in an integral to try to estimate its behavior can be used in several occasion and it is worth becoming familiar with them.

Let us now go back to the physics behind (1.71). The form of the potential is known as the Yukawa potential. One sees that the Coulomb potential in a solid is not long range anymore, but decays extremely rapidly beyond the length λ, called the screening length. This drastic change of behavior comes from the fact that electrons being mobile can come and surround the external charge Q. As long as this charge produces a visible potential it will attract or repel the electrons, until their cloud of charge exactly compensates the external charge. We thus have the paradoxical result that in a solid the Coulomb interaction is short range, and of range λ. This means that two charges that are beyond the length λ will essentially not see each other. As can be expected λ is again proportional to the density of states at the Fermi level: one needs to have electrons that can be excited to be able to screen.

Let us estimate λ. We can use the fine structure constant

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137} \tag{1.77}$$

to obtain

$$\lambda^{-2} = 4\pi\alpha\hbar c \mathcal{N}(E_{\mathrm{F}}) = 4\pi\alpha\hbar c \frac{3n}{2E_{\mathrm{F}}} \tag{1.78}$$

using (1.14). Using $E_{\mathrm{F}} = \hbar v_{\mathrm{F}} k_{\mathrm{F}}$, and $6\pi^2 n = k_{\mathrm{F}}^3$ one gets

$$\lambda^{-2} = \frac{1}{\pi}\alpha\frac{c}{v_{\mathrm{F}}}k_{\mathrm{F}}^2 \tag{1.79}$$

Since $c/v_{\mathrm{F}} \sim 10^2$ in most systems, one sees that $k_{\mathrm{F}}\lambda \sim 1$. In other words the screening length is of the order of the inverse Fermi length, i.e. essentially the lattice spacing in normal metals. This is a striking result, which means that not only is the Coulomb interaction screened, but that the screening is so efficient that the interaction is practically local ! Of course one could then question the precise approximations that we have used to establish this formula but the order of magnitude will certainly remain.

1.3 Importance of the interactions

One could thus hope from the previous chapter that the Coulomb interaction plays a much minor role than initially anticipated. Let us estimate what is its order of magnitude compared to the kinetic energy. The interaction between two particles can be written as

$$H_{\mathrm{int}} = \int d\boldsymbol{r} d\boldsymbol{r}' V(\boldsymbol{r} - \boldsymbol{r}')\rho(\boldsymbol{r})\rho(\boldsymbol{r}') \tag{1.80}$$

Since the interaction is screened it will be convenient to replace it by a local interaction. We will assume based on the results of the previous chapter that the screening length λ is roughly the interparticle spacing a. Let us look at the effective potential seen at point $\boldsymbol{r}$ by one particle

$$\int d\boldsymbol{r}' V(\boldsymbol{r} - \boldsymbol{r}')\rho(\boldsymbol{r}') \tag{1.81}$$

we can consider that due to screening we can only integrate within a ball of radius a around the point r. Assuming that the density is roughly constant one obtains

$$\int_{|\boldsymbol{r}-\boldsymbol{r}'|<a} d\boldsymbol{r} \frac{e^2}{4\pi\epsilon_0|\boldsymbol{r}-\boldsymbol{r}'|}\rho_0 \sim \frac{e^2\rho_0 S_d a^{d-1}}{4(d-1)\pi\epsilon_0} \tag{1.82}$$

and using $\rho_0 \sim 1/a^d$ and (1.77) one gets

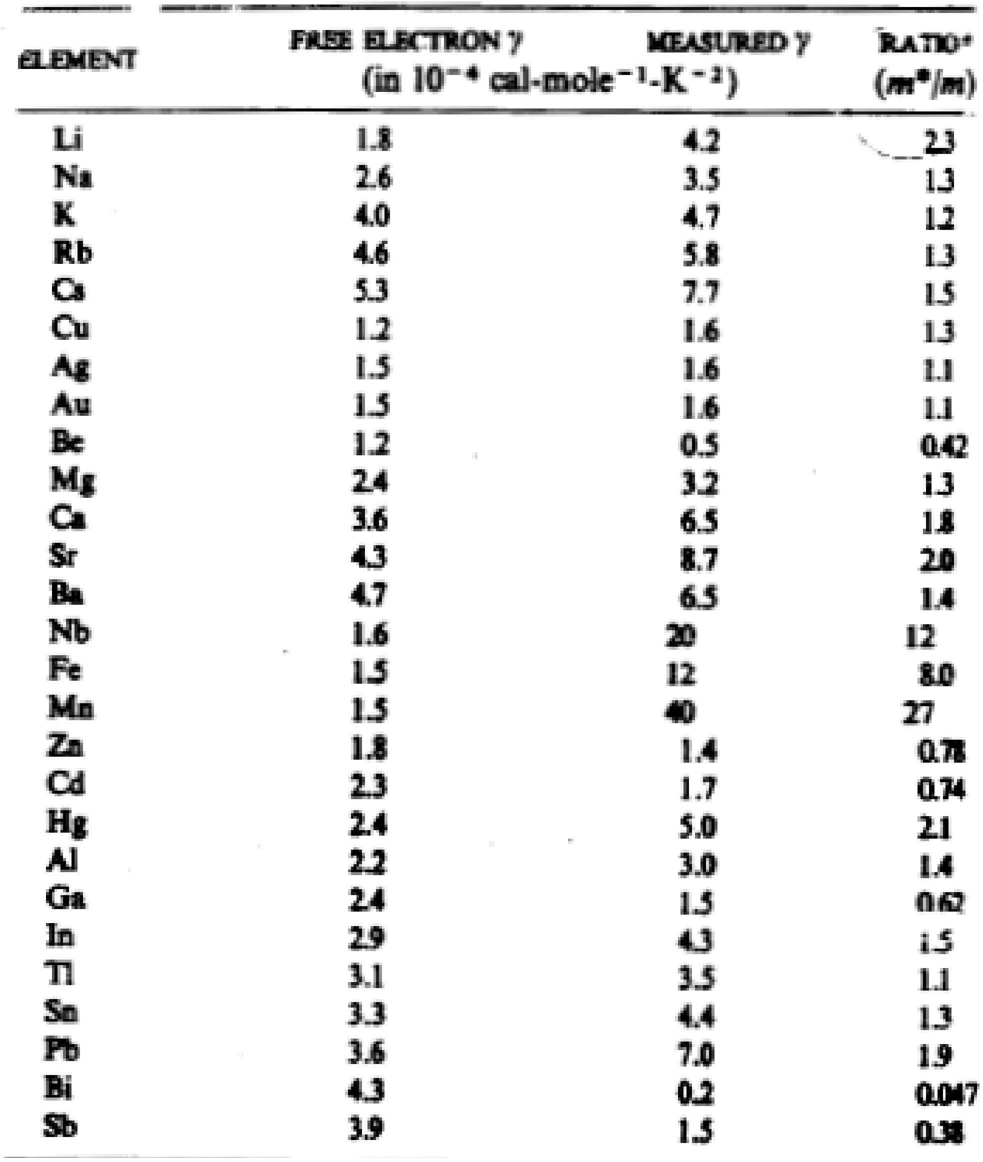
$$\frac{S_d \alpha \hbar c}{(d-1)a} \tag{1.83}$$

Since this is the potential acting on a particle, this has to be compared with the kinetic energy of this particle at the Fermi level which is $E_{\rm F} = \hbar v_{\rm F} k_{\rm F}$. Since $k_{\rm F} \sim a^{-1}$ one sees that one has again to compare α and $c/v_{\rm F}$ which are about the same order of magnitude ! The Coulomb energy, even if screened, is thus of the *same* order than the kinetic energy even in good metals. This means energies of the order of the electron volt.

1.4 Theoretical assumptions and experimental realities

How much of the previous estimates and calculation corresponds to the actual solids. Let us start with measurements of the specific heat. Results are shown in Fig. 1.8 where the coefficient of the linear term of the specific heat is given for simple elements. The first observation is that even for the realistic systems the specific heat is still linear in temperature. This is already a little bit surprising since the linear behavior of the temperature is coming from the existence of a sharp discontinuity at the Fermi surface. One could have naively expected that since the energy of the interaction is of the order of the Fermi energy, the probability of having occupied states is now spread over energies of the order of the electron Volt, as indicated in Fig. 1.9 It is thus surprising to still have a linear T dependence of the specific heat. The independent electron results seem to be much more robust than anticipated. One can nevertheless see from Fig. 1.8 that although the picture of independent electrons works qualitatively it does not work quantitatively and that the coefficient γ can be quite different from the one from the free electron picture.

For the electron gas various factors can enter in this change of γ. First the bandstructure of the material can lead, as we saw, to a strong change of the dispersion relation, and thus to a quite different γ. Second to estimate the effects of the interactions is difficult given their long range nature (with the screening) in solids. An very nice alternative to electrons is provided by ^{3}He. Indeed the ^{3}He atom is a fermion, since it is made of three nucleons. It is neutral, and since the scattering potentials of two ^{3}He atoms are very well known the interactions are short range and perfectly characterized. In addition the kinetic energy is simply of the form $P^2/(2M)$ so the density of states at the Fermi level are perfectly known. The specific heat coefficient, compressibility and spin susceptibility are shown in Fig. 1.10 Here again one has the surprising result that the independent fermion theory works qualitatively very well. In addition to the specific heat that is linear in temperature, the compressibility is a constant at low temperatures and the spin susceptibility is also a constant. Both these last properties are also strongly dependent on the existence of a sharp discontinuity at the Fermi surface at zero temperature, and it is thus very surprising to see the hold in the presence of interactions. But as for the electron case, one sees that the values of these three quantities are *not* given by the independent fermion theory, where these three quantities are simply the density of states at the Fermi level. Here we have three independent numbers, which clearly vary as a function of the interaction, as can be seen by the pressure dependence of these quantities. Indeed increasing the pressure changes the density of particles, and thus the interaction between them (the change



ELEMENT	FREE ELECTRON γ (in 10^{-4} cal-mole^{-1}-K^{-2})	MEASURED γ (in 10^{-4} cal-mole^{-1}-K^{-2})	RATIO* (m^*/m)
Li	1.8	4.2	2.3
Na	2.6	3.5	1.3
K	4.0	4.7	1.2
Rb	4.6	5.8	1.3
Cs	5.3	7.7	1.5
Cu	1.2	1.6	1.3
Ag	1.5	1.6	1.1
Au	1.5	1.6	1.1
Be	1.2	0.5	0.42
Mg	2.4	3.2	1.3
Ca	3.6	6.5	1.8
Sr	4.3	8.7	2.0
Ba	4.7	6.5	1.4
Nb	1.6	20	12
Fe	1.5	12	8.0
Mn	1.5	40	27
Zn	1.8	1.4	0.78
Cd	2.3	1.7	0.74
Hg	2.4	5.0	2.1
Al	2.2	3.0	1.4
Ga	2.4	1.5	0.62
In	2.9	4.3	1.5
Tl	3.1	3.5	1.1
Sn	3.3	4.4	1.3
Pb	3.6	7.0	1.9
Bi	4.3	0.2	0.047
Sb	3.9	1.5	0.38

Figure 1.8: Coefficient γ of the linear term in temperature for the specific heat, both from free electron calculations and measured for simple elements (From [AM76]). This shows that for realistic systems, the specific heat still exhibits a linear temperature dependence at low temperatures, just like for free electrons. The slope γ is different from the one of free electrons and allows to define an effective mass m^*.

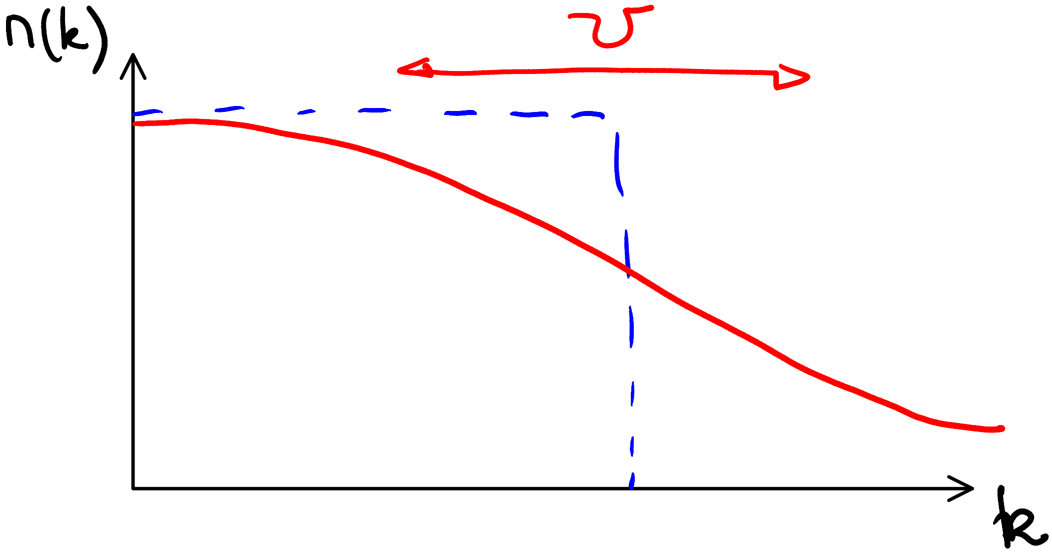


Figure 1.9: Cartoon of the expected occupation factor, for an interaction energy U of the order of the Fermi energy. One expects the occupation factor $n(\boldsymbol{k})$ to be spread over energies of the order of the interaction, leading to a washing out of the Fermi surface singularity.

in kinetic energy and density of states can be computed very precisely in that case). We will thus have to understand this very puzzling experimental fact.

Let us now look at the excitations in a more microscopic way. A remarkable experimental technique to look at the single particle excitations is provided by the photoemission technique. We will come back in more details on this technique but a sketch is shown in Fig. 1.11 Photons are send on the sample, and they kick out an electron from the solid. By measuring both the energy of the outgoing electron and its momentum, one can reconstruct from the knowledge of the energy and momentum of the initial photon, the energy and momentum of the electron inside the solid. The measured signal gives thus directly access to the probability $A(E,\boldsymbol{k})$ to find in the solid an electron with the energy E and the momentum $\boldsymbol{k}$. For free electrons this is simply a delta function $A(E,\boldsymbol{k}) = \delta(E - \varepsilon(\boldsymbol{k}))$. Since electrons can only be extracted if they are actually in the solid, this expression is limited to the occupied states and thus for free electrons simply cut by the Fermi function as shown in Fig. 1.12. More details and references on the photoemission technique can be found in [DHS03]. Of course integrating over energies gives the momentum distribution $n(\boldsymbol{k})$ and integrating over momenta give the probability $n(E)$ of finding an electron at energy E, which for independent electrons is simply the Fermi function $f_{\mathrm{F}}(E)$. For the interacting case we would expect again the electrons to be able to exchange an energy of the order of the interaction. The delta peak $\delta(E - \varepsilon(\boldsymbol{k}))$ should naively be broadened by ΔE which is the typical energy due to the interactions. Given the value of ΔE here, this would lead to extremely broad peaks as shown in Fig. 1.12. In the same way $n(E)$ should be extremely flat and loose its discontinuity.

As can be seen from the experimental data in Fig. 1.13 what really happens is completely different from these naive expectations. There are still quite sharp peaks that are visible in the spectral function. Moreover the width of the peaks seems to be getting smaller and smaller and the peaks sharper and sharper when one approaches the Fermi level. The bottom part of

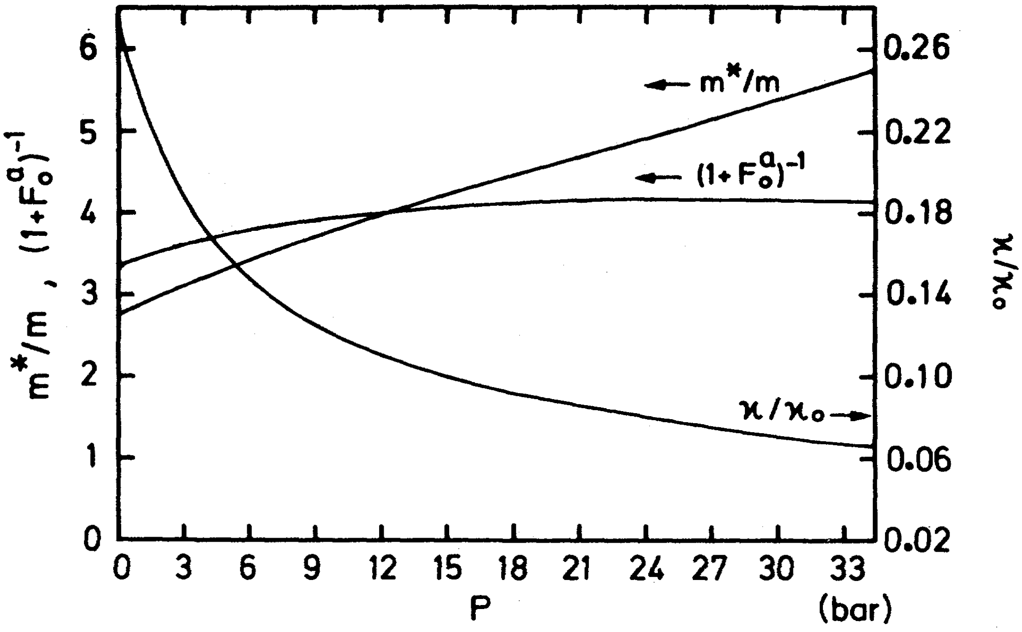


Figure 1.10: Effective mass m^* normalized to the bare mass m of the ^{3}He atom, as extracted from the specific heat measurements and as a function of pressure P. The notation $(1 + F_0^a)^{-1}$ is the ratio between the spin susceptibility and the effective mass. κ/κ_0 is the normalized (with respect to free particles) compressibility. One sees that although the basic properties of free fermions still work (specific heat linear in temperature, constant compressibility, constant susceptibility at low temperature), the coefficients characterizing these three quantities are three different numbers, dependent on the pressure, hence on the interactions.

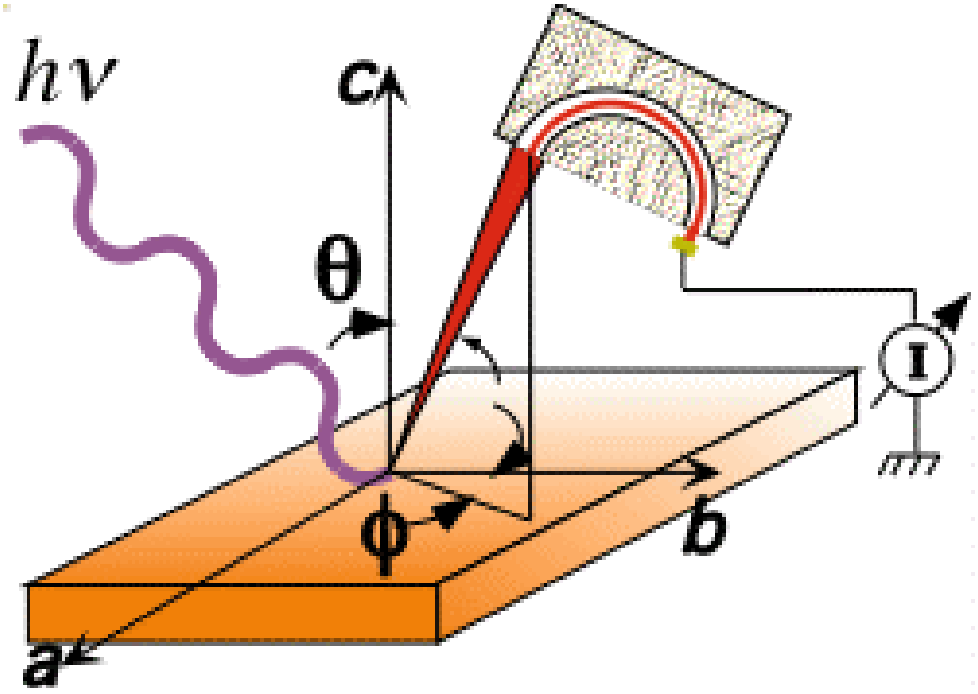


Figure 1.11: Cartoon of the photoemission technique. Photons are sent on a solid and extract electrons. The measurements of the momentum and energy of the extracted electrons allows to measure the spectral function, i.e. the probability of finding an single particle state with both a given momentum $\boldsymbol{k}$ and energy ω.

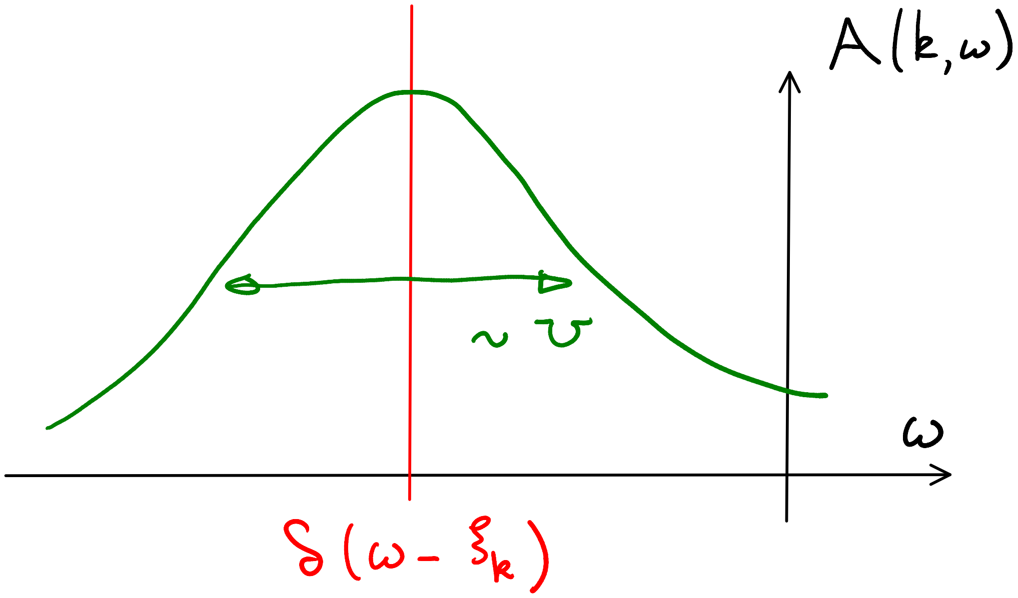


Figure 1.12: (red) Cartoon of the spectral function for independent electrons. Since for a free electron the energy for a given momentum is $\xi(\boldsymbol{k})$ the spectral function is a Dirac peak at $\omega = \xi(\boldsymbol{k})$. (green) Naively expected spectral function for interacting electrons. Since the electrons can exchange energies of the order of the interaction U one would naively expect the peak to be spread over energies of order U. Given the fact that the interaction is usually of the same order than the typical energies of the peak, this would mean that the peak is practically completely washed out.

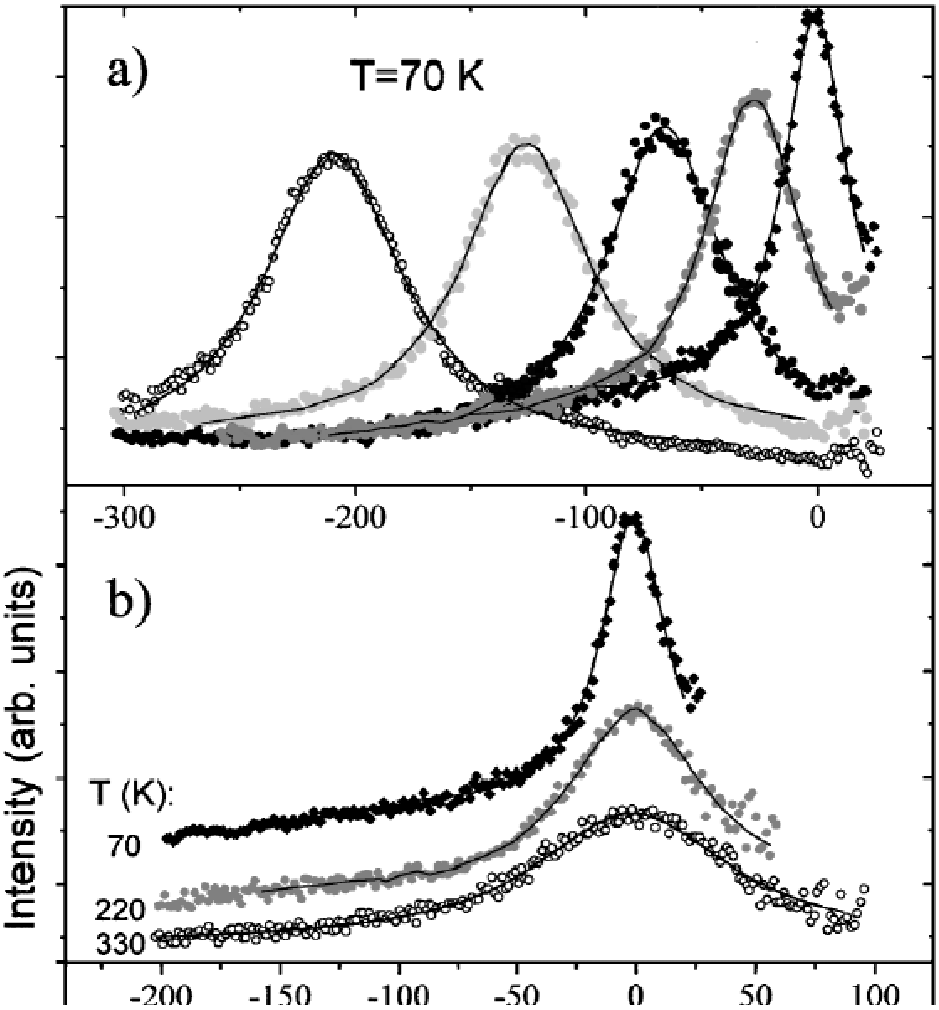


Figure 1.13: Photoemission data from a Mo (110) surface [VFJH99]. (Top) The spectral function $A(\boldsymbol{k}, \omega)$ is plotted as a function of ω. The zero denotes the Fermi level. Different peaks corresponds to different values of $\boldsymbol{k}$. One sees that the peaks in the spectral function are becoming, contrarily to naive expectations narrower and narrower and narrower. (Bot) Width of a peak close to the Fermi level as a function of the temperature T. One sees that the width of the peak is controlled in a large part by the temperature, which corresponds to energies several order of magnitude smaller than the typical energy of the interactions.

Fig. 1.13 shows that for peaks that are very close to the Fermi level the width of the peak is controlled by the temperature and not by the much larger energy that would correspond to the typical interaction energy. It just looks as if a certain “fraction” of the electrons was not seeing the interaction and was remaining free. This is of course consistent, even if unexplained, with the results of the compressibility, susceptibility and specific heat that are qualitatively, if not quantitatively in agreement with the free electron picture.

It is thus necessary to find the explanation for these mysterious facts.

CHAPTER 2

Linear response

Let us now see how we can compute observables for a quantum problem. Such observables correspond always to the average of some operator. In quantum mechanics the averages are taken in the ground state of the system. In solid state we always work at finite temperature so one has to generalize this concept to finite temperatures, and we will see how later.

Forgetting this question for the moment computing averages is easy if we know well enough the Hamiltonian to diagonalize it or at least if the system is in thermodynamic equilibrium. However this is often not the case: to probe a system one exert on it small perturbations such as the application of a small magnetic field to see how it magnetizes, a small voltage to see whether it conducts, etc. . One is thus very often faced with the situation of trying to study a problem which is described by an equilibrium (time independent) Hamiltonian H to which one will add in general a time dependent perturbation H_{pert}. Computing the full properties of the time dependent Hamiltonian would be a formidable task. However if the perturbation is small (in a sense to be defined) then one can hope to compute the observable in a development in the perturbing Hamiltonian.

This is what is called the linear response theory, and we will examine how one can make such a calculation for a quantum system.

2.1 Brief reminder of quantum mechanics

For a quantum mechanical system in a pure state $|\psi\rangle$, any observable can be measures by computing the average of the corresponding operator in the state $|\psi\rangle$

$$\mathcal{O} = \langle\psi|O|\psi\rangle \tag{2.1}$$

where here O is an operator that represents the observable we want to measure and $\mathcal{O}$ is the value (thus a number) corresponding to the result of the average of the operator. In what follows we will use the same symbol for the operators and the average value, the context making it clear whether one deals with an operator or with a number. If there is a possibility of confusion the average value will be denoted $\langle O\rangle$

A priori the function $|\psi\rangle$ can be time dependent which we will denote as $|\psi(t)\rangle$. If this is the case the average depends on time and this will be denoted by the various notations

$$O(t) = \langle O\rangle_t = \langle\psi(t)|O|\psi(t)\rangle \tag{2.2}$$

This is the standard Schrödinger representation. The operators are time independent and all the time dependence is put in the wave function that obeys the Schrödinger equation

$$i\partial_t\,|\psi\rangle = H\,|\psi\rangle \tag{2.3}$$

For and Hamiltonian that is not explicitly dependent on time, the equation has the formal solution

$$|\psi(t)\rangle = e^{-iHt} |\psi(t=0)\rangle \tag{2.4}$$

This allows for an alternative representation of the physical observables known as the Heisenberg representation. An observable at time t reads

$$O = \langle\psi(t)|O|\psi(t)\rangle = \langle\psi(t=0)|e^{iHt}Oe^{-iHt}|\psi(t=0)\rangle \tag{2.5}$$

One can thus consider that the wavefunctions are time independent and characterize the state of the system and that all the time dependence is due to *time dependent* operators. These operators are given by

$$O_H(t) = e^{iHt}O_S e^{-iHt} \tag{2.6}$$

where $O_H(t)$ and O_S denote respectively the operators in the Heisenberg and Schrödinger representation. The indices H and S are here added to emphasize the two representations. In the following, operators in the Schrödinger representation will be denoted without any special notation, and when an explicit time dependence will be noted for an operator it will mean that this is the operator in the Heisenberg representation.

The average of a physical quantity is thus given in the Heisenberg representation by

$$O(t) = \langle\psi_0|O_H(t)|\psi_0\rangle \tag{2.7}$$

The definition of the operators in the Heisenberg representation (2.6) can be rewritten in a different form.

$$\begin{aligned}\frac{dO(t)}{dt} &= (iH)e^{iHt}Oe^{-iHt} - e^{iHt}Oe^{-iHt}(iH) \\ &= i[H, O_H(t)]\end{aligned} \tag{2.8}$$

Note that the Hamiltonian is time independent both in the Schrödinger and Heisenberg representation $H_H = H_S$. This representation will be useful when we will deal with the second quantization in the next chapter.

For systems which are not in pure states the average is a superposition of the averages in pure states with the corresponding probabilities. Thus if p_i are the probabilities of finding the system in the state $|\psi_i\rangle$, a physical observable is given by

$$\langle O\rangle = \sum_i p_i\langle\psi_i|O|\psi_i\rangle \tag{2.9}$$

It is more convenient to introduce a density matrix that describes the system. The density matrix for the above average is given by

$$\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i| \tag{2.10}$$

and the average of the observable is now given generally by

$$\langle O\rangle = \mathrm{Tr}[\rho\, O] \tag{2.11}$$

where $\mathrm{Tr}[A]$ denotes the trace of the operator A (see Appendix A.4). Note that the density matrix can of course be time dependent if for example the functions $|\psi_i(t)\rangle$ are time dependent.

For a quantum system, with a time independent Hamiltonian H in equilibrium with a bath at temperature T, a very natural density matrix corresponds to a superposition state where each level $|n\rangle$ of energy E_n corresponding to the eigenstates of the Hamiltonian is occupied with a probability given by the Boltzmann factor

$$\rho = \sum_n \frac{e^{-\beta E_n}}{Z} |n\rangle\langle n| \tag{2.12}$$

where $Z = \sum_n e^{-\beta E_n}$ ensures that the probabilities are normalized, and $\beta = 1/(k_{\rm B}T)$ is the inverse temperature. It is easy to check that the density matrix (2.12) can be simply rewritten as

$$\rho = \frac{e^{-\beta H}}{Z} \tag{2.13}$$

with

$$Z = \mathrm{Tr}[e^{-\beta H}] \tag{2.14}$$

being the partition function of the system. All averages at finite temperature can thus be computed as

$$\langle O \rangle = \frac{\mathrm{Tr}[e^{-\beta H}\, O]}{\mathrm{Tr}[e^{-\beta H}]} \tag{2.15}$$

2.2 Linear response

Let us start with a system described by an Hamiltonian H_0 which is time independent and add to the Hamiltonian of the system a perturbation, a priori time and space dependent

$$H_{\rm pert} = \int d\boldsymbol{r} h(\boldsymbol{r}, t) O(\boldsymbol{r}) \tag{2.16}$$

where $h(\boldsymbol{r}, t)$ is some external field (magnetic field, electric field, pressure, etc.), and $O(\boldsymbol{r})$ the operator to which it couples (magnetization, current, density, etc.). We consider that this operator can depend on space. We choose O such that in the absence of perturbation its average is zero $\langle O(\boldsymbol{r})\rangle = 0$, since one can always subtract this average value. Since the Schrödinger equation give the time evolution of the quantum system, we are now in a position to compute the response to a *time dependent* perturbation. Let us emphasize that here H_0 does designate a totally general (interacting etc.) Hamiltonian, as long as this Hamiltonian does not contain an explicit time dependence. Typically H_0 is the full Hamiltonian of the system whose properties one is trying to probe by the perturbation.

Let us consider an observable described by an operator $A(\boldsymbol{r})$. As for O we choose the operator A such that its average in the unperturbed system vanishes. We want to compute the this observable at a given point $\boldsymbol{r}_0$ and at given time t_0. As can be readily guessed computing the full response for the complete Hamiltonian $H_0 + H_{\rm pert}$ is hopeless. However if the perturbation $h(\boldsymbol{r}, t)$ is small we can compute the average of A in a perturbation expansion in the perturbation. By definition of A the term of order zero vanishes. The most general term of order one (linear response) is of the form

$$A(\boldsymbol{r}_0, t_0) = \langle A(\boldsymbol{r}_0)\rangle_{t_0} \simeq \int d\boldsymbol{r} dt \chi(\boldsymbol{r}_0, \boldsymbol{r}; t_0, t) h(\boldsymbol{r}, t) \tag{2.17}$$

$\chi(\boldsymbol{r}_0, \boldsymbol{r}; t_0, t)$ is the susceptibility measuring the way the system responds to the perturbation $h(\boldsymbol{r}, t)$. Because we have performed an expansion in powers of h and confined to the linear term, χ depends *only* on the unperturbed Hamiltonian H_0. One can thus exploit the existing symmetries of H_0 to simplify the expression for χ. First H_0 does not explicitly depends on time, thus the susceptibility that measures the response between a perturbation at time t and a response at time t_0 can only depend on the difference between the two times. If in addition the Hamiltonian H_0 is also invariant by translation then the susceptibility will only depend on the space difference $\boldsymbol{r}_0 - \boldsymbol{r}$. The second point depends on the precise Hamiltonian H_0 but the first one is totally general.

We want now to explicitly compute the susceptibility χ. To do so we first need to know how to compute the average for a system which is time dependent since the full Hamiltonian $H(t) = H_0 + H_{\text{pert}}$ contains an *explicit* time dependence. Clearly we need to define what is the density matrix of the system at time t, to be able to use (2.11).

$$\langle A \rangle_t = \text{Tr}[\rho(t)A] \tag{2.18}$$

A naive possibility would be to assume that at each moment the density matrix is $\rho = e^{-\beta H(t)}$. In order to know if this is the correct let us look at the time evolution of the density matrix. Let us assume that at time $t = -\infty$ the perturbation vanishes $[h(\boldsymbol{r}, t \to -\infty) \to 0]$. In that case the system is fully determined by the time independent Hamiltonian H_0 and we know the density matrix which is simply given by $\rho_0 = e^{-\beta H_0}/Z_0$ or the expression (2.12). If now we switch on the perturbation the time evolution of the functions $|n\rangle$ are easy (formally) to compute since they simply obey the Schrödinger equation. However we have to decide how the *probabilities* p_n that the system can be found in the state $|n\rangle$ should evolve with time. We will assume that the time evolution is solely given by the evolution of the wavefunctions. This amounts to say that the coefficients c_n are not changing as the system is evolving with time, and thus the populations of the levels. The time evolution is thus supposed to be adiabatic. In other words the thermal bath is introduced at a given time when the perturbation does not exist (in particular $t = -\infty$) and the levels are populated according to the (time independent) distribution $e^{-\beta E_n}$. The thermal bath is then removed and the perturbation switched on slowly, so that the wavefunction evolves. Using the Schrödinger equation and (2.12) it is easy to show that

$$\frac{\partial \rho(t)}{\partial t} = -i[H(t), \rho(t)] \tag{2.19}$$

To obtain the linear response we consider that $\rho(t) = \rho_0 + f(t)$ where $f(t)$ is the part proportional to perturbation h. Keeping only the linear terms, (2.19) becomes

$$i\frac{\partial f(t)}{\partial t} = [H_0, \rho_0] + [H_0, f(t)] + [H_{\text{pert}}, \rho_0] \tag{2.20}$$

Since $\rho_0 = \frac{1}{Z_0}e^{-\beta H_0}$, one has $[H_0, \rho_0] = 0$. One can transform (2.20) into

$$e^{-iH_0 t}\left[i\frac{\partial}{\partial t}\left(e^{iH_0 t}f(t)e^{-iH_0 t}\right)\right]e^{iH_0 t} = [H_{\text{pert}}(t), \rho_0] \tag{2.21}$$

This expression becomes

$$i\frac{\partial}{\partial t}\left(e^{iH_0 t}f(t)e^{-iH_0 t}\right) = [e^{iH_0 t}H_{\text{pert}}(t)e^{-iH_0 t}, \rho_0] \tag{2.22}$$

Note that the time dependence in $H_{\text{pert}}(t)$ comes from the explicit dependence of $h(\boldsymbol{r}, t)$ in time. (2.22) is easily integrated, using the fact that the perturbation is absent at time $t = -\infty$ and thus $f(-\infty) = 0$ to give

$$f(t) = -ie^{-iH_0 t}\int_{-\infty}^{t} dt'[e^{iH_0 t'}H_{\text{pert}}(t')e^{-iH_0 t'}, \rho_0]e^{iH_0 t} \tag{2.23}$$

Using (2.18) we can now compute the average values

$$\langle A \rangle_t = \text{Tr}[(\rho_0 + f(t))A] \tag{2.24}$$

The term $\mathrm{Tr}[\rho_0 A]$ vanishes since we have taken an operator with zero average value in the absence of a perturbation. In that case

$$\begin{aligned}\langle A\rangle_t &= \mathrm{Tr}[Af(t)] \\ &= -i\,\mathrm{Tr}[\int_{-\infty}^{t} dt' e^{-iH_0 t}[e^{iH_0 t'} H_{\text{pert}}(t') e^{-iH_0 t'}, \rho_0] e^{iH_0 t} A]\end{aligned} \tag{2.25}$$

Using the cyclic invariance of the trace (A.18) and the explicit expression of H_{pert} one obtains

$$\begin{aligned}\langle A\rangle_t &= -i\,\mathrm{Tr}[\int_{-\infty}^{t} dt' [e^{iH_0 t'} H_{\text{pert}}(t') e^{-iH_0 t'}, \rho_0] e^{iH_0 t} A e^{-iH_0 t}] \\ &- i\,\mathrm{Tr}[\int_{-\infty}^{t} dt' [e^{iH_0 t'} \int d\boldsymbol{r}' h(\boldsymbol{r}', t') O(\boldsymbol{r}) e^{-iH_0 t'}, \rho_0] e^{iH_0 t} A e^{-iH_0 t}]\end{aligned} \tag{2.26}$$

One can now recognize the Heisenberg expression of the operators A and O in the above formula. We can thus rewrite it in the more compact form

$$\langle A\rangle_t = -i \int_{-\infty}^{t} dt' \int d\boldsymbol{r} h(\boldsymbol{r}', t')\,\mathrm{Tr}[[O(\boldsymbol{r}', t'), \rho_0] A(t)] \tag{2.27}$$

Of course the operator A can (and in general will) have an explicit spatial dependence in which case we compute $\langle A(\boldsymbol{r})\rangle_t$. Using

$$\mathrm{Tr}[[A, B]C] = \mathrm{Tr}[ABC - BAC] = \mathrm{Tr}[B[C, A]] \tag{2.28}$$

one can rewrite (2.26) as

$$\begin{aligned}\langle A(\boldsymbol{r})\rangle_t &= -i \int_{-\infty}^{t} dt' \int d\boldsymbol{r}'\,\mathrm{Tr}[\rho_0 [A(\boldsymbol{r}, t), O(\boldsymbol{r}', t')]] h(\boldsymbol{r}', t') \\ &= -i \int_{-\infty}^{t} dt' \int d\boldsymbol{r}' \langle [A(\boldsymbol{r}, t), O(\boldsymbol{r}', t')]\rangle_0 h(\boldsymbol{r}', t')\end{aligned} \tag{2.29}$$

where $\langle\rangle_0$ denotes averages taken with the unperturbed Hamiltonian H_0, and the time dependence of the operators coming from the Heisenberg representation is also computed with the Hamiltonian H_0. Comparing with the expression (2.17) one obtains for the susceptibility

$$\chi(\boldsymbol{r}, \boldsymbol{r}'; t, t') = -i\theta(t - t')\langle [A(\boldsymbol{r}, t), O(\boldsymbol{r}', t')]\rangle_0 \tag{2.30}$$

The θ function is due to the causality and expresses the fact that a measure at time t can only depend on the perturbation at anterior times. Such form of correlations that are non zero only when $t > t'$ are called retarded correlation functions. They directly correspond to physically observable quantities.

Rather than working in space time, it is better to go to Fourier space. Indeed since H_0 is time independent the Fourier transform over time of (2.17) will be diagonal. We also assume that H_0 is invariant by translation in space which ensures that the Fourier transform will be diagonal in momentum. One thus gets

$$\langle A(\boldsymbol{q}, \omega)\rangle = \chi(\boldsymbol{q}, \omega) h(\boldsymbol{q}, \omega) \tag{2.31}$$

with

$$\chi(\boldsymbol{q}, \omega) = \int d\boldsymbol{r} dt e^{-i(\boldsymbol{q}\boldsymbol{r} - \omega t)} \chi(\boldsymbol{r}, t) \tag{2.32}$$

Using (2.29) gives

$$\begin{aligned}\chi(\boldsymbol{q}, \omega) &= \int d\boldsymbol{r} \int dt \chi(\boldsymbol{r}, t) e^{-i(\boldsymbol{q}\boldsymbol{r} - \omega t)} \\ &= -i \int d\boldsymbol{r} \int_{0}^{+\infty} dt e^{-i(\boldsymbol{q}\boldsymbol{r} - \omega t)} \langle [A(\boldsymbol{r}, t), O(0, 0)]\rangle_0\end{aligned} \tag{2.33}$$

In order to perform the Fourier transform an important point has to be noticed. The response is defined (only time dependence is indicated)

$$\langle A\rangle_t = \int dt'\chi(t-t')h(t') \tag{2.34}$$

where $\chi(t-t')$ is the retarded correlation function. However, we also want the perturbation $h(t')$ to vanish at time $t' \to -\infty$ in order to be able to use the fact that we know the density matrix at $t' = -\infty$. Although of course one can implement directly this in the functions $h(t)$ one considers for perturbations, it is simpler to assume that the functions $h(t)$ are totally free (and in particular can be of the form $h(t) = e^{i\omega t}$, i.e. purely periodic signals) and to multiply $h(t)$ by a small convergence factor that ensures that the perturbation does indeed vanish. It is thus convenient to assume that the perturbation is of the form

$$h(t)e^{\delta t} \tag{2.35}$$

where $\delta = 0^+$ is an infinitesimal positive number, and $h(t)$ is now a function that does not necessarily needs to vanish at time $t \to -\infty$. As a consequence the linear response formula becomes

$$\begin{aligned}\langle A\rangle_t &= \int dt'\chi(t-t')h(t')e^{\delta t'} \\ &= e^{\delta t}\int dt'\chi(t-t')e^{\delta(t'-t)}h(t')\end{aligned} \tag{2.36}$$

where we have extracted the term $e^{\delta t}$ which shows explicitly that the response also vanish as it should when $t \to -\infty$. The consequence of (2.36) is that when one performs the Fourier transform of the retarded correlation function one should in fact, to ensure to proper vanishing of the perturbation at time $t' \to -\infty$ the Fourier transform of $\chi(t-t')e^{\delta(t'-t)}$. The formula (2.33) should thus be modified into (remember that we show here only the Fourier transform over time)

$$\chi(\omega) = \int_0^{+\infty} dt e^{i\omega t}e^{-\delta t}\langle[A(t), O(0)]\rangle_0 \tag{2.37}$$

The factor $e^{-\delta t}$ which is physically necessary to ensure that the perturbation vanishes in the distant past, is thus here able to ensure the convergence of the integral giving the Fourier transform of the susceptibility at large time even in cases when the correlation function $\langle[A(t), O(0)]\rangle_0$ would not decrease fast enough. Of course if this correlation decreases fast, then it simply means that one can take simply the limit $\delta \to 0^+$. But in general, and we will see examples below, it is necessary to always keep this factor in doing the integral and taking the limit $\delta \to 0^+$ in the end. This can lead to the presence of distributions in the expression of the Fourier transform. One can also view the presence of this convergence factor as if one was not doing the Fourier transform with a real frequency ω but with a frequency $\omega + i\delta$ containing an infinitesimal positive imaginary part.

2.3 Fluctuation dissipation theorem

To get a complete physical understanding of the meaning of the susceptibility χ let us examine the change of energy of the system. In the absence of an external perturbation the energy is conserved. This is not the case any more when the system is subject to a time dependent external potential, and some energy is injected in the system. The energy of the system at time t is given by

$$E(t) = \mathrm{Tr}[\rho(t)H(t)] \tag{2.38}$$

and thus the change of energy is

$$\frac{dE(t)}{dt} = \mathrm{Tr}[\rho(t)\frac{dH(t)}{dt}] + \mathrm{Tr}[\frac{d\rho(t)}{dt}H(t)] \tag{2.39}$$

Using the equation of evolution (2.19) for $\rho(t)$ one can rewrite the second term in (2.39) as

$$-i\,\mathrm{Tr}[[H(t),\rho(t)]H(t)] = -i\,\mathrm{Tr}[\rho(t)[H(t),H(t)] = 0 \tag{2.40}$$

using the cyclic invariance of the trace. Thus

$$\frac{dE(t)}{dt} = \mathrm{Tr}[\rho(t)\frac{dH(t)}{dt}] = \langle\frac{dH(t)}{dt}\rangle \tag{2.41}$$

Let us consider a simple sinusoidal perturbation of the form

$$H_{\text{pert}} = Ohe^{i\omega t} + O^\dagger h^* e^{-i\omega t} \tag{2.42}$$

In that case

$$\frac{dE(t)}{dt} = i\omega[\langle O\rangle_t he^{i\omega t} - \langle O^\dagger\rangle_t h^* e^{-i\omega t}] \tag{2.43}$$

Using linear response one has

$$\langle O(t)\rangle = \int dt' \chi_{OO}(t-t')he^{i\omega t'} + \chi_{OO^\dagger}(t-t')h^* e^{-i\omega t'} \tag{2.44}$$

Rather than compute the change in energy at a given time, since we deal here with a sinusoidal perturbation we can average over one period (we assume $\omega > 0$)

$$\overline{\frac{dE(t)}{dt}} = \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}=2\pi/\omega} dt \frac{dE(t)}{dt} \tag{2.45}$$

Using (2.43) and (2.44) one gets

$$\begin{aligned}\overline{\frac{dE(t)}{dt}} &= i\omega[\chi_{OO^\dagger}(\omega) - \chi_{O^\dagger O}(-\omega)]hh^* \\ &= \omega i[\chi_{OO^\dagger}(\omega) - \chi_{O^\dagger O}(-\omega)]hh^* \end{aligned} \tag{2.46}$$

Using the definition (2.33) one obtains

$$\begin{aligned}\chi(\omega)^* &= +i\int_0^{+\infty} dt \langle[O(t),O^\dagger(0)]\rangle^* e^{-i\omega t} \\ &= -i\int_0^{+\infty} dt\langle[O^\dagger(t),O(0)]\rangle e^{-i\omega t} \\ &= \chi(-\omega)\end{aligned} \tag{2.47}$$

as it should be for an hermitian operator. Thus (2.46) becomes

$$\begin{aligned}\overline{\frac{dE(t)}{dt}} &= \omega i hh^*[\chi_{OO^\dagger}(\omega) - \chi_{OO^\dagger}(\omega)^*] \\ &= -2\omega hh^*\,\mathrm{Im}\,\chi_{OO^\dagger}(\omega)\end{aligned} \tag{2.48}$$

Thus the imaginary part of the susceptibility controls the dissipation of energy in the system. (2.33) relates the response of the system to an external perturbation (and thus the dissipation of the energy brought by this perturbation) to a correlation function of the system in equilibrium. This relation is known as the fluctuation-dissipation theorem. It is a very important relation since it relies *only* on two very general assumptions: (i) that we restrict to linear response; (ii) more importantly that the system is in thermodynamic equilibrium. The fluctuation dissipation theorem is thus a very powerful tool since it quantify the fact that by slightly perturbing the system we can probe the various correlations and physical properties of the unperturbed system. In the opposite direction it provides a very practical and powerful way to compute the response of a system to an arbitrary time and space dependent perturbation as soon as we are able to compute the correlations of the system in equilibrium.

2.4 Spectral representation, Kramers-Kronig relations

The result (2.33) is very general. Let us now examine some of the properties of the retarded correlation function. To do so it is very convenient to introduce a formal decomposition known as the spectral representation. Let us again introduce a complete basis $|n\rangle$ constituted of the eigenstates of the Hamiltonian H

$$H_0 \, |n\rangle = E_n \, |n\rangle \tag{2.49}$$

It is important to realize that except for very simple Hamiltonians determining the eigenstates $|n\rangle$ and eigenvalues E_n is a formidable problem. In general we are thus not able to compute those explicitly, but in what follows we will simply use these quantities to derive formally a series of relations, and it is simply sufficient to know that such a basis exists without having to know it explicitly.

Let us first rewrite the retarded correlation

$$\chi(t) = -i\,\theta(t)\langle[A(t), O(0)]\rangle_0 \tag{2.50}$$

where the operators A and O can also depend on other quantum numbers such as the positions. We focuss here on the time dependence of the operators. It reads

$$\begin{aligned} \chi(t) &= -i\,\theta(t)\frac{1}{Z}\,\mathrm{Tr}[e^{-\beta H_0}[A(t), O(0)]] = \frac{1}{Z}\,\mathrm{Tr}[e^{-\beta H_0}(A(t)O(0) - O(0)A(t))] \\ &= -i\,\theta(t)\frac{1}{Z}\sum_n \langle n| e^{-\beta H_0}(A(t)O(0) - O(0)A(t))\,|n\rangle \\ &= -i\,\theta(t)\frac{1}{Z}\sum_{n,m} \langle n| e^{-\beta H_0} A(t)\,|m\rangle\,\langle m| O(0)\,|n\rangle - \langle n| e^{-\beta H_0} O(0)\,|m\rangle\,\langle m| A(t)\,|n\rangle \end{aligned} \tag{2.51}$$

where we have introduced the closure relation $1 = \sum_m |m\rangle\langle m|$. Using the definition of the Heisenberg operators (2.6) and the fact that $|n\rangle$ is an eigenstate of the Hamiltonian H_0, one gets

$$\chi(t) = -i\,\theta(t)\frac{1}{Z}\sum_{n,m} \langle n| e^{-\beta E_n} e^{i(E_n - E_m)t} A\,|m\rangle\,\langle m| O\,|n\rangle - e^{-\beta E_n}\langle n| O\,|m\rangle\, e^{i(E_m - E_n)t}\langle m| A\,|n\rangle \tag{2.52}$$

It is convenient to relabel n and m in the second term using the fact that one is summing over a complete basis, to get

$$\chi(t) = -i\,\theta(t)\frac{1}{Z}\sum_{n,m} e^{i(E_n - E_m)t}\langle n| A\,|m\rangle\,\langle m| O\,|n\rangle\,(e^{-\beta E_n} - e^{-\beta E_m}) \tag{2.53}$$

This allows to directly write the Fourier transform as

$$\begin{aligned}\chi(\omega) &= \int dt e^{i(\omega+i\delta)t}\chi(t) \\ &= -i\int_0^{+\infty} dt e^{i(\omega+i\delta)t}\frac{1}{Z}\sum_{n,m} e^{i(E_n-E_m)t}\langle n|A\,|m\rangle\,\langle m|O\,|n\rangle\,(e^{-\beta E_n}-e^{-\beta E_m}) \\ &= \frac{1}{Z}\sum_{n,m}\langle n|A\,|m\rangle\,\langle m|O\,|n\rangle\,\frac{e^{-\beta E_n}-e^{-\beta E_m}}{\omega+E_n-E_m+i\delta}\end{aligned} \tag{2.54}$$

The above expression is quite remarkable since the full time dependence has now been performed. We see that the frequency dependence consists in a series of poles, when the frequency is in resonance with an energy difference between two energy levels. This is particularly transparent on the imaginary part of the retarded correlation function. Using (A.10) one gets

$$\begin{aligned}\operatorname{Im}\chi(\omega) &= -\frac{\pi}{Z}\sum_{n,m}\langle n|A\,|m\rangle\,\langle m|O\,|n\rangle\,(e^{-\beta E_n}-e^{-\beta E_m})\delta(\omega+E_n-E_m) \\ &= -\frac{\pi}{Z}(1-e^{-\beta\omega})\sum_{n,m}\langle n|A\,|m\rangle\,\langle m|O\,|n\rangle\,e^{-\beta E_n}\delta(\omega+E_n-E_m)\end{aligned} \tag{2.55}$$

This last form becomes particularly transparent if we consider the absorption of energy in presence of a perturbation O, as seen in the previous section. In that case $A = O^\dagger$ and the expression becomes

$$\operatorname{Im}\chi(\omega) = -\frac{\pi}{Z}(1-e^{-\beta\omega})\sum_{n,m}|\langle m|O\,|n\rangle\,|^2 e^{-\beta E_n}\delta(\omega+E_n-E_m) \tag{2.56}$$

which has a very simple interpretation in terms of the Fermi golden rule. A transition to a perturbation with the frequency ω occurs when the system can absorb the quantum of energy $\hbar\omega$ to make a transition between two states. Using the Fermi golden rule we see that the probability of transition is simply given by

$$\sum_m |\langle m|O\,|n\rangle\,|^2\delta(\omega+E_n-E_m) \tag{2.57}$$

where as usual one has to sum over all possible final states the initial state $|n\rangle$ can make the transition to. The probability of transition is proportional to the square of the matrix element coupling the two states $|\langle m|O\,|n\rangle\,|^2$. The energy must be conserved in the transition which is ensured by the term $\delta(\omega + E_n - E_m)$. Finally because we are at finite temperature $|n\rangle$ is not just one state but we can make the transition from any possible initial state which will be occupied with a probability $e^{-\beta E_n}$ since $|n\rangle$ is an eigenstate of H with energy E_n. This leads back essentially to formula (2.56).

We see also quite generally that $\operatorname{Im}\chi(\omega) < 0$ for $\omega > 0$ and $\operatorname{Im}\chi(\omega) > 0$ for $\omega < 0$. This is mandatory since as we saw in the previous section $-\omega\operatorname{Im}\chi(\omega)$ is proportional to the energy absorbed by the system in an oscillatory field. We also see directly that for $\omega = 0$ one has $\operatorname{Im}\chi(\omega = 0) = 0$ and thus in particular no absorption or dissipation of energy is possible with a static field.

The expression (2.54) allows to immediately generalize the correlation function to any complex number z by

$$\chi(z) = \frac{1}{Z}\sum_{n,m}\langle n|A\,|m\rangle\,\langle m|O\,|n\rangle\,\frac{e^{-\beta E_n}-e^{-\beta E_m}}{z+E_n-E_m} \tag{2.58}$$

As can be seen this function is an analytic function for any z not on the real axis. Quite generally it will have a cut on the real axis. The (physical) retarded correlation is obtained by $\chi_{\mathrm{ret}}(\omega) = \chi(z \to \omega + i\delta)$ and thus looking at the function $\chi(z)$ just above the real axis.

The comparison between (2.58) and (2.55) allows to immediately show the very general and remarkable relation

$$\chi(z) = \frac{-1}{\pi} \int d\omega' \frac{1}{z - \omega'} \operatorname{Im} \chi_{\text{ret}}(\omega') \tag{2.59}$$

This equation shows that the correlation function in the *whole* complex plane is totally determined by the value of the imaginary part of the retarded correlation. In particular it means that the retarded correlation, obtained by replacing $z \to \omega + i\delta$ obeys

$$\chi_{\text{ret}}(\omega) = \frac{-1}{\pi} \int d\omega' \frac{1}{\omega - \omega' + i\delta} \operatorname{Im} \chi_{\text{ret}}(\omega') \tag{2.60}$$

Using (A.10) one sees that (2.60) gives and identity for the imaginary part and gives, for the real part

$$\operatorname{Re} \chi_{\text{ret}}(\omega) = \frac{-1}{\pi} \int d\omega' \mathcal{P} \left(\frac{1}{\omega - \omega'} \right) \operatorname{Im} \chi_{\text{ret}}(\omega') \tag{2.61}$$

This relation that related the real and imaginary parts of the response function is known as the Kramers-Kronig relation. As we saw it is fully general ((2.59) is an even more useful and compact form of it), and is the direct consequence of causality of the response function.

One was to physically understand this remarkable constraint on the response function and the existence of the Kramers-Kronig relation is to realize that because of the causality the response function obeys the identity

$$\chi_{\text{ret}}(t) = f(t) \chi_{\text{ret}}(t) \tag{2.62}$$

where $f(t)$ is *any* function of time that obeys $f(t > 0) = 1$. This seemingly trivial point gives an enormous constraint on the Fourier transform, which is now a convolution product of the two functions

$$\chi_{\text{ret}}(\omega) = \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} f(\omega - \omega') \chi_{\text{ret}}(\omega') \tag{2.63}$$

In particular one can choose for the function $f(t)$ the step function $f(t) = \theta(t)$. Its Fourier transform is

$$f(\omega + i\delta) = \int dt\, \theta(t) e^{i(\omega + i\delta)t} = \int_0^{+\infty} e^{i(\omega + i\delta)t} = \frac{-1}{i(\omega + i\delta)} \tag{2.64}$$

Injecting in (2.63) one obtains

$$\chi_{\text{ret}}(\omega) = \int \frac{d\omega'}{2\pi} \frac{-1/i}{\omega - \omega' + i\delta} \chi_{\text{ret}}(\omega') \tag{2.65}$$

Which using (A.10) again leads to

$$\chi_{\text{ret}}(\omega) = \frac{-1}{i\pi} \int d\omega' \mathcal{P} \left(\frac{1}{\omega - \omega'} \right) \chi_{\text{ret}}(\omega') \tag{2.66}$$

which upon separation of the real and imaginary parts leads directly to the Kramers-Kronig relations.

CHAPTER 3

Second quantization

3.1 Why not Schrödinger

We want to deal with the case of several quantum particles. Normally this is a problem we know very well how to treat in quantum mechanics. If we know the Hilbert space H_1 of a single particle and a complete basis $|\alpha\rangle$, we know that for $\mathcal{N}$ particles we have a Hilbert space which is

$$H_{\mathcal{N}} = \bigotimes_{i=1}^{\mathcal{N}} H_i \tag{3.1}$$

and that a complete basis of such a space is simply

$$|\alpha, \beta, \ldots, \omega) = |\alpha\rangle \otimes |\beta\rangle \cdots |\gamma\rangle \tag{3.2}$$

We then have to solve a Schrödinger equation with a wavefunction depending on the $\mathcal{N}$ variables corresponding to the $\mathcal{N}$ particles $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \ldots, \boldsymbol{r}_{\mathcal{N}})$ for example.

Although this is a perfectly acceptable program when the number of particles is small it is particularly ill suited to tackle the case of many interacting quantum fermions or bosons, for several reasons.

The first one has to do with the indiscernibility of the particles. *Even* if the particles are free, not all states are acceptable for the wavefunctions of $\mathcal{N}$ indiscernible particles. Indeed only the totally symmetric (for bosons) or antisymmetric (for fermions) wavefunctions are allowed. It means that even for noninteracting particles we cannot directly use wavefunctions of the form (3.2). In a way the fact that we have to deal with indistinguishable particles already introduces correlations in the wavefunction even when interactions are not present in the systems.

The wavefunction become rather complicated. They should be properly (anti-)symmetrized and normalized which make them very heavy to use. For example for two particles one has

$$|\alpha, \beta\rangle = \frac{1}{\sqrt{2}} \left[|\alpha, \beta) \pm |\beta, \alpha) \right] \tag{3.3}$$

where the + sign is for bosons and the − one for fermions. In the above formula we denote by $|\alpha, \beta\rangle$ properly symmetrized and normalized kets for indistinguishable particles. We denote the *ordered* kets by

$$|\alpha, \beta) = |\alpha) \otimes |\beta) \tag{3.4}$$

where the first particles is in the state α and the second β. We will always (regardless of whether this is a bra or a ket) number the particles from left to right.

When the number of particle grows the need to use a symmetrized wavefunction becomes more and more heavy since the number of terms in the wavefunction grows as $\mathcal{N}!$. Indeed the wavefunction reads in general

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \ldots, \boldsymbol{r}_{\mathcal{N}}) = C \sum_P (\pm 1)^{s(P)} \psi_1(\boldsymbol{r}_{P(1)}) \psi_2(\boldsymbol{r}_{P(2)}) \cdots \psi_{\mathcal{N}}(\boldsymbol{r}_{P(\mathcal{N})}) \tag{3.5}$$

where P are the permutations of the set of numbers from 1 to $\mathcal{N}$, $s(P)$ is the signature of the permutation P (i.e. the number of transpositions in P) and as usual the $+$ sign is for bosons and the $-$ one for fermions. The constant C has to be determined such that the wavefunction is normalized. For fermions one can rewrite the wavefunction as a determinant (known as a Slater determinant)

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \ldots, \boldsymbol{r}_{\mathcal{N}}) = \begin{vmatrix} \psi_1(\boldsymbol{r}_1) & \cdots & \psi_1(\boldsymbol{r}_{\mathcal{N}}) \\ \vdots & & \vdots \\ \psi_{\mathcal{N}}(\boldsymbol{r}_1) & \cdots & \psi_{\mathcal{N}}(\boldsymbol{r}_{\mathcal{N}}) \end{vmatrix} \tag{3.6}$$

which helps a little bit for practical computations, but not very much. In fine, one would have to deal even for noninteracting electrons with wavefunctions containing $10^{23}!$ terms, which is really unpleasant.

The second problem is linked to the way we represent the operators in the standard expression of quantum mechanics. If we consider for example an operator measuring the total momenta of the particles, it has to be a sum of operators acting on each particles individually. This means that such an operator would write as

$$P_{\text{tot}} = \sum_{i=1}^{\mathcal{N}} P_i \tag{3.7}$$

where P_i is the operator acting on the i-th particle. Note that this is an abuse of notation since in fact one should note that

$$P_i = 1 \otimes 1 \otimes \ldots P \otimes \ldots \otimes 1 \tag{3.8}$$

where 1 is the identity and P is inserted at the i-th position. The operator and the wavefunctions thus depend *explicitly* on the particle number. One should thus completely change the whole calculation depending on whether we look at 2 or 20000 particles, which is again singularly unpleasant. It also prevent to take in a straightforward manner the thermodynamic limit $\mathcal{N} \to \infty$ when the volume of the systems is also going to infinity. Given the very large number of particles it is clear that taking this limit is highly desirable and will simplify a lot the calculations.

For these reasons the standard quantum mechanical representation (also known as first quantization) of systems of several indistinguishable particles is not very well adapted. One would like a system that takes care automatically of the following points

1. The antisymmetrization is taken care of in a painless manner without having to explicitly having to deal with the $\mathcal{N}!$ terms.

2. The way to describe the system is not explicitly dependent on the number of particles present in the system. This should allow to take the thermodynamic limit easily and also to deal to more general situations where the number of particles can change (such as e.g. in a photoemission experiment).

This is provided by the so-called “second quantization” method that we will now describe.

Figure 3.1: The state of a system with and arbitrary number of indistinguishable particles is fully known if one knows how many particles are in a given quantum state.

3.2 Fock space

The idea is to turn the fact that the particles are indiscernible into an advantage. Indeed, if this is the case it means that one does not need to know the quantum state of *each* individual particle, but simply how many particles are in a given quantum state. Let us assume that one has a complete basis $|\alpha\rangle$ of states for a single particle. Quite generally this basis is infinite, but if the system is put into a box one can quantize the states (for example by quantizing the momentum) and have a finite number of states (generally proportional to the volume of the system Ω). Let us thus denote all the states in this basis as

$$|\alpha_1\rangle, |\alpha_2\rangle, \ldots, |\alpha_\Omega\rangle \tag{3.9}$$

where we have used the notation $|\alpha_\Omega\rangle$ for the last state to remind that in general the total number of states in this basis is growing with the volume of the system. Note that the size of the basis is unrelated to the *number* of particles that are present in the system. For bosons for example one could have a complete basis of the one particle states that contain two states only and have 10000 bosons present in the system (since several of them can go in the same quantum state). For fermions of course the total number of particles is always smaller than the total number of available states because of the Pauli principle. As shown in Fig. 3.1 we can fully describe the system and reconstruct its wavefunction if we know the number of particles in a given state $|\alpha_i\rangle$ of the complete basis of single particle states. We can thus fully characterize the wave function of the system by the set of numbers $n_1, n_2, \ldots, n_\Omega$. The total number of particles in the system is of course $\mathcal{N} = n_1 + n_2 + \cdots + n_\Omega$, and can vary if one varies one of the n_i.

Let us thus define a space in which an arbitrary number of particles can exist. If we call $H_\mathcal{N}$ the Hilbert space with $\mathcal{N}$ particles ($H_k = \bigotimes_{j=1}^{k} H_j$), we can define

$$\mathfrak{F} = \bigoplus_{j=0}^{+\infty} H_j \tag{3.10}$$

which is the direct sum of all Hilbert spaces with 0, 1, 2, etc. particles. Such a space is called the Fock space. In this space let us now define the state

$$|n_1, n_2, n_3, \ldots, n_\Omega\rangle \tag{3.11}$$

where as shown in Fig. 3.1, the n_i are the number of particles in the quantum state $|\alpha_i\rangle$. We can fully give the expression of the wavefunction corresponding to the state (3.11). If we call

$\boldsymbol{r}_1$, $\boldsymbol{r}_2$, ..., $\boldsymbol{r}_{\mathcal{N}}$ the positions of each particles (keep in mind that $\mathcal{N} = n_1 + n_2 + \cdots + n_\Omega$) we have

$$\begin{aligned}\langle \boldsymbol{r}_1 \boldsymbol{r}_2, \ldots, \boldsymbol{r}_{\mathcal{N}} | n_1, n_2, \ldots, n_\Omega \rangle = \frac{\sqrt{\mathcal{N}!}}{\sqrt{n_1!} \ldots \sqrt{n_\Omega!}} S_\pm \Bigg[& \overbrace{\alpha_1(\boldsymbol{r}_1)\alpha_2(\boldsymbol{r}_2)\cdots\alpha_2(\boldsymbol{r}_{n_1})}^{n_1 \text{terms}} \\ & \overbrace{\alpha_2(\boldsymbol{r}_{n_1+1})\alpha_2(\boldsymbol{r}_{n_1+2})\cdots\alpha_2(\boldsymbol{r}_{n_1+n_2})}^{n_2 \text{terms}} \ldots \\ & \overbrace{\alpha_\Omega(\boldsymbol{r}_{n_1+\ldots+n_{\Omega-1}+1})\alpha_2(\boldsymbol{r}_{n_1+\ldots+n_{\Omega-1}+2})\cdots\alpha_2(\boldsymbol{r}_{n_1+\ldots+n_{\Omega-1}+n_\Omega})}^{n_\Omega \text{terms}} \Bigg]\end{aligned} \tag{3.12}$$

where the prefactors ensure that the wavefunction is suitably normalized. The (anti-)symmetrizer $S_\pm$ is defined as

$$S_\pm \left[\phi_1(\boldsymbol{r}_1) \ldots \phi_{\mathcal{N}}(\boldsymbol{r}_{\mathcal{N}}) \right] = \frac{1}{\mathcal{N}!} \sum_P (\pm 1)^{s(P)} \psi_1(\boldsymbol{r}_{P(1)}) \psi_2(\boldsymbol{r}_{P(2)}) \cdots \psi_{\mathcal{N}}(\boldsymbol{r}_{P(\mathcal{N})}) \tag{3.13}$$

Two states (3.11) having a different number of particles $\mathcal{N}$ belong to two different Hilbert spaces and are thus obviously orthogonal in the Fock space. For systems with the same total number of particles one can check by using the wavefunction (3.12) that the states (3.11) for an orthogonal and normalized basis

$$\langle n_1, n_2, \ldots, n_\Omega | n'_1, n'_2, \ldots, n'_\Omega \rangle = \delta_{n_1, n'_1} \delta_{n_2, n'_2} \cdots \delta_{n_\Omega, n'_\Omega} \tag{3.14}$$

We can thus use the basis (3.11) to characterize every operator and matrix element in the Fock space. As mentioned before this basis is extremely convenient since it relies on the minimal amount of information needed to describe a system of indistinguishable particles. In particular the number of "counters", n_i needed does not grow with the total number of particles.

3.3 Creation and destruction operators

Let us introduce operators that will allow us to generate all the elements of the above mentioned basis. For each state α_i of the single particle complete basis, we define a creation and destruction operator that will increase or decrease by one the number of particles in this particular state. We will thus be able to use these operators to modify the counter n_i giving the number of particles in a given quantum state, and thus span the whole Fock space. The practical definition of these operators is different depending on the statistics of the particles.

3.3.1 Bosons

Let us define the creation $a_i^\dagger$ and destruction a_i operators by their action on all the states of a complete basis in the Fock space

$$\begin{aligned} a_i^\dagger \left| n_1, \ldots, n_i, \ldots, n_\Omega \right\rangle &= \sqrt{n_i + 1} \left| n_1, \ldots, n_i + 1, \ldots, n_\Omega \right\rangle \\ a_i \left| n_1, \ldots, n_i, \ldots, n_\Omega \right\rangle &= \sqrt{n_i} \left| n_1, \ldots, n_i - 1, \ldots, n_\Omega \right\rangle \end{aligned} \tag{3.15}$$

These definitions completely define the operators by their matrix elements between all the elements of a complete basis. Let us check that the operators $a_i^\dagger$ and a_i are indeed hermitian conjugate. The only non-zero matrix element for $a_i^\dagger$ is

$$\langle n_1, \ldots, n_i + 1, \ldots, n_\Omega | a_i^\dagger | n_1, \ldots, n_i, \ldots, n_\Omega \rangle = \sqrt{n_i + 1} \tag{3.16}$$

Taking the complex conjugate of the above expression one thus gets

$$\langle n_1, \ldots, n_i, \ldots, n_\Omega | a_i | n_1, \ldots, n_i + 1, \ldots, n_\Omega \rangle = \sqrt{n_i + 1} \tag{3.17}$$

which is indeed exactly the definition of the operator a_i in (3.15) (with the replacement of n_i by $n_i + 1$). Another important properties of the operators, is that they only span the Fock space. Indeed although it seems formally from (3.15) that the operator a_i could operate on a state that has $n_i = 0$ particles in the state α_i the prefactor in the definition ensures that the corresponding matrix element is zero

$$a_i \, |n_1, \ldots, n_i = 0, \ldots, n_\Omega \rangle = 0 \tag{3.18}$$

and thus if one tries to apply the destruction operator on a state that has no particle in the corresponding quantum state one cannot generate unphysical states.

