

12.6 Origin of magnetic moments: coupling of angular momenta

In order to understand the magnetic response of a system it is important to understand how the magnetic moments of each individual atom is determined before we may understand how different magnetic moments couple in a many-body system.

The simplest example is a set of two electrons such as in a helium atom.

12.6.1 Helium atom

The helium atom is composed of a nucleus with two protons and two electrons. Since the nucleus has a much larger mass than the electrons it is usually a good approximation to consider that the nucleus is at rest. Neglecting relativistic effects such as the spin-orbit interaction and the effects of the current due to the motion of one electron on the other we may write the Hamiltonian as

$$H = \frac{1}{2m} \mathbf{p}_1^2 + \frac{1}{2m} \mathbf{p}_2^2 - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (12.32)$$

This can be written as

$$H = H^{(1)} + H^{(2)} + V, \quad (12.33)$$

where

$$H^{(i)} = \frac{1}{2m} \mathbf{p}_i^2 - \frac{Ze^2}{r_i} \quad (12.34)$$

and

$$V = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \quad (12.35)$$

the interaction between the electrons, may be considered as a perturbation on the separate motion of each electron around the nucleus. Neglecting the effect of the perturbation each subsystem is equivalent to an hydrogenoid atom with charge $Z = 2$. In this case, the wave function of the system may be written as a product of the individual wave functions in the form

$$u(\mathbf{r}_1, \mathbf{r}_2) = \phi_{n_1 l_1 m_1}(r_1) \phi_{n_2 l_2 m_2}(r_2) \quad (12.36)$$

being solutions of the simplified Schrödinger equation

$$[H^{(1)} + H^{(2)}] u(\mathbf{r}_1, \mathbf{r}_2) = E u(\mathbf{r}_1, \mathbf{r}_2). \quad (12.37)$$

with eigenvalue

$$E_{n_1 n_2} = E_{n_1} + E_{n_2}, \quad (12.38)$$

where $E_n = -\frac{mc^2}{2} \frac{(Z\alpha)^2}{n^2}$. The lowest energy state is

$$E_{11} = 2E_1 = -mc^2(2\alpha)^2 = -108,8 \text{ eV}. \quad (12.39)$$

and the first excited state is

$$E_{12} = E_{21} = E_1 + E_2 = -68,0 \text{ eV}. \quad (12.40)$$

The ionization energy is the energy required to remove one electron and is given by $E_{ion} = (E_1 + E_\infty) - 2E_1 = -E_1 = 54,4 \text{ eV}$ and it sets the threshold to the states continuum. Note that the state $n_1 = 2, n_2 = 2$ has energy $2E_2 = -27,2 \text{ eV}$ which is larger than the ionization energy and is a discrete state inside the continuum.

Actually, since the electrons are identical, its wavefunction has to be antisymmetric. The wavefunction has a spatial part and a spin part. Two possibilities arise: the spatial part is symmetric to the interchange of the two electrons and the spin part is antisymmetric or the other way around.

The groundstate is particular since the spatial part of the wavefunction is necessarily symmetric which implies a spin singlet

$$u_0(\mathbf{r}_1, \mathbf{r}_2) = \phi_{100}(\mathbf{r}_1)\phi_{100}(\mathbf{r}_2) X_{singlet}, \quad (12.41)$$

where, as usual,

$$X_{singlet} = \frac{1}{\sqrt{2}}(\chi_+^{(1)}\chi_-^{(2)} - \chi_-^{(1)}\chi_+^{(2)}). \quad (12.42)$$

or, in other notation,

$$X_{singlet} = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \quad (12.43)$$

Clearly, the spatial part is symmetric to the change of the coordinates of the two electrons and the spin part is antisymmetric.

The first excited state has two possible solutions

$$\boxed{u_1^{(s)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}[\phi_{100}(\mathbf{r}_1)\phi_{2lm}(\mathbf{r}_2) + \phi_{2lm}(\mathbf{r}_1)\phi_{100}(\mathbf{r}_2)] X_{singlet}} \quad (12.44)$$

$$u_1^{(t)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} [\phi_{100}(\mathbf{r}_1) \phi_{2lm}(\mathbf{r}_2) - \phi_{2lm}(\mathbf{r}_1) \phi_{100}(\mathbf{r}_2)] X_{triplet}, \quad (12.45)$$

where

$$X_{triplet} = \begin{cases} \chi_+^{(1)} \chi_+^{(2)} \\ \frac{1}{\sqrt{2}} (\chi_+^{(1)} \chi_-^{(2)} + \chi_-^{(1)} \chi_+^{(2)}) \\ \chi_-^{(1)} \chi_-^{(2)} \end{cases} . \quad (12.46)$$

Let us now add the interaction between the two electrons as a perturbation. Using time independent perturbation theory the leading correction to the energy of the groundstate is given by

$$\Delta E_0 = \int d^3r_1 d^3r_2 u_0^*(\mathbf{r}_1, \mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} u_0(\mathbf{r}_1, \mathbf{r}_2). \quad (12.47)$$

$$\Delta E_0 = \int d^3r_1 d^3r_2 |\phi_{100}(\mathbf{r}_1)|^2 \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} |\phi_{100}(\mathbf{r}_2)|^2. \quad (12.48)$$

This expression has a clear (classical-like) physical meaning. Noting that $e|\phi_{100}(\mathbf{r}_i)|^2$ is the charge electronic density at point \mathbf{r}_i , this is the standard Coulomb interaction between two charge densities at different spatial points, \mathbf{r}_1 and \mathbf{r}_2 . Since the wave function ϕ_{100} is known we may calculate the energy correction. The result is

$$\Delta E_0 = \frac{5}{8} \frac{Ze^2}{a_0} = \frac{5}{4} Z \left(\frac{1}{2} mc^2 \alpha^2 \right). \quad (12.49)$$

Taking $Z = 2$ we get

$$\Delta E_0 = +34 \text{ eV}. \quad (12.50)$$

which gives

$$E'_{11} \simeq -74,8 \text{ eV}. \quad (12.51)$$

which may be compared with the experimental value $E_{exp} = -78,975 \text{ eV}$. The agreement is not good. One reason is that the screening on the potential the nucleus exerts on one electron is partially shielded by the motion of the other electron. In some sense, the effective nucleus charge that one electron "sees" is smaller. This effect becomes even more important when the atom has more electrons.

12.6.2 Direct and exchange terms. Heisenberg interaction

Let us now consider the correction to the first excited state. Since the perturbation commutes with L_z we may select any value of the orbital projection. Specifically, we may choose $m = 0$. The energy correction is given by

$$\begin{aligned}\Delta E_1^{(s,t)} &= \int d^3r_1 \int d^3r_2 u_1^{(s,t)*}(\mathbf{r}_1, \mathbf{r}_2) V u_1^{(s,t)}(\mathbf{r}_1, \mathbf{r}_2) \\ &= \frac{1}{2}e^2 \int d^3r_1 \int d^3r_2 [\phi_{100}(\mathbf{r}_1)\phi_{210}(\mathbf{r}_2) \pm \phi_{210}(\mathbf{r}_1)\phi_{100}(\mathbf{r}_2)]^* \\ &\quad \times \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} [\phi_{100}(\mathbf{r}_1)\phi_{210}(\mathbf{r}_2) \pm \phi_{210}(\mathbf{r}_1)\phi_{100}(\mathbf{r}_2)].\end{aligned}\quad (12.52)$$

where we have the two possibilities, singlet and triplet,

$$\begin{aligned}\Delta E_1^{(s,t)} &= e^2 \int d^3r_1 \int d^3r_2 |\phi_{100}(\mathbf{r}_1)|^2 |\phi_{210}(\mathbf{r}_2)|^2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ &\quad \pm e^2 \int d^3r_1 \int d^3r_2 \phi_{100}^*(\mathbf{r}_1)\phi_{210}^*(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \phi_{210}(\mathbf{r}_1)\phi_{100}(\mathbf{r}_2).\end{aligned}$$

(12.53)

Note that since the Coulomb potential does not depend on spin, the overlap in spin space is trivial and chosen to be unity due to the normalization of the spin wave functions. The \pm signs are for the singlet and triplet spin states (symmetric and antisymmetric spatial wave functions, respectively). At this point it is worthwhile to stress that, even though the unperturbed hamiltonian and the perturbation do not depend on the spin, the energy correction depends. Also note that the raising of the degeneracy of the two spin states by the perturbation, is originated by a Coulomb-like energy interaction.

The first term of the energy correction is of the same form as the correction to the groundstate energy. It may be understood as an interaction between two charge densities. It is called a direct term. However, the second term has no classical interpretation. Note that it is not possible to rewrite the wave functions as two density distributions. For the same set of quantum numbers one wave function is associated with one electron and the complex conjugate with the other electron. It is called an exchange term and has no classical correspondence.

In general we may write that

$$\begin{aligned}\Delta E_{n,l}^{(s)} &= J_{n,l} + K_{n,l}, \\ \Delta E_{n,l}^{(t)} &= J_{n,l} - K_{n,l},\end{aligned}\tag{12.54}$$

Even though the perturbation does not depend on the spin, the energies of the two spin states are different. This is a consequence of the symmetry of the wave function. Here

$$J_{n,l} = e^2 \int d^3r_1 \int d^3r_2 |\phi_{100}(\mathbf{r}_1)|^2 |\phi_{nl0}(\mathbf{r}_2)|^2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\tag{12.55}$$

and

$$K_{n,l} = e^2 \int d^3r_1 \int d^3r_2 \phi_{100}^*(\mathbf{r}_1) \phi_{nl0}^*(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \phi_{nl0}(\mathbf{r}_1) \phi_{100}(\mathbf{r}_2).\tag{12.56}$$

It is easy to see that $J_{n,l} > 0$. If $K_{n,l} > 0$ then

$$J_{n,l} - K_{n,l} < J_{n,l} + K_{n,l} \Leftrightarrow \boxed{\Delta E_{n,l}^{(t)} < \Delta E_{n,l}^{(s)}}\tag{12.57}$$

It can be shown that is so. Therefore, the system prefers to be in a triplet state. The perturbation lifts the degeneracy between the two spin states. This may be understood from a simple argument. The spin triplet is associated with a spatial wavefunction that is antisymmetric. Therefore, there is a small probability that the electrons are both in the neighborhood of the nucleus. The electrons are attracted to the nucleus but repel each other. It seems reasonable therefore that it will be favorable energetically if the electrons are far apart. In this way the repulsive energy is minimized, while the attraction to the nucleus is still in place. A symmetric wavefunction would imply a higher probability to find the electrons near the nucleus and a larger repulsive contribution to the energy of the system. Therefore, in this simple case, the spin takes the maximal value.

A similar result holds also in atoms with a larger number of electrons. This constitutes the first Hund rule which states that the electrons organize themselves in such a way as to maximize the spin.