If we define the state that contain no particles in any of the quantum states (sometimes referred to as the vacuum)

$$|\varnothing\rangle = |n_1 = 0, n_2 = 0, \ldots, n_\Omega = 0\rangle \tag{3.19}$$

it is easy to see that from this vacuum $|\varnothing\rangle$ and the operators $a_i^\dagger$ we can construct all the vectors of the complete basis of the Fock space, since

$$|n_1, \ldots, n_i, \ldots, n_\Omega \rangle = \frac{(a_1^\dagger)^{n_1} \ldots (a_\Omega^\dagger)^{n_\Omega}}{\sqrt{n_1!} \ldots \sqrt{n_\Omega!}} \, |\varnothing\rangle \tag{3.20}$$

Thus one can completely describe the Fock space from the single state $|\varnothing\rangle$ and the creation (and destruction since they are hermitian conjugate) operators. Note that the vacuum verifies the property that for any i

$$a_i \, |\varnothing\rangle = 0 \tag{3.21}$$

Note also that one should not mix up the vacuum $|\varnothing\rangle$ which is a vector of the Fock space, and one which operators can act to give other states of the Fock space with 0.

The creation and destruction operators constitute thus a very convenient way of describing the Fock space. Rather than defining them from their matrix elements in a given basis such as (3.15), it is more convenient do define them from an intrinsic properties. We will show that the definition (3.15) implies that the operators $a_i^\dagger$ and a_i have certain specific commutation relations. Conversely if these commutation relations are obeyed, then the corresponding operators, and the corresponding vacuum, defined by (3.19), will define a Fock space from (3.20) in which they will have the matrix elements (3.15).

Let us first look at (here $i \neq j$)

$$\begin{aligned} a_i^\dagger a_j^\dagger \, |n_1, \ldots, n_i, \ldots, n_j, \ldots, n_\Omega \rangle &= a_i^\dagger \sqrt{n_j + 1} \, |n_1, \ldots, n_i, \ldots, n_j + 1, \ldots, n_\Omega \rangle \\ &= \sqrt{n_i + 1}\sqrt{n_j + 1} \, |n_1, \ldots, n_i + 1, \ldots, n_j + 1, \ldots, n_\Omega \rangle \end{aligned} \tag{3.22}$$

It is easy to see that the action of $a_j^\dagger a_i^\dagger$ would produce exactly the same result. Thus for any element of the basis one has

$$[a_i^\dagger, a_j^\dagger] \, |n_1, \ldots, n_i, \ldots, n_j, \ldots, n_\Omega \rangle = 0 \tag{3.23}$$

which means that

$$[a_i^\dagger, a_j^\dagger] = 0 \tag{3.24}$$

Since an operator commutes with itself this is also true when $i = j$. The Hermitian conjugation of (3.24) implies that

$$[a_i, a_j] = 0 \tag{3.25}$$

Let us now look at $(i \neq j)$

$$\begin{aligned} a_i^\dagger a_j \left|n_1, \ldots, n_i, \ldots, n_j, \ldots, n_\Omega\right\rangle &= a_i^\dagger \sqrt{n_j} \left|n_1, \ldots, n_i, \ldots, n_j - 1, \ldots, n_\Omega\right\rangle \\ &= \sqrt{n_i + 1}\sqrt{n_j} \left|n_1, \ldots, n_i + 1, \ldots, n_j - 1, \ldots, n_\Omega\right\rangle \end{aligned} \tag{3.26}$$

and in a similar way the action of $a_j a_i^\dagger$ (with $i \neq j$) would give the same result. One has thus $[a_i^\dagger, a_j] = 0$ when $i \neq j$. The case $i = j$ is special. One has

$$\begin{aligned} a_i^\dagger a_i \left|n_1, \ldots, n_i, \ldots, n_\Omega\right\rangle &= a_i^\dagger \sqrt{n_i} \left|n_1, \ldots, n_i - 1, \ldots, n_\Omega\right\rangle \\ &= \sqrt{(n_i - 1) + 1}\sqrt{n_i} \left|n_1, \ldots, n_i, \ldots, n_\Omega\right\rangle \\ &= n_i \left|n_1, \ldots, n_i, \ldots, n_\Omega\right\rangle \end{aligned} \tag{3.27}$$

On the other hand

$$\begin{aligned} a_i a_i^\dagger \left|n_1, \ldots, n_i, \ldots, n_\Omega\right\rangle &= a_i \sqrt{n_i + 1} \left|n_1, \ldots, n_i + 1, \ldots, n_\Omega\right\rangle \\ &= \sqrt{n_i + 1}\sqrt{n_i + 1} \left|n_1, \ldots, n_i, \ldots, n_\Omega\right\rangle \\ &= (n_i + 1) \left|n_1, \ldots, n_i, \ldots, n_\Omega\right\rangle \end{aligned} \tag{3.28}$$

and thus one has

$$[a_i, a_i^\dagger] \left|n_1, \ldots, n_i, \ldots, n_\Omega\right\rangle = \left|n_1, \ldots, n_i, \ldots, n_\Omega\right\rangle \tag{3.29}$$

and thus summarizing the two results

$$[a_i, a_j^\dagger] = \delta_{i,j} \tag{3.30}$$

One can thus summarize the properties of the creation and destruction operators by

$$\begin{aligned} [a_i, a_j^\dagger] &= \delta_{i,j} \\ [a_i^\dagger, a_j^\dagger] &= 0 \\ [a_i, a_j] &= 0 \end{aligned} \tag{3.31}$$

These are the fundamental properties that the operators must obey. They are equivalent to the definition of the matrix elements (3.15).

It means that if we give ourselves

1. A complete basis α_i of the single particle states (and the corresponding wavefunctions $|\alpha_i\rangle$).
2. Creation and destruction operators a_i, for each one of these states that obey the canonical commutation relations (3.31).
3. A vacuum $|\varnothing\rangle$ that is destroyed by the destruction operators $a_i |\varnothing\rangle = 0$

we can fully construct a Fock space for bosons. The idea is thus to directly exploit the above properties and use the canonical commutation relations between the bosons to compute the physical properties. Representing properties using the above ingredients rather than the wavefunctions is known as the second quantization.

Let us give some examples. Let us take for example the following function

$$|\psi\rangle = a_1^\dagger a_2^\dagger |\varnothing\rangle \tag{3.32}$$

Obviously the function contains two bosons. One can reconstruct the expression of the wavefunction using (3.12). One obtains

$$\langle \boldsymbol{r}_1 \boldsymbol{r}_2 | \psi\rangle = \frac{1}{\sqrt{2}}[\alpha_1(\boldsymbol{r}_1)\alpha_2(\boldsymbol{r}_2) + \alpha_1(\boldsymbol{r}_2)\alpha_2(\boldsymbol{r}_1)] \tag{3.33}$$

where $\alpha_i(\boldsymbol{r}_i) = \langle \boldsymbol{r}_i|\alpha_i\rangle$. This is the properly symmetrized function describing two bosons. The interest of the second quantization however is to stick with the operators and avoid coming back to the quite untractable wavefunctions. We see that we can deduce many things directly from the commutation relations. For example since $a_1^\dagger a_2^\dagger = a_2^\dagger a_1^\dagger$ from the commutation relation we see that

$$a_1^\dagger a_2^\dagger \,|\varnothing\rangle = a_2^\dagger a_1^\dagger \,|\varnothing\rangle \tag{3.34}$$

and thus the wavefunction $|\psi\rangle$ is obviously symmetric by permutation of the particles. The creation and destruction operators are thus directly ingeniered to take properly care of the symmetrization of the wavefunctions and the indiscernibility of the particles. One can in fact extract a lot more directly from the commutation relations. In particular averages can be computed easily directly.

Let us illustrate it by looking at the normalization of the function $|\psi\rangle$. We want to compute

$$\langle\psi|\psi\rangle = \langle\varnothing|a_2 a_1 a_1^\dagger a_2^\dagger|\varnothing\rangle \tag{3.35}$$

although this is here a specific example, we will see that quite generally all physical observables reduce to the average in the vacuum of a certain product of creation and destruction operators, so the method illustrated here can be applied generally. In order to compute the average, the only thing we need to use is the fact that the vacuum is destroyed by the a_i. We should thus bring back, using the commutation relations the operator a_i to act on the vacuum. Here we use $a_1 a_1^\dagger = 1 + a_1^\dagger a_1$ from the commutation relation. One has thus

$$\begin{aligned}\langle\psi|\psi\rangle &= \langle\varnothing|a_2(1 + a_1^\dagger a_1)a_2^\dagger|\varnothing\rangle \\ &= \langle\varnothing|a_2 a_2^\dagger|\varnothing\rangle + \langle\varnothing|a_2 a_1^\dagger a_1 a_2^\dagger|\varnothing\rangle\end{aligned} \tag{3.36}$$

The second term is simple. We can use the commutation relation $a_1 a_2^\dagger = a_2^\dagger a_1$ to rewrite the second term as $\langle\varnothing|a_2 a_1^\dagger a_2^\dagger a_1|\varnothing\rangle$ which immediately gives zero. $\langle\psi|\psi\rangle$ is thus only given by the first term. Using again the commutation relations one obtains

$$\begin{aligned}\langle\psi|\psi\rangle &= \langle\varnothing|a_2 a_2^\dagger|\varnothing\rangle = \langle\varnothing|(1 + a_2^\dagger a_2)|\varnothing\rangle \\ &= \langle\varnothing|1|\varnothing\rangle \\ &= 1\end{aligned} \tag{3.37}$$

Although the calculations can become tedious when the number of operator grows, the mechanics is always the same. With a little bit of habit one can speed up the calculations as will be seen in the exercices.

3.3.2 Fermions

Let us now turn to the creation and destruction operators for Fermions. In a similar way than for bosons (3.15) we define

$$\begin{aligned}c_i^\dagger \,|n_1, \ldots, n_i, \ldots, n_\Omega\rangle &= (1 - n_i)(-1)^{\epsilon_i} \,|n_1, \ldots, n_i + 1, \ldots, n_\Omega\rangle \\ c_i \,|n_1, \ldots, n_i, \ldots, n_\Omega\rangle &= n_i(-1)^{\epsilon_i} \,|n_1, \ldots, n_i - 1, \ldots, n_\Omega\rangle\end{aligned} \tag{3.38}$$

where $\epsilon_i = \sum_{j=1}^{i-1} n_j$ and $\epsilon_1 = 0$. Note that the order of the elements in the basis, has to be fixed once and for all, but is of course arbitrary.

In the above definitions some terms are quite transparent. Since for Fermions the Pauli principle prevents two fermions to be in the same quantum number it means that the n_i are restricted to be either 0 or 1. It is thus important that the creation operator does is not able to create two particles in one state. This is ensured by the factor $1 - n_i$ that ensures that if $c_i^\dagger$ acts on a state with $n_i = 1$ then the action of the operator will give zero. In a similar way the n_i factor ensures that the destruction operator cannot destroy a particle in the state for which $n_i = 0$. The physics of the strange factor $(-1)^{\epsilon_i}$ is for the moment not obvious, and one could have been tempted to define the operators without such a phase factor. We will see its use a little bit later.

We now proceed exactly as with the bosons. We first can check that the operators $c_i^\dagger$ and c_i are indeed hermitian conjugate. Fermions are in fact simpler in that respect, since for each state α_i there are only two possibilities $n_i = 0$ or $n_i = 1$ for the corresponding state. The only non zero matrix element for the operator $c_i^\dagger$ is

$$\langle n_1, \ldots, n_i = 1, \ldots, n_\Omega | c_i^\dagger | n_1, \ldots, n_i = 0, \ldots, n_\Omega \rangle = (-1)^{\epsilon_i} \tag{3.39}$$

while for c_i the only non zero matrix element is

$$\langle n_1, \ldots, n_i = 0, \ldots, n_\Omega | c_i | n_1, \ldots, n_i = 1, \ldots, n_\Omega \rangle = (-1)^{\epsilon_i} \tag{3.40}$$

which is obviously the complex conjugate of the other one.

In order to proceed with the commutations relations, and understand the role of the coefficients $(-1)^{\epsilon_i}$ let us first look at the action of $c_i c_i^\dagger$. Since this only affects state α_i we can simply consider the action on the two states with $n_i = 0$ and $n_i = 1$

$$\begin{aligned} c_i c_i^\dagger \, |n_1, \ldots, n_i = 0, \ldots, n_\Omega\rangle &= (-1)^{\epsilon_i} c_i \, |n_1, \ldots, n_i = 1, \ldots, n_\Omega\rangle \\ &= (-1)^{2\epsilon_i} \, |n_1, \ldots, n_i = 0, \ldots, n_\Omega\rangle \\ &= |n_1, \ldots, n_i = 0, \ldots, n_\Omega\rangle \end{aligned} \tag{3.41}$$

On the contrary

$$c_i^\dagger c_i \, |n_1, \ldots, n_i = 0, \ldots, n_\Omega\rangle = 0 \tag{3.42}$$

Notice that in this result the terms $(-1)^{\epsilon_i}$ do not play any role, and we could have defined the operators without including them. In a similar way

$$\begin{aligned} &c_i c_i^\dagger \, |n_1, \ldots, n_i = 1, \ldots, n_\Omega\rangle = 0 \\ &c_i^\dagger c_i \, |n_1, \ldots, n_i = 1, \ldots, n_\Omega\rangle = |n_1, \ldots, n_i = 1, \ldots, n_\Omega\rangle \end{aligned} \tag{3.43}$$

Thus one sees that $[c_i, c_i^\dagger]$ does not have any simple expression. On the contrary the *anticommutator*

$$[c_i, c_i^\dagger]_+ = c_i c_i^\dagger + c_i^\dagger c_i \tag{3.44}$$

leads to

$$[c_i, c_i^\dagger]_+ \, |n_1, \ldots, n_i, \ldots, n_\Omega\rangle = |n_1, \ldots, n_i, \ldots, n_\Omega\rangle \tag{3.45}$$

and thus

$$[c_i, c_i^\dagger]_+ = 1 \tag{3.46}$$

On can thus guess that the anticommutator will play a major role. The role of the factor $(-1)^{\epsilon_i}$, will thus be to ensure that for the other combinations one also get simple relations for the anticommutator. Let us illustrate it on the action of $c_i c_j^\dagger$ with $i \neq j$. One has (assuming that $i < j$),

$$\begin{aligned} c_i c_j^\dagger \, |n_1, \ldots, n_i, \ldots, n_j, \ldots, n_\Omega\rangle &= (1 - n_j)(-1)^{\epsilon_j} c_i \, |n_1, \ldots, n_i, \ldots, n_j + 1, \ldots, n_\Omega\rangle \\ &= (1 - n_j)(-1)^{\epsilon_j} n_i (-1)^{\epsilon_i} \, |n_1, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_\Omega\rangle \end{aligned} \tag{3.47}$$

On the contrary

$$\begin{aligned} c_j^\dagger c_i \left|n_1,\ldots,n_i,\ldots,n_j,\ldots,n_\Omega\right\rangle &= n_i(-1)^{\epsilon_i} c_j^\dagger \left|n_1,\ldots,n_i-1,\ldots,n_j,\ldots,n_\Omega\right\rangle \\ &= (1-n_j)(-1)^{\epsilon'_j} n_i (-1)^{\epsilon_i} \left|n_1,\ldots,n_i-1,\ldots,n_j+1,\ldots,n_\Omega\right\rangle \end{aligned} \tag{3.48}$$

The term ϵ'_j corresponds to the phase factor with a state with $n_i - 1$ instead of n_i. Thus $\epsilon'_j = \epsilon_j - 1$. In the absence of such phase terms the two expressions (3.47) and (3.48) would be identical and one would have $[c_i, c_j^\dagger] = 0$. Thanks to the phase factors ϵ_j we have now a minus sign between the two terms and the relation now becomes

$$[c_i, c_j^\dagger]_+ = 0 \tag{3.49}$$

which will allow to define the c_i operators only in terms of their anticommutators.

It is easy to check the other relations, and thus one has, in a similar way than for the bosons (3.31)

$$\begin{aligned} [c_i, c_j^\dagger]_+ &= \delta_{i,j} \\ [c_i^\dagger, c_j^\dagger]_+ &= 0 \\ [c_i, c_j]_+ &= 0 \end{aligned} \tag{3.50}$$

In the same way than for the bosons the c_i and $c_i^\dagger$ can be used to construct all the states of the Fock space from a vacuum $|\varnothing\rangle$ which is destroyed by all the c_i ($c_i |\varnothing\rangle = 0$) using the relation (3.20).

Wavefunctions and averages can be computed by exactly the same techniques that were given for the bosons. As an example let us look at the function with two fermions in the states α_1 and α_2

$$|\psi\rangle = c_1^\dagger c_2^\dagger |\varnothing\rangle \tag{3.51}$$

From (3.12) the wavefunction is

$$\langle \boldsymbol{r}_1 \boldsymbol{r}_2 | \psi \rangle = \frac{1}{\sqrt{2}} [\alpha_1(\boldsymbol{r}_1)\alpha_2(\boldsymbol{r}_2) - \alpha_1(\boldsymbol{r}_2)\alpha_2(\boldsymbol{r}_1)] \tag{3.52}$$

which is of course the properly antisymmetrized wavefunction for Fermions. Without going to the wavefunction one can directly see the antisymmetrization at the operator level. Using the anticommutation relation $[c_1, c_2]_+ = 0$ one sees that

$$c_1^\dagger c_2^\dagger |\varnothing\rangle = -c_2^\dagger c_1^\dagger |\varnothing\rangle \tag{3.53}$$

and thus the wavefunction $|\psi\rangle$ is obviously antisymmetric by permutation of the particles.

The fact that the antisymmetrization is taken care of automatically by the operator c_i makes it very convenient to write even complicated functions. For example the Fermi sea corresponds to a state where all states with a lower energy that the Fermi energy are occupied. To describe the one particle state we introduce a complete basis. For the orbital part we can take the momentum basis $|\boldsymbol{k}\rangle$. The electron having a spin 1/2 we need a complete basis for the spin sector. Let us recall that the spin in given in terms of the Pauli matrices σ_l by

$$S_l = \frac{1}{2}\sigma_l \tag{3.54}$$

and the three Pauli matrices are given by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.55}$$

We can take the two eigenstates of S_z denoted either $|\uparrow\rangle$ or $|\downarrow\rangle$ or $|\sigma_z = \pm\rangle$, for the complete basis in the spin sector. The states α are thus $\alpha = (\boldsymbol{k}, \sigma_z = \pm)$. We can define the corresponding creation operators $c^\dagger_{\boldsymbol{k},\sigma}$ which creates an electron with the momentum $\boldsymbol{k}$ and the spin $\sigma = \uparrow, \downarrow$. In terms of these operators the Fermi sea is simply

$$|\mathrm{F}\rangle = \prod_{\boldsymbol{k}, \varepsilon(\boldsymbol{k}) < E_\mathrm{F}} c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{\boldsymbol{k}\downarrow} |\varnothing\rangle \tag{3.56}$$

Averages in the void can be computed by exactly the same technique than described for the bosons. For example, if we take $|\psi\rangle = c^\dagger_1 |\varnothing\rangle$, then (using the anticommutation relations)

$$\begin{aligned} \langle\psi|\psi\rangle &= \langle\varnothing|c_i c_i^\dagger|\varnothing\rangle \\ &= \langle\varnothing|1 - c_i^\dagger c_i|\varnothing\rangle \\ &= \langle\varnothing|1|\varnothing\rangle = 1 \end{aligned} \tag{3.57}$$

Generalizing the above calculation we see that $\langle\mathrm{F}|\mathrm{F}\rangle = 1$.

3.4 One body operators

Now that we have operators that allow to construct the whole Fock space, what remains to be done is to express the physical observables we want to compute in terms of these operators. In order to do that we have to take into account that the physical observables have to act on indistinguishable particles, which sets some constraints on what they can be. To give the expression of the observables in second quantization we have to sort out observables in terms of how many particles are involved. Indeed there are physical observables that measure only the quantum numbers of *one* particle at a time (such as measuring the momentum, density, etc. of the particles) and others that need to deal with the quantum numbers of two of the particles to determine the matrix elements. This is for example the case of the operator measuring the interactions between the particles. The first type is called one body operators, while the second one is two body operators. One can have in principle operators that involve more than two particles to get the matrix elements (such as three body collisions and up) but they are of little use in practice in solid state physics, so we will mostly discuss here the one and two body operators. The formulas given here can be easily generalized if need be.

3.4.1 Definition

Let us first start with one body operators. Quite generally let us call O an operator which involves only the measurement of one particle at a time. Of course if O acts in space with $\mathcal{N}$ particles it must do the measurement on each particle of the system. Let us call $O^{(1)}$ the operator acting in the Hilbert space of a single particle, the operator O must thus be

$$O = O^{(1)}_1 \otimes 1_2 \otimes \ldots \otimes 1_\mathcal{N} + 1_1 \otimes O^{(1)}_2 \otimes \ldots \otimes 1_\mathcal{N} + \ldots + 1_1 \otimes \ldots \otimes O^{(1)}_\mathcal{N} \tag{3.58}$$

where $O^{(1)}_i$ is the operator acting on the i-th particle. The fact that in the above sum, all coefficients are identical, is the direct consequence of the fact that the particles are indistinguishable, and we cannot now in a physical measurement if we are measuring the quantum numbers of the, say 3-rd or 120-th particle. The form (3.58) is thus the most general form, for indistinguishable particles, of a one body operator.

In order to express (3.58) in second quantization, we will first start to look at what happens if we have a system with only one particle in it (if there is no particle a one body operator gives

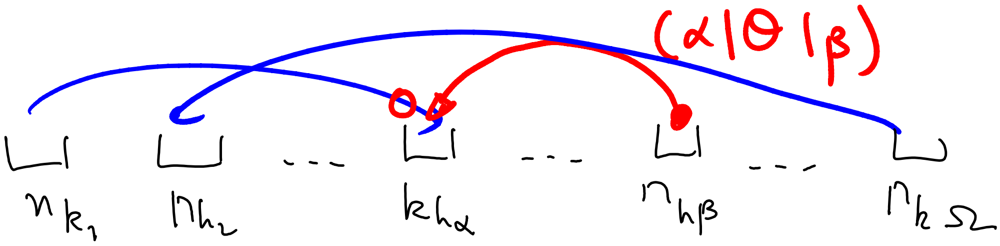


Figure 3.2: The one body operator in second quantization. The operator $c^\dagger_\alpha c_\beta$ changes the quantum state of any particle in the system from the state β to the state α. The amplitude element of such process is given by the matrix elements of the single particle operator $O^{(1)}$: $\langle\alpha|O^{(1)}|\beta\rangle$. This ensure that the operator $O^{(1)}$ is operating on each particle in turn.

obviously zero). In that case $O = O^{(1)}$ and using the complete basis α we can write

$$O = \sum_{\alpha,\beta} |\alpha\rangle\langle\alpha|O^{(1)}|\beta\rangle\langle\beta| \tag{3.59}$$

We can now use the fact that $|\alpha\rangle = c^\dagger_\alpha |\varnothing\rangle$ to obtain

$$O = \sum_{\alpha,\beta} \langle\alpha|O^{(1)}|\beta\rangle c^\dagger_\alpha|\varnothing\rangle\langle\varnothing|c_\beta \tag{3.60}$$

The physical interpretation of this formula is quite simple. The operator c_β destroys a particle in a state β. Since we have only one particle in the system we are forced to go to the vacuum, then from the vacuum the operator $c^\dagger_\alpha$ recreates the particle in the state α. The net result is that we have still one particle in the system but it has changed its quantum state going from the state β to the state α. The amplitude of such transition is given by the matrix elements of the operator $O^{(1)}$ between the states β and α.

If instead of one particle we had now an arbitrary number of particles in the system, we would have exactly the same thing to do for each one of the particles, leaving the quantum numbers of the others invariant as is obvious from (3.58) and do the sum. An operator that accomplishes this is the operator

$$O = \sum_{\alpha,\beta} \langle\alpha|O^{(1)}|\beta\rangle c^\dagger_\alpha c_\beta \tag{3.61}$$

which is identical to (3.60) except that one is not forced to go to the vacuum after the destruction of the particle in the state β. Indeed if there are several particles the operator $c^\dagger_\alpha c_\beta$ will change the quantum number of one particle from state β to state α, and leave the quantum numbers of all the other particles in the system untouched. However the operator c_β will operate on *all* the particles in the system, and thus make such a transition for the 1-st, 2-nd etc. realizing automatically the sum in (3.58). The interpretation of (3.61) is show in Fig. 3.2

The expression (3.61) thus allows to represent any single-body operator in second quantization, with the sole knowledge of the operator $O^{(1)}$ acting in the space of one particle only. Note that the wavefunctions coming from the choice of the complete basis α only intervene in the calculation of the matrix elements $\langle\alpha|O^{(1)}|\beta\rangle$. Once these matrix elements are computed, the whole operator is reduced to a linear combination of creation and destruction operators, and thus all physical averages can be computed by the techniques described in the previous section, without ever having to go back to wavefunctions. Of course all the aspects of symmetrization or antisymmetrization are automatically taken care of by the nature of the creation or destruction operators.

3.4.2 Examples

Let us start with the operator measuring the density of particles at a point $\boldsymbol{r}_0$. The operator giving such a density for one particle is

$$\rho^{(1)}(\boldsymbol{r}_0) = |\boldsymbol{r}_0\rangle\langle\boldsymbol{r}_0| \tag{3.62}$$

In first quantization $\langle\psi|\rho^{(1)}(\boldsymbol{r}_0)|\psi\rangle = |\psi(\boldsymbol{r}_0)|^2$. In second quantization the form of the operator will depend on the choice of the complete basis α we take. Let us start by taking the basis of the position $|\boldsymbol{r}\rangle$. In that case the operator $c^\dagger_{\boldsymbol{r}}$ is the operator creating a particle at point $\boldsymbol{r}$. Using this basis and the relation (3.61) one obtains

$$\begin{aligned}\rho(\boldsymbol{r}_0) &= \sum_{\boldsymbol{r}\boldsymbol{r}'} \langle\boldsymbol{r}|\boldsymbol{r}_0\rangle\langle\boldsymbol{r}_0|\boldsymbol{r}'\rangle c^\dagger_{\boldsymbol{r}} c_{\boldsymbol{r}'} \\ &= \sum_{\boldsymbol{r}\boldsymbol{r}'} \delta(\boldsymbol{r}-\boldsymbol{r}_0)\delta(\boldsymbol{r}_0-\boldsymbol{r}') c^\dagger_{\boldsymbol{r}} c_{\boldsymbol{r}'} \\ &= c^\dagger_{\boldsymbol{r}_0} c_{\boldsymbol{r}_0}\end{aligned} \tag{3.63}$$

The expression $c^\dagger_{\boldsymbol{r}_0} c_{\boldsymbol{r}_0}$ is particularly simple to understand. The operator $c^\dagger_{\boldsymbol{r}_0} c_{\boldsymbol{r}_0}$ destroys and recreate a particle in the same quantum state. Thus it has changed nothing on the system. However the action of the operator $c_{\boldsymbol{r}_0}$ will give zero if there is no particle in the corresponding quantum state (here a particle at the point $\boldsymbol{r}_0$) to destroy. The operator $c^\dagger_{\boldsymbol{r}_0} c_{\boldsymbol{r}_0}$ thus gives zero if there is no particle in the corresponding quantum state and one if there is one particle. It thus simply counts the number of particles at the point $\boldsymbol{r}_0$. Quite generally the operator $c^\dagger_\alpha c_\alpha$ simply counts the number of particles in the state α. The total number of particles in the system is simply given by

$$N = \int d\boldsymbol{r}\, c^\dagger_{\boldsymbol{r}} c_{\boldsymbol{r}} \tag{3.64}$$

The generalization to the case of particles with spins is obvious. In that case the complete basis will be $\alpha = (\boldsymbol{r}, \sigma)$. The density operator only acts on the spatial part thus

$$\rho^{(1)}(\boldsymbol{r}_0) = |\boldsymbol{r}_0\rangle\langle\boldsymbol{r}_0| \otimes 1_{\text{spin}} \tag{3.65}$$

and thus (3.61) gives

$$\begin{aligned}\rho(\boldsymbol{r}_0) &= \sum_{\boldsymbol{r}\sigma,\boldsymbol{r}'\sigma'} \langle\boldsymbol{r}\sigma|\boldsymbol{r}_0\rangle\langle\boldsymbol{r}_0|\boldsymbol{r}'\sigma'\rangle c^\dagger_{\boldsymbol{r}\sigma} c_{\boldsymbol{r}'\sigma'} \\ &= \sum_{\boldsymbol{r}\sigma,\boldsymbol{r}'\sigma'} \delta(\boldsymbol{r}-\boldsymbol{r}_0)\delta(\boldsymbol{r}_0-\boldsymbol{r}')\delta_{\sigma\sigma'} c^\dagger_{\boldsymbol{r}\sigma} c_{\boldsymbol{r}',\sigma'} \\ &= c^\dagger_{\boldsymbol{r}_0\uparrow} c_{\boldsymbol{r}_0\uparrow} + c^\dagger_{\boldsymbol{r}_0\downarrow} c_{\boldsymbol{r}_0\downarrow}\end{aligned} \tag{3.66}$$

For particles with spins we could compute the spin density along z at the point $\boldsymbol{r}_0$. In that case the operator is

$$\sigma^{(1)}_z(\boldsymbol{r}_0) = |\boldsymbol{r}_0\rangle\langle\boldsymbol{r}_0| \otimes S_z \tag{3.67}$$

And thus using (3.61) one gets

$$\begin{aligned}\sigma_z(\boldsymbol{r}_0) &= \sum_{\boldsymbol{r}\sigma,\boldsymbol{r}'\sigma'} \langle \boldsymbol{r}\sigma|\boldsymbol{r}_0\rangle\langle \boldsymbol{r}_0| \otimes S_z|\boldsymbol{r}'\sigma'\rangle c^\dagger_{\boldsymbol{r}\sigma} c_{\boldsymbol{r}'\sigma'} \\ &= \sum_{\sigma\sigma'} \langle\sigma|S_z|\sigma'\rangle c^\dagger_{\boldsymbol{r}_0\sigma} c_{\boldsymbol{r}_0\sigma'} \\ &= \frac{1}{2}(c^\dagger_{\boldsymbol{r}_0\uparrow}c_{\boldsymbol{r}_0\uparrow} - c^\dagger_{\boldsymbol{r}_0\downarrow}c_{\boldsymbol{r}_0\downarrow})\end{aligned} \tag{3.68}$$

In a similar way the spin density along the x direction would give

$$\begin{aligned}\sigma_x(\boldsymbol{r}_0) &= \sum_{\boldsymbol{r}\sigma,\boldsymbol{r}'\sigma'} \langle \boldsymbol{r}\sigma|\boldsymbol{r}_0\rangle\langle \boldsymbol{r}_0| \otimes S_x|\boldsymbol{r}'\sigma'\rangle c^\dagger_{\boldsymbol{r}\sigma} c_{\boldsymbol{r}'\sigma'} \\ &= \sum_{\sigma\sigma'} \langle\sigma|S_x|\sigma'\rangle c^\dagger_{\boldsymbol{r}_0\sigma} c_{\boldsymbol{r}_0\sigma'} \\ &= \frac{1}{2}(c^\dagger_{\boldsymbol{r}_0\uparrow}c_{\boldsymbol{r}_0\downarrow} + c^\dagger_{\boldsymbol{r}_0\downarrow}c_{\boldsymbol{r}_0\uparrow})\end{aligned} \tag{3.69}$$

Alternatively we could have used the basis of the eigenstates of the momentum $|\boldsymbol{k}\rangle$ where

$$\langle \boldsymbol{r}|\boldsymbol{k}\rangle = \frac{1}{\sqrt{\Omega}} e^{i\boldsymbol{k}\boldsymbol{r}} \tag{3.70}$$

Since the spin and orbital part are independent in the above we will just give the expressions for the spinless case, the addition of the spin sector being exactly as above. The operator $c_{\boldsymbol{k}}$ thus now destroys a particle with momentum $\boldsymbol{k}$ (i.e. in a plane wave state with momentum $\boldsymbol{k}$). Using (3.61) one gets

$$\begin{aligned}\rho(\boldsymbol{r}_0) &= \sum_{\boldsymbol{k}_1\boldsymbol{k}_2} \langle \boldsymbol{k}_1|\boldsymbol{r}_0\rangle\langle \boldsymbol{r}_0|\boldsymbol{k}_2\rangle c^\dagger_{\boldsymbol{k}_1} c_{\boldsymbol{k}_2} \\ &= \frac{1}{\Omega}\sum_{\boldsymbol{k}_1\boldsymbol{k}_2} e^{-i\boldsymbol{k}_1\boldsymbol{r}_0} e^{i\boldsymbol{k}_2\boldsymbol{r}_0} c^\dagger_{\boldsymbol{k}_1} c_{\boldsymbol{k}_2}\end{aligned} \tag{3.71}$$

The expression (3.71) is not as simple as (3.63) since the density operator is not diagonal in the momentum basis. However both (3.71) and (3.63) represent the same operator. This gives us a direct connection between the operators creating a particle at point $\boldsymbol{r}$ and the ones creating a particle with momentum $\boldsymbol{k}$. Comparing the expressions (3.71) and (3.63) one gets

$$c_{\boldsymbol{r}} = \frac{1}{\sqrt{\Omega}} \sum e^{i\boldsymbol{k}\boldsymbol{r}} c_{\boldsymbol{k}} \tag{3.72}$$

Using the expression (3.71) we can also compute the total number of particles in the system

$$\begin{aligned}N &= \int d\boldsymbol{r} \frac{1}{\Omega} \sum_{\boldsymbol{k}_1\boldsymbol{k}_2} e^{-i\boldsymbol{k}_1\boldsymbol{r}} e^{i\boldsymbol{k}_2\boldsymbol{r}} c^\dagger_{\boldsymbol{k}_1} c_{\boldsymbol{k}_2} \\ &= \sum_{\boldsymbol{k}_1\boldsymbol{k}_2} \delta_{\boldsymbol{k}_1\boldsymbol{k}_2} c^\dagger_{\boldsymbol{k}_1} c_{\boldsymbol{k}_2} \\ &= \sum_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}} c_{\boldsymbol{k}}\end{aligned} \tag{3.73}$$

Keeping in mind that $c^\dagger_{\boldsymbol{k}} c_{\boldsymbol{k}}$ simply counts the number of particles in the quantum state $\boldsymbol{k}$, one gets again that the total number of particles is the sum of all numbers of particles in all possible quantum states.

Finally one can use (3.71) to get a simple expression of the Fourier transform of the density

$$\begin{aligned}\rho(\boldsymbol{q}) &= \int d\boldsymbol{r} e^{-i\boldsymbol{q}\boldsymbol{r}}\rho(\boldsymbol{r}) \\ &= \int d\boldsymbol{r} e^{-i\boldsymbol{q}\boldsymbol{r}}\frac{1}{\Omega}\sum_{\boldsymbol{k}_1\boldsymbol{k}_2} e^{-i\boldsymbol{k}_1\boldsymbol{r}}e^{i\boldsymbol{k}_2\boldsymbol{r}}c^\dagger_{\boldsymbol{k}_1}c_{\boldsymbol{k}_2} \\ &= \sum_{\boldsymbol{k}_1\boldsymbol{k}_2}\delta_{\boldsymbol{k}_2,\boldsymbol{k}_1+\boldsymbol{q}}c^\dagger_{\boldsymbol{k}_1}c_{\boldsymbol{k}_2} \\ &= \sum_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}-\boldsymbol{q}}c_{\boldsymbol{k}}\end{aligned} \tag{3.74}$$

Another important operator is of course the kinetic energy of the particles. For one particles one has $H^{(1)} = \frac{P^2}{2m}$. More generally we could have any function of the momentum $H^{(1)} = E(P)$. It is thus very convenient to use the momentum basis. The kinetic energy is thus expressed as

$$\begin{aligned}H &= \sum_{\boldsymbol{k}_1\boldsymbol{k}_2}\langle\boldsymbol{k}_1|\varepsilon(P)|\boldsymbol{k}_2\rangle c^\dagger_{\boldsymbol{k}_1}c_{\boldsymbol{k}_2} \\ &= \sum_{\boldsymbol{k}_1\boldsymbol{k}_2}\delta_{\boldsymbol{k}_1\boldsymbol{k}_2}\varepsilon(\boldsymbol{k}_2)c^\dagger_{\boldsymbol{k}_1}c_{\boldsymbol{k}_2} \\ &= \sum_{\boldsymbol{k}}\varepsilon(\boldsymbol{k})c^\dagger_{\boldsymbol{k}}c_{\boldsymbol{k}}\end{aligned} \tag{3.75}$$

which has the simple interpretation that the total kinetic energy of the system is the sum of the kinetic energy $E(\boldsymbol{k})$ of particles with momentum $\boldsymbol{k}$, times the number of particles in such a state $\boldsymbol{k}$ (which is given by $c^\dagger_{\boldsymbol{k}}c_{\boldsymbol{k}}$). The generalization for system with spin is obvious and in general one has

$$H = \sum_{\boldsymbol{k}\sigma}\varepsilon(\boldsymbol{k})c^\dagger_{\boldsymbol{k}\sigma}c_{\boldsymbol{k}\sigma} \tag{3.76}$$

assuming that the kinetic energy does not depend on spin (no spin-orbit coupling). Note that since the total number of particles is $N = \sum_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}}c_{\boldsymbol{k}}$ adding a chemical potential $-\mu N$ does not change the form of the Hamiltonian

$$H = \sum_{\boldsymbol{k}\sigma}\xi(\boldsymbol{k})c^\dagger_{\boldsymbol{k}\sigma}c_{\boldsymbol{k}\sigma} \tag{3.77}$$

but simply replaces $\varepsilon(\boldsymbol{k})$ by $\xi(\boldsymbol{k}) = \varepsilon(\boldsymbol{k}) - \mu$. At zero temperature the energy $\xi(\boldsymbol{k})$ is zero at the Fermi level, negative below and positive above.

3.5 Two body operators

Let us now look at operators that involve two particles to define their matrix elements. This is in particular the case of the interaction between two particles.

3.5.1 Definition

In a similar spirit than for single body operators, let us call $O^{(2)}$ the corresponding operator that operates in the space of only two particles. The two body operator that will operate with $\mathcal{N}$ particles must thus have the form

$$O = \sum_{i<j} O^{(2)}_{i,j}\bigotimes_{k\neq i,j} 1_k = \frac{1}{2}\sum_{i\neq j} O^{(2)}_{i,j}\bigotimes_{k\neq i,j} 1_k \tag{3.78}$$

so that the operator $O^{(2)}$ can operate on each pairs of particles in the system. In a similar way than for single body operators, the coefficients in the above sum must all be equal otherwise it would mean that the particles could be distinguished.

To understand how to write O in second quantization let us look at the case where there are exactly two particles in the system. We must define the operator O by its matrix elements in the physical space of the (anti)symmetrized functions $|\alpha,\beta\rangle$. It means we should know all the elements

$$\langle\alpha,\beta|O^{(2)}|\gamma,\delta\rangle \tag{3.79}$$

Using (3.3) to express $|\alpha,\beta\rangle$ in terms of the ordered kets (3.4), and the fact that the particles are indistinguishable thus

$$(\alpha,\beta|O^{(2)}|\gamma,\delta) = (\beta,\alpha|O^{(2)}|\delta,\gamma) \tag{3.80}$$

since this is simply the exchange of the two particles, we obtain that

$$\langle\alpha,\beta|O^{(2)}|\gamma,\delta\rangle = (\alpha,\beta|O^{(2)}|\gamma,\delta) \pm (\alpha,\beta|O^{(2)}|\delta,\gamma) \tag{3.81}$$

We should now find in second quantization an operator that reproduces these matrix elements, and of course works for $\mathcal{N}$ instead of two. One can check that

$$O = \frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}(\alpha,\beta|O^{(2)}|\gamma,\delta)c^\dagger_\alpha c^\dagger_\beta c_\delta c_\gamma \tag{3.82}$$

works both for fermions and bosons. We will not prove the relation in general but simply check that it works for $\mathcal{N}=2$ particles. We compute from (3.82) the matrix elements

$$\langle\alpha_0,\beta_0|O|\gamma_0,\delta_0\rangle = \frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}(\alpha,\beta|O^{(2)}|\gamma,\delta)\langle\alpha_0,\beta_0|c^\dagger_\alpha c^\dagger_\beta c_\delta c_\gamma|\gamma_0,\delta_0\rangle \tag{3.83}$$

Since $|\alpha_0,\beta_0\rangle = c^\dagger_{\alpha_0}c^\dagger_{\beta_0}$ we have to compute averages of the form

$$\langle\varnothing|c_{\beta_0}c_{\alpha_0}c^\dagger_\alpha c^\dagger_\beta c_\delta c_\gamma c^\dagger_{\gamma_0}c^\dagger_{\delta_0}|\varnothing\rangle \tag{3.84}$$

which can be easily computed by the technique consisting to bringing the destruction operators to act on the vacuum. This gives

$$\langle\varnothing|c_{\beta_0}c_{\alpha_0}c^\dagger_\alpha c^\dagger_\beta c_\delta c_\gamma c^\dagger_{\gamma_0}c^\dagger_{\delta_0}|\varnothing\rangle = [\delta_{\alpha_0,\alpha}\delta_{\beta_0,\beta} \pm \delta_{\alpha_0,\beta}\delta_{\beta_0,\alpha}][\delta_{\gamma_0,\gamma}\delta_{\delta_0,\delta} \pm \delta_{\gamma_0,\delta}\delta_{\delta_0,\gamma}] \tag{3.85}$$

The $+$ sign is as usual for bosons and the $-$ one for fermions. Physically it means that when the destruction operators act

$$c_\delta c_\gamma\,|\gamma_0,\delta_0\rangle \tag{3.86}$$

they have to destroy the two particles in the two possible quantum states and thus δ has to be one of the states and γ the other with the proper sign depending on the (anti)symmetry of the wavefunction. Using (3.85) in (3.83) we indeed recover the same matrix elements than (3.81).

Physically the formula (3.82) has a similar interpretation than the one for the single body operators. The term $c^\dagger_\alpha c^\dagger_\beta c_\delta c_\gamma$ destroys two particles with the quantum numbers γ and δ. Note that two particles must be present in the system for this operator to act, which is what it should be for a two body operator. It then recreates the two particles with two new quantum numbers α and β. The amplitude for the transition is the matrix elements of the operator $O^{(2)}$ in a transition where the first particle goes from state γ to state α and the second from state δ to state β. Note that the matrix element can be written for ordered kets, the creation and destruction operators taking care of all the permutations and of doing this transition for any pair of particles in the system.

3.5.2 Examples

Let us show some examples, more will be seen in the exercices. The most common interaction between the electrons is one that depends on the distance between the two particles. The two body operators of such an interaction is thus

$$O^{(2)} = V(\boldsymbol{R}_1 - \boldsymbol{R}_2) \tag{3.87}$$

where $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are the *operators* measuring the position of the first and the second particle. The function V depends on the interaction. For the Coulomb interaction it is

$$V(\boldsymbol{r}) = \frac{e^2}{4\pi\epsilon_0 r} \tag{3.88}$$

but other types of interactions such as a local interaction $V(\boldsymbol{r}) = U\delta(\boldsymbol{r})$ are also possible choices. We will keep V as a general function in what follows.

To express the operator, we have again the choice for the basis. Because the operator $V(\boldsymbol{R}_1 - \boldsymbol{R}_2)$ is diagonal in the position basis, let us start with this one. Using (3.82) and the fact that α is the position basis we obtain

$$\begin{aligned} H_{\text{int}} &= \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 d\boldsymbol{r}_3 d\boldsymbol{r}_4 (\boldsymbol{r}_3\boldsymbol{r}_4|V(\boldsymbol{R}_1 - \boldsymbol{R}_2)|\boldsymbol{r}_1\boldsymbol{r}_2) c^\dagger_{\boldsymbol{r}_3} c^\dagger_{\boldsymbol{r}_4} c_{\boldsymbol{r}_2} c_{\boldsymbol{r}_1} \\ &= \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 d\boldsymbol{r}_3 d\boldsymbol{r}_4 V(\boldsymbol{r}_1 - \boldsymbol{r}_2)\delta(\boldsymbol{r}_3 - \boldsymbol{r}_1)\delta(\boldsymbol{r}_4 - \boldsymbol{r}_2) c^\dagger_{\boldsymbol{r}_3} c^\dagger_{\boldsymbol{r}_4} c_{\boldsymbol{r}_2} c_{\boldsymbol{r}_1} \\ & \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2) c^\dagger_{\boldsymbol{r}_1} c^\dagger_{\boldsymbol{r}_2} c_{\boldsymbol{r}_2} c_{\boldsymbol{r}_1} \end{aligned} \tag{3.89}$$

If one includes spin the complete basis becomes $\alpha = (\boldsymbol{r}, \sigma)$ and since the operator $V(\boldsymbol{R}_1 - \boldsymbol{R}_2)$ is the identity in the spin sector, one gets

$$H_{\text{int}} = \frac{1}{2}\sum_{\sigma_1\sigma_2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2) c^\dagger_{\boldsymbol{r}_1\sigma_1} c^\dagger_{\boldsymbol{r}_2\sigma_2} c_{\boldsymbol{r}_2\sigma_2} c_{\boldsymbol{r}_1\sigma_1} \tag{3.90}$$

The expression (3.90) can be cast in a more familiar form using the (anti)commutation relations. For fermions

$$\begin{aligned} c^\dagger_{\boldsymbol{r}_1\sigma_1} c^\dagger_{\boldsymbol{r}_2\sigma_2} c_{\boldsymbol{r}_2\sigma_2} c_{\boldsymbol{r}_1\sigma_1} &= -c^\dagger_{\boldsymbol{r}_1\sigma_1} c^\dagger_{\boldsymbol{r}_2\sigma_2} c_{\boldsymbol{r}_1\sigma_1} c_{\boldsymbol{r}_2\sigma_2} \\ &= -c^\dagger_{\boldsymbol{r}_1\sigma_1}(\delta_{\boldsymbol{r}_1,\sigma 1;\boldsymbol{r}_2\sigma_2} - c_{\boldsymbol{r}_1\sigma_1} c^\dagger_{\boldsymbol{r}_2\sigma_2}) c_{\boldsymbol{r}_2\sigma_2} \\ &= -\delta_{\boldsymbol{r}_1,\sigma 1;\boldsymbol{r}_2\sigma_2} c^\dagger_{\boldsymbol{r}_1\sigma_1} c_{\boldsymbol{r}_2\sigma_2} + c^\dagger_{\boldsymbol{r}_1\sigma_1} c_{\boldsymbol{r}_1\sigma_1} c^\dagger_{\boldsymbol{r}_2\sigma_2} c_{\boldsymbol{r}_2\sigma_2} \\ &= -\delta_{\boldsymbol{r}_1,\sigma 1;\boldsymbol{r}_2\sigma_2} \rho_{\sigma_1}(\boldsymbol{r}_1) + \rho_{\sigma_1}(\boldsymbol{r}_1)\rho_{\sigma_2}(\boldsymbol{r}_2) \end{aligned} \tag{3.91}$$

There is a similar expression for bosons with a + sign. The second term would lead to the interaction

$$H_{\text{int}} = \frac{1}{2}\sum_{\sigma_1\sigma_2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2)\rho_{\sigma_1}(\boldsymbol{r}_1)\rho_{\sigma_2}(\boldsymbol{r}_2) \tag{3.92}$$

This is the familiar expression of the interaction between two densities of particles (or charges) at two different points. The difference is that now the ρ are *operators* measuring the density instead of classical variables. The first term gives

$$\sum_{\sigma}\int d\boldsymbol{r} V(\boldsymbol{r} = 0)\rho_{\sigma_1}(\boldsymbol{r}_1) = V(\boldsymbol{r} = 0)N \tag{3.93}$$

is simply a chemical potential term. Notice that it can be infinite for some interactions such as the Coulomb interaction. This first terms is there to correct that the expression (3.92) contrarily to (3.90) does not contain *only* the interaction between two different particles. Indeed (3.90) has two destruction operators on the right, which means that the operators can only act on states containing two particles. One the contrary (3.92) is of the form

$$c^\dagger_{\boldsymbol{r}_1\sigma_1} c_{\boldsymbol{r}_1\sigma_1} c^\dagger_{\boldsymbol{r}_2\sigma_2} c_{\boldsymbol{r}_2\sigma_2} \tag{3.94}$$

and can thus act even if there is only one particle in the system. It thus contains a fake "self-interaction" of the particle with itself. It is this interaction that leads to the chemical potential (3.93) that need to be properly included together with (3.92). Note however that if one fixes the chemical potential by ensuring that the total number of particles is fixed, then this modification is irrelevant since it is simply absorbed in the redefinition of the chemical potential and one can use (3.90) or (3.92) indifferently.

Let us now rewrite the interaction in the momentum basis. Using (3.82) and a basis $\alpha = (\boldsymbol{k}, \sigma)$ one has

$$\begin{aligned}
H_{\text{int}} &= \frac{1}{2} \sum_{\substack{\boldsymbol{k}_1\sigma_1, \boldsymbol{k}_2\sigma_2, \\ \boldsymbol{k}_3\sigma_3, \boldsymbol{k}_4\sigma_4}} (\boldsymbol{k}_3\sigma_3, \boldsymbol{k}_4\sigma_4 | V(\boldsymbol{R}_1 - \boldsymbol{R}_2) | \boldsymbol{k}_1\sigma_1, \boldsymbol{k}_2\sigma_2) c^\dagger_{\boldsymbol{k}_3\sigma_3} c^\dagger_{\boldsymbol{k}_4\sigma_4} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \\
&= \frac{1}{2} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4, \\ \sigma_1\sigma_2}} (\boldsymbol{k}_3\boldsymbol{k}_4 | V(\boldsymbol{R}_1 - \boldsymbol{R}_2) | \boldsymbol{k}_1\boldsymbol{k}_2) c^\dagger_{\boldsymbol{k}_3\sigma_1} c^\dagger_{\boldsymbol{k}_4\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \\
&= \frac{1}{2} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4, \\ \sigma_1\sigma_2}} \int d\boldsymbol{r}_1\, d\boldsymbol{r}_2 (\boldsymbol{k}_3\boldsymbol{k}_4 | V(\boldsymbol{R}_1 - \boldsymbol{R}_2) | \boldsymbol{r}_1\boldsymbol{r}_2)(\boldsymbol{r}_1\boldsymbol{r}_2 | \boldsymbol{k}_1\boldsymbol{k}_2) c^\dagger_{\boldsymbol{k}_3\sigma_1} c^\dagger_{\boldsymbol{k}_4\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \\
&= \frac{1}{2\Omega} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4, \\ \sigma_1\sigma_2}} \int d\boldsymbol{r}_1\, d\boldsymbol{r}_2 (\boldsymbol{k}_3\boldsymbol{k}_4 | V(\boldsymbol{R}_1 - \boldsymbol{R}_2) | \boldsymbol{r}_1\boldsymbol{r}_2) e^{i(\boldsymbol{k}_1\boldsymbol{r}_1 + \boldsymbol{k}_2\boldsymbol{r}_2)} c^\dagger_{\boldsymbol{k}_3\sigma_1} c^\dagger_{\boldsymbol{k}_4\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \\
&= \frac{1}{2\Omega} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4, \\ \sigma_1\sigma_2}} \int d\boldsymbol{r}_1\, d\boldsymbol{r}_2 (\boldsymbol{k}_3\boldsymbol{k}_4 | \boldsymbol{r}_1\boldsymbol{r}_2) V(\boldsymbol{r}_1 - \boldsymbol{r}_2) e^{i(\boldsymbol{k}_1\boldsymbol{r}_1 + \boldsymbol{k}_2\boldsymbol{r}_2)} c^\dagger_{\boldsymbol{k}_3\sigma_1} c^\dagger_{\boldsymbol{k}_4\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \\
&= \frac{1}{2\Omega^2} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4, \\ \sigma_1\sigma_2}} \int d\boldsymbol{r}_1\, d\boldsymbol{r}_2 e^{-i(\boldsymbol{k}_3\boldsymbol{r}_1 + \boldsymbol{k}_4\boldsymbol{r}_2)} V(\boldsymbol{r}_1 - \boldsymbol{r}_2) e^{i(\boldsymbol{k}_1\boldsymbol{r}_1 + \boldsymbol{k}_2\boldsymbol{r}_2)} c^\dagger_{\boldsymbol{k}_3\sigma_1} c^\dagger_{\boldsymbol{k}_4\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1}
\end{aligned} \tag{3.95}$$

One can use the new variables of center of mass $\boldsymbol{R} = (\boldsymbol{r}_1 + \boldsymbol{r}_2)/2$ and relative motion $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ to reexpress $\boldsymbol{r}_1 = \boldsymbol{R} + \boldsymbol{r}/2$ and $\boldsymbol{r}_2 = \boldsymbol{R} - \boldsymbol{r}/2$, to obtain

$$\begin{aligned}
H_{\text{int}} &= \frac{1}{2\Omega^2} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4, \\ \sigma_1\sigma_2}} \int d\boldsymbol{R} d\boldsymbol{r} e^{i(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{k}_3 - \boldsymbol{k}_4)\boldsymbol{R}} V(\boldsymbol{r}) e^{i(\boldsymbol{k}_1 - \boldsymbol{k}_3 - \boldsymbol{k}_2 + \boldsymbol{k}_4)\boldsymbol{r}/2} c^\dagger_{\boldsymbol{k}_3\sigma_1} c^\dagger_{\boldsymbol{k}_4\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \\
&= \frac{1}{2\Omega} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4, \\ \sigma_1\sigma_2}} \delta_{\boldsymbol{k}_1 + \boldsymbol{k}_2, \boldsymbol{k}_3 + \boldsymbol{k}_4} \int d\boldsymbol{r} V(\boldsymbol{r}) e^{i(\boldsymbol{k}_1 - \boldsymbol{k}_3)\boldsymbol{r}} c^\dagger_{\boldsymbol{k}_3\sigma_1} c^\dagger_{\boldsymbol{k}_4\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \\
&= \frac{1}{2\Omega} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4, \\ \sigma_1\sigma_2}} \delta_{\boldsymbol{k}_1 + \boldsymbol{k}_2, \boldsymbol{k}_3 + \boldsymbol{k}_4} V(\boldsymbol{q} = \boldsymbol{k}_3 - \boldsymbol{k}_1) c^\dagger_{\boldsymbol{k}_3\sigma_1} c^\dagger_{\boldsymbol{k}_4\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1}
\end{aligned} \tag{3.96}$$

Let us comment this expression. The integration over $\boldsymbol{R}$ gives the term $\delta_{\boldsymbol{k}_1 + \boldsymbol{k}_2, \boldsymbol{k}_3 + \boldsymbol{k}_4}$ which expresses the conservation of the momenta of the two particles before and after the interaction. This is the direct consequence of the fact that we have chosen an interaction term that is translationally invariant $V(\boldsymbol{R}_1 - \boldsymbol{R}_2)$ and thus the total momentum ($\boldsymbol{k}_1 + \boldsymbol{k}_2$ and $\boldsymbol{k}_3 + \boldsymbol{k}_4$) should

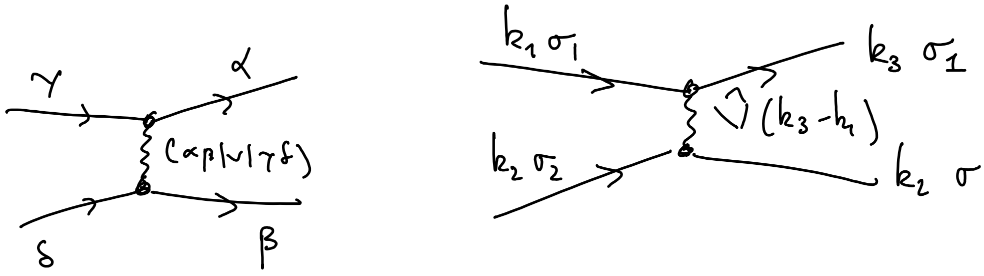


Figure 3.3: Pictorial visualization of the term (3.97). Each destruction operator is represented by an incoming arrow, each creation one by an outgoing one. One sees that the interaction can be viewed as the scattering of one particle going from state $\boldsymbol{k}_1\sigma_1$ to $\boldsymbol{k}_1 + \boldsymbol{q}, \sigma_1$ by another one going from state $\boldsymbol{k}_2\sigma_2$ to state $\boldsymbol{k}_2 - \boldsymbol{q}, \sigma_2$. The amplitude of such matrix elements is the Fourier transform of the interaction potential $V(\boldsymbol{q})$. Since the potential is invariant by translation in space, the momentum is conserved across the interaction. Since the potential does not depend on the spin degrees of freedom the spin of each particle is individually conserved by the interaction. This representation is known as Feynman diagrams. It is extremely useful when constructing the perturbation theory.

be conserved. The integral over the relative coordinates leads directly to the Fourier transform of the interaction potential with a wavevector that corresponds to momentum transferred on one of the particles during the interaction. One can rewrite the term taking into account the $\delta_{\boldsymbol{k}_1+\boldsymbol{k}_2,\boldsymbol{k}_3+\boldsymbol{k}_4}$ constraint as

$$H_{\text{int}} = \frac{1}{2\Omega} \sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{q},\\ \sigma_1\sigma_2}} V(\boldsymbol{q}) c^\dagger_{\boldsymbol{k}_1+\boldsymbol{q},\sigma_1} c^\dagger_{\boldsymbol{k}_2-\boldsymbol{q},\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \tag{3.97}$$

This term can be represented visually as shown in Fig. 3.3.

3.6 Solving with second quantization

We now have the tools to express all the operators we need, either for the Hamiltonian or other physical observables, in second quantization. We also saw how to compute averages of an arbitrary number of such creation and destruction operators in the vacuum. However one important question that remains is how to solve practically a problem, in second quantization, when we know the Hamiltonian. In the first quantization we write the Schrödinger equation, and from that find both the eigenvalues and eigenvectors. But the very essence of the second quantization is to avoid to ever deal with wavefunction so we want to follow another route to obtain the eigenvalues and eigenvectors. How to do this is what we examine now.

3.6.1 Eigenvalues and Eigenstates

Let us first look if we can find the eigenvalues or eigenvector of some simple Hamiltonian. Let us start with a general quadratic Hamiltonian

$$H = \sum_\alpha A_\alpha c^\dagger_\alpha c_\alpha \tag{3.98}$$

where α is some complete basis, and the coefficients A_α are arbitrary numbers. Several Hamiltonians are of such forms, for example the kinetic energy (3.77). For quadratic and diagonal Hamiltonians of the form (3.98) the problem is solved. Indeed each vector of the form

$$c^\dagger_{\alpha_1} c^\dagger_{\alpha_2} c^\dagger_{\alpha_3} \dots c^\dagger_{\alpha_p} |\varnothing\rangle \tag{3.99}$$

is an eigenvector of H with an eigenvalue

$$E = \sum_{i=1}^{p} A_i \tag{3.100}$$

To show this let us illustrate the calculation with two terms $|\psi\rangle = c^\dagger_{\alpha_1} c^\dagger_{\alpha_2} |\varnothing\rangle$ (for fermions, an analogous calculation can be performed for bosons)

$$\begin{aligned}
H c^\dagger_{\alpha_1} c^\dagger_{\alpha_2} |\varnothing\rangle &= \left(\sum_\alpha A_\alpha c^\dagger_\alpha c_\alpha\right) c^\dagger_{\alpha_1} c^\dagger_{\alpha_2} |\varnothing\rangle \\
&= \sum_\alpha A_\alpha c^\dagger_\alpha (\delta_{\alpha,\alpha_1} - c^\dagger_{\alpha_1} c_\alpha) c^\dagger_{\alpha_2} |\varnothing\rangle \\
&= A_{\alpha_1} |\psi\rangle - \sum_\alpha A_\alpha c^\dagger_\alpha c^\dagger_{\alpha_1} c_\alpha c^\dagger_{\alpha_2} |\varnothing\rangle \\
&= A_{\alpha_1} |\psi\rangle - \sum_\alpha A_\alpha c^\dagger_\alpha c^\dagger_{\alpha_1} (\delta_{\alpha,\alpha_2} - c^\dagger_{\alpha_2} c_\alpha) |\varnothing\rangle \\
&= A_{\alpha_1} |\psi\rangle - A_{\alpha_2} c^\dagger_{\alpha_2} c^\dagger_{\alpha_1} |\varnothing\rangle \\
&= A_{\alpha_1} |\psi\rangle + A_{\alpha_2} |\psi\rangle
\end{aligned} \tag{3.101}$$

The physics of this result is simple to understand. The operator $c^\dagger_\alpha c_\alpha$ counts the particles in the state α. Thus if in $|\psi\rangle$ there is a particle in such a state it will return 1 and the corresponding energy will be counted in H.

We thus see that if we have an Hamiltonian that is in a diagonal quadratic form such as (3.98) then we can get all the eigenvalues and eigenvectors of the system. At zero temperature the ground state will simply consist (for fermions) in occupying all the states with the minimum energy according to the number of particles in the system.

$$|\mathrm{F}\rangle = \prod_{i=1}^{\mathcal{N}} c^\dagger_{\alpha_i} |\varnothing\rangle \tag{3.102}$$

if $E_1 \leq E_2 \leq \ldots \leq E_\Omega$. Note that the Fermi sea is a particular case of (3.102).

At finite temperature we can also compute the averages of many operators. One important case is the operator giving the number of particles in the state α_p

$$\begin{aligned}
\langle c^\dagger_{\alpha_p} c_{\alpha_p} \rangle &= \frac{\mathrm{Tr}[e^{-\beta H} c^\dagger_{\alpha_p} c_{\alpha_p}]}{\mathrm{Tr}[e^{-\beta H}]} \\
&= \frac{\sum_{n_1,\dots,n_\Omega} \langle n_1,\dots,n_\Omega | e^{-\beta \sum_\alpha A_\alpha c^\dagger_\alpha c_\alpha} c^\dagger_{\alpha_p} c_{\alpha_p} | n_1,\dots,n_\Omega\rangle}{\sum_{n_1,\dots,n_\Omega} \langle n_1,\dots,n_\Omega | e^{-\beta \sum_\alpha A_\alpha c^\dagger_\alpha c_\alpha} | n_1,\dots,n_\Omega\rangle}
\end{aligned} \tag{3.103}$$

Using the fact that (both for fermions and bosons) $[c^\dagger_\alpha c_\alpha, c_\gamma] = 0$ if $\alpha \neq \gamma$ and a similar relation for $c^\dagger_\gamma$, we see that the term $e^{-\beta H}$ factorizes

$$e^{-\beta H} = \prod_{j=1}^{\Omega} e^{-\beta A_{\alpha_j} c^\dagger_{\alpha_j} c_{\alpha_j}} \tag{3.104}$$

Since in the trace each term n_i is independent the average factorizes. The numerator becomes

$$\left(\sum_{n_{\alpha_p}} \langle n_{\alpha_p}| e^{-\beta A_{\alpha_p} c^\dagger_{\alpha_p} c_{\alpha_p}} c^\dagger_{\alpha_p} c_{\alpha_p} |n_{\alpha_p}\rangle \right) \prod_{j \neq p} \left(\sum_{n_{\alpha_j}} \langle n_{\alpha_j}| e^{-\beta A_{\alpha_j} c^\dagger_{\alpha_j} c_{\alpha_j}} |n_{\alpha_j}\rangle \right) \tag{3.105}$$

All the terms for $j \neq p$ are identical in the numerator and denominator and cancel each other. The trace thus reduces to

$$\langle c^\dagger_{\alpha_p} c_{\alpha_p} \rangle = \frac{\sum_{n_{\alpha_p}} \langle n_{\alpha_p}| e^{-\beta A_{\alpha_p} c^\dagger_{\alpha_p} c_{\alpha_p}} c^\dagger_{\alpha_p} c_{\alpha_p} |n_{\alpha_p}\rangle}{\sum_{n_{\alpha_p}} \langle n_{\alpha_p}| e^{-\beta A_{\alpha_p} c^\dagger_{\alpha_p} c_{\alpha_p}} |n_{\alpha_p}\rangle} \tag{3.106}$$

which is quite obvious physically. Indeed the Hamiltonian being diagonal in α only the state α_p can contribute to the average of an operator only involving the state α_p.

Since $c^\dagger_{\alpha_p} c_{\alpha_p} |n_p\rangle = n_p |n_p\rangle$ one simply has

$$\langle c^\dagger_{\alpha_p} c_{\alpha_p} \rangle = \frac{\sum_{n_{\alpha_p}} e^{-\beta A_{\alpha_p} n_p} n_p}{\sum_{n_{\alpha_p}} e^{-\beta A_{\alpha_p} n_p}} \tag{3.107}$$

So far all what we did is independent on having bosons or fermions. The final result however will depend on what are the allowed values of n_p

For fermions only $n_p = 0$ and $n_p = 1$ are in the sum. Thus

$$\langle c^\dagger_{\alpha_p} c_{\alpha_p} \rangle = \frac{e^{-\beta A_{\alpha_p}}}{1 + e^{-\beta A_{\alpha_p}}} = \frac{1}{1 + e^{\beta A_{\alpha_p}}} \tag{3.108}$$

and one recovers the Fermi factor. We see that this is a totally general result (not limited to eigenstates of momentum) as soon as one has a quadratic Hamiltonian and one is in thermal equilibrium.

For bosons $n_p = 0, \ldots, +\infty$, and thus the sum becomes

$$\begin{aligned}
\langle c^\dagger_{\alpha_p} c_{\alpha_p} \rangle &= -\frac{\partial}{\partial \beta} \log[\sum_{n_p=0}^{+\infty} e^{-\beta n_p A_{\alpha_p}}] \\
&= -\frac{\partial}{\partial \beta} \log\left[\frac{1}{1 - e^{-\beta A_{\alpha_p}}}\right] \\
&= \frac{e^{-\beta A_{\alpha_p}}}{1 - e^{-\beta A_{\alpha_p}}} \\
&= \frac{1}{e^{\beta A_{\alpha_p}} - 1}
\end{aligned} \tag{3.109}$$

and one recovers the Bose factor.

Se we see that with quadratic diagonal Hamiltonians we can compute essentially all the physical quantities we need. Of course in general the Hamiltonian of the system will be neither quadratic nor diagonal. So solving in second quantization means essentially that we have to find a transformation of the operators c and $c^\dagger$ that put the Hamiltonian in a quadratic diagonal form. Although any transformation is in principle possible not all transformations are good. We want the new operators d and $d^\dagger$ that are the results of the transformation to still spawn the Fock space. It means that we can only consider transformation that preserve the canonical commutation relations. Of course finding such transformations is in general a formidable task. There is however a very important class of transformations when the Hamiltonian is still a quadratic form, but is not diagonal, that we will examine in the next section.

Before doing so, let us finally remark that even without solving the Hamiltonian one can exploit the freedom of choosing different creation and destruction operators to use a more convenient representation. As already mentioned *every* transformation that preserves the canonical commutation relations is allowed. Let us a simple example, more examples will be seen in the next section and in the exercises.

The simplest transformation is the particle-hole transformation

$$\begin{aligned} c_\alpha^\dagger &= d_\alpha \\ c_\alpha &= d_\alpha^\dagger \end{aligned} \tag{3.110}$$

For fermions is it easy to check, by substitution of the d operators that they verify the canonical anticommutation relations. For example

$$[d_\alpha, d_\beta^\dagger]_+ = [c_\alpha^\dagger, c_\beta]_+ = \delta\alpha, \beta \tag{3.111}$$

If the operators c_α and $c_\alpha^\dagger$ respectively destroy and create an electron in the state α, the operator d_α and $d_\alpha^\dagger$ are also destruction and creation operators of "something else" that has also a fermionic statistics and can thus be used to build a Fock space. In that particular case the operator d_α destroys a hole with state α (which is identical to creating an electron) and the operator $d_\alpha^\dagger$ creates a hole (which is the same thing than destroying an electron).

One important point when making the transformation is not to forget to modify the vacuum as well. Indeed the vacuum for the particles d is the not same than the vacuum for the particles c. One has thus $|\varnothing_c\rangle$ and $|\varnothing_d\rangle$. The vacuum of d particles is defined as usual by

$$d_\alpha |\varnothing_d\rangle = 0 \tag{3.112}$$

for all states α. It is easy to check using the relation (3.110) that

$$|\varnothing_d\rangle = \prod_\alpha c_\alpha^\dagger |\varnothing_c\rangle \tag{3.113}$$

3.6.2 Bogoliubov transformation

In principle we know also how to solve Hamiltonians of the form

$$H = \sum_{\alpha,\beta} A_{\alpha,\beta} c_\alpha^\dagger c_\beta \tag{3.114}$$

Indeed in that case the Hamiltonian is quadratic but not diagonal. This is of course not the most general form of the Hamiltonian since there are terms that can be different from a quadratic form (such as an interaction term for example). However this class of Hamiltonian already covers a very large number of cases. We can bring the matrix $A_{\alpha,\beta}$ into a diagonal form by making the appropriate linear combination of the operators c_α. The operators d_α that are the linear combinations of the c_α that brings the matrix under diagonal form allow thus to get all the eigenvectors of the problem. Of course in general the matrix $A_{\alpha,\beta}$ would be of size $\Omega \times \Omega$ and the diagonalization would be very difficult to perform. There are however simple cases where the physics of the problem helps us to perform the diagonalization.

Let us illustrate it with the tight binding Hamiltonian (1.18). This will also allows us to write this Hamiltonian in second quantized form. A complete basis is provided by the states on each site $|i\rangle$ and we can thus define the creation and destruction operators associated with it, i.e. $c_i^\dagger$ is the operator creating a particle on site i. Using (1.18), the second quantized form reads

$$H = \sum_{i,j} (i|H^{(1)}|j) c_i^\dagger c_j \tag{3.115}$$

where $H^{(1)}$ is the Hamiltonian (1.18). We thus immediately get

$$H = E_0 \sum_i c_i^\dagger c_i - t \sum_{\langle i,j\rangle} c_i^\dagger c_j \tag{3.116}$$

The second term describes a process where a particle on site i to reappear on the neighboring site j, and vice versa. This Hamiltonian is obviously quadratic but not diagonal. In the language of (3.114) it corresponds to a tridiagonal matrix. In order to diagonalize this matrix we can use the physical input that since the Hamiltonian is invariant by translation, momentum must be a good quantum number. We thus have interest to deal with creation and destruction operators that are linear combination of the c_i that correspond to a Fourier transform. Of course this is exactly the same reasoning than the one leading to (1.19) and this operators will correspond to the state $|k\rangle$. But let us simply treat it as a linear combination of operators. We can define

$$d_k^\dagger = \frac{1}{\sqrt{N_s}} \sum_{j=0}^{N_s-1} e^{ikr_j} c_j^\dagger \tag{3.117}$$

where we have used a different name d to emphasize that this is new operators (in the future we will use the same letter the index r or k, making it clear that one is the Fourier transform of the other. Note the analogy between (3.117) and (3.72). One can check that the operators d_k satisfy the canonical commutation rules. We will just verify one of the anticommutators and leave the other relations as an exercise. One has

$$\begin{aligned}
[d_{k_1}, d_{k_2}^\dagger]_+ &= \frac{1}{N_s} \sum_{j,l} e^{-ik_1 r_j} e^{ik_2 r_l} [c_j, c_l^\dagger]_+ \\
&= \frac{1}{N_s} \sum_{j,l} e^{-ik_1 r_j} e^{ik_2 r_l} \delta_{j,l} \\
&= \frac{1}{N_s} \sum_{j} e^{i(k_2-k_1) r_j} \\
&= \delta_{k_1,k_2}
\end{aligned} \tag{3.118}$$

The operators d_k are thus acceptable Fermions operators. There are exactly N_s different operators (the size of the Hilbert space cannot change) and k is confined inside the first Brillouin zone $k \in [-\pi/a, \pi/a]$ as discussed for the first quantization solution. Moreover, as is obvious from the definition (3.117), $|\varnothing_d\rangle = |\varnothing_c\rangle$. The transformation (3.117) is easily inverted

$$c_j^\dagger = \frac{1}{\sqrt{N_s}} \sum_k e^{-ikr_j} d_k^\dagger \tag{3.119}$$

and thus replacing the c_j in (3.116) and after a little bit of algebra, one finds

$$H = E_0 \sum_k d_k^\dagger d_k - \sum_k 2t\cos(ka) d_k^\dagger d_k \tag{3.120}$$

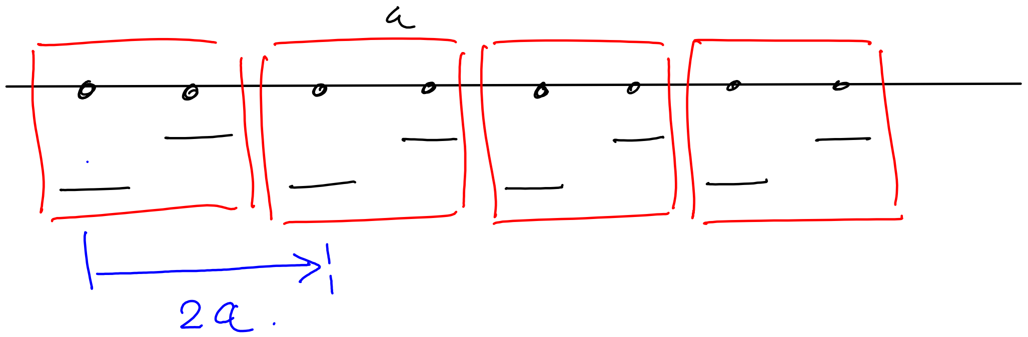


Figure 3.4: A tight binding model with a periodic potential of the form $(-1)^i$. One can still see the problem as a system invariant by translation with a unit cell that now contains two atoms. In this reduced scheme zone, the momentum is still a conserved quantum number and can be used to severely reduce the size of the matrix to diagonalize.

This expression can be compared with the first quantized form (1.29). The Hamiltonian being now diagonal we can use the operators d_k to express simply the ground state and the various averages. On the physical level we have used the fact that since momentum is conserved one can diagonalize simultaneously the momentum operators and the Hamiltonian. The Hamiltonian is thus block diagonal in the basis of eigenvectors of the momentum operator. Since this basis is of size N_s (N_s different k values in the first Brillouin zone) we are left for each value of k with a 1×1 matrix to diagonalize. This is why we find that all the job is done.

Let us now look at the slightly more complicated problem of the tight binding Hamiltonian (3.116) to which we have added a periodic potential, show in Fig. 3.4. The potential is in first quantization

$$H^{(1)}_{\text{pot}} = V \sum_i (-1)^i |i\rangle\langle i| \tag{3.121}$$

which immediately gives us in second quantization

$$H_{\text{pot}} = V \sum_i (-1)^i c^\dagger_i c_i \tag{3.122}$$

Since we know how to diagonalize (3.116) using the operators d_k, it is natural to express H_{pot} in this basis. One has, using that $(-1)^j = e^{i\frac{\pi}{a} r_j}$

$$H_{\text{pot}} = V \sum_j e^{i\frac{\pi}{a} r_j} \frac{1}{N_s} \sum_{k_1,k_2} e^{-ik_1 r_j} e^{-ik_2 r_j} d^\dagger_{k_1} d_{k_2} \tag{3.123}$$

where the two vectors k_1 and k_2 belong to the first Brillouin zone $k \in [-\pi/a, \pi/a]$. The sum over j can be easily performed and gives the constraint that $k_1 = k_2 + \pi/a$ modulo $2\pi/a$. The potential becomes

$$H_{\text{pot}} = V \sum_k c^\dagger_{k+\pi(\mathrm{mod} 2\pi)/a} c_k \tag{3.124}$$

the momentum $k + \pi/a$ has to be brought back in the first Brillouin zone by the proper translation by $\pm 2\pi/a$. The potential thus couples two different k values. Thus the tight binding Hamiltonian is diagonal in neither basis. In the real space basis the tight binding Hamiltonian (3.116) is non diagonal while the potential is, while in the k basis the tight binding Hamiltonian is diagonal but the potential is not.

To diagonalize the total Hamiltonian we can notice that although a state k is coupled to a state $k+\pi/a$ the coupling stops there since the state $k+\pi/a$ is coupled to $k+2\pi/a \equiv k$ and thus linear combinations of d_k and $d_{k+\pi}$ should be able to diagonalize the full Hamiltonian. We can also use the more physical approach to use the symmetries of the problem. Even in the presence of the potential the system is still invariant by translation provided that one considers a unit cell with two atoms per unit cell as shown in Fig. 3.4. For such a system where the lattice spacing is $2a$ the Brillouin zone $Z'_B = [-\pi/(2a), \pi/(2a)]$. On has thus two perfectly equivalent ways of looking at this problem

1. The original lattice spacing: in that case one has one atom per unit cell, i.e. a single band in the system. The set of allowed k values spans the Brillouin zone $Z_B = [-\pi/a, \pi/a]$. In such a representation the momentum is not a good quantum number since the potential breaks the invariance by translation.

2. A reduced zone scheme: in that case one has two atoms per unit cell, i.e. two bands in the system. The set of allowed k values spans the reduced Brillouin zone $Z'_B = [-\pi/(2a), \pi/(2a)]$ associated with a unit cell $2a$. Of course we see that although the number of allowed values of k has been divided by two the total number of states is conserved since there are now two bands corresponding to each value of k instead of one. In this reduced scheme, the system is invariant by translation and thus the momentum is a good quantum number. We can thus be sure that the Hamiltonian will be block diagonal in each value of k.

Let us use the reduced zone scheme. We need to express the operators d_k with $k \in Z_B$ in terms of new operators expressing the degrees of freedom in the reduced zone. This is illustrated in Fig. 3.5. Let us consider the mapping (shown in Fig. 3.5

$$\begin{aligned} d_k &= \alpha_k \quad k \in [-\pi/(2a), \pi/(2a)] \\ d_k &= \beta_{k-\pi/a} \quad k \in [\pi/(2a), \pi/a] \\ d_k &= \beta_{k+\pi/a} \quad k \in [-\pi/a, -\pi/(2a)] \end{aligned} \tag{3.125}$$

It is easy to check that the operators α_k and β_k are defined for $k \in Z'_B$. They also obey the canonical commutation rules, and are thus faithful representations of fermions. In terms of this operator one can rewrite the Hamiltonian. Let us start with the tight binding part [we use $\xi(k) = -2t\cos(ka)$]

$$\begin{aligned} \sum_{k\in Z_B} \xi(k) c^\dagger_k c_k &= \sum_{k\in Z'_B} \xi(k) c^\dagger_k c_k + \sum_{k\in[\pi/(2a),\pi/a]} \xi(k) c^\dagger_k c_k + \sum_{k\in[-\pi/a,-\pi/(2a)]} \xi(k) c^\dagger_k c_k \\ &= \sum_{k\in Z'_B} \xi(k) \alpha^\dagger_k \alpha_k + \sum_{k\in[-\pi/(2a),0]} \xi(k+\pi/a) \beta^\dagger_k \beta_k + \sum_{k\in[0,\pi/(2a)]} \xi(k-\pi/a) \beta^\dagger_k \beta_k \\ &= \sum_{k\in Z'_B} \xi(k)(\alpha^\dagger_k \alpha_k - \beta^\dagger_k \beta_k) \end{aligned} \tag{3.126}$$

having used $\xi(k \pm \pi/a) = -\xi(k)$. In the formulation (3.126) the two bands are clearly visible. Of course since the tight binding Hamiltonian is also invariant by a one unit cell translation the energy dispersion has no gap as the zone boundary, since folding the zone is merely an arbitrary

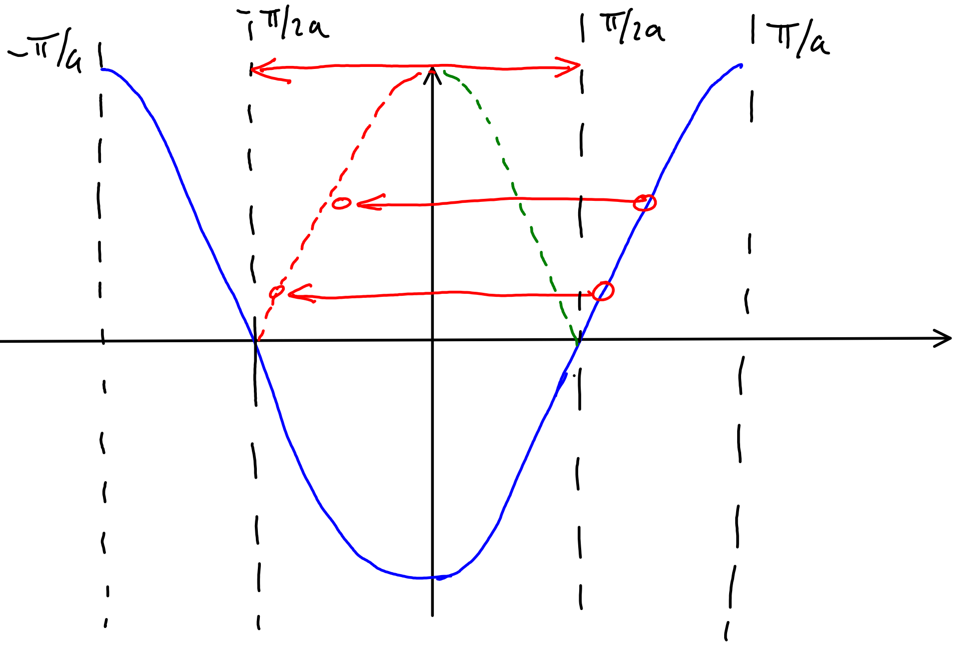


Figure 3.5: One can go from the extended scheme zone to the reduced one. There are now two species of fermions corresponding to the lower and upper band.

choice here. We can now express the potential

$$\begin{aligned}
\sum_{k\in Z_B} c^\dagger_{k+\pi/a(\text{mod}2\pi/a)}c_k = & \sum_{k\in[-\pi/a,-\pi/(2a)]} c^\dagger_{k+\frac{\pi}{a}(\text{mod}\frac{2\pi}{a})}c_k + \sum_{k\in[-\pi/(2a),0]} c^\dagger_{k+\frac{\pi}{a}(\text{mod}\frac{2\pi}{a})}c_k \\
& + \sum_{k\in[0,\pi/(2a)]} c^\dagger_{k+\frac{\pi}{a}(\text{mod}\frac{2\pi}{a})}c_k + \sum_{k\in[\pi/(2a),\pi/a]} c^\dagger_{k+\frac{\pi}{a}(\text{mod}\frac{2\pi}{a})}c_k \\
= & \sum_{k\in[0,\pi/(2a)]} \alpha^\dagger_k\beta_k + \sum_{k\in[-\pi/(2a),0]} \beta^\dagger_k\alpha_k \\
& + \sum_{k\in[0,\pi/(2a)]} \beta^\dagger_k\alpha_k + \sum_{k\in[-\pi/(2a),0]} \alpha^\dagger_k\beta_k \\
= & \sum_{k\in Z'_B} (\alpha^\dagger_k\beta_k + \beta^\dagger_k\alpha_k)
\end{aligned} \tag{3.127}$$

The final result is remarkably simple

$$H = \sum_{k\in Z'_B} [A(k)(\beta^\dagger_k\beta_k - \alpha^\dagger_k\alpha_k) + V(\alpha^\dagger_k\beta_k + \beta^\dagger_k\alpha_k) \tag{3.128}$$

were we have introduced $A(k) = -\xi(k)$ which is a positive quantity for $k \in Z'_B$ We see that the Hamiltonian is diagonal in k in the reduced zone, as it should since the system is invariant by translation. However, it is not fully diagonal since the lower band α is coupled to the upper band β by the potential V. We can rewrite the Hamiltonian in a matrix form

$$H = \sum_{k\in Z'_B} \begin{pmatrix} \alpha^\dagger_k & \beta^\dagger_k \end{pmatrix} \begin{pmatrix} -A(k) & V \\ V & A(k) \end{pmatrix} \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} \tag{3.129}$$

This structure makes it clear that the Hamiltonian is diagonal in k, but that there are two residual degrees of freedom that are coupled by the potential making the matrix of the quadratic form non diagonal. It also makes it obvious that all we have to do is to diagonalized the matrix by the appropriate linear combination of the operators α and β. This type of Hamiltonian where two degrees of freedom are coupled is specially important and arises in several occasions, for example as a result of a mean field approximation of the real Hamiltonian as we will see. It is thus important to see the general procedure to solve it. It is known as the Bogoliubov transformation and is nothing but the above mentioned diagonalization of the matrix.

On the form (3.129) we can read directly the eigenvalues. They are given by the equation

$$(A(k) - E)(-A(k) - E) - V^2 = 0 \tag{3.130}$$

and thus

$$E(k) = \pm\sqrt{A(k)^2 + V^2} \tag{3.131}$$

As can be seen on Fig. 3.6 there is now a gap $2V$ that opens at the reduced zone boundary. To determine the eigenvectors we must make combination of the operators α and β.

$$\begin{pmatrix} \gamma_{k-} \\ \gamma_{k+} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} \tag{3.132}$$

which defines new operators γ_{k+} and γ_{k-}. These operators result from the combination of the top and bottom band or in the original language of states with momentum k and momentum $k + \pi/a$ and describe the interference effect between these two states due to the scattering potential. Since the new operators γ must verify the canonical commutation relations, the

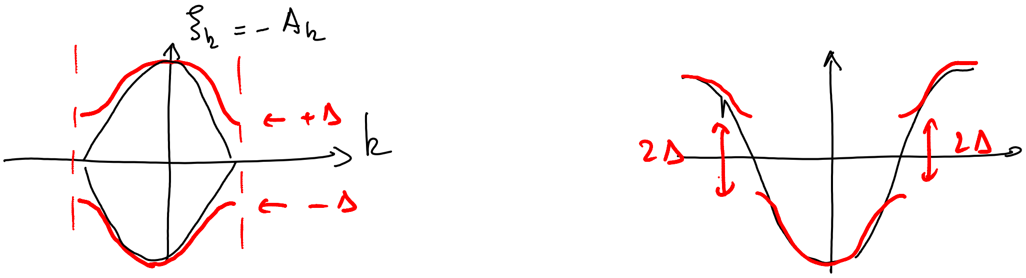


Figure 3.6: The dispersion relation $E(k)$. Because of the potential there is now a gap $2V$ at the zone edge. This gap results from the interferences of states of momentum k and $k + \pi/a$ on the periodic potential.

coefficients a, b, c, d cannot be arbitrary. In order to ensure $[\gamma_{k+}, \gamma_{k-}^\dagger]_+ = 0$ and the other commutators, one can check that is is necessary to take the coefficients as

$$\begin{pmatrix} \gamma_{k-} \\ \gamma_{k+} \end{pmatrix} = \begin{pmatrix} u_k & -v_k \\ v_k & u_k \end{pmatrix} \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} \tag{3.133}$$

moreover one must have $u_k^2 + v_k^2 = 1$. Note that here we have chosen real coefficients but one can generalize the transformation to complex ones. The last condition comes for example from

$$\begin{aligned} [\gamma_{k-}, \gamma_{k-}^\dagger]_+ &= [u_k\alpha_k - v_k\beta_k, u_k\alpha_k^\dagger - v_k\beta_k^\dagger]_+ \\ &= u_k^2[\alpha_k, \alpha_k^\dagger]_+ + v_k^2[\beta_k, \beta_k^\dagger]_+ + u_kv_k([\beta_k, \alpha_k^\dagger]_+ + [\alpha_k, \beta_k^\dagger]_+) \\ &= u_k^2 + v_k^2 \end{aligned} \tag{3.134}$$

This quite natural since the transformation (3.133) (the Bogoliubov tranformation) is an orthogonal transformation with two vectors that are orthogonal and normed. It is thus a “rotation” in the space spawned by the vectors α_k and β_k. Note that in general the coefficients of the transformation depends on k. Given the constrain one often uses the parametrization

$$u_k = \cos\theta_k \quad , \quad v_k = \sin\theta_k \tag{3.135}$$

The transformation is easily inverted

$$\begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} = \begin{pmatrix} u_k & v_k \\ -v_k & u_k \end{pmatrix} \begin{pmatrix} \gamma_{k-} \\ \gamma_{k+} \end{pmatrix} \tag{3.136}$$

Inserting (3.136) in (3.129) one gets

$$H = \sum_{k\in Z_B'} \begin{pmatrix} \gamma_{k-}^\dagger & \gamma_{k+}^\dagger \end{pmatrix} \begin{pmatrix} -[A(k)(u_k^2 - v_k^2) + 2Vu_kv_k] & V(u_k^2 - v_k^2) - 2A(k)u_kv_k \\ V(u_k^2 - v_k^2) - 2A(k)u_kv_k & +[A(k)(u_k^2 - v_k^2) + 2Vu_kv_k] \end{pmatrix} \begin{pmatrix} \gamma_{k-} \\ \gamma_{k+} \end{pmatrix} \tag{3.137}$$

This leads to the condition

$$V(u_k^2 - v_k^2) - 2A(k)u_kv_k = 0 \tag{3.138}$$

to diagonalize the matrix, while

$$E_\pm(k) = \pm[A(k)(u_k^2 - v_k^2) + 2Vu_kv_k] \tag{3.139}$$

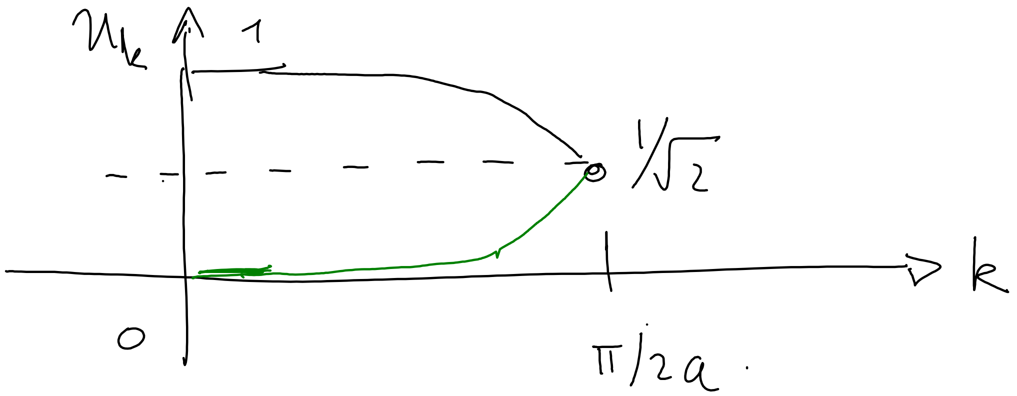


Figure 3.7: Bogoliubov parameters u_k and v_k as a function of the energy $A(k)$.

are the two eigenvalues. To solve (3.138) one can use the parametrization (3.135). The solution is

$$\begin{aligned} u_k &= \left[\frac{1}{2}\left(1 + \frac{A(k)}{\sqrt{A(k)^2 + V^2}}\right)\right]^{1/2} \\ v_k &= \left[\frac{1}{2}\left(1 - \frac{A(k)}{\sqrt{A(k)^2 + V^2}}\right)\right]^{1/2} \end{aligned} \tag{3.140}$$

which also gives

$$E_\pm(k) = \pm[A(k)(u_k^2 - v_k^2) + 2Vu_kv_k] = \pm[A(k)^2 + V^2]^{1/2} \tag{3.141}$$

which are of course the same values than the one we obtained by computing directly the eigenvalues.

The expressions (3.140) and the energies (3.141) are shown in Fig. 3.7 One sees that far from the zone boundary where $A(k) = 0$ one has $u_k \sim 1$ and $v_k \sim 0$. The new operators $\gamma_\pm$ are thus practically the old ones $\gamma_{k-} \simeq \alpha_k$ and $\gamma_{k+} \simeq \beta_k$. For these wavevectors the potential is not able to act efficiently and produce interferences. The wavefunctions of the eigenstates are thus essentially simply the old plane waves. One the other hand close to the zone boundary the interferences produced by the potential will be maximal. One can look at the wavefunctions corresponding to the new operators

$$\begin{aligned} \langle r_i|\gamma_{k-}^\dagger|\varnothing\rangle &= \langle r_i|(u_k\alpha_k^\dagger - v_k\beta_k^\dagger)|\varnothing\rangle \\ &= \langle r_i|(u_kc_k^\dagger - v_kc_{k+\pi/a}^\dagger)|\varnothing\rangle \\ &= u_k\frac{1}{\sqrt{\Omega}}e^{ikr_i} - v_k\frac{1}{\sqrt{\Omega}}e^{i(k+\pi/a)r_i} \\ &= \frac{1}{\sqrt{\Omega}}e^{ikr_i}(u_k - (-1)^iv_k) \end{aligned} \tag{3.142}$$

and

$$\begin{aligned} \langle r_i|\gamma^\dagger_{k+}|\varnothing\rangle &= \langle r_i|(v_k\alpha^\dagger_k + u_k\beta^\dagger_k)|\varnothing\rangle \\ &= \frac{1}{\sqrt{\Omega}}e^{ikr_i}(v_k + (-1)^i u_k) \\ &= \frac{1}{\sqrt{\Omega}}e^{i(k+\pi/a)r_i}(u_k + (-1)^i v_k) \end{aligned} \tag{3.143}$$

we thus see that when $u_k \sim 1$ and $v_k \sim 0$ the two operators γ_{k-} and γ_{k+} are creating respectively essentially a plane wave with momentum k and $k + \pi/a$, and the potential has little action. But when the wavevector of these plane waves approaches $\pm\pi/(2a)$, i.e. the zone boundary in the reduced scheme zone for which $A(k) = 0$, the potential that has a wavevector $Q = \pi/a$ can induce strong interferences between the two states $\pm\pi/(2a)$. The eigenstates are thus transformed because of these interferences into stationary waves. The amplitude of the wave caused by γ_{k-} is essentially zero on the even sites and maximal on the odd sites. Since the potential is $V(r_i) = V(-1)^i$ we see that such a wave gain an energy of order $-V$, which is indeed what its eigenvalue is. On the contrary the stationary wave created by γ_{k+} is maximal on the even sites and thus looses the energy $+V$. The Bogoliubov transformation allows to go beyond this cartoon and gives the full eigenstates for all values of k.

To finish let us illustrate how we can use the transformation to compute any observable. The observables are easily expressed in terms of the original operators c, while we know the ground state or other thermodynamic averages easily in terms of the operators γ. The strategy can work two ways: either we express the ground state back in terms of the operators c, and then compute the observables by the usual way, or we do the opposite and reexpress the observables in terms of the operators γ. The second method is usually simpler, but both give of course equivalent results. Let us illustrate it with the calculation of the average, at finite temperature, of the density of particles at the point $r_i = 0$. Using (3.71) the corresponding operator is

$$\rho(r = 0) = \frac{1}{\Omega}\sum_{k_1,k_2} c^\dagger_{k_1}c_{k_2} \tag{3.144}$$

the sum over momentum runs in Z_B. We now restrict the momenta to Z'_B and express the operators c in terms of the operators α and β. One has

$$\begin{aligned} \rho(r = 0) &= \frac{1}{\Omega}\sum_{(k_1,k_2)\in Z'_B}(\alpha^\dagger_{k_1}\alpha_{k_2} + \beta^\dagger_{k_1}\beta_{k_2} + \alpha^\dagger_{k_1}\beta_{k_2} + \beta^\dagger_{k_1}\alpha_{k_2}) \\ &= \frac{1}{\Omega}\sum_{(k_1,k_2)\in Z'_B}\begin{pmatrix} \alpha^\dagger_{k_1} & \beta^\dagger_{k_1}\end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1\end{pmatrix}\begin{pmatrix}\alpha_{k_2} \\ \beta_{k_2}\end{pmatrix} \\ &= \frac{1}{\Omega}\sum_{(k_1,k_2)\in Z'_B}\begin{pmatrix} \gamma^\dagger_{k_1-} & \gamma^\dagger_{k_1+}\end{pmatrix}\begin{pmatrix} (u_{k_1} - v_{k_1})(u_{k_2} - v_{k_2}) & (u_{k_1} - v_{k_1})(u_{k_2} + v_{k_2}) \\ (u_{k_1} + v_{k_1})(u_{k_2} - v_{k_2}) & +(u_{k_1} + v_{k_1})(u_{k_2} + v_{k_2})\end{pmatrix}\begin{pmatrix}\gamma_{k_2-} \\ \gamma_{k_2+}\end{pmatrix} \end{aligned} \tag{3.145}$$

Computing the average of the operator is now reduced to computing averages of the form $\gamma^\dagger_{k_1\pm}\gamma_{k_2\pm}$ which we now well since the Hamiltonian is diagonal in terms of the operators γ. Let us start with the case at $T = 0$. Given the eigenvalues (3.141) the ground state of the system consists in filling first the states γ_{k-} and then the states γ_{k+}. The filling depends on the total number of particles. For example for a case when one has one particle every two sites (i.e. and half filled band since we did not put spin for the particles) all states γ_{k-} will be filled and all states γ_{k+} will be empty. The ground state is thus

$$|G\rangle = \prod_{k\in Z'_B}\gamma^\dagger_{k-}|\varnothing\rangle \tag{3.146}$$

In that case the only non zero average for the operators $\gamma^\dagger_{k_1\pm}\gamma_{k_2\pm}$ is

$$\langle G|\gamma^\dagger_{k_1\pm}\gamma_{k_2\pm}|G\rangle = \delta_{k_1,k_2} \tag{3.147}$$

and the average of the operator is

$$\begin{aligned}\langle G|\rho(r=0)|G\rangle &= \frac{1}{\Omega}\sum_{k\in Z'_B}(u_k - v_k)^2 \\ &= \frac{1}{\Omega}\sum_{k\in Z'_B}\left[1 - \frac{V}{\sqrt{A(k)^2+V^2}}\right]\end{aligned} \tag{3.148}$$

which we can compute explicitly. For $V = 0$, we get back that $\rho(r = 0) = 1/2$ as it should since the number of k values in Z'_B is half of the ones in the total zone. We also see that when $V \to \infty$, $\rho(r = 0) \to 0$ since the particles tend to avoid the site where there is the gigantic potential $+V$.

Thanks to the fact that γ are eigenstates of the Hamiltonian we can even compute easily the averages at *finite* temperature. In that case, using (3.108) we see that the only non zero averages are

$$\begin{aligned}\langle\gamma^\dagger_{k_1-}\gamma_{k_2-}\rangle &= \delta_{k_1,k_2} f_{\rm F}(-E(k)) \\ \langle\gamma^\dagger_{k_1+}\gamma_{k_2+}\rangle &= \delta_{k_1,k_2} f_{\rm F}(+E(k))\end{aligned} \tag{3.149}$$

The average of the density at $r = 0$ thus becomes

$$\begin{aligned}\langle\rho(r=0)\rangle &= \frac{1}{\Omega}\sum_{k\in Z'_B}(u_k - v_k)^2 f_{\rm F}(-E(k)) + (u_k + v_k)^2 f_{\rm F}(+E(k)) \\ &= \frac{1}{\Omega}\sum_{k\in Z'_B}\left(1 - \frac{V}{\sqrt{A(k)^2+V^2}}\tanh(\frac{\beta E(k)}{2})\right)\end{aligned} \tag{3.150}$$

When $\beta \to 0$ we obviously recover the previous result, while when $\beta \to 0$ we see that $\langle\rho(r = 0)\rangle = 1/2$ since the potential or kinetic energy do not matter any more when compared to the thermal energies.

The Bogoliubov transformation is thus a very useful tool to solve *quadratic* but non diagonal Hamiltonians in second quantization. One of its remarkable uses is that the one is not limited to Hamiltonians of the form $c^\dagger c$ but can treat forms that contain $c^\dagger c^\dagger$ terms as well. For example the Hamiltonian

$$H_{BCS} = \sum_k \xi(k)(c^\dagger_{k\uparrow}c_{k\uparrow} + c^\dagger_{k\downarrow}c_{k\downarrow}) + \Delta\sum k(c^\dagger_{k\uparrow}c^\dagger_{-k\downarrow} + c_{-k\downarrow}c_{k\uparrow}) \tag{3.151}$$

is the so called Bardeen-Cooper-Schrieffer Hamiltonian which is the basis for the theory of superconductivity. The fist term is simply the kinetic energy, the second term represents the creation and destruction of pairs of electrons. Note that this Hamiltonian does not conserve the number of particles and has thus no simple expression in first quantization. However, since it is quadratic one can treat this Hamiltonian by the Bogoliubov transformation. In fact using the canonical transformation

$$\begin{aligned}c^\dagger_{k\uparrow} &= \alpha^\dagger_k \\ c^\dagger_{-k\downarrow} &= \beta_k\end{aligned} \tag{3.152}$$

Note that this transformation mixes creation and destruction operators. This is one great advantage of the second quantization to allow easily for such operations. In the transformation it is also important to remember that

$$|\varnothing_{\alpha,\beta}\rangle = \prod_k c^\dagger_{k\downarrow} \left|\varnothing_{c_\uparrow,c_\downarrow}\right\rangle \tag{3.153}$$

In terms of the operators α and β the BCS Hamiltonian becomes

$$\begin{aligned} H_{BCS} &= \sum_k \xi(k)\alpha^\dagger_k\alpha_k + \xi(-k)\beta_k\beta^\dagger_k + \Delta\sum_k(\alpha^\dagger_k\beta_k + \beta^\dagger_k\alpha_k) \\ &= \sum_k \xi(k)\alpha^\dagger_k\alpha_k + \xi(-k)(1-\beta^\dagger_k\beta_k) + \Delta\sum_k(\alpha^\dagger_k\beta_k + \beta^\dagger_k\alpha_k) \\ &= \sum_k \xi(k)(\alpha^\dagger_k\alpha_k - \beta^\dagger_k\beta_k) + \Delta\sum_k(\alpha^\dagger_k\beta_k + \beta^\dagger_k\alpha_k) + \sum_k \xi(k) \end{aligned} \tag{3.154}$$

since for most dispersions $\xi(k) = \xi(-k)$. This is, up to a simple constant, exactly the Hamiltonian we already examined, and it thus can be solved by exactly the same transformations.

CHAPTER 4

Fermi liquid theory

Now that we have the tools to tackle interacting systems, let us examine the properties of interacting quantum particles.

4.1 Interaction Hamiltonian

Let us first write the general interaction Hamiltonian that we will use. The first term is the kinetic energy of the system of the form

$$H_{\text{kin}} = \sum_{\boldsymbol{k}\sigma} \varepsilon(\boldsymbol{k}) c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} \tag{4.1}$$

where the sum over k runs in general over the first Brillouin zone. One usually incorporates the chemical potential in the energy $\xi(\boldsymbol{k}) = \varepsilon(\boldsymbol{k}) - E_{\text{F}}$ to make sure that $\xi(\boldsymbol{k}) = 0$ at the Fermi level. In addition to the kinetic energy one has the standard interaction term (3.97)

$$H_{\text{int}} = \frac{1}{2\Omega} \sum_{\substack{\boldsymbol{k}_1 \boldsymbol{k}_2 \boldsymbol{q}, \\ \sigma_1 \sigma_2}} V(\boldsymbol{q}) c^\dagger_{\boldsymbol{k}_1+\boldsymbol{q},\sigma_1} c^\dagger_{\boldsymbol{k}_2-\boldsymbol{q},\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \tag{4.2}$$

In a solid the potential $V(r)$ is in general the Coulomb potential

$$V(\boldsymbol{r}) = \frac{e^2}{4\pi\epsilon r^2} \tag{4.3}$$

However, both for systems such as ^{3}He and if one wants to take simply into account the screening of the Coulomb interaction one can consider a short range interaction. In that case for simplicity one very often takes a contact interaction

$$V(\boldsymbol{r}) = U\delta(\boldsymbol{r}) \tag{4.4}$$

Although for such an interaction one could keep the above form of the Hamiltonian it is important to notice that for a contact interaction the terms in (6.34) with equal spins become in real space (see e.g. (3.90) :

$$\begin{aligned} \int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2) c^\dagger_{\boldsymbol{r}_1\sigma_1} c^\dagger_{\boldsymbol{r}_2\sigma_1} c_{\boldsymbol{r}_2\sigma_1} c_{\boldsymbol{r}_1\sigma_1} \\ = U \int d\boldsymbol{r} c^\dagger_{\boldsymbol{r}_1\sigma_1} c^\dagger_{\boldsymbol{r}_1\sigma_1} c_{\boldsymbol{r}_1\sigma_1} c_{\boldsymbol{r}_1\sigma_1} \\ = 0 \end{aligned} \tag{4.5}$$

because of the Pauli principle that prevent to identical electrons to be at the same point. We thus see that for a *contact* interaction, even if the interaction is full spin rotation invariant the terms in the interaction for which the spins are parallel disappear. It is thus natural for such interactions to only include the interaction term for the two opposite spins. This avoids mistakes when one treats the interaction term approximately (e.g. as we will see using a mean field approximation in the next chapter), since in that case one may miss the subtle cancellation due to the Pauli principe that should be obeyed in an exact treatment of the problem. It is thus simpler to separate from the start in the interaction the components for which the two spins are parallel and to denote the corresponding term $V_{\|}$ and for which they are opposite and to use $V_{\perp}$. We will thus use the form

$$H_{\text{int}} = \frac{1}{2\Omega} \sum_{\substack{\boldsymbol{k}_1 \boldsymbol{k}_2 \boldsymbol{q}, \\ \sigma_1}} [V_{\|}(\boldsymbol{q}) c^{\dagger}_{\boldsymbol{k}_1+\boldsymbol{q},\sigma_1} c^{\dagger}_{\boldsymbol{k}_2-\boldsymbol{q},\sigma_1} c_{\boldsymbol{k}_2\sigma_1} c_{\boldsymbol{k}_1\sigma_1} + V_{\perp}(\boldsymbol{q}) c^{\dagger}_{\boldsymbol{k}_1+\boldsymbol{q},\sigma_1} c^{\dagger}_{\boldsymbol{k}_2-\boldsymbol{q},-\sigma_1} c_{\boldsymbol{k}_2,-\sigma_1} c_{\boldsymbol{k}_1\sigma_1}] \quad (4.6)$$

or its equivalent form in real space

$$H_{\text{int}} = \frac{1}{2} \sum_{\sigma_1} \int d\boldsymbol{r}_1 d\boldsymbol{r}_2 [V_{\|}(\boldsymbol{r}_1 - \boldsymbol{r}_2) c^{\dagger}_{\boldsymbol{r}_1\sigma_1} c^{\dagger}_{\boldsymbol{r}_2\sigma_1} c_{\boldsymbol{r}_2\sigma_1} c_{\boldsymbol{r}_1\sigma_1} + V_{\perp}(\boldsymbol{r}_1 - \boldsymbol{r}_2) c^{\dagger}_{\boldsymbol{r}_1\sigma_1} c^{\dagger}_{\boldsymbol{r}_2,-\sigma_1} c_{\boldsymbol{r}_2,-\sigma_1} c_{\boldsymbol{r}_1\sigma_1}] \quad (4.7)$$

For the Coulomb interaction we take $V_{\|} = V_{\perp}$ while for a contact interaction we can set $V_{\|} = 0$ without affecting the result.

One important particular case of the above Hamiltonian is the case of a contact interaction for a system on a lattice. In that case the kinetic energy (4.1) can be represented by a tight-binding Hamiltonian (3.116). The contact interaction (4.7) becomes simply

$$\begin{aligned} H_{\text{int}} &= \frac{U}{2} \sum_{\sigma_1} \sum_{i} c^{\dagger}_{i\sigma_1} c^{\dagger}_{i,-\sigma_1} c_{i,-\sigma_1} c_{i\sigma_1} \\ &= \frac{U}{2} \sum_{\sigma_1} \sum_{i} c^{\dagger}_{i\sigma_1} c_{i\sigma_1} c^{\dagger}_{i,-\sigma_1} c_{i,-\sigma_1} \\ &= \frac{U}{2} \sum_{\sigma_1} \sum_{i} n_{i,\sigma_1} n_{i,-\sigma_1} \\ &= U \sum_{i} n_{i,\uparrow} n_{i,-\downarrow} \end{aligned} \quad (4.8)$$

the physical interpretation of this term is very simple. Since the interaction is a contact interaction, the electrons of the same spin cannot feel it since they cannot be on the same sites of the lattice. Thus only the interaction between electrons of opposite spins remains. Two electron of opposite spins only feel the interaction when they are sitting on the same site, thus when both the density of spin up $n_{i\uparrow}$ and the density of spin down $n_{i\downarrow}$ on this site verify $n_{i\uparrow} = n_{i\downarrow} = 1$, as shown in Fig. 4.1 The total Hamiltonian for such a model is thus

$$H = -t \sum_{\langle i,j \rangle} c^{\dagger}_i c_j + U \sum_{i} n_{i,\uparrow} n_{i,-\downarrow} \quad (4.9)$$

This Hamiltonian, known as the Hubbard Hamiltonian, is remarkably simple and yet contains all the essential effect necessary to describe interacting electrons in a solid. It contains the band via the presence of the lattice and the tight binding Hamiltonian and thus, via the filling of the band, can describe bands insulators. The contact interaction is the simplest one can put. Thus this model plays for interacting quantum system the same role that the Ising model played to describe phase transitions. Despite its apparent simplicity this model is essentially still unsolved. We will tackle the properties of this model in details in the second part of these lectures.

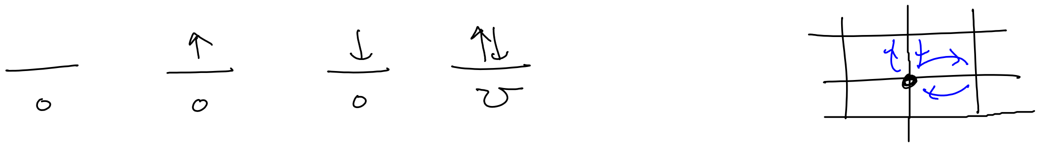


Figure 4.1: (left) A contact interaction on a lattice. The Fock space on a single state has four possible states. The interaction only acts when a site is doubly occupied. (right) A cartoon of the Hubbard model, where electrons hop on a lattice via a tight binding Hamiltonian and only feel a local interaction.

4.2 Single particle excitations: photoemission

4.3 Free electron excitations

In order to understand the differences between interacting systems and free electrons, let us recall first the salient properties of the free electron system (or any free quantum fermionic system). We mostly focuss here on the $T = 0$ properties, that fully characterize the quantum state of the system.

A free fermionic system has a ground states, that is a Fermi sea of occupied states, where the lowest kinetic energy states are fully occupied

$$|\mathrm{F}\rangle = \prod_{k,\varepsilon(k)\leq E_\mathrm{F}} c^\dagger_{k,\uparrow} c^\dagger_{k\downarrow} |\varnothing\rangle \tag{4.10}$$

There is (at $T = 0$) a sharp separation between the occupied and non-occupied states in the ground state and in particular $n(k)$ the probability that a state k is occupied has a sharp discontinuity at $k = k_\mathrm{F}$. The amplitude of this discontinuity is 1.

The excitations of the system above the ground state are excitations that consist either in creating particles in a given state k above the Fermi level

$$|\psi_{k_1,k_2,\ldots}\rangle = c^\dagger_{k_1} c^\dagger_{k_2} \ldots |\mathrm{F}\rangle \tag{4.11}$$

or destroying particles (creating holes) among the states that are in the Fermi sea

$$|\psi'_{k_1,k_2,\ldots}\rangle = c_{k'_1} c_{k'_2} \ldots |\mathrm{F}\rangle \tag{4.12}$$

The individual excitations are thus particles with all the quantum numbers of the electrons (or holes), in particular with a well defined momentum k and energy $\xi(k)$. The also carry the other quantum numbers of the electrons (or hole), i.e. a charge $\pm e$, and a spin $1/2$. The wavefunction of such excitations corresponds to Bloch waves with the momentum k.

In the interacting system one can expect the nature of such excitations to change. We would thus like to have a correlation function that tells us how much such excitations look like or not to independent particles.

4.3.1 Single particle Green's function

One way to realize this is to construct the correlation function that would correspond to the gedanken experiment described in Fig. 4.2. One injects a particle at point r_1. If $|G\rangle$ is the

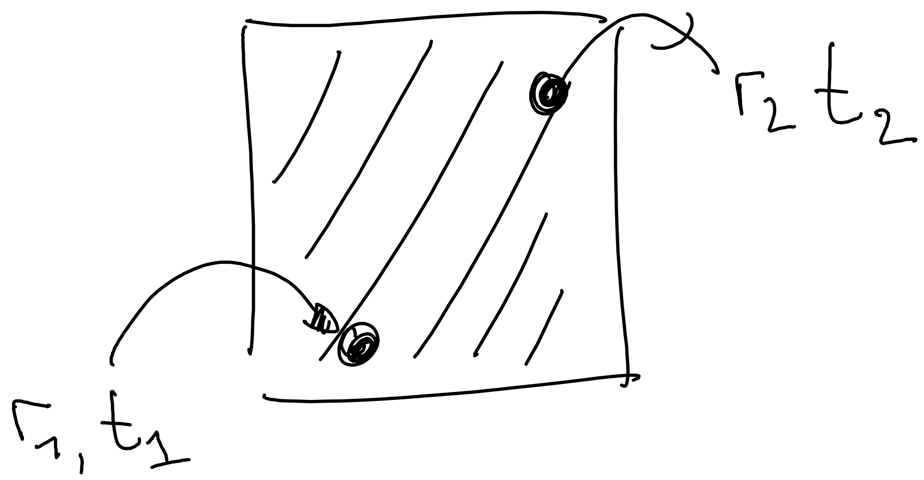


Figure 4.2: One way to understand the nature of the excitations in an interacting system is to inject a particle at point r_1 and time $t = 0$. One lets the resulting system evolves, and then one destroys a particle at time t and point r_2. The amplitude of such a process indicates how much the propagation of the excitation between the two points and times is similar to the case of non interacting electrons or not.

ground state of the system (not necessarily the interacting Fermi sea), the wavefunction of the system becomes

$$|\psi\rangle = c^\dagger_{r_1}|G\rangle \tag{4.13}$$

one then let this wavefunction with $\mathcal{N}+1$ particles evolve until the time t

$$|\psi(t)\rangle = e^{-iHt}c^\dagger_{r_1}|G\rangle \tag{4.14}$$

The question is how much this function resembles the one to which one would have simply added one free particle on the top of the ground state. For that we can define the corresponding reference function: we let the ground state of the system evolve with time t

$$|\psi'(t)\rangle = c^\dagger_{r_1}e^{-iHt}|G\rangle \tag{4.15}$$

the function $|\psi'(t)\rangle$ correspond to the addition of a free particle on the top of the ground state (at time t). The overlap between $|\psi(t)\rangle$ and $|\psi'(t)\rangle$ will thus tell how much the particle that has been injected in the system at point r_1 and time $t = 0$ has evolved freely and still resembles a free particle at point r_2 and time t. This overlap is

$$\langle\psi'(t)|\psi(t)\rangle = \langle G|e^{iHt}c_{r_2}e^{-iHt}c^\dagger_{r_1}|G\rangle \tag{4.16}$$

This correlation can also be viewed as

$$\langle\psi'(t)|\psi(t)\rangle = \langle G|c_{r_2}(t)c^\dagger_{r_1}(t=0)|G\rangle \tag{4.17}$$

using the Heisenberg representation for the operators. It consists in the process described in Fig. 4.2 where one creates a particle at point $(r_1, t = 0)$ and destroys it at point (r_2, t) to leave the system in the same state it started from.

In fact in order to have a more physical correlation function one could imagine, in the same spirit than what we did for the linear response, that one adds to the Hamiltonian a "perturbation"

$$H_{\text{pert}} = \int dr(\lambda(r,t)c_r^\dagger + \lambda^*(r,t)c_r) \tag{4.18}$$

This perturbation corresponds to "source" terms λ and λ^* that would create and destroy particles at given points. Although such a perturbation does not conserve the number of particles it is perfectly allowed in second quantization. One could for example on a more physical basis imagine a Scanning Tunnelling Microscope (STM) tip that could either inject of remove electrons from the system, but for the moment let us view it as a formal perturbation. We could measure, in response to this perturbation the average of the operators destroying one particle at point (r_2, t)

$$\langle c_{r_2}\rangle_t \tag{4.19}$$

This average exactly describes the physical information we are looking for and that we described above, namely if one injects particle in the system, how much do they still resemble or not free fermions at point r_2 and time t. Based on the general theory of linear response we would thus have

$$\langle c_{r_2}\rangle_t = \int dr_1 dt_1 G(r_2 - r_1, t_2 - t_1)\lambda(r_1, t_1) \tag{4.20}$$

where

$$G(r_2 - r_1, t_2 - t_1) = -i\,\theta(t_2 - t_1)\langle [c_{r_2,t_2}, c^\dagger_{r_1,t_1}]\rangle \tag{4.21}$$

This correlation is essentially (up to the commutator) the one we wrote above, based on physical considerations.

In fact there is an important correction that must be made to the correlation (4.21) for it to represent faithfully a linear response process. The problem with the expression (4.21), is that contrarily with the perturbations we examined in Section 2 the perturbation (4.18) is not allowed if the λ are simple numbers. Indeed in order to figure in an Hamiltonian operators must have a well defined classical limit to give the classical energy. This means that each term in the Hamiltonian should tend to a simple number with the appropriate $\hbar \to 0$ limit. In particular if $O_1(r_1)$ and $O_2(r_2)$ are two parts in the Hamiltonian that are roughly local around the points r_1 and r_2 it means that when the separation between the points r_1 and r_2 become large the two operators must commute

$$\lim_{r_1 - r_2 \to \infty} [O_1(r_1), O_2(r_2)] = 0 \tag{4.22}$$

It of course does not mean that the operators O_1 and O_2 commute in general but simply that the non zero terms in their commutator should only involve the short distance (i.e.g $r_1 \sim r_2$) behavior. This is obviously *not* the case for c_{r_1} and $c^\dagger_{r_2}$ that have no simple commutation relation regardless of the distance $r_1 - r_2$. The term (4.18) is thus not something that can figure in an Hamiltonian if the λ is a classical field. The only way to get an acceptable Hamiltonian from (4.18) is if the fields λ in fact also anticommute. In that case

$$[\lambda(r_1)c_{r_1}, \lambda^*(r_2)c^\dagger_{r_2}] = \lambda(r_1)\lambda^*(r_2)[c_{r_1}, c^\dagger_{r_2}]_+ \to 0 \tag{4.23}$$

when $r_1 - r_2 \to \infty$. Such fields exists and are called Grassmann variables. Physically it means that a fermion operator must be hidden in the field λ since one cannot simply create or destroy a fermion. The consequence of having λ that anticommute is that when one repeats the steps that led to the linear response formula we will now have an anticommutator instead of a commutator. The good correlation function to consider is thus

$$G(r_2 - r_1, t_2 - t_1) = -i\,\theta(t_2 - t_1)\langle [c_{r_2,t_2}, c^\dagger_{r_1,t_1}]_+\rangle \tag{4.24}$$

This correlation function, called the single particle Green's function, is the one we were looking for. It is the physical response, since it corresponds to the linear response to source terms. It satisfies all the requirements of causality etc. that a physical correlation function must satisfy. We will see in the following section how to relate it to well defined experimental quantities (by opposition to gedanken ones).

Instead of using the basis of position we can of course use the momentum basis. One can rewrite the Green's function as

$$G(r_2 - r_1, t_2 - t_1) = -i\,\theta(t_2 - t_1)\frac{1}{\Omega}\sum_{k_1,k_2} e^{-ik_1r_1}e^{ik_2r_2}\langle[c_{k_2,t_2}, c^\dagger_{k_1,t_1}]_+\rangle \tag{4.25}$$

The correlation $\langle[c_{k_2,t_2}, c^\dagger_{k_1,t_1}]_+\rangle$ corresponds to the creation of a particle with momentum k_1 and the destruction of a particle with momentum k_2. If the system is invariant by translation and G only depends on $r_2 - r_1$ then it is clear that the momentum is conserved and the only non zero correlation must be the one for which $k_1 = k_2$. Thus one can define the Green's function in momentum space by

$$G(k, t_2 - t_1) = -i\,\theta(t_2 - t_1)\langle[c_{k,t_2}, c^\dagger_{k,t_1}]_+\rangle \tag{4.26}$$

This correlation creates a particle in a well defined momentum state k at time t_1, let it propagate and then tries to destroy a particle in a well defined momentum state k at time t_2. It thus measures how well in the interacting system the single particle excitations still resemble Bloch waves, i.e. independent particles.

4.3.2 Properties and spectral function

Let us first compute the Green function for independent electrons. We consider the Hamiltonian

$$H = \sum_k \xi(k) c^\dagger_k c_k \tag{4.27}$$

One can easily compute

$$c_k(t) = e^{iHt}c_k e^{-iHt} \tag{4.28}$$

by looking at the action of the operator on the two states $|0_k\rangle$ and $|1_k\rangle$ with zero and one particle in the state k. These two states are the complete Hilbert space for fermions.

$$\begin{aligned} c_k(t)\,|0_k\rangle &= 0 \\ c_k(t)\,|1_k\rangle &= e^{iHt}c_k e^{-i\xi(k)t}\,|1_k\rangle = e^{-i\xi(k)t}\,|0_k\rangle \end{aligned} \tag{4.29}$$

which immediately leads to

$$c_k(t) = e^{-i\xi(k)t}c_k \tag{4.30}$$

The Green function (4.26) thus becomes

$$\begin{aligned} G^0(k,t) &= -i\,\theta(t)\langle[c_{k,t}, c^\dagger_{k,0}]_+\rangle \\ &= -i\,\theta(t)e^{-i\xi(k)t}\langle[c_k, c^\dagger_k]_+\rangle \\ &= -i\,\theta(t)e^{-i\xi(k)t} \end{aligned} \tag{4.31}$$

The Fourier transform is thus

$$\begin{aligned} G^0(k,\omega) &= \int dt e^{i(\omega+i\delta)t}G^0(k,t) \\ &= -i\int_0^{+\infty} dt e^{i(\omega+i\delta)t}e^{-i\xi(k)t} \\ &= \frac{1}{\omega - \xi(k) + i\delta} \end{aligned} \tag{4.32}$$

The imaginary part of the Green function takes a specially simple form. Let us introduce

$$A(k,\omega) = \frac{-1}{\pi} \operatorname{Im} G(r,\omega) \tag{4.33}$$

The function $A(k,\omega)$ is known as the spectral function and its physical meaning will become clear in a moment. As usual one can express the full Green function as a function of $A(k,\omega)$ using the spectral representation (2.59). For free electrons the spectral function is

$$A^0(k,\omega) = \delta(\omega - \xi(k)) \tag{4.34}$$

and is thus non zero only when the frequency is given by the dispersion relation of the particle $\xi(k)$.

In order to understand the meaning of the function $A(k,\omega)$ let us use the spectral representation as in Section 2.4. Using again a complete basis $|n\rangle$ of the eigenstates of the Hamiltonian H, with energies E_n one can rewrite the correlation function (4.26) as

$$G(k,\omega) = \frac{1}{Z}\sum_{n,m} \langle n|c_k|m\rangle \langle m|c_k^\dagger|n\rangle \frac{e^{-\beta E_n} + e^{-\beta E_m}}{\omega + E_n - E_m + i\delta} \tag{4.35}$$

The only difference compared to the formula (2.54) is the plus sign in the numerator instead of the minus sign, due to the anticommutator instead of the commutator. The spectral function is thus

$$A(k,\omega) = \frac{1}{Z}\sum_{n,m} \langle n|c_k|m\rangle \langle m|c_k^\dagger|n\rangle (e^{-\beta E_n} + e^{-\beta E_m})\delta(\omega + E_n - E_m) \tag{4.36}$$

This can be rewritten as

$$A(k,\omega) = \frac{1+e^{-\beta\omega}}{Z}\sum_{n,m} |\langle m|c_k^\dagger|n\rangle|^2 e^{-\beta E_n}\delta(\omega + E_n - E_m) \tag{4.37}$$

This describes all the possible transitions where one creates a particle with a well defined momentum k from the initial state $|n\rangle$. The only allowed transitions are the ones that conserve energy $E_m = \omega + E_n$. The initial state $|n\rangle$ is weighted with the proper factor $e^{-\beta E_n}$ if one is not at zero temperature to take into account all possible initial states. The spectral function thus measures the possibility to make a transition from a state $|n\rangle$ to a state $|m\rangle$ by creating a single particle excitation with a well definite momentum k and a well defined energy ω. The spectral function thus measures the probability to find in the system such an excitation. It gives the probability to find a single particle excitation with *both* momentum k and energy ω. Note that the spectral function $A(k,\omega)$ is indeed a positive function as can be seen from the expression (4.37).

Since we interpret the spectral function as a probability to find a single particle excitation with the frequency ω and the momentum k, one would expect the sum rule

$$\int_{-\infty}^{+\infty} d\omega A(k,\omega) = 1 \tag{4.38}$$

to be obeyed irrespective of the Hamiltonian H. From (4.36) one has

$$\begin{aligned}
\int -\infty^{+\infty} d\omega A(k,\omega) &= \frac{1}{Z}\sum_{n,m} \langle n|c_k|m\rangle \langle m|c_k^\dagger|n\rangle (e^{-\beta E_n} + e^{-\beta E_m}) \\
&= \frac{1}{Z}\sum_n e^{-\beta E_n}\langle n|c_k c_k^\dagger|n\rangle + \frac{1}{Z}\sum_m e^{-\beta E_m}\langle m|c_k^\dagger c_k|m\rangle \\
&= \frac{1}{Z}\sum_n e^{-\beta E_n}\langle n|[c_k, c_k^\dagger]_+|n\rangle \\
&= 1
\end{aligned} \tag{4.39}$$

and the sum rule is indeed satisfied.

This interpretation of the spectral function is perfectly consistent with the result (4.34) for free fermions. Indeed in that case the only possibility to find a single particle excitation is when the energy ω is equal to the energy of the particle with the fixed momentum $\xi(k)$, hence the δ function in (4.34). In the presence of interactions one would expect that even when the momentum k is well fixed, the particle can exchange energy with the other particles due to the interaction. One would thus expect a spread of the δ peak for free particles, over a width in energy of the order of the typical energy that can be exchanged due to the interactions. This is schematically represented in Fig. 1.12.

4.3.3 Connection with photoemission

We saw in Section 1.4 the photoemission technique where one can use photons to kick out electrons out of a system. One measures the energy and momenta of the outgoing electron. Let us now show that the photoemission technique is (in an idealized world) a direct measure of the spectral function $A(k, \omega)$.

In the experiment one starts with the system in the state $|m\rangle$. If one is at $T = 0$ $|m\rangle$ is the ground state, at finite temperature $|m\rangle$ is as usual distributed with the probability $e^{-\beta E_m}$. The photon removes one particle with momentum k (directly measured) from the system, and thus induces a transition to the state $|n\rangle$ (which contains $\mathcal{N} - 1$ electrons). Using the Fermi Golden rule the probability of transition from state $|m\rangle$ to $|n\rangle$ is thus

$$P_{m \to n} = \sum_n |\langle n|c_k|m\rangle|^2 \delta(E_n - E_m - E) \tag{4.40}$$

where E is the energy of the outgoing particle. If one considers all possible transitions from all possible initial states, the total probability of transition is thus

$$P = \frac{1}{Z} sum_{n,m} |\langle n|c_k|m\rangle|^2 e^{-\beta E_m} \delta(E_n - E_m - E) \tag{4.41}$$

This probability of transition is the one measured in photoemission since it will directly give the number of outgoing electrons with a momentum k and the energy E for a fixed influx of photons.

This expression is to be compared with (4.36). Using the δ function (4.36) can be rewritten as

$$\begin{aligned} A(k, \omega) &= \frac{e^{\beta\omega} + 1}{Z} \sum_{n,m} |\langle n|c_k\,|m\rangle\,|^2 e^{-\beta E_m} \delta(\omega + E_n - E_m) \\ &= \frac{1}{f_{\mathrm{F}}(\omega)} \frac{1}{Z} \sum_{n,m} |\langle n|c_k\,|m\rangle\,|^2 e^{-\beta E_m} \delta(\omega + E_n - E_m) \end{aligned} \tag{4.42}$$

Thus, up to a factor $f_{\mathrm{F}}(\omega)$, that can be easily taken into account since one works at fixed energy ω, the photoemission experiment directly measures the spectral function $A(k, \omega)$. It will thus be an invaluable tool to analyze the nature of the single particle excitations.

Note in addition that one has

$$f_{\mathrm{F}}(\omega) A(k, \omega) = \frac{1}{Z} \sum_{n,m} |\langle n|c_k\,|m\rangle\,|^2 e^{-\beta E_m} \delta(\omega + E_n - E_m) \tag{4.43}$$

and integrating over ω gives

$$\begin{aligned}
\int d\omega f_{\mathrm{F}}(\omega) A(k,\omega) &= \frac{1}{Z}\sum_{n,m} |\langle n|c_k\,|m\rangle\,|^2 e^{-\beta E_m} \\
\int d\omega f_{\mathrm{F}}(\omega) A(k,\omega) &= \frac{1}{Z}\sum_{m} \langle m|c_k^\dagger c_k|m\rangle e^{-\beta E_m} \qquad (4.44) \\
\int d\omega f_{\mathrm{F}}(\omega) A(k,\omega) &= \langle c_k^\dagger c_k\rangle
\end{aligned}$$

which directly relates the spectral function to the occupation factor $n(k)$. Of course for free electrons $A(k,\omega) = \delta(\omega - \xi(k))$ and one recovers

$$n(k) = f_{\mathrm{F}}(\xi(k)) \tag{4.45}$$

4.4 Phenomenology of the spectral function

We now have to analyze the single particle Green function (as well as other physical properties) for the interacting problem. This is clearly a a very difficult question. The time evolution of the operators such as in (4.30) becomes highly non trivial since there is no simple commutator of H and the operators c_k anymore. In the same vein since in general the interaction Hamiltonian is non quadratic, there is no simple transformation, analogous to a Bogoliubov transformation, that can bring is to a diagonal form.

We thus have to rely to different approximate techniques to tackle this problem. Indeed even if one takes an Hamiltonian as simple as the Hubbard model there is despite about 50 years of research no exact solution except in dimension $d = 1$ and $d = \infty$. Contrarily to other models such as the Ising model, there is still no agreement on what the physics of the model is for intermediate dimensions, specially for $d = 2$.

One method that could be used to tackle the interacting problem is to perform perturbation in the interaction term H_{int}. Indeed we know how to compute the free Green function and other observable so this is a quite logical approach. Even if in real solids, as we saw the interaction is not effectively small, one can still expect the perturbation theory to give us information on the qualitative effects of the interactions. However performing such a perturbation is in itself a formidable task. Indeed the perturbation must be made at two places: both in the time evolution e^{iHt} and in the Bolzmann factor $e^{-\beta H}$. There are two difficulties. One is purely technical and liked to the fact that H_{kin} and H_{int} are operators and that these operators do not commute (in general). So performing the expansion of the exponentials is clearly more complicated than for simple numbers. The second difficulty is much more profound in nature. In the time evolution, we are in principle interested in computing the value of the operators at all times, thus in principle one can have $t \to \infty$. It is thus unclear whether one can perform an expansion of the exponentials

$$e^{-i(H_0+H_1)t} \simeq e^{-iH_0 t} e^{-iH_1 t}\overline{?}\simeq e^{-iH_0 t}[1 - iH_1 t + \cdots] \tag{4.46}$$

or how to do it. Fortunately these problems can be solved and one can construct a perturbation method for quantum problems at finite temperature. This is the method known as Feynmann diagrams generalized to finite temperature. We will not pursue more in this direction and refer the reader to [Mah81] for more details on this technique.

Instead of following a systematic route we will play with the single particle Green function and determine which parameters can control its shape. It will thus give us, in a phenomenological way, the properties of interacting systems.

4.4.1 Self energy: lifetime and effective mass

Let us first consider the single particle Green function for an interacting problem. One can always write it as

$$G(k,\omega) = \frac{1}{\omega - \xi(k) - \Sigma(k\omega) + i\delta} \tag{4.47}$$

where $\Sigma(k,\omega)$ called the self energy is a certain function of momenta and frequency. The relation (4.47) in fact defines the function Σ. Given the form (4.47) and the expressions (4.32) one sees that Σ must go to zero in the absence of interactions. To make connection with the perturbation method $\Sigma(k,\omega)$ is thus an object that is expected to have a well behaved perturbation expansion in powers of the interactions. However we will not attempt here to compute the self energy Σ but simply to examine how it controls the spectral function. We will absorb the small imaginary part $i\delta$ in Σ for simplicity, since one can expect in general Σ to have a finite imaginary part as well. The spectral function is

$$A(k,\omega) = -\frac{1}{\pi}\frac{\operatorname{Im}\Sigma(k,\omega)}{(\omega - \xi(k) - \operatorname{Re}\Sigma(k,\omega))^2 + (\operatorname{Im}\Sigma(k,\omega))^2} \tag{4.48}$$

and we thus see that $\operatorname{Im}\Sigma$ and $\operatorname{Re}\Sigma$ have very different actions on the spectral function. We also see that quite generally (4.48) imposes that $\operatorname{Im}\Sigma(k,\omega) < 0$ to get a positive spectral function.

Imaginary part: lifetime

Let us fist assume that the real part of the spectral function $\operatorname{Re}\Sigma(k,\omega)$ is zero to investigate the consequences of the existence of an imaginary part. Note that doing this kind of crude approximation might violate some relations between the real and imaginary part of the self energy, which in general should be related by the Kramers-Kronig relations. There is thus a chance that we will get some absurd, or incorrect results, due to that, but we will face this problem if or when it occurs. Moreover let us assume that the imaginary part is a constant.

With these approximations the spectral function would simply be

$$A(k,\omega) = -\frac{1}{\pi}\frac{\operatorname{Im}\Sigma}{(\omega - \xi(k))^2 + (\operatorname{Im}\Sigma)^2} \tag{4.49}$$

As shown in Fig. 4.3 a finite imaginary part in the self energy thus gives a Lorentzian shape for the peaks. The peaks are still centered at $\omega = \xi(k)$ but have now a finite width and height instead of being δ functions. The width is given by the imaginary part of the self energy and the height in $1/\operatorname{Im}\Sigma$. As usual with a Lorentzian the total spectral weight in the peak is a constant, hence the height inversely proportional to the width. As we see the peaks are sharper and sharper when the $\operatorname{Im}\Sigma$ becomes smaller and smaller. In the limit when the $\operatorname{Im}\Sigma \to 0$ one recovers, as one should a δ function peak, which is indeed the limit of a zero width Lorentzian.

We thus see that the imaginary part of the self energy controls the spread of the energies of the particles. One can see the spectral function as a particle which has an average energy $\omega = \xi(k)$, related to its momentum, but with a certain spread $\operatorname{Im}\Sigma$ in energy. To understand the physics of this spread let us consider the Green function of a free particle in real time (4.31). This function when Fourier transform would give the perfect δ function. However if one modifies it by

$$G(k,t) = -i\,\theta(t)e^{-i\xi(k)t}e^{-t/\tau} \tag{4.50}$$

the Fourier transform becomes

$$G(k,\omega) = \frac{1}{\omega - \xi(k) + i/\tau} \tag{4.51}$$

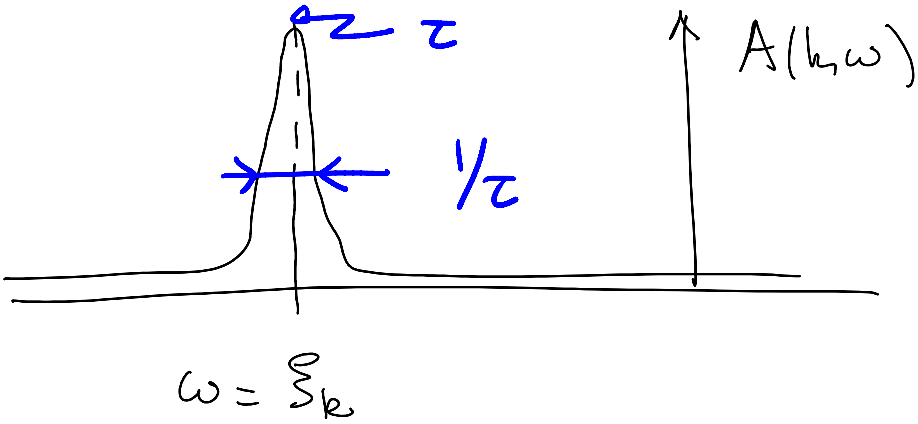


Figure 4.3: A finite imaginary part in the self energy gives a Lorentzian peak. The width of the Lorentzian is $\operatorname{Im}\Sigma$ and the height $1/\operatorname{Im}\Sigma$. The smaller the imaginary part, the sharper the peaks. The imaginary part of the self energy has thus the meaning of an inverse lifetime for a particle with momentum k.

and the spectral function is

$$A(k,\omega) = \frac{1/\tau}{(\omega - \xi(k))^2 + (1/\tau)^2} \tag{4.52}$$

which is exactly the one we are considering with the identification

$$\frac{1}{\tau} = \operatorname{Im}\Sigma \tag{4.53}$$

We thus see from (4.50) that a Lorentzian spectral function corresponds to a particle with a well defined energy $\xi(k)$ which defines the center of the peak, but also with a finite *lifetime* τ. Of course this does not mean that the electron physically disappears, but simply that it does not exist as an excitation with the given quantum number k. This is indeed an expected effect of the interaction since the particle will exchange momenta with the others particles and thus is able to change its quantum state.

To go back to the more general form of the self energy, which depends on k and ω we see that we can keep this interpretation in terms of a lifetime, if the peak is narrow enough. Indeed in that case what will matter is the self energy at the position of the peak $\operatorname{Im}\Sigma(k, \omega = \xi(k))$ if one assumes that the self energy varies slowly enough with ω compared to $\omega - \xi(k)$.

Real part: effective mass and quasiparticle weight

Let us now turn to the real part. Now that we understand that the imaginary part provides a lifetime for the particle let us turn the imaginary part to zero to focuss on the effects of the real part of the self energy. Of course in doing so we strongly violated the Kramers-Kronig relation since the real part should have been zero as well. But this simplification is only to replace the

Lorentzian peaks by sharp δ functions for simplicity so we do not expect it to drastically affect the physics driven by the real part.

If we only consider the real part the spectral function becomes

$$A(k,\omega) = \delta(\omega - \xi(k) - \mathrm{Re}\,\Sigma(k,\omega)) \tag{4.54}$$

The role of the real part of the self energy is thus to modify the position of the peak. Instead of having the peak at $\omega = \xi(k)$, one has now a new dispersion relation $E(k)$ which is defined by

$$E(k) - \xi(k) - \mathrm{Re}\,\Sigma(k, \omega = E(k)) = 0 \tag{4.55}$$

The relation (4.55) defines the new dispersion relation. The interactions, via the real part of the self-energy are thus leading to a modification of the energy of single particle excitations. Although we can in principle compute the whole dispersion relation $E(k)$, in practice we do not need it to characterize the system. Indeed we are only interested in low energy excitations close to the Fermi level. Close to the Fermi level the energy, with a suitable subtraction of the chemical potential is zero. One can thus expand it in powers of k. For free electrons with $\xi(k) = \frac{k^2}{2m} - \frac{k_\mathrm{F}^2}{2m}$ the corresponding expansion would give

$$\xi(k) = \frac{k_\mathrm{F}}{m}(k - k_F) \tag{4.56}$$

A similar expansion for the new dispersion $E(k)$ gives

$$E(k) = 0 + \frac{k_\mathrm{F}}{m^*}(k - k_\mathrm{F}) \tag{4.57}$$

which defines the coefficient m^*. Comparing with (4.56) we see that m^* has the meaning of a mass. This is an effective mass which traduces the fact that the dispersion relation has been changed by the interactions. We thus see that close to the Fermi level we only need to compute the effective mass m^* to fully determine (at least for a spherical Fermi surface) the effects of the interactions on the energy of single particle excitations. To relate the effective mass to the self energy one computes from (4.55)

$$\frac{dE(k)}{dk} = \frac{d\xi(k)}{dk} + \left.\frac{\partial\,\mathrm{Re}\,\Sigma(k,\omega)}{\partial k}\right|_{\omega=E(k)} + \left.\frac{\partial\,\mathrm{Re}\,\Sigma(k,\omega)}{\partial \omega}\right|_{\omega=E(k)} \frac{dE(k)}{dk} \tag{4.58}$$

which can be solved to give (using the fact that one the Fermi

$$\frac{k_\mathrm{F}}{m^*} = \frac{\frac{k_\mathrm{F}}{m} + \left.\frac{\partial\,\mathrm{Re}\,\Sigma(k,\omega)}{\partial k}\right|_{\omega=E(k)}}{1 - \left.\frac{\partial\,\mathrm{Re}\,\Sigma(k,\omega)}{\partial \omega}\right|_{\omega=E(k)}} \tag{4.59}$$

or in a more compact form

$$\frac{m}{m^*} = \frac{1 + \frac{m}{k_\mathrm{F}} \left.\frac{\partial\,\mathrm{Re}\,\Sigma(k,\omega)}{\partial k}\right|_{\omega=E(k)}}{1 - \left.\frac{\partial\,\mathrm{Re}\,\Sigma(k,\omega)}{\partial \omega}\right|_{\omega=E(k)}} \tag{4.60}$$

To determine the effective mass these relations should be computed on the Fermi surface $E(k_\mathrm{F}) = 0$. The equation (4.60) indicates how the self energy changes the effective mass of the particles. We thus see that although one can keep single particle excitations they will have in general, due to interactions, a different mass than the one of independent electrons. This renormalization of the mass by interaction is well consistent with the experimental findings of

Section 1.4 where we saw that in the specific heat one had something that was resembling the behavior of free electrons but with a different mass m^*.