If we couple two spins, $\mathbf{s}_1, \mathbf{s}_2$, the total, \mathbf{S} , spin satisfies

$$\mathbf{S}^2 = \mathbf{s}_1^2 + \mathbf{s}_2^2 + 2\mathbf{s}_1 \cdot \mathbf{s}_2.\tag{12.58}$$

Acting on the states $|s_1 s_2 S m\rangle$ we may rewrite the equation replacing the diagonal operators by their eigenvalues as

$$s(s+1)\hbar^2 = \frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2 + 2\mathbf{s}_1 \cdot \mathbf{s}_2,$$

which implies

$$\frac{2\mathbf{s}_1 \cdot \mathbf{s}_2}{\hbar^2} = s(s+1) - \frac{3}{2} = \begin{cases} \frac{1}{2} & \text{triplet } (s=1) \\ -\frac{3}{2} & \text{singlet } (s=0) \end{cases}. \quad (12.59)$$

Therefore,

$$\frac{1}{2}\sigma_1 \cdot \sigma_2 = \begin{cases} \frac{1}{2} & \text{triplet} \\ -\frac{3}{2} & \text{singlet} \end{cases}. \quad (12.60)$$

and we may rewrite

$$\boxed{\Delta E_{n,l} = J_{n,l} - \frac{1}{2}(1 + \sigma_1 \cdot \sigma_2)K_{n,l}}. \quad (12.61)$$

This way to rewrite the energy correction is chosen to highlight that there is a spin dependence and suggests some effective interaction between the spins of the two electrons. This type of expression has the general designation of Heisenberg exchange interaction. As will be seen later, this is a form that is supposed to be the basis for the magnetic properties of interacting spin systems.

12.6.3 Variational method: screening

A better approximation may be attempted using a variational method. A simple choice of the variational wave function is to select

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \phi_{100}(\mathbf{r}_1)\phi_{100}(\mathbf{r}_2), \quad (12.62)$$

where

$$\left(\frac{\mathbf{p}^2}{2m} - \frac{Z^* e^2}{r}\right)\phi_{100}(\mathbf{r}) = \epsilon \phi_{100}(\mathbf{r}), \quad (12.63)$$

with $\epsilon = -\frac{1}{2}mc^2(Z^*\alpha)^2$ are the eigenfunctions of a hydrogenoid atom with charge Z^* taken as the variational parameter. Minimizing the average value of the Hamiltonian

in this state given by

$$\langle \psi | H | \psi \rangle = \int d^3 r_1 \int d^3 r_2 \phi_{100}^*(\mathbf{r}_1) \phi_{100}^*(\mathbf{r}_2) \left(\frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right) \phi_{100}(\mathbf{r}_1) \phi_{100}(\mathbf{r}_2). \quad (12.64)$$

leads to

$$\langle \psi | H | \psi \rangle = -\frac{1}{2} mc^2 \alpha^2 \left(4ZZ^* - 2Z^{*2} - \frac{5}{4} Z^* \right). \quad (12.65)$$

Minimizing determines the value of the effective charge of the nucleus

$$Z^* = Z - \frac{5}{16}, \quad (12.66)$$

which leads to a slightly improved value of the groundstate energy

$$E_0 \leq -\frac{1}{2} mc^2 \alpha^2 \left[2 \left(Z - \frac{5}{16} \right)^2 \right] = -77,38 \text{ eV}. \quad (12.67)$$

12.6.4 Hartree method

A way to improve the variational method and approach the exact groundstate energy is to increase the size of the variational space, for instance selecting trial wave functions with larger freedom increasing the number of variational parameters. An alternative way is Hartree's method where the functional form of the trial wave functions is to be determined.

Consider an atom with Z electrons. Neglecting the motion of the nucleus Schrödinger's equation may be written as

$$\left(\sum_{i=1}^Z \frac{\mathbf{p}_i^2}{2m} - \frac{Ze^2}{r_i} + \sum_{i>j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) = E \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) \quad (12.68)$$

This is a complex problem that in general requires a numerical solution. Due to the arbitrarily large number of electrons screening is important and a perturbative method will not yield good results.

Let us choose a variational method where the trial wave function has the form

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) = \phi_1(\mathbf{r}_1) \phi_1(\mathbf{r}_2) \cdots \phi_1(\mathbf{r}_Z) \quad (12.69)$$

where the functions $\phi_i(\mathbf{r}_i)$ are to be determined. The average value of the Hamiltonian in this state is given by

$$\begin{aligned} \langle \psi | H | \psi \rangle &= \sum_{i=1}^Z \int d^3 r_i \phi_i^*(\mathbf{r}_i) \left(-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{Ze^2}{r_i} \right) \phi_i(\mathbf{r}_i) \\ &+ e^2 \sum_{i < j} \sum_j \int d^3 r_i \int d^3 r_j \frac{|\phi_i(\mathbf{r}_i)|^2 |\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \end{aligned} \quad (12.70)$$

If a given set of functions $\phi_i(\mathbf{r}_i)$ minimizes the average value of the Hamiltonian in state $|\psi\rangle$ then if we consider a shift of the wave function in the form $\phi_i(\mathbf{r}_i) \rightarrow \phi_i(\mathbf{r}_i) + \lambda f_i(\mathbf{r}_i)$ where λ is a small parameter and $f_i(\mathbf{r}_i)$ is an arbitrary function, the average value of the Hamiltonian differs from the minimal value by terms that are at least of the order of λ^2 . In other words, the term linear in λ must vanish since we are expanding around a minimum. Also, the shift must keep the normalization of the wave functions. This implies that

$$\int d^3 r_i |\phi_i(\mathbf{r}_i) + \lambda f_i(\mathbf{r}_i)|^2 = 1. \quad (12.71)$$

To leading order in λ this leads to

$$\int d^3 r_i [\phi_i^*(\mathbf{r}_i) f_i(\mathbf{r}_i) + \phi_i(\mathbf{r}_i) f_i^*(\mathbf{r}_i)] = 0 \quad (12.72)$$

The shift of the wave functions may now be inserted in $\langle \psi | H | \psi \rangle$. The condition for minimum is that the term linear in λ vanishes, subject to the restriction 12.72.

The kinetic energy term is written as

$$\lambda \sum_i \int d^3 r_i \left[f_i^*(\mathbf{r}_i) \left[-\frac{\hbar^2}{2m} \nabla_i^2 \phi_i(\mathbf{r}_i) \right] + c.c. \right]$$

where we have integrated by parts twice, imposing that $f_i(\mathbf{r}_i)$ must vanish at infinity and where *c.c.* stands for complex conjugate.

The Coulomb interaction with the nucleus is written as

$$-\lambda \sum_i \int d^3 r_i \left[f_i^*(\mathbf{r}_i) \frac{Ze^2}{r_i} \phi_i(\mathbf{r}_i) + c.c. \right]$$

Finally, the term that describes the interactions between the electrons:

$$\lambda e^2 \sum_{i>j} \sum_j \int d^3 r_i \int d^3 r_j \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \times \\ [f_i^*(\mathbf{r}_i)\phi_i(\mathbf{r}_i) + f_i(\mathbf{r}_i)\phi_i^*(\mathbf{r}_i)] |\phi_j(\mathbf{r}_j)|^2 + [f_j^*(\mathbf{r}_j)\phi_j(\mathbf{r}_j) + f_j(\mathbf{r}_j)\phi_j^*(\mathbf{r}_j)] |\phi_i(\mathbf{r}_i)|^2$$

Collecting the various linear terms and introducing the restriction with a Lagrange multiplier, $-\epsilon_i$, we obtain that

$$\begin{aligned} & \sum_i \int d^3 r_i f_i^*(\mathbf{r}_i) \left[-\frac{\hbar^2}{2m} \nabla_i^2 \phi_i(\mathbf{r}_i) - \frac{Ze^2}{r_i} \phi_i(\mathbf{r}_i) \right] \\ & + e^2 \sum_{i \neq j} \sum_j \int d^3 r_i \int d^3 r_j f_i^*(\mathbf{r}_i) \frac{|\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \phi_i(\mathbf{r}_i) \\ & - \epsilon_i \int d^3 r_i f_i^*(\mathbf{r}_i) \phi_i(\mathbf{r}_i) = 0 \end{aligned} \quad (12.73)$$

and an equivalent equation that is the complex conjugate. Since the function $f_i^*(\mathbf{r}_i)$ is arbitrary this implies that the integrand vanishes. This leads to the Hartree equations

$$\left[-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{Ze^2}{r_i} + e^2 \sum_{j \neq i} \int d^3 r_j \frac{|\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \phi_i(\mathbf{r}_i) = \epsilon_i \phi_i(\mathbf{r}_i) \quad (12.74)$$

The first thing to note about these equations is that they are eigenvalue equations, where the Lagrange multiplier is the eigenvalue and the wavefunctions are the eigenfunctions. The second aspect is that these are self-consistent equations for the wavefunctions. If the interaction between the electrons is neglected, the Lagrange multipliers are simply the eigenenergies of the electrons, in the presence of a nucleus of charge Ze . We expect therefore a similar meaning for the parameters ϵ_i when the electron repulsion is taken into account. The term due to this interaction is interesting. Within this approximation, it may be interpreted as the mean-field due to the electrons $j \neq i$, on electron i . Note that it is the sum over all other electrons of a term that is the Coulomb potential, due to some charge density at point \mathbf{r}_j and then summed over all $j \neq i$. Therefore, Hartree's variational approach leads to some sort of mean-field approximation, where the equation for the wave-functions of a given

electron i involves all the remaining electrons interacting with electron i , but where the interactions among the j electrons do not appear explicitly, in contrast with the Hamiltonian of the full system.

To some extent, Hartree's approximation reduces the many-body interacting problem to a single particle problem in the presence of an effective potential

$$V_{eff}^i(\mathbf{r}_i) = -\frac{Ze^2}{r_i} + e^2 \sum_{j \neq i} \int d^3r_j \frac{|\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (12.75)$$

although self-consistent, since the potential itself depends on the solution of the wavefunctions of the other electrons. This is typical of a mean-field approximation as will be seen in later sections.

If the effective potential has spherical symmetry, we may integrate over the solid angle and the problem simplifies to a problem with a central potential. We may therefore focus only on the radial part and introduce the usual set of quantum numbers characteristic of the hydrogen atom.

12.6.5 Hartree-Fock equations

The description we have just considered is however incomplete. Since the electrons are fermions their wavefunction should be completely anti-symmetric. The trial wavefunction should be taken as the Slater determinant of the trial functions $\phi_i(\mathbf{r}_j)$. The minimization of the average value of the Hamiltonian in the Slater determinant leads to a set of equations that have an extra term due to the exchange between the electrons. This term is called the Fock term and the new set of equations are called Hartree-Fock equations and can be written as

$$\left[-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{Ze^2}{r_i} + e^2 \sum_{j \neq i} \int d^3r_j \frac{|\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \phi_i(\mathbf{r}_i) - e^2 \sum_{j \neq i} \int d^3r_j \frac{\phi_j^*(\mathbf{r}_j) \phi_i(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} \phi_i(\mathbf{r}_i) = \epsilon_i \phi_i(\mathbf{r}_i) \quad (12.76)$$

The solution of these equations is in general more complex than the Hartree equations. In the context of atoms the improvement that results from the addition of the Fock term is not very significant. Often an analysis of the solutions using the Hartree approximation and taking into account Pauli's exclusion principle, leads to

a good description of the energy levels of atoms. This problem of atomic physics is beyond the scope of this book and a simplified description is enough for our purposes. In some cases their solution in the context of solid state physics is possible and later on we will consider some examples of their solution.

12.6.6 Hund rules and magnetic moments in a crystal

To a good approximation, the single particle states of an atom with Z electrons can be characterized by the same quantum numbers as those of an hydrogen-like atom with a nuclear charge $|Ze|$. The distribution of the electrons among the available single-particle states requires a calculation of the level degeneracies. The quantum numbers are traditionally taken as n, l, m_l, m_s , where n is the principal quantum number, $l = 0, \dots, n-1$ is the angular momentum, $m_l = -l, -l+1, \dots, l-1, l$ is the angular momentum projection along some quantization axis, and $m_s = -1/2, 1/2$ is the spin angular momentum projection of each electron. In the absence of external fields the energy eigenvalues are only a function of n . The degeneracy of each energy level is $g = 2n^2$ since

$$\sum_{l=0}^{n-1} 2(2l+1) = 2n^2$$

Each orbital with a given pair of quantum numbers n, l has $2(2l+1)$ states. If all m_l states are occupied with an electron the particle distribution is given by

$$\sum_{m_l=-l}^l |R_{nl}(r)|^2 |Y_{lm_l}(\theta, \varphi)|^2 = \frac{2l+1}{4\pi}$$

(a constant) and therefore is spherically symmetric.