However the interactions have another effects. Indeed if we try to write the relation (4.54) in the canonical form $\delta(\omega - E(k))$ that we would naively expect for a free particle with the dispersion $E(k)$ we see that we cannot do it. Instead using (A.7) we obtain

$$A(k,\omega) = Z_k \delta(\omega - E(k)) \tag{4.61}$$

with

$$\begin{aligned} Z_k &= \left[\frac{\partial}{\partial\omega}(\omega - \xi(k) - \operatorname{Re}\Sigma(k,\omega))\Big|_{\omega=E(k)} \right]^{-1} \\ &= \frac{1}{1 - \frac{\partial \operatorname{Re}\Sigma(k,\omega)}{\partial\omega}\Big|_{\omega=E(k)}} \end{aligned} \tag{4.62}$$

Because of the frequency dependence of the real part of the self energy, we see that the total spectral weight in the peak is not one any more but the total weight is now Z_k, which is in general a number smaller than one. It is thus as if not the whole electron (or rather the total spectral weight of an electron) was converted into something that looks like a free particle with a new dispersion relation, but only a faction Z_k of it. With our crude approximation the rest of the spectral function has totally vanished and the sum rule (4.38) is violated. This is the consequence of our crude approximation for the self energy that violates the Kramers-Kronig relation. However the effect that we found is quite real, and what becomes of the remaining spectral weight will be described in the next section.

To conclude we see that the real part of the self energy controls the dispersion relation and the total weight of excitations which in the spectral function would produce peaks exactly like free particles. The frequency and momentum dependence of the real part of the self energy lead to the two independent quantities m^* the effective mass of the excitations and Z_k the weight. In the particular case when the momentum dependence of the self energy is small on can see from (4.62) and (4.60)

$$\frac{m}{m^*} = Z_{k_\mathrm{F}} \tag{4.63}$$

4.5 Fermi liquid theory

4.5.1 Landau quasiparticles

From the previous analysis of the spectral function and its connection with the self energy we have a schematic idea of the excitations as summarized in Fig. 4.4. Quite generally we can thus distinguish two parts in the spectral function. There is a continuous background, without any specific feature for which the probability to find a particle with energy ω is practically independent of its momentum k. This part of the spectrum cannot be easily identified with excitations resembling free or quasi-free particles. One the other hand, in addition to this part, which carries a total spectral weight $1 - Z_k$, another part of the excitations gives a spectral weight with a lorentzian peak, well centered around a certain energy $E(k)$. This part of the spectrum can thus be identified with a “particle”, called Landau quasiparticle, with a well defined relation between its momentum k and energy $\omega = E(k)$, but which has a only a finite lifetime, which determined the width and height of the peak. The dispersion relation and the total weight of the quasiparticle peak are controlled by the real part of the self energy, while the lifetime is inversely proportional to the imaginary part. Depending on the self energy, and

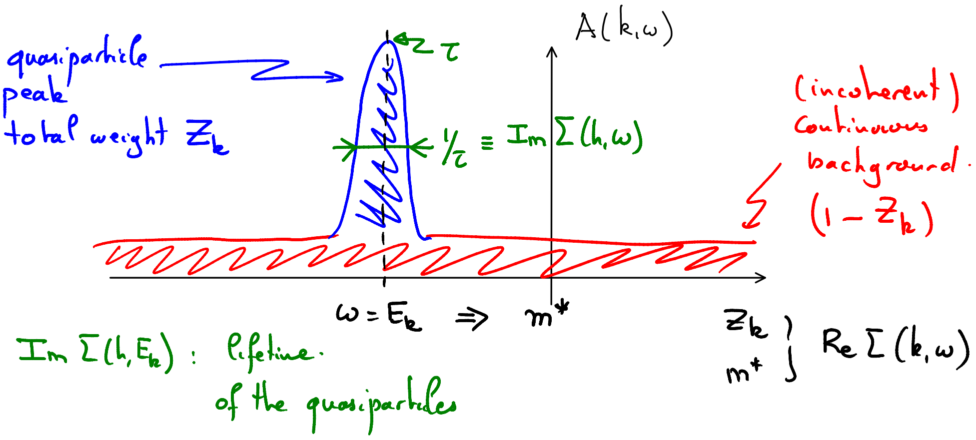


Figure 4.4: A cartoon of the spectral function for interacting particle. One can recognize several features. There is a continuous background of excitations of total weight $1 - Z_k$. This part of the spectrum cannot be identified with excitations that resemble quasi-free particles. In addition to this continuous background there can be a quasiparticle peak. The total weight of the peak is Z_k determined by the real part of the self energy. The center of the peak is at energy $E(k)$ which is renormalized by the interactions compared to the independent electron dispersion $\xi(k)$. This change of dispersion defines an effective mass m^* determined also by the real part of the self energy. The quasiparticle peak has a lorentzian lineshape that traduces the finite lifetime of the quasiparticles. The lifetime is inversely proportional to the imaginary part of the self energy.

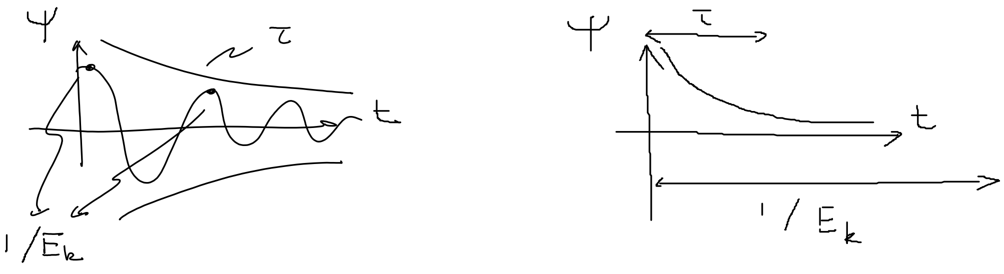


Figure 4.5: For particles with an energy $E(k)$ and a finite lifetime τ, the energy controls the oscillations in time of the wavefunction. In order to properly identify an excitation as a particle it is mandatory that the wavefunction can oscillate several time before being damped by the lifetime, otherwise it is impossible to precisely define the frequency of the oscillations. This is illustrated on the left part of the figure. On the contrary if the damping is too fast, one cannot define an average energy and thus identify the excitation with a particle.

thus the interactions, we thus see that we can still have objects that we could identify with "free" particles, solving our problem of why the free electron picture works qualitatively so well with just a renormalization of the parameters such as the mass into an effective mass.

However it is not clear that in the presence of interactions one can have sharp quasiparticles. In fact one would naively expect exactly the opposite. Indeed we would like to identify the peak in the spectral function with the existence of a quasiparticle. The energy of this excitation is $E(k)$ which of course tends towards zero at the Fermi level, while the imaginary part of the self energy is the inverse lifetime $1/\tau$. Since $E(k)$ gives the oscillations in time of the wavefunction of the particle $e^{-iE(k)t}$, in order to be able to identify properly a particle it is mandatory, as shown in Fig. 4.5 that there are many oscillations by the time the lifetime has damped the wavefunction. This imposes

$$E(k) \gg \tau \tag{4.64}$$

Since $1/\tau$ is the imaginary part of the self energy and controlled by energy scales of the order of the interactions, one would expect the life time to be roughly constant close to the Fermi level. On the other hand one has always $E(k) \to 0$ when $k \to k_{\rm F}$, and thus the relation (4.64) to be violated when one gets close to the Fermi level. This would mean that for weak interactions one has perhaps excitations that resemble particles far from the Fermi level, but that this becomes worse and worse as one looks at low energy properties, with finally all the excitations close to Fermi level being quite different from particles.

Quite remarkably, as was first shown by Landau, this "intuitive" picture is totally incorrect and the lifetime has a quite different behavior when one approaches the Fermi level.

4.5.2 Lifetime

In order to estimate the lifetime let us look at what excitations can lead to the scattering of a particle from the state k to another state. Let us start from the non interacting ground state in the spirit of a perturbative calculation in the interactions. As shown in Fig. 4.6 a particle coming in the system with an energy ω and a momentum k can excite a particle-hole excitation, taking a particle below the Fermi surface with an energy ω_1 and putting it above

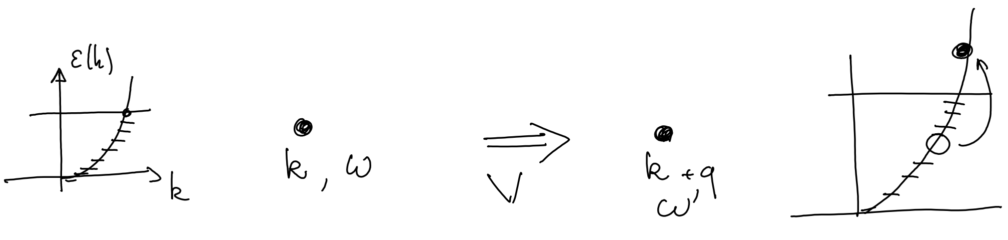


Figure 4.6: Cartoon of the lifetime of a particle with energy ω. It can interact with the ground state of the system, which has all single particle states filled below the Fermi energy E_{F}. The excitations are thus particle hole excitations where a particle is promoted from below the Fermi level to above the Fermi level. Due to the presence of the sharp Fermi level, the phase space available for making such particle hole excitations is severely restricted.

the Fermi level with an energy ω_2. The process is possible if the initial state is occupied and the final state is empty. One can estimate the probability of transition using the Fermi golden rule. The probability of the transition gives directly the inverse lifetime of the particle, and thus the imaginary part of the self energy. We will not care here about the matrix elements of the transition, assuming that all possible transitions will effectively happen with some matrix element. The probability of transition is thus the sum over all possible initial states and final states that respect the constraints (energy conservation and initial state occupied, final state empty). Since the external particle has an energy ω it can give at most ω in the transition. Thus $\omega_2 - \omega_1 \leq \omega$. This implies also directly that the initial state cannot go deeper below the Fermi level than ω otherwise the final state would also be below the Fermi level and the transition would be forbidden. The probability of transition is thus

$$P \propto \int_{-\omega}^{0} d\omega_1 \int_{0}^{\omega+\omega_1} d\omega_2 = \frac{1}{2}\omega^2 \tag{4.65}$$

One has thus the remarkable result that because of the discontinuity due to the Fermi surface and the Pauli principle that only allows the transitions from below to above the Fermi surface, the inverse lifetime behaves as ω^2. This has drastic consequences since it means that contrarily to the naive expectations, when one considers a quasiparticle at the energy ω, the lifetime grows much *faster* than the period $\mathcal{T} \sim 1/\omega$ characterizing the oscillations of the wavefunction. In fact

$$\frac{\tau}{\mathcal{T}} = \frac{1}{\omega} \to \infty \tag{4.66}$$

when one approaches the Fermi level. In other words the Landau quasiparticles become *better and better defined* as one gets closer to the Fermi level. This is a remarkable result since it confirms that we can view the system as composed of single particle excitations that resemble the original electrons, but with renormalized parameters (effective mass m^* and quasiparticle weight Z_k). Other quantum numbers are the same than the ones of an electron (charge, spin). Note that this does *not* mean that close to the Fermi level the interactions are disappearing from the system. They are present and can be extremely strong, and can affect both the effective mass and quasiparticle weight very strongly. It is only the scattering of the quasiparticles that is going to zero when one is going close to the Fermi level. This is thus a very unusual situation quite different from what would happen in a classical gas. In such a case diluting the gas would thus reduce both the interaction between the particles and also their scattering in essentially the same proportion, as schematically represented in Fig. 4.7. On the contrary in a Fermi liquid

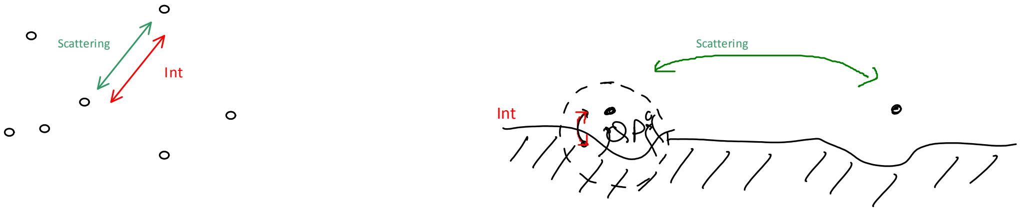


Figure 4.7: (left) A classical gas: in that case both the interactions among the particles and their scattering are changed similarly, for example by diluting the gas. (right) In a Fermi liquid, the few excitations above the ground state can interact strongly with all the other electrons present in the ground state. The effect of such interactions is strong and lead to a strong change of the parameters compared to free electrons. The quasiparticles have thus characteristics depending strongly on the interactions. However the scattering of the quasiparticles is blocked by the Pauli principle leaving a very small phase space of scattering. The lifetime of the quasiparticles is thus extremely large. This is the essence of the Fermi liquid theory.

there are many $\mathcal{N} \to \infty$ electrons in the ground state, which are in principle strongly affected by the interactions. Note that computing the ground state would be a very complicated task. However there are very few excitations above this ground state at low energy. These excitations ca interact strongly with the other electrons in the soup of the ground state, leading to a very strong change of the characteristics compared to free electron excitations. This can lead to very large effective masses or small quasiparticle weight. On the other hand the lifetime of the quasiparticles is controlled by a totally different mechanism since it is blocked by the Pauli principle. Thus even if the interaction is strong the *phase space* available for such a scattering is going to zero close to the Fermi level, making the quasiparticle in practice infinitely long lived particles, and allowing to use them to describe the system. This image also gives us a description of what a quasiparticle is: this is an electron that is surrounded by a cloud of particle-hole excitations, or in other words density fluctuations since $c^\dagger_{k+q}c_k$ is typically the type of terms entering the density operator. Such density fluctuations are of course neutral and do not change the spin. This composite object electron+density fluctuation cloud, thus represent a tight object (just like an electron does dress with a cloud of photons in quantum electrodynamics), that is the Landau quasiparticle. Since the electron when moving must carry with it its polarization cloud, one can guess that its effective mass will indeed be affected.

The Fermi liquid theory is a direct explanation of the fact that "free" electrons theory works very well *qualitatively* (such as the specific heat linear in temperature) even when the change of parameters can be huge. In addition to the parameters already shown in Fig. 1.10 we show in Fig. 4.8 the case of systems where the renormalization of the mass is about $m^* \sim 10^3 m$ indicating very strong interactions effects. Nevertheless we see that the specific heat varies linearly with temperature just like for free electrons. The prediction for the quasiparticle peaks fits very well with the photoemission data of Fig. 1.13, in which one clearly sees the peaks becoming sharper as one approaches the Fermi level. There is another direct consequence of the prediction for the lifetime. At finite temperature one can expect the lifetime to vary as $\tau \sim 1/T^2$ since T is the relevant energy scale when $T \gg \omega$. If we put such a lifetime in the Drude formula for the conductivity we get

$$\sigma(T) = \frac{ne^2\tau}{m} \propto \frac{1}{T^2} \tag{4.67}$$

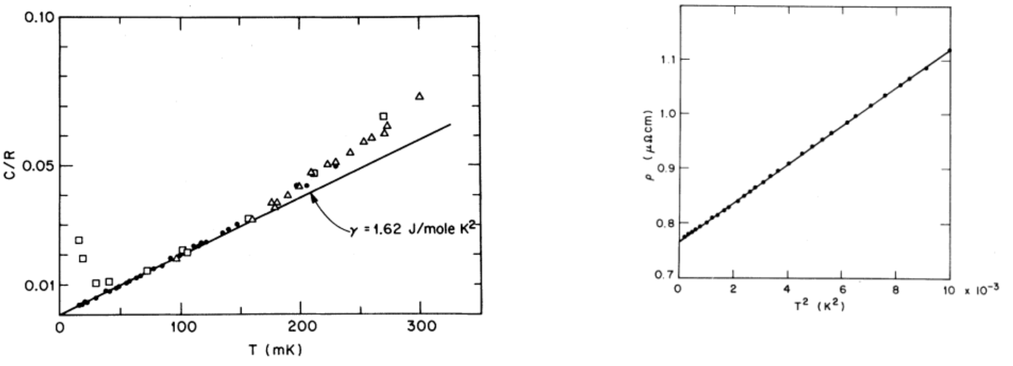


Figure 4.8: Physical properties of the compound $\mathrm{CeAl_3}$. (left) The specific heat is linear in temperature T, but the slope gives an effective mass of about 10^3m showing extremely strong interaction effects, showing that the Fermi liquid theory applies even when the interaction effects are strong. (right) The resistivity varies as T^2 in very good agreement with the Fermi liquid theory. After [AGO75].

This result can be confirmed by a full calculation. This shows that the electron-electron interactions give an intrinsic contribution to the resistivity that varies as $\rho(T) \sim T^2$, and which also can be taken as a characteristic of Fermi liquid behavior. This is however difficult to test since this temperature dependence can easily be masked by other scattering phenomena (impurities, scattering by the phonons etc.) that must be added to the electron-electron scattering and that have quite different temperature dependence. Nevertheless there are some materials where the T^2 law can be well observed as shown in Fig. 4.8. Another interesting consequence can be deduced by looking at the occupation factor $n(k)$. Using the representation (4.44) we see that if we represent the spectral function as

$$A(k, \omega) = Z_k \delta(\omega - E(k)) + A_{\text{inc}}(k, \omega) \tag{4.68}$$

where the incoherent part is a smooth flattish function without any salient feature, then the $n(k)$ becomes

$$n(k) = Z_k f_{\text{F}}(E(k)) + Cste \tag{4.69}$$

Thus even in the presence of interaction there is still a discontinuity at the Fermi level, that is only rounded by the temperature. Contrarily to the case of free electron the amplitude of the singularity at $T = 0$ is not 1 anymore but is now $Z_{k_{\text{F}}} < 1$. The existence of this discontinuity *if* quasiparticle exists tells us directly that the Fermi liquid theory is internally consistent since the very existence of the quasi particles (namely the large lifetime) was heavily resting on the existence of such a discontinuity at the Fermi level. One can thus in a way consider that the existence of a sharp discontinuity at the Fermi level is a good order parameters to characterize the existence of a Fermi liquid.

One important question is when the Fermi liquid theory does apply. This is of course a very delicate issue. One can see both from the arguments given above, and from direct perturbative calculations that when the interactions are weak the Fermi liquid theory will in general be valid. There are some notable exceptions that we will examine in the second part of the notes, and for which the phase space argument given above fails. However the main interest of the Fermi liquid theory is that it does not rest on the fact that the interactions are small and, as we have seen through examples, works also remarkably well for the case of strong interactions, even when all perturbation theory fails to be controlled. This is specially important for realistic systems since, as we showed, the interaction is routinely of the same order than the kinetic energy even in very good metals. The Fermi liquid theory has thus been the cornerstone of our description of most condensed matter systems in the last 50 years or so. Indeed it tells us that we can "forget" (or easily treat) the main perturbation, namely the interaction among electrons, by simply writing what is essentially a free electron Hamiltonian with some parameters changed. It is not even important to compute microscopically these parameters, but one can simply extract them from one experiment and then use them consistently in the others. This allows to go much further and treat effects caused by much smaller perturbations that one would otherwise have been totally unable to take into account. One of the most spectacular examples is the possibility to now look at the very tiny (compared to the electron-electron interactions) electron-phonon coupling, and to obtain from that the solution to the phenomenon of superconductivity.

Of course not all materials follow the Fermi liquid theory. Those who do not are commonly called "non Fermi liquids" a term that hides our poor knowledge of their properties, since of course in their case the question of the effects of the electron-electron interactions becomes again a formidable problem. Most of the actual research in now devoted to such non fermi liquid systems. There are fortunately some situation where one can understand the physics and we will examine such cases in the second part of these lectures.

CHAPTER 5

Collective modes

As we saw in the previous section the Fermi liquid theory allows for a very successful understanding of the single particle excitations in a solid. For free electrons, single particle excitations are the only excitations that can exist as we discussed in Section 4.3. However we can immediately see that this cannot be the only excitations that exists when interactions are present. If this were the case, we would get the same density of states (even if renormalized by the interactions) that would enter in the specific heat, charge compressibility, spin susceptibility (see Section 1.1.3). As we saw experimentally (see Fig. 1.10) his is not the case, and although the compressibility and susceptibility are constants at low temperatures, just as for free electrons, their value are different, and also different from the one one would get from the single particles excitations.

This strongly suggests that in addition to the single particle excitations that are the quasiparticles, there exits additional modes, absent for the non interacting system. This is the question that we will investigate in this chapter

5.1 Mean field approximation

As discussed in Section 2.2, we would like to compute the response of the full interacting Hamiltonian to an external perturbation, for example coupling to the charge or spin density. One can of course use the general linear response formula (2.29) to obtain the charge and spin susceptibilities.

Let us take for example a perturbation coupling to the total charge density.

$$\int d\boldsymbol{r}\lambda(\boldsymbol{r},t)[\rho(\boldsymbol{r})-\rho_0] \tag{5.1}$$

where $\rho = \rho_\uparrow + \rho_\downarrow$ is the total charge density, and ρ_0 the average density. Let us denote the deviation from the average value

$$\overline{\rho}(\boldsymbol{r}) = \rho(\boldsymbol{r}) - \rho_0 \tag{5.2}$$

In the absence of the perturbation λ one has $\langle\overline{\rho}(\boldsymbol{r})\rangle = 0$. One can relate the response in the density of electrons to this perturbation by the standard linear response formula

$$\langle\overline{\rho}(\boldsymbol{r})\rangle_t = \int d\boldsymbol{r}_1 dt_1 \chi_{\rho\rho}(\boldsymbol{r}-\boldsymbol{r}_1, t-t_1)\lambda(\boldsymbol{r}_1,t_1) \tag{5.3}$$

where $\chi_{\rho\rho}$ is the charge-charge correlation function

$$\chi_{\rho\rho}(\boldsymbol{r},t) = -i\,\theta(t)\langle[\overline{\rho}(\boldsymbol{r},t),\overline{\rho}(0,0)]\rangle_H \tag{5.4}$$

This expression is exact, but of little use since, when H is a fully interacting Hamiltonian we do not really know how to compute the susceptibility (5.4). As we saw in exercise we however know

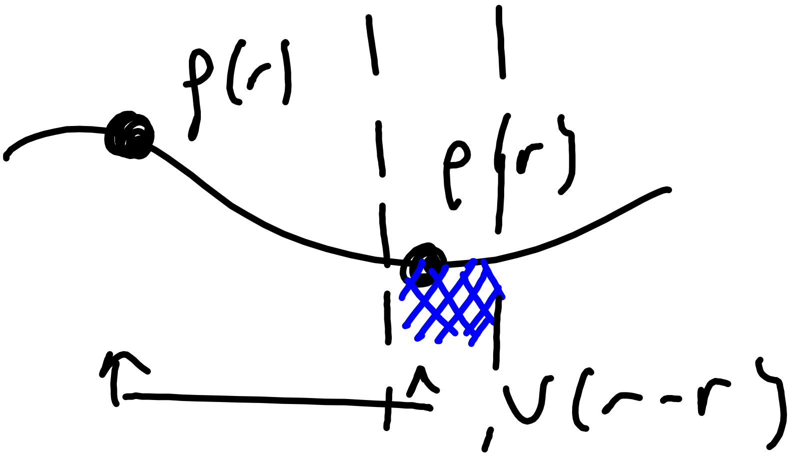


Figure 5.1: In the presence of an external perturbation the operator representing the density (or any other observable) will acquire a certain average value. The interaction term gives thus rise to an effective potential corresponding to the average action at point $\boldsymbol{r}_1$ of the other sites $\boldsymbol{r}_2$ mediated by the interaction. The mean-field approximation consists in keeping only this part an neglecting the fluctuations around this average value.

how to compute $\chi = \chi_0$ when the Hamiltonian H is just the kinetic energy of non-interacting electrons. We will thus try to find an approximate expression of (5.3) that we can compute using this fact.

5.1.1 Method

The idea is to find an approximate expression for the interaction part of the Hamiltonian. Up to a chemical potential term the interaction can be written

$$H_{\text{int}} = \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2)\overline{\rho}(\boldsymbol{r}_1)\overline{\rho}(\boldsymbol{r}_2) \tag{5.5}$$

$\overline{\rho}$ is an operator so the above expression is a full interaction term. However, in the absence of external perturbation $\langle\overline{\rho}\rangle = 0$. One can thus expect that although $\overline{\rho}$ fluctuates, its fluctuations will take place around an average values that is non zero only in the presence of the external perturbation. The interaction will thus contain, as shown in Fig. 5.1, a term coming from this average value. One can indeed write

$$\overline{\rho}(\boldsymbol{r}) = \langle\overline{\rho}(\boldsymbol{r})\rangle_t + \delta\overline{\rho}(\boldsymbol{r}) \tag{5.6}$$

where

$$\delta\overline{\rho}(\boldsymbol{r}) = \overline{\rho}(\boldsymbol{r}) - \langle\overline{\rho}(\boldsymbol{r})\rangle_t \tag{5.7}$$

Note that although the operator itself does not depend on time, the average value in presence of a time dependent external perturbation can in principle depend on time. We can thus rewrite the interaction term

$$H_{\text{int}} = \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2)[\langle\overline{\rho}(\boldsymbol{r}_1)\rangle_t + \delta\overline{\rho}(\boldsymbol{r}_1)][\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t + \delta\overline{\rho}(\boldsymbol{r}_2)] \tag{5.8}$$

The mean field approximation consists in assuming that the fluctuations around the average value are small, or fast, or uncorrelated enough so that one can neglect the second order term in the fluctuations. Keeping the first order term is necessary since otherwise the interaction Hamiltonian would just become a constant in the energy and would not affect the average of the operators. In the mean field approximation the interaction Hamiltonian thus becomes:

$$\begin{aligned}
H_{\text{int}} &\simeq \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2)[\langle\overline{\rho}(\boldsymbol{r}_1)\rangle_t\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t + \delta\overline{\rho}(\boldsymbol{r}_1)\rangle_t\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t + \langle\overline{\rho}(\boldsymbol{r}_1)\rangle_t\delta\overline{\rho}(\boldsymbol{r}_2)] \\
&= \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2)[\overline{\rho}(\boldsymbol{r}_1)\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t + \langle\overline{\rho}(\boldsymbol{r}_1)\rangle_t\overline{\rho}(\boldsymbol{r}_2) - \langle\overline{\rho}(\boldsymbol{r}_1)\rangle_t\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t \\
&= \int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2)\overline{\rho}(\boldsymbol{r}_1)\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t - \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 V(\boldsymbol{r}_1 - \boldsymbol{r}_2)\langle\overline{\rho}(\boldsymbol{r}_1)\rangle_t\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t
\end{aligned} \tag{5.9}$$

The last term in the above equation is a simple constant in the energy. Although it is important if one wants to compute the energy of the ground state or some similar quantity it simply disappears when one wants to compute averages of any operators. One can thus forget about this term. We thus see that in the mean field approximation the real interaction between the electrons is replaced by

$$H_{\text{MF}} = \int d\boldsymbol{r}\overline{\rho}(\boldsymbol{r})\int d\boldsymbol{r}_2 V(\boldsymbol{r} - \boldsymbol{r}_2)\rangle_t\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t \tag{5.10}$$

This term is a *one body* Hamiltonian of electrons subjected to the external potential

$$\gamma(\boldsymbol{r}, t) = \int d\boldsymbol{r}_2 V(\boldsymbol{r} - \boldsymbol{r}_2)\rangle_t\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t \tag{5.11}$$

which is created be the *average* of the deviations in the density in all the points in the system. Note of course that we do not know this potential explicitly since we do not know what the average $\rangle_t\langle\overline{\rho}(\boldsymbol{r}_2)\rangle_t$. However we have now replaced the full interacting Hamiltonian by an Hamiltonian of electrons subjected to an external potential $\gamma(\boldsymbol{r}, t)$. Moreover this potential is small if the external perturbation $\lambda(\boldsymbol{r}, t)$ is also small. The full Hamiltonian in presence of the perturbation is thus

$$H = H_0 + \int d\boldsymbol{r}\overline{\rho}(\boldsymbol{r})\gamma(\boldsymbol{r}, t) + \int d\boldsymbol{r}\overline{\rho}(\boldsymbol{r})\lambda(\boldsymbol{r}, t) \tag{5.12}$$

We thus see that the mean field approximation replaces the problem of *interacting* electrons subjected to an *external* perturbation λ by a problem of *non-interacting* electrons subjected to an *effective* perturbation $\lambda + \gamma$. The field γ is the induced perturbation and comes from the fact that due to the interactions between particles there will be an induced field that will try to modify the deviations in the density that would otherwise be created by the external field. The mean field approximation is a very useful and general approximation that applies to many physical situations. In many situations it will indeed capture the main effects of the feedback due to the interactions. For the Hamiltonian (5.12) we can easily compute the average of the density using linear response

$$\langle\overline{\rho}(\boldsymbol{r})\rangle_t = \int d\boldsymbol{r}_2 dt_2 \chi_0(\boldsymbol{r} - \boldsymbol{r}_2, t - t_2)[\lambda(\boldsymbol{r}_2, t_2) + \gamma(\boldsymbol{r}_2, t_2)] \tag{5.13}$$

Since γ depends on $\langle\overline{\rho}(\boldsymbol{r})\rangle_t$ we have an integral self consistent equation that will allow to determine the average value. This is easily done by going in Fourier space

$$\langle\overline{\rho}(\boldsymbol{q},\omega)\rangle = \chi_0(\boldsymbol{q},\omega)[\lambda(\boldsymbol{q},\omega)+\gamma(\boldsymbol{q},\omega)] \tag{5.14}$$

Using (5.11) one gets

$$\gamma(\boldsymbol{q},\omega) = V(\boldsymbol{q})\langle\overline{\rho}(\boldsymbol{q},\omega)\rangle \tag{5.15}$$

which leads to

$$\langle\overline{\rho}(\boldsymbol{q},\omega)\rangle = \frac{\chi_0(\boldsymbol{q},\omega)\lambda(\boldsymbol{q},\omega)}{1-V(\boldsymbol{q})\chi_0(\boldsymbol{q},\omega)} \tag{5.16}$$

Comparing (5.16) with (5.3) one obtains an approximate expression for the charge susceptibility of the interacting electron system

$$\chi_{\rho\rho}(\boldsymbol{q},\omega) = \frac{\chi_0(\boldsymbol{q},\omega)}{1-V(\boldsymbol{q})\chi_0(\boldsymbol{q},\omega)} \tag{5.17}$$

This formula is quite remarkable since now the charge susceptibility is expressed in terms of known quantities namely the Fourier transform of the interaction potential $V(\boldsymbol{q})$ and the *independent* electrons susceptibility $\chi_0(\boldsymbol{q},\omega)$. This formula will thus allow us to explore in details the consequences of the interactions between particles on the various susceptibilities.

5.1.2 Spin and charge

Before exploring such physics let us derive both the charge and spin susceptibilities. We start from the general interaction (4.7) and consider a perturbation that couples separately to the spin up and down

$$H_{\text{pert}} = \int d\boldsymbol{r}[\lambda_\uparrow(\boldsymbol{r},t)\overline{\rho}_\uparrow(\boldsymbol{r}) + \lambda_\downarrow(\boldsymbol{r},t)\overline{\rho}_\downarrow(\boldsymbol{r})] \tag{5.18}$$

It is obvious that with such a perturbation we can reproduce the standard coupling to charge and spin. If one takes

$$\lambda_\uparrow = \lambda_\downarrow = -\mu \tag{5.19}$$

we couple to the total charge density and μ is the chemical potential. On the other hand

$$\lambda_\uparrow = -\lambda_\downarrow = \frac{-h}{2} \tag{5.20}$$

couples to the spin density along z and h is an external magnetic field applied along the z direction. It is obvious that with the two independent fields $\lambda_\uparrow$ and $\lambda_\downarrow$ we can generate any combination of chemical potential and external magnetic field along z.

Performing the mean field approximation on the Hamiltonian (4.7) we get

$$H_{\text{MF}} = \int d\boldsymbol{r}\overline{\rho}_\uparrow(\boldsymbol{r},t)\gamma_\uparrow(\boldsymbol{r},t) + \overline{\rho}_\downarrow(\boldsymbol{r},t)\gamma_\downarrow(\boldsymbol{r},t) \tag{5.21}$$

where

$$\begin{aligned}\gamma_\uparrow(\boldsymbol{r},t) &= \int d\boldsymbol{r}_2 V_{\|}(\boldsymbol{r}-\boldsymbol{r}_2)\langle\overline{\rho}_\uparrow(\boldsymbol{r}_2)\rangle_t + V_\perp(\boldsymbol{r}-\boldsymbol{r}_2)\langle\overline{\rho}_\downarrow(\boldsymbol{r}_2)\rangle_t \\ \gamma_\downarrow(\boldsymbol{r},t) &= \int d\boldsymbol{r}_2 V_\perp(\boldsymbol{r}-\boldsymbol{r}_2)\langle\overline{\rho}_\uparrow(\boldsymbol{r}_2)\rangle_t + V_{\|}(\boldsymbol{r}-\boldsymbol{r}_2)\langle\overline{\rho}_\downarrow(\boldsymbol{r}_2)\rangle_t\end{aligned} \tag{5.22}$$

As usual the expressions are much simpler in Fourier space

$$\begin{aligned}\gamma_\uparrow(\boldsymbol{q},\omega) &= V_{\|}(\boldsymbol{q})\langle\overline{\rho}_\uparrow(\boldsymbol{q},\omega)\rangle_t + V_\perp(\boldsymbol{q})\langle\overline{\rho}_\downarrow(\boldsymbol{q},\omega)\rangle_t \\ \gamma_\downarrow(\boldsymbol{q},\omega) &= V_\perp(\boldsymbol{q})\langle\overline{\rho}_\uparrow(\boldsymbol{q},\omega)\rangle_t + V_{\|}(\boldsymbol{q},\omega)\langle\overline{\rho}_\downarrow(\boldsymbol{q},\omega)\rangle_t\end{aligned} \tag{5.23}$$

The response can thus be written as a response to the total field $\lambda + \gamma$ for the free electron gas. Since for the free electrons gas there is no response of the down spins when there is a perturbation applied to the up spins and vice versa, one has

$$\begin{pmatrix} \langle \overline{\rho}_\uparrow(\boldsymbol{q},\omega)\rangle \\ \langle \overline{\rho}_\downarrow(\boldsymbol{q},\omega)\rangle \end{pmatrix} = \begin{pmatrix} \chi_0(\boldsymbol{q},\omega) & 0 \\ 0 & \chi_0(\boldsymbol{q},\omega) \end{pmatrix} \begin{pmatrix} \lambda_\uparrow(\boldsymbol{q},\omega) + \gamma_\uparrow(\boldsymbol{q},\omega) \\ \lambda_\downarrow(\boldsymbol{q},\omega) + \gamma_\downarrow(\boldsymbol{q},\omega) \end{pmatrix} \tag{5.24}$$

where $\chi_0(\boldsymbol{q},\omega)$ is the susceptibility of e.g. spin up particles

$$\chi_0(\boldsymbol{r},t) = -i\,\theta(t)\langle[\overline{\rho}_\uparrow(\boldsymbol{r},t), \overline{\rho}_\uparrow(0,0)]\rangle_{H_0} \tag{5.25}$$

As we saw in the exercices

$$\chi_0(\boldsymbol{q},\omega) = \frac{1}{\Omega}\sum_{\boldsymbol{k}} \frac{f_{\rm F}(\xi(\boldsymbol{k})) - f_{\rm F}(\xi(\boldsymbol{k}+\boldsymbol{q}))}{\omega + \xi(\boldsymbol{k}) - \xi(\boldsymbol{k}+\boldsymbol{q}) + i\delta} \tag{5.26}$$

Using the explicit expressions of the γ fields the expression (5.24) becomes

$$\begin{pmatrix} \langle \overline{\rho}_\uparrow(\boldsymbol{q},\omega)\rangle \\ \langle \overline{\rho}_\downarrow(\boldsymbol{q},\omega)\rangle \end{pmatrix} = \begin{pmatrix} \chi_0(\boldsymbol{q},\omega) & 0 \\ 0 & \chi_0(\boldsymbol{q},\omega) \end{pmatrix} \begin{pmatrix} \lambda_\uparrow(\boldsymbol{q},\omega) \\ \lambda_\downarrow(\boldsymbol{q},\omega) \end{pmatrix} + \chi_0(\boldsymbol{q},\omega) \begin{pmatrix} V_\parallel(\boldsymbol{q}) & V_\perp(\boldsymbol{q}) \\ V_\perp(\boldsymbol{q}) & V_\parallel(\boldsymbol{q}) \end{pmatrix} \begin{pmatrix} \langle \overline{\rho}_\uparrow(\boldsymbol{q},\omega)\rangle \\ \langle \overline{\rho}_\downarrow(\boldsymbol{q},\omega)\rangle \end{pmatrix} \tag{5.27}$$

The equation for the average densities can easily be solved by inverting the matrix. This gives

$$\begin{pmatrix} \langle \overline{\rho}_\uparrow\rangle \\ \langle \overline{\rho}_\downarrow\rangle \end{pmatrix} = \frac{\chi_0}{[1 - V_\parallel\chi_0]^2 - [V_\perp\chi_0]^2} \begin{pmatrix} 1 - V_\parallel\chi_0 & V_\perp\chi_0 \\ V_\perp\chi_0 & 1 - V_\parallel\chi_0 \end{pmatrix} \begin{pmatrix} \lambda_\uparrow \\ \lambda_\downarrow \end{pmatrix} \tag{5.28}$$

where everything is taken at $(\boldsymbol{q},\omega)$ and these indices have been omitted for clarity.

Thus if one takes $\lambda_\uparrow = \lambda_\downarrow = -\mu$ one finds that $\langle\overline{\rho}_\uparrow\rangle - \langle\overline{\rho}_\downarrow\rangle = 0$. This means that even with interactions there is no spin response to a change in chemical potential as is obvious intuitively. The total charge response is

$$\langle \rho_\uparrow\rangle + \langle\rho_\downarrow\rangle = \frac{\chi_0}{1 - [V_\parallel + V_\perp]\chi_0}\lambda_\uparrow \tag{5.29}$$

The charge susceptibility is thus

$$\chi_\rho(\boldsymbol{q},\omega) = \frac{-\chi_0(\boldsymbol{q},\omega)}{1 - [V_\parallel(\boldsymbol{q}) + V_\perp(\boldsymbol{q})]\chi(\boldsymbol{q},\omega)} \tag{5.30}$$

The spin response is obtained by setting $\lambda_\uparrow = -\lambda_\downarrow$. In the same way $\langle\overline{\rho}_\uparrow\rangle + \langle\overline{\rho}_\downarrow\rangle = 0$ showing that there is no charge response if one applies a magnetic field. The spin response is

$$\langle \rho_\uparrow\rangle - \langle\rho_\downarrow\rangle = \frac{\chi_0}{1 - [V_\parallel - V_\perp]\chi_0}\lambda_\uparrow \tag{5.31}$$

which gives for the spin susceptibility

$$\chi_\sigma(\boldsymbol{q},\omega) = \frac{-\chi_0(\boldsymbol{q},\omega)/4}{1 - [V_\parallel(\boldsymbol{q}) - V_\perp(\boldsymbol{q})]\chi(\boldsymbol{q},\omega)} \tag{5.32}$$

5.1.3 Static fields: thermodynamic response

The formulas (5.30) and (5.32) show important properties.

Up to non important numbers, linked to units and to the precise definitions of the charge and spin densities, we see that for non interacting electrons

$$\chi_\rho(\boldsymbol{q},\omega) \equiv \chi_\sigma(\boldsymbol{q},\omega) \tag{5.33}$$

This is true for all wavevectors, frequencies and temperatures.

To compute the thermodynamic susceptibilities we can use (5.26). Since we want a *thermodynamic* quantity we must consider the response to a *static* (i.e. with $\omega = 0$ field). Indeed if the field has a time dependence the system does not conserve energy and one is not measuring a thermodynamic observable, i.e. and observable that can be obtained by differentiating the free energy of the system using the standard thermodynamic relations. On the other hand this static field can depend a priori on space. So in order to compute *thermodynamic* quantities one needs to compute

$$\chi(\boldsymbol{q}, \omega = 0) \tag{5.34}$$

If one wants the response to a uniform field, as is often the case, one lets then $\boldsymbol{q} \to 0$. Taking the proper order of limits is here *crucial* since as we will see in the next section the limits $\boldsymbol{q} \to 0$ and $\omega \to 0$ do not commute. For $\omega = 0$ (5.26) becomes

$$\chi_0(\boldsymbol{q}, \omega = 0) = \frac{1}{\Omega} \sum_{\boldsymbol{k}} \frac{f_{\mathrm{F}}(\xi(\boldsymbol{k})) - f_{\mathrm{F}}(\xi(\boldsymbol{k}+\boldsymbol{q}))}{\xi(\boldsymbol{k}) - \xi(\boldsymbol{k}+\boldsymbol{q})} \tag{5.35}$$

The imaginary part is obviously zero. This can be shown by a direct calculation and is also obvious from the fact that for a static field no dissipation can occur in the system. There is no need to put the principal part in the above expression since the numerator cancels at the same place than the denominator. When $\boldsymbol{q} \to 0$ one can expand the difference in energy

$$\begin{aligned} \chi_0(\boldsymbol{q} \to 0, \omega = 0) &= \frac{1}{\Omega} \sum_{\boldsymbol{k}} \frac{(\xi(\boldsymbol{k}) - \xi(\boldsymbol{k}+\boldsymbol{q}))\frac{\partial f_{\mathrm{F}}(\xi)}{\partial \xi}}{\xi(\boldsymbol{k}) - \xi(\boldsymbol{k}+\boldsymbol{q})} \\ &= \frac{1}{\Omega} \sum_{\boldsymbol{k}} \frac{\partial f_{\mathrm{F}}(\xi)}{\partial \xi} \\ &= \int d\xi n(\xi) \frac{\partial f_{\mathrm{F}}(\xi)}{\partial \xi} \end{aligned} \tag{5.36}$$

at low temperatures $\frac{\partial f_{\mathrm{F}}(\xi)}{\partial \xi} = -\delta(\xi)$ and

$$\chi_0(\boldsymbol{q} \to 0, \omega = 0) = -n(\xi = 0) \tag{5.37}$$

which is the density of states at the Fermi level. One recovers directly from the linear response that for noninteracting electrons both the charge and spin susceptibilities are equal, and also equal to the density of states at the Fermi level, just as was the linear coefficient of the specific heat.

For interacting electrons the situation changes. As we can see from (5.30) and (5.32) the interactions enter in a *different* way in the charge and spin susceptibilities. There is thus no reason that these quantities remain equal. For the static and uniform susceptibilities, since (5.37) holds, we see that even when interactions are present the spin and charge susceptibilities go to constants at low temperatures. However this constant is not simply the density of states anymore but is modified by the interactions according to (5.30) and (5.32). Note that the interactions enter in a quite different way in the two functions. We thus recover the important experimental fact that for real systems one has still a behavior similar to the one of free electrons (namely constant spin and charge susceptibilities) but with different parameters. Note that of course the value of the susceptibilities has no reason to be identical to the change of effective mass due to the interactions. One has thus three independent numbers as is the case experimentally (see e.g. Fig. 1.10).

Another important point to notice is that, even if χ_0 is featureless because of the denominator in (5.30) and (5.32), the susceptibilities can in principle diverge. This can happen for either specific values of the frequency ω or momentum $\boldsymbol{q}$ or at a given temperature T. For the static response $\omega = 0$, we will examine in the second part of these lectures the physical meaning of such a divergence in the susceptibilities. We will examine in the next section what happens when the divergence occurs at a finite frequency ω.

5.2 Collective modes

Let us examine the case when one of the susceptibilities is divergent. We look in this section to the case at $T = 0$, and when the divergence occurs at a finite frequency ω. The physics of such situation is easy to understand. The susceptibility (say of charge) relates the response of the system $\overline{\rho}(\boldsymbol{r}, t)$ to an excitation $\lambda(\boldsymbol{r}, t)$. If the excitations is at a given momentum and frequency

$$\lambda(\boldsymbol{r}, t) = \lambda_0 \cos(\boldsymbol{q}\boldsymbol{r} - \omega t) \tag{5.38}$$

it means that the response is of the same form

$$\overline{\rho}(\boldsymbol{r}, t) = \overline{\rho}_0 \cos(\boldsymbol{q}\boldsymbol{r} - \omega t) \tag{5.39}$$

but with an amplitude that is given by

$$\overline{\rho}_0 = \chi(\boldsymbol{q}, \omega)\lambda_0 \tag{5.40}$$

Thus if the susceptibility $\chi(\boldsymbol{q}, \omega)$ is very large it means that even if the amplitude of an external perturbation was extremely small there would be a sizeable fluctuation of the density that would exist in the system. If the susceptibility diverges it means that even if the external perturbation vanishes there would be a mode of oscillations of density that would exist in the solid: i.e. even if $\lambda_0 = 0$, $\overline{\rho}_0$ would be finite.

The divergence of a susceptibility thus signals the existence of a *collective* mode. The charge and spin susceptibilities give rise to the two collective modes of oscillations of density of charge and density of spins respectively. A priori other susceptibilities can diverge leading to other collective modes but the two previous ones are the ones that will in principle occur systematically in an interacting electronic system. We can thus identify the collective modes by examining the expressions (5.30) and (5.32) and finding the values $(\boldsymbol{q}, \omega)$ for which they diverge. This condition will give us the dispersion relation of the collective mode. Note that such collective modes represent oscillations in density (or spin density) of the electron gas. They are thus quite different from the single particle excitations that we saw in the previous chapter. In a Fermi liquid we will thus have quite generally three types of excitations:

1. The single particles excitations: they are the Landau quasiparticles, and are characterized by their effective mass m^* and quasiparticle weight Z_k. They carry a charge e and a spin 1/2, and essentially resemble individual electronic excitations.

2. A collective mode of charge. This is a density oscillation. The total number of particle is constant but the density redistributes in the system. This is a wave of density with a wavevector $\boldsymbol{q}$ and a frequency ω that are related by a dispersion relation $\omega(\boldsymbol{q})$ characteristic of the corresponding mode and given by the divergence of the charge susceptibility. There is no disturbance in the spin density

3. A collective mode of spin. This is a spin density wave. The charge is uniform and undisturbed. The dispersion relation of this density wave $\omega(\boldsymbol{q})$ is given by the divergence of the spin susceptibility

We can easily write the condition for the divergence of the susceptibility. For example for the charge susceptibility one finds from (5.30)

$$\begin{aligned} 1-[V_{\parallel}(\boldsymbol{q})+V_{\perp}(\boldsymbol{q})]\operatorname{Re}\chi_0(\boldsymbol{q},\omega)&=0\\ \operatorname{Im}\chi_0(\boldsymbol{q},\omega)&=0 \end{aligned} \tag{5.41}$$

and the equivalent relation for the spin density waves

$$\begin{aligned} 1-[V_{\parallel}(\boldsymbol{q})-V_{\perp}(\boldsymbol{q})]\operatorname{Re}\chi_0(\boldsymbol{q},\omega)&=0\\ \operatorname{Im}\chi_0(\boldsymbol{q},\omega)&=0 \end{aligned} \tag{5.42}$$

We see that the condition $\operatorname{Im}\chi_0(\boldsymbol{q},\omega)=0$ is in any case mandatory for the susceptibility to diverge. The physical meaning of this condition is clear. Imposing $\operatorname{Im}\chi_0(\boldsymbol{q},\omega)$ is equivalent to imposing that $\operatorname{Im}\chi(\boldsymbol{q},\omega)=0$ as can be directly seen by writing the imaginary part of χ. It is thus equivalent to imposing that there is no dissipation taking place in the system. This is indeed mandatory for a mode to proceed without damping. Indeed in the absence of and external excitation, if there is dissipation taking place it means that the energy must be taken from the density oscillation and that the collective mode will cease to propagate undamped. We thus see that the condition

$$\operatorname{Im}\chi_0(\boldsymbol{q},\omega)=0 \tag{5.43}$$

corresponds to the absence of damping for the collective mode. We will come back to this condition in the section Section 5.2.3. Let us ignore the possibility of damping of the mode for the moment and focuss on the first equation which gives the dispersion relation of the collective mode.

We see that two cases will have to be distinguished depending on whether the interaction is long range or short range. If the interaction is short range, then $V(\boldsymbol{q})$ is convergent when $\boldsymbol{q}\to 0$. In that case if we are interested in the behavior at small $\boldsymbol{q}$ and ω we can essentially replace $V(\boldsymbol{q})$ by $V(\boldsymbol{q}\to 0)$ which is a constant in the above expressions. This is extremely similar to what one has if one considers a contact interaction just as was the case in the Hubbard model. Note that in that case $V_{\parallel}=0$ and $V_{\perp}(\boldsymbol{q})=U$. This would give for the two conditions for the charge and spin collective modes

$$\begin{aligned} 1-U\operatorname{Re}\chi_0(\boldsymbol{q},\omega)&=0 \quad (\text{charge, shortrange})\\ 1+U\operatorname{Re}\chi_0(\boldsymbol{q},\omega)&=0 \quad (\text{spin}) \end{aligned} \tag{5.44}$$

This is for example what happens in systems such as ^{3}He where the atoms being neutral the interaction is short range. Cold atomic gases in optical lattices are also remarkable systems to realize such a short range interaction.

On the other hand if the interaction has a long range part, as is the case for the electron system, the results are different. Indeed in that case $V(\boldsymbol{q}\to 0)$ is divergent. For the Coulomb interaction one has

$$V_{\parallel}(\boldsymbol{q})=V_{\perp}(\boldsymbol{q})=\frac{e^2}{\epsilon q^2} \tag{5.45}$$

We see that the long range part of the interaction drops out of the spin susceptibility. This is quite natural on a physical basis since a disturbance in spin density does not induce an imbalance on charge, and thus cannot feel the Coulomb interaction. For the spin susceptibility one is thus always controlled by the short range part of the interaction leading back to the above condition. For the charge susceptibility on the contrary the condition becomes

$$1-\frac{2e^2}{\epsilon q^2}\operatorname{Re}\chi_0(\boldsymbol{q},\omega)=0 \quad (\text{charge, longrange}) \tag{5.46}$$

We now examine the various solutions for the charge sector.

5.2.1 Short range interaction: zero sound

Let us start with the condition

$$1 - U \operatorname{Re} \chi_0(\boldsymbol{q}, \omega) = 0 \tag{5.47}$$

In order to satisfy this equation we need to find lines of constant $\operatorname{Re} \chi_0(\boldsymbol{q}, \omega)$ in the $(\boldsymbol{q}, \omega)$ plane.

$$\operatorname{Re} \chi_0(\boldsymbol{q}, \omega) = \frac{1}{U} \tag{5.48}$$

The susceptibility can be easily calculated for a dispersion relation of free electrons $\xi(\boldsymbol{k}) = \frac{k^2}{2m} - \frac{k_\mathrm{F}^2}{2m}$ and the calculation is performed as an exercise. We will quote here only the result. If one considers small $\boldsymbol{q}$ and ω such that the ratio

$$s = \frac{m\omega}{k_\mathrm{F} q} = \frac{\omega/q}{v_\mathrm{F}} \tag{5.49}$$

is a constant. In that case for small ω the susceptibility χ_0 goes to a constant that is only dependent of the number s. One has with the condition (5.49)

$$\begin{aligned} \operatorname{Re} \chi_0(\boldsymbol{q}, \omega) &\simeq \frac{m k_\mathrm{F}}{\pi^2} \left[-1 + \frac{s}{2} \ln \left| \frac{1+s}{1-s} \right| \right] \\ \operatorname{Im} \chi_0(\boldsymbol{q}, \omega) &\simeq \begin{cases} 0 & , \quad s < 1 \\ \frac{-m k_\mathrm{F} s}{2\pi} & , \quad s > 1 \end{cases} \end{aligned} \tag{5.50}$$

We thus see that a mode that satisfies (5.49) obeys the equation (5.48), which now fixes the ratio s. We thus see that the collective mode has a dispersion relation for which the frequency ω is proportional to the wavevector q. This is identical to what one expects for a sound wave propagating in the system. Accordingly this collective mode is called zero sound. We will later compare it to the propagation of a normal sound in the system. If we denote c the sound velocity of this mode we see that

$$s = \frac{c}{\boldsymbol{k}_\mathrm{F} v_\mathrm{F}} \tag{5.51}$$

is simply the ratio between the (zero) sound velocity and the Fermi velocity. The equation (5.49) can easily be solved graphically as indicated on Fig. 5.2. We see that two solutions could in principle exist for repulsive interactions $U > 0$. However, only one of the mode has a velocity larger than v_F. According to the second condition (5.50) the mode for which $c < v_\mathrm{F}$ leads to an imaginary part, and it thus damped.

We thus see that due to the interactions a modulation of the density $\overline{\rho}(r, t) = \overline{\rho}_0 \cos(\boldsymbol{q}\boldsymbol{r} - \omega t)$ can propagate in the system. In order to understand precisely the nature of such wave of density, let us compare it with the one that would expect for a normal sound in a gas, which also corresponds to a wave of density and leads to a mode with a finite velocity. The velocity of the normal sound is determined by the two equations coming from the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = 0 \tag{5.52}$$

where j is the current and the fundamental equation of mechanics

$$m\rho \frac{\partial \boldsymbol{v}}{\partial t} = F = -\nabla P \tag{5.53}$$

where P is the pressure. If one assumes small deviations, compared to the equilibrium situation where $\rho = \rho_0$ and $v = 0$, one has (we take for simplicity a one dimensional situation)

$$\begin{aligned} \rho(x, t) &= \rho_0 + \overline{\rho}(x, t) \\ j(x, t) &= \rho(x, t) v(x, t) \simeq \rho_0 v(x, t) \end{aligned} \tag{5.54}$$

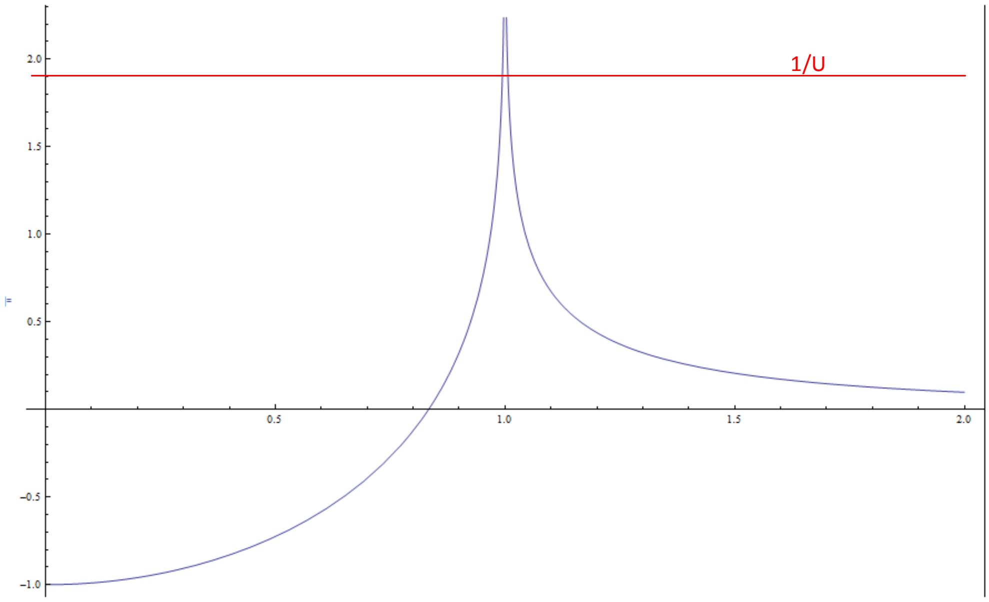


Figure 5.2: Graphical solutions for the equation (5.48). For repulsive interactions two solutions are a priori possible but only one of the modes will be undamped and thus corresponds to the zero sound. When the interaction becomes very small the zero sound velocity tends to $\boldsymbol{F}$. The zero sound velocity increases as the repulsion increases and the system becomes stiffer and stiffer.

Using the definition (1.39) of the compressibility one can relate the pressure and the density. Indeed one has

$$\frac{1}{\kappa}\delta\Omega = -\omega\delta P \tag{5.55}$$

from (1.39). Since $\omega = \mathcal{N}/\rho$ the relation becomes in the lowest order in $\delta\rho$

$$\frac{1}{\kappa}\frac{-\delta\rho}{\rho_0^2}\mathcal{N} = -\Omega\delta P \tag{5.56}$$

which leads to

$$\delta P = \frac{1}{\kappa\rho_0}\delta\rho \tag{5.57}$$

Injecting this relation in the above equations one obtains

$$\begin{aligned} \partial_t\overline{\rho}(x,t) + \rho_0\partial_x v(x,t) &= 0 \\ m\rho_0\partial_t v(x,t) &= -\frac{1}{\kappa\rho_0}\partial_x\overline{\rho}(x,t) \end{aligned} \tag{5.58}$$

which leads to the equation

$$\partial_t^2\overline{\rho}(x,t)\frac{1}{m\rho_0\kappa}\partial_x^2\overline{\rho}(x,t) = 0 \tag{5.59}$$

The normal sound has thus a speed of sound

$$c_s = \frac{1}{\sqrt{\kappa m\rho_0}} \tag{5.60}$$

In order to compare the normal sound (sometimes also called first sound) and the zero sound let us look at the case when the interactions become very small. In that case one sees that the speed of the zero sound tends to

$$c \to v_{\mathrm{F}} \tag{5.61}$$

since $s \to 1$ as can be seen from Fig. 3.1. Using the expression

$$\frac{1}{\kappa} = \frac{4}{9} n(E_{\mathrm{F}}) E_{\mathrm{F}}^2 \tag{5.62}$$

derived in the exercises and the expression (1.14) for the density of states, one gets

$$\frac{1}{\kappa} = \frac{2}{3} \rho_0 E_{\mathrm{F}} \tag{5.63}$$

Which leads for the sound velocity to

$$c_s = \frac{v_{\mathrm{F}}}{\sqrt{3}} \tag{5.64}$$

We thus see that the zero sound and normal sound correspond to two different velocities. They thus corresponds to two different internal ways of having a wave of density propagating in the system. The zero sound is thus a novel mode characteristics of a Fermi liquids and not existing in a normal gas.

To understand the difference between the two type of sound requires not only a determination of the velocity, and the knowledge that both modes corresponds to a propagating wave of density but also on how exactly the excitations are distributed in the system to create this type of density wave. The precise calculation would carry us too far (see [Leg65] for details), so we will only describe the results qualitatively. The main difference between a Fermi liquid and a normal gas, is the separation that exists between interaction acting on the quasiparticles and the scattering among the quasiparticles. In a normal system, both effects are directly related. In a perfect gas there are no interactions and no scattering, but as the interactions are reintroduced in a real gas both the interaction effects and the scattering of the particle grow essentially in the same way. This is not the case, as we already saw, in a Fermi liquid. Since the excitations interact with the soup of the ground state, one can have strong interaction effects on the quasiparticles while being in an essentially collisionless regime for the quasiparticles. It means that contrarily to the case of a normal gas one can have excitations around the Fermi surface that will be of a particular shape and that will be able to retain their shape without being "redistributed" by collision. This is summarized in Fig. 5.3. This difference is at the essence of the difference between the normal sound and the zero sound. The normal sound corresponds to a situation where the scattering is high. Thus in each portion of the system, excitations very rapidly reach the only possible equilibrium distribution which is an isotropic one around the Fermi surface. On the other hand the zero sound corresponds to the limit of low scattering, and other modes are possible. It is possible to show, either from a phenomenological description as was initially done by Landau, or more microscopically, that the distribution of excitation corresponding to the collective mode is of the form

$$\delta n_0(\boldsymbol{k}) = C \delta(\xi(\boldsymbol{k})) \frac{\cos\theta}{\cos\theta - s} \tag{5.65}$$

where θ is the azimuthal angle with respect to the vector $\boldsymbol{q}$ giving the direction of the wave propagation. This indicates that the excitations are localized close to the Fermi level but correspond to a redistribution of particles around the Fermi surface that is anisotropic and very seriously so if s is close to one. The propagation of this wave is the zero sound. A summary is shown in Fig. 5.4 Of course if the propagation is too long, the zero sound will turn into the normal sound as the scattering even if small will finish by killing the anisotropic shape. In the same way if the temperature is too high the scattering is too high (as we saw it essentially grows as T^2) and only the normal sound will be able to propagate.

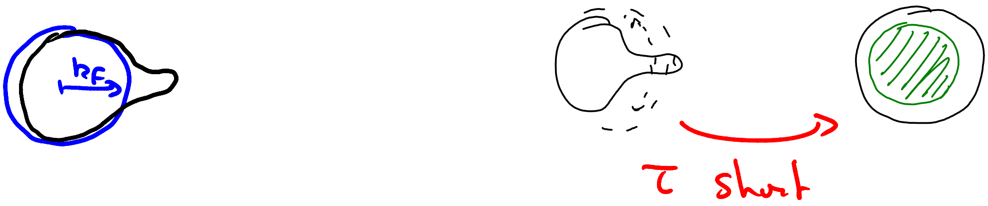


Figure 5.3: (left) In a Fermi liquid a distribution of excitations can take a well defined shape around the Fermi surface. Because of the absence of collisions this shape can be retained. (right) In a normal gas the collisions are high, thus even if some initial anisotropic distribution of excitations was created, it would rapidly be brought back to an isotropic one by scattering of the excitations.

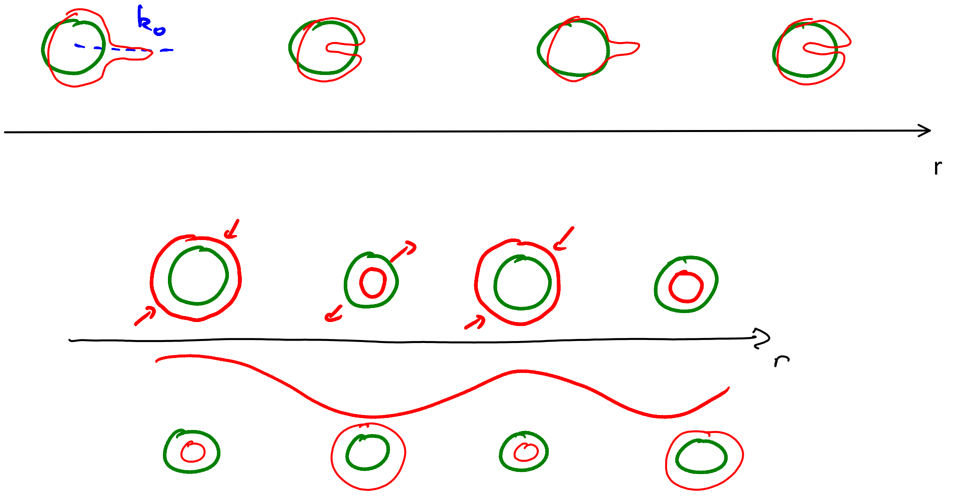


Figure 5.4: (top) The zero sound corresponds to an anisotropic deformation of the Fermi surface that will retain its shape and propagate. (bot) In the normal sound all anisotropies are washed out by the scattering of the quasiparticles, and thus only expansions and contractions of the Fermi surface are possible.

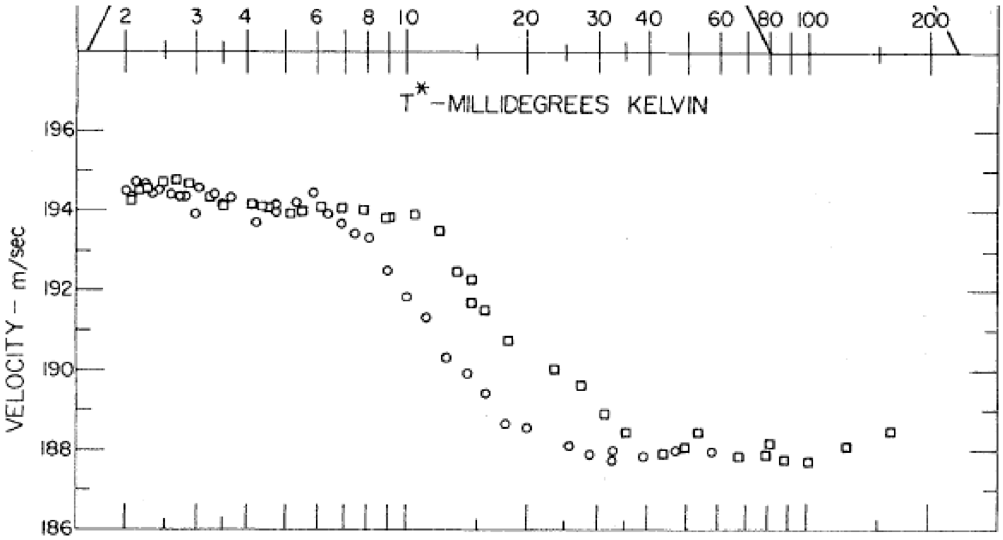


Figure 5.5: Observation of the zero sound in ^{3}He. At high temperature the velocity is the one of the normal sound, and as the temperature is lowered and the scattering decreases, the velocity increases to the one of the zero sound. After [AAW66].

The existence of the zero sounds is thus a remarkable prediction and a strong characteristics of the Fermi liquid. Although the difference of velocity between the two modes is small it can be observed in ^{3}He for which the parameters are controlled enough. This is shown in Fig. 5.5

Let us finish this section by noting that having a "sound" mode with the energy ω going to zero as $\boldsymbol{q}$ goes to zero is in fact a quite general property. This will occur each time the Hamiltonian has a continuous symmetry (here the translation symmetry) and that we consider excitations that corresponds to a spontaneously breaking of this symmetry. In this case one can show that there must be low energy modes (so called Goldstone modes) that have to exist. We will come back on this point in the second part of these lectures.

5.2.2 Long range interations: Plasmon

Let us now consider the case of long range Coulomb interaction. This case is directly relevant for solids. In that case one has to satisfy the condition (5.46), and one immediately sees that searching for modes for which $\omega = cq$ would not works since $\chi_0(\boldsymbol{q},\omega)$ constant would not satisfy the equation. Instead we must find a mode such that $\chi_0(\boldsymbol{q},\omega) \sim q^2$. A quite complete solution of this problem has been given in the exercises. We will just follow here an alternative route to get part of the results in a simple, if less complete, way.

In order to have $\chi_0 \to 0$ when $q \to 0$ one must search a mode for which ω remains finite when $q \to 0$. It is then obvious from (5.26) that this would work. Expanding (5.26) to first order in q would give

$$\operatorname{Re}\chi_0(\boldsymbol{q},\omega) \simeq \frac{1}{\Omega}\sum_{\boldsymbol{k}} \frac{(\xi(\boldsymbol{k}) - \xi(\boldsymbol{k}+\boldsymbol{q}))\frac{\partial f_{\mathrm{F}}(\xi)}{\partial \xi}}{\omega} \tag{5.66}$$

Since

$$\xi(\boldsymbol{k}) - \xi(\boldsymbol{k}+\boldsymbol{q}) = \boldsymbol{q}\cdot\nabla\xi \tag{5.67}$$

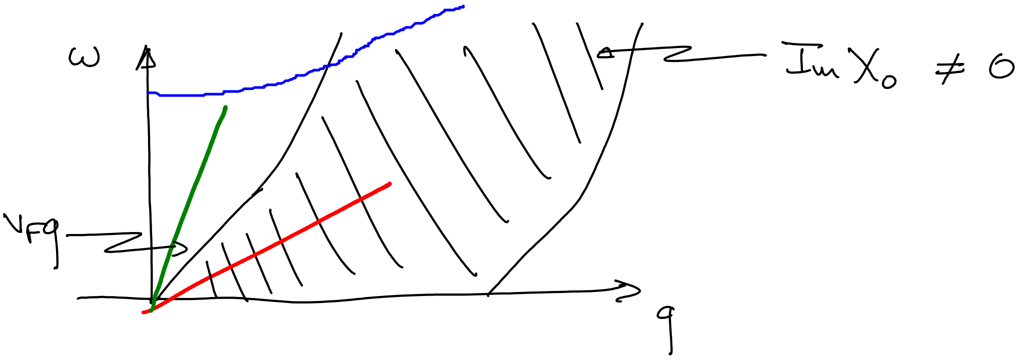


Figure 5.6: Plasmon mode (blue) for systems with long range Coulomb interaction. The dispersion relation of the mode tends to a finite frequency ω_p when the density distortion becomes uniform, indicating an oscillation of the density of charge with time. Not that at small q this mode is indeed undamped. This is to be contrasted with the zero sound mode (green) that exists for short range interactions. This mode is only undamped if its velocity is large than v_{F}. Otherwise (red curve) the mode is damped all the way to $q \to 0$.

this expression looks naively to be of the first order in q. However this is not the case since each term q_α corresponds to a term *odd* in k_α in the integral because of the derivative of the even function of $\boldsymbol{k}$ $\xi(\boldsymbol{k})$. Because for each systems with time invariant symmetry both states k and $-k$ have the same energy, the integral over an odd function of k must vanish. The expansion of $\chi_0(\boldsymbol{q} \to 0, \omega)$ thus starts at order q^2, which is exactly what we were looking for.

For a free electron dispersion the calculation has been performed in the exercises and give

$$\mathrm{Re}\,\chi_0(\boldsymbol{q} \to 0, \omega) \to \frac{k_{\mathrm{F}}^3}{3\pi^2 m}\frac{q^2}{\omega^2} \tag{5.68}$$

Using (5.46) we obtain for the frequency ω of the mode

$$\omega_p^2 = \frac{4\pi e^2 \rho_0}{m} \tag{5.69}$$

which means that even when $q \to 0$ there is an oscillation of the charge of the system at the frequency ω_p. When q becomes finite this mode disperses in a way that was determined in the exercises and is schematically represented in Fig. 5.6. Note that at small momentum q this mode is indeed undamped since the condition $\omega(q)/q > v_{\mathrm{F}}$ is indeed satisfied. We refer the reader to the exercises for a more detailed discussion of the dispersion relation and of the physical properties of this mode.

Let us finish by noting that, as for the zero sound, the existence of the Plasmon is a quite general and important phenomena. It corresponds to a situation of a system that has a continuum symmetry but is in addition coupled to a gauge field (here the electromagnetic field, since there is the Coulomb interaction). This situation occurs in many physical situation, such as for superconductivity ($U(1)$ symmetry of changing the phase of the wave function)) or particle physics for the standard model. In that case, it is possible to show that the Goldstone modes that one would expect in the absence of gauge field do not appear, but that they are absorbed by the gauge field that becomes massive. This mechanism was identified by P. W. Anderson as

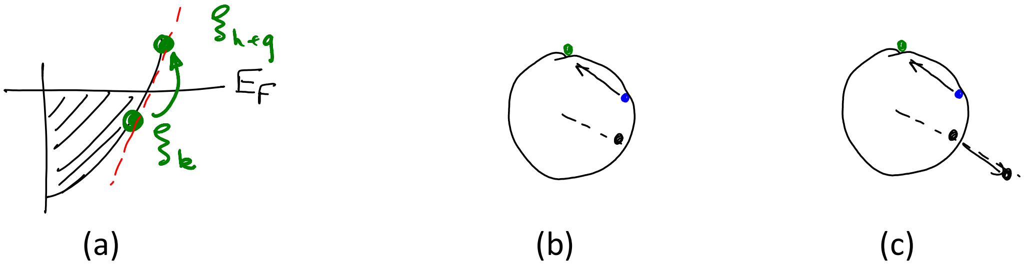


Figure 5.7: (a) An excitation close to $q \sim 0$ costing the maximum energy. One uses all the $\boldsymbol{q}$ to raise in energy away from the Fermi surface. (b) Excitations of essentially zero energy take an electron just below the Fermi surface and recreate it just above the Fermi surface with a different $\boldsymbol{q}$ this is possible for $q < 2k_{\mathrm{F}}$. (c) For $q > 2k_{\mathrm{F}}$ the remaining of the wavevector must be used to increase the energy of the particle giving a minimum value before particle-hole excitations can be created.

being responsible for the existence of the Meissner effect in a superconductor (massive electromagnetic field). It has been extended by Higgs to general symmetry group and is at the heart of the properties of the standard model. It is widely knows as the Higgs (or Anderson-Higgs) mechanism, and we will encounter it in several other situations.