Given an atom with Z electrons the groundstate of the atom is obtained distributing the electrons among the available single-particle states appropriately. The problem is complex but some empirical rules are helpful. These rules are summarized in Hund's rules.

Let us consider a free atom where all energy levels are filled except a higher energy one which is partially filled. Assume this last orbital has a given angular momentum, l . In addition to the potential energy due to the nucleus, there are other contributions. As we have seen above there is the interaction between the electrons. In a first approximation we are assuming that the Hartree and Fock terms are a

small enough perturbation that even though shifts the energy levels, does not lead to level crossings and mixings such that we are able to still label the states with the quantum numbers of the hydrogen atom. However, there are other corrections to the hydrogen atom itself, that are important as the number of electrons increases such as the spin-orbit interaction.

Hund first rule: The first empirical rule has already been discussed. The electrons organize themselves to be in a state of maximal total spin, S (total spin is a spin that is the sum of all the individual spins of each electron). Note that the total spin of a filled orbital vanishes, due to Pauli's principle.

Hund second rule: The total angular momentum, L , has the maximal value that is compatible with the first rule and with Pauli's principle.

Hund third rule: the total angular momentum, J , where $\mathbf{J} = \mathbf{L} + \mathbf{S}$, is given by $J = |L - S|$ if the orbital is less than half-full, and $J = L + S$ if the orbital is more than half-full. This rule is the result of the spin-orbit interaction written as $\lambda_{SO}(\mathbf{S} \cdot \mathbf{L})$. If $\lambda_{SO} > 0$ the spin and angular momentum minimize the energy if they are parallel and if $\lambda_{SO} < 0$ they tend to be anti-parallel. If the orbital is less than half-full the first case occurs and vice-versa if the band is more than half-full.

As examples we will consider the rare-earth materials of the lanthanide and actinide families and the iron group elements. First recall that the various orbitals with different l values are denoted as s, p, d, f, \dots when $l = 0, 1, 2, \dots$, respectively. The degeneracy of each orbital is $2, 6, 10, 14, \dots$, respectively. For instance the lanthanides have in common a set of filled levels that share with the rare gas Xe . In addition they have an electronic configuration such that there are 2 electrons in the $n = 5, l = 0$ orbital and six electrons in the $n = 5, l = 1$ orbital. That is they have a configuration $5s^2 5p^6$. The various elements along the lanthanide series include the elements $La^{57}; Ce^{58}; Pr^{59}; Nd^{60}; \dots; Yb^{70}; Lu^{71}$. Their electronic configurations include as well the following electrons: $(5d, 6s^2); (4f^2, 6s^2); (4f^3, 6s^2); (4f^4, 6s^2); \dots; (4f^{14}, 6s^2); (4f^{14}, 5d, 6s^2)$. The main difference between the various lanthanides is the filling of the $4f$ orbital. In comparison with the $n = 5$ and $n = 6$ orbitals, the $4f$ orbital is of shorter range and more internal. The actinides differ by the filling of the $5f$ orbital and a similar property may be invoked. When several atoms are brought together, such as in the solid state, the spatial extent of these internal partially filled orbitals is typically smaller than the distance between the atoms. Therefore, the neighborhood of each atom does not affect in an important way the electrons that occupy these orbitals.

The case of the iron group elements is different. For instance considering the elements Fe ; Co ; Ni ; Cu and the electronic configurations of the higher energy orbitals $(3d^6, 4s^2)$; $(3d^7, 4s^2)$; $(3d^8, 4s^2)$; $(3d^10, 4s)$ a similar trend is verified. In this case they differ by the filling of the $3d$ orbital. Note that in Cu this orbital is completely filled and therefore is expected to be inert. An important difference with respect to the lanthanides or the actinides is that the spatial extent of the $3d$ orbital is similar to the radius of the atom. Therefore, the presence of neighboring atoms will affect the electrons of the $3d$ orbital in a significant way.

As mentioned above a filled orbital has no total spin and therefore no spin magnetic moment. The possible contribution to the spin, and total magnetic momentum of the atom is originated in the partially filled orbitals. As an exercise let us use Hund rules to calculate the angular momenta of various atoms. As atoms are brought together their valencies will change depending on the neighboring atoms.

Consider for instance Mn . Its partially filled level has the configuration $(3d^5, 4s^2)$. The two electrons in the $4s$ orbital have $S = 0, L = 0$. The electrons in the $3d$ orbital have $S = 5/2, L = 0$. The same happens if the valence state is Mn^{2+} where the configuration is simply $3d^5$. Therefore $J = S = 5/2$. Another example is Ce^{3+} with electronic configuration $4f^1$. It is easy to see that $S = 1/2, L = 3$ and $J = 5/2$. For instance Fe leads to $S = 2, L = 2$ and $J = 4$. The same result holds for a divalent ion where the two $4s$ electrons are lost.

A given total angular momentum leads as usual to a finite magnetic moment given by

$$\mu = -g\mu_B \mathbf{J} \quad (12.77)$$

where g is the Landé factor

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \quad (12.78)$$

and μ_B is the Bohr magneton. The energy of a magnetic moment in an external magnetic field \mathbf{B} is given by

$$U = -\mu \cdot \mathbf{B} = m_J g \mu_B B \quad (12.79)$$

where m_J is the projection of the total angular momentum along the direction of the applied magnetic field.

We may define the effective number of Bohr magnetons as

$$p = g\sqrt{J(J+1)} \quad (12.80)$$

This number can be measured directly through the magnetic susceptibility at high temperatures. As we have shown previously, the susceptibility at high temperatures follows Curie's law $\chi = C/T$, where T is the temperature and C is Curie's constant given by

$$C = \frac{Np^2\mu_B^2}{3k_B} \quad (12.81)$$

The experimental results for the effective number of Bohr magnetons when the atom is in the solid state are somewhat unexpected. In the case of rare earths the agreement between the calculated value and the experimental value is quite good. The calculated value uses Hund rules whose validity holds for an isolated atom. In the solid state the atomic states give place to different Bloch states together with more atomic-like states associated with the orbitals deep inside the atom cores. As discussed above, the orbitals responsible for the magnetic moment of the rare earths, are quite confined around each atom which explains the good agreement with the experimental value. In the case of the iron group elements, the agreement is much worse. It turns out that a good fit of the experimental results is obtained if we take that

$$p = 2\sqrt{S(S+1)} \quad (12.82)$$

This expression is obtained from the general case if $L = 0$, which implies $J = S$ and $g = 2$. We conclude that in the crystal the orbital angular momentum is quenched to zero due to the effects of the neighborhood of an atom.

In the solid state the outer orbitals overlap significantly and their shape changes considerably. In a metal some electrons become completely delocalized in extended states leaving behind ions whose charge distributions create a crystalline field that acts on the remaining electrons of the neighboring lattice sites. Due to the lattice structure the crystalline field is in general non-homogeneous and not spherically symmetric. As a consequence, it tends to distort the orbitals. Also, energetic considerations imply that some orientations are preferred and degeneracies are lifted due to the crystalline field. The energies involved may be competitive with the energies associated with Hund rules. Typically the first and second Hund rules are very robust but the competition between the spin-orbit interaction and the crystalline field has to be taken into account. For instance for the elements *Fe, Co, Ni* typically the crystalline field is more important than the spin-orbit interaction, but as we move to atoms with a higher number of electrons, such as *Ru, Rh, Pd* or *Os, Ir, Pt* the

spin-orbit interaction becomes dominant over the crystalline field. The outcome is still non-trivial since the energies are comparable.

As a simple example of the effect of the crystal on the orbitals consider a planar structure such that along a given direction (call it z) there is a positive electric field both above and below the plane. In a free atom the three p -orbitals ($l = 1$) are degenerate. However, in the planar structure it is clearly preferable energetically that the electrons are oriented along the z direction. The orbital p_z has a lower energy with respect to the degenerate p_x, p_y orbitals and the three-degeneracy is partially lifted by the field.

Consider now a crystal with orthorhombic symmetry such that the neighboring ions produce an electrostatic potential on a given site in the form

$$e\varphi = Ax^2 + By^2 - (A + B)z^2 \quad (12.83)$$

Note that this potential satisfies Laplace's equation $\nabla^2\varphi = 0$. The $l = 1$ wavefunctions in free space can be written as

$$u_x = xf(r), u_y = yf(r), u_z = zf(r) \quad (12.84)$$

Note that $L^2u_i = l(l+1)u_i = 2u_i$, in units of \hbar^2 and where $i = x, y, z$. Now it is easy to show by symmetry that

$$\langle u_x | e\varphi | u_y \rangle = \langle u_x | e\varphi | u_z \rangle = \langle u_y | e\varphi | u_z \rangle = 0 \quad (12.85)$$

On the other hand the diagonal matrix elements do not vanish. For instance,

$$\langle u_x | e\varphi | u_x \rangle = \int dx dy dz |f(r)|^2 [Ax^4 + By^2x^2 - (A + B)x^2z^2] = A(I_1 - I_2) \quad (12.86)$$

where

$$\begin{aligned} I_1 &= \int dx dy dz |f(r)|^2 x^4 \\ I_2 &= \int dx dy dz |f(r)|^2 x^2 y^2 \end{aligned} \quad (12.87)$$

Similarly,

$$\begin{aligned} \langle u_y | e\varphi | u_y \rangle &= B(I_1 - I_2) \\ \langle u_z | e\varphi | u_z \rangle &= -(A + B)(I_1 - I_2) \end{aligned} \quad (12.88)$$

As a consequence, the electrostatic field lifts the degeneracy between the three orbitals.

We may now calculate the average value of L_z in the three orbitals. Using that $L_z = -i\hbar(x\partial/\partial y - y\partial/\partial x)$ it is easy to see that

$$\langle u_x | L_z | u_x \rangle = \langle u_y | L_z | u_y \rangle = \langle u_z | L_z | u_z \rangle = 0 \quad (12.89)$$

and the off-diagonal matrix elements, such as $\langle u_x | L_z | u_y \rangle$ do not vanish. Since the degeneracy is lifted by the crystalline field the average value of the angular momentum is taken on each orbital separately leading to a vanishing value, and therefore quenching of the angular momentum. If the degeneracy is not lifted, for instance in a cubic crystal the electrostatic field $e\varphi = 0$, the average value of the angular momentum has to be evaluated in a state that is a linear combination of the three degenerate states. Even though the diagonal terms vanish there is a contribution from the off-diagonal matrix elements that, therefore, lead to a non-vanishing average value of the angular momentum.

Note however, that even if there is cubic symmetry quenching may occur. As we saw a distortion of the orbitals leads to a lifted degeneracy. This may lead to a lower energy state if the occupation of the degenerate levels is such that the energy is smaller. It may be then energetically favorable for the system to distort spontaneously. This is called a Jahn-Teller effect. As a consequence there may be again quenching of the angular momentum.

As a consequence, the electrostatic field lifts the degeneracy between the three orbitals.

We may now calculate the average value of L_z in the three orbitals. Using that $L_z = -i\hbar(x\partial/\partial y - y\partial/\partial x)$ it is easy to see that

$$\langle u_x | L_z | u_x \rangle = \langle u_y | L_z | u_y \rangle = \langle u_z | L_z | u_z \rangle = 0 \quad (12.89)$$

and the off-diagonal matrix elements, such as $\langle u_x | L_z | u_y \rangle$ do not vanish. Since the degeneracy is lifted by the crystalline field the average value of the angular momentum is taken on each orbital separately leading to a vanishing value, and therefore quenching of the angular momentum. If the degeneracy is not lifted, for instance in a cubic crystal the electrostatic field $e\varphi = 0$, the average value of the angular momentum has to be evaluated in a state that is a linear combination of the three degenerate states. Even though the diagonal terms vanish there is a contribution from the off-diagonal matrix elements that, therefore, lead to a non-vanishing average value of the angular momentum.

Note however, that even if there is cubic symmetry quenching may occur. As we saw a distortion of the orbitals leads to a lifted degeneracy. This may lead to a lower energy state if the occupation of the degenerate levels is such that the energy is smaller. It may be then energetically favorable for the system to distort spontaneously. This is called a Jahn-Teller effect. As a consequence there may be again quenching of the angular momentum.

12.6.7 Hydrogen molecule

We want now to see the effect on the magnetic moments of several atoms brought together. The goal is to understand a solid of many atoms where each one has partially filled orbitals and, therefore, a potential magnetic moment. As a first example we consider two hydrogen atoms brought together to form an hydrogen molecule.

To simplify things further we consider the ionized hydrogen molecule where one electron has been removed, H_2^+ . Let us call \mathbf{R} the coordinate of one proton with respect to the other and \mathbf{r} the coordinate of the electron with respect to the middle point between the two protons. Schrödinger equation may be written as

$$\left(-\frac{\hbar^2}{2m} \nabla_{\mathbf{r}}^2 - \frac{\hbar^2}{2m} \nabla_{\mathbf{R}}^2 - \frac{e^2}{|\mathbf{r} - \mathbf{R}/2|} - \frac{e^2}{|\mathbf{r} + \mathbf{R}/2|} + \frac{e^2}{R} \right) \psi(\mathbf{r}, \mathbf{R}) = E \psi(\mathbf{r}, \mathbf{R}). \quad (12.90)$$

Here $M = M_p/2$, where M_p is the mass of each proton. Since the mass of the protons is much larger than the mass of the electron (and in general the mass of the nuclei are much larger than the mass of the electron) the nuclei move much slower than the electrons. It is usually a good approximation to neglect the motion of the nuclei. The electrons move in an effective field of the nuclei, for a given set of positions of the nuclei (in the case of the hydrogen molecule for a given value of R , the distance between the protons). As the nuclei change their relative positions the electrons are assumed to move fast enough that adapt instantaneously to the new nuclei configuration. Specifically, in the case of the hydrogen molecule, the distance R may be considered as a parameter and we may neglect the kinetic energy of the nuclei. This is usually called the Born-Oppenheimer approximation.

Neglecting the kinetic energy we get a simplified electronic Hamiltonian written as

$$\begin{aligned} H_0 u_0(\mathbf{r}, \mathbf{R}) &= \left(\frac{p_e^2}{2m} - \frac{e^2}{|\mathbf{r} - \mathbf{R}/2|} - \frac{e^2}{|\mathbf{r} + \mathbf{R}/2|} + \frac{e^2}{R} \right) u_0(\mathbf{r}, \mathbf{R}) \\ &= \epsilon_0(R) u_0(\mathbf{r}, \mathbf{R}). \end{aligned} \quad (12.91)$$

The first term is the kinetic energy of the motion of the electron, the next two terms are the attractive interactions between the electron and the protons and the last term is the Coulomb repulsion between the protons, assumed static. The problem of the electron and one proton is just the hydrogen atom. We may attempt a variational approach where we use two groundstate eigenfunctions of the hydrogen atom, specifically the functions

$$\psi_1(\mathbf{r}, \mathbf{R}) = \left(\frac{1}{\pi a_0^3} \right)^{1/2} e^{|\mathbf{r} - \mathbf{R}/2|/a_0} \quad (12.92)$$

and

$$\psi_2(\mathbf{r}, \mathbf{R}) = \left(\frac{1}{\pi a_0^3} \right)^{1/2} e^{|\mathbf{r} + \mathbf{R}/2|/a_0}, \quad (12.93)$$

which may be interpreted as the two possibilities that either the electron is attached to one proton or the other.

As a trial solution for the ionized hydrogen molecule we may consider a symmetric and an antisymmetric linear combination of these two wavefunctions. Since the Hamiltonian ($\mathbf{p}_e \rightarrow -\mathbf{p}_e$, $\mathbf{r} \rightarrow -\mathbf{r}$, $\mathbf{R} \rightarrow -\mathbf{R}$), we may use $\psi_1(\mathbf{r}, \mathbf{R})$ and $\psi_2(\mathbf{r}, \mathbf{R})$:

$$\boxed{\begin{aligned} \psi_p(\mathbf{r}, \mathbf{R}) &= C_+(R)[\psi_1(\mathbf{r}, \mathbf{R}) + \psi_2(\mathbf{r}, \mathbf{R})], \\ \psi_i(\mathbf{r}, \mathbf{R}) &= C_-(R)[\psi_1(\mathbf{r}, \mathbf{R}) - \psi_2(\mathbf{r}, \mathbf{R})], \end{aligned}} \quad (12.94)$$

The normalizations are defined by

$$\frac{1}{C_{\pm}} = \langle \psi_1 \pm \psi_2 | \psi_1 \pm \psi_2 \rangle \quad (12.95)$$

Defining

$$S(R) = \int d^3r \psi_1(\mathbf{r}, \mathbf{R}) \psi_2(\mathbf{r}, \mathbf{R}) \quad (12.96)$$

we can calculate

$$\begin{aligned} \langle H_0 \rangle_{p,i} &= \frac{1}{2[1 \pm S(R)]} \langle \psi_1 \pm \psi_2 | H_0 | \psi_1 \pm \psi_2 \rangle \\ &= \frac{1}{2[1 \pm S(R)]} [\langle \psi_1 | H_0 | \psi_1 \rangle + \langle \psi_2 | H_0 | \psi_2 \rangle \pm \langle \psi_1 | H_0 | \psi_2 \rangle \pm \langle \psi_2 | H_0 | \psi_1 \rangle] \\ &= \frac{\langle \psi_1 | H_0 | \psi_1 \rangle \pm \langle \psi_1 | H_0 | \psi_2 \rangle}{1 \pm S(R)}, \end{aligned} \quad (12.97)$$

The calculation leads to

$$\begin{aligned} \langle H_0 \rangle_{p,i} &= \frac{1}{1 \pm S(R)} \left[E_1 + \frac{e^2}{R} \left(1 + \frac{R}{a_0} \right) e^{-\frac{2R}{a_0}} \right. \\ &\quad \left. \pm \left(E_1 + \frac{e^2}{R} \right) S(R) \mp \frac{e^2}{a_0} \left(1 + \frac{R}{a_0} \right) e^{-\frac{R}{a_0}} \right] \end{aligned} \quad (12.98)$$

A plot of the two energy solutions shows that, as a function of R , the even solution corresponds to a bound state and the odd solution leads to an unbound state. Therefore, the symmetric solution is the favored one energetically. However, the agreement with the exact solution is not very good, particularly when the distance between the protons is small, small R . As $R \rightarrow 0$ the wave function should tend to the ionized helium atom, He^+ .

Let us now consider the neutral hydrogen molecule. We have now two electrons and two protons. We denote the two electrons as 1 and 2 and the two protons as a and b . The distance between the protons is denoted R_{ab} . The distance between the electrons is r_{12} . The distance from electron i to proton α is $r_{\alpha i}$, where $i = 1, 2$ and $\alpha = a, b$. The Hamiltonian may be written as

$$H = H_1 + H_2 + \frac{e^2}{r_{12}} + \frac{e^2}{R_{ab}}, \quad (12.99)$$

with

$$H_i = \frac{\mathbf{p}_i^2}{2m_e} - \frac{e^2}{r_{ai}} - \frac{e^2}{r_{bi}}, \quad (12.100)$$

The previous example of the ionized hydrogen molecule is described by the Hamiltonian

$$\bar{H}_i = H_i + \frac{e^2}{R_{ab}}, \quad (12.101)$$

and the total Hamiltonian may be written as

$$H = \bar{H}_1 + \bar{H}_2 + \frac{e^2}{r_{12}} - \frac{e^2}{R_{ab}}. \quad (12.102)$$

We may now use as a variational wave function the one for the ionized hydrogen molecule

$$\psi_p(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2(1 + S(R_{ab}))} [\psi_a(\mathbf{r}_1) + \psi_b(\mathbf{r}_1)] [\psi_a(\mathbf{r}_2) + \psi_b(\mathbf{r}_2)] \chi_{singlet}, \quad (12.103)$$

Each electron is therefore associated with both protons. Note that the spin part is a singlet since the orbital part is symmetric. This type of trial wavefunction is a product of molecular orbitals.

The average of the Hamiltonian in this wave function can be evaluated

$$\begin{aligned} \langle \psi_p | H | \psi_p \rangle &= \langle \psi_p | \bar{H}_1 + \bar{H}_2 + \frac{e^2}{r_{12}} - \frac{e^2}{R_{ab}} | \psi_p \rangle = \\ &= E(R_{ab}) + E(R_{ab}) + \langle \psi_p | \frac{e^2}{r_{12}} | \psi_p \rangle - \frac{e^2}{R_{ab}} = \\ &= 2E(R_{ab}) - \frac{e^2}{R_{ab}} + \langle \psi_p | \frac{e^2}{r_{12}} | \psi_p \rangle, \end{aligned} \quad (12.104)$$

$$\begin{aligned} E_l &= -2,68 \text{ eV}, \\ R &= 0,85 \text{ \AA}. \end{aligned} \quad (12.105)$$

Minimizing with respect to the distance between the protons we can get an estimate of the energy and the distance, leading to

$$\begin{aligned} E_l^{exp} &= -4,75 \text{ eV}, \\ R^{exp} &= 0,74 \text{ \AA}, \end{aligned} \quad (12.106)$$

which compare with the experimental results reasonably. The main reason for the discrepancy lies with the poor treatment of the small distance regime.

Note now that we may expand the product of the molecular orbitals in the following way

$$\begin{aligned} & [\psi_a(\mathbf{r}_1) + \psi_b(\mathbf{r}_1)][\psi_a(\mathbf{r}_2) + \psi_b(\mathbf{r}_2)] = \\ & = [\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_b(\mathbf{r}_2)] + [\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)]. \end{aligned} \quad (12.107)$$

The first term is an ionic term that describes two electrons in one proton or the other. The second term is a covalent term that describes one electron in one proton and the other in the other proton. The covalent term is a description in terms of linear combinations of atomic orbitals (LCAO). The wave function we are using, in addition to the problem at small distances, implies that for large distances between the protons the molecule has equal probability to dissociate into ions H^+ and H^- as to dissociate in two hydrogen atoms. This is clearly not true since the second case is preferred over the first. We may therefore select a more appropriate trial wavefunction using the Heitler-London approach by selecting the LCAO. The normalized wavefunction is then selected as

$$\psi_{LV}(\mathbf{r}_1, \mathbf{r}_2) = \left\{ \frac{1}{2(1 + S^2(R_{ab}))} \right\}^{\frac{1}{2}} [\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)] \chi_{singlet}. \quad (12.108)$$

A possible reasoning for this choice lies on the fact that the spatially symmetric nature of the wavefunction may lead to a high probability to find the electrons in the region shared by the two protons, specifically in the middle point of the molecule. In contrast, an antisymmetric choice for the spatial part implies a vanishing probability to find the electrons in the middle point and expectedly a small probability in the middle region. Even though the electrons repel themselves we may expect that a symmetric arrangement of the electrons is favored energetically since the attractive interaction with *both* protons is maximized. As a consequence, the spin part is antisymmetric and it is favorable for the spins of the two electrons to be aligned in opposite directions to form the spin singlet state.