5.2.3 Landau Damping

To finish this chapter on collective modes let us go back to the condition giving the damping of the collective modes

$$\operatorname{Im}\chi_0(\boldsymbol{q},\omega) \neq 0 \tag{5.70}$$

Although one can compute it explicitly for some specific dispersion relationξ as we did in the previous section, let us examine it in general. Writing explicitly the imaginary part one gets

$$\operatorname{Im}\chi_0(\boldsymbol{q},\omega) = \frac{\pi}{\Omega}\sum_{\boldsymbol{k}}[f_{\mathrm{F}}(\xi(\boldsymbol{k}+\boldsymbol{q})) - f_{\mathrm{F}}(\xi(\boldsymbol{k}))]\delta(\omega + \xi(\boldsymbol{k}) - \xi(\boldsymbol{k}+\boldsymbol{q})) \tag{5.71}$$

If one is at $T = 0$ one thus sees that the imaginary part consists in creating particle hole excitations where one take a particle with momentum $\boldsymbol{k}$ below the Fermi surface, and brings it at a higher energy $\xi(\boldsymbol{k}+\boldsymbol{q}) = \xi(\boldsymbol{k}) + \omega$ above the Fermi surface. All other excitations are blocked by the Pauli principle. Since the sum over $\boldsymbol{k}$ is made of positive terms, to identify the region in the plane (q,ω) where the imaginary part is non zero one must thus find for which values of ω and q one can make such particle-hole excitations. It is easy to see, as shown in Fig. 5.7. that the worse one can make in terms of energy is to use all the available wavevector to "climb" in energy. The maximum difference in energy is in that case

$$\xi\boldsymbol{k} + \boldsymbol{q} - \xi(\boldsymbol{k}) \simeq v_{\mathrm{F}}(k + q - k) = v_{\mathrm{F}}q \tag{5.72}$$

This provides an upper limit above which particle-hole excitations of small momentum cannot be created. Note that this condition coincides with the one we found for the free electrons dispersions relation (5.50). This result is in fact quite general. Then it is clear that, as shown in Fig. 5.7 one can make for $0 < q < 2k_{\mathrm{F}}$ excitations of arbitrarily low energy simply by bringing the particle from one point of the Fermi surface to the other. When $q > 2k_{\mathrm{F}}$ one is in a similar

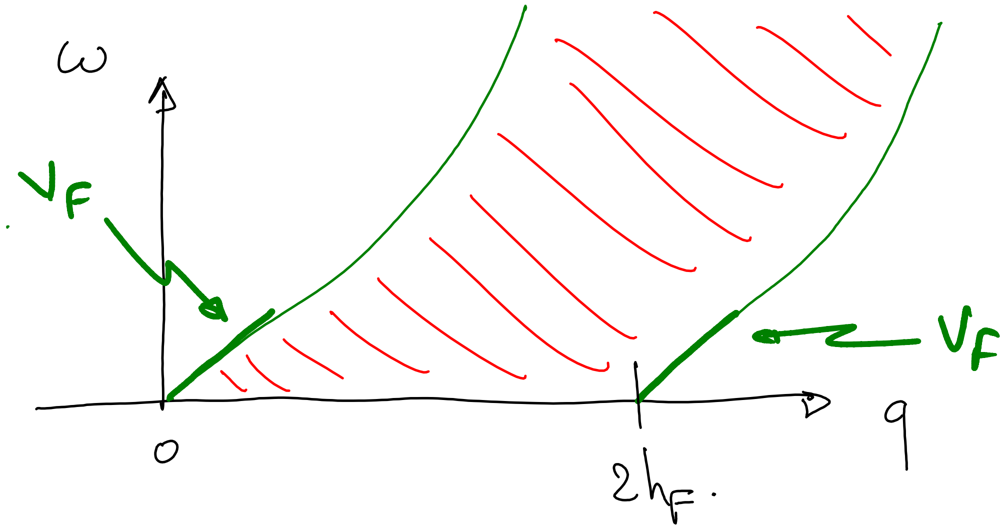


Figure 5.8: Particle-hole continuum. In this region single particle excitations can be excited. If a collective mode enters this zone, it will thus be damped by Landau damping by exciting particle hole pairs.

situation than the one for $q \sim 0$ since in that case $q - 2k_\text{F}$ can only be used to "climb" away from the Fermi surface. This time this gives a minimum value for the energy below which a particle-hole excitation cannot exist. These criteria define the particle-hole continuum showed in Fig. 5.8. In this region the imaginary part of χ_0 is non zero and a collective mode entering this region will thus be damped.

On could however wonder on why a collective mode would be damped at all. Indeed we only considered interaction that conserves energy so it is unclear where such a damping could come on a physical basis, even if the formalism tells us that it should occur. The mechanism of damping of such a collective mode is in fact quite general, and has been explained by Landau and thus termed Landau damping. We can give a very simple physical description of it. If there is a collective mode propagating in the system at a velocity c it can potentially couple to single particle excitations if such *single particle* excitations are also present in the system, and thus exchange (gain or loose) energy with the single particle excitations. In order to get a coupling between the collective mode and the single particles it is necessary for the single particle excitations to be able to "surf" on the collective mode to ensure the time for the energy transfer. This needs that the velocity of the single particle excitation match the ones of the collective mode. The single particle excitations have a velocity v_s, but they need not go in the same direction than the collective mode. Thus the effective velocity seen by the collective mode will be $v < v_s$. In order to ensure the coupling on must thus have

$$c < v_s \tag{5.73}$$

otherwise the collective mode will go too fast for even the fastest quasiparticles. This is exactly the condition that we have found, at small q and ω for a collective mode to enter the quasiparticle continuum and give $\text{Im}\,\chi_0 \neq 0$. We thus see that this physical argument gives back the correct criterion on the velocities. Now we have to check that when there is coupling of the collective modes to the single particle excitations this indeed corresponds to a damping.

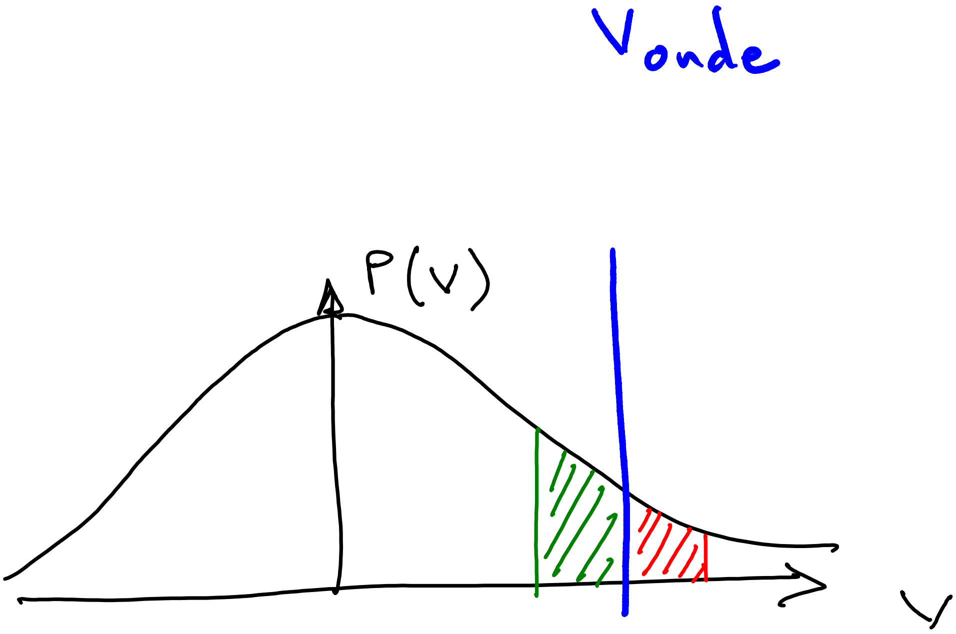


Figure 5.9: Some equilibrium distribution function for the velocities of the single particle excitations. The probability of finding an excitation decreases with the velocity in reasonable equilibrium distributions. There are thus more particles that will get energy from a collective mode at velocity c than particles that can give energy to it, leading to a damping of the collective mode, despite having only energy conserving processes. (red) If the distribution was not remaining at equilibrium the transfer of energy would accelerate particles and change the distribution till a point when the transfer of energy between the single particles and collective mode would be equilibrated.

The collective mode will gain or loose energy from the single particle excitations depending on their respective velocities. If $c > v_s$ the collective mode will accelerate the single particle excitations and thus loose energy, on the other hand if the collective mode is moving slower $c < v_s$ then the single particle excitations will slow down and give energy to the collective mode. We thus see that the total energy in the system is perfectly conserved. However damping comes from the fact that we assume that the system remains in thermodynamic equilibrium and that the single particle excitations distribution function remains the equilibrium one. Given the large number of single particle excitations this is a reasonable physical assumption. If this is the case then the distribution function for the velocity of quasiparticles must decrease with velocity as shown in Fig. 5.9. There will thus be more particles getting energy from the collective modes than particles able to give energy to it. This leads to a damping of the collective mode. Note that the Landau damping is a very general process since the only assumption there is that the distribution of the single particle excitations remains at equilibrium. It can thus occur in many physical situations when a single mode can couple to many single particle excitations. If the distribution was not remaining at equilibrium the slow going particles would be accelerated leading to an flattening of the velocity distribution, up to a point when the transfer of energy between the single particle excitations and the collective mode would stop.

Part II

Strongly correlated systems

CHAPTER 6

Instabilities of the Fermi liquid

As we saw in the previous section the Fermi liquid theory allows for a very successful understanding of the single particle excitations in a solid. We will now look on how the Fermi liquid can have a phase transition into another state, usually with some degree of order.

6.1 Susceptibilities and phase transitions

In section Section 5 we saw how the susceptibilities give access to the collective modes of the Fermi liquid. Susceptibilities also give an indication on the possible phase transitions that the Fermi liquid can undergo. Indeed is there is a certain perturbation h which induces a variation of a certain quantity O (charge, spin, current, etc.) the susceptibility gives

$$\langle O \rangle_{q,\omega} = \chi(q,\omega) h(q,\omega) \tag{6.1}$$

If we look at the $\omega = 0$ part of the perturbation and response, it means that we have applied a static perturbation and that we are looking at the static response. We are thus computing the thermodynamic state of the system. An increasing susceptibility would thus mean that even an extremely small perturbation h would be able to induce a sizeable response in the system. The temperature T_c at which the susceptibility $\chi(q, \omega = 0, T)$ diverges, would thus be the one at which even a perturbation of infinitesimal amplitude would be able to induce a finite response in the system. In other words the system wants to spontaneously order, and for $T < T_c$ would have a spontaneous symmetry breaking with a finite value of $\langle O \rangle$, even in the absence of an external perturbation.

Calculation of the susceptibilities thus allows to predict starting from the Fermi liquid phase, where no symmetry is broken to which state the system would like to go when the temperature is lowered. In general, as indicated in Fig. 6.1 various instabilities can diverge. The phase which is realized is the one with the higher T_c. Below this critical temperature the ground state breaks now a symmetry. Thus the calculation of all the other susceptibilities that was made in the Fermi liquid phase is not valid any more. The other susceptibilities must thus be recomputed, taking into account the new ground state, in general a much more complicated calculation than in the symmetric Fermi liquid phase. This can change completely the behavior of the susceptibilities, and remove or enhance divergence.

Many order are possible. The most common ones involve a modulation of the spin or charge density of the system. The wavevector of the order parameter q is determined by the one of the most divergent susceptibility. Some examples are indicated in Fig. 6.2. The most common type of order corresponding respectively to a divergence of the spin or charge susceptibilities are the ferro- or antiferro-magnetic order. These correspond respectively to a divergence of the $q = 0$ or $q = \pi/a$ (or the suitable generalization to high dimension) of the spin susceptibility. Charge instabilities corresponding to a modulation of the charge density ("charge density wave") are

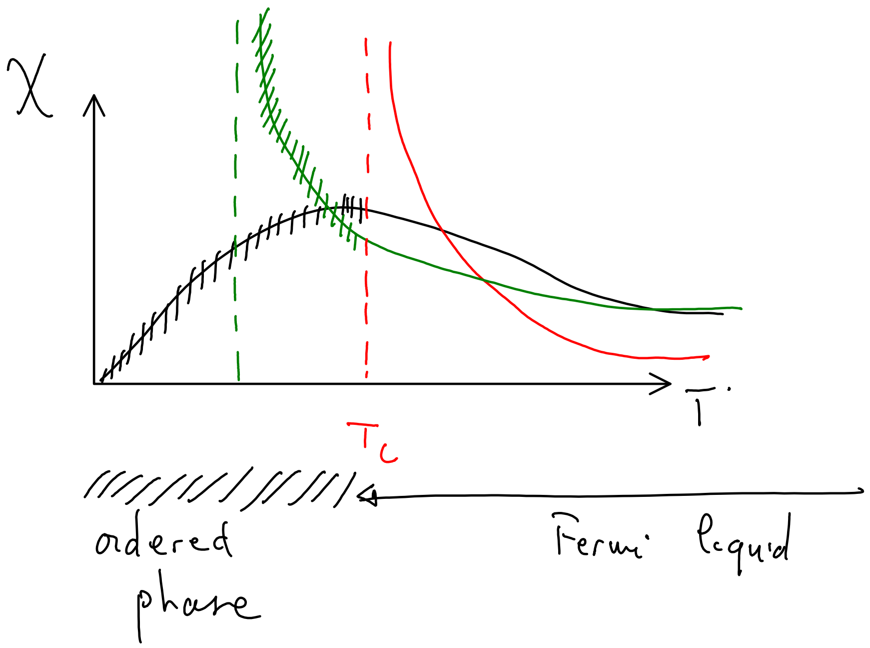


Figure 6.1: Static ($\omega = 0$) susceptibilities can diverge when the temperature is lowered. In general different susceptibilities (shown by different colors) can diverge. This signals a phase transition towards an ordered state. The only one that one can take into consideration is the one with the higher critical temperature. Below this first critical temperature, the ground state breaks a symmetry, and thus all the calculation of the other susceptibilities that were made in the Fermi liquid phase, without such a symmetry breaking must be redone.

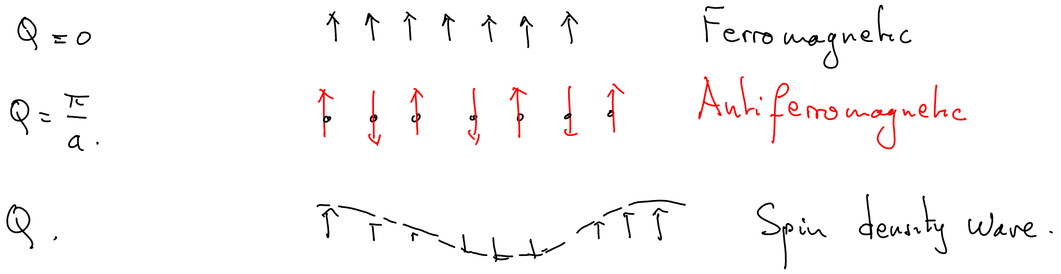


Figure 6.2: Some examples of ordered states. If the spin susceptibility at $q = 0$ diverges one has a ferromagnetic state. If $q = \pi/a$ then the spin order parameter changes sign on each site of the lattice and this is an antiferromagnetic state. More general modulations of the spin density with an incommensurate wavevector q are possible and are generally called “spin density waves”. Similarly if the charge susceptibility diverges at a wavevector q then one has a modulation of the charge density also called “charge density wave”. Other types of order involving other physical observables than charge or spin are of course possible.

also possible and correspond to a divergence of the charge susceptibility. We will also examine the susceptibilities corresponding to more exotic orders. For example one could measure the order parameter corresponding to the existence of pairs in the system

$$\langle \sum_k c^\dagger_{k,\uparrow} c^\dagger_{-k,\downarrow} \rangle \tag{6.2}$$

The ordered state corresponding such an order is the superconducting ground state.

6.2 Spin and charge instabilities

6.2.1 Ferromagnetism

Let us start by first looking at the possibility of spin instabilities. Since getting the exact spin susceptibility of an interacting system is extremely difficult, we will take as an approximation the RPA result that was obtained in (5.32), and that we already used to compute the collective modes. This time we need to evaluate this expression at $\omega = 0$ and finite temperature T. For simplicity we will consider that we deal with a purely local interaction U, thus leading to $V_\parallel = 0$ and $V_\perp(\boldsymbol{q}) = U$ as was discussed when the criterion for instability (5.44) was established. The criterion for a divergence of the static spin susceptibility is thus

$$1 + U \operatorname{Re} \chi_0(\boldsymbol{q}, \omega = 0, T_c) = 0 \tag{6.3}$$

where the free susceptibility is given in (5.35).

In order to determine the nature of the instability we should in principle compute all the susceptibilities as a function of the wavevector $\boldsymbol{q}$ and then determine the one which solves (6.3) with the higher temperature. Since the denominator in (5.35) grows when $\boldsymbol{q}$ grows one can naively expect that the largest χ_0 will be the one with $\boldsymbol{q} = 0$, and thus that the ferromagnetic instability would be the most favored of all the spin ones. Let us thus first start by examining the possibilities to get such a ferromagnetic instability, and we will come back the other $\boldsymbol{q}$ instabilities later.

As was shown in (5.36), one has

$$\chi_0(\boldsymbol{q} \to 0, \omega = 0, T) = \int d\xi n(\xi) \frac{\partial f_{\mathrm{F}}(\xi)}{\partial \xi} \tag{6.4}$$

At $T = 0$ $\frac{\partial f_{\mathrm{F}}(\xi)}{\partial \xi} = -\delta(\xi)$ and the above expression leads readily to

$$\chi_0(\boldsymbol{q} \to 0, \omega = 0, T \to 0) = -n(0) \tag{6.5}$$

More generally at finite temperature one has

$$\chi_0(\boldsymbol{q} \to 0, \omega = 0, T) = -\int d\xi n(\xi) \frac{\beta}{4\cosh^2(\beta\xi/2)} \tag{6.6}$$

which can in principle be evaluated for a realistic density of states as a function of the temperature. In order to make a crude estimate of the susceptibility we will make the approximation that the density of state is totally flat within the band $n(\xi) = n(0)$, and we call Λ the bandwidth of the system. This leads to the approximation

$$\int d\xi n(\xi) \simeq n(0) \int_{-\Lambda}^{\Lambda} d\xi \tag{6.7}$$

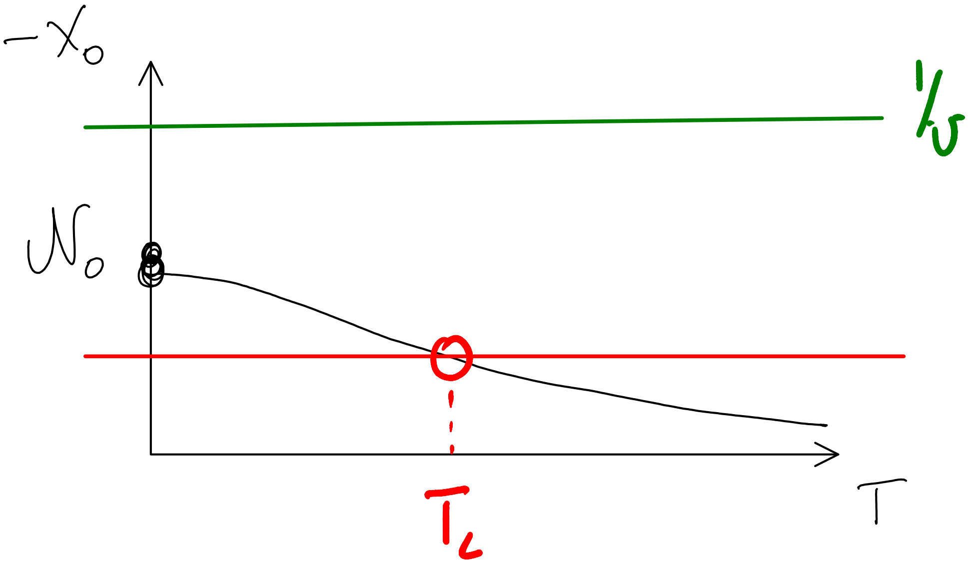


Figure 6.3: The ferromagnetic critical temperature as given by the RPA approximation. If U is not large enough, there is no solution and the system remains paramagnetic. Above a certain threshold of interaction there is a finite temperature T_c below which the system develops a ferromagnetic instability.

This allows to evaluate

$$\begin{aligned}\chi_0(\boldsymbol{q}\to 0,\omega=0,T) &\simeq n(0)\int_{-\Lambda}^{\Lambda} d\xi \frac{\partial f_{\mathrm{F}}(\xi)}{\partial \xi}\\ &= n(0)[f_{\mathrm{F}}(\Lambda)-f_{\mathrm{F}}(-\Lambda)]\\ &= -n(0)\tanh(\frac{\beta\Lambda}{2})\end{aligned} \tag{6.8}$$

The solution of the equation (6.3) is thus given graphically as shown in Fig. 6.3. Clearly an instability only exits for repulsive interactions. There is no solution to the equation is the strength of the interaction is below

$$U < U_c = 1/n(0) \tag{6.9}$$

If this is the case the system has no ferromagnetic instability and thus remains paramagnetic down to zero temperature (provided no other instability occurs). Above this value $U > U_c$, a ferromagnetic instability does occur, starting at $T = 0$ for $U = U_c$, and with a growing T_c and the interaction is growing. This criterion deciding on the existence of a ferromagnetic instability is known as the Stoner criterion.

What is the reason for such an instability ? We can understand the physics behind it by a more primitive argument. Let us compare the energy of an unpolarized Fermi sea with the one in which one has partly polarized it, as shown in Fig. 6.4. In the absence of interactions the minimum of energy is clearly reached if one fills in an equal way the up and down Fermi sea. If a local repulsion exists between the particles such a state will lead to a large repulsion cost since spin up and down electrons can be at the same place and thus feel the repulsion.

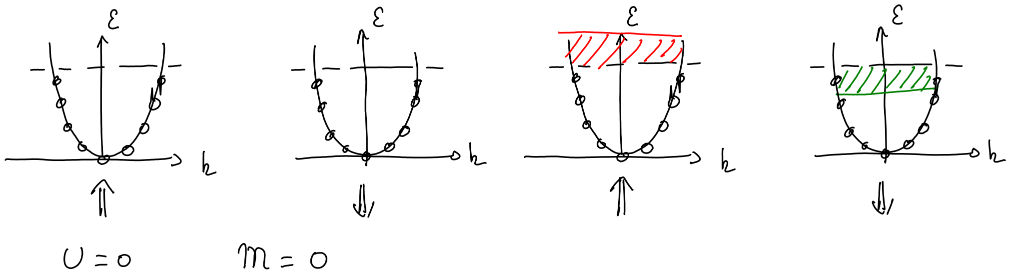


Figure 6.4: In the absence of repulsion the unpolarized Fermi sea allows to reach the minimum of energy while obeying to the Pauli exclusion principle. To avoid a local repulsion between particles one solution is to polarize the Fermi sea, since the Pauli principle will now prevent particles to be at the same place. This however costs kinetic energy. The competition between these two effects determines wether a ferromagnetic instability occurs.

One solution to avoid this would be to have only spin up particles since in that case the Pauli principle would prevent them to be at the same place and thus to feel the repulsion. However doing so would force to fill much more one of the Fermi sea and would thus cost a lot of kinetic energy. To estimate this competition let us assume that one put a small magnetic field on the system which leads to an imbalance of chemical potential $\delta\mu$ for the spin up and down electrons. The excess of kinetic energy per unit volume reads

$$\begin{aligned} H_K/\Omega &= \int_0^{\delta\mu} d\epsilon n\epsilon\epsilon - \int_0^{\delta\mu} d\epsilon n\epsilon \\ &\simeq n02 \int_0^{\delta\mu} d\epsilon n\epsilon\epsilon \\ &= n0(\delta\mu)^2 \end{aligned} \tag{6.10}$$

The interaction cost per unit volume is

$$\begin{aligned} H_U/\Omega &= \frac{U}{\Omega} \int dx \langle (\rho_\uparrow(x) - \rho_0)(\rho_\downarrow(x) - \rho_0) \rangle \\ &= U(+\int_0^{\delta\mu} d\epsilon n\epsilon)(-\int_0^{\delta\mu} d\epsilon n\epsilon) \\ &\simeq -Un0^2(\delta\mu)^2 \end{aligned} \tag{6.11}$$

Both terms vary (at the lowest order) as $(\delta\mu)^2$. Thus two cases are possible If the cost in kinetic energy is the largest, then one indeed needs to put an external magnetic field to polarize the system, but the minimum of energy in the absence of external field remains the unpolarized state. On the contrary if the gain in interaction dominates, then there is clearly an instability of the unpolarized state and the system will spontaneously polarize. We just computed the lowest order term in $\delta\mu$ and thus can only compute the instability but higher order term will lead to a finite value of the polarization for a given interaction. The instability occurs when

$$Un0^2 = n0 \tag{6.12}$$

which is exactly the Stoner criterion that we found by the more general method of the divergence of the susceptibilities.

The existence of such a ferromagnetic instability illustrate a quite important concept that we will encounter several times in correlated materials. The naive expectation for a state that has some degree of spin order would be that such an order results from a direct magnetic exchange between the magnetic moments carried by the electrons in the system. However it is easy to see that such a direct magnetic interaction cannot be responsible for the magnetic order seen in solids. Indeed, the direct magnetic exchange between the moments in a solid is the dipole-dipole interactions

$$H = \frac{\mu_0}{4\pi r^3}[3(m_1\hat{r})(m_2\hat{r}) - m_1 m_2] \tag{6.13}$$

where $m_i = g\mu_B S_i$ is the magnetic moment corresponding to the spin S. To estimate the order of magnitude of this term we take $g = 2$, the magnitude of the spin $S = 1/2$ and a distance between the spins $r \sim a$ where a is the interatomic distance. Rather than using directly $\mu_B = e\hbar/(2m_e)$ it is simpler to use the fine structure constant $\alpha = e^2 c\mu_0/(2h)$. One has thus

$$H \sim \frac{\hbar^3\alpha}{4cr^3m_e^2} \tag{6.14}$$

One can use the fact that in solids $k_{\rm F} \sim \pi/a$ one can reexpress the above as

$$H \sim \frac{\alpha}{2\pi^3}\frac{\boldsymbol{k}_{\rm F}}{c}E_{\rm F} \tag{6.15}$$

Even if the π^3 in the denominator should not be taken too seriously the other factors give already a factor $\sim 10^{-4}$ showing that the dipolar interaction is well below an energy scale of $\sim$ 1K. It is thus impossible that such an interaction can lead to ordered states able to survive at the experimentally observed order (ferromagnetism, antiferromagnetism etc.). The source of magnetic interaction must thus have another source. The above calculation show what the source is. It is the combination of the kinetic energy of the electrons and the interaction between them. Indeed due to the *Pauli principle* a short range interaction will not play similar roles for two spin up and a spin up and a spin down, *even* if the interaction itself only depends on the density of the particles and thus fully respects the spin rotation symmetry. The competition between the kinetic energy and the interaction will thus lead to spin interaction term and thus allow for *magnetically* ordered state. Since the typical energy of the kinetic energy and the interactions is of the order of the electron Volt, it means that the typical scale for these magnetic interactions will be similar, leading to a quite high ordering temperature for ferromagnetic states. We will see that a similar calculation can lead to other type of magnetically ordered state, such as antiferromagnetism.

Finally we need to discuss whether the instability at $q = 0$, namely the ferromagnetism is the dominant one. For a free electron dispersion $\xi(k) = \frac{\hbar^2k^2}{2m} - \frac{\hbar^2k_{\rm F}^2}{2m}$ the free susceptibility (5.35) can be computed exactly. The calculation is left for the exercises. This confirms the intuitive result we were already mentioning above, namely that $\chi_0(q, \omega = 0, T)$ decreases with increasing q. As shown in Fig. 6.5, this ensures that the ferromagnetic instability is the most favored of all spin instabilities.

6.2.2 Charge instabilities

Let us now briefly discuss whether one can get a charge instability, namely a modulation of the charge with a wavevector q, also named a charge density wave (CDW).

The charge susceptibility is given by (5.30). Given the change of sign in the denominator compared to the spin one, no charge instability can occur when $U > 0$. When $U < 0$ one could in principle imagine to have a charge instability, the calculation would be totally similar in that case to the one for the spin. With a free dispersion relation the vector $q = 0$ would be

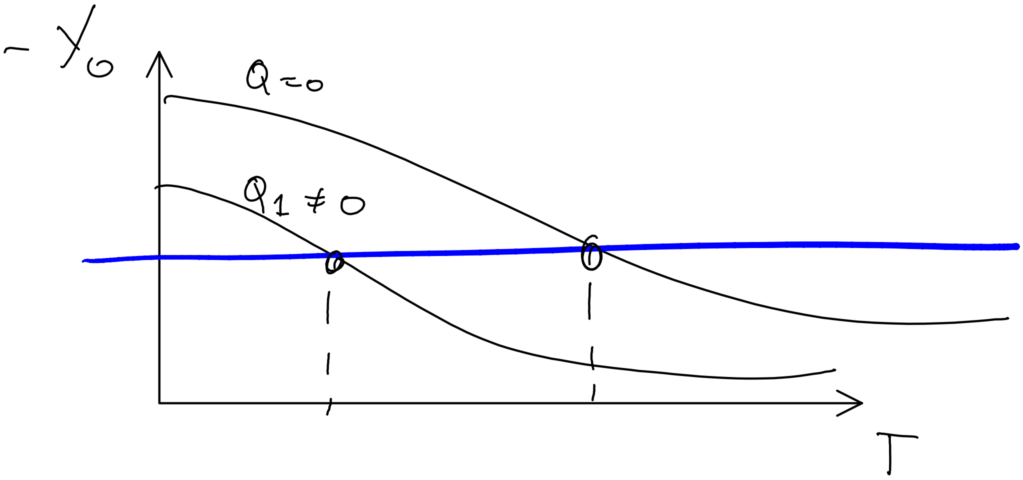


Figure 6.5: For a free dispersion relation $\chi_0(q, \omega = 0, T)$ is a decreasing function of q. The ferromagnetic instability is thus, among all the spin instabilities, the one that will occur at the higher critical temperature. This result is drastically modified if the dispersion relation has nesting properties.

the most favored which would correspond to a global shift of the density which is clearly non physical. As we will see below, for $U < 0$ another instability (pair instability) occurs at a higher temperature for $U < 0$, and thus such a corresponding divergence of the charge susceptibility is meaningless.

However this will not be the case when the energy $\xi(k)$ has special properties as we will now see.

6.3 Nesting of the Fermi surface

So far we have focussed on the susceptibilities computed for a dispersion relation close to the free electron dispersion.

6.3.1 Basic concept

Let us now suppose that the energy obeys the following property: there exists a given vector $\boldsymbol{Q}$ such that for an ensemble of vector $\boldsymbol{k}$ (around the Fermi wavevector) one has

$$\xi(\boldsymbol{Q} + \boldsymbol{k}) = -\xi(\boldsymbol{k}) \tag{6.16}$$

In particular this means that if one has a portion of the Fermi surface $\xi(\boldsymbol{k}) = 0$ this portion can be translated by the vector $\boldsymbol{Q}$ to fall on another portion of the Fermi surface $\xi(\boldsymbol{Q} + \boldsymbol{k}) = -\xi(\boldsymbol{k}) = 0$. Hence the name of the property (6.16), which is the nesting of the Fermi surface. It is important to keep in mind that the full property (6.16) implies much more than just the fact that one can superimpose two portions of the Fermi surface.

Some examples of nesting are indicated in Fig. 6.6 As is quite obvious in Fig. 6.6, the free

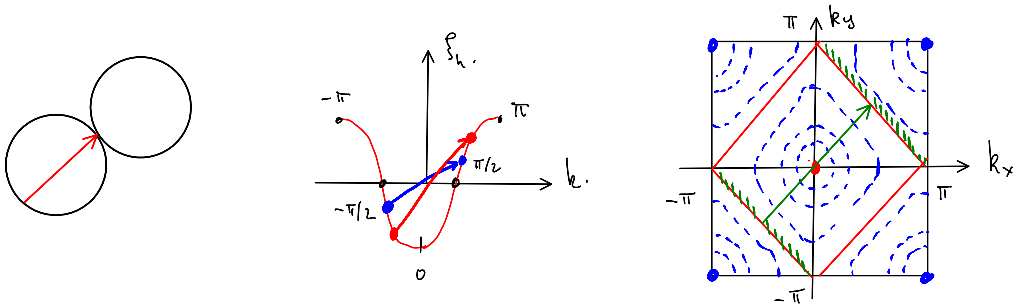


Figure 6.6: Some examples of nesting of Fermi surface. a) for a free electron dispersion there is no nesting of the Fermi surface. b) In one dimension the Fermi "surface" is in fact the two points at $\pm k_{\rm F}$, except for pathological dispersion relations the nesting is always taking place close to the Fermi surface with $Q = 2k_{\rm F}$. c) In some cases, in higher dimensions nesting occurs. Here is the example of a two dimensional tight binding dispersion relation $\xi(k) = -t\cos(k_x a) - t\cos(k_y a)$. The nesting vector is $Q = (\pi/a, \pi/a)$.

dispersion relation $\xi(\boldsymbol{k}) = \boldsymbol{k}^2/(2m) - k_{\rm F}^2/(2m)$ has no good nesting properties. One the contrary if one looks in one dimension, the Fermi surface is reduced to two points. If one considers the expansion of the dispersion close to the Fermi points one has

$$\xi_+(k) = \epsilon(k) - \epsilon(k_{\rm F}) \simeq a(k - k_{\rm F}) = a\delta k_+ \tag{6.17}$$

around $k \sim k_{\rm F}$, whith $\delta k_+ = k - k_{\rm F}$ and

$$\xi_-(k) = \epsilon(k) - \epsilon(-k_{\rm F}) \simeq -a(k + k_{\rm F}) = -a\delta k_- \tag{6.18}$$

and $\delta k_- = k - (-k_{\rm F})$ around $k \sim -k_{\rm F}$. For all systems with an inversion center $\epsilon(k) = \epsilon(-k)$ and thus the linear coefficient a is the same in the two cases. It is thus easy to see that for example if $k \sim -k_{\rm F}$ (see Fig. 6.6)

$$\xi(k + 2k_{\rm F}) = \xi(\delta k_- - k_{\rm F} + 2k_{\rm F}) = \xi(\delta k_- + k_{\rm F}) = a\delta k_- \tag{6.19}$$

and one has thus $\xi(k + 2k_{\rm F}) = -\xi(k)$. The nesting is thus the rule (for any dispersion that allows a linear expansion close to the Fermi surface) in one dimension.

In higher dimensions special dispersions can also lead to nesting. Let us consider for example the case of a two dimensional tight binding dispersion relation

$$\epsilon(\boldsymbol{k}) = -2t_x\cos(k_x a) - 2t_y\cos(k_y a) \tag{6.20}$$

Clearly such a dispersion relation verifies $\xi(k_x + \pi/a, k_y + \pi/a) = -\xi(k_x, k_y)$. The possibility to have nesting or not will thus depend on the filling since $\xi(\boldsymbol{k}) = \epsilon(\boldsymbol{k}) - E_{\rm F}$. Thus if $E_{\rm F} = 0$, namely if the band is half filled, one has a perfect nesting (working for all $\boldsymbol{k}$) in the dispersion relation. If $E_{\rm F} \neq 0$ then the nesting does not occur, or rather becomes worse and worse as one deviates from half filling.

Let us now investigate the consequence of the existence of nesting on the various susceptibilities.

The free susceptibility (5.26) becomes at the nesting vector $\boldsymbol{Q}$

$$\begin{aligned}\chi_0(\boldsymbol{Q},\omega=0,T) &= \frac{1}{\Omega}\sum_{\boldsymbol{k}}\frac{f_{\mathrm{F}}(\xi(\boldsymbol{k}))-f_{\mathrm{F}}(\xi(\boldsymbol{k}+\boldsymbol{Q}))}{\xi(\boldsymbol{k})-\xi(\boldsymbol{k}+\boldsymbol{Q})} \\ &= \frac{1}{\Omega}\sum_{\boldsymbol{k}}\frac{f_{\mathrm{F}}(\xi(\boldsymbol{k}))-f_{\mathrm{F}}(-\xi(\boldsymbol{k}))}{2\xi(\boldsymbol{k})} \\ &= \int d\xi n(\xi)\frac{-\tanh(\beta\xi/2)}{2\xi} \\ &\simeq -n(0)\int_{-\Lambda}^{\lambda}d\xi\frac{-\tanh(\beta\xi/2)}{2\xi} = -n(0)\int_0^{\lambda}d\xi\frac{\tanh(\beta\xi/2)}{\xi}\end{aligned} \tag{6.21}$$

where one has made the usual approximation of a constant density of states over a band of width 2Λ. To evaluate the integral one notice that in the absence of the tanh the integral would be divergent at small ξ. One thus split the integral in two parts $\beta\xi \gg 1$ where the tanh can be approximated by 1 and small values of ξ such that $\beta\xi \ll 1$ for which $\tanh(x) \sim x$. One has thus (with C a constant of order one)

$$\begin{aligned}\chi_0(\boldsymbol{Q},\omega=0,T) &= -n(0)\int_0^{\lambda}d\xi\frac{\tanh(\beta\xi/2)}{\xi} \\ &\simeq -n(0)[\int_0^{2C/\beta}d\xi\beta/2+\int_{2C/\beta}^{\Lambda}d\xi\frac{1}{\xi} \\ &= -n(0)[C+\log(\Lambda\beta/(2C))] \\ &= -n(0)\log(\Lambda\beta/(2C'))\end{aligned} \tag{6.22}$$

where C' is a number of order one. The main consequence is that now, contrarily to what happened for the free dispersion relation (except in $d=1$), in presence of nesting, the free susceptibility is *divergent* when the temperature is lowered. The difference is emphasized in Fig. 6.7

6.3.2 Spin susceptibility

Since in presence of nesting the free susceptibility is divergent at Q, this is usually the wavevector for which the instability will occur. The instability equation (6.3) now becomes

$$1+-Un(0)\log(\Lambda\beta/(2C'))=0 \tag{6.23}$$

Since the free susceptibility is divergent, now one as a finite T_c regardless of the strength of the interactions. The critical temperature is given by (absorbing the various constants in the bandwidth Λ)

$$T_c \sim \Lambda e^{-1/(n(0)U)} \tag{6.24}$$

This is a remarkable result since it shows that now a spin instability can develop even for a very weak interaction. Of course (6.24) shows that the critical temperature will fall very fast when the interaction is weak compared to the density of states. If one keeps in mind that roughly $n(0) \sim 1/E_{\mathrm{F}}$, since one spread one or two states over a spread of energy of order E_{F}, the characteristic ratio is U/E_{F}. Because of the exponential dependence in this parameter this explain why one can also find critical temperatures in a solid which are several orders of magnitude smaller than the typical energy scales of the interactions or kinetic energies. This is in particular the case for the superconducting instability on which we will come back later. Note also that (6.24) is non analytic in U and thus cannot be obtained by any simple perturbative expansion in the coupling U.

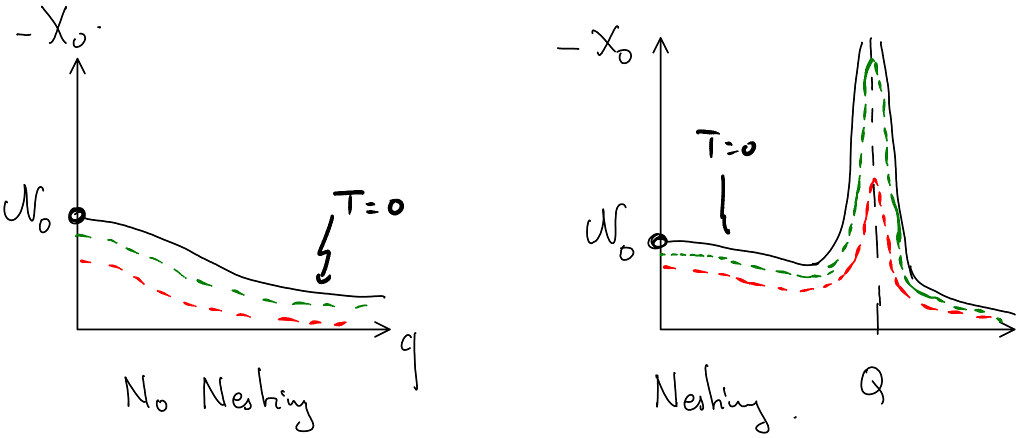


Figure 6.7: a) In the absence of nesting the susceptibility is in general maximum at $Q = 0$ and when the temperature is lowered saturates to a maximum value. One thus needs a critical value of the interaction to get an instability b) In presence of nesting there is a divergence of the free susceptibility when the temperature is lowered at the nesting vector Q. Since the susceptibility is unbounded at that wavevector it means that in general instabilities will always occur at the nesting wavevector, regardless of the strength of the interactions.

Physically this means that systems for which there is a nesting of the Fermi surface will have a spin instability at the nesting wavevector Q. This wavevector can be incommensurate with the lattice of the system, for example for one dimensional systems for which $Q = 2k_{\mathrm{F}}$. In that case one has a modulation of the spin density, and one speaks of a spin density wave (SDW). In many case the nesting vector is $Q = (\pi/a, \pi/a, \ldots)$. This is for example the case of the half filled band on a square lattice. The spin modulation is thus of the form $\cos(\boldsymbol{Q}\boldsymbol{r})$. Since $r = (n_x a, n_y a, \ldots)$, the magnetization is thus simply

$$S = (-1)_x^n(-1)_y^n \cdots \tag{6.25}$$

and thus one has an *antiferromagnetic* order, for which the uniform magnetization remains zero, but there is a staggered magnetization that appears.

As a rule of thumb one can thus consider that in the absence of nesting, a ferromagnetic instability would rather be favored by repulsive interactions, while it would be antiferromagnetic in the presence of nesting. Of course this is an oversimplification and one has to compute the susceptibilities for a given system to know the precise type of order.

6.3.3 Charge instability

In the presence of nesting one has, in a similar way than for repulsive interactions, the possibility to get a charge instability at the vector Q for attractive interactions. The calculations are exactly the same so I will not repeat them here. The phase that is realized is a phase in which the charge density is spatially modulated at the wavevector Q (a charge density wave). Because of the nesting this instability is strongly reinforced and can now beat the pair formation (superconducting instability) that we will examine later.

Physically an attractive interaction can be mediated by the vibration of the lattice of the material (phonons) in a similar way than the attractive interactions which is at the root of the superconductivity. In the case of nesting superconducting and charge density wave instabilities will thus be in competition. Which one is realized depends in a detailed way on the material.

6.4 Pair instabilities

Let us now examine a very special (but very important) case of instability. This one is less obvious than an order in charge or spin and corresponds to the tendency that the Fermi liquid would have to form electron pairs.

6.4.1 Susceptibility

Let us first look at the susceptibility of the free electron gas for a perturbation that would like to create pairs. We consider the Hamiltonian

$$H = H_0 + [\lambda(t) \sum_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} + \text{h.c.}] \tag{6.26}$$

the perturbation $\lambda(t)$ is thus creating (or destroying for the hermitian conjugate) pairs. We want to compute, in linear response the average number of pairs

$$\langle \sum_k c_{-k\downarrow} c_{k\uparrow} \rangle_{t_0} = \int dt \bar{\chi}_p(t_0 - t)\lambda(t) \tag{6.27}$$

Using linear response the susceptibility to pair creation $\bar{\chi}_p$ is given by

$$\bar{\chi}^0_p(t) = -i\,\theta(t) \langle [\sum_k c_{-k\downarrow}(t) c_{k\uparrow}(t), \sum_{k'} c^\dagger_{k'\uparrow} c^\dagger_{-k'\downarrow}] \rangle_{H_0} \tag{6.28}$$

In order to compute this susceptibility we use the time dependence of the destruction operators for free particles (4.30) to obtain

$$c_k(t) = e^{-i\xi(k)t} c_k \tag{6.29}$$

This leads for the susceptibility (6.28)

$$\begin{aligned} \bar{\chi}^0_p(t) &= -i\,\theta(t) \sum k, k' e^{-i(\xi(k)+\xi(-k))t} \langle [c_{-k\downarrow} c_{k\uparrow}, c^\dagger_{k'\uparrow} c^\dagger_{-k'\downarrow}] \rangle_{H_0} \\ &= -i\,\theta(t) \sum k e^{-i(\xi(k)+\xi(-k))t} [\langle c_{k\downarrow} c^d agger_{k\downarrow} \rangle - \langle c^d agger_{k\uparrow} c_{kup} \rangle] \\ &= -i\,\theta(t) \sum k e^{-i(\xi(k)+\xi(-k))t} [1 - 2 f_{\text{F}}(\xi(k))] \\ &= -i\,\theta(t) \sum k e^{-i2\xi(k)t} \tanh(\beta\xi(k)/2) \end{aligned} \tag{6.30}$$

where we have used that $\xi(k) = \xi(-k)$ for systems with an inversion center. The Fourier transform can be easily performed to give

$$\begin{aligned} \bar{\chi}^0_p(\omega) &= \int_0^\infty dt \chi_p(t) e^{i(\omega + i\delta)t} \\ &= \sum_k \frac{\tanh(\beta\xi(k)/2)}{\omega - 2\xi(k) + i\delta} \end{aligned} \tag{6.31}$$

and the susceptibility per unit volume is

$$\chi_p(\omega) = \frac{1}{\Omega} \sum_k \frac{\tanh(\beta\xi(k)/2)}{\omega - 2\xi(k) + i\delta} \tag{6.32}$$

From this form we can compute the static susceptibility. As usual at $\omega = 0$ only the real part remains. One can push the calculation by again assuming that the density of states is roughly constant over a band of width Λ. One thus obtains

$$\begin{aligned}
\chi_p^0(\omega = 0, T) &= -\frac{1}{\Omega}\sum_k \frac{\tanh(\beta\xi(k)/2)}{2\xi(k)} \\
&= -\int d\xi n(\xi)\frac{\tanh(\beta\xi/2)}{2\xi} \\
&\simeq -n(0)\int_{-\Lambda}^{\Lambda} d\xi \frac{\tanh(\beta\xi/2)}{2\xi} \\
&= -n(0)\int_0^{\Lambda} d\xi \frac{\tanh(\beta\xi/2)}{\xi} \\
&= -n(0)\log(\Lambda\beta/(2C'))
\end{aligned} \tag{6.33}$$

One recognizes exactly the expression that was obtained for the free charge or spin susceptibility in presence of nesting (6.22). This is not a coincidence and has deep reasons. One can view the pair susceptibility as resulting from the nesting "time" of the dispersion relation, namely that two states that are related by time reversal symmetry have the same energy. This is the nesting that is important for a particle-particle susceptibility. As one can readily guess such a nesting always take place, at least for non pathological systems, regardless of the precise form of $\xi(k)$. This ensures that the pair susceptibility will always diverge, showing the natural tendency of the free electron gas to such an instability. By contrast the divergence of a charge or spin susceptibility, which corresponds to a particle-hole excitation, needs a special dispersion form $\xi(k)$ verifying the nesting condition. This is why the divergence of the charge or spin susceptibility only occurs for special forms of the Fermi surface and dispersion.

The fact that the pair susceptibility for the free electron gas diverges, means that the electron gas is on the verge of such an instability. Some interactions should thus be able to induce a true phase transition to a pair ordered state. This is the point we now examine.

6.4.2 Attractive interaction, BCS instability

Let us now consider an interacting electronic system. We take the standard interaction (6.34)

$$H_{\text{int}} = \frac{1}{2\Omega}\sum_{\substack{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{q},\\ \sigma_1\sigma_2}} V(\boldsymbol{q}) c^\dagger_{\boldsymbol{k}_1+\boldsymbol{q},\sigma_1} c^\dagger_{\boldsymbol{k}_2-\boldsymbol{q},\sigma_2} c_{\boldsymbol{k}_2\sigma_2} c_{\boldsymbol{k}_1\sigma_1} \tag{6.34}$$

which for simplicity we reduce to a contact interaction. In that case the interaction rewrites, as discussed in Section 4.1 as

$$H_U = \frac{U}{\Omega}\sum_{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{q}} c^\dagger_{\boldsymbol{k}_1+\boldsymbol{q},\uparrow} c^\dagger_{\boldsymbol{k}_2-\boldsymbol{q},\downarrow} c_{\boldsymbol{k}_2\downarrow} c_{\boldsymbol{k}_1\uparrow} \tag{6.35}$$

Let us now treat this interaction in a way that will illustrate how we can make use of the Fermi liquid concept. We will split it in two parts $H_U = H_{BCS} + H_{\text{rest}}$

$$H_{BCS} = \frac{U}{\Omega}\sum_{\boldsymbol{k}_2=-\boldsymbol{k}_1,\boldsymbol{q}} c^\dagger_{\boldsymbol{k}_1+\boldsymbol{q},\uparrow} c^\dagger_{\boldsymbol{k}_2-\boldsymbol{q},\downarrow} c_{\boldsymbol{k}_2\downarrow} c_{\boldsymbol{k}_1\uparrow} \tag{6.36}$$

In the first term we have restricted the interaction to terms such that $\boldsymbol{k}_2 = -\boldsymbol{k}_1$, the second term represent all the remaining interaction process. As we will see in a moment the first term is specially important since it will directly be the interaction that will correspond to the coupling between the pairs. Given that the corresponding susceptibility in the free electron gas is divergent it is important to treat this term extremely carefully. The other process H_{rest} are also important, but we know from Landau Fermi liquid theory that if we consider $H_0 + H_{\text{rest}}$ the whole excitations of this Hamiltonian will be Landau quasiparticles. We will thus consider that we get rid of H_{rest}. We can thus consider now that the system is described *only* by $H_0 + H_{BCS}$. Of course now H_0 does not represent "true" free electrons anymore, but the Landau quasiparticles. The difference however is only hidden in the parameters m^* etc. This allows us to focuss on the part of the interaction that is producing a singular effect due to the particularly large susceptibility of the free system for pairs.

We can rewrite H_{BCS} as

$$H_{BCS} = \frac{U}{\Omega}(\sum_{\boldsymbol{k}_1} c^\dagger_{\boldsymbol{k}_1,\uparrow} c^\dagger_{-\boldsymbol{k}_1,\downarrow})(\sum_{\boldsymbol{k}_2} c_{-\boldsymbol{k}_2\downarrow} c_{\boldsymbol{k}_2\uparrow}) \tag{6.37}$$

which makes apparent the operator creating pairs

$$O_p^\dagger = \sum_{\boldsymbol{k}_1} c^\dagger_{\boldsymbol{k}_1,\uparrow} c^\dagger_{-\boldsymbol{k}_1,\downarrow} \tag{6.38}$$

In order to study the response of the system described by $H_0 + H_{BCS}$ to the perturbation (6.26) we follow the same mean field method than already described in Section 5.1.1. We approximate the interaction term by its mean-field expression

$$H_{BCS} \simeq \frac{U}{\Omega}[O_p^\dagger \langle O_p \rangle + \langle O_p^\dagger \rangle O_p] \tag{6.39}$$

The full Hamiltonian in presence of the pair perturbation (6.26) is thus

$$H_0 + [\lambda(t) + \frac{U}{\Omega}\langle O_p \rangle] O_p^\dagger + \text{h.c.} \tag{6.40}$$

where of course the averages have to be determined self-consistently. One can then, in a similar way than in Section 5.1.1 use the linear response with the *free* (or rather Fermi liquid) Hamiltonian H_0 to obtain

$$\langle O_p \rangle_\omega = \bar{\chi}_p^0(\omega)[\lambda(\omega) + \frac{U}{\Omega}\langle O_p \rangle_\omega] \tag{6.41}$$

where we denoted by the subscript ω the Fourier transform of the time dependent functions. Solving the self consistent equation leads to

$$\langle O_p \rangle_\omega = \frac{\bar{\chi}_p^0(\omega)}{1 - \frac{U}{\Omega}\bar{\chi}_p^0(\omega)} \tag{6.42}$$

The full (in the mean field approximation) susceptibility per unit volume is thus given by

$$\chi_p(\omega) = \frac{\chi_p^0(\omega)}{1 - U\chi_p^0(\omega)} \tag{6.43}$$

We thus see that interactions are in principle able to induce a phase transition towards an ordered state in which the pair operator has a non zero average value. This state is the superconducting state. From the expression (6.43) and the free susceptibility (6.33) we see that the pair instability can occur when the interaction is attractive. Moreover since the free susceptibility is divergent it will *always* occur in that case, unless superseded by another instability that could occur at a higher temperature, such as a charge instability if the Fermi surface is nested. An attractive interaction can be mediated via the electron-phonon coupling but we will not enter into more details on that mechanism at that point and refer the reader to the literature for more on that point [Tin75]. The critical temperature to pair formation is given by the expression (6.24). For the particular case of the electron-phonon attractive interaction the bandwidth Λ has to be replaced by a smaller energy scale which is the Debye energy since the interaction is not attractive beyond such a scale. Other systems, in particular cold atomic system can lead to a realization of Fermions with attractive interaction without phonons [BDZ08]. The expression (6.24) is exactly the BCS form for the superconducting temperature.

This example thus illustrated the general procedure that one can follow to determine the various instabilities of the Fermi liquid. Of course in more complex cases it is not so easy to determine which order parameter can occur or to compute the corresponding susceptibilities.

6.5 Ordered state

So far we started from the high temperature unbroken Fermi liquid phase and by computing the susceptibilities looked at the temperature at which such a phase becomes unstable. Since we were mostly using linear response, it is impossible to describe what happens in the ordered regime, in which the order parameter is non-zero. In this section we will illustrate on one example, namely the antiferromagnetic instability, how one can extend the calculation to describe, within some mean-field approximation, the ordered state as well.

Let us consider the Hubbard Hamiltonian (4.9) on a cubic lattice. The dispersion relation of the tight binding model on such a lattice is

$$\epsilon(\boldsymbol{k}) = -2t \sum_{j=1}^{d} \cos(k_j a) \tag{6.44}$$

where d is the spatial dimension and we have for simplicity considered that all hopping integrals t are identical. Depending on the filling of the band one has a chemical potential μ such that $\xi(\boldsymbol{k}) = \epsilon(\boldsymbol{k}) - \mu$. It is easy to check that for the free Hamiltonian $\mu = 0$ corresponds to a half-filled band. Compared to (4.9) we rewrite the interaction term in the form

$$H_U = U \sum_i (n_{i\uparrow} - 1/2)(n_{i\downarrow} - 1/2) \tag{6.45}$$

IT is easy to see that the difference between this form and the form (4.9) are simply chemical potential and constant energy terms, that can be safely be absorbed in the definition of μ.

We will in the following focuss on the case where the band is half-filled, so that the average density per spin on each site is $\langle n_{i\sigma} \rangle = 1/2$. Since, as discussed in Section 6.3.1 the free electron dispersion has for half-filling of the band a nested dispersion relation with a nesting wavevector $Q = (\pi/a, \pi/a, \ldots)$, one can expect, for repulsive interactions that this model will have an antiferromagnetic instability at a certain temperature T_c. It means that in the antiferromagnetic case, the average total density on each site remains unchanged while the spin density varies. One thus has

$$\begin{aligned} \langle n_{j\uparrow} - 1/2 \rangle &= (-1)^j m_0 \\ \langle n_{j\downarrow} - 1/2 \rangle &= -(-1)^j m_0 \end{aligned} \tag{6.46}$$

where $(-1)^j = (-1)^{j_x + j_y + \cdots}$ changes sign when one goes from one site to the next. m_0 is the staggered magnetization, since

$$m_z = \frac{1}{2}(n_{j\uparrow} - n_{j\downarrow}) = (-1)^j m_0 \tag{6.47}$$

Using the fact that such order parameter is realized in the ordered phase we make the usual mean-field approximation to the interaction (6.45)

$$\begin{aligned} H_U &\simeq U \sum_i (n_{i\uparrow} - 1/2)\langle n_{i\downarrow} - 1/2\rangle + \langle n_{i\uparrow} - 1/2\rangle(n_{i\downarrow} - 1/2) \\ &= U \sum_i [(-1)^i m_0 (n_{i\downarrow} - 1/2) - (-1)^i m_0 (n_{i\uparrow} - 1/2)] \end{aligned} \tag{6.48}$$

After this mean field approximation the full Hamiltonian $H_0 + H_U$ decouples into two separate Hamiltonians involving separately (but self consistently) the spin up and down

$$\begin{aligned} H_\uparrow &= \sum_k \xi(k) c^\dagger_{k\uparrow} c_{k\uparrow} - U m_0 \sum_i (-1)^i m_0 (n_{i\uparrow} - 1/2) \\ H_\downarrow &= \sum_k \xi(k) c^\dagger_{k\downarrow} c_{k\downarrow} + U m_0 \sum_i (-1)^i m_0 (n_{i\downarrow} - 1/2) \end{aligned} \tag{6.49}$$

Each spin species sees thus a *periodic* potential that is self-consistently generated by the interaction with the other species. Such a periodic potential would naturally favor the presence of spin up on even sites and reduce their presence on odd sites, while for the spin down it will be the opposite since the amplitude of the two periodic terms for the two spin species are opposite. We now need to solve the problem of free electrons in such a periodic potential and then solve the self consistent equations giving the amplitude of the potential as a function of the average occupation on each site.

Fortunately this is a problem we already solved in Section 3.6.2, so we can simply follow exactly the steps of the solution. Let us solve for the spin down, that have exactly the same Hamiltonian than studied in Section 3.6.2, with $V = U m_0$. The solution for the spin up can trivially be obtained by replacing $m_0 \to -m_0$. Given the periodicity we have to reduce the zone and thus introduce the new operators in the reduced zone α_k, β_k related to the original operators by the relation (3.125). Then in order to diagonalize the Hamiltonian we perform the Bogoliubov transformation (3.133), leading to the operators $\gamma_{k\pm}$ with the respective eigenvalues $E_\pm(\boldsymbol{k})$ given by (3.141). Since we now know the eigenvalues and operators diagonalizing the Hamiltonian we can compute the average number of particle per site for the spin up

$$\langle n_{j\downarrow} - 1/2\rangle \tag{6.50}$$

In order to compute the order parameter we simply need the value for $j = 0$. One can of course check directly because of the invariance by translation of he Hamiltonian in the reduced zone (i.e. with a doubled unit cell) that the j dependence of the average has exactly the periodic form expected for (6.46). Using the result (3.150) we get

$$\langle n_{0\downarrow} - 1/2\rangle = \frac{1}{\Omega} \sum_{k \in Z'_B} \left(1 - \frac{U m_0}{\sqrt{\xi(k)^2 + (U m_0)^2}} \tanh(\frac{\beta E(k)}{2})\right) - 1/2 \tag{6.51}$$

Using the fact that the reduced zone contains exactly $\Omega/2$ states, one can rewrite the above expression as

$$\langle n_{0\uparrow} - 1/2\rangle = -\frac{1}{\Omega} \sum_{k \in Z'_B} \frac{U m_0}{\sqrt{\xi(k)^2 + (U m_0)^2}} \tanh(\frac{\beta E(k)}{2}) \tag{6.52}$$

where the energy $E(\boldsymbol{k})$ is given by (3.141)

$$E(\boldsymbol{k}) = [\xi(\boldsymbol{k})^2 + (U m_0)^2]^{1/2} \tag{6.53}$$

The self consistent equation for the order parameter m_0 is thus, using (6.46)

$$-m_0 = -\frac{1}{\Omega} \sum_{k \in Z'_B} \frac{U m_0}{\sqrt{\xi(k)^2 + (U m_0)^2}} \tanh(\frac{\beta E(k)}{2}) \tag{6.54}$$

It can be simply rewritten in terms of the quantity $\Delta = Um_0$ as

$$\Delta = \frac{U}{\Omega}\sum_{k\in Z'_B}\frac{\Delta}{\sqrt{\xi(k)^2+\Delta^2}}\tanh(\frac{\beta[\xi(\boldsymbol{k})^2+\Delta^2]^{1/2}}{2}) \tag{6.55}$$

Such an equation and the Bogoliubov transformation that give now all the physical quantities in terms of the parameter Δ allow to get the physical properties in the ordered state.

Several important physical quantities can be extracted from (6.55). First one note that $\Delta = 0$ is always a solution of this equation. This corresponds to the phase without the antiferromagnetic order, and is clearly the correct solution of the equation at high temperature. Of course in that phase the mean field approximation treats poorly the interaction since the mean-field Hamiltonian with $\Delta = 0$ would simply become the free Hamiltonian. This is clearly an artefact of the mean-field approximation used, but the fact that at high temperature we do recover a phase for which $m_0 = 0$ is quite natural. When the temperature is lowered another solution of the equation has to exist for the antiferromagnetic phase to be possible. One has then in principle to verify which one of the two solutions $\Delta = 0$ and $\Delta \neq 0$ has the lowest free energy to know which phase is realized. Since we have all the eigenvalues of the Hamiltonian, such calculation is easy and we will leave it as an exercise.

In the phase for which $\Delta \neq 0$ we can extract the physics from the solution that we already obtained in Section 3.6.2. The periodic potential, that is here the result of the antiferromagnetic order and the interaction between the spin species, open a gap in the spectrum. We see that below T_c a system of half filled electrons with an interaction on a square (cubic, etc.) lattice is thus *not* a metal as the band theory would naively have predicted but an *insulator*. We have here the first example of drastic modifications that interactions can give to the transport and physical properties of the system, compared to free electrons or Fermi liquids. This question of whether interactions can turn an half filled system from a metal into an insulator is an extremely important and general phenomenon that we will examine in details in the next chapter. Within our mean-field approximation we know for this problem the gap between the two bands, which is simply 2Δ. Δ is thus a physically important quantity. Let us again emphasize that this particular gap is the direct consequence of the interactions between the particles. If we look for a solution with $\Delta \neq 0$ the equation simplifies to

$$\frac{1}{U} = \frac{1}{\Omega}\sum_{k\in Z'_B}\frac{1}{\sqrt{\xi(k)^2+\Delta^2}}\tanh(\frac{\beta[\xi(\boldsymbol{k})^2+\Delta^2]^{1/2}}{2}) \tag{6.56}$$

from which we can extract several interesting quantities. First if we are right at the critical temperature we want the order parameter to go to zero. This is assuming that the order parameter is continuous at the transition (second order transition), and can be confirmed by a full solution of the equation. If we use that $\Delta(T_c) \to 0$ then we have the equation giving the critical temperature as a function of the interaction. Replacing in (6.56) we get

$$\begin{aligned}\frac{1}{U} &= \frac{1}{\Omega}\sum_{k\in Z'_B}\frac{1}{\sqrt{\xi(k)^2}}\tanh(\frac{\beta[\xi(\boldsymbol{k})^2]^{1/2}}{2}) \\ &= \frac{1}{\Omega}\sum_{k\in Z'_B}\frac{\tanh(\frac{\beta_c\xi(\boldsymbol{k})}{2})}{\xi(k)}\end{aligned} \tag{6.57}$$

which is exactly the equation for the critical temperature that we obtained from the susceptibility criterion using (6.3) and (6.21) – the factor of 2 coming from the difference over the Brillouin zones over which the $\boldsymbol{k}$ sum is performed. We thus see that the mean field approximation gives us the same T_c. We have however now the possibility to track down the order parameter as the temperature is lowered.

In particular we can compute Δ at zero temperature. In that case the equation (6.56) becomes

$$\begin{aligned}\frac{1}{U} &= \frac{1}{\Omega}\sum_{k\in Z'_B}\frac{1}{\sqrt{\xi(k)^2+\Delta^2}}\\ &= \int d\xi n'(\xi)\frac{1}{\sqrt{\xi^2+\Delta^2}}\end{aligned} \tag{6.58}$$

This equation can be easily solved numerically but we can easily some limiting cases. Fist let us assume that Δ is small. In that case the integral is dominated by the singularity at small ξ. As usual we make the approximation that the density of states is roughly constant, which leads to

$$\frac{1}{U} = n(0)\int_0^{\Lambda'} d\xi \frac{1}{\sqrt{\xi^2+\Delta^2}} \tag{6.59}$$

which we can solve by splitting the integral as we did before

$$\begin{aligned}\int_0^{\Lambda'} d\xi\frac{1}{\sqrt{\xi^2+\Delta^2}} &\simeq \int_0^{C\Delta} d\xi\frac{1}{\Delta} + \int_{C\Delta}^{\Lambda'} d\xi\frac{1}{\xi}\\ &= C + \log(\Lambda'/(C\Delta)) = \log(\Lambda''/\Delta)\end{aligned} \tag{6.60}$$

$\Delta(T=0)$ is thus given, for small Δ by

$$\delta(T=0) \simeq \Lambda'' e^{-1/(n(0)U)} \tag{6.61}$$

The condition of small Δ is obviously realized when U is small. As discussed before $n(0) \sim E_{\mathrm{F}} \sim \Lambda$ since one is at half-filling. On thus see that when the interaction is small compared to the bandwidth of the system, although the system becomes immediately an insulator, regardless of the strength of the interaction U the gap is exponentially small in the interaction. It becomes progressively of the order of the bandwidth itself when $U \sim \Lambda$. Note that the functional form of the zero temperature order parameter (6.61) is the same than the one for the critical temperature (6.24). A more precise calculation of the integrals shows that one has indeed $2\Delta(T=0) = 3.52T_c$ a relation known as the BCS relation [Tin75], and is valid if the mean-field theory is an adequate description of the transition and the order parameter. In general this ratio can deviate from the value indicated here.

For small interaction our calculation of Δ immediately indicates that we have an exponentially small, but finite staggered (antiferromagnetic) magnetization, since $m_0 = \Delta/U$. Clearly the exponentially small behavior of Δ dominates.

Let us now turn to the case of very large interactions. In that case one can expect Δ to be large compared to all the energies of the band. One can thus simply replace the denominator in (6.58) by Δ to obtain

$$\frac{1}{U} = \frac{1}{\Omega}\sum_{k\in Z'_B}\frac{1}{\Delta} = \frac{1}{2\Delta} \tag{6.62}$$

and thus $\Delta = U/2$, and thus $m_0 = \Delta/U = 1/2$. This means that in the large U limit one has a localized spin 1/2 on each site. This is a quite physical result: if the interaction is very large the particles will sit on the sites of the lattice the spin up on the even sites and the spin down on the odd sites. Creating an excitation would bring two spins on one site and this would cost an energy U which is exactly the gap 2Δ that we have computed. A schematic plot of the gap and of the order parameter is indicated on Fig. 6.8.

We thus found that interactions are able to turn an half filled metal into an antiferromagnetic insulator. In the language of the Bogoliubov transformation this corresponds to a state where all γ_{k-} states have been filled. There is now a gap to start filling the band of the γ_{k+} states.

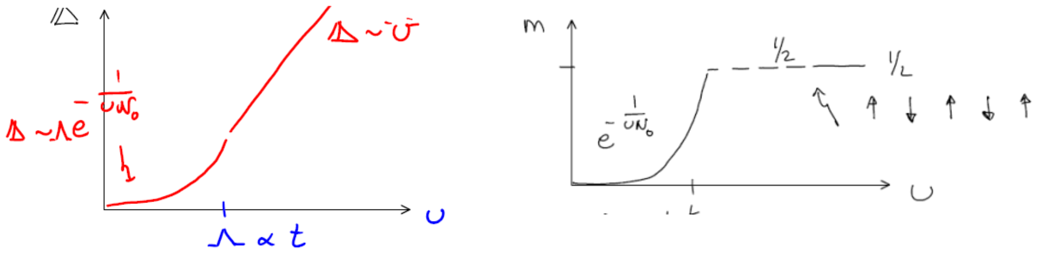


Figure 6.8: Gap 2Δ and staggered magnetization m_0 as a function of the interaction U. Two main regimes must be distinguished. If the interaction is smaller than the bandwidth $U \ll \Lambda$, there is an exponentially small gap and staggered magnetization. Although the system is an antiferromagnetic insulator spin up and down are essentially sharing all the sites of the lattice, practically as in a normal metal. For interactions larger than the bandwidth $U \gg \Lambda$ the gap grows as U and the order parameter is very close to $m_0 = 1/2$ indicating that the spins are essentially localized on each site of the lattice. The resulting state is a Mott insulator as we will discuss in the next chapter.

This solution gives us also an indication of what happens when the system is not exactly half filled. Because we are now starting to fill the band of γ_{k+} states or making holes in the band of γ_{k-} we now recover a gapless ground state, where the excitations have both a good momentum k and a good energy $E_+(k)$ (e.g. assuming we are filling the upper band). Let us look for example what would happen in one dimension. The energy of these states is

$$E_+(k) = \sqrt{4t^2\cos^2(k_x a) + \Delta^2} \tag{6.63}$$

one thus start filling the band around $k = \pm\pi/2a$ which is the lowest energy state. If one sets $k = \pi/(2a) + \delta k$ one has

$$E_+(\delta k) \simeq \Delta + \frac{2a^2t^2(\delta k)^2}{\Delta} \tag{6.64}$$

One thus sees that one can define an effective mass $m^* \sim 1/\Delta$ for these particles. The properties of the doped system are thus very close to the ones of a band insulator where only the extra carriers compared to $n = 1$ will matter. These particles (or holes) can have a large effective mass if the gap is large.

CHAPTER 7

Mott transition

The previous chapter has shown that, because of interactions, the Fermi liquid could be destroyed and lead at low temperatures to an ordered state, with charge, magnetic, superconducting or even potentially more exotic order. One important result was the realization that, because of the interaction an half filled system could develop an antiferromagnetic order, open a gap at the Fermi level and turn into an insulator.

This is the first encounter with an interaction mechanism that can change the credo of band theory that systems with empty or filled band are band insulators while systems with partial band filling are metals or semiconductors. Let us explore this type of physics more in details in this chapter and see how much of this can be generic.

7.1 Basic idea

We obtained in the previous chapter the remarkable result that on a cubic (or hypercubic) lattice an half filled system of interacting fermions would develop an antiferromagnetic order below a certain temperature T_N and as a result would open a gap at the Fermi level and become an insulator. This is a very appealing result since from an experimental point of view, many materials with reasonably strong interactions such as oxides indeed exhibit an insulating behavior for a half filled band. One of the historic examples, shown in Fig. 7.1 is vanadium oxides, whole insulating behavior at half filling is in complete disagreement with the bandstructure predictions. The fact that these properties are linked to interaction effects can be seen experimentally since these insulating properties disappear when the system is put under pressure. The pressure brings the atoms slightly closer together. Its main effect is thus to increase the wavefunction overlap between atomic orbitals on different atoms. As a result the hopping amplitude increases and thus the kinetic energy of the system increases. It also can modify slightly the interactions but the change of wavefunction overlap is much more sensitive to distance and is thus the main effect. Upon pressure the ratio of interactions versus kinetic energy is thus decreased, leading to a less strongly correlated system. At high pressure one recovers a metal, with reasonably normal Fermi liquid behavior.

Although at first glance the solution we found in the previous chapter could seem an explanation of this type of phenomena, it is clear that there must be more to it that this simple solution. Indeed the solution we found has very distinctive features:

1. The insulating character and the existence of an antiferromagnetic ground states are totally linked. It is the existence of the antiferromagnetic order that leads to the doubling of the unit cell and to the opening of an interaction driven gap at the Fermi level. Although the large majority of the systems that are insulating at half filling show indeed an antiferromagnetic order, this is not necessarily the case for all materials. In particu-

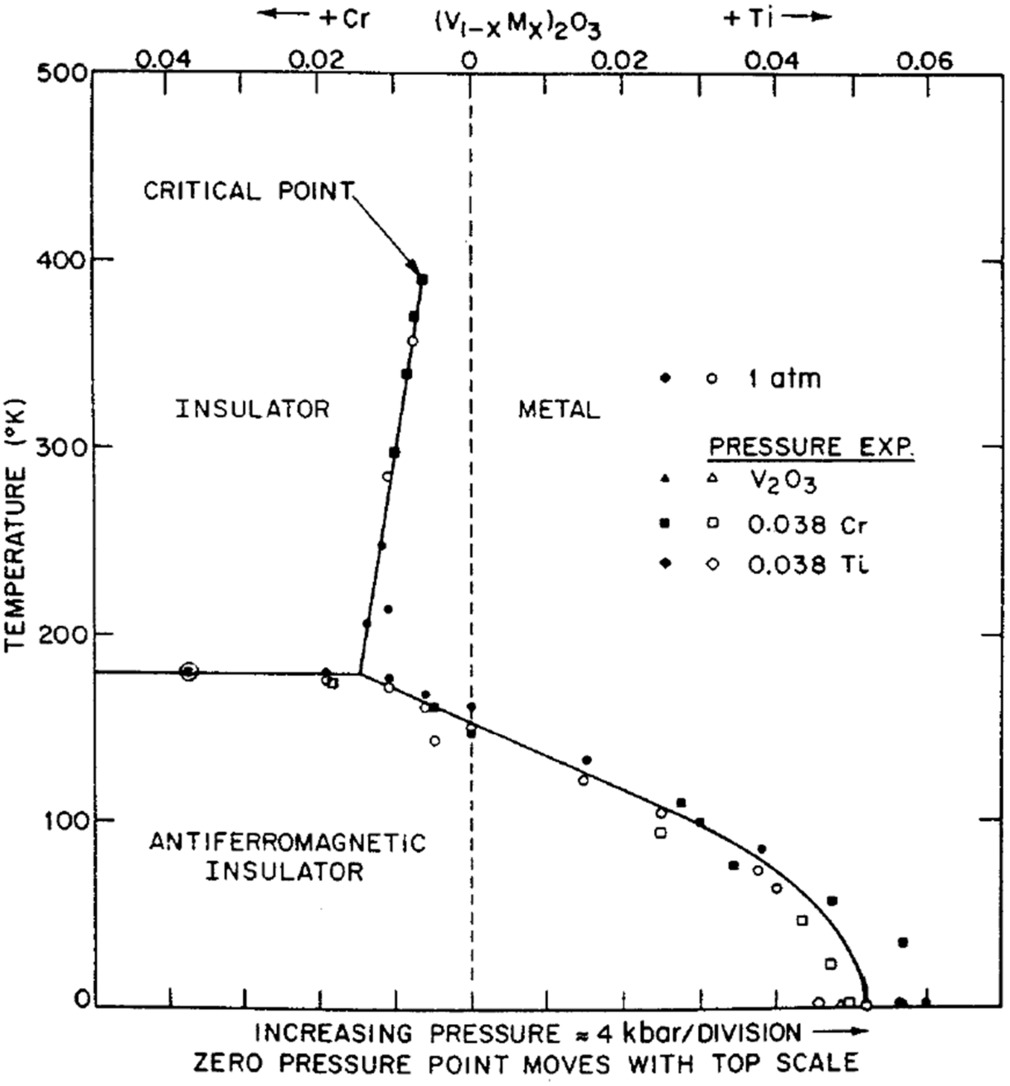


Figure 7.1: Phase diagram of Vanadium oxide as a function of pressure. The filling of the system correspond to half filling. At variance with the predictions of bandstructure theory the system can become an insulator with or without magnetic order.

lar for Vanadium oxides, a clear insulating behavior is observed while the system is still paramagnetic. It is thus clear that there must be a more general mechanism in which the magnetic order and insulating or metallic character are no necessarily linked.

2. The solution we have found was (at least in our derivation) drastically resting on the fact that we had a cubic (or hypercubic) lattice. For other systems (such as a triangular lattice) it is not obvious on how to build such a solution. This does not prevent systems with triangular lattice to be excellent insulators at half filling (for example the cobaltates).

3. The solution we found lead to an insulator for all repulsive interactions (even if the gap is quite small for $U/t \ll 1$. From an experimental point of view it seems that at least for some compounds a critical and relatively large value of the interactions is needed to get the insulating behavior. This for example obvious with the vanadium oxides.

We thus need to find a more general theory that could explain why systems with a commensurate filling (for example one particle per site – half filling) would turn into insulators. Such a theory was given by Sir Nevill Mott, and the corresponding insulators are known as Mott insulators. It rests on the fact that two particles would pay a heavy price in energy if they are on the same site. Such a situation occurs often if the particles are delocalized, as in a metal. To avoid paying such a penalty another state can be more favorable where the particles are localized on each site. This is a mechanism which is totally dominated by the charge properties of the system. In order to illustrate it on a simple example let us consider the Hubbard model (4.9). Indeed this model contains the minimum necessary ingredients to describe the effects of interactions in solids

1. the tight-binding description contains the essential qualitative properties of the bands: finite bandwidth, filling of the band, proper analytic structure of the energy dispersion.

2. the interaction, which is the simplest that one can have for a two species fermionic system.

This is not to say that this Hamiltonian is realistic nor describes all the situations, but it is a very good starting point to understand these phenomena.

In order to have a qualitative idea of what happens let us compare the energy of two different types of wavefunction, which could be potential candidates for the ground state of such a system. The idea is that the wavefunction which has the lowest energy is the one that has the largest overlap with the group state. The first wavefunction is the Fermi sea (3.56). This is obviously the exact ground state if $U = 0$. Let us compute the energy of such a state. The average energy is given by

$$E_{\mathrm{FS}} = \langle \mathrm{F}|H\,|\mathrm{F}\rangle = \langle \mathrm{F}|H_{\mathrm{kin}}\,|\mathrm{F}\rangle + U\sum_i \langle \mathrm{F}|n_{i\uparrow}n_{i\downarrow}\,|\mathrm{F}\rangle \tag{7.1}$$

The kinetic energy is easy to evaluate and is simply

$$E_{\mathrm{kin}} = 2\sum_{k\in \mathrm{FS}} \xi(k) = 2\int^{E_{\mathrm{F}}} d\xi \mathcal{N}(\xi)\xi \tag{7.2}$$

where the factor of 2 comes from the spin summation. The precise value depends on the dispersion relation, but some general feature can be easily seen. Let us take for simplicity a constant density of sites $\mathcal{N}(\xi) = Nn(0)$ if $\xi \in [-W, W]$, where N is the number of sites. This mimics the presence of a finite bandwidth while keeping the calculations simple. The kinetic energy is thus

$$E_{\mathrm{kin}} = N(E_{\mathrm{F}}^2 - W^2)n(0) \tag{7.3}$$

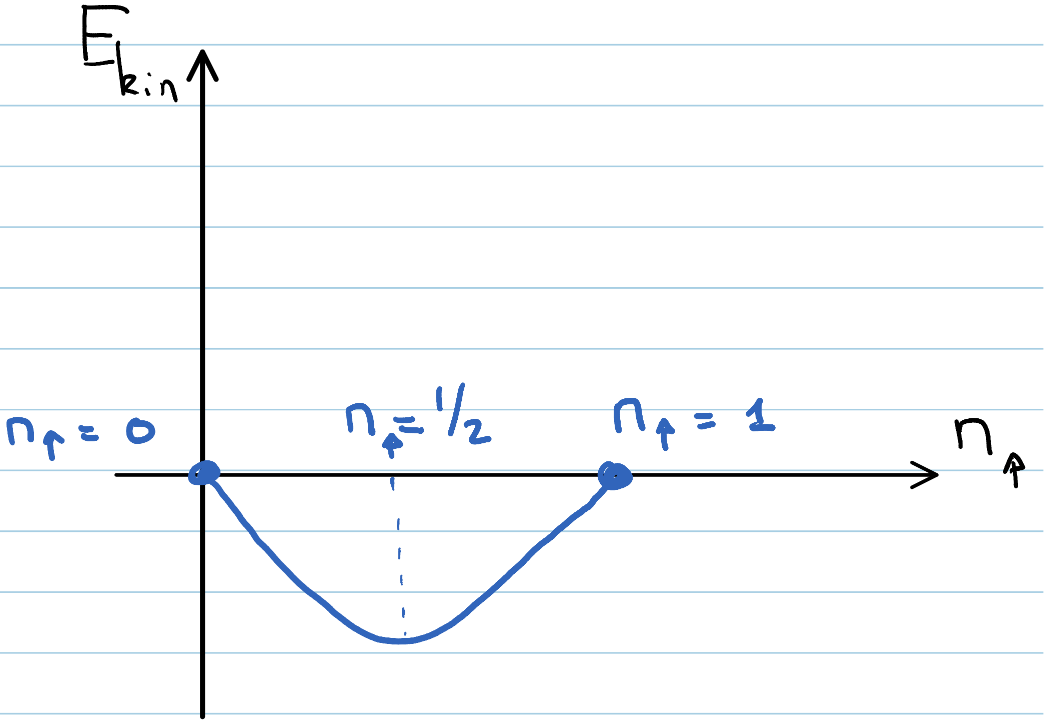


Figure 7.2: Kinetic energy of free fermions with a constant density of states as a function of the filling of one of the spin species n. The kinetic energy is negative indicating the gain due to the delocalization of the particles. The gain is maximum for half filling $n = 1/2$ and zero for an empty band or a filled band.

a simpler expression can be obtained by the number of particles of one spin species per site. Using again the simplified from of the density of states, the total number of particles is given by

$$N_{tot} = N(E_\mathrm{F} + W)n(0) \tag{7.4}$$

thus if we define the filling of one spin species by $n_\uparrow = N_{tot}/N$ we obtain $n_\uparrow = n(0)(E_\mathrm{F} + W)$. From this we can determine the density of state since we want that for a filled band ($E_\mathrm{F} = W$) we have $n_\uparrow = 1$. This imposes $n(0) = 1/(2W)$. One thus obtains

$$E_\mathrm{kin}/N = 2Wn_\uparrow(n_\uparrow - 1) \tag{7.5}$$

Thus the kinetic energy has the form indicated in Fig. 7.2. It is zero for an empty ($n = 1$) or filled ($n = 1$) band since in that case there is no energy gain due to the delocalization of the particles. For all other filling he average energy is negative, which is due to the delocalization of the particles compared to independent sites. The maximum gain is $n = 1/2$, i.e. half filling.

The interaction energy can also be easily computed since the particles are independent. The energy per site is

$$E_U/N = U\langle \mathrm{F}|n_{i\uparrow}n_{i\downarrow}|\mathrm{F}\rangle = U\langle \mathrm{F}|n_{i\uparrow}|\mathrm{F}\rangle\langle \mathrm{F}|n_{i\downarrow}|\mathrm{F}\rangle = Un^2 \tag{7.6}$$

The particles, being in plane wave states are fully delocalized and there is a high probability to find at the same site both a spin up and a spin down. There is a thus an interaction price to pay for that. The total energy looks at depicted in Fig. 7.3. Let us now consider another type

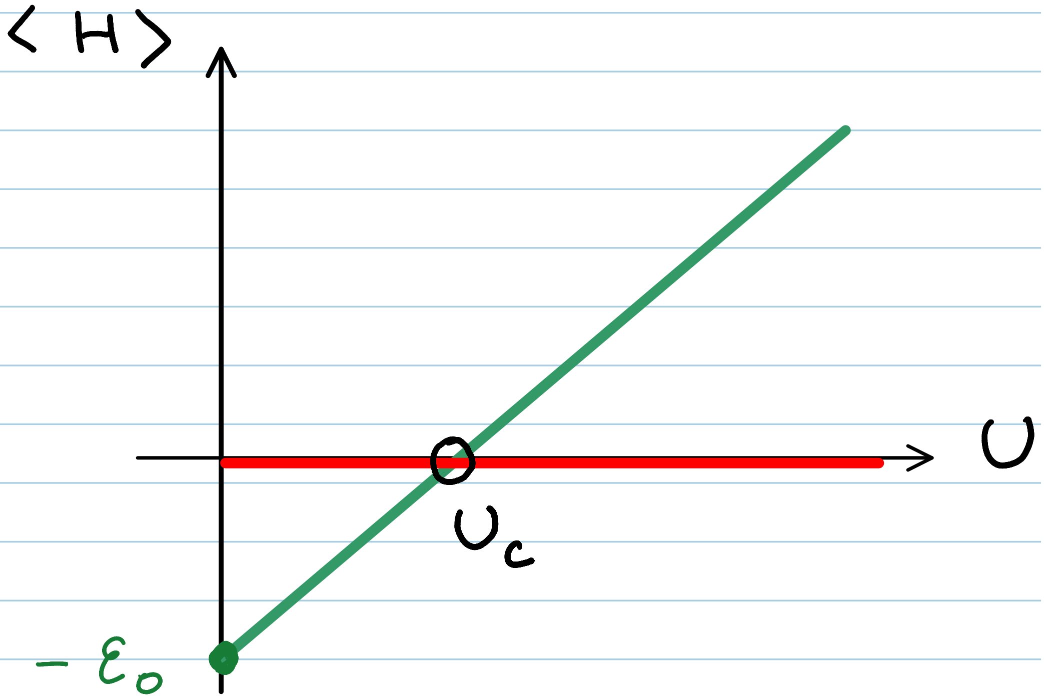


Figure 7.3: Total energy for the Hubbard model for two type of wavefunction. Green: the usual Fermi sea which is the ground state for $U = 0$. The particle gain kinetic energy because of the delocalization, but at the same time have to pay a price $\sim U$ because there are many doubly occupied states. The energy of this function thus becomes increasingly bad as U increases. One the contrary a totally localized wave function (red) has a worse energy at $U = 0$ because particles have to stay on each site, but does not pay the repulsion. The energy of this localized state becomes thus better when U is large. This suggests the possibility of a metal-insulator transition for a critical value of U.

of wavefunction where the particles are localized on each site. For the moment let us restrict to one particle per site ($n = 1/2$). Such a wavefunction is

$$|\mathrm{M}\rangle = \prod_{i=1}^{N} c^\dagger_{i,\sigma_i} |\varnothing\rangle \tag{7.7}$$

Note that there are many such functions ($\mathcal{C}^N_{N/2}$) depending on the choice one makes for the spin orientations σ_i. In the same way one can compute the energy of this function. The kinetic energy is clearly zero since the kinetic term is make of terms such as $c^\dagger_j c_i$ and is moving a particle from one site to the neighboring one. If one applies this term to a function with exactly one particle per site is it clear that one has cannot go back to a function with exactly one particle per site and thus the average of the kinetic energy is zero

$$\langle \mathrm{M}| H_{\text{kin}} |\mathrm{M}\rangle = 0 \tag{7.8}$$

This is traducing the fact that in a localized wavefunction no kinetic energy is gained by the delocalization of the particle. Clearly for zero interaction this state has a much worse energy than the Fermi sea. However because the particles are localized one at each site, there is no interaction due to double occupation. One thus has

$$\langle \mathrm{M}| H |\mathrm{M}\rangle = 0 \tag{7.9}$$

As shown in Fig. 7.3, this function becomes more favorable energetically than the Fermi sea one when

$$U_c = 2W \tag{7.10}$$

The precise value depends of course on the precise evaluation of the kinetic energy, but all results will be of the order of the bandwidth of the system. If we trust our very primitive variational estimate this strongly suggests that the ground state of the system would have a character that would be similar to the one of the Fermi sea, namely a metallic Fermi liquid like behavior when $U < 2W$, while the system would be described by a localized wave function where the particle are well positioned one on each side to avoid repulsion. This describes an insulator. This thus strongly suggests that a metal-insulator transition could occur as a function of the strength of the interactions when there is one particle per site in the system. Note that this transition would require, at variance with the derivation of the previous chapters, a finite value of the interaction of the order of the bandwidth of the system. It is also important to note that the transition never needed to know what are the magnetic properties of the system, and in particular does not need a specific long range magnetic order to be present in the system. It is purely based on the fact that two particle do not like to be on the same site because of the repulsion they feel.

It seems that we are thus on the right tracks since these arguments properly address the concerns and deviations from the experimental facts that the previous derivation encountered. Note that this does not mean that this derivation was incorrect, simply that it might describe a particular case, and that a more general Mott mechanism does exist. We will come back to that point.

7.2 Mott transition

Although the previous arguments are appealing there are of course many objections to them. In particular we have treated very badly the Fermi liquid side by just considering the free Fermi sea as a tentative wavefunction. One could easily imagine that a more complex wavefunction exists in the presence of interactions which manages to avoid paying the interaction price, but at the same time remains metallic. In order to decide whether the Mott transition exists we thus have to use more sophisticated approximations than just the comparison of two wavefunctions. It is of course a very complex problem and no generic solution exists. Strictly speaking the problem is only solved in $d = 1$ and $d = \infty$, and has been tackled by a large variety of numerical and analytical methods. In this course we will examine of of such approximate method, known as the Gutzwiller approximation. It allows relatively simply to illustrate some of the concepts with moderately complicated calculations. Of course more sophisticated approaches exist.

7.2.1 Gutzwiller wavefunction

Let us introduce a better wavefunction than the free Fermi sea. For this we follow the proposal of M. Gutzwiller. The idea is that any wavefunction can be decomposed on the basis of eigenstates of the operator counting the number of doubly occupied sites. Of course such a complete basis is at that stage purely formal since it is extremely complicated to determine explicitly. Nevertheless it allows to represent any wavefunction, so in particular the Fermi sea is

$$|\mathrm{F}\rangle = \sum_{d=0,\nu}^{d=\infty} A_{d,\nu} |d,\nu\rangle \tag{7.11}$$

where ν is the set of quantum numbers besides the double occupation number d in order to have complete basis and the A as the proper coefficients. Now if we want to form a wavefunction in which we want to keep the character of the Fermi sea but at the same time avoid double occupation, the idea would be to simply reduce the weight of the states with a large number of doubly occupied sites in the above decomposition. One way of doing it would be to take the following wavefunction

$$= \sum_{d=0,\nu}^{d=\infty} A_{d,\nu} g^d |d,\nu\rangle \tag{7.12}$$

If $g = 1$ one recovers the Fermi sea, while if $g < 1$ the weight of doubly occupied state is strongly reduced in the wave function. If $g = 0$ no doubly occupied state survives, so for one particle per site this function describes the above insulating state. This wavefunction, known as Gutzwiller wavefunction is thus a nice way to interpolate as a function of g between the free Fermi sea and the insulating state. One can write it in a formal way using the double occupation operator

$$N_d = \sum_i n_{i\uparrow} n_{i\downarrow} \tag{7.13}$$

as

$$= g^{N_d} |\mathrm{F}\rangle \tag{7.14}$$

One can thus in principle undertake a variational study as before but with this more sophisticated function

$$E(g) = H \tag{7.15}$$

and then determine the optimal g by minimizing $E(g)$. If $g > 0$ the system is metallic. If above a certain value of U the optimal solution is $g = 0$ it means one has a metal-insulator transition. The previous variational study we did was limited to a comparison between $g = 1$ and $g = 0$. Note again that there is no reason why this function should be exact nor give the right physics for this transition. It is a convenient wavefunction to try and one must then compare the results one obtains with other techniques or experiments.

Computing the above average is extremely complicated. We will see in the next section how to approximate the wavefunction in order to be able to do the the calculation analytically. But one can illustrate on this wavefunction a numerical technique known as variational monte-carlo which will be explained in the next chapter. In practice the evaluation of the average can be done quite accurately. The result is that g is a smooth function tending to zero when $U \to \infty$ so based on the wavefunction itself there would be no metal-insulator transition. This is clearly an artefact of this wavefunction. However the curve shows a sharp drop of g nearly to zero at $U_c \sim 2W$ as expected. Let us now look at another way of formulating the same problem which will allow us to extract an analytical solution.

7.2.2 Gutzwiller approximation (Nozières formulation)

The average energy with the Gutzwiller wavefunction is too difficult to evaluate analytically. One can make additional approximations (known as Gutzwiller approximation) which allow to perform the calculation. Quite funnily the additional approximation gives results that are more physical than the exact treatment of the original wavefunction. This illustrates again that a variational calculation is not necessarily giving the right physics and has always to be checked by other methods, good common sense or comparison with experiments.

Instead of proceeding with the historical line, and make further approximations on the wavefunction, let us follow an alternative derivation, much more transparent, due to P. Nozires. The idea is to use the density matrix of the system. Indeed any observable can be computed as

$$\langle O \rangle = \mathrm{Tr}[\rho O] = \sum_\alpha \langle \alpha | \rho O | \alpha \rangle \tag{7.16}$$

where $|\alpha\rangle$ is a complete basis of the system, which we will take as the position of all the particles and ρ the density matrix of the system. Introducing another closure relation leads to

$$\langle O \rangle = \sum_{\alpha,\beta} \rho(\alpha,\beta) \langle \beta | O | \alpha \rangle \tag{7.17}$$

where

$$\rho(\alpha,\beta) = \langle \alpha | \rho | \beta \rangle \tag{7.18}$$

are the matrix elements of the density matrix.

Of course we do not know the density matrix of the interacting system, so we need to find an approximation for it. We will build this approximation by imposing that the approximate density matrix must give exactly the averages of the operators that are *diagonal* operators in the basis of position of the particles i.e. the operators for which

$$\langle \alpha | \rho O | \alpha \rangle = O_\alpha \delta_{\alpha,\beta} \tag{7.19}$$

This class of operators includes the number of particles of spin up $N_\uparrow$ the number of particles of spin down $N_\downarrow$ and quite importantly the number of doubly occupied sites N_d. Note that for such an operator

$$\langle O \rangle = \sum_\alpha \rho(\alpha,\alpha) O_\alpha = \sum_\alpha p(\alpha) O_\alpha \tag{7.20}$$

where $p(\alpha) = \rho(\alpha,\alpha)$ is the probability to have the configuration $|\alpha\rangle$.

Since one of the density matrix we know very well is the density matrix for free particles ρ_0 one will simply take it as the basis of our approximation for ρ. One possible ansatz is

$$\rho(\alpha,\beta) = \rho_0(\alpha,\beta) \sqrt{\frac{p(\alpha)p(\beta)}{p_0(\alpha)p_0(\beta)}} \tag{7.21}$$

In this ansatz one keeps for the density matrix ρ the phases of the free one. The amplitudes of each elements have been modified such that one the diagonal one has

$$\rho(\alpha,\alpha) = \rho_0(\alpha,\alpha) \frac{p(\alpha)}{p_0(\alpha)} = p(\alpha) \tag{7.22}$$

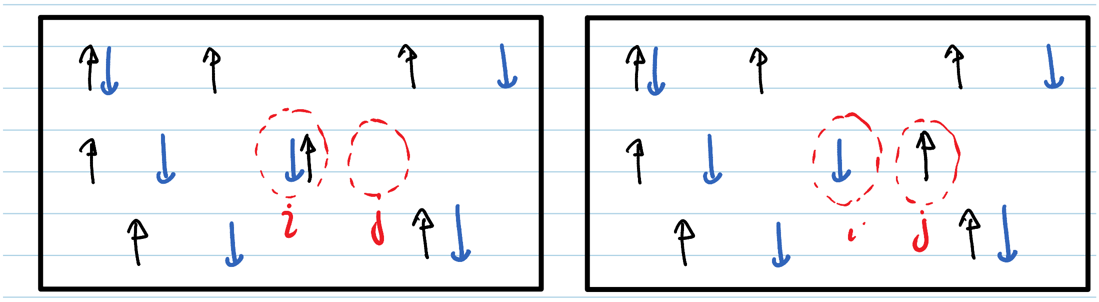


Figure 7.4: The kinetic energy terms only couple configurations α and β which differ by the position of a single particle of, say spin up. The opposite spins are left untouched.

ensuring that the diagonal elements are exact.

Let us now compute the average of the total energy with such density matrix. The interaction energy is very easy since diagonal operators such as the number of doubly occupied sites are exacts

$$\langle H_U \rangle = U \sum_i \langle N_d \rangle = UNd \tag{7.23}$$

where d is the *exact* (and for the moment unknown) number of doubly occupied sites. The kinetic energy is more complex to compute. Each term of the kinetic energy is of the form

$$c^\dagger_{i\uparrow} c_{j\uparrow} \tag{7.24}$$

where i and j are two neighboring sites and there is of course a corresponding term for spin down. The average of such a term

$$\langle c^\dagger_{i\uparrow} c_{j\uparrow} \rangle = \sum_{\alpha,\beta} \rho(\alpha,\beta) \langle \beta | c^\dagger_{i\uparrow} c_{j\uparrow} | \alpha \rangle \tag{7.25}$$

Thus as indicated in Fig. 7.4, the two configurations α and β are identical for all the spin down and differ by one spin up which has been moved to a neighboring site. Let us compute, for the above mentioned configurations α and β which are coupled in (7.25)

$$A = \rho_0(\alpha,\beta) \frac{1}{\sqrt{p_0(\alpha) p_0(\beta)}} \tag{7.26}$$

Because we are dealing with the density matrix of *free* particles, matrix density factorize in a product of the density matrices for the spin up and the spin down. Since the spin down configuration is identical in α and β, the spin down part is purely diagonal and thus cancels between the numerator and denominator. In a similar way, for the spin up, all the sites which are different from i and j are identical in α and β and cancel in the same way. One can thus compute (7.26) by just considering the two sites i and j. Note that this is of course only true because we only consider quantities corresponding to free particles. For the two sites i and j one has the configuration described in Fig. 7.5. One has thus

$$p_0(\alpha) = p_0(\beta) n_\uparrow (1 - n_\uparrow) \tag{7.27}$$

where $n_\uparrow$ is the faction of the sites occupied by a spin up.

In order to finish the calculation let us distinguish in the configurations the parts corresponding to spin up and down. Let us denote $\alpha = \alpha_\uparrow \alpha_\downarrow$ and a similar expression for β. The $\alpha_\uparrow$ and $\alpha_\downarrow$

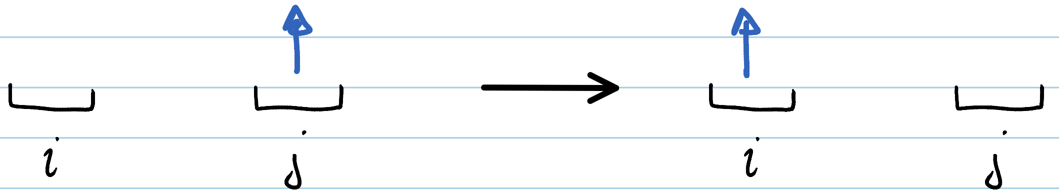
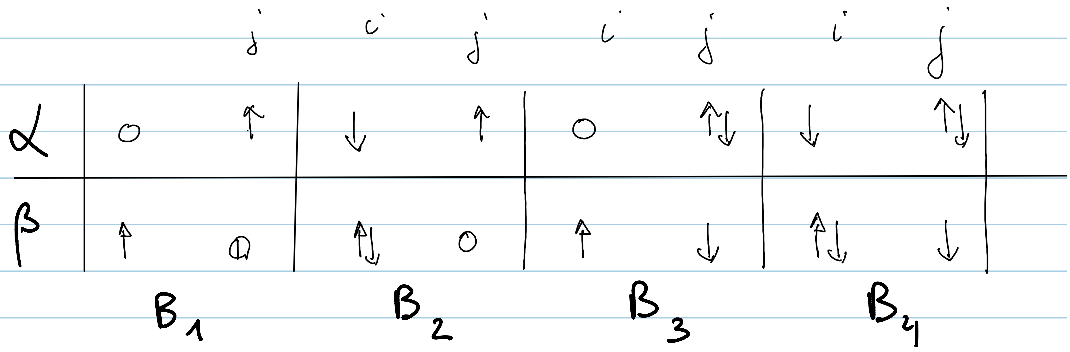


Figure 7.5: Configurations that are coupled by the subpart of the kinetic energy entering in the calculation of the term A (see text).



	B_1		B_2		B_3		B_4	
		j	i	j	i	j	i	j
α	0	↑	↓	↑	0	↑↓	↓	↑↓
β	↑	0	↑↓	0	↑	↓	↑↓	↓

Figure 7.6: All the configurations corresponding to the two sites for which a single spin up has been moved.

denote respectively the positions of the up and down particles respectively. One has

$$\begin{aligned}
\langle c^\dagger_{i\uparrow} c_{j\uparrow}\rangle &= \sum_{\alpha,\beta} \rho(\alpha,\beta)\langle\beta|c^\dagger_{i\uparrow}c_{j\uparrow}|\alpha\rangle \\
&= \sum_{\alpha_\uparrow\alpha_\downarrow,\beta_\uparrow\beta_\downarrow} \rho(\alpha_\uparrow\alpha_\downarrow,\beta_\uparrow\beta_\downarrow)\langle\beta_\uparrow\beta_\downarrow|c^\dagger_{i\uparrow}c_{j\uparrow}|\alpha_\uparrow\alpha_\downarrow\rangle \\
&= \sum_{\alpha_\uparrow,\beta_\uparrow\alpha_\downarrow} \rho(\alpha_\uparrow\alpha_\downarrow,\beta_\uparrow\alpha_\downarrow)\langle\beta_\uparrow|c^\dagger_{i\uparrow}c_{j\uparrow}|\alpha_\uparrow\rangle \\
&= \sum_{\alpha_\uparrow,\beta_\uparrow} \frac{\rho_0(\alpha_\uparrow,\alpha_\downarrow)}{n_\uparrow(1-n_\uparrow)}\langle\beta_\uparrow|c^\dagger_{i\uparrow}c_{j\uparrow}|\alpha_\uparrow\rangle \left(\sum_{\alpha_\downarrow}\sqrt{p(\alpha_\uparrow\alpha_\downarrow)p(\beta_\uparrow\alpha_\downarrow)}\right)
\end{aligned} \tag{7.28}$$

The crucial quantity to compute is, for the two configurations for which a single spin up has been moved to compute

$$C = \sum_{\alpha_\downarrow}\sqrt{p(\alpha_\uparrow\alpha_\downarrow)p(\beta_\uparrow\alpha_\downarrow)} \tag{7.29}$$

Note that because now we are dealing with the density matrix of the interacting system the spin up and the spin down are *not* independent. The probabilities will thus depend on both the $\alpha_\uparrow$, β_up and $\alpha_\downarrow$ configurations. We thus need to enumerate, for the two sites i and j all the possible configurations. This enumeration is shown in Fig. 7.6. We need to compute the factor (7.29) for each of the possibilities shown in Fig. 7.6. We will express the corresponding probabilities in terms of $n_\uparrow$ (resp. $n_\downarrow$) the fraction of sites with a spin up (resp. down), and d

the fraction of double occupied sites. Note that since the system is interacting $d \neq n_\uparrow n_\downarrow$ and is an independent parameter. The different factors B entering in the sum in (7.29) corresponding the four cases of Fig. 7.6 are (noting $n = n_\uparrow + n_\downarrow$)

$$\begin{aligned}
B_1 &= (n_\uparrow - d)(1 - n + d) \\
B_2 &= \sqrt{(n_\uparrow - d)(n_\downarrow - d)d(1 - n + d)} \\
B_3 &= \sqrt{(n_\uparrow - d)(n_\downarrow - d)d(1 - n + d)} \\
B_4 &= d(n_\downarrow - d)
\end{aligned} \tag{7.30}$$

These factors and thus the sum (7.29) is totally independent of the sites i and j. Thus one can simplify (7.28)

$$\begin{aligned}
\langle c^\dagger_{i\uparrow} c_{j\uparrow}\rangle &= \sum_{\alpha_\uparrow,\beta_\uparrow} \frac{\rho_0(\alpha_\uparrow,\alpha_\downarrow)}{n_\uparrow(1-n_\uparrow)} \langle \beta_\uparrow | c^\dagger_{i\uparrow} c_{j\uparrow} | \alpha_\uparrow \rangle \left(\sum_{\alpha_\downarrow} \sqrt{p(\alpha_\uparrow \alpha_\downarrow) p(\beta_\uparrow \alpha_\downarrow)} \right) \\
&= \frac{C}{n_\uparrow(1-n_\uparrow)} \sum_{\alpha_\uparrow,\beta_\uparrow} \rho_0(\alpha_\uparrow,\alpha_\downarrow) \langle \beta_\uparrow | c^\dagger_{i\uparrow} c_{j\uparrow} | \alpha_\uparrow \rangle \\
&= \frac{C}{n_\uparrow(1-n_\uparrow)} \langle c^\dagger_{i\uparrow} c_{j\uparrow}\rangle_{\text{free}}
\end{aligned} \tag{7.31}$$

where $\langle\rangle_{\text{free}}$ denote the averages for free particles. We thus obtained the remarkable results that with our approximation for the density matrix the kinetic energy is identical to the kinetic energy of free particles but with a renormalized amplitude. For the spin up it is multiplied by a factor

$$q_\uparrow = \frac{C}{n_\uparrow(1-n_\uparrow)} = \frac{\left[\sqrt{(1-n+d)(n_\uparrow - d)} + \sqrt{d(n_\downarrow - d)}\right]^2}{n_\uparrow(1-n_\uparrow)} \tag{7.32}$$

and of course a similar expression for the spin down, in which the roles of $n_\uparrow$ and $d_\downarrow$ have been exchanged. It is easy to check that if one sets in the above expression $d = n_\uparrow n_\downarrow$ which would be the case for free particles then one indeed recovers from (7.32) that $q_\uparrow = 1$, as it should since one should recover the kinetic energy of free particles.

7.2.3 Half-filling

The above expressions are quite general. Let us now specialize to the case of half filling ($n = 1$) and unpolarized systems $n_\uparrow = n_\downarrow$). In that case the expressions simplify considerably and one has

$$q_\uparrow = q_\downarrow = q = 8d(1 - 2d) \tag{7.33}$$

The parameters $n_\uparrow = n_\downarrow = 1/2$ are fixed. In order to determine the last unknown parameter, namely the fraction of doubly occupied sites d, we will minimize the total energy of the system with respect to d. We thus see that the formulation with the density matrix is a variational calculation similar to the Gutzwiller approach. Instead of using the factor g in the wavefunction (7.14) as a variational parameter, one uses directly the number of doubly occupied sites. Of course if $d > 0$ the system will stay metallic while a state with $d = 0$ denotes an insulator.

The total energy of the system is given by

$$\langle H\rangle = q\langle H_{\text{kin}}\rangle_0 + UNd \tag{7.34}$$

An estimate of the kinetic energy for free particles was done in (7.5). For half filling this crude estimate gives

$$\langle H_{\text{kin}}\rangle_0 = -NW/2 = -N\epsilon_0 \tag{7.35}$$

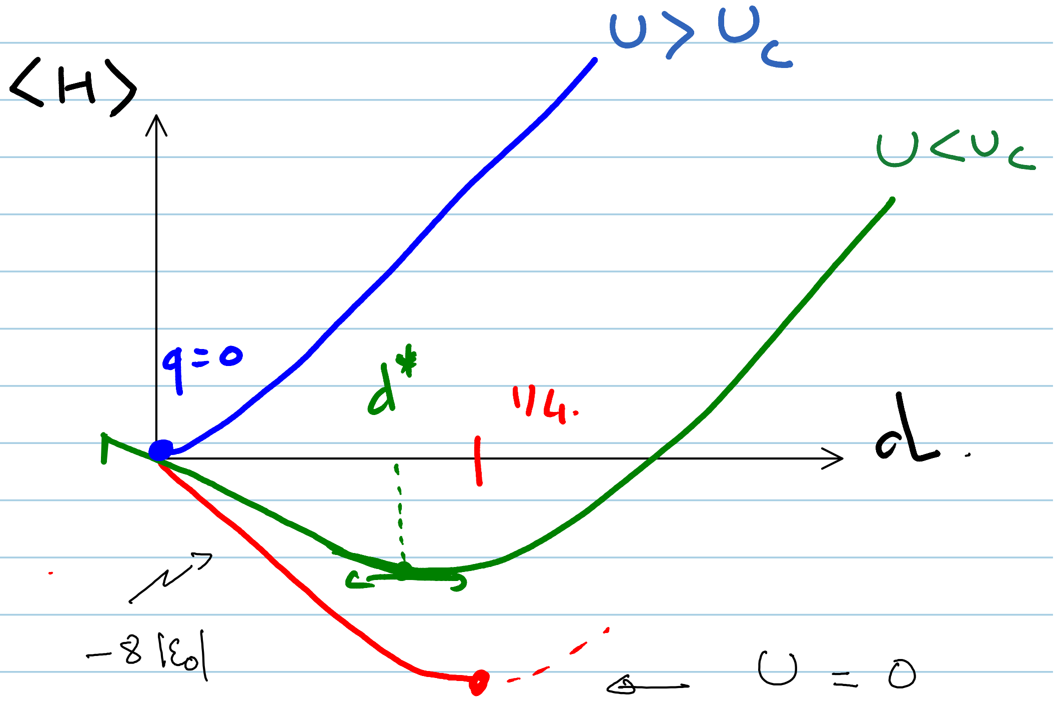


Figure 7.7: The total energy of an interacting system as a function of the fraction d of doubly occupied sites. Depending on the value of U a minimum of the energy exists for $d > 0$, or above $U = U_c$ the minimum only exists for $q = 0$. This signals a metal-insulator transition. With this particular approximation scheme the transition is continuous.

where W is the half bandwidth. We will use quite generally ϵ_0 for this gain in kinetic energy coming from the delocalization of the particles. ϵ_0 is of the order of the bandwidth of the system (quarter bandwidth with the above approximations)

Thus the energy per site is given by

$$\langle H \rangle / N = -8d(1-2d)\epsilon_0 + Ud \tag{7.36}$$

The corresponding curve is shown in Fig. 7.7. For a non interacting system $U = 0$, the minimum of energy is located at $d = 1/4$ and one recovers the expected result for noninteracting particles. For finite U two cases must be distinguished: i) if $U < U_c$ the interaction term is shifting the minimum to a smaller value of d but $d > 0$. One thus has a system in which the number of doubly occupied sites is finite. This is a metallic state since in that case the factor $q > 0$, and kinetic energy remains; ii) if $U > U_c$ then the minimum occurs at $d = 0$. In that case one has a localized wavefunction where one has exactly one particle per site. The factor $q = 0$ and there is no kinetic energy, the system is an insulator. Our Gutzwiller approximation scheme thus gives a metal-insulator transition, known as the Brinkmann-Rice transition. Within this approximation scheme the transition is continuous and the number of doubly occupied sites d can be used as an order parameter, in the same way than for example the magnetization would be an order parameter of the ferromagnetic transition: d is zero in the insulating region and finite in the metallic one.

The critical value U_c can be determined by looking at the slope of the energy around $d = 0$ in

(). One finds

$$U = 8\epsilon_0 \tag{7.37}$$

as expected the metal-insulator transition occurs when the value of the interactions is of the order of the bandwidth of the system. It is interesting to note several points concerning the above results:

1. We have thus a convincing study showing that even if one takes into account interactions in the metallic state, above a certain critical value of the interaction U_c and for a filling of one particle per site the system will turn into an insulator. This shows that interaction will indeed lead to Mott insulators, and put serious flesh on the bones of our first oversimplified comparison. The physics is however essentially the same.

2. The mechanism does not rely in any essential way to the presence or not of a magnetic order, but based on the competition between kinetic energy and the repulsion energy. This is in agreement with the experimental result. The presence of a critical value U_c seems also to be compatible with several experimental results. We will comment more on the connection between this solution of the Mott transition towards the end of this chapter.

3. This corresponding transition is different from the usual transitions. In particular this ia a transition that occurs at $T = 0$ upon variation of a parameter of the system (for example the interaction). It thus does not result from an energy-entropy competition as usual phase transitions in classical systems, but directly from an energy-energy competition. This type of transition is called a quantum phase transition to distinguish it from the "classical" phase transition. We will come back on such transitions in the next chapter.

Using our approximate density matrix we can now compute the physical properties. One example is the momentum distribution

$$\begin{aligned}
n_\uparrow(k) = \langle c^\dagger_\uparrow(k) c_\uparrow(k)\rangle &= \frac{1}{\Omega}\sum_{i,j} e^{ik(r_i-r_j)}\langle c^\dagger_{i\uparrow} c_{j\uparrow}\rangle \\
&= \frac{1}{\Omega}\sum_{i\neq j} e^{ik(r_i-r_j)}\langle c^\dagger_{i\uparrow} c_{j\uparrow}\rangle + \frac{1}{\Omega}\sum_{i}\langle c^\dagger_{i\uparrow} c_{i\uparrow}\rangle \\
&= \frac{1}{\Omega}\sum_{i\neq j} e^{ik(r_i-r_j)}\langle c^\dagger_{i\uparrow} c_{j\uparrow}\rangle + n_\uparrow \\
&= \frac{q_\uparrow}{\Omega}\sum_{i\neq j} e^{ik(r_i-r_j)}\langle c^\dagger_{i\uparrow} c_{j\uparrow}\rangle_0 + n_\uparrow \\
&= \frac{q_\uparrow}{\Omega}\sum_{i,j} e^{ik(r_i-r_j)}\langle c^\dagger_{i\uparrow} c_{j\uparrow}\rangle_0 + n_\uparrow(1-q_\uparrow) \\
&= q_\uparrow n^0_\uparrow(k) + n_\uparrow(1-q_\uparrow)
\end{aligned} \tag{7.38}$$

where $n^0_\uparrow(k)$ is the momentum distribution of free particle, i.e. at $T = 0$ the Fermi step. Since everything is independent of the spin orientation the spin down distribution is identical. We thus see that below k_{F} one has for $T = 0$ (recall that $n_\uparrow = 1/2$):

$$\begin{aligned}
n_\uparrow(k < k_{\mathrm{F}}) &= 1/2 + q_\uparrow/2 \\
n_\uparrow(k > k_{\mathrm{F}}) &= 1/2 - q_\uparrow/2
\end{aligned} \tag{7.39}$$

As shown in Fig. 7.8 the discontinuity on the Fermi surface is now $q_\uparrow$. We thus see that although the system remains a metal, it feels strongly the effects of the interactions and become more and more correlated. Remember from (4.69) that the jump is directly related to the quasiparticle

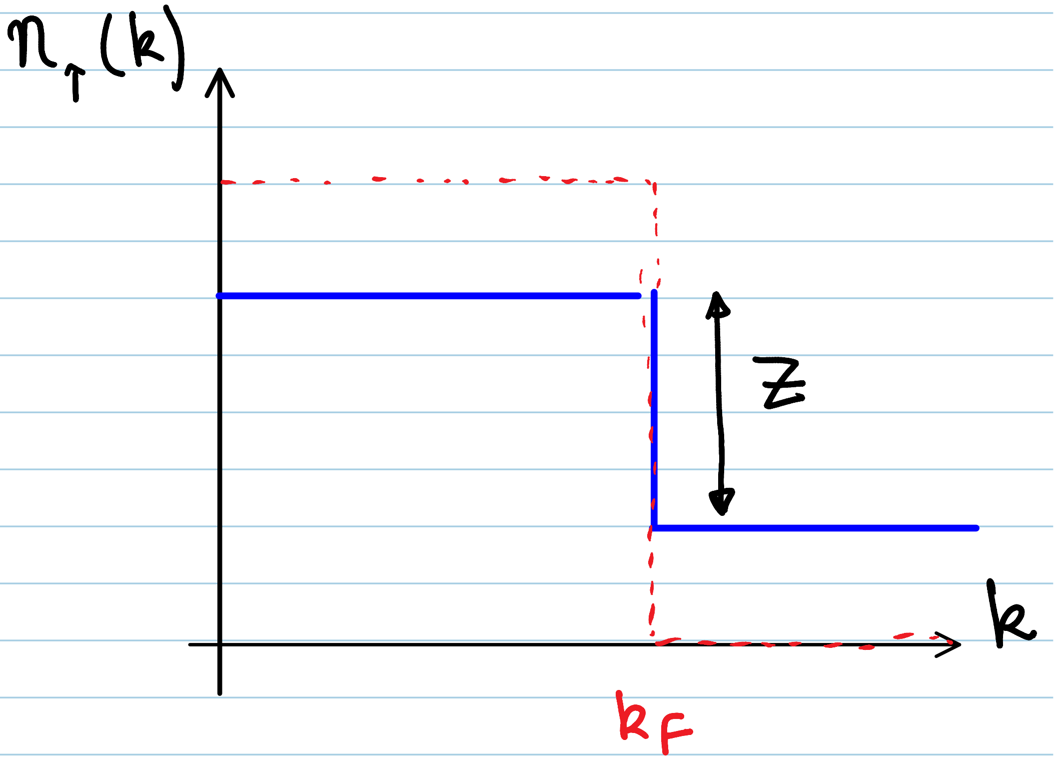


Figure 7.8: Momentum distribution for one spin species in the Brinkmann-Rice approximation. The jump at $k = k_\mathrm{F}$ gets smaller and smaller leading to a smaller quasiparticle residue until the Metal-Insulator transition is reached. As the result the metal becomes more and more correlated.

residue in the Fermi liquid One has thus $Z_\uparrow = q_\uparrow$. Since $q_\uparrow \to 0$ at the metal insulator transition we see that the metal becomes a less and less good Fermi liquid when approaching the critical point. We have not examined here the momentum dependence of the self-energy but the relation (4.63) which would be valid if the self energy is reasonably momentum independent strongly suggests that the effective mass will diverge at the transition as $1/q_\uparrow$.

In the insulator, one has localized particles on each site. If one tries to create and excitation by putting two particles on each site, then there is a finite energy price to pay. For very large repulsion, this difference of energy between the ground state and the first excited state is $\sim U$. When one approaches the transition it is not U but a certain function of U that we would call Δ. This quantity is called the Mott gap. In the previous chapter using our mean-field theory we had computed such a quantity. The fact that in the insulator there is a finite energy difference between the ground state and the first excited state, and in particular that adding a particle costs an energy Δ means that to add one particle one has to change the chemical potential μ by a finite amount. This implies that

$$\frac{dN}{d\mu} = 0 \tag{7.40}$$

and the system is incompressible. We will summarize the properties in Section 7.2.5.

7.2.4 Doped case

Let us now look what would happen if instead of being exactly at half filling we were slightly doped, i.e. if $n = 1 + \delta$ with $\delta \ll 1$. An interesting question is of course whether such a state would be a metal or an insulator. If $U < U_c$ we can readily guess that the system is very likely to stay a metal and thus a little bit of charges is not doing much.

One would thus have a "normal" albeit strongly correlated Fermi liquid, where the total number of charges is n. In that regime there is not much difference between $n = 1$ and $n = 1 + \delta$. The situation is very different for $U > U_c$ in that case we know that for $n = 1$ the system is an insulator. In order to get an idea of the physics let us evaluate the factor (7.32) for $n = 1\delta$ assuming that U is large enough so that $d = 1$. One gets for $\delta \ll 1$

$$q_\uparrow = \frac{(1-n)n_\uparrow}{n_\uparrow(1-n_\uparrow)} \simeq 2\delta \tag{7.41}$$

This results immediately suggests that even for $U > U_c$ the system would remain a *metal.* But we see that the kinetic energy would be strongly reduced and correspond only to the excess part of particles (or holes depending on the sign) δ. One can push the calculation to the first order in d and get

$$q_\uparrow = \frac{2\delta}{1+\delta} + \frac{4\sqrt{\delta}}{1+\delta}\sqrt{d} \tag{7.42}$$

Minimizing the total energy (7.34) leads to

$$d^* = \frac{4\delta\epsilon_0^2}{U^2} \tag{7.43}$$

Thus there is a (very) small amount of doubly occupied states, but since $d^* \sim 1/U^2$ it means that at large U the contribution of the interaction energy Ud^* in the total energy tends to zero. In the same way the contribution of the second term in (7.42) tends to zero. For large U we can thus see the system as a system of free particles where *only* the excess particles δ exist and are behaving essentially as free objects. This is a very important result, since it shows that a doped Mott insulator will behave as a metal, but with a number of charges that is fixed by δ and *not* the total number of particles. This is something that is directly testable experimentally by using for example a Hall measurement which is giving access to the number of carriers (although one has to keep in mind that Hall effect in a correlated system might be more complex).

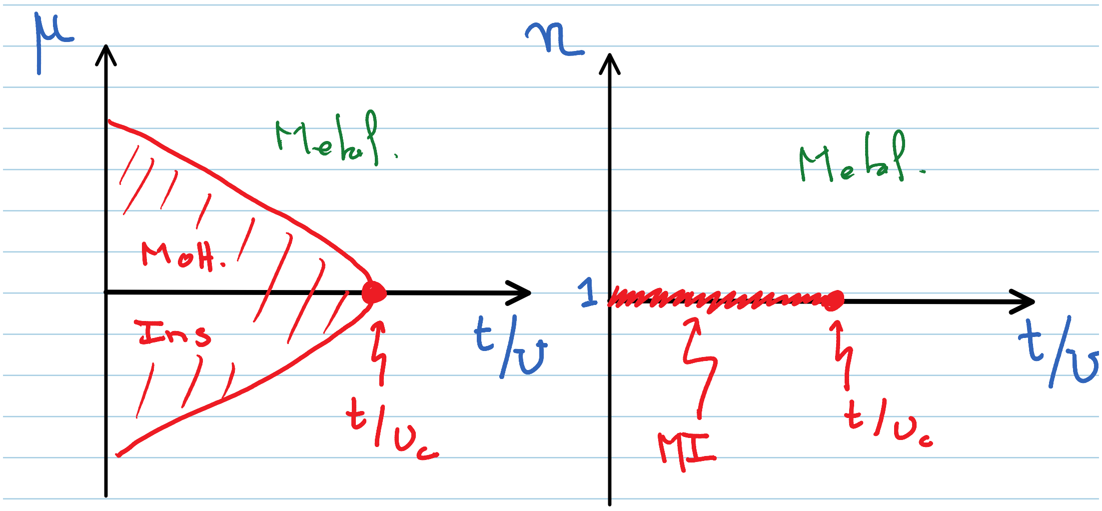


Figure 7.9: Phase diagram as a function of ratio of hopping to interactions t/U, and chemical potential (left) and number of particles (right). In the Mott insulator phase, there is a gap Δ to excitations.

7.2.5 Summary of the properties

Let us summarize here the properties of Mott insulators We show in Fig. 7.9 the phase diagram of the system both as a function of the chemical potential and as a function of the number of particles. In the metallic phase, there is no gap to excitations it is thus possible to change the number of particles by an infinitesimal variation of the chemical potential. On the contrary in the Mott phase there is a gap Δ towards excitations and the system is incompressible. The relation between the chemical potential μ and the number of particles n is shown in Fig. 7.10.

Based on the above it is important to note that there are two very different type of transitions that are linked to the Mott phenomenon and that are in two different universality classes.

1. The transition that occurs at commensurate filling $n = 1$, upon increasing the interactions or reducing the kinetic energy.
2. The transition that occurs for $U > U_c$ when going from a state with $n = 1$ to a state with a different number of particles $n = 1 + \delta$

Finally let us comment on the connection between the result that we have obtained in this chapter and the ones we got using the mean-field theory on the Hubbard model in Section 6.5. The two have of course many similarities and must be connected. In the previous study we also find that for $n = 1$ we had an insulator, but for any value of U. However the gap was exponentially small for small U while it was $\sim U$ for large U. It thus seems that at large U the two studies are essentially the same, In addition upon doping our system was also becoming a metal again, just as we have found there. The two gaps are represented in Fig. 7.11. In fact the case we considered in this chapter is the generic case, The Mott transition is driven by the charge fluctuations and does not care in any essential way on what the spin degrees of freedom are doing. In that case we have a critical value of U to get the Mott insulating state. The insulator can or cannot acquire magnetic properties, and the question of the magnetic features of the Mott insulator is something we will address in the next chapter. On the other hand on some special lattices, which lead to a nesting of the Fermi surface, we know from the

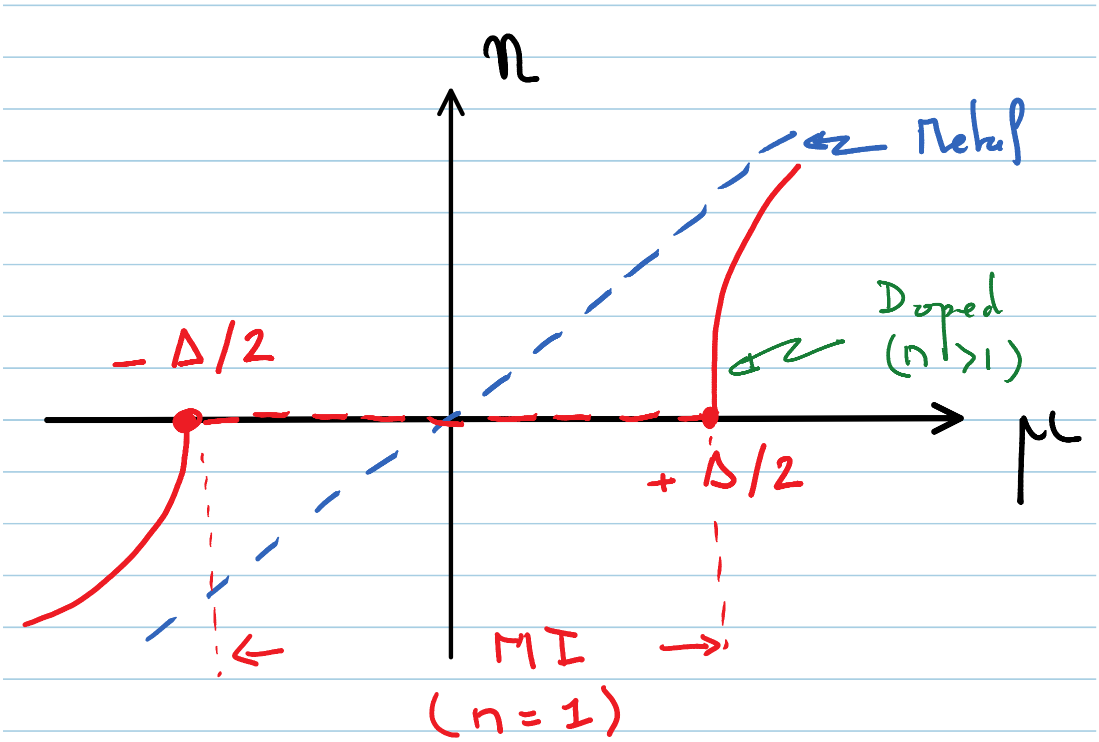


Figure 7.10: Relation between the number of particles n and the chemical potential μ. In the metal (dashed line) there is no gap to excitations and the system is compressible. In the Mott insulator there is a finite gap Δ to excitations so the number of particles does not change until the chemical potential exceeds the gap, and the system is incompressible.

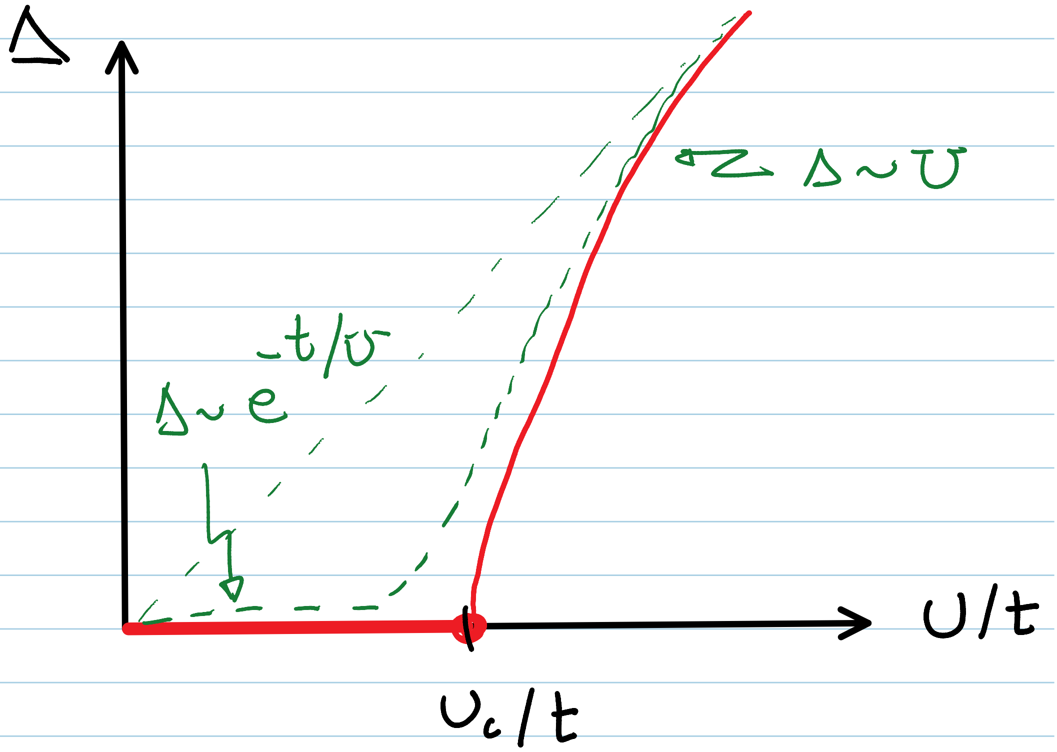


Figure 7.11: Schematic of the gap as a function of the interaction U. For a generic system (full line), one needs a finite value of the interaction to open the Mott gap. On some special lattices (dashed line), the nesting of the Fermi surface allows the immediate appearance of magnetic order and thus the appearance of a Mott gap even for small U.

results of the previous chapter that the spin susceptibility at the nesting vector diverges. If this vector corresponds to a half-filled Fermi surface there will thus be an instability even for infinitesimal U. In that case the appearance of the gap is directly linked to this magnetic (in general antiferromagnetic) instability. This in these cases (for example for the square, cubic etc.) lattices, one gets an antiferromagnetic Mott insulator immediately. As U increases charge is more and more frozen independently of the spin and at large U one recovers the same physics than the one we discussed in this chapter.

CHAPTER 8

Localized magnetism

8.1 Basic Ideas

In the previous chapter we have seen that a new form of insulators, driven by the interactions, could exist: the Mott insulators. The existence of such systems opens the way to new magnetic phases. Indeed if we consider band insulators, they occur when the band is either empty or totally filled, which means that there are two electrons per quantum state of the band. It is thus obvious that such an insulator has no magnetic properties as well. For an empty band this is trivial, and for a filled band because there are two electrons in each one of the quantum number, the Pauli principle which imposes that the wavefunction is totally antisymmetric imposes thus to put the two electrons in an antisymmetric state of spin. The corresponding state is thus the singlet state

$$c^\dagger_{k,\uparrow} c^\dagger_{k,\downarrow} |\varnothing\rangle = \frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \tag{8.1}$$

A singlet state corresponds to $S = 0$ and thus to an absence of magnetic properties. Band insulators have thus both no charge properties and spin properties.

The situation is quite different for Mott insulators. Indeed in that case there is one electron per site which is localized. It means that on each site there is potentially a spin $1/2$ that remains. Thus although a Mott insulator has no charge excitations (below the Mott gap Δ_M), there remains the possibility of highly not trivial *magnetic* excitations at low energy since on each site the spin can be up or down. There is thus a very large Hilbert space of low energy magnetic excitations of size 2^N. It is thus important to understand the processes that control the physics of these magnetic excitations. This field, in which the charge degrees of freedom are essentially blocked and only the localized magnetic remains is an important branch of solid state physics and goes under the name of localized magnetism.

8.1.1 Reminder of spin properties

Before we determine the Hamiltonian of the system in the low energy sector as shown in Fig. 8.1 let us remind some of the basics of spin 1/2.

On each site the system is described by a spin 1/2. This corresponds to a magnetic moment given by

$$M^\alpha = g\mu_B S^\alpha = g\mu_B \frac{1}{2}\sigma^\alpha \tag{8.2}$$

where $\alpha = x, y, z$ are the three components of the moment and spin, g is the gyromagnetic factor, μ_B the Bohr magneton, and the σ^α are the three Pauli matrices.

$$S^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, S^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, S^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{8.3}$$

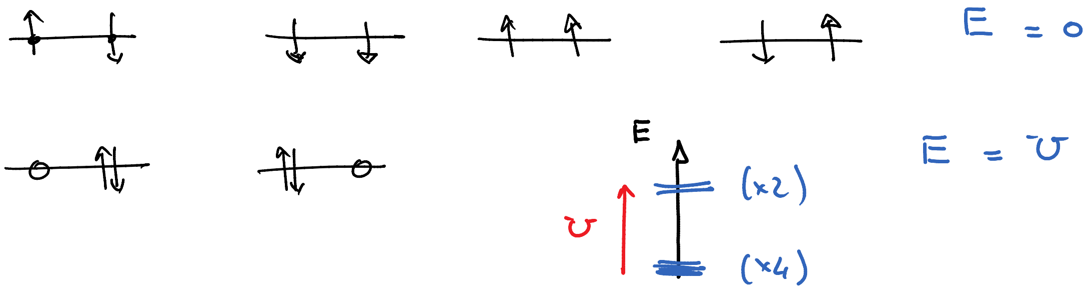


Figure 8.1: In a Mott insulator the charge excitations need a finite energy of the order of the Mott gap Δ. On the contrary since there is on each site a residual spin 1/2 there is a low energy sector of spin excitations corresponding to a Hilbert space of size 2^N. The degeneracy that would occur if each particle was totally localized on a site is lifted by the superexchange (see text).

The three component of the spin obey the commutation relations

$$[S^\alpha, S^\beta] = i\epsilon_{\alpha\beta\gamma} S^\gamma \tag{8.4}$$

where $\epsilon_{\alpha\beta\gamma}$ is the fully antisymmetric tensor (i.e. $[S^x, S^y] = iS^z$ and circular permutations).

A complete basis of the Hilbert space is provided by the two eigenstates of one of the spin components, for example the one of the S^z operator $|\uparrow\rangle$ and $|\downarrow\rangle$.

It is convenient to introduce the raising and lowering spin operators, which are hermitian conjugate operators defined by

$$S^\pm = S^x \pm iS^y \tag{8.5}$$

These operators are raising or lowering the spin. For example

$$S^+ |\uparrow\rangle = 0 \quad , \quad S^+ |\downarrow\rangle = |\uparrow\rangle \tag{8.6}$$

These operators also obey

$$[S^+, S^-] = 2S^z \tag{8.7}$$

as can be computed using the commutation relations (8.4).

8.1.2 Superexchange

As shown in Fig. 8.1 we have, in a Mott insulator, a spin 1/2 on each site of a lattice. If the wavefunction of each electron is fully localized on a site, each spin is independent and one has to face a huge degeneracy of the ground state since each spin can be either up or down. It is thus important to understand the properties of such a system and in particular the interactions that could exist between the spins.

The most naive interacting that could be taken into account is the direct magnetic moment interaction between the magnetic moments located on the different sites. This is the standard dipolar interaction

$$H_D = \frac{\mu_0}{4\pi}\left[\frac{\boldsymbol{M}_1 \cdot \boldsymbol{M}_2}{r^3} - \frac{3(\boldsymbol{M}_1 \cdot \boldsymbol{r})(\boldsymbol{M}_2 \cdot \boldsymbol{r})}{r^5}\right] \tag{8.8}$$

This interaction would lead to a an (anisotropic) interaction between the spins. However it is easy to see by putting for r the typical interatomic distance that the typical value of such

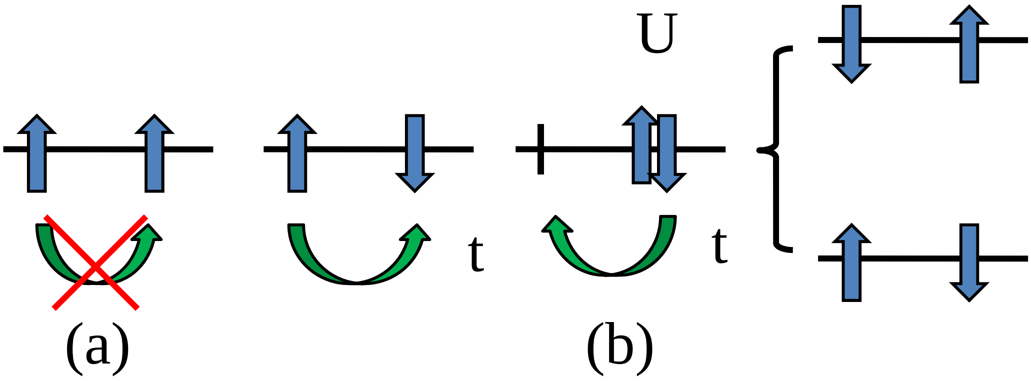


Figure 8.2: (a) The hopping between two neighboring identical spins is totally blocked by the Pauli principle. The system cannot gain some residual kinetic energy in this case. (b) If the spins are antiparallel then the kinetic energy can allow virtual hopping through an intermediate state of energy U. This state comes back either to the original state or to a configuration in which the spins have been exchanged. This is the mechanism of superexchange leading to dominant antiferromagnetic correlations for fermionic Mott insulators.

interaction would correspond to energies of the order to less than a Kelvin. Such interaction, although present, would not explain the various magnetic properties of Mott insulators that are observed at temperatures of the order of hundredth of Kelvins.

Another interaction must therefore exist. As we saw for the case of the itinerant ferromagnetism in Section 6.2.1 the source of such interaction between the spins will have its roots in the competition between the Pauli principle, kinetic energy and interaction. This is a mechanism called superexchange.

In order to describe such an interaction, let us examine the case of two sites. The total Hilbert space is

$$|\uparrow,\downarrow\rangle\,,|\downarrow,\uparrow\rangle\,,|\uparrow\downarrow,0\rangle\,,|0,\uparrow\downarrow\rangle\,,|\uparrow,\uparrow\rangle\,,|\downarrow,\downarrow\rangle \tag{8.9}$$

Since the states are composed of Fermions one should be careful with the order of operators to avoid minus signs. Let us take the convention that

$$\begin{aligned} |\uparrow,\downarrow\rangle = c^\dagger_{1\uparrow}c^\dagger_{2\downarrow}\,|\varnothing\rangle \quad &, \quad |\downarrow\uparrow\rangle = c^\dagger_{1\downarrow}c^\dagger_{2\uparrow}\,|\varnothing\rangle \\ |\uparrow\downarrow,0\rangle = c^\dagger_{1\uparrow}c^\dagger_{1\downarrow}\,|\varnothing\rangle \quad &, \quad |0,\uparrow\downarrow\rangle = c^\dagger_{2\uparrow}c^\dagger_{2\downarrow}\,|\varnothing\rangle \end{aligned} \tag{8.10}$$

The states with two particles per site are states of energy $\sim U$ and therefore strongly suppressed. We thus need to find what is the form of the Hamiltonian when restricted to the states with only one particle per site. It is easy to check that the two states $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ are eigenstates of H

$$H\,|\uparrow\uparrow\rangle = 0 \tag{8.11}$$

and a similar equation for $|\downarrow\downarrow\rangle$. The reason why the kinetic energy does not act on such a state is shown in Fig. 8.2.

The Pauli principle block the hopping if the two spins are equal. On the contrary, if the two spins are opposite the particles can make a virtual jump on the neighboring site. Since this state is of high energy U the particles must come back to the original position, or the two particles can exchange leading to a term similar to a spin exchange (see Fig. 8.2). The *kinetic energy* thus leads to a *magnetic* exchange, named superexchange, that will clearly favor configurations with opposite neighboring spins, namely will be antiferromagnetic.

Let us now quantify the mechanism. The Hamiltonian can be written in the basis (8.9) (only the action on the first four states is shown since $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ are eigenstates and thus are uncoupled to the other ones)

$$H = \begin{pmatrix} 0 & 0 & -t & -t \\ 0 & 0 & t & t \\ -t & t & U & 0 \\ -t & t & 0 & U \end{pmatrix} \tag{8.12}$$

This Hamiltonian couples the low energy states with one particle per site to high energy states of energy U. We now need to find the restriction of this Hamiltonian to the low energy sector that does not involve excitations of order U. Since the ground state is degenerate we cannot use the usual perturbation theory formulas. Let us thus adapt them to our case.

In order to find the restriction of H to the low energy sector, let us first make a canonical transformation of H

$$H' = e^{iS} H e^{-iS} \simeq H + i[S, H] + \frac{i^2}{2}[S, [S, H]] + \cdots \tag{8.13}$$

where the matrix S is expected to be perturbative in t/U. Such a canonical transformation, being unitary will not affect the energies of the system but merely rotate the vectors of the system by

$$|\psi\rangle \to e^{iS} |\psi\rangle \tag{8.14}$$

In this transformation one wants to chose the matrix S such that

$$H_t + i[S, H_U] = 0 \tag{8.15}$$

This will ensure that H' has no elements connecting the low energy sector with the sector of energy U. The restriction of H' to the low energy sector will thus be the Hamiltonian we need to diagonalize to find the spin properties.

Using the condition (8.15) and keeping only terms up to order t^2/U it is easy to check that

$$H' = H_U + \frac{i}{2}[S, H_t] \tag{8.16}$$

Since H is block diagonal one can easily determine S to be

$$S = \frac{i}{U} \begin{pmatrix} 0 & 0 & -t & -t \\ 0 & 0 & t & t \\ t & -t & 0 & 0 \\ t & -t & 0 & 0 \end{pmatrix} \tag{8.17}$$

This leads using (8.16) to

$$H' = \frac{4t^2}{U}\left[\frac{1}{2}[|\uparrow\downarrow\rangle\langle\downarrow\uparrow| + |\downarrow\uparrow\rangle\langle\uparrow\downarrow|] - \frac{1}{2}[|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow|]\right] \tag{8.18}$$

This Hamiltonian must be complemented by the fact that $H'|\uparrow\uparrow\rangle = 0$. Since there is one particle per site, it can now be represented by a spin 1/2 operator as defined in the Section 8.1.1.

One obtains in terms of the spin operators

$$H' = \frac{4t^2}{U}\left[\frac{1}{2}[S_1^+ S_2^- + S_1^- S_2^+] + S_1^z S_2^z - \frac{1}{4}\right] \tag{8.19}$$

Up to a constant which is a simple shift of the origin of energies, the final Hamiltonian is thus the Heisenberg Hamiltonian

$$H = J \sum_{\langle ij \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j \tag{8.20}$$

where the magnetic exchange is $J = 4t^2/U$. Note that the exchange J is positive and thus the system has dominant antiferromagnetic correlations.

The formula (8.20) shows that the combination of Pauli principle, kinetic energy and interaction lead to an effective interaction between the spins that take the form of a spin spin interaction. Since the typical values of t are $\sim 1eV$ while in a Mott insulator $U \sim 10eV$ one can expect values for the (super)-exchange $J \sim 0.4eV$ which corresponds to about $J \sim 5000K$. Such value of the exchange can easily explain the strong antiferromagnetic tendency that is found in many of the Mott insulators.

Of course the final properties of the system depend on the lattice and thus the Hamiltonian (8.20) can potentially lead to very rich physics [Aue98]. It is important to note that we are dealing here with quantum spins for which $[S_x, S_y] = iS_z$ and thus the three component of the spins cannot be determined simultaneously. Quantum fluctuation will thus drastically affect the possible spin orders. Depending on the lattice various ground states are possible ranging from spin liquids to ordered states. This is the physics that we will examine in the coming sections.

8.2 Dimer systems

Let us start by investigating the simplest quantum system, made of two spins. In a way this is also one of the ones on which the quantum effects will manifest themselves in the strongest way and thus will serve to clearly illustrate the differences that can exist between classical and quantum spins.