Using once again the variational method with this trial wavefunction leads to a slight improvement of the energy: $E = -3.14\text{eV}$ and an equilibrium distance between the protons of $R = 0.87\text{\AA}$. The result is still not very impressive since the problem at low values of R_{ab} remains. Of course, if we increase the number of variational

parameters or in general the space in which the trial function resides, we may get an arbitrarily good agreement with the experimental results.

12.6.8 Wigner theorem and Lieb and Mattis theorem

The results of the previous section seem to indicate that in the hydrogen molecule the lowest energy state has spin zero. The same result is obtained if we consider a system with two electrons not necessarily bound to two protons but that are free to move in a d -dimensional space. These electrons have kinetic energy and may interact with each other via an arbitrary potential that we may assume is symmetric under the change of the two electrons coordinates. Then it can be shown with some generality that the groundstate of this system has a spatial wavefunction that is symmetric and therefore the spin part is antisymmetric. This is called Wigner theorem.

The theorem may be shown in the following way. Consider an Hamiltonian of the general type

$$H = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} + V(\mathbf{r}_1, \mathbf{r}_2) \quad (12.109)$$

Choose the potential to be real and symmetric. The two electrons have spin $s = 1/2$. Therefore, the pair has total spin $S = 0, 1$. We want now to show that the groundstate has $S = 0$. Since the potential is real we may take the wavefunctions of the Hamiltonian to be real. The groundstate satisfies $H\psi_0 = E\psi_0$. We want to show that a $S = 0$ wavefunction written as

$$\psi_0 = \phi_s(\mathbf{r}_1, \mathbf{r}_2) [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \quad (12.110)$$

is the groundstate. A $S = 1$ wavefunction may be written as

$$\psi_{S=1} = \phi_{as}(\mathbf{r}_1, \mathbf{r}_2) \chi_{S=1} \quad (12.111)$$

Let us now suppose that the lowest energy state has $S = 1$. Consider then the trial function

$$\tilde{\psi}_0 = |\phi_{as}| [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \quad (12.112)$$

This trial wave function has $S = 0$ but is antisymmetric. Pauli's principle is therefore obeyed. We may now calculate

$$\begin{aligned} E_{S=1} &= \langle \psi_{S=1} | H | \psi_{S=1} \rangle \\ \tilde{E}_{S=0} &= \langle \tilde{\psi}_0 | H | \tilde{\psi}_0 \rangle \end{aligned} \quad (12.113)$$

Explicitly

$$\begin{aligned} E_{S=1} &= \int d^3r_1 \int d^3r_2 \phi_{as}^*(\mathbf{r}_1, \mathbf{r}_2) H \phi_{as}(\mathbf{r}_1, \mathbf{r}_2) \\ \tilde{E}_{S=0} &= \int d^3r_1 \int d^3r_2 |\phi_{as}^*(\mathbf{r}_1, \mathbf{r}_2)| H |\phi_{as}(\mathbf{r}_1, \mathbf{r}_2)| \end{aligned} \quad (12.114)$$

Clearly $E_{S=1} = \tilde{E}_{S=0}$. This implies that there is always an eigenfunction with $S = 0$ with an energy that is, at least, as low as the lowest energy with $S = 1$. But $\tilde{\psi}_0$ is not an eigenfunction, unless the potential is quite special such as an infinitely repulsive core in one-dimension. Therefore, we may use it as a trial function in a variational procedure, from which we expect to find some lower energy average value $E(S = 0) < E(S = 1)$ because the variational method will lead to a smaller energy. Therefore the lowest energy has $S = 0$.

This result may be extended to an arbitrary number of electrons on a one-dimensional system. It may also be shown that the groundstate has zero total spin. This is called Lieb and Mattis theorem.

These results seem to indicate a tendency for antiparallel alignment of spins between neighboring points. This contrasts to some extent with the result expressed by Hund rule where the spin of an atom is maximized. (Recall, however, that the spin also vanished for the groundstate of the Helium atom. In that case, the wavefunction of the two electrons had a symmetrical spatial part and, consequently, spin zero).

Capítulo 11

Linear response theory

The action of a time dependent perturbation on a system is described by the solution of the time dependent Schrödinger equation

$$i\hbar \frac{\partial \psi(t)}{\partial t} = H\psi(t) \quad (11.1)$$

The solution can be written formally as

$$|t\rangle = U(t, t_0)|t_0\rangle \quad (11.2)$$

in terms of the unitary time evolution operator, $U(t, t_0)$, between an initial time, t_0 , and a given instant, t . This operator satisfies the relations

$$\begin{aligned} U(t_0, t_0) &= 1 \\ U(t_2, t_0) &= U(t_2, t_1)U(t_1, t_0) \end{aligned} \quad (11.3)$$

Taking $t_1 = t_2 - \delta t$, where δt is small it can be shown that the operator satisfies an evolution equation

$$i\hbar \frac{\partial U(t)}{\partial t} = H(t)U(t) \quad (11.4)$$

which naturally leads back to Schrödinger equation.

If the Hamiltonian is time independent, $H(t) = H$, the solution is straightforward

$$|t\rangle = e^{-\frac{i}{\hbar}H(t-t_0)}|t_0\rangle \quad (11.5)$$

If the Hamiltonian is time dependent the solution can be written as

$$\begin{aligned} U(t, t_0) &= 1 - \frac{i}{\hbar} \int_{t_0}^t dt' H(t') U(t', t_0) \\ &= 1 + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar} \right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H(t_1) H(t_2) \cdots H(t_n) \end{aligned} \quad (11.6)$$

and involves a time ordered product (recall that the same operator at different times in general does not commute). This solution is not very helpfull since it involves infinite terms. Each term involves the total Hamiltonian of the system which includes the original Hamiltonian plus the time dependent perturbation. Even if this extra term is small, the series has to be summed since the terms of increasing order are not of smaller amplitude. It is therefore convenient to obtain an expansion that only involves the perturbation. Using the fact that this term is small we may hope to truncate the series at some desired order.

The standard Schrödinger equation is written in the so-called Schrödinger picture. In the standard approach the time-dependence is contained in the quantum states and the operators are in general time-independent. For instance the time-dependent matrix element of an operator A can be written as $\langle t|A|t \rangle$. Using that $|t \rangle = U|t_0 \rangle$ we can write that

$$\langle t|A_S|t \rangle = \langle t_0|U^\dagger A_S U|t_0 \rangle = \langle t_0|A_H|t_0 \rangle \quad (11.7)$$

where we have added the designation A_S to emphasize that the (time-independent) operator is in the Schrödinger picture and we have defined a time dependent operator A_H where H stands for Heisenberg picture. This operator is defined by

$$A_H = U^\dagger A_S U \quad (11.8)$$

If the Hamiltonian is time independent $H(t) = H$ we can write that

$$A_H(t) = e^{\frac{i}{\hbar} H t} A_S e^{-\frac{i}{\hbar} H t} \quad (11.9)$$

(taking $t_0 = 0$, to simplify). We obtain therefore that

$$\begin{aligned} \frac{dA_H}{dt} &= \frac{i}{\hbar} H e^{\frac{i}{\hbar} H t} A_S e^{-\frac{i}{\hbar} H t} \\ &\quad - \frac{i}{\hbar} e^{\frac{i}{\hbar} H t} A_S H e^{-\frac{i}{\hbar} H t} \\ &\quad + e^{\frac{i}{\hbar} H t} \frac{\partial A_S}{\partial t} e^{-\frac{i}{\hbar} H t} \end{aligned} \quad (11.10)$$

including a possible explicit time dependent of the operator in the Schrödinger picture. Interestingly this equation can be written in a way that is formally equivalent to the Ehrenfest theorem for the time average of an operator

$$i\hbar \frac{d}{dt} \langle t | A_S | t \rangle = \langle t | [A_S, H] | t \rangle + i\hbar \left\langle \frac{\partial A_S}{\partial t} \right\rangle \quad (11.11)$$

Indeed we can write that the operator in the Heisenberg picture obeys

$$i\hbar \frac{dA_H(t)}{dt} = [A_H(t), H(t)] + i\hbar \frac{\partial A_H}{\partial t} \quad (11.12)$$

Consider now that to a given Hamiltonian H_{0S} we add a perturbation V_S . The time evolution of a state in Schrödinger picture is governed by the full Hamiltonian $H_S = H_{0S} + V_S$ as

$$i\hbar \frac{\partial}{\partial t} |t\rangle_S = H_S |t\rangle_S \quad (11.13)$$

In the spirit of time dependent perturbation theory assume we know the solution of

$$i\hbar \frac{\partial}{\partial t} |t\rangle_S = H_{0S} |t\rangle_S \quad (11.14)$$

and that we want to know the effect of the extra term, V_S , assumed to be small compared to H_{0S} . One defines an unitary operator

$$\bar{U} = e^{\frac{i}{\hbar} H_{0S} t} \quad (11.15)$$

that leads to a state in the so-called interaction picture

$$|t\rangle_I = \bar{U} |t\rangle_S \quad (11.16)$$

and an operator

$$A_I = \bar{U} A_S \bar{U}^{-1} \quad (11.17)$$

At the initial time $t = 0$ the interaction picture coincides with the Schrödinger picture and if the extra term $V_S = 0$ the interaction picture coincides with the Heisenberg picture. Therefore, it is also called intermediate picture.

Note that $H_{0S} = H_{0I} = H_0$. We conclude that

$$i\hbar \frac{\partial}{\partial t} |t\rangle_I = e^{\frac{i}{\hbar} H_0 t} V_S |t\rangle_S \quad (11.18)$$

and

$$i\hbar \frac{\partial}{\partial t} |t\rangle_I = V_I |t\rangle_I \quad (11.19)$$

Also, we may obtain that

$$i\hbar \frac{d}{dt} A_I = [A_I, H_0] + i\hbar \frac{\partial A_I}{\partial t} \quad (11.20)$$

Using that $|t\rangle_S = U_S |t_0\rangle_S$ and defining the operator U_I as

$$|t\rangle_I = U_I |t_0\rangle_I \quad (11.21)$$

we get that

$$U_I = \bar{U} U_S \quad (11.22)$$

We may calculate the time evolution of the state in the interaction picture

$$i\hbar \frac{\partial}{\partial t} |t\rangle_I = V_I(t) U_I(t) |t_0\rangle_I \quad (11.23)$$

We obtain therefore that

$$i\hbar \frac{\partial}{\partial t} U_I(t) = V_I(t) U_I(t) \quad (11.24)$$

The solution of this equation follows similar steps to those for the evolution operator U_S . Using that $U_I(t_0, t_0) = 1$ we obtain that

$$U_I(t, t_0) = 1 + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar} \right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n V_I(t_1) V_I(t_2) \cdots V_I(t_n) \quad (11.25)$$

This is also a time ordered product. A significant improvement is that it involves powers of the interaction, assumed small. Therefore we may truncate the series to a few terms since the solution is expected to converge fast.

11.1 Susceptibilities and linear response theory

Let us consider now that we apply an external field to the system and we want to study the response of the system to the external probe. Let us then apply a field which in general will be time dependent. We have then to solve Schroedinger's

equation or same equivalent alternative way. Recall that we can define an operator in the so-called Heisenberg representation as

$$A_H(t) = U^{-1}(t)A_SU(t) \quad (11.26)$$

here H stands for Heisenberg, S for Schroedinger and $U(t)$ is the time evolution operator. (Recall that in the Schroedinger picture the states have all the time dependence and the operators have no explicit time dependence; in the Heisenberg representation the time dependence is in the operators and the states do not evolve in time). Let us now consider that at a given time, t_i , we apply an external field such that the Hamiltonian is given by

$$\mathcal{H}_H(t) = \mathcal{H} + \delta\mathcal{H} \quad (11.27)$$

where, as an example, $\delta\mathcal{H}$ is the result of applying a magnetic field, $\vec{h}(\vec{r}, t)$ like

$$\delta\mathcal{H} = - \int d^3r \vec{h}(\vec{r}, t) \cdot \vec{S}(\vec{r}) \quad (11.28)$$

here $\vec{S}(\vec{r})$ is the operator spin density. The magnetization is therefore given by

$$\vec{m}(\vec{r}, t) = \langle \vec{S}_H(\vec{r}, t) \rangle = \text{Tr} \left(\rho_H \vec{S}_H(\vec{r}, t) \right) \quad (11.29)$$

Here $\rho_H = \rho^{eq} = \frac{1}{Z} e^{-\beta\mathcal{H}}$ is the canonical distribution. Then

$$\vec{S}_H(\vec{r}, t) = U_{\mathcal{H}+\delta\mathcal{H}}^{-1}(t, t_i) \vec{S}_S(\vec{r}) U_{\mathcal{H}+\delta\mathcal{H}}(t, t_i) \quad (11.30)$$

If $\delta\mathcal{H} = 0$ and \mathcal{H} is independent of time the evolution operator is simply given by

$$U_{\mathcal{H}}(t, t_i) = e^{-\frac{i}{\hbar}\mathcal{H}(t-t_i)}$$

It is convenient to introduce the so-called interaction representation where

$$\begin{aligned} \vec{S}_I(\vec{r}, t) &= e^{\frac{i}{\hbar}\mathcal{H}(t-t_i)} \vec{S}_S(\vec{r}) e^{-\frac{i}{\hbar}\mathcal{H}(t-t_i)} \\ U_{\mathcal{H}+\delta\mathcal{H}}^{-1}(t, t_i) &= e^{-\frac{i}{\hbar}\mathcal{H}(t-t_i)} U_I(t, t_i) \end{aligned} \quad (11.31)$$

Using the equation of motion for the evolution operator

$$i\hbar \frac{dU}{dt} = (\mathcal{H} + \delta\mathcal{H})U \quad (11.32)$$

we obtain easily the equation of motion for the operator U_I in the form

$$i\hbar \frac{dU_I}{dt} = \delta\mathcal{H}_I U_I \quad (11.33)$$

This equation can be integrated

$$U_I(t, t_i) = 1 - \frac{i}{\hbar} \int_{t_i}^t dt' \delta\mathcal{H}_I(t') U_I(t', t_i) \quad (11.34)$$

Iterating to first order gives

$$U_I(t, t_i) = 1 - \frac{i}{\hbar} \int_{t_i}^t dt' \delta\mathcal{H}_I(t') + \dots \quad (11.35)$$

The magnetization can then be written as

$$\begin{aligned} \vec{m}(\vec{r}, t) &= \text{Tr} \left\{ \frac{1}{Z} e^{-\beta\mathcal{H}} U_{\mathcal{H}+\delta\mathcal{H}}^{-1} \vec{S}_S U_{\mathcal{H}+\delta\mathcal{H}} \right\} \\ &= \text{Tr} \left\{ \frac{1}{Z} e^{-\beta\mathcal{H}} U_I^{-1} \vec{S}_I U_I \right\} \end{aligned} \quad (11.36)$$

Replacing the expression for the evolution operator U_I and its inverse to linear order in the external field we obtain that

$$\vec{m}(\vec{r}, t) = \langle \vec{S}_I(\vec{r}, t) \rangle_{eq} + \frac{i}{\hbar} \int_{t_i}^t dt' \langle [\delta\mathcal{H}_I(t'), \vec{S}_I(\vec{r}, t)] \rangle_{eq} \quad (11.37)$$

Taking the fluctuation of the magnetization with respect to the equilibrium value without external field we get the response of the system to the external stimulus. This is given by

$$\delta m^p(\vec{r}, t) = \int_{t_i}^{\infty} \int d^3r' \sum_l \chi^{pl}(\vec{r}, \vec{r}'; t - t') h^l(\vec{r}', t') \quad (11.38)$$

where we have introduced the susceptibility $\chi^{pl}(\vec{r}, \vec{r}'; t - t')$ which the response function which relates the external field to the response of the system which in this example is the change in the magnetization. The susceptibility is given by

$$\chi^{pl}(\vec{r}, \vec{r}'; t - t') = \frac{i}{\hbar} \theta(t - t') \langle [S_I^p(\vec{r}, t), S_I^l(\vec{r}', t')] \rangle_{eq} \quad (11.39)$$

The subscript eq means that the average values are to be evaluated using the Boltzmann weight. The θ (step) function reflects the causality of the problem: there is no response before the external field is switched on. Note that in the time integration we can extend the lower limit to $-\infty$ since before t_i the perturbation is zero anyway. In equilibrium $\delta\mathcal{H} = 0$ and therefore the operators in the interaction and in the Heisenberg representations are equal. Therefore we may also rewrite

$$\chi^{lm}(\vec{r}, \vec{r}'; t - t') = \frac{i}{\hbar} \theta(t - t') < [S_H^l(\vec{r}, t), S_H^m(\vec{r}', t')] >_{eq} \quad (11.40)$$

It is easy to see that as expected

$$\chi^{lm}(\vec{r}, \vec{r}'; t - t') = \frac{\partial m_l(\vec{r}, t)}{\partial h_m(\vec{r}', t')} \quad (11.41)$$

11.1.1 Kramers-Kronig relations

Let us write

$$\begin{aligned} \chi^{lm}(\vec{r}, t; \vec{r}', t') &= 2i\theta(t - t') \chi_{lm}''(\vec{r}, t; \vec{r}', t') \\ \chi_{lm}''(\vec{r}, t; \vec{r}', t') &= \frac{1}{2\hbar} < [S_H^l(\vec{r}, t), S_H^m(\vec{r}', t')] >_{eq} \end{aligned} \quad (11.42)$$

Assuming that the system is uniform in space and time we get that

$$\begin{aligned} \chi_{lm}(\vec{r}, t; \vec{r}', t') &= \chi_{lm}(\vec{r} - \vec{r}'; t - t') \\ \chi_{lm}''(\vec{r}, t; \vec{r}', t') &= \chi_{lm}''(\vec{r} - \vec{r}'; t - t') \end{aligned} \quad (11.43)$$

Taking $t_i \rightarrow -\infty$ and performing a Fourier transform we can write that

$$\chi_{ij}(\vec{k}, \omega) = \int d(t - t') \int d^3(x - x') e^{-i\vec{k} \cdot (\vec{r} - \vec{r}') + i\omega(t - t')} \chi_{ij}(\vec{r} - \vec{r}'; t - t') \quad (11.44)$$

and

$$\delta m_i(\vec{k}, \omega) = \sum_j \chi_{ij}(\vec{k}, \omega) h_j(\vec{k}, \omega) \quad (11.45)$$

To simplify notation let us focus now our attention on the time dependence of the susceptibilities and let us drop the matrix indices. We write then $\chi(t - t') =$

$2i\theta(t-t')\chi''(t-t')$. We can write the following sequence of identities using that the function is causal

$$\begin{aligned}
\chi(\omega) &= 2i \int_0^\infty d(t-t') e^{i\omega(t-t')} \chi''(t-t') \\
&= 2i \int_0^\infty dt e^{i\omega t} \chi''(t) \\
&= 2i \int_0^\infty dt e^{i\omega t} \int_{-\infty}^\infty \frac{d\omega'}{2\pi} e^{-i\omega' t} \chi''(\omega') \\
&= 2i \int_{-\infty}^\infty \frac{d\omega'}{2\pi} \chi''(\omega') \int_0^\infty dt e^{-i(\omega' - \omega - i\eta)t}, \eta \rightarrow 0^+ \\
&= 2i \int_{-\infty}^\infty \frac{d\omega'}{2\pi} \chi''(\omega') \frac{1}{i(\omega' - \omega - i\eta)} \\
\chi(\omega) &= \int_{-\infty}^\infty \frac{d\omega'}{\pi} \frac{\chi''(\omega')}{\omega' - \omega - i\eta}
\end{aligned} \tag{11.46}$$

Using that

$$\frac{1}{x \pm i\eta} = P \frac{1}{x} \mp i\pi \delta(x) \tag{11.47}$$

where P stands for principal part and $\delta(x)$ is the Dirac delta function written as

$$\begin{aligned}
P \frac{1}{x} &= \lim_{\eta \rightarrow 0^+} \frac{x}{x^2 + \eta^2} \\
\delta(x) &= \lim_{\eta \rightarrow 0^+} \frac{1}{\pi} \frac{\eta}{x^2 + \eta^2}
\end{aligned}$$

we can write that

$$\chi(\omega) = P \int \frac{d\omega'}{\pi} \frac{\chi''(\omega')}{\omega' - \omega} + i\chi''(\omega) \tag{11.48}$$

If $\chi''(\omega)$ is real (as we will show later) then

$$\begin{aligned}
Im\chi(\omega) &= \chi''(\omega) \\
Re\chi(\omega) &= \chi'(\omega) = P \int \frac{d\omega'}{\pi} \frac{\chi''(\omega')}{\omega' - \omega}
\end{aligned} \tag{11.49}$$

which implies a relation between the real and the imaginary parts of the dynamical susceptibility $\chi(\omega)$. This relation is the result of the causality of the response function. This constitutes one of the Kramers-Kronig relations that relate the real and imaginary parts of the frequency part of a causal function.