Let us consider a system made of two spins. In that case the Heisenberg Hamiltonian (8.20) simply reduces to

$$H = J[S_1^x S_2^x + S_1^y S_2^y + S_1^z S_2^z] \tag{8.21}$$

Let us first consider the classical version of such a system. In that case it is easy to see that if the exchange J is positive then the lowest energy of obtained if the two spins point in opposite direction. In that case the energy is $E = -J/4$. Note also that if the system has a ferromagnetic exchange one can go directly from the antiferromagnetic ground state to the ferromagnetic one by simply reversing one of the two spins

$$S_2^\alpha \to -S_2^\alpha \tag{8.22}$$

and the energy is still $E = -|J|/4$. For a classical system of dimer there is thus no essential difference between the solution of the ferromagnetic ground state and the antiferromagnetic one, since one can simply go from one to the other by reversing one of the spins.

8.2.1 Ground state

The situation changes drastically when one as to consider quantum spins. Let us first look at the ferromagnetic case which is the simplest. In that case $J < 0$. Note that in the previous section we showed that for fermions in the simple case of one particle per site we obtained the opposite sign for the exchange. We will thus not dwell here on which type of systems one could get a ferromagnetic J but simply consider it as a mathematical case. There are however systems in which the intra-atomic interactions and the orbital structure is such that the superexchange mechanism can also lead to ferromagnetic exchange, thus corresponding to physical situations. In that case we have to diagonalize the Hamiltonian in the complete Hilbert space of two spins, which contains now four states

$$|\uparrow\uparrow\rangle\,,|\uparrow\downarrow\rangle\;|\downarrow\uparrow\rangle\;|\downarrow\downarrow\rangle \tag{8.23}$$

It is easy to check that for $J < 0$ the two states in which the spins are parallel is an eigenstate of the problem with the energy $-|J|/4$. To do so let us rewrite the Hamiltonian as

$$H = \frac{J}{2}[S_1^+ S_2^- + S_1^- S_2^+] + J S_1^z S_2^z \tag{8.24}$$

and it is easy now to check that

$$H\,|\uparrow\uparrow\rangle = -|J|/4\,|\uparrow\uparrow\rangle \tag{8.25}$$

Thus if the exchange is ferromagnetic the quantum solution and the classical one look very much alike and lead to the same energy and also the same naive representation of the ground state in which the spins point in parallel direction.

Note however that one cannot go as naively from the ferromagnetic solution to the antiferromagnetic one. Indeed the transformation (8.22) is forbidden since it would lead to a violation of the canonical commutation relation of the spins (8.4). It would be only possible to change *two* of the spin coordinates into their opposite but not the three of them. One has thus to examine the antiferromagnetic case separately.

In order to understand the antiferromagnetic side, let us perform a full diagonalization of the Hamiltonian (8.21). This will serve to illustrate some general principles. Since the full Hilbert space is of size 4 one would have normally to diagonalize a 4 × 4 matrix. But doing so would be totally inefficient and would not use the symmetries of the problem. Indeed one can use the fact that if we find an operator that commutes with the Hamiltonian then we can diagonalize both this operator and the Hamiltonian in the same basis. In other words the Hamiltonian will be block diagonal in the basis of eigenstates of this operator. The Hamiltonian (8.21) is clearly invariant by spin rotation. We can thus use the fact that the total spin $S = S_1 + S_2$ of the system is conserved. This means in particular that the z component of this operator is a conserved quantity. One can also use the fact that S^2 is also conserved.

Let us start with the second quantity. The total spin resulting from the addition of two spin 1/2 can take two values $S = 0$ and $S = 1$. The first case $S = 0$ the singlet has only one possible eigenvector, and thus this eigenvector must also be an eigenvector of the total Hamiltonian. The singlet is

$$|S = 0, S_z = 0\rangle = \frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \tag{8.26}$$

In order to find the energy of this state let us apply the Hamiltonian (8.21) or more simply (8.24). One obtains

$$H\,|S = 0, S_z = 0\rangle = -\frac{3J}{4}\,|S = 0, S_z = 0\rangle \tag{8.27}$$

This confirms that the singlet is indeed an eigenstate of H. The triplet state corresponds to $S = 1$. Such state can have three values of the component S_z of the spin leading to three different states $|S = 1, S_z = 1, 0, -1\rangle$. Since S_z is also conserved each one of this state is also an eigenstate of H since they are the only one with these quantum numbers. The corresponding eigenvalues is, applying again the Hamiltonian H

$$H\,|S = 1, S_z = 1, 0, -1\rangle\;\frac{J}{4}\,|S = 1, S_z = 1, 0, -1\rangle \tag{8.28}$$

This leads to the eigenstates and eigenenergies shown in Fig. 8.3 Note that even without having the conservation of the total spin, the simple conservation of the S^z component would simplify

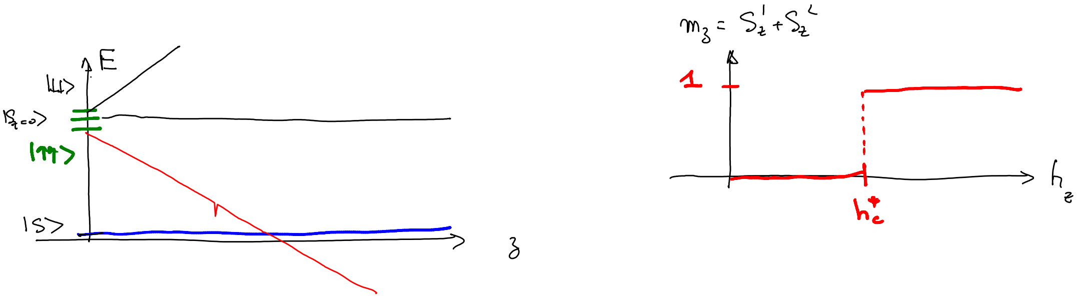


Figure 8.3: The energies and eigenvectors of the dimer Hamiltonian. The ground state is a singlet, and there is a triply degenerate excited state of $S = 1$ triplet. (b) When a magnetic field is applied the triplet splits. The ground state remains non magnetic until a critical field is reached, at which point the system goes to the triplet state. The magnetization jumps from $m = 0$ to $m = 1$ at that field.

considerably the diagonalization of the Hamiltonian. Indeed one can immediately identify the two unique states of $S_z = \pm 1$ as $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$, which are thus eigenstates of the Hamiltonian. This simply leaves the two states with $S_z = 0$, $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$. One has thus a single 2×2 matrix to diagonalize to find the two eigenstates ($S = 0$ and $S = 1$, $S_z = 0$).

The above results for the antiferromagnetic case are quite remarkable and, contrarily to the ferromagnetic case, quite different from the classical ones. One sees that the antiferromagnetic ground state has no real classical equivalent. It is made from a quantum superposition of the two states with opposite spins $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$, with the proper phase. The net result is a state without any magnetic moment $S = 0$. Quite surprisingly the energy of the resulting quantum state is much lower ($E = -3J/4$) than the one one would have expected from the corresponding classical state $E_c = -J/4$ obtained by putting the spins in opposite direction along one given direction. It is more or less as if the spin could be opposite along the three directions simultaneously leading to a further lowering of the energy. It is also important to note that the first excitations above this non-magnetic singlet state are separated by a finite energy gap $\Delta = J$.

8.2.2 Magnetic field

The difference between the quantum spin system an classical ones is even more apparent when one applies a magnetic field to the system. Let us add to (8.21) a term

$$H_z = -g\mu_B h_z(S_1^z + S_2^z) = -g\mu_B h_z S^z \tag{8.29}$$

Such a term corresponds to the coupling of an external magnetic field h_z along a direction that we will choose as the z axis to the two magnetic moments that are corresponding to the two spins. Since each one of the eigenstates found in the previous section is also an eigenstate of S^z it is very easy to take this term into account. The effects are shown in Fig. 8.3. The singlet state, which corresponds to $S^z = 0$ is totally insensitive to the magnetic field and its energy does not change. The three state of the triplet corresponding respectively to $S^z = 1, 0, -1$ see their energy shifted by the term (8.29) in which S^z has been replaced by the corresponding eigenvalue. The resulting energy diagram is shown in Fig. 8.3. As a result the ground state remains the singlet one up to a critical field which corresponds to the energy gap between the

Figure 8.4: Two classical antiferromagnetically coupled spins orient perpendicularly to an externally applied magnetic field. This allows to cant the spins by an angle θ to generate a ferromagnetic component while retaining most of the antiferromagnetic coupling. The angle varies continuously as a function of the external magnetic field. This is at strong variance with the quantum solution in which the non-magnetic ground state remains totally insensitive to the magnetic field up to a critical field.

singlet and triplet state.

$$J = g\mu_b h_z^c \tag{8.30}$$

Above this critical field $h_z > h_z^c$ the ground state becomes the lowest triplet $|\uparrow\uparrow\rangle$. The magnetization along the z direction thus jumps discontinuously from $S^z = 0$ for $h_z < h_z^c$ to $S^z = 1$ for $h_z > h_z^c$.

Let us compare these results with the ones one would get for a classical antiferromagnetic dimer. In that case the two spins are opposite. The system being invariant by spin rotation the pair can point in any direction. If a magnetic field is present it is clear that pointing the pair in the magnetic field direction would be extremely unfavorable since in that case the magnetic field contribution would simply cancel between the two spins. It is much more favorable to put the pair essentially *perpendicular* to the magnetic field and tilt the two spins by an angle θ as indicated on Fig. 8.4 We thus have $S_1 = 1/2(\cos(\theta), \sin(\theta))$ and $S_1 = 1/2(-\cos(\theta), \sin(\theta))$, which corresponds to the energy

$$\begin{aligned} H &= \frac{J}{4}[\sin^2(\theta) - \cos^2(\theta)] - g\mu_B h_z \sin(\theta) \\ &= \frac{J}{4}[2\sin^2(\theta) - 1] - g\mu_B h_z \sin(\theta) \end{aligned} \tag{8.31}$$

The optimal angle is determined by minimizing the energy with respect to the angle:

$$J \sin(\theta) = g\mu_B h_z \tag{8.32}$$

Thus for $h_z < h_z^c$ there is a solution for which the angle θ goes from $\theta = 0$ to the full polarization $\theta = \pi/2$. The magnetization being proportional to $\sin(\theta)$ varies continuously with the applied magnetic field. At very small field

$$m = \frac{(g\mu_B)^2}{J} h_z \tag{8.33}$$

The classical system has a finite susceptibility to an external magnetic field. As can be expected the susceptibility is inversely proportional to the strength of the antiferromagnetic exchange but remains finite nevertheless. On the contrary the quantum problem has a *zero* susceptibility. As already indicated above this is true to *all* orders in the magnetic field. This is quite a remarkable result and difference between the classical and quantum behavior of the spins.

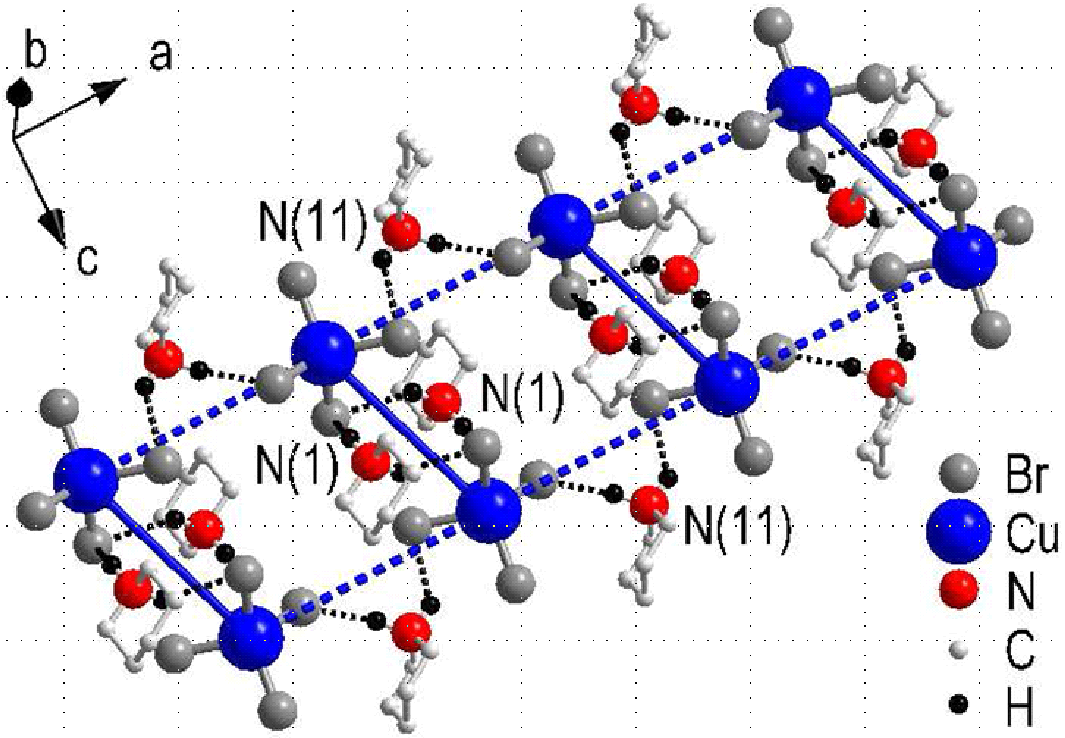


Figure 8.5: An example of a material made of dimers with a strong antiferromagnetic exchange $J_\perp$. These dimers are weakly coupled by another and weaker antiferromagnetic exchange J. In the compound shown $J_\perp \sim 13K$ while $J \sim 3K$. The magnetization is shown on the right. One note that there is indeed a critical field below which the magnetization remains strictly zero. Instead of the abrupt jump in magnetization the coupling between dimers split the jump in a continuous growth for $h_{c1} < h_z < h_{c2}$.

The properties of the dimer are one of the striking examples of quantum magnetism. Quite remarkable some compounds are a very good realization of such systems. An example is shown in Fig. 8.5 together with the corresponding magnetization curve. For more details on such systems we refer the reader to [GRT08].

8.3 Macroscopic number of spins

Although we could solve fully the case of the quantum dimer, if one has to deal with N coupled spins it is clear that one has to resort to more sophisticated methods to find the ground state and excitations of the system. Spins being neither fermions nor bosons, one cannot use all the the many body formulas that have been so useful is solving interacting quantum problems. It is thus very interesting to see if one could find ways to relate the spins to standard fermionic or bosonic "particles". We will see such methods in the present section and then use them to study a large number of spins coupled with ferromagnetic and then antiferromagnetic interactions. As we did with the dimer we will compare our results with the classical equivalents such as Ising or Heisenberg classical models.

8.3.1 Holstein-Primakoff representation

Spins have strange commutation relations since they commute on different sites, a little bit like bosons, but on the same site they have strange commutation relations which are neither bosonic nor fermionic. For a general angular moment (spin S) one has:

$$\begin{aligned} &[S_i^\alpha, S_j^\beta] = 0 \quad , \quad i \neq j \\ &[S_i^+, S_i^-] = 2S_i^z \end{aligned} \tag{8.34}$$

In order to avoid this problem, a very nice representation for the spins has been introduced by Holstein and Primakov [HP40] in term of bosons. For a spin S the representation goes as follows. The fully polarized state $|S_z = S\rangle$ is considered as the vacuum of bosons $|\varnothing\rangle$. Each decrease by one of the S_z value corresponds to the addition of a boson. It is thus easy to check that one must have essentially

$$\begin{aligned} S_i^+ &\to b_i \\ S_i^- &\to b_i^\dagger \\ S_i^z &= S - \mathcal{N}_i = S - b_i^\dagger b_i \end{aligned} \tag{8.35}$$

However such a representation would be incorrect. Indeed, the Hilbert space of the bosons is infinite while it is impossible for the S_z component to go below $-S$. It is thus imperative to add factors that will make sure that $S^- |{-S}\rangle = 0$. Let us check that the representation

$$S_i^+ \to \sqrt{2S - \mathcal{N}_i}\, b_i \tag{8.36}$$

$$S_i^- \to b_i^\dagger \sqrt{2S - \mathcal{N}_i} \tag{8.37}$$

$$S_i^z = S - \mathcal{N}_i = S - b_i^\dagger b_i \tag{8.38}$$

does indeed respect all commutation relations and all constraints on the limitations on the Hilbert space with b being bosonic operators.

It is clear from (8.36) and (8.37) that

$$\begin{aligned} S_i^+ |S_i^z = S\rangle &= \sqrt{2S - \mathcal{N}_i}\, b_i |\varnothing\rangle = 0 \\ S_i^- |S_i^z = -S\rangle &= b_i^\dagger \sqrt{2S - \mathcal{N}_i}\, |\mathcal{N}_i = 2S\rangle = 0 \end{aligned} \tag{8.39}$$

as one should for angular momentum operators. Thus if one starts from any state that belongs to the proper Hilbert space of angular momenta (i.e. $0 \leq n_b \leq 2S$) it is impossible with the representation (8.36-8.38) to leave this Hilbert space. One should thus now check that the proper commutation relations (8.34) are obeyed. Clearly the spins on different sites commute since the boson operators on different sites commute. Let us thus check only the second relation on the same site. One has (using the canonical commutations relations for boson operators):

$$\begin{aligned} S_i^+ S_i^- &= \sqrt{2S - \mathcal{N}_i}\, b_i b_i^\dagger \sqrt{2S - \mathcal{N}_i} \\ &= \sqrt{2S - \mathcal{N}_i}(1 + \mathcal{N}_i)\sqrt{2S - \mathcal{N}_i} \\ &= (2S - \mathcal{N}_i)(1 + \mathcal{N}_i) \end{aligned} \tag{8.40}$$

and

$$\begin{aligned} [S_i^+, S_i^-] &= (2S - \mathcal{N}_i)(1 + \mathcal{N}_i) - b_i^\dagger (2S - \mathcal{N}_i) b_i \\ &= 2S - \mathcal{N}_i(1 + \mathcal{N}_i) + b_i^\dagger b_i^\dagger b_i b_i \\ &= 2S - \mathcal{N}_i(1 + \mathcal{N}_i) + b_i^\dagger (b_i b_i^\dagger - 1) b_i \\ &= 2S - \mathcal{N}_i - \mathcal{N}_i = 2S_i^z \end{aligned} \tag{8.41}$$

which is the correct commutation relation. One can thus fully represent the spins in terms of boson operators, which opens the way to use all the second quantization formalism that we introduced in the first part. The price to pay is of course that the representation is quite complicated in terms of bosons.

For spin 1/2 there is a considerable simplification to this representation [MM56]. In that case the Hilbert space can only contain either zero bosons or one boson per site, since the S^z component can be only $+1/2$ or $-1/2$. As a result the factors $\sqrt{2S - \mathcal{N}_i}$ can only take the value 1 within the Hilbert space. The mapping between spins and bosons thus becomes, for spin 1/2

$$\begin{aligned} S_i^+ &\to b_i \\ S_i^- &\to b_i^\dagger \\ S_i^z &= 1/2 - \mathcal{N}_i \end{aligned} \tag{8.42}$$

and the constraint that there can be at most one boson on each site. To remind of this constraint one sometimes speaks of “hard core” bosons to indicate bosons which repel so strongly that only one boson can exist on a site. This constraint can also be taken into account by putting an on-site repulsion of the Hubbard form

$$H_U = \frac{U}{2} \sum_i \mathcal{N}_i(\mathcal{N}_i - 1) \tag{8.43}$$

which ensures that zero or one boson per site do not pay any repulsion, while two, three, etc. bosons per site woudl pay the repulsion U, 3U, etc. Letting $U \to \infty$ thus ensure the hard core constraints.

Using the Matsubara and Mastuda representation one can get a very physically transparent representation of the Heisenberg model in term of bosons. Let us write the Heisenberg model as

$$H = \frac{J_{xy}}{2} \sum_{\langle i,j \rangle} [S_i^+ S_j^- + S_i^- S_j^+] + J_z \sum_{\langle i,j \rangle} S_i^z S_j^z \tag{8.44}$$

where we hare made explicit the separation between the x, y component and the z one. Of course in the isotropic Heisenberg model $J_{xy} = J_z = J$. Here we have introduced the possibility to get different exchanges in the xy plane and along the z axis (single axis anisotropy). In addition to the fact that in a solid such an anisotropy can exist, it is interesting from the theoretical point of view, since as will be seen, in the boson language the x, y and z parts of the Hamiltonian have quite different physical interpretations. Sometimes such a model is called the XXZ model. If we use the representation (**??**), the XXZ Hamiltonian becomes

$$\begin{aligned} H = \frac{J_{xy}}{2} \sum_{\langle i,j \rangle} [b_i^\dagger b_j + b_i^\dagger b_j] \\ + J_z \sum_{\langle i,j \rangle} (\mathcal{N}_i - 1/2)(\mathcal{N}_j - 1/2) \\ + U \sum_i \mathcal{N}_i(\mathcal{N}_i - 1) \end{aligned} \tag{8.45}$$

The first line in (8.45) is thus tight binding bosons jumping on the sites of the lattice with a hopping amplitude $\frac{J_{xy}}{2}$. Note that if $J_{xy} < 0$ (ferromagnetic case) the minimum of energy of this part will be at $k = 0$ while it would be at $k = \pi$ for $J_{xy} > 0$. In the later case it will thus be useful to first make a canonical transformation to bring back the minimum around $k = 0$. We will come back to this point later. The second line comes from the J_z term. In the boson language this is a nearest neighbor repulsion (or attraction depending on the sign of J_z)

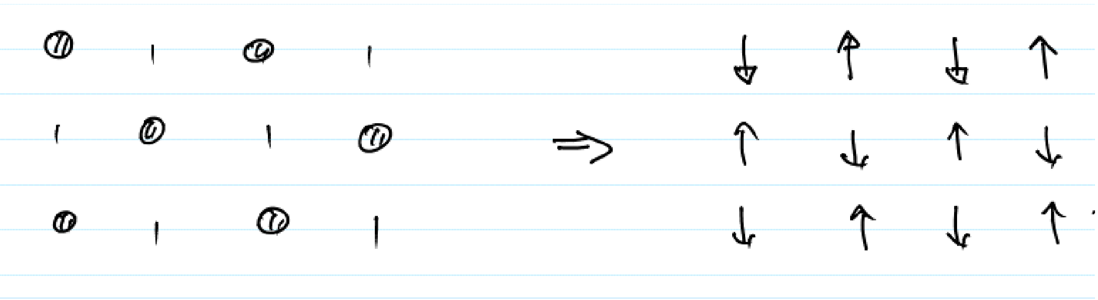


Figure 8.6: As an example the classical antiferromagnetic state with the S^z component of the spins alternating maps on a crystal of bosons.

betweem the bosons. We thus see that the representation has quite different physical meanings for the J_{xy} and J_z terms, one being the kinetic energy of the bosons, and the second their nearest neighbor interactions. Of course this is simply due to the explict choice of quantization axis and for $J_{xy} = J_z$ the model is still fully symmetric by spin rotation. But such a symetry is more difficult to see in the boson representation. On the other hand since we now very well how to solve free bosons on a lattice, one can guess that models with small J_z or for which we can make approximations for the J_z term shoulb tractable in such a representation. Finally the third line, with $U \to \infty$ is implementing the hard core constraint for the bosons. Since this interaction is getting very large, it is clearly a term that will pose difficulties to treat analytically.

Finally let us consider the magnetization along the z direction. If the system is non magnetized

$$\langle S_i^z \rangle = \langle \mathcal{N}_i - 1/2 \rangle = 0 \tag{8.46}$$

which means for the bosons $\mathcal{N}_i = 1/2$. The system is thus at half filling. A magnetic field in the spin representation gives

$$H = -g\mu_B h \sum_i S_i^z = +g\mu_B \sum_i (\mathcal{N}_i - 1/2) \tag{8.47}$$

and thus acts as a chemical potential for the bosons. Similarly the various states can easily be mapped onto one another. For example, as shown in Fig. 8.6 the classical antiferromagnetic state where the spins are alternating from one site of the lattice to the next along the z direction, maps onto a crystal state of bosons. One and thus use the intuition and knowledge of the phases of one representation to deduce possible phases and properties in the other language.

Using such a representation we are now able to deal with many spins coupled together.

8.3.2 Localized Ferromagnetism

Let us first consider the ferromagnetic case $J = -J_f < 0$. We will consider the generic case of a spin S. In that case it is quite natural to check if the "classical" ferromagnetic state $|Fer\rangle = |\uparrow, \uparrow, \uparrow, \ldots, \uparrow\rangle$ would be the ground state of the system described by the quantum Hamiltonian (8.44).

Let us first check that it is an eigenstate. Using (8.44) and $J_{xy} = J_z = -J_f$ one has clearly

$$H\,|Fer\rangle = -J_f \sum_{\langle i,j \rangle} S_i^z S_j^z \,|fer\rangle = -J_f \frac{N_s \mathcal{Z}}{2} S^2 \,|Fer\rangle \tag{8.48}$$

where N_s is the number of sites of the lattice, $\mathcal{Z}$ the coordination number (i.e. the number of neighbors of one given site – 2 in dimension one, 4 in dimension two, 6 in dimension three etc. for a square lattice). This state is thus an eigenstate. Proving that this is the state of lowest energy is a little bit more tricky [Aue98] and we will not do it here, but this is the case. We have thus found "the" ground state of the ferromagnetic system. This answer is physically intuitive, this is indeed the state in which all the spins are aligned along the z direction. This is the state of total spin $\mathcal{S}^z = N_s S$ and thus of total spin $N_s S$. Although the problem is fully quantum is identical to what we would have for a fully classical version of the Heisenberg Hamiltonian. Of course having said that one realizes immediately that the ground state cannot be unique. Indeed the choice of the quantization axis is totally arbitrary and thus we can have the total spin pointing in any direction. Said differently any (quantized) rotation of the state $|Fer\rangle$ is also an eigenstate with the same energy since the Heisenberg Hamiltonian is fully invariant by spin rotation. This huge degeneracy of the ground state is quite rare in quantum mechanics for coupled systems since usually any coupling between two states leads to a splitting of the levels and a lift of the degeneracy.

Having now found the ground state(s) we need to understand the excitations and the first excited states. For that we will assume that we choose the quantization axis along the direction of the total spin, so that the ground state is indeed $|Fer\rangle$ where all the spins point up in the z direction. To analyze the excitations we will make use of the representation (8.36-8.38). We will furthermore assume that we are only interested in states of energy close to the ground state, and thus hopefully states for which the spins have not deviated too much from the direction of the ground state. This is an hypothesis that will need to be checked at the end, but let us proceed with it for the moment. In the boson language it means that we want the number of bosons on each site to be small compared to S. This allows to considerably simplify (8.36) which becomes

$$S_i^+ \simeq \sqrt{2S} b_i \tag{8.49}$$

This allows to rewrite the xy part of the Heisenberg Hamiltonian as

$$H_{xy} = -\frac{J_f(2S)}{2} \sum_{\langle i,j \rangle} [b_i^\dagger b_j + b_j^\dagger b_i] \tag{8.50}$$

which is simply free tight binding bosons moving on a lattice. One notices that this term is of order $b^\dagger b$. One will thus need to be carefull in the term J_z to retain the *linear* order in $b^\dagger b$ (i.e. for this term the $\mathcal{N}_i$ terms) to be consistent. This term thus becomes:

$$\begin{aligned} H_z &= -J_f \sum_{\langle i,j \rangle} (S - \mathcal{N}_i)(S - \mathcal{N}_j) \\ &\simeq -J_f \sum_{\langle i,j \rangle} (S^2 - 2S\mathcal{N}_i) \\ &= E_{Fer} + \frac{J_f \mathcal{Z}(2S)}{2} \sum_i \mathcal{N}_i \end{aligned} \tag{8.51}$$

where E_{Fer} is the energy of the ground state as defined in (8.48). The finally gives for the full Hamiltonian

$$H = -(J_f S) \sum_{\langle i,j \rangle} [b_i^\dagger b_j + b_j^\dagger b_i] + (J_f \mathcal{Z} S) \sum_i \mathcal{N}_i + E_{Fer} \tag{8.52}$$

We thus have, with the approximation of a small number of bosons, a tight binding Hamiltonian of free bosons moving on the lattice, with a chemical potential for the bosons which is given by the J_z term. Since the system is invariant by translation, a Fourier transform is the good way to diagonalize the Hamiltonian. Let us do it on a square lattice in d dimensions. We use the

Fourier transform (3.72) which in this case reads

$$b_i = \frac{1}{\sqrt{N_s}} \sum_{\boldsymbol{k}} b_{\boldsymbol{k}} e^{i\boldsymbol{k}\boldsymbol{r}} \tag{8.53}$$

and thus

$$H = -\frac{J_f S}{2} \sum_{\boldsymbol{z}} \sum_{\boldsymbol{k}} 2\cos(\boldsymbol{k}\boldsymbol{z}) b_k^\dagger b_k + (J_f \mathcal{Z} S) \sum_{\boldsymbol{k}} b_k^\dagger b_k + E_{Fer} \tag{8.54}$$

where $\boldsymbol{z}$ is the vector connecting to the neighbors of a given site (and thus there are $\mathcal{Z}$ such vectors). If one is on a square lattice these vectors are simply the unit vectors of the lattice along the $x, y, z, \ldots$ directions. One has thus (a being the lattice spacing)

$$H = (J_f S \mathcal{Z}) \sum_{\boldsymbol{k}} [1 - \frac{2}{\mathcal{Z}} (\cos(k_x a) + \cos(k_y a) + \ldots)] b_k^\dagger b_k + E_{Fer} \tag{8.55}$$

The excitations are thus free bosonic particles with the momentum $\boldsymbol{k}$ and the energy $E(\boldsymbol{k})$ above the ground state energy, where

$$E(\boldsymbol{k}) = (J_f S 2d)[1 - \frac{1}{d} \sum_{j=1}^{d} \cos(k_j a)] \tag{8.56}$$

This is a remarkable result since we have established that the spin excitations in a ferromagnet are bosonic in nature, they carry the quantum number $\Delta S^z = 1$ since they correspond to the creation of a boson they lower the total spin along z by one unit, and they have a well defined dispersion (i.e. relation between the momentum of the excitation and the energy) given by (8.56). The dispersion is represented, for $d = 1$ on Fig. 8.7. We notice several remarkable feature. The first one is that the energy is minimum at $\boldsymbol{k} = 0$. Quite remarkably the difference of energy between such excitation with $\Delta S^z = -1$ and the ground state is exactly zero. This sounds surprising but is in fact a consequence of the invariance by rotation of the Hamiltonian. Indeed let us consider the state that corresponds to such an excitation. It is given by

$$\begin{aligned} b^\dagger_{\boldsymbol{k}=0} |\varnothing\rangle &= \frac{1}{\sqrt{N}} \sum_i S_i^- |S,S,S,S,S,\ldots,S\rangle \\ &= \frac{1}{\sqrt{N}} [|S-1,S,S,S,S,\ldots,S\rangle + |S,S-1,S,S,S,\ldots,S\rangle + \cdots + |S,S,S,S,S,\ldots,S-1\rangle] \end{aligned} \tag{8.57}$$

which is a uniform superposition of all states with one spin tilted by one quantum of angular momentum. Such a state in in fact the one that corresponds to an infinitesimal rotation applied to the ground state. A rotation of the angular momentum around the vector $\boldsymbol{n}$ can be expressed as [CTDL77]

$$R(\boldsymbol{n}, \alpha) = e^{-i\alpha \boldsymbol{n}\boldsymbol{L}} \tag{8.58}$$

where α is the angle and $\boldsymbol{L}$ the angular momentum operator. Let us take the x axis and do an infinitesimal rotation on each site by the same angle. In that case it is clear that the state

$$|\alpha\rangle = |R(\boldsymbol{x},\alpha)S, R(\boldsymbol{x},\alpha)S, \ldots, R(\boldsymbol{x},\alpha)S\rangle \tag{8.59}$$

has the same energy than the ground state. Expanding (8.58) one finds

$$|\alpha\rangle \simeq |S,S,\ldots,S\rangle - i\alpha[|S-1,S,\ldots,S\rangle + |S,S-1,\ldots,S\rangle + \cdots + |S,S,\ldots,S-1\rangle] \tag{8.60}$$

and thus for this state to have the energy E_{Fer} the action of the Hamiltonian on the second term must give an eigenstate of energy E_{Fer}. This is an important point on which we will come

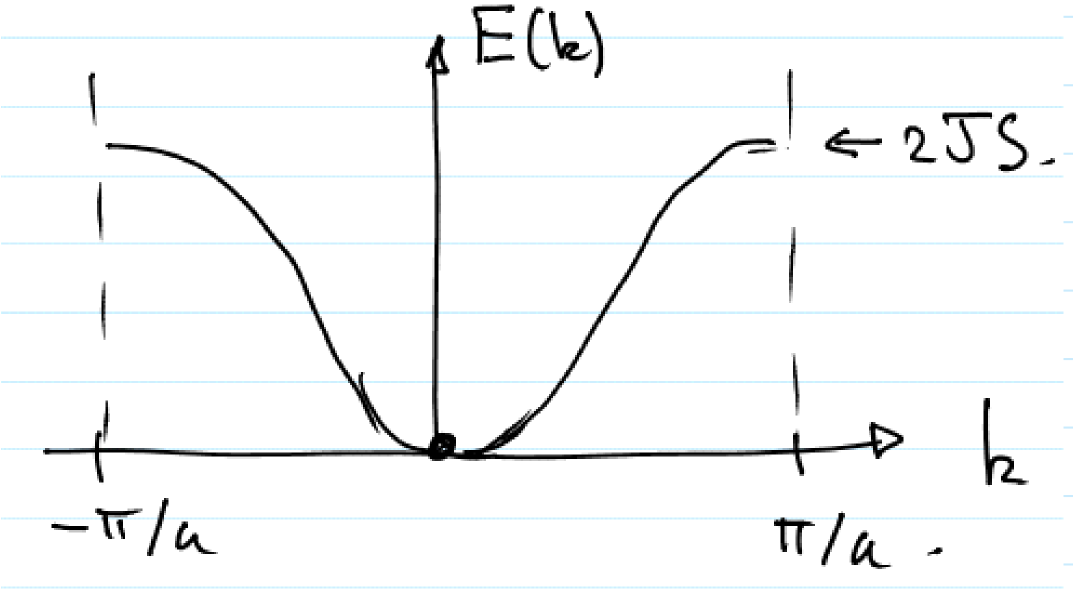


Figure 8.7: Dispersion relations of the spin excitations (magnons) in a ferromagnet. The minimum of energy is at $k = 0$ and is zero since such a uniform excitation would simply correspond to a rotation of the ground state. For small k the energy grows quadratically with k (in any dimension).

back. There are cases for which there is a symmetry of the Hamiltonian which is spontaneously broken in the ground state. This is the case for ferromagnetism since when we pick one of the ground states we break the invariance by rotation. In such a situation applying the symmetry uniformly to the ground state must give a *different* state with the same energy. This means that there exists uniform "excitations" that have zero energy. If now instead of doing the excitation with $\boldsymbol{k} = 0$ one does it with a long wavelength ($\boldsymbol{k}$ small) the energy must be small and tend to zero. This is obviously the case for the magnons here. It is easy to check from (8.56) that for small $\boldsymbol{k}$ the energy behaves as

$$E(\boldsymbol{k}) = (J_f S)\boldsymbol{k}^2 a^2 \tag{8.61}$$

and is thus quadratic in $\boldsymbol{k}$.

The magnon excitations will thus control the properties of localized ferromagnets. Within the approximations that we have made, they are non-interacting and thus any excited state is obtained by exciting a certain number of such magnons. The number will condition the deviations from the fully polarized $\langle S_i^z \rangle = S$ value since each magnon decreases the total magnetization by -1. We will postpone the question of the observation of such excitations until Section 8.5.

One final remaining step is to check the self consistency of our assumptions, namely that the number of magnons is small. Of course at zero temperature the system is in its ground state, but finite temperature allows to thermally excite magnons. Let us thus compute the resulting

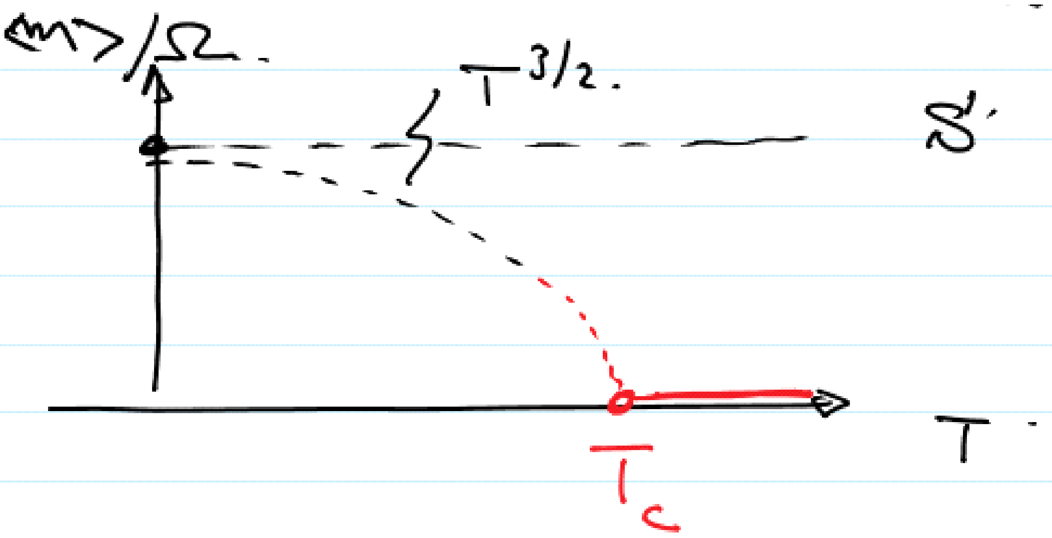


Figure 8.8: Schematic temperature dependence of the magnetization for a ferromagnet. At $T = 0$ the system is fully polarized and the magnetization is the maximum value of the spin S per site. For $d = 1$ and $d = 2$ the magnetization is zero at any finite temperature. For $d = 3$ the magnetization decreases with temperature. Although this is beyond the range of the calculation this strongly suggests the existence of a critical temperature T_c above which the magnetization is zero and the system is in the paramagnetic state. In $d = 3$ a spontaneously broken symmetry state with a finite ferromagnetic order can thus exist at low enough temperature.

magnetization. We have

$$\begin{aligned}\langle S^z_{i=0}\rangle &= S - \frac{1}{N}\sum_{\boldsymbol{k}}\langle b^\dagger_{\boldsymbol{k}} b_{\boldsymbol{k}}\rangle \\ &= S - \frac{1}{(2\pi)^d}\int_{BZ} d^d k \frac{1}{e^{E(\boldsymbol{k})-1}}\end{aligned} \tag{8.62}$$

The integral is limited on the Brillouin zone, thus the only possible divergence is coming from small momenta. Since the integrand is behaving as $1/\boldsymbol{k}^2$ the behavior is strongly dependent on the dimension of the system. For $d = 1$ the integral diverges. This means that our assumptions of a small number of magnon are violated as soon as the temperature is infinitesimal. Within our assumptions the number of magnons would diverge. Of course this divergence is artificial and will be stopped when we take again into account the terms $\mathcal{N}_i\mathcal{N}_j$ that were neglected in (8.51). Although it is of course not possible to rigorously extract from the present calculation what will happen, the proliferation of magnon that one finds strongly suggests that the ordered state is unstable and thus that at *any* finite temperature $\langle S^z_i\rangle = 0$. It is thus impossible to get an ordered antiferromagnetic state at any finite temperature in $d = 1$. In $d = 2$ a similar effect occurs, although the divergence of the integral is only logarithmic. An ordered ferromagnetic state is also forbidden in $d = 2$. In $d = 3$ on the contrary the integral is convergent. It is very easy to analyze the temperature dependence of the integral for very low temperatures by simple dimensional analysis. Rescaling each k component by $1/\sqrt{\beta}$ to eliminate the temperature in the bose factor one finds

$$\langle S^z_{i=0}\rangle \sim S - T^{3/2} Cste \tag{8.63}$$

As can be expected and is schematically shown in Fig. 8.8 the magnetization decreases with temperature since magnons are thermally created. At a finite temperature T_c the magnetization will go to zero (this is beyond the scope of (8.63)) and the system becomes paramegnetic. We will come back in Section 8.4 on the generality of these results.

8.3.3 Localized antiferromagnetism

Now that we understand the ferromagnetic case let us turn to the more complex antiferromagnetic situation $J > 0$. For the sake of simplicity we will restrict ourselves in this course to the case of bipartite lattices, such as the square or the hexagonal lattice in which no frustration is present.

Let us try to follow the same route than for the ferromagnet and determine first the ground state. A good candidate would be the classical antiferromagnetic state where the S^z component of the magnetization alternates from one site to the next:

$$|AF_c\rangle = |S, -S, S, -S, \dots, S\rangle \tag{8.64}$$

Let us check whether such a state is an eigenstate by applying (8.44) on it. Although the J_z part of the Hamiltonian clearly leave the state unchanged it is clear that in the xy part terms such as $S_i^- S_j^+$ will not destroy the state, as was the case for the ferromagnet, but will give

$$S_i^- S_j^+ |S, -S, S, -S, \dots, S\rangle = |S-1, -S+1, S, -S, \dots, S\rangle \tag{8.65}$$

and many similar terms. The classical antiferromagnetic state is thus not an eigenstate of the Hamiltonian. We thus need to find both the ground state and the excitations of the systems. This difference between the ferromagnetic case and the antiferro one is again an illustration of the fact that for a quantum system one cannot simply change $S \to -S$ (on every other site) and thus go from a ferromagnetic coupling to the antiferromagnetic one by this simple transformation. Quantum ferromagnetic states and quantum antiferromagnetic ones are thus fundamentally different.

In order to study the antiferromagnetic case, we will still follow the same route and map the system using the Holstein-Primakoff representation. In order to avoid dealing with an order parameter which oscillate from site to site we will do transformation (only possible to do on a bipartite lattice)

$$\begin{aligned} S_i^x &\to \mathscr{S}_i^x \\ S_i^y &\to (-1)^i \mathscr{S}_i^y \\ S_i^z &\to (-1)^i \mathscr{S}_i^z \end{aligned} \tag{8.66}$$

This transformation which respects the canonical commutation relations for the spins changes the classical antiferromagnetic state for the spins S $|S, -S, S, -S, \dots, -S\rangle$ into a ferromagnetic state for the spins $\mathscr{S}$ namely $|S, S, S, S, \dots, S\rangle$. It will thus provide a starting point, in the language of the spins $\mathscr{S}$ which is invariant by translation and thus on which it will be more easy to make approximations. Note that in fourier transform it means simply that we have shifted the momentum of the y and z components of the spins by π. Let us write the Hamiltonian in term of the new spins.

$$\begin{aligned} H &= J \sum_{\langle i,j \rangle} [S_i^x S_j^x + S_i^y S_j^y] + J \sum_{\langle i,j \rangle} S_i^z S_j^z \\ &= J \sum_{\langle i,j \rangle} [\mathscr{S}_i^x \mathscr{S}_j^x - \mathscr{S}_i^y \mathscr{S}_j^y] - J \sum_{\langle i,j \rangle} \mathscr{S}_i^z \mathscr{S}_j^z \\ &= \frac{J}{2} \sum_{\langle i,j \rangle} [\mathscr{S}_i^+ \mathscr{S}_j^+ + \mathscr{S}_i^- \mathscr{S}_j^-] - J \sum_{\langle i,j \rangle} \mathscr{S}_i^z \mathscr{S}_j^z \end{aligned} \tag{8.67}$$

We see that in the new spins $\mathscr{S}$ the Hamiltonian takes a new form. The J_z part of the Hamiltonian is now the conventional ferromagnetic coupling between the z components of two neighboring spins. This term will clearly favor a classical ferromagnetic ground state. The first part however is now able to flip up or down two neighboring spins, making clear that the action of such a term on a fully polarized state would not give zero. As for the ferromagnetic case we now proceed with the Holstein-Primakoff mapping. In a similar way we will assume that the number of bosons in the ground state is small. This was clear at zero temperature for the ferromagnetic case, since this number was zero. For the antiferromagnetic one, this is clearly not the case, and we can expect that a *finite* number of bosons will be present in the ground state *even* at zero temperature. However the classical state should not be a too bad starting point, and thus we can expect this number to be reasonably small. We will thus proceed with this assumption and test it at the end. Using the representation (8.49) one has:

$$H = JS \sum_{\langle i,j \rangle} [b_i^\dagger b_j^\dagger + b_i b_j] + (J_f \mathcal{Z} S) \sum_i \mathcal{N}_i + E_{Fer} \tag{8.68}$$

using the same linearization of the J_z term than for the ferromagnetic case, to be consistent with the first term. Since we have now thanks to the transformation (8.66) a starting point which is invariant by translation we can proceed with the Fourier transform and find using (8.53)

$$H = \frac{JS}{2} \sum_{\boldsymbol{k},\boldsymbol{z}} [e^{-i\boldsymbol{k}\boldsymbol{z}} b^\dagger_{\boldsymbol{k}} b^\dagger_{-\boldsymbol{k}} + e^{i\boldsymbol{k}\boldsymbol{z}} b_{-\boldsymbol{k}} b_{\boldsymbol{k}}] + (J \mathcal{Z} S) \sum_{\boldsymbol{k}} b^\dagger_{\boldsymbol{k}} b_{\boldsymbol{k}} + E_{Fer} \tag{8.69}$$

This expression shows that our approximation reduce the Hamiltonian to a quadratic form for the bosons. However contrarily to the case of ferromagnetism this form is not diagonal and has terms that can create or destroy two bosons. This will clearly lead to a ground state containing a finite number of bosons. Nevertheless since the form is quadratic it can still be diagonalized by a Bogoliubov transformation. This will allow us to determine the excitations of the system, and thus also the ground state. If order to perform the transformation let us first rewrite the Hamiltonian (8.69). Clearly the transformation will have to couple terms $b_{\boldsymbol{k}}$ and $b^\dagger_{-\boldsymbol{k}}$. Since in the $\sum_{\boldsymbol{k}}$ such terms would be counted twice, it is important to make sure that such double counting will not occur. We will thus restrict the sum over $\boldsymbol{k}$ to the "positive" values. In one dimension this means that we restrict the sum to $k \in [0, \pi]$. For higher dimensions we simply take only half of the Brillouin zone such that if a point $\boldsymbol{k}$ is kept, the point vk is not excluded. Using this restricted zone one obtains

$$H = 2JS \sum_{\boldsymbol{k}}{}' \sum_{j=1}^{d} \cos(k_j a) [b^\dagger_{\boldsymbol{k}} b^\dagger_{-\boldsymbol{k}} + b_{-\boldsymbol{k}} b_{\boldsymbol{k}}] + (J2dS) \sum_{\boldsymbol{k}}{}' [b^\dagger_{\boldsymbol{k}} b_{\boldsymbol{k}} + b^\dagger_{-\boldsymbol{k}} b_{-\boldsymbol{k}}] + E_{Fer} \tag{8.70}$$

where $\sum'$ denotes the restricted sum. This can be put in a matrix form

$$H = (JS2d) \sum_{\boldsymbol{k}}{}' \begin{pmatrix} b^\dagger_{\boldsymbol{k}} & b_{-\boldsymbol{k}} \end{pmatrix} \begin{pmatrix} 1 & \gamma(\boldsymbol{k}) \\ \gamma(\boldsymbol{k}) & 1 \end{pmatrix} \begin{pmatrix} b_{\boldsymbol{k}} \\ b^\dagger_{-\boldsymbol{k}} \end{pmatrix} - (J2dS) \frac{N_s}{2} + E_{Fer} \tag{8.71}$$

where

$$\gamma(\boldsymbol{k}) = \frac{1}{d} \sum_{j_1}^{d} \cos(k_j a) \tag{8.72}$$

Note that in (8.71) the operators $b_{\boldsymbol{k}}$ and $b_{-\boldsymbol{k}'}$ now correspond to different values of their index regardless of the values of $\boldsymbol{k}$ and $\boldsymbol{k}'$, and thus the corresponding operators commute. We can thus make a Bogoliubov transformation of the form

$$\begin{pmatrix} b_{\boldsymbol{k}} \\ b^\dagger_{-\boldsymbol{k}} \end{pmatrix} = \begin{pmatrix} u_{\boldsymbol{k}} & -v_{\boldsymbol{k}} \\ -v_{\boldsymbol{k}} & u_{\boldsymbol{k}} \end{pmatrix} \begin{pmatrix} \alpha_{\boldsymbol{k}} \\ \beta^\dagger_{\boldsymbol{k}} \end{pmatrix} \tag{8.73}$$

One should keep in mind that $\boldsymbol{k}$ is restricted to half the Brillouin zone in the above expressions. One could thus have also used $\alpha_{-\boldsymbol{k}}$ instead of $\beta_{\boldsymbol{k}}$. The above expression emphasizes the difference between the two operators α and β (as the original operators $b_{\boldsymbol{k}}$ and $b^\dagger_{-\boldsymbol{k}}$ which correspond to two different objects). If one wants that the new excitations $\alpha_{\boldsymbol{k}}$ and $\beta_{\boldsymbol{k}}$ to have standard bosonic commutation relations, then it is easy to check that the commutator $[b_{\boldsymbol{k}}, b^\dagger_{\boldsymbol{k}}] = 1$ imposes $u^2_{\boldsymbol{k}} - v^2_{\boldsymbol{k}} = 1$ if one takes the u and v reals. The other relations are automatically satisfied by the form (8.73). One can thus also use the parametrization

$$u_{\boldsymbol{k}} = \cosh(\theta_{\boldsymbol{k}}) \quad , \quad v_{\boldsymbol{k}} = \sinh(\theta_{\boldsymbol{k}}) \tag{8.74}$$

which automatically ensures the constraint. Note the difference between this transformation and the one we used in Section 3.6.2. The difference is mainly due to the difference of statistics (bosons vs fermions) in the operators.

Using the transformation (8.73) we obtain for the matrix in the Hamiltonian

$$\begin{pmatrix} \cosh(2\theta_{\boldsymbol{k}}) - \sinh(2\theta_{\boldsymbol{k}})\gamma(\boldsymbol{k}) & -\sinh(2\theta_{\boldsymbol{k}}) + \cosh(2\theta_{\boldsymbol{k}})\gamma(\boldsymbol{k}) \\ -\sinh(2\theta_{\boldsymbol{k}}) + \cosh(2\theta_{\boldsymbol{k}})\gamma(\boldsymbol{k}) & \cosh(2\theta_{\boldsymbol{k}}) - \sinh(2\theta_{\boldsymbol{k}})\gamma(\boldsymbol{k}) \end{pmatrix} \tag{8.75}$$

This gives the condition

$$\tanh(2\theta_{\boldsymbol{k}}) = \gamma(\boldsymbol{k}) \tag{8.76}$$

This condition can always be satisfied since we see from (8.72) that $-1 \le \gamma(\boldsymbol{k}) \le 1$. Using this condition we obtain for the Hamiltonian

$$H = (JS2d) \sum_{\boldsymbol{k}}{}' \sqrt{1 - \gamma^2(\boldsymbol{k})} [\alpha^\dagger_{\boldsymbol{k}}\alpha_{\boldsymbol{k}} + \beta_{\boldsymbol{k}}\beta^\dagger_{\boldsymbol{k}}] - (J2dS)\frac{N_s}{2} + E_{Fer} \tag{8.77}$$

One can rewrite this expression in a more compact form by reintroducing the $-\boldsymbol{k}$ excitations $\alpha_{-\boldsymbol{k}} = \beta_{\boldsymbol{k}}$. Using the fact that $\sum'_{\boldsymbol{k}} 1 = \frac{1}{2}\sum_{\boldsymbol{k}} 1$ one can go back to a sum over the whole Brillouin zone with "only" excitations $\alpha_{\boldsymbol{k}}$ (where now $\boldsymbol{k}$ is unrestricted) to obtain

$$\begin{aligned} H &= (JS2d) \sum_{\boldsymbol{k}} \sqrt{1 - \gamma^2(\boldsymbol{k})} [\alpha^\dagger_{\boldsymbol{k}}\alpha_{\boldsymbol{k}} + \frac{1}{2}] - (J2dS)\frac{N_s}{2} + E_{Fer} \\ &= (JS2d) \sum_{\boldsymbol{k}} \sqrt{1 - \gamma^2(\boldsymbol{k})} \alpha^\dagger_{\boldsymbol{k}}\alpha_{\boldsymbol{k}} + E_{Fer} \end{aligned} \tag{8.78}$$

The $\alpha_{\boldsymbol{k}}$ operators and the Hamiltonian (8.78) fully define the excitations and the ground state of the antiferromagnet. The ground state $|AF\rangle$ is defined by

$$\alpha_{\boldsymbol{k}} |AF\rangle = 0 \quad , \quad \forall \boldsymbol{k} \tag{8.79}$$

since the energies of the $\alpha_{\boldsymbol{k}}$ excitations are positive for the whole Brillouin zone. Computing $|AF\rangle$ explicitly is not obvious but as usual we don't need to do it, since we know the action of the operators α on it, which is sufficient to compute any observable of the problem. What we can say, given the fact that the operator α contains both creation and destruction of bosons, is that the ground state contains already a certain number of bosons. We thus why the fully polarized classical ground state $|AF_c\rangle$ was not a good solution. There are quantum fluctuations that make the spin fluctuate and partly reduce the polarization of the staggered order parameter on each site. The above solution allows us to take this effect into account.

In addition to the ground state we get the excitations. We see that, in the same way that for the ferromagnet the excitations (also called spin-waves or magnons) are free bosons. Their relation dispersion is $E(\boldsymbol{k}) = \sqrt{1 - \gamma^2(\boldsymbol{k})}$ which is shown in Fig. 8.9 The first thing we can note from Fig. 8.9 is the periodicity of the energy dispersion. It is not periodic with period π instead of 2π for the ferromagnet. This halving of the Brillouin zone is quite logical since the presence of

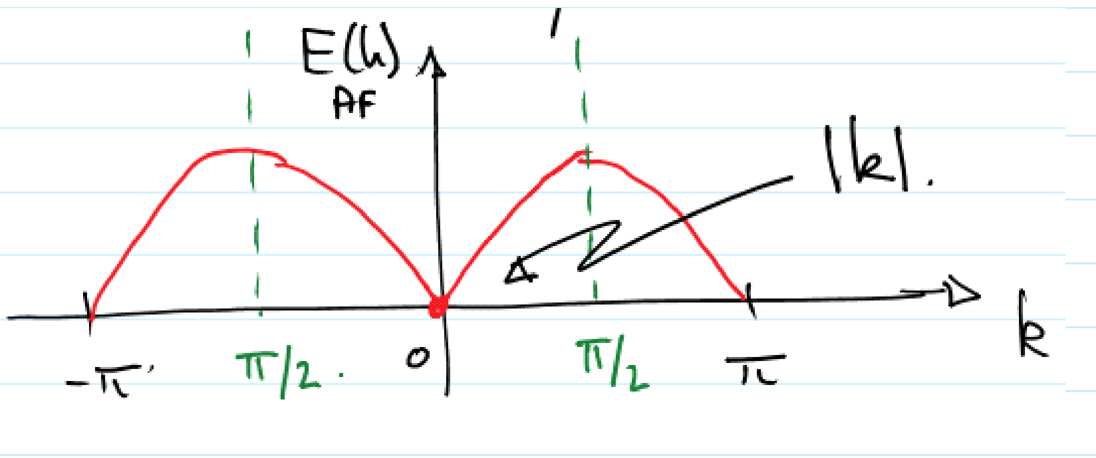


Figure 8.9: Dispersion relation for an antiferromagnet. Because of the doubling of the zone due to the antiferromagnetic order the energy is now periodic with a period π/a, instead of $2\pi/a$. The energy goes to zero at $\boldsymbol{k} = 0$ again as a consequence of the invariance by spin rotation of the original model. However for an antiferromagnet the dispersion relation is *linear* around the point of zero energy.

an antiferromagnetic order imposes a doubling of the minimal periodic cell in real space. We also note that the energy goes to zero at $k = 0$ and $k = \pi/a$ (which are now equivalent). This is again a consequence of the full rotation symmetry of the Hamiltonian and the fact that the ground state breaks this symmetry. We thus recover the presence of Goldstone modes. Around the zero energy point the dispersion relation is linear $E(\boldsymbol{k}) \sim V|\boldsymbol{k}|$ where V has the dimensions of a velocity. For example in one dimension $V = JS2a$ and is thus proportional to J. The spin excitations are thus, for small $\boldsymbol{k}$ (or $\boldsymbol{k} \sim \pi$) bosonic particles with a Dirac-like dispersion. The larger the exchange is the faster this particles can move. Compare to the ferromagnet for which the dispersion was $\sim \boldsymbol{k}^2$ we thus see that in a way the antiferromagnet is more "rigid" since creating excitation of the same wavelength would cost a higher energy. As for the ferromagnet, any excited state can thus be realized by creating several magnon excitations, which within our approximations behave as free particles.

Let us now turn to the check of our initial hypothesis, namely that the number of bosons $\mathcal{N}$ is small in the ground state $|AF\rangle$. To do so let us compute the staggered magnetization in the ground state. Given the invariance by translation we can simply compute the magnetization at site $i = 0$. It is given by

$$\langle S^z_{i=0}\rangle = S - \langle AF|b^\dagger_0 b_0|AF\rangle \tag{8.80}$$

The number of bosons must be reexpressed in terms of the α and β excitations

$$\begin{aligned}
b^\dagger_0 b_0 &= \frac{1}{N_s}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2} b^\dagger_{\boldsymbol{k}_1} b_{\boldsymbol{k}2} \\
&= \frac{1}{N_s}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2}{}' (b^\dagger_{\boldsymbol{k}_1} + b^\dagger_{-\boldsymbol{k}_1})(b_{\boldsymbol{k}2} + b_{-\boldsymbol{k}2}) \\
&= \frac{1}{N_s}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2}{}' [u_{\boldsymbol{k}_1}(\alpha^\dagger_{\boldsymbol{k}_1} + \beta^\dagger_{\boldsymbol{k}_1}) - v_{\boldsymbol{k}_1}(\beta_{\boldsymbol{k}_1} + \alpha_{\boldsymbol{k}_1})][u_{\boldsymbol{k}_2}(\alpha_{\boldsymbol{k}_2} + \beta_{\boldsymbol{k}_2}) - v_{\boldsymbol{k}_2}(\beta^\dagger_{\boldsymbol{k}_2} + \alpha^\dagger_{\boldsymbol{k}_2})]
\end{aligned} \tag{8.81}$$

Using the fact that the ground state $|AF\rangle$ is annihilated by the α and β destruction operators one obtains for the average in the ground state

$$
\begin{aligned}
\langle AF|b_0^\dagger b_0|AF\rangle &= \frac{1}{N_s}\sideset{}{'}\sum_{\boldsymbol{k}_1} v_{\boldsymbol{k}_1}^2 2 \\
&= \frac{1}{N_s}\sideset{}{'}\sum_{\boldsymbol{k}} \left[\frac{1}{\sqrt{1-\gamma^2(\boldsymbol{k})}} - 1\right]
\end{aligned}
\tag{8.82}
$$

Note that the finite temperature could have been obtained in the same way using that $\langle \alpha_{\boldsymbol{k}}^\dagger \alpha_{\boldsymbol{k}}\rangle = f_{\mathrm{B}}(E(\boldsymbol{k}))$.

Based on (8.82) we can now check what happens as a function of the dimensionality. For $d = 1$ the integral in (8.82) is divergent which means that the number of bosons *in the ground state* would diverge. This is of course an artefact of our approximations but it strongly indicates, as for the ferromagnet, that the staggered magnetization will be reduced to zero. There is thus no possibility to have a spontaneously broken symmetry state with antiferromagnetic order in one dimension. Note that here we are talking about *zero temperature*. The presence of excitations is not due to thermal excitations of the magnons but to the presence of the quantum fluctuations. These fluctuations are enough to prevent the AF order to appear. In $d \geq 2$ the integral is convergent and thus AF order can appear at $T = 0$. Of course a finite temperature will contribute to generating more excitations and can also destroy the order. Here we see a new phenomenon: for quantum problems, the fact that there are uncertainty relations and thus quantum fluctuations can prevent the presence of order *even* at $T = 0$. The effect of the quantum fluctuations is thus a little bit similar to the one of thermal fluctuations for a classical system. This strongly suggests that there is the possibility of having quantum phase transitions, even at $T = 0$, by varying parameters in the Hamiltonian. We will come back to that point in the next section.

A remarkable consequence of the previous calculation is the fact that even at $T = 0$ the staggered magnetization in the ground state is reduced by the quantum fluctuations. For spin 1/2 systems this effect can be quite important, and for example on a two dimensional square lattice the actual staggered magnetization is only about 60% of the fully saturated one. On other lattices this reduction of the AF order can make other type of spin order more competitive than one would naively think from looking at the corresponding classical system.

8.4 Broken symetries for classical and quantum systems

In the previous sections we have explicitly computed the properties of ferromagnets and antiferromagnets and found that depending on the dimension, order could exist or not. In addition we saw that to some extent the presence of quantum fluctuations had striking similarities in their effect on the presence of long range order to the ones of thermal fluctuations. These results are in fact rather general and go beyond the precise models that we have examined here. We will cast them in this chapter in a more general framework.

8.4.1 Symetries and presence of long range order

One important question is whether we can predict the presence or not of long range order. As was investigated by Landau the presence of second order phase transitions is linked to the existence of a spontaneous symmetry breaking. The Hamiltonian possesses a certain symmetry but the configurations of the system that are relevant choose to spontaneously violate this symmetry. One can then define an order parameter which is non zero in the phase with the broken symmetry and zero otherwise. The simplest example of this is for the classical Ising model, where on each site there is a spin $\sigma_i = \pm 1$ and the Hamiltonian is

$$H = -J \sum_{\langle i,j, \rangle} \sigma_i \sigma_j \tag{8.83}$$

In that case the Hamitlonian is invariant by the *discrete* symmetry $\sigma_i \to -\sigma_i$. One the other hand if a spontaneous magnetization

$$m = \frac{1}{N_s} \sum_i \sigma_i \tag{8.84}$$

develops it clearly means that this symmetry has been violated, since otherwise we would have necessarily $m = 0$ by summing over all configurations given the fact that two configurations with all $\sigma_i \to -\sigma_i$ have the same energy and thus the same Boltzmann weight. Such spontaneous symmetry breaking can of course only occur in the thermodynamic limit $N_s \to \infty$. It is thus natural to ask whether the possibility of having long range order depends on the nature of such symmetries.

A simple argument, to address this question, is to start from the assumption that at zero temperature (we are dealing here with classical problems) an ordered state exists that breaks the symmetry and to examine the free energy of defects in this state, which would tend to restore the symmetry. If such a free energy is negative, then defects will proliferate and the broken symmetry state will be destroyed. If on the other hand the free energy of defects is positive, then defects will not occur and the broken symmetry state can exist.

Let us look fist at the case of a system with a discrete symmetry such as the Ising model. In that case at $T = 0$ we can start from the fully polarized state. If we want to create a defect that reduces the magnetization we can invert a block of spins of linear size L as shown in Fig. 8.10. In order for the defect to have an impact on the magnetization in the thermodynamic limit, the size L must be macroscopic and thus grow together with the total size of the system L_t. We can thus consider than $L \propto L_t \to \infty$. Because all the spins are turned inside the block the energy cost of such a defect in the block is zero (this is a consequence of the symmetry of the Hamiltonian). However at the surface of the defects we have replaces a bond with an energy $-J$ by a bond with an energy $+J$ in comparison with the ground state. The energy cost ofthe defect is thus $2J$ times the surface of the defect which for a large block is growing as L^{d-1} where d is the spatial dimension. Thus

$$\Delta E \propto J L^{d-1} \tag{8.85}$$

The fact that we have a *brutal* variation of the order parameter is a consequence of the *discrete* symmetry of the Hamiltonian. If we have created such a defect its entropy is the log of the number of configurations which is roughly the number of positions that this defect can take in the whole system. This entropy is thus growing as

$$\Delta S \propto \log(L_t^d) \propto \log(L) \tag{8.86}$$

Putting the two results together we find for the variation of the free energy due to the presence of a defect

$$\Delta F \simeq J L^{d-1} - T \log(L) \tag{8.87}$$

We thus see from (8.87) that in one dimension the energy cost is constant while the entropy grows logarithmically. There will thus be at *any* finite temperature proliferation of defect (and defects within defects etc.) until the initial order is completely destroyed. We thus immediately recover the fact that at any finite temperature a one dimensional system (without long range interactions) cannot have a finite order parameter. On the other hand we see that for $d = 2$

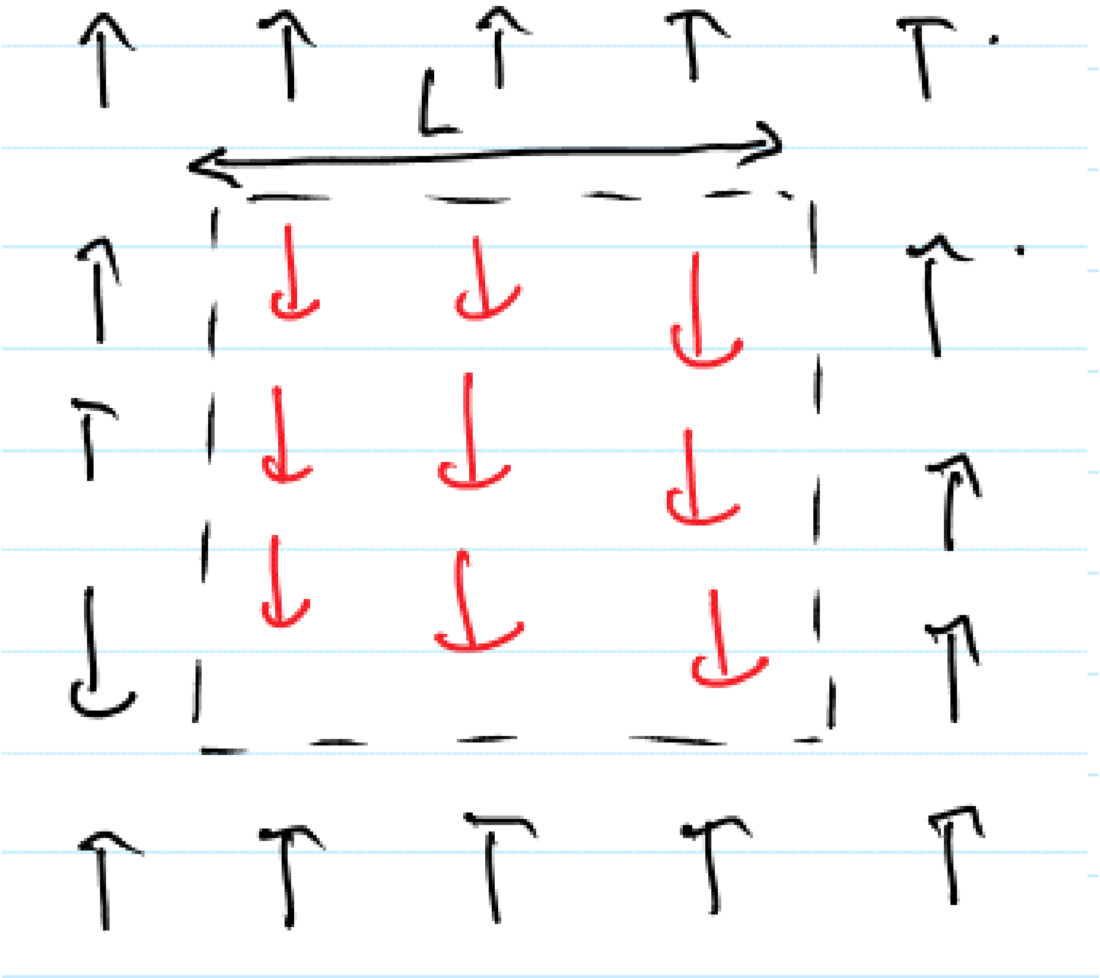


Figure 8.10: A defect of size L is created in the fully polarized state of the Ising model. Since the order parameter must vary brutally at the boundary because of the discrete symmetry, the cost of this defect is proportional to its surface L^{d-1}.

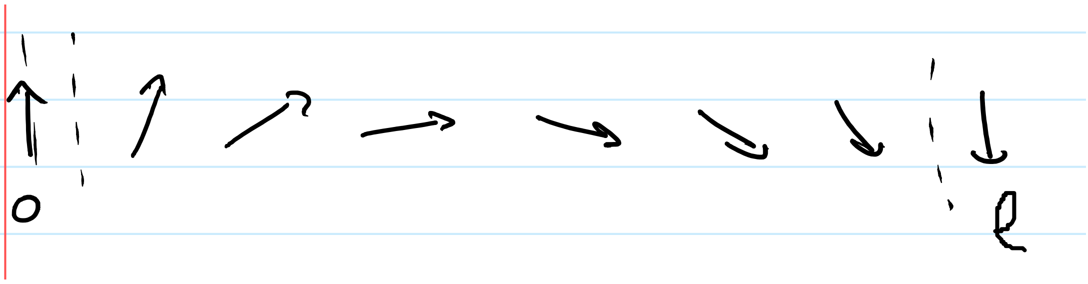


Figure 8.11: For systems with a continuous symmetry there is the possibility to spread the surface of the defect. This leads to a considerable reduction of the energy cost.

and above the energy cost of the defect wins over the entropy gain, and thus provided the temperature is small enough, the ordered state is stable. There is thus for $d \geq 2$ the possibility for systems with a *discrete* symmetry to have a finite order parameter at low temperature.

The situation is different if the Hamiltonian has a *continuous* symmetry. Let us take as an example the so-called XY model which corresponds to spins with two components $\boldsymbol{S} = (\cos(\theta), \sin(\theta))$. In that case the Hamitlonian is

$$H = -J \sum_{\langle i,j, \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j = -J \sum_{\langle i,j, \rangle} \cos(\theta_i - \theta_j) \tag{8.88}$$

and the system is invariant under the continuous symmetry of spin rotation. This is also the case of the Heisenberg model. In that case there is the possibility to spread the surface of the defect by turning the spin gradually from the configuration outside the defect to the one inside. Let us see the energy cost if we perform such a gradual rotation. We look at a single bond, for which the spins goes from fully polarized up to fully polarized down over a length l as shown in Fig. 8.11 For the XY model this would mean that $\theta_i = 0 + \frac{i}{l}\pi$. If l is large we can expand the cosine to get

$$\begin{aligned} \Delta E &= -J \sum_{i=0} l - 1[\cos(\frac{1}{l}\pi) - 1] \\ &= \frac{J\pi^2}{2} \sum_{i=0} l - 1 \frac{1}{l^2} \\ &\simeq \frac{J}{l} \end{aligned} \tag{8.89}$$

We thus see that the presence of the continuous symmetry leads to a drastic reduction of the energy cost at the surface of the defect which is now of order J/l instead of J for a brutal variation of the order parameter. We can thus spread the surface over a size which is of order L which is the largest size we can reach for a defect of size L. In that case the energy cost for the defect becomes $\Delta E \propto JL^{d-2}$ leading to a free energy

$$\Delta F \simeq JL^{d-2} - T\log(L) \tag{8.90}$$

We thus see that for systems with a continuous symmetry, both $d = 1$ and $d = 2$ situations lead to an absence of a finite order parameter. The XY model, as the Heisenberg one cannot have an ordered state in $d = 1$ and $d = 2$ dimensions. It is only in $d = 3$ and above that a finite order parameter can exist at sufficiently low temperatures.

We thus see that the nature of the symmetries have drastic consequences for the presence or not of an order parameter. For classical systems the above arguments allow to classify this physics. For quantum problems the situation is more complex due to the presence of quantum fluctuations. We will thus now examine how one can relate such quantum problems to the classical ones and thus use similar arguments.

8.4.2 Link between classical and quantum systems

The results of the previous chapters suggest that there should be some deep links between the behavior of classical and of quantum systems. Indeed these links can be put on a formal basis in a quite general way. Let us first recall what looking at a classical system means. We have a system described by an Hamiltonian $H_{cl}[\{\sigma\}]$ which is a function of an infinite number of parameters $\{\sigma\} = (\sigma_1, \sigma_2, \dots, \sigma_{N_s})$. The partition function of the problem consists in integrating over this infinite set of variables the Boltzmann weight of each configuation

$$Z_{cl} = \sum_{\{\sigma\}} e^{-\frac{1}{T_{cl}} H_{cl}[\{\sigma\}]} \tag{8.91}$$

The important parameters are thus the size if the system in all dimensions $L_x, L_y, L_z, \dots$, the classical temperature T_{cl} and of course the Hamiltonian H itself giving the energy of a configuration.

For quantum problems the phase transition properties are also determined by looking at the partition function which is now

$$Z = \mathrm{Tr}[e^{-\beta_Q H_Q}] \tag{8.92}$$

the trace is taken on a set of variables (for example the positions of the particles) $|\{x\}\rangle$. Of course H_Q is an operator containing usually non commuting terms. Let us assume for the sake of simplicity that $H_Q = H_p + H_x$ where H_p only depends on momenta, while H_x only depends on positions. This is in fact the case for many Hamiltonians such as the Hubbard model for example. Using this it is possible to bring the quantum partition function in a form that is very similar to a classical one. The idea is to cut β_Q in slices

$$\begin{aligned} Z &= \mathrm{Tr}[e^{-\frac{\beta_Q}{N} H_Q} \dots e^{-\frac{\beta_Q}{N} H_Q}] \\ &= \sum_{\{x_0\},\{x_1\},\dots,\{x_{N-1}\}} \langle\{x_0\}|e^{-\frac{\beta_Q}{N} H_Q}|\{x_1\}\rangle\langle\{x_1\}|e^{-\frac{\beta_Q}{N} H_Q}|\{x_2\}\rangle \dots \langle\{x_{N-1}\}|e^{-\frac{\beta_Q}{N} H_Q}|\{x_0\}\rangle \end{aligned} \tag{8.93}$$

If one chooses N large enough (in fact $N \to \infty$) then $\epsilon = \beta_Q/N$ is small enough so that one can make the approximation

$$e^{-\epsilon H_Q} \simeq e^{-\epsilon H_p} e^{-\epsilon H_x} \tag{8.94}$$

this neglects terms of higher order in ϵ and allows now to evaluate the matrix elements by introducing a complete basis of the momenta

$$A = \langle\{x_0\}|e^{-\epsilon H_p} e^{-\epsilon H_x}|\{x_1\}\rangle = \sum_{\{p\}_0} \langle\{x_0\}|e^{-\epsilon H_p}|\{p_0\}\rangle\langle\{p_0\}|e^{-\epsilon H_x}|\{x_1\}\rangle \tag{8.95}$$

Now in this term all the matrix elements can be easily computed since the Hamiltonian acts on a state made of eigenstates of either position or momenta. One can thus simply replace the operators by the corresponding eigenvalues. The final result is thus

$$A = \langle\{x_0\}|\{p_0\}\rangle e^{-\epsilon H_p[\{p_0\}]} \langle\{p_0\}|\{x_1\}\rangle e^{-\epsilon H_x[\{p_0\}]} \tag{8.96}$$

where one has only numbers. One can thus recast it as a kind of Boltzmann weight

$$A = e^{[} - \epsilon \mathcal{H}[\{x_0\}, \{p_0\}, \{x_1\}] \tag{8.97}$$

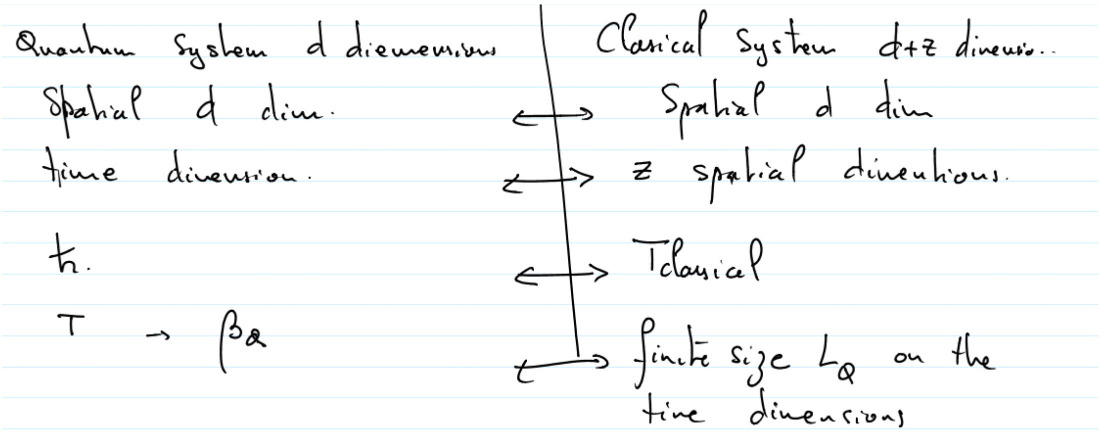


Figure 8.12: Correspondence between classical and quantum problems.

One thus sees that by cutting the "temperature" direction in slices, one can reexpress the partition function of a quantum problem as the partition function of some kind of classical problem. The partition function now reads

$$Z = \sum_{\{x_0\},\{p_0\},\{x_1\},\{p_1\},\ldots,\{x_{N-1}\},\{p_{N-1}\}} e^{-\epsilon\mathcal{H}[\{x_0\},\{p_0\},\{x_1\}]} e^{-\epsilon\mathcal{H}[\{x_1\},\{p_1\},\{x_2\}]} \cdots e^{-\epsilon\mathcal{H}[\{x_{N-1}\},\{p_{N-1}\},\{x_0\}]} \tag{8.98}$$

It is like of the temperature direction was acting a an additional dimension for the the classical problem. Of course the Hamiltonian of the classical problem is not the original H_Q but the modified function $\int_{\tau=0}^{\beta_Q} \mathcal{H}[\tau]$. We thus see that the correspondence is rather subtle and is detailed in Fig. 8.12. Let us detail our findings. A quantum problem in spatial dimension d is thus equivalent to a classical problem in a priori $d+1$ dimensions, where the extra dimension is coming from the slicing of β_Q in the partition function. In fact one has to be more careful. Indeed β_Q has the dimensions of an energy, and in general this "direction" is not equivalent to a simple spatial dimension. This will depend between the relation between the energy and the momenta, i.e. of the dispersion relation of the excitations. Let us for simplicity assume that the dispersion relation is of the form

$$E(\boldsymbol{k}) \sim \boldsymbol{k}^z \tag{8.99}$$

then it means that space and time (inverse of energy) are related (using the uncertainty relations) by

$$L_\tau \sim L_x^z \tag{8.100}$$

The slicing of β_Q thus gives in fact z additional directions, z itself depending on the nature of the quantum problem. We can thus say that a quantum problem of d spatial dimensions will be equivalent to a classical problem of $d+z$ dimensions, z being called the dynamical exponent and being directly determined by the nature of the excitations of the quantum problem. Moreover because of the trace, the slicing of β_Q has to be periodical, which means that for the classical problem these additional z direction are of *finite size* which is determined by the quantum temperature β_Q. The classical problem is defined on a torus along those directions. On the contrary the equivalent of the *classical* temperature is played by $1/\hbar$ in the classical-quantum mapping. We thus see that this mapping makes precisely the connection that we had obtained intuitively before that the quantum fluctuations were playing at $\beta_Q = \infty$ a little bit the same role than the classical thermal fluctuations.

The above mapping is extremely powerful and allows us to relate quantum problems both at zero and finite temperatures, to classical problems. We can thus borrow from the vast knowledge of phase transitions for classical problems to study the properties of the associated quantum problem. As a simple example let us recover the existence or absence of magnetic order that we had derived directly.

For a ferromagnetic model, we can map it onto a classical model with $d+2$ dimension since the dispersion relation is quadratic. The resulting classical model will of course also have a continuous symmetry corresponding to the spin rotation symmetry for the quantum one. We can thus immediately infer that at $\beta_Q = \infty$ (zero temperature for the quantum problem), this will correspond to classical problems in $d_{cl} = 3, 4, \ldots$ for $d = 1, 2, \ldots$ for the quantum problem. There is thus the possibility to have a non zero order parameter. This is what we found by the explicit calculations of the previous section. On the contrary when the quantum problem is at finite temperature, the 2 extra dimensions become of finite size. They thus essentially do not contribute for a critical phenomenon, and we are back to $d_{cl} = 1, 2, 3$ and find that no order parameter can exist below three dimensions, which was exactly our results. For the antiferromagnet, we can make the mapping to a classical model with $d+1$ dimensions because of the linear dispersion relation. As a result we see that even at $\beta_Q = \infty$ the quantum model would correspond to a $d = 2$ classical model with a continuous symmetry and thus must loose the order parameter. For $d = 2$ and above one can have a broken symmetry ground state as we had found. At finite temperature the extra dimension, being of finite size disappears. As a consequence we can have an ordered antiferromagnet at finite (quantum) temperature only for dimension three and above.

Since varying the $1/\hbar$ is equivalent to varying the temperature of the associated classical problem we see that we can obtain quantum phase transitions at $\beta_Q = \infty$ which will be the equivalent of classical phase transitions as a function of the temperature for the associated classical model. We will not enter too deep in this direction and refer the reader to [Sac00] for more details.

Let us finish this section by mentioning a few points. The mapping we have shown is a particular case of something names path integral which allows a representation of quantum mechanics in term of the action. This formalism developed by Feynman [FH65, NO87] is extremely useful and allows many impressive treatments of quantum and classical problems alike. It would however take us too far from the main topic of this book and we refer the reader to the existing literature. One of the very useful consequences of the formula (8.98) is to open the path to a numerical evaluation of the partition function of quantum problems. Indeed we see that the problem reduces to a large but straightforward integral which can be computed, just as for classical problems, by a monte-Carlo evaluation as will be detailed in the next part. For bosons the procedure is working extremely well. For fermions however there are complications due to the fact that the weights are in fact a priori not positive for the antisymmetric configurations that are acceptable for the fermions. Since the symmetric configurations have a lower energy, the monte-Carlo procedure must eliminate all those. It means that the signal to noise ratio of such a calculation is decreasing as

$$e^{-\beta\Delta E} \tag{8.101}$$

where β is the inverse temperature and ΔE is the difference in energy between a symmetric ground state and an antisymmetric one, since this corresponds roughly to the probability to draw an antisymmetric configuration (useful) among all the other ones. This number is usually quite large. For example for free bosons the minimum of energy is zero (all bosons in $k = 0$ momentum state, while for fermions it is $E_{\rm F}$. It means that the error of the Monte-Carlo procedure grows exponentially fast when lowering the temperature, needing exponentially large computational time to compensate. For realistic numbers (e.g. $E_{\rm F} = 10000K$ and $T = 100$ this is humongously long times, beyond the actual possibilities. This is something known as the sign problem, which limits the possibility to use these methods for fermions. Nevertheless depending on the temperature one can obtain, just as for classical problems, extremely useful information from such a procedure.

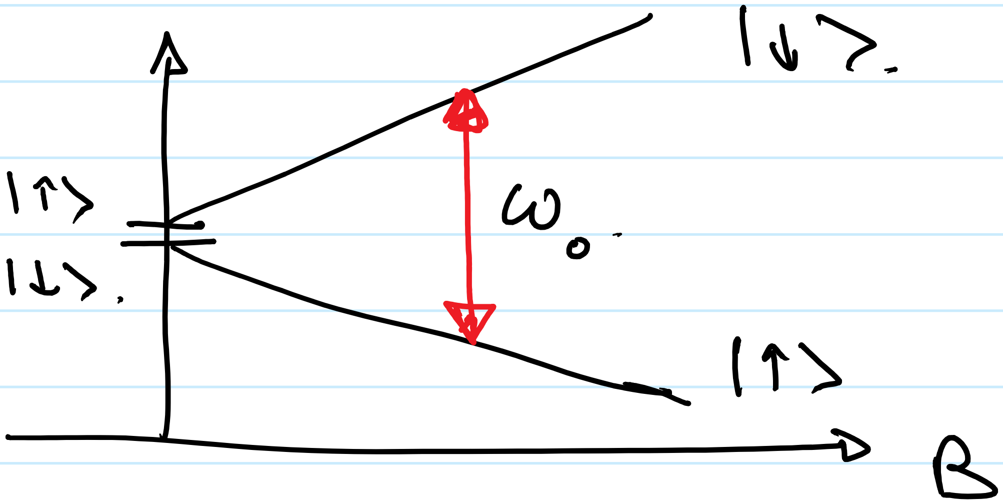


Figure 8.13: Principle of the NMR method. In an external magnetic field it is possible to induce rabi oscillations between the two level of a nuclear spin. Measurement of the spitting frequency gives information on the local magnetic field.

8.5 Experimental probes

Let us now briefly look at some experimental probes for the localized magnetism. We will not discuss the macroscopic ones, but rather focuss on the microscopic ones able to give information at the level of the lattice spacing. We will mostly concentrate on two: the Nuclear Magnetic Resonance (NMR) and the neutron scattering.

8.5.1 Nuclear Magnetic Resonance

We will not enter in details but just give the main results. We refer the reader to [Abr61, Sli89] for more details.

The basic idea is to use the presence of a nuclear spin. Such a spin is weakly coupled to the magnetic fields in the material and in particular the ones created by the electronic spins, which make it an ideal probe for magnetism in solids. Let us for example consider a nuclear spin which would be a spin $1/2$. In that case, as indicated in Fig. 8.14 under an external field the two spin level would split. It is then possible to perform Rabi oscillations between these two levels by sending microwaves on the system. Typically this occurs at frequencies which are of the order of $1/2$ GHz for external field of the order of $10T$, depending on the material. If the two levels are separated enough, the thermal population of the upper level is small enough so that the population is determined by the rabi oscillations. The measure of the oscillation frequency is directly giving the field which is acting on the nuclear spin. Such as field is the sum of the externally applied field B_{ext} and the local static field existing in the system and which is influencing the nuclear spin. For sake of simplicity let us assume that only magnetic field created on the same atom can influence the nuclear spin, the measurement of the resonance frequency is thus a direct measure of the static *internal* magnetic field which exists in a solid. Some cases are shown in Fig. **??**. If the material is paramagnetic then no local static field exists. Then the resonnance occurs at the place one would expect for an isolated nuclear spin. On the contrary if the material is ferromagnetic, then each atom sees an additional magnetic field, proportional to the order parameter. The resonance is thus shifted compared to the paramagnetic case. This shift, known as Knight shift, is a direct measure of the local field. In that case it is directly proportional to the uniform magnetization. It is thus easy with NMR to extract, e.g. the temperature dependence of the magnetization. If the material is antiferromagnetic then half of the atoms will see a positive field, half a negative field. The line will thus split and the splitting is proportional to the staggered magnetization. It is thus very convenient to extract such order parameters from the NMR. An example is shown in Fig. 8.15.

In addition to the Knight shift one can extract other information from the NMR signal. One useful information is the so-called relaxation time $1/T_1$. Let us assume that the system starts in the lowest level $|\uparrow\rangle$ as shown in Fig. 8.15. One can then apply a Rabi pulse that leaves the system in an intermediate state $|\uparrow\rangle + |\downarrow\rangle$ by timing the time of the pulse (so called $\pi/2$ pulses). The nuclear system would then precess in the XY plane and thus have no S^z component any more. However, if electronic spins can flip an interaction of the form

$$H = \boldsymbol{S}_{el} \cdot \boldsymbol{S}_{nuc} \tag{8.102}$$

would allow to let the nuclear spin relax towards the polarized most stable configuration. One should thus have

$$M_{nuc}(t) = M_{nuc}^{max}[1 - e^{-t/T_1}] \tag{8.103}$$

the characteristic time $1/T_1$ will tell how many electronic spin flips can occur by unit time. One can show that

$$\frac{1}{T_1} \propto T \frac{1}{\omega} \operatorname{Im}(-i \int_{t=0}^{\infty} e^{i\omega t} \langle [S^+(r=0,t), S^-(r=0,0)] \rangle \tag{8.104}$$

In the above equation the last term is the retarded transverse spin-spin correlation function. The fact that it is local in space comes from our approximation that only electronic spins on the same atom can influence the nuclear spin. This formula should be taken at the frequency which corresponds to the NMR frequency. However since we are dealing here with the electronic spins, and that the NMR frequencies are usually only in the MHz range, it is a very good approximation to simply replace the frequency by zero. This gives the canonical formula for the relaxation time

$$\frac{1}{T_1} = T \lim_{\omega \to 0} \frac{\operatorname{Im} \chi_{+-}(r=0,\omega)}{\omega} \tag{8.105}$$

The $1/T_1$ thus gives information on how the electronic spins are correlated in time. We will examine several examples in the exercises.

8.5.2 Neutron diffraction

Another important probe for magnetism is provided by neutron diffraction. The idea is exactly similar to the one of X-ray diffraction to probe the nature and periodicity of crystals. This time we want a probe which is not sensitive to the charge of the electrons but to their spin. The perfect particle/wave to use for that is the neutron. Being neutral it will not be sensitive to the charge distributions inside the solids, but having a spin 1/2 will be diffracted by the electronic (and nuclear) spins.

The principle of the measurement is simple and shown in Fig. 8.16. One sends a beam of neutrons of known momentum (and thus energy since it is in the vaccuum) and also potentially of known spin orientation. These neutrons enter the solid and interact with the electronic spins. As a result of this interaction they change their quantum numbers. One measures the neutrons

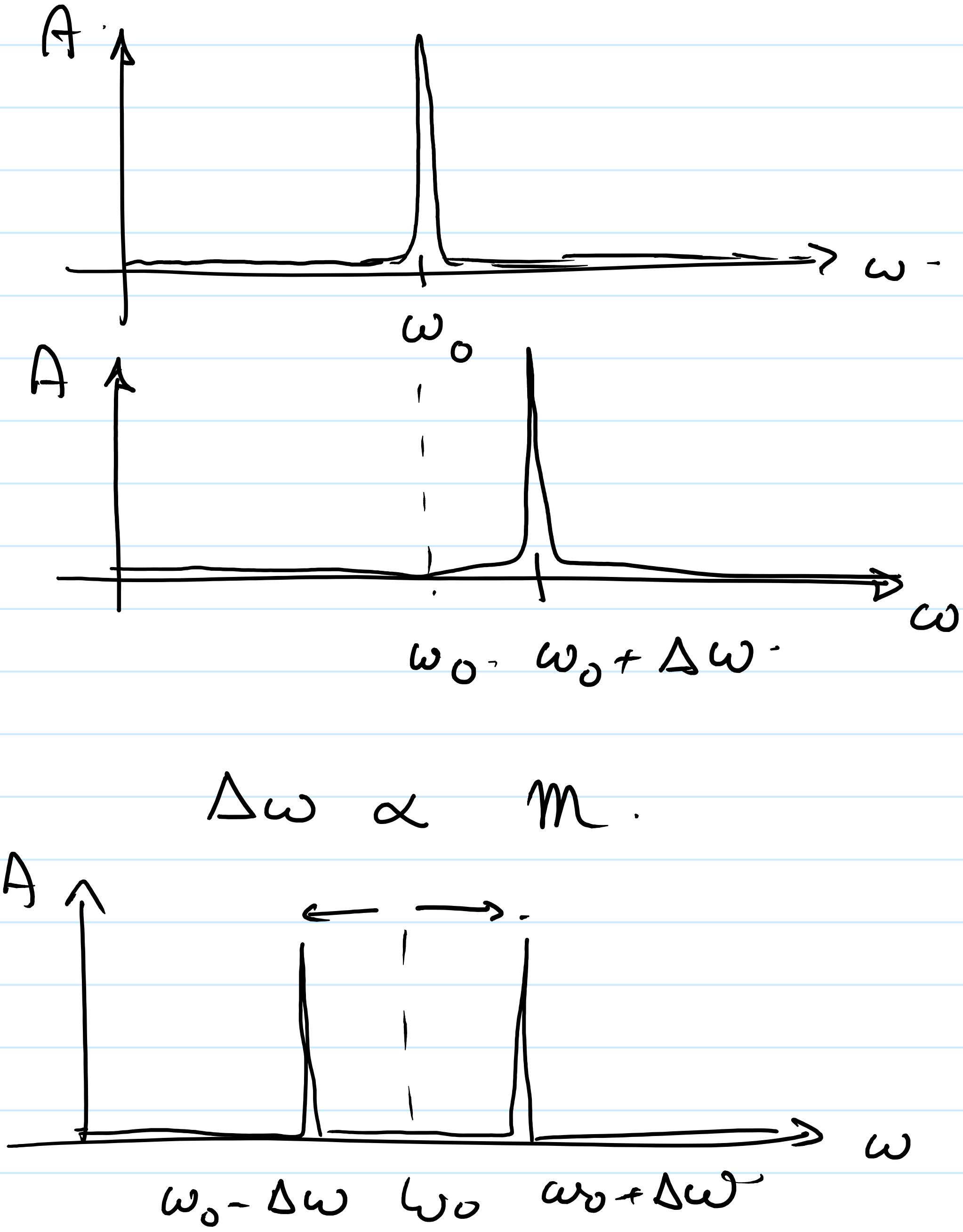


Figure 8.14: Knight shift in NMR. a) the material is paramagnetic, the resonance occurs at the expected place. b) the material is ferromagnetic. Each atom sees the same magnetic field, then the resonance is shifted. The shift is proportional to the magnetization. c) the material is antiferromagnetic. In that case the line splits depending on whether the local magnetic field is positive or negative. The splitting is proportional to the staggered magnetization.

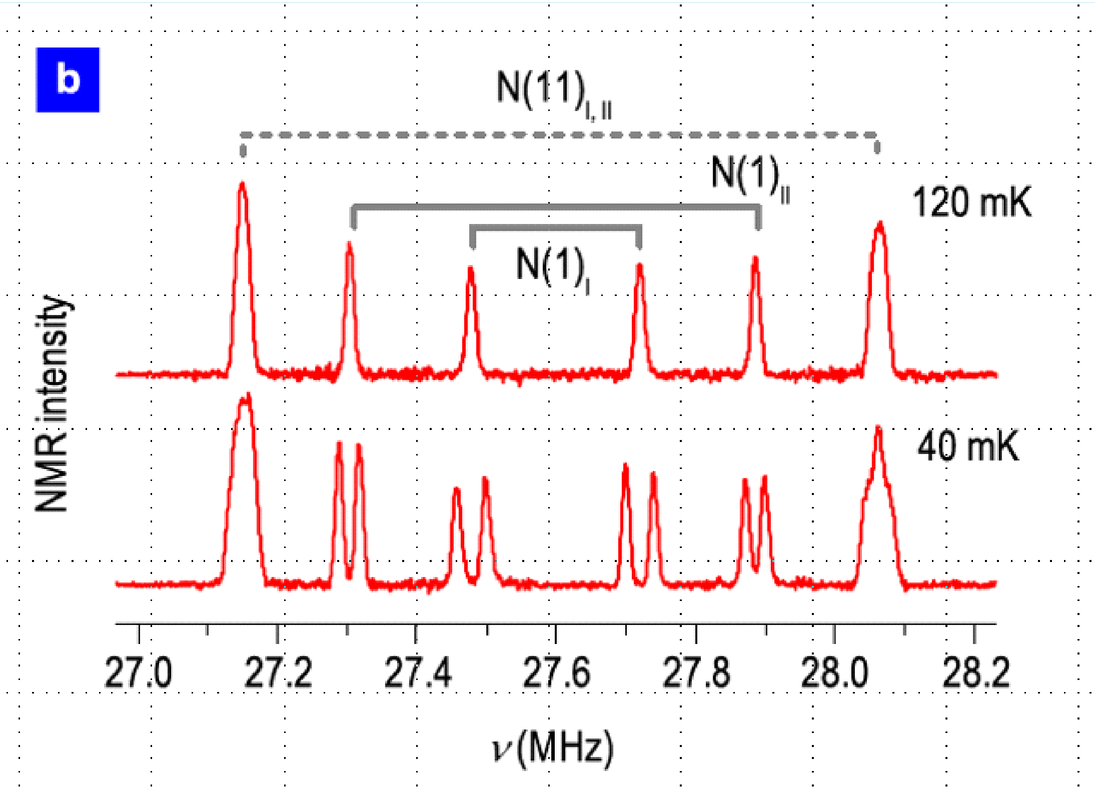


Figure 8.15: An example of NMR signal on the various N sites of the ladder compound BPCB discussed previously. One sees clearly the splitting of the lines at low temperatures indicating the presence of an antiferromagnetic phase transition. From [KMB$^+$08].

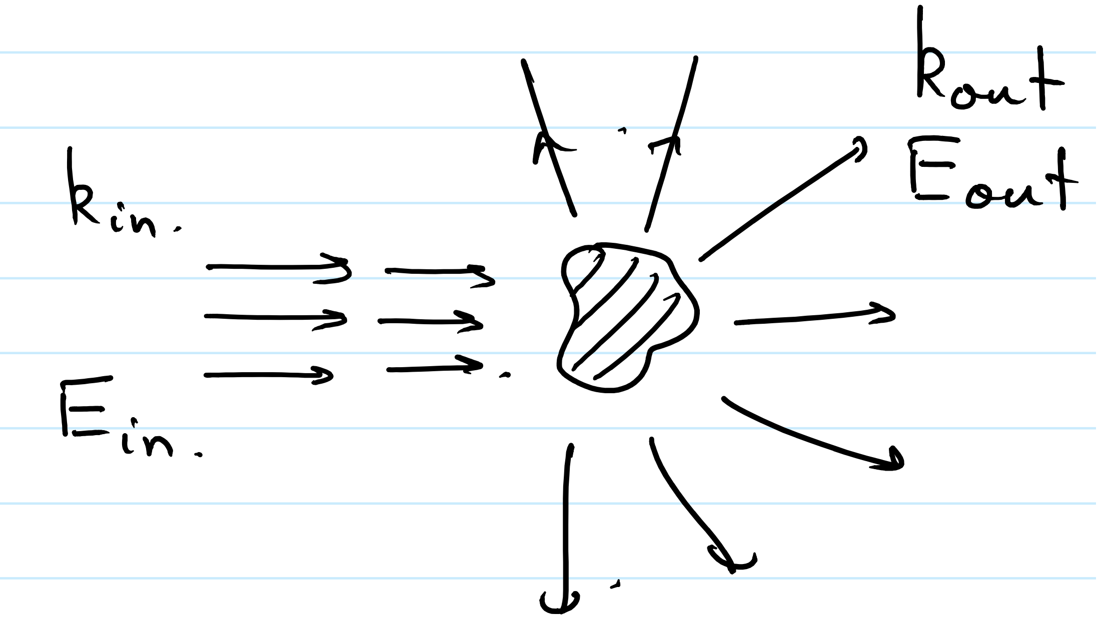


Figure 8.16: Principle of neutron diffraction. A monochromatic neutron beam of known spin orientation is sent to the target. Neutrons diffracted in a certain direction, with a certain spin and a certain energy is measured. This gives the scattering cross section for a change of momentum, energy and spin of the neutrons that gives information on the corresponding excitations in the solid.

that are diffracted with a given momentum (usually given from the diffraction angle), a given energy and a given spin orientation. This is giving the cross section of diffraction by the solid. The change of quantum numbers of the neutrons reflects which excitations they have created in the solid and thus give information on the magnetism.

The scattering cross section can be computed from the Fermi golden rule. It is proportional (up to terms of incident flux etc.) to the probability of transition

$$P(\boldsymbol{k}_{in} \to \boldsymbol{k}_{out}, E^n_{in} \to E^n_{out}, \boldsymbol{S}^n_{in} \to \boldsymbol{S}^n_{out}) = \sum_{f,i} |\langle f|H|i\rangle|^2 \delta(E^n_{out} - E^n_{in} + E^{sol}_{out} - E^{sol}_{in}) \quad (8.106)$$

where $|f\rangle$ and $|i\rangle$ are the final and initial state of the combined neutron-solid system. The interaction Hamiltonian between the neutron and the spins in the solid is simply the dipolar one

$$H = \int d\boldsymbol{R} d\boldsymbol{r} \frac{\boldsymbol{S}^n(\boldsymbol{R}+\boldsymbol{r}) \cdot \boldsymbol{S}(\boldsymbol{R})}{\boldsymbol{r}^3} - \frac{\boldsymbol{S}^n(\boldsymbol{R}+\boldsymbol{r}) \cdot \boldsymbol{r})(\boldsymbol{S}(r_2) \cdot \boldsymbol{r}}{(\boldsymbol{r}^5} \quad (8.107)$$

which can be put in a diagonal form by a Fourier transform

$$H = \sum_{\alpha\beta} \int d\boldsymbol{q} A^{\alpha\beta}[\boldsymbol{q}] S^n_\alpha(\boldsymbol{q}) S_\beta(-\boldsymbol{q}) \quad (8.108)$$

where $\alpha, \beta = x, y, z$ are the three spin components. We will not explicit the form factor $A^{\alpha\beta}[\boldsymbol{q}]$ which is easily computable.

To compute (8.106) we realize that the initial and final state can be factorized into a part that contains only the solid and one only the neutrons The matrix element can thus be simplified considerably

$$\langle f|H|i\rangle = \sum_{\alpha\beta} \int d\boldsymbol{q} A^{\alpha\beta}[\boldsymbol{q}] \langle f^n|S^n_\alpha(\boldsymbol{q})|i^n\rangle \langle f^{sol}|S_\beta(-\boldsymbol{q})|i^{sol}\rangle \quad (8.109)$$

since we know the initial and final state of the neutron momentum we have immediately $\boldsymbol{q} = \boldsymbol{k}_f - \boldsymbol{k}_i$ which fixes fully the momentum. In the same way knowing the initial and final state of the spin of the neutron fixes essentially α and thus the corresponding sums that must be performed on β. We will assume here for simplicity (not true in general) that the form factor is diagonal and only retain one value for β. In general several values are possible and must be summed accordingly. This is however a purely technical complication. The probability of transition would thus become

$$P(\boldsymbol{k}_{in} \to \boldsymbol{k}_{in} - \boldsymbol{q}, E^n_{in} \to E^n_{in} - \omega, \alpha) = \sum_{f^{sol}} A^{\alpha\alpha}[\boldsymbol{q}] |\langle \nu^{sol}|S^\alpha(\boldsymbol{q})|GS\rangle|^2 \delta(\omega + E_{GS} - E_\nu) \quad (8.110)$$

where $|GS\rangle$ is the ground state of the solid. $|\nu\rangle$ is one of the excited states of energy E_ν, and $\alpha = x, y, z$ one of the spins orientations. The physical interpretation of the formula (8.110) is quite transparent. The neutrons deposit a momentum $\boldsymbol{q}$ and an energy ω in the solid. In addition they had to create a spin excitation of component α, since this corresponds to the difference of quantum numbers measured for the outgoing neutron. It is thus necessary to create in the solid a spin excitation of momentum $\boldsymbol{q}$ and of component α. The probability of transition is thus given by the probability of creating such an excitation. The amplitude is given by the matrix element of transition from the ground state to an excited state $\langle \nu^{sol}|S^\alpha(\boldsymbol{q})|GS\rangle$ whose energy matches the energy conservation. Finally as with any diffraction experiment the form factor of the potential A modifies the final result. This last point must be taken into account to compare with experiment but is not fundamental. The important part is the previous one. The neutron are thus measureing the probability to create a spin excitation with a given momentum and given energy. This can simply be related to the retarded spin-spin correlations

$$\chi(\alpha, \beta, \boldsymbol{q}, \omega) = -i \int_{t=0}^{\infty} dt e^{i\omega t} \langle [S^\alpha(-\boldsymbol{q}, t), S^\beta(\boldsymbol{q}, 0)] \rangle \quad (8.111)$$

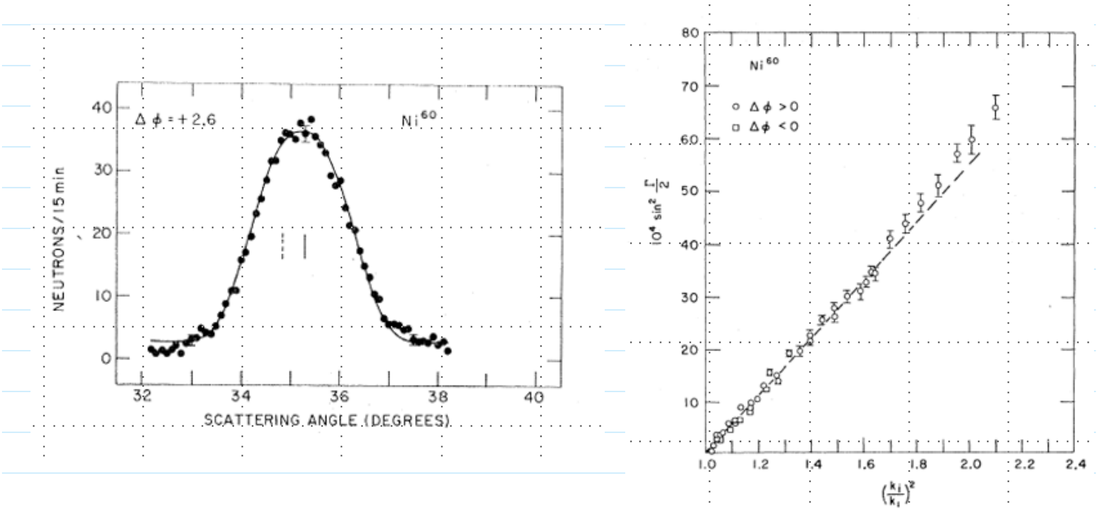


Figure 8.17: Neutron scattering on ferromagnetic nickel. left: peak of diffraction obtained by varying the angle of the detector and thus momentum. right: energy of the peak as a function of the momentum at which the maximum occurs. One see that the relation is quadratic, which is consistent with the existence of the sharply defined magnon excitations in a ferromagnet. Note on the left the neutron count, and gasp at what must have been the signal to noise ratio and experimental time then. From [PAM$^+$67].

We thus see that neutron and NMR measure essentially the same objects. The main difference is that the NMR measures the local correlation (hence the sum over all $\boldsymbol{q}$) at zero frequency, while the neutrons have in principle both momentum and frequency resolution. However the resolution is not as good in the later case and it is difficult to reach below $0.2meV$ with neutrons. Another limitation is that it is usually inconvenient (for reasons of flux and signal to noise ratio) to polarize the incoming neutron beam. It means that an average of all spin orientations is present, which means that one also has to average over all the correlation functions. This is not necessarily a very serious limitation for systems for which the correlation functions can be obtained by a symmetry operation.

As can be seen from (8.110) the neutrons will tell us where in energy-momentum space the spin excitations can live. If there are sharply defined excitations (a good relation between the momentum and energy) then the δ-function constraint on the conservation of the energy can only be satisfied for well given pairs of momenta-energy. We thus expect in that case a narrow peak positioned on the dispersion relation. The weight in the peak will depend on the matrix elements, but the position of the peak itself is a direct measure of the dispersion relation. For more complex case, for which there are no sharply defined spin excitations one can have a continuum of absorption whose intensity will be modulated by the form factors.

Let us see two examples in connection with our study of ferromagnetism and antiferromagnetism. The first spectrum shown in Fig. 8.18 we see peaks which signal sharp magnetic excitations. Their dispersion relation which is quadratic in momentum is perfectly consistent with the magnons in a ferromagnet.

The second curves shown in Fig. **??** This time the energy shows minima in other positions than $\boldsymbol{q} = (0, 0)$ the position of the minima at (π, π) indicate and antiferromagnetic ordering. In addition the dispersion shows clearly the linear dispersion that one expects for the spin

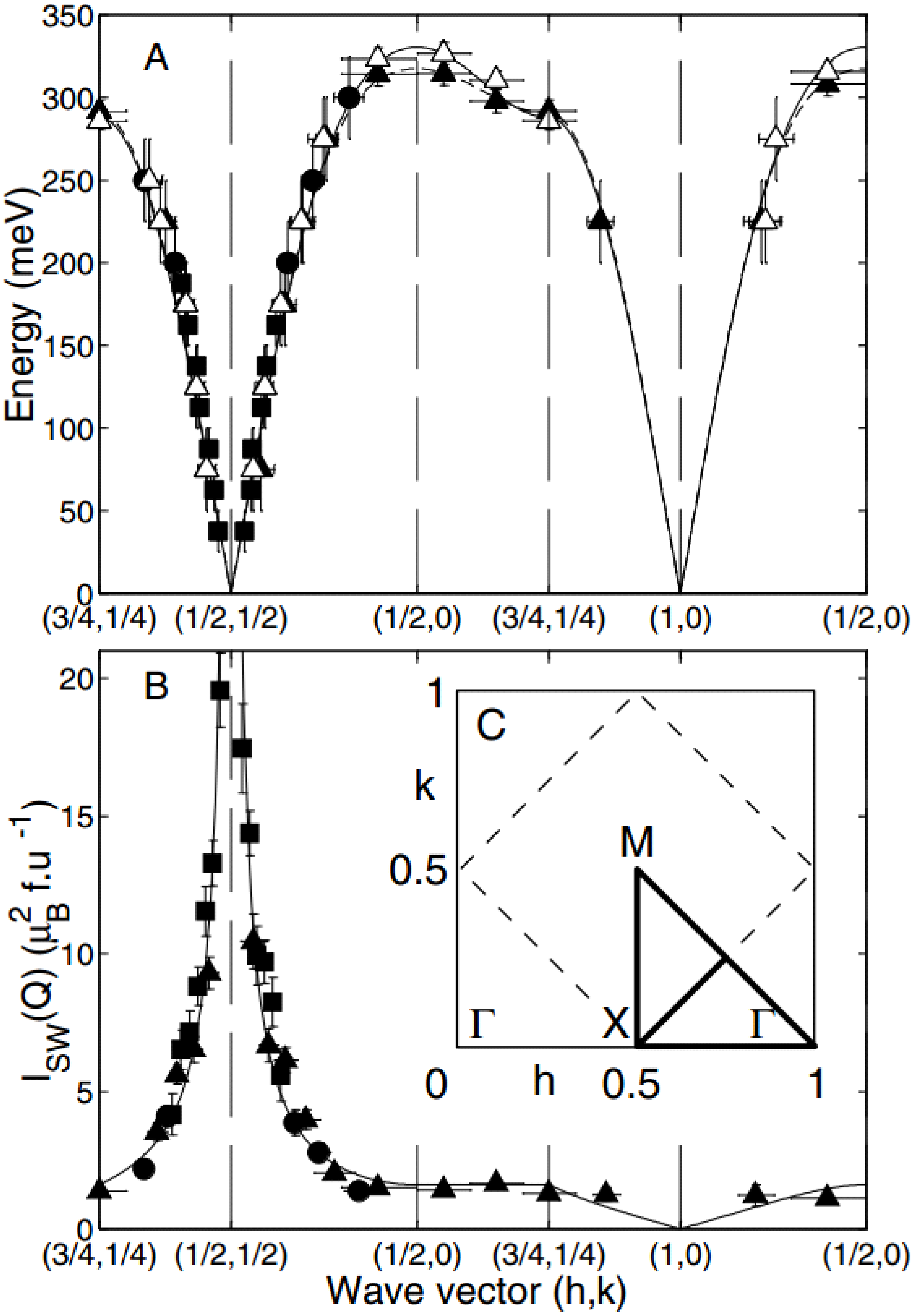


Figure 8.18: Neutron diffraction in the La_2CuO_4 two dimensional system. One sees clearly the zero of the energy at (π, π) as well as the linear dispersion relation compatible with the existence of sharply defined magnetic excitations in an antiferromagnet. From [CHA+01].

excitations inside an antiferromagnet.

Part III

Numerical methods

CHAPTER 9

Monte-Carlo method

The goal of this chapter is to introduce some of the numerical methods that are useful in dealing with interacting quantum systems. Before embarking on the more refined techniques for such systems, let us first consider classical systems and illustrate on them one of the most powerful numerical techniques, the Monte-Carlo technique. A good description of such algorithms can be found in [Kra06].

9.1 Purpose and basic idea

Let us consider the calculation of an integral $I = \int_a^b dx f(x)$. One efficient numerical method is to approximate the integral by the Riemann sum

$$I_N = \sum_{i=0}^{N-1} \left(\frac{b-a}{N}\right) f(x_i = a + i\frac{b-a}{N}) \tag{9.1}$$

If the function is a smooth one, even this crude approximation is excellent since the error is of order

$$|I_N - I| \sim \frac{1}{N} \tag{9.2}$$

Of course more refined versions (middle point sums etc.) can even improve on this result.

Problems with this way of evaluating integrals appear when the dimension of space increases. For example if one has to evaluate a two dimensional integral then one would cover the interval with a grid of $N = N_x^2$ points, and then sum the discrete elements on the grid

$$I_N = \sum i = 0, N - 1 \Delta S f(r_i) \tag{9.3}$$

where ΔS is the surface of one little square in the grid and r_i is one point inside such a little square. Unfortunately, as can be readily seen by taking a function $f(x, y)$ that is strictly constant along e.g. y, the error that is made on the integral is a function of the size of the grid along one of the directions. In other words

$$|I_N - I| \sim \frac{1}{N_x} = \frac{1}{\sqrt{N}} \tag{9.4}$$

and of course the generalization to an integral in d dimensions is obvious

$$|I_N - I| \sim \frac{1}{N_x} = \frac{1}{N^{1/d}} \tag{9.5}$$

It is thus easy to see that if one has to evaluate an integral by the Riemann method in a space of dimension, say 100 and one would like an reasonable error of 10% one would need a number of points in the grid to be 10^100 an astronomical number and it would take forever to compute the integral in such a deterministic way. Unfortunately such large dimensional integrals are the rule when one is dealing with statistical systems. Let us for example consider an evaluation of the partition function of the Ising model. It is given by

$$Z = \sum \sigma_1 = \pm, \sigma_2 = \pm, \ldots, \sigma_N = \pm e^{-\beta H[\{\sigma\}]} \tag{9.6}$$

where N is the number of spins σ in the system and the Hamiltonian is

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j \tag{9.7}$$

If one wants to deal with a system that is reasonably close to a a thermodynamic limit one needs a "large" number N of spins, which means that partition function can be viewed as a result of a sum, in a space of very large dimension (here N). It is thus totally unrealistic to try to evaluate this sum directly by enumerating all the configurations. This is also the case if one wants to evaluate one physical observable. For example the average of the spin on one site i_0 would be given by

$$\langle \sigma_{i_0} \rangle = \frac{1}{Z} \sum_{\sigma_1=\pm,\sigma_2=\pm,\ldots,\sigma_N=\pm} \sigma_{i_0} e^{-\beta H[\{\sigma\}]} \tag{9.8}$$

A smarter way to evaluate such an integral would be to evaluate the integral around the points where the function is the largest. This would ensure that with a reasonable number of points one has an a good evaluation of the integral. Of course finding such regions is not so easy and is the purpose of the Monte-Carlo technique that we now describe.

9.2 Monte-Carlo technique: basic principle

Let us assume that we want to evaluate a sum or an integral (we will use the integral notations, but of course this covers discrete sums as well). Let us take for example

$$\int_a^b dx f(x) \tag{9.9}$$

We introduce a probability weight $P(x)$ which is positive over the interval. One can always rewrite the integral as

$$\int_a^b dx P(x) \frac{f(x)}{P(x)} \tag{9.10}$$

Given that $P(x)$ is a probability it satisfies

$$\int_a^b dx P(x) = 1 \tag{9.11}$$

At that stage the choice of $P(x)$ is totally arbitrary. If one considers the Ising model average (9.8), one possible choice would be to take

$$P(x) \frac{1}{Z} e^{-\beta H[\{\sigma\}]} \tag{9.12}$$

since in that case $f(x)/P(x)$ is simply the spin σ_{i_0}. But of course other choices are in principle possible.

With such a probability distribution, let us assume that we have a random number generator that can generate random numbers with the probability $P(x)$. In that case, if one generates N random numbers x_i with the distribution $P(x)$

$$I_N = \frac{1}{N} \sum i = 0^{N_1} \frac{f(x_i)}{P(x_i)} \tag{9.13}$$

is a random variable. Its average is given by

$$\begin{aligned} \overline{I_N} &= \frac{1}{N} \sum i = 0^{N_1} \overline{\frac{f(x_i)}{P(x_i)}} \\ &= \frac{1}{N} \sum i = 0^{N_1} \int_a^b dx P(x) \frac{f(x)}{P(x)} \\ &= \int_a^b dx P(x) \frac{f(x)}{P(x)} \end{aligned} \tag{9.14}$$

The average of the random variable I_N is thus exactly the integral I that we want to evaluate. Even more interesting is the distribution itself of this random variable I_N. It is made by the sum of N random variables, for whom the distribution is difficult to evaluate. This is however not important since the central limit theorem tells us that regardless of the distribution of the random variable $f(x_i)/P(x_i)$, if we sum many of such variables they will have a gaussian distribution. Let us briefly derive this result again without any purpose of mathematical rigor. If we have variables x_i with the probability distribution $P(x)$, let us compute the probability distribution $\mathcal{P}(X)$ where $X = \frac{1}{N}\sum_{i=1}^N x_i$. For simplicity we assume that the x_i are of zero average $\overline{x_i} = 0$. The probability distribution $\mathcal{P}(X)$ is given by

$$\begin{aligned} \mathcal{P}(X) &= \prod_i \int dx_i P(x_i) \delta(X - \frac{1}{N}\sum_{i=1}^N x_i) \\ &= \int \frac{d\lambda}{2\pi} \prod_i \int dx_i P(x_i) e^{i\lambda(X - \frac{1}{N}\sum_{i=1}^N x_i)} \\ &= \int \frac{d\lambda}{2\pi} e^{i\lambda X} \prod_i \left(\int dx_i P(x_i) e^{-i\frac{\lambda}{N} x_i} \right) \\ &= \int \frac{d\lambda}{2\pi} e^{i\lambda X} \left(\int dx P(x) e^{-i\frac{\lambda}{N} x} \right)^N \end{aligned} \tag{9.15}$$

One has thus the Fourier transform $P(q = \frac{\lambda}{N})$ of the distribution function, at very small argument λ/N if the number of terms in the sum is large. One can thus expand the Fourier transform around $q = 0$. Since $\int dx P(x) = P(q = 0) = 1$, and assuming the function $P(x)$ is even to ensure that $\overline{x} = 0$ one has

$$\begin{aligned} P(q = \frac{\lambda}{N})^N &= \left[1 - \left(\frac{\lambda}{2N} \right)^2 P''(0) \right]^N \\ &= e^{-\frac{\lambda^2}{2N} P''(0)} \end{aligned} \tag{9.16}$$

Inserting in (9.15) one gets

$$\mathcal{P}(X) = \int \frac{d\lambda}{2\pi} e^{i\lambda X - \frac{\lambda^2}{2N} P''(0)} = \sqrt{\frac{N}{2\pi P''(0)}} e^{-\frac{N X^2}{2P''(0)}} \tag{9.17}$$

As can be seen the standard deviation of around the average value is decreasing as $1/\sqrt{N}$, namely

$$\frac{\sqrt{\overline{[I_N - I]^2}}}{I} \sim \frac{1}{\sqrt{N}} \tag{9.18}$$

In practice this means that the probability that the difference between the exact result I and the approximate result I_N can be larger than $1/\sqrt{N}$ is quite small. One can thus consider as rule of thumb consider that approximate sum I_N is getting the value of the integral with an error of the order of $\pm 1/\sqrt{N}$. Note that such an error is independent of the dimension of the integral or the sum. This statistic method of evaluating integrals thus gives an error which is much less good than the simple Riemann sums for low dimensional integral but which will be vastly superior when the dimension of the integral will be large. One can control de accuracy of the result by getting more terms in the sum.

Of course one of the main question is to generate the random numbers with the probability $P(x)$. In general it is easy to find pseudo-random number generators with flat distributions, but getting numbers with an arbitrary distribution is exactly the point we examine in the next section.

9.3 Markov's chains

In order to determine the algorithm to compute our random numbers let us introduce the concept of Markov's chains.

9.3.1 Definition

Let us give ourselves a random variable X_t depending on an index t and an evolution law which gives the variable X_{t+1} if we know the variable X_t. This defines a set of random variables known as a Markov chain.

Note that the evolution law can be probabilistic as well. Loosely speaking the index t can be viewed as a kind of discrete time. To define the evolution law we give $P(X \to Y)$ the probability to go from the variable X to the variable Y. Since P are probabilities one has $0 \leq P(X \to Y) \leq 1$ and

$$\sum_Y P(X \to Y) = 1 \tag{9.19}$$

where the sum can be replaced by an integral if the variables are continuous. If we call $\pi_t(X)$ the probability to find the value X for the random variable in the Markov chain at "time" t then one must have

$$\pi_{t+1}(Y) = \sum_X \pi_t(X) P(X \to Y) \tag{9.20}$$

In the following we will restrict to a special class of Markov chains, which obey a more restrictive condition known as detailed balance condition, which reads (for all X, Y)

$$\pi_t(X) P(X \to Y) = \pi_{t+1}(Y) P(Y \to X) \tag{9.21}$$

This condition expresses the fact that in the evolution the flow of probability going from X to Y is compensated by the opposite flow. This let us expect that a steady state can easily be reached by such Markov chains.

9.3.2 Steady state

Let us now examine whether the Markov chain can reach a steady state. This means that the probability distribution $\pi_t(X)$ will converge to a steady state probability distribution $\mathcal{P}(X)$ as the time is growing.

Let us assume that such a distribution does exist. If this is the case it must obey the detailed balance condition (9.21) which becomes

$$\mathcal{P}(X)P(X \to Y) = \mathcal{P}(Y)P(Y \to X) \tag{9.22}$$

Let us thus show that a distribution that obeys such a condition corresponds indeed to a steady state distribution for the Markov chain. If the distribution at time t is given by $\mathcal{P}(X)$ then the evolution condition (9.20) gives us for the distribution at time $t+1$

$$\pi_{t+1}(Y) = \sum_X \mathcal{P}(X)P(X \to Y) \tag{9.23}$$

Using the condition (9.22) this becomes

$$\pi_{t+1}(Y) = \sum_X \mathcal{P}(X)P(X \to Y) = \sum_X \mathcal{P}(Y)P(Y \to X) = \mathcal{P}(Y) \tag{9.24}$$

which proves that the distribution $\mathcal{P}(X)$ that obeys the condition (9.22) is indeed a steady state of the Markov chain. It also means that we can determine the steady state of the Markov chain by simply looking at the condition (9.22).

Of course showing that such a steady state distribution exists, and we also have to show that the probability distribution $\pi_t(X)$ does converge to it. For this let us define a "distance" to such a steady state. Several choices are possible but let us take for simplicity the distance $\mathcal{D}$ defined as

$$\mathcal{D}_t = \sum_X |\pi_t(X) - \mathcal{P}(X)| \tag{9.25}$$

Let us now examine the distance at time $t+1$. It is given by

$$\mathcal{D}_{t+1} = \sum_X |\pi_{t+1}(X) - \mathcal{P}(X)| \tag{9.26}$$

By definition of the evolution and the fact that $\mathcal{P}$ is a steady state distribution one has

$$\begin{aligned} \pi_{t+1}(X) &= \sum_X \pi_t(X)P(X \to Y) \\ \mathcal{P}(X) &= \sum_X \mathcal{P}(X)P(X \to Y) \end{aligned} \tag{9.27}$$

which inserted in (9.26) give

$$\begin{aligned} \mathcal{D}_{t+1} &= \sum_X |\sum_Z (\pi_t(Z) - \mathcal{P}(Z))P(Z \to X)| \\ &\leq \sum_X \sum_Z |(\pi_t(Z) - \mathcal{P}(Z))|P(Z \to X) \\ &\leq \sum_Z |(\pi_t(Z) - \mathcal{P}(Z))| \\ &\leq \mathcal{D}_t \end{aligned} \tag{9.28}$$

the distance is thus reducing at each step of the Markov chain. The distribution $\pi_t(X)$ will thus converge to the steady state distribution that obeys the detailed balance condition (9.22).

We now have a practical algorithm to generate random numbers with *any* distribution $\mathcal{P}(X)$. The algorithm goes as follows:

1. We choose an evolution rule $P(X \to Y)$ such that the condition (9.22) is fulfilled. For a given distribution $\mathcal{P}(X)$ there can be of course many choices for such evolution.
2. We initialize the Markov chain with an arbitrary initial condition and let it evolve according to the evolution rule $P(X \to Y)$ chosen above.
3. After a certain time t_0 the probability distribution $\pi_t(X)$ of the Markov chain should be close enough to the distribution we want $\mathcal{P}(X)$. It means that the random numbers that are generated by the evolution of the Markov chain will be generated according to the distribution $\mathcal{P}(X)$. We can thus use these numbers.

Of course in practice there are a lot of subtle points that can influence the quality and speed with which one can obtain the result with a given accuracy: the initial value of the Markov chain, how many configurations are discarded initially etc.

9.4 Metropolis algorithm and simple example

Let us illustrate the above concepts on the very simple example of the calculation of

$$A_p = \frac{\sum_n n^p e^{-n^2}}{\sum_n e^{-n^2}} \tag{9.29}$$

where p is a fixed integer and n runs on all the relative integer numbers. Of course doing such a simple sum using the Monte-Carlo technique is not the most efficient way to do it, but this example is a baby version of the kind of calculations one would have to do in the more realistic case of the Ising model.

In order to evaluate A_p we can use the formula (9.10) with the probability

$$P(n) = \frac{e^{-n^2}}{\sum_n e^{-n^2}} \tag{9.30}$$

It means we have to generate integer random numbers with this probability. Once we have a set $n_1, n_2, \ldots, n_N$ of such numbers we can use them to compute

$$S_p = \frac{1}{N}\sum_{i=1}^{N} n_i^p \tag{9.31}$$

which will be an approximation of the sum A_p.

Let us thus create our Markov chain. We introduce at each "time" a number n_t that will be the variable. We can start with any value for this number. Clearly the terms with weight in the sum are for numbers not too different from zero so it might be a good idea to start with $n_0 = 0$ but this is absolutely not mandatory. Then we have to make sure that we have an evolution that can explore *all* the possible configurations, otherwise there would be terms in our sum that we would simply and totally miss. There are of course many ways to ensure that, so let us just choose one. We will first decide wether the system wants to increase or decrease n_t by one unity. For that we can use a uniform random number generator to decide on which direction

($n_t + 1$ or $n_t - 1$) the system would like to move. We generate a random number uniformly distributed between say $x \in [0, 1[$ and then move on the left if $0 \leq x < 1/2$ and on the right otherwise. Clearly such moves allow to reach any value of n after a certain number of steps.

To decide on what is the evolution law between n_t and n_{t+1} we want it to obey (9.22), namely

$$e^{-n_t^2} P(n_t \to n_{t+1}) = e^{-n_{t+1}^2} P(n_{t+1} \to n_t) \tag{9.32}$$

Note that the normalization condition of the probability disappear, which simplifies considerably the condition. There are several ways to obey this condition. We will give here the one that was introduced by Metropolis et al. Let us assume that we make a move from a configuration X to a configuration Y, with their respective probabilities $\pi(X)$ and $\pi(Y)$. There are two possibilities

1. $\pi(Y) > \pi(X)$: the new configuration is more probable. In that case one makes systematically the move. In other words $P(X \to Y) = 1$ in that case.

2. $\pi(Y) < \pi(X)$: the new configuration is less probably than the old one. In that case one makes the move with the probability

$$P(X \to Y) = \frac{\pi(Y)}{\pi(X)} \tag{9.33}$$

It is easy to see that this choice of evolution directly satisfy the condition

$$\pi(X)P(X \to Y) = \pi(Y)P(Y \to X) \tag{9.34}$$

In practice one performs the move in the following way. One computes the ratio of the probabilities $r = \pi(Y)/\pi(X)$. If it is greater than one the move is done. If not, one draws again a random number x uniformly in the interval $x \in [0, 1[$, and one compares the two numbers r and x. The move is done if $x \leq r$, which indeed corresponds to a probability r.

The appendix Appendix B give an extremely simple and non-optimized version of a program implementing this calculation.

A more sophisticated example is to look at the Ising model. The Hamiltonian is

$$H = -J \sum_{i,z} \sigma_i \sigma_{i+z} \tag{9.35}$$

where z is the set of vectors of nearest neighbors. and one wants to compute observables such as the magnetization on one site which is given by

$$\langle \sigma_{i_0} \rangle = \frac{1}{Z} \sum \{\sigma_i\} \sigma_{i_0} e^{-\beta H} \tag{9.36}$$

One can compute the average by a monte-carlo technique. The idea is a simple generalization of the above simple example

1. Pick a spin at random

2. Flip it and compute the difference of energy between the two configurations. Note that this difference of energy only involves the spin and its neighbors and not the whole lattice. It can thus be evaluated very quickly.

3. if the move is energetically favorable, do it, if not do it with the proper probability as defined by the Metropolis algorithm

4. use these configurations to compute the sum $\frac{1}{T} \sum_{t=1} T \sigma_{i_0}(t)$.

9.5 Variational monte-Carlo

One can directly apply the above concepts to the quantum world. The simplest is the evaluation of variational wavefunctions. This is known as variational monte-carlo. We will just describe here the basic ideas. Let us imagine that one gives oneself a wavefunction $|\psi\rangle$. One example is the Gutzwiller wavefunction (7.14). One needs to compute the average of some operator, for example the Hamiltonian

$$E = \langle\psi| H |\psi\rangle \tag{9.37}$$

In order to do it let us introduce a complete basis $|\alpha\rangle$ of the states of the system. A convenient basis is for example the position of all the particles. In that case the above expressions can be rewritten as

$$E = \sum_\alpha \langle\psi|\alpha\rangle\langle\alpha|H|\psi\rangle = \sum_\alpha |\langle\psi|\alpha\rangle|^2 \frac{\langle\alpha|H|\psi\rangle}{\langle\alpha|\psi\rangle} \tag{9.38}$$

One thus sees that it is perfectly possible to evaluate the sum over all the elements of the complete basis $|\alpha\rangle$ by a monte-carlo procedure. One can take $|\langle\psi|\alpha\rangle|^2$ as the probability and $\frac{\langle\alpha|H|\psi\rangle}{\langle\alpha|\psi\rangle}$ as the quantity that ones wants to compute. Each one of these quantity is quite simple to compute in itself. For example if one has for $|\psi\rangle$ the wavefunction of free fermions, it is simply a Slater determinant, which can be computed by a standard procedure in a time growing as N^3 where N is the number of particles (in fact here a trick allows to do the calculation faster).

The calculation is thus essentially exact, even if lengthy, since the evaluation of the probability will cost a computer time which will grow quite fast with the number of particles.

CHAPTER A

Useful formulas

We recall in this appendix some *very* useful formulas.

A.1 Normalizations in a finite volume

We consider a system in a box of volume $\Omega = L_x \cdots L_z$ in a space of d dimensions. In the thermodynamic limit the boundary conditions are irrelevant and it is convenient to take periodic ones. In that case the eigenstates for free particles are plane waves, whose wave-function satisfies

$$(\boldsymbol{r}|\boldsymbol{k}) = \frac{1}{\sqrt{\Omega}} e^{i\boldsymbol{r}\cdot\boldsymbol{k}} \tag{A.1}$$

ensuring the normalization of the wave function over the volume Ω. To satisfy the periodic condition one must have for each direction

$$k_i L_i = 2\pi p \tag{A.2}$$

where p is a relative integer. This allows to go from a discrete sum over momenta to a continuous sum using

$$\sum_{\boldsymbol{k}} \to \frac{\Omega}{(2\pi)^d} \int d^d k \tag{A.3}$$

The allowed values of $\boldsymbol{k}$ depend on the nature of $\boldsymbol{x}$. In the continuum all the values of $\boldsymbol{k}$ satisfying the quantization condition are allowed. On a discrete lattice were $\boldsymbol{x}$ is stemming from a basis vector set (i.e. on a square lattice $\boldsymbol{x} = (n_x a_x, n_y a_y, \cdots, n_z a_z)$ where the n_i are relative integers and a_i the lattice spacing) one must make sure not to double count. Only the $\boldsymbol{k}$ belonging to the Brillouin zone should be counted. For the above square lattice this means $k_i \in [-\frac{\pi}{a_i}, \frac{\pi}{a_i}]$.

In a same way using this representation we can define the Fourier transform over a finite volume. If we have a function $V(\boldsymbol{x})$ then

$$V(\boldsymbol{k}) = \int d^d x, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} V(\boldsymbol{x}) \tag{A.4}$$

with $\boldsymbol{k}$ spanning the first Brillouin zone. The inverse Fourier transform is defined by

$$V(\boldsymbol{x}) = \frac{1}{\Omega} \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} V(\boldsymbol{k}) \tag{A.5}$$

A.2 Delta functions

One has often to use the δ function in various forms. Here are some useful formulas with it.

For a nice function $f(x)$ one has

$$\int f(x)\delta(x) = f(0) \tag{A.6}$$

To make change of variables

$$\delta(f(x)) = \sum_i \frac{1}{|f'(x_i)|}\delta(x - x_i) \tag{A.7}$$

where the x_i are the zeros of the function $f(x)$. i.e. $f(x_i) = 0$.

The δ function has several integral representations, one of the most useful being

$$\int e^{i\lambda x} = (2\pi)\delta(x) \tag{A.8}$$

Essentially any function that has a peak whose height is going to infinity and width is going to zero as a function of a parameter tends to a δ function (with the proper weight). This is in particular the case of the Lorentzian

$$\lim_{\lambda\to} \frac{\lambda}{\lambda^2 + x^2} \to (2\pi)\delta(x) \tag{A.9}$$

Finally one has

$$\frac{1}{x + i\lambda} = \mathcal{P}\frac{1}{x} - i\pi\delta(x) \tag{A.10}$$

when $\lambda = 0^+$. $\mathcal{P}f(x)$ denotes the principal part of the function $f(x)$ (see below).

A.3 Complex plane integration

One often uses complex plane integration to compute several integrals difficult to compute directly using standard integration.

The most useful formula is the Cauchy theorem. If one has a function $f(z)$ of a complex variable z. Then

$$\oint_C dz f(z) = \pm 2i\pi \sum_{z_i} \text{Res.} f(z_i) \tag{A.11}$$

where C is a contour of integration in the complex plane that does not encounter any singularity of the function $f(z)$. In other words $f(z)$ is an analytical function on the contour C. The sign plus in the above formula is if the contour is done in the trigonometric direction, and the sign minus if it is done in the anti-trigonometric (clockwise) direction. Even if $f(z)$ is analytic on the contour C is can have singularities (poles) elsewhere on the complex plane. In the above sum, z_i are the poles of the function $f(z)$. Only the poles z_i that are inside the area delimited by the contour C must be counted in the sum. Res. $f(z_i)$ is the residue of the function $f(z)$ at the pole z_i. Its value depends on the nature of the pole.

1. Simple pole: in that case in the vicinity of the pole z_i, $f(z)$ can be cast in the form

$$f(z) = \frac{g(z)}{z - z_i} + f_{\text{reg}}(z) \tag{A.12}$$

where $g(z)$ and $f_{\rm reg}(z)$ are regular functions around z_i. In that case

$$\text{Res.}f(z_i) = g(z_i) \tag{A.13}$$

2. Double pole: in that case in the vicinity of the pole z_i, $f(z)$ can be cast in the form

$$f(z) = \frac{g(z)}{(z - z_i)^2} + f_{\rm reg}(z) \tag{A.14}$$

where $g(z)$ and $f_{\rm reg}(z)$ are regular functions around z_i. In that case

$$\text{Res.}f(z_i) = g'(z_i) \tag{A.15}$$

Similar formulas can be derived for higher order poles. The residue for a pole of order n being

$$\frac{1}{(n-1)!} \left. \frac{d^{n-1}g(z)}{dz^{n-1}} \right|_{z_i} \tag{A.16}$$

A.4 Some properties of operators

Let us denote $|\alpha)$ a complete basis of a vectorial space. The trace of the operator $\hat{A}$ is defined by

$$\text{Tr}[\hat{A}] = \sum_\alpha (\alpha|\hat{A}|\alpha) \tag{A.17}$$

the value of the trace does not depend on the complete basis used. Operator order is important in the trace, but the trace is invariant by cyclic permutation

$$\text{Tr}[\hat{A}\hat{B}\hat{C}] = \text{Tr}[\hat{B}\hat{C}\hat{A}] \tag{A.18}$$

One denotes the commutators and anticommutator of two operators by

$$[\hat{A}, \hat{B}]_\pm = \hat{A}\hat{B} \pm \hat{B}\hat{A} \tag{A.19}$$

where the minus sign is for the commutator and the plus sign for the anticommutator. The commutator is commonly denoted $[,]$ while the anticommutator will be denoted $[,]_+$. Commutators and anticommutators verify the simple relations

$$\begin{aligned} [A, B]_\pm &= \pm[B, A]_\pm \\ [AB, C]_\pm &= A[B, C]_\pm + [A, C]_\pm B \\ [AB, C] &= A[B, C]_+ - [A, C]_+ B \end{aligned} \tag{A.20}$$

CHAPTER B

Example of Monte Carlo program

Here is an example, in *C*, of a Monte Carlo program computing

$$A_f = \frac{\sum_n f(n) e^{-n^2}}{\sum_n e^{-n^2}} \tag{B.1}$$

In the example given below $f(n) = \sqrt{|n|}$.

```
#include <cstdio>
#include <cstdlib>
#include <iostream>
#include <stdlib.h>
#include <time.h>
#include <math.h>

using namespace std;

double func(long);

int main ()
{
    long n,nn,aa;
    long disc,count,i;
    double a,value;
    char quit;

    n = 0;
    disc = 20000;
    count = 100000000;
    value = 0.0;
    srand( (unsigned)time( NULL ) ); // init Rand() function


    for(i=0;i < disc;i++)
    {aa = rand() % 2;
     if(aa == 0){
          nn = n+1;}
     else{
          nn = n-1;
     }
     if(func(nn) > func(n)){n = nn;}
```

```
     else{a = (double)rand() / (double)RAND_MAX;
            if(a < func(nn)/func(n)){n = nn;}
     }
    }

    for(i=0;i < count;i++)
    {aa = rand() % 2;
     if(aa == 0){
          nn = n+1;}
     else{
          nn = n-1;
     }
     if(func(nn) > func(n)){n = nn;}
     else{a = (double)rand() / (double)RAND_MAX;
            if(a < func(nn)/func(n)){n = nn;}
     }
      value = value + sqrt(abs(n));
    }

    value = value/count;
    cout << "Integrale = " << value << endl;

    cin >> quit;
    return 0;
}

double func(long n)
{   double a;
    a = n;
    return exp(-a*a);
}
```

Bibliography

[AAW66] W. R. Abel, A. C. Anderson, and J. C. Wheatley. *Physical Review Letters*, 17:74, 1966.

[Abr61] A. Abragam. *Principles of Nuclear Magnetism.* Clarendon Press, Oxford, 1961.

[AGD63] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski. *Methods of Quantum Field Theory in Statistical Physics.* Dover, New York, 1963.

[AGO75] K. Andres, J. E. Graebner, and H. R. Ott. *Physical Review Letters*, 35:1779, 1975.

[AM76] N. W. Ashcroft and N. D. Mermin. *Solid State Physics.* Saunders College, Philadelphia, 1976.

[Aue98] A. Auerbach. *Interacting Electrons and Quantum Magnetism.* Springer, Berlin, 1998.

[BDZ08] I. Bloch, J. Dalibard, and W. Zwerger. Many body physics with ultracold gases. *Reviews of Modern Physics*, 80:885, 2008.

[CHA$^+$01] R. Coldea, S. M. Hayden, G. Aeppli, T. G. Perring, C. D. Frost, T. E. Mason, S.-W. Cheong, and Z. Fisk. Spin waves and electronic interactions in $\mathrm{la_2cuo_4}$. *Physical Review Letters*, 86:5377, 2001.

[CTDL77] C. Cohen-Tanoudji, B. Diu, and F. Laloë. *Quantum Mechanics.* Wiley Interscience, New York, 1977.

[DHS03] A. Damascelli, Z. Hussain, and Z.-X. Shen. *Reviews of Modern Physics*, 75:473, 2003.

[FH65] R. P. Feynmann and A. R. Hibbs. *Path Integrals and Quantum Mechanics.* McGraw Hill, New York, 1965.

[GRT08] T. Giamarchi, C. Ruegg, and O. Tchernyshyov. Bose-einstein condensation in magnetic insulators. *Nature Physics*, 4:198, 2008.

[HP40] T. Holstein and H. Primakoff. *Phys. Rev.*, 58:1098–1113, 1940.

[Kit88] C. Kittel. *Introduction to Solid State Physics*, chapter Superconductivity. WHATEVER, SOMEHWHERE, 6th edition, 1988.

[KMB$^+$08] M. Klanjšek, H. Mayaffre, C. Berthier, M. Horvatić, B. Chiari, O. Piovesana, P. Bouillot, C. Kollath, E. Orignac, R. Citro, and T. Giamarchi. Controlling luttinger liquid physics in spin ladders under a magnetic field. *Physical Review Letters*, 101:137207, 2008.

[Kra06] Werner Krauth. *Statistical Mechanics: Algorithms and Computations.* Oxford University Press, 2006.

[Leg65] A. J. Leggett. *Physical Review*, 140:A1869, 1965.

[Mah81] G. D. Mahan. *Many Particle Physics.* Plenum, New York, 1981.

[MM56] T. Matsubara and H. Matsuda. *prog. Theor. Phys.*, 16:569, 1956.

[NO87] J.W. Negele and H. Orland. *Quantum Many Particle systems.* Frontiers in Physics. Addison-Wesley, Reading, Mass., 1987.

[Noz61] P. Nozieres. *Theory of Interacting Fermi Systems.* Benjamin, New York, 1961.

[PAM+67] S. J. Pickart, H. A. Alperin, V. J. Minkiewicz, R. Nathans, G. Shirane, and O. Steinsvoll. *Physical Review*, 156:623, 1967.

[PN66] D. Pines and P. Nozières. *The Theory of Quantum Liquids.* W. A. Benjamin, New York, 1966.

[Sac00] Subir Sachdev. *Quantum Phase Transitions.* Cambridge University Press, Cambridge, UK, 2000.

[Sli89] C.P. Slichter. *Principles of Magnetic Resonance.* Springer Verlag, Berlin, 3rd edition, 1989.

[Tin75] M. Tinkham. *Introduction to Superconductivity.* McGraw Hill, New York, 1975.

[VFJH99] T. Valla, A. V. Fedorov, P. D. Johnson, and S. L. Hulbert. Many-body effects in angle-resolved photoemission: Quasiparticle energy and lifetime of a mo(110) surface state. *Physical Review Letters*, 83:2085, 1999.

[Zim72] J. M. Ziman. *Principles of the Theory of Solids.* Cambridge University Press, Cambridge, 1972.