Since $\chi(t) = 2i\theta(t)\chi''(t)$ we can write that

$$\chi(\omega) = 2i \int_0^\infty dt \chi''(t) e^{i\omega t} \quad (11.50)$$

Taking the frequency as a complex variable we can write that $\omega = \omega_r + i\omega_i$. Than we conclude that

$$|e^{i(\omega_r + i\omega_i)t}| = |e^{i\omega_r t} e^{-\omega_i t}| = e^{-\omega_i t}$$

Therefore, if $\chi(\omega)$ exists for real frequency ($\omega = \omega_r$) than it also exists in the upper complex plane ($\omega_i > 0$). Therefore, $\chi(\omega)$ is analytical in the upper half-plane. Considering a contour, $C = C_1 + C_2$, in the complex plane that includes the real axis, C_1 , and a semi-circle of infinite radius in the upper half-plane, C_2 , we may write

$$\frac{1}{2\pi i} \oint_C \frac{\chi(z')}{z' - z} dz' = \chi(z) \quad (11.51)$$

If $\chi(z)$ decays at least as fast as $1/z$ as $|z| \rightarrow \infty$ then the integral C_2 vanishes and we obtain that

$$\chi(z) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} d\omega' \frac{\chi(\omega')}{\omega' - z} \quad (11.52)$$

where ω' is the real axis. Taking $z = \omega + i\eta$ and $\eta \rightarrow 0$ we obtain that

$$\frac{1}{2\pi i} P \int_{-\infty}^{\infty} d\omega' \frac{\chi(\omega')}{\omega' - \omega} + \frac{1}{2} \chi(\omega) = \chi(\omega) \quad (11.53)$$

and

$$\chi(\omega) = -iP \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi(\omega')}{\omega' - \omega} \quad (11.54)$$

Writing as before $\chi(\omega) = \chi'(\omega) + i\chi''(\omega)$ we arrive at the Kramers-Kronig relations between the real and imaginary parts of the susceptibility

$$\begin{aligned} \chi'(\omega) &= P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi''(\omega')}{\omega' - \omega} \\ \chi''(\omega) &= -P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi'(\omega')}{\omega' - \omega} \end{aligned} \quad (11.55)$$

11.1.3 Dissipation

Let us calculate the energy dissipated in the system due to the presence of the external perturbation. The coupling of the external source to the system degrees of freedom may be written as

$$H_{int} = - \sum_i A_i F_i(t) \quad (11.71)$$

where we consider that the external sources represented by the terms $F_i(t)$ are taken as classical and the internal variables are represented by the hermitian operators A_i . The energy dissipated in the system is given by

$$W = \int_{-\infty}^{\infty} dt \left\langle \frac{\partial H}{\partial t} \right\rangle_{n-e} = - \int_{-\infty}^{\infty} dt \sum_i \left\langle A_i \frac{\partial F_i(t)}{\partial t} \right\rangle = \int_{-\infty}^{\infty} dt \sum_i \left\langle \frac{\partial A_i}{\partial t} \right\rangle F_i(t) \quad (11.72)$$

The averages are taken in a non-equilibrium system due to the presence of the time-dependent perturbation. We have integrated by parts assuming that the sources vanish at $t = \pm\infty$. In linear response theory we know that

$$\langle A_i(t) \rangle_{n-e} = \langle A_i(t) \rangle_{eq} + \sum_j \int_{-\infty}^{\infty} dt' \chi_{ij}(t-t') F_j(t') \quad (11.73)$$

The average in equilibrium is time independent. Therefore,

$$\begin{aligned} W &= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' \sum_{ij} F_i(t) \frac{\partial}{\partial t} \chi_{ij}(t-t') F_j(t') \\ &= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' \sum_{ij} F_i(t) \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')} (-i\omega) \chi_{ij}(\omega) F_j(t') \\ &= \sum_{ij} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F_i^*(\omega) (-i\omega) \chi_{ij}(\omega) F_j(\omega) \end{aligned} \quad (11.74)$$

Since $F_i(t)$ is real it implies that $F_i^*(\omega) = F_i(-\omega)$. Making the substitutions $\omega \rightarrow -\omega$ and $i \leftrightarrow j$ we may as well write that

$$W = \sum_{ij} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F_i^*(\omega) i\omega \chi_{ji}(-\omega) F_j(\omega) \quad (11.75)$$

which leads to

$$W = \sum_{ij} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F_i^*(\omega) i\omega \frac{1}{2} (-\chi_{ij}(\omega) + \chi_{ji}(-\omega)) F_j(\omega) \quad (11.76)$$

We may now write that

$$\begin{aligned} \chi_{ji}(-\omega) &= P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi_{ji}''(\omega')}{\omega' + \omega} + i\chi_{ji}''(-\omega) \\ &= -P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi_{ji}''(-\omega')}{\omega' - \omega} + i\chi_{ji}''(-\omega) \\ &= -P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi_{ij}''(\omega')}{\omega' - \omega} + i\chi_{ji}''(-\omega) \end{aligned} \quad (11.77)$$

This implies that

$$\frac{1}{2} (-\chi_{ij}(\omega) + \chi_{ji}(-\omega)) = \frac{1}{2} (-i\chi_{ij}''(\omega) + i\chi_{ji}''(-\omega)) = -i\chi_{ij}''(\omega)$$

Therefore, we arrive at the result that

$$W = \sum_{ij} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F_i^*(\omega) \omega \chi_{ij}''(\omega) F_j(\omega) \quad (11.78)$$

which shows that the dissipation is controlled by the imaginary part of the susceptibility, as expected.

Consider now that $F_i = a_i e^{i\omega t}$. We obtain then

$$\sum_{ij} a_i^* \omega \chi_{ij}''(\omega) a_j \geq 0 \quad (11.79)$$

Therefore,

$$\omega \chi_{ii}'' \geq 0 \quad (11.80)$$

The definition of susceptibility may be extended to two operators as

$$\chi_{AB}''(t - t') = \frac{1}{2\hbar} \langle [A(t), B(t')] \rangle \quad (11.81)$$

It follows that

$$\chi''_{AB}(t - t') = -\chi''_{BA}(t' - t) \quad (11.82)$$

and

$$\chi''_{AB}(\omega) = -\chi''_{BA}(-\omega) \quad (11.83)$$

If the two operators A and B are hermitian

$$\begin{aligned} (\chi''_{AB}(t - t'))^* &= \frac{1}{2\hbar} \langle [A(t), B(t')] \rangle^* \\ &= \frac{1}{2\hbar} \text{Tr} [\rho [A(t), B(t')]]^\dagger \\ &= \frac{1}{2\hbar} \text{Tr} [[A(t), B(t')]^\dagger \rho^\dagger] \\ &= \frac{1}{2\hbar} \text{Tr} [\rho [B(t'), A(t)]] \\ &= \chi''_{BA}(t' - t) \end{aligned} \quad (11.84)$$

Also,

$$\begin{aligned} (\chi''_{AB}(\omega))^* &= \int d(t - t') e^{-i\omega(t - t')} (\chi''_{AB}(t - t'))^* \\ &= \int d(t' - t) \chi''_{BA}(t' - t) e^{i\omega(t' - t)} \\ &= \chi''_{BA}(\omega) \end{aligned} \quad (11.85)$$

Therefore,

$$(\chi''_{AB}(\omega))^* = \chi''_{BA}(\omega) = -\chi''_{AB}(-\omega) \quad (11.86)$$

If the two operators are the same, $B = A$, we get that χ''_{AA} is real, as we saw above.

11.1.4 Fluctuation-dissipation theorem

The dissipation is associated with the imaginary part of the susceptibility which, as we found before, is given by a commutator of two operators

$$\chi''_{AB}(t - t') = \frac{1}{2\hbar} \langle [A(t), B(t')] \rangle \quad (11.87)$$

On the other hand, a perturbation leads to a response in the system as measured in a dispersion experiment. Typically this involves correlation functions of the type

$$\bar{S}_{AB}(t-t') = \langle A(t)B(t') \rangle \quad (11.88)$$

These two types of correlation functions are related leading to a correlation between dissipation in a system and its response in a scattering experiment.

In thermal equilibrium the density matrix, ρ , is given by the Boltzmann exponential and we obtain

$$\begin{aligned} \bar{S}_{AB}(t-t') &= \frac{1}{Z} \text{Tr} e^{-\beta H} A(t) B(t') \\ &= \frac{1}{Z} \text{Tr} e^{-\beta H} A(t) e^{\beta H} e^{-\beta H} B(t') \\ &= \frac{1}{Z} \text{Tr} e^{-\beta H} e^{\frac{i}{\hbar} H t} A(0) e^{-\frac{i}{\hbar} H t} e^{\beta H} e^{-\beta H} B(t') \\ &= \frac{1}{Z} \text{Tr} e^{-\beta H} B(t') A(t + i\beta\hbar) \\ \bar{S}_{AB}(t-t') &= \bar{S}_{BA}(t' - t - i\beta\hbar) \end{aligned} \quad (11.89)$$

Carrying a Fourier transform we get that

$$\bar{S}_{AB}(\omega) = \int_{-\infty}^{\infty} d(t-t') e^{i\omega(t-t')} \bar{S}_{BA}(t' - t - i\beta\hbar) \quad (11.90)$$

Making a change of variables $t' - t \rightarrow t$, and then another transformation $t \rightarrow t + i\beta\hbar$ we get that

$$\begin{aligned} \bar{S}_{AB}(\omega) &= e^{\beta\hbar\omega} \int_{-\infty}^{\infty} dt e^{-i\omega t} \bar{S}_{BA}(t) \\ &= e^{\beta\hbar\omega} \bar{S}_{BA}(-\omega) \end{aligned} \quad (11.91)$$

Note that

$$\chi''_{AB}(t-t') = \frac{1}{2\hbar} \langle [A(t), B(t')] \rangle = \frac{1}{2\hbar} [\bar{S}_{AB}(t-t') - \bar{S}_{BA}(t'-t)] \quad (11.92)$$

In terms of their Fourier transforms this can be expressed

$$\begin{aligned} \chi''_{AB}(\omega) &= \frac{1}{2\hbar} [\bar{S}_{AB}(\omega) - \bar{S}_{BA}(-\omega)] \\ \chi''_{AB}(\omega) &= \frac{1}{2\hbar} (1 - e^{-\beta\hbar\omega}) \bar{S}_{AB}(\omega) \end{aligned} \quad (11.93)$$

which relates the susceptibility and the scattering cross section of the system degrees of freedom.

Consider now the fluctuation of an operator around its mean value, $\delta A = A - \langle A \rangle$. We may define a correlation function of the fluctuations of two operators A and B and take their symmetrized product

$$S_{AB}(t - t') = \frac{1}{2} \langle \{ \delta A(t), \delta B(t') \} \rangle \quad (11.94)$$

It is straightforward to see that

$$S_{AB}(t - t') = \frac{1}{2} (\bar{S}_{AB}(t - t') + \bar{S}_{BA}(t' - t)) \quad (11.95)$$

plus a possible constant term. In Fourier space we get that

$$\begin{aligned} S_{AB}(\omega) &= \frac{1}{2} (\bar{S}_{AB}(\omega) + \bar{S}_{BA}(-\omega)) \\ &= \frac{1}{2} (1 + e^{-\beta\hbar\omega}) \bar{S}_{AB}(\omega) \end{aligned} \quad (11.96)$$

Using the relation to the imaginary part of the susceptibility we get that

$$\begin{aligned} S_{AB}(\omega) &= \frac{1}{2} (1 + e^{-\beta\hbar\omega}) 2\hbar \frac{1}{1 - e^{-\beta\hbar\omega}} \chi''_{AB}(\omega) \\ S_{AB}(\omega) &= \hbar \coth \frac{\beta\hbar\omega}{2} \chi''_{AB}(\omega) \end{aligned} \quad (11.97)$$

This relation is called the fluctuation-dissipation theorem since it relates the fluctuations expressed in $S_{AB}(\omega)$ with the dissipation of the system expressed in $\chi''_{AB}(\omega)$.

11.2 Conductivity. Kubo formulas

Linear response theory may also be used to determine the current of a system of particles in the presence of an external field. An example is the charge current due to the presence of an electric field. The electric field may be the result of a potential gradient or the result of a time varying vector potential.

In the presence of a vector potential the minimal coupling of a particle of charge e_i to the electromagnetic field leads to the change $\vec{p} \rightarrow \vec{p} - e_i/c\vec{A}$. The Hamiltonian

Capítulo 16

16.1 Quantificação do campo electromagnético

Comecemos por recordar o caso clássico em que uma partícula de carga q e massa m está na presença de um campo electromagnético. A quantificação deste problema servirá de exemplo de base para a (segunda) quantificação de outros sistemas, como veremos adiante. A sua equação de movimento é dada por

$$m \frac{d\mathbf{v}}{dt} = q \left(\mathbf{E}(\mathbf{r}) + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right) \quad (16.1)$$

em que \mathbf{E} é o campo eléctrico e \mathbf{B} é o campo magnético. Os campos eléctrico e magnético obedecem às equações de Maxwell

$$\begin{aligned} \nabla \cdot \mathbf{B} &= 0 \\ \nabla \cdot \mathbf{E} &= 4\pi\rho \\ \nabla \times \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} &= \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi\mathbf{J}}{c} \end{aligned} \quad (16.2)$$

em que ρ é a densidade de carga e \mathbf{J} a densidade de corrente. Os campos eléctrico e magnético podem ser derivados a partir dos potenciais escalar e vector na forma

$$\begin{aligned} \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \nabla\Phi \\ \mathbf{B} &= \nabla \times \mathbf{A} \end{aligned} \quad (16.3)$$

Como recordámos anteriormente, os campos são definidos a menos de uma transformação de gauge, visto \vec{E} e \vec{B} permanecerem invariantes para uma transformação simultânea da forma

$$\begin{aligned}\Phi &\rightarrow \Phi - \frac{1}{c} \frac{\partial f}{\partial t} \\ \mathbf{A} &\rightarrow \mathbf{A} + \nabla f\end{aligned}\quad (16.4)$$

em que $f(\mathbf{r}, t)$ é uma função arbitrária das coordenadas espaciais e do tempo. Existe por isso uma infinidade de escolhas possíveis para os potenciais consistentes com as equações de Maxwell. Uma dada escolha corresponde a uma dada escolha de gauge.

Uma escolha conveniente corresponde a tomar

$$\nabla \cdot \mathbf{A} = 0$$

chamada gauge de Coulomb, também chamada de gauge transversa. Com esta escolha o potencial escalar é instantâneo e obedece à equação

$$\Phi(\mathbf{r}, t) = \int d^3r' \frac{\rho(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} \quad (16.5)$$

Por seu turno o potencial vector obedece à equação

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{A} = -\frac{4\pi}{c} \mathbf{J}_t \quad (16.6)$$

em que \mathbf{J}_t é a parte transversa da corrente definida por

$$4\pi \mathbf{J}_t = \nabla \times \nabla \times \int d^3r' \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

Outra gauge usualmente escolhida é a gauge de Lorentz em que

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} = 0 \quad (16.7)$$

As equações de movimento de um sistema de partículas na presença de um campo electromagnético podem ser obtidas a partir do Lagrangeano

$$L = \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \int \frac{d^3r}{8\pi} (\mathbf{E}(\mathbf{r})^2 - \mathbf{B}(\mathbf{r})^2) - \sum_i q_i \Phi(\mathbf{r}_i) + \sum_i \frac{q_i}{c} \mathbf{r}_i \cdot \mathbf{A}(\mathbf{r}_i) \quad (16.8)$$

em que q_i é a carga da partícula i . As equações de Euler-Lagrange para as coordenadas $q_j = \Phi, \mathbf{A}, \mathbf{r}$

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{q}_j} + \sum_{\alpha=1}^3 \frac{d}{dr_\alpha} \frac{\delta L}{\partial(\partial q_j / \partial r_\alpha)} - \frac{\delta L}{\delta q_j} = 0 \quad (16.9)$$

conduzem à equação, variando em ordem a $q_j = \Phi$

$$\nabla \cdot \mathbf{E} = -4\pi\rho \quad (16.10)$$

Variando em ordem a $q_j = \mathbf{A}$ obtém-se

$$\nabla \times \mathbf{B} = \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi \mathbf{J}}{c} \quad (16.11)$$

Usou-se que $\rho(\mathbf{r}) = \sum_i q_i \delta(\mathbf{r} - \mathbf{r}_i)$ e que os momentos conjugados definidos da forma usual

$$P_{q_j} = \frac{\partial L}{\partial \dot{q}_j}$$

são dados por

$$\begin{aligned} P_\Phi &= 0 \\ P_{\mathbf{A}} &= -\frac{\mathbf{E}}{4\pi c} \end{aligned} \quad (16.12)$$

Variando o Lagrangeano em ordem às coordenadas espaciais, \mathbf{r} , obtém-se a equação de movimento das partículas sob a acção da força de Lorentz.

A construção do Hamiltoniano segue os passos usuais. Define-se o Hamiltoniano pela relação

$$H = \sum_i (\dot{q}_j)_i P_{q_j} - L \quad (16.13)$$

O Hamiltoniano pode ser escrito na forma

$$H = \frac{1}{2m} \sum_i \left(\mathbf{p}_i - \frac{q_i}{c} \mathbf{A}(\mathbf{r}_i) \right)^2 + \int \frac{d^3r}{8\pi} (\mathbf{E}^2 + \mathbf{B}^2) \quad (16.14)$$

O primeiro termo descreve a parte da energia cinética das partículas e a sua interacção com o campo electromagnético e o segundo termo descreve a energia do campo electromagnético.

Pretendemos agora quantificar este problema e em particular quantificar o campo electromagnético. Uma forma é proceder à quantificação canónica, conforme descrito na secção 5.3. Para quantificar o campo impomos relações de comutação entre as coordenadas generalizadas e os momentos conjugados. Como vimos o momento associado ao potencial escalar é nulo e por isso não é quantificado. No entanto o momento associado ao potencial vector é proporcional ao campo eléctrico. Este resultado implica que as componentes do potencial vector têm relações de comutação com as componentes do campo eléctrico. Vamos então impor as relações de comutação

$$\boxed{[A_\alpha(\mathbf{r}, t), -\frac{E_\beta}{4\pi c}(\mathbf{r}', t)] = i\hbar\delta_{\alpha,\beta}\delta(\mathbf{r} - \mathbf{r}')} \quad (16.15)$$

Esta relação de comutação é automaticamente satisfeita com a escolha

$$A_\alpha(\mathbf{r}, t) = \sum_{\mathbf{k}, \lambda} \left(\frac{2\pi\hbar c^2}{v\omega_k} \right)^{1/2} u_\alpha(\mathbf{k}, \lambda) e^{i\mathbf{k}\cdot\mathbf{r}} \left(a_{\mathbf{k}, \lambda} e^{-i\omega_k t} + a_{-\mathbf{k}, \lambda}^\dagger e^{i\omega_k t} \right) \quad (16.16)$$

em que λ representa os estados de polarização da onda (dois modos transversais) e os operadores $a_{\mathbf{k}}$ obedecem as relações de comutação

$$\boxed{[a_{\mathbf{k}, \lambda}, a_{\mathbf{k}', \lambda'}^\dagger] = \delta_{\mathbf{k}, \mathbf{k}'} \delta_{\lambda, \lambda'}} \quad (16.17)$$

e

$$\boxed{[a_{\mathbf{k}, \lambda}, a_{\mathbf{k}', \lambda'}] = 0} \quad (16.18)$$

O campo eléctrico obtém-se do potencial vector usando

$$E_\alpha(\mathbf{r}, t) = -\frac{1}{c} \dot{A}_\alpha - \nabla_\alpha \Phi \quad (16.19)$$

Usando a relação de comutação entre as componentes do potencial vector e do campo eléctrico obtemos as relações de comutação

$$[A_\alpha(\mathbf{r}, t), \dot{A}_\beta(\mathbf{r}', t)] = 4\pi c^2 i \hbar \delta_{\alpha,\beta} \delta(\mathbf{r} - \mathbf{r}') \quad (16.20)$$

As energias $\omega_k = ck$ são as energias dos fotões. O potencial vector é assim expandido nos modos do campo electromagnético. A quantificação destes modos conduz a uma expansão dos campos (agora operadores) em termos dos operadores de criação e

destruição do oscilador harmónico, que como veremos são bósons. Portanto cada estado de fótons é caracterizado pelos números de ocupação de cada estado individual na forma usual

$$|n_{\mathbf{k},\lambda}\rangle = \frac{(a_{\mathbf{k},\lambda}^\dagger)^{n_{\mathbf{k},\lambda}}}{(n_{\mathbf{k},\lambda})^{1/2}} |0\rangle \quad (16.21)$$

Os operadores $(a_{\mathbf{k},\lambda}^\dagger)$ criam um fóton com momento \mathbf{k} e polarização λ , de acordo com a interpretação de Einstein, como vimos no contexto do corpo negro. A energia total do campo de fótons é então dada por

$$\sum_{\mathbf{k},\lambda} \hbar\omega_k(n_{\mathbf{k},\lambda} + \frac{1}{2})$$

A energia total do sistema é então dada por

$$H = \sum_i \frac{1}{2m} \left(\mathbf{p}_i - \frac{q_i}{c} \mathbf{A}(\mathbf{r}_i) \right)^2 + \frac{1}{2} \sum_{i,j} \frac{q_i q_j}{r_{ij}} + \sum_{\mathbf{k},\lambda} \hbar\omega_k(n_{\mathbf{k},\lambda} + \frac{1}{2}) \quad (16.22)$$

16.2 Quantificação da função de onda

Uma outra forma de entender a segunda quantificação consiste no seguinte. Recordemos que numa formulação variacional da equação de Schrödinger dependente do tempo concluímos que o campo conjugado da função de onda $\psi(\mathbf{r})$ é $i\hbar\psi^*(\mathbf{r})$ (ver Eq. (13.160)). Vamos então quantificar a função de onda na forma canónica usual, ou seja, vamos impôr a relação de comutação

$$[\psi(\mathbf{r}, t), i\hbar\psi^\dagger(\mathbf{r}', t)] = i\hbar\delta(\mathbf{r} - \mathbf{r}') \quad (16.23)$$

e portanto

$$[\psi(\mathbf{r}, t), \psi^\dagger(\mathbf{r}', t)] = \delta(\mathbf{r} - \mathbf{r}'). \quad (16.24)$$

Em geral, a função de onda pode-se expandir numa base completa de funções $\Phi_\lambda(\mathbf{r})$, com coeficientes $a_\lambda(t)$

$$\psi(\mathbf{r}, t) = \sum_\lambda a_\lambda(t) \Phi_\lambda(\mathbf{r}) \quad (16.25)$$

Da mesma forma

$$\psi^*(\mathbf{r}, t) = \sum_{\lambda} a_{\lambda}^*(t) \Phi_{\lambda}^*(\mathbf{r}) \quad (16.26)$$

Ao impormos as relações de comutação à função de onda transformamos a função de onda num operador $\psi \rightarrow \hat{\psi}$ (a este processo designa-se por isso de segunda quantificação. A primeira quantificação resulta de promover os observáveis clássicos descritos por variáveis clássicas, comutativas, a operadores que em geral não comutam entre si). Como a função de onda passa a ser interpretada como um operador, os coeficientes $a_{\lambda}(t)$ passam eles próprios a operadores. A relação de comutação 16.24 impõe relações de comutação entre estes operadores na forma

$$\begin{aligned} [a_{\lambda}(t), a_{\lambda'}^{\dagger}(t)] &= \delta_{\lambda, \lambda'} \\ [a_{\lambda}(t), a_{\lambda'}(t)] &= 0 \\ [a_{\lambda}^{\dagger}(t), a_{\lambda'}^{\dagger}(t)] &= 0 \end{aligned} \quad (16.27)$$

Consideremos então, por exemplo, a energia e o operador que lhe corresponde o Hamiltoniano, \mathcal{H} . A energia de um dado estado descrito pela função de onda $\psi(\mathbf{r})$ em primeira quantificação é dada pela forma usual

$$E = \int d^3r \psi^*(\mathbf{r}) \mathcal{H} \psi(\mathbf{r}) \quad (16.28)$$

Define-se então um operador Hamiltoniano em segunda quantificação, \hat{H} , fazendo a transposição

$$\hat{H} = \int d^3r \hat{\psi}^{\dagger}(\mathbf{r}) \mathcal{H} \hat{\psi}(\mathbf{r}) \quad (16.29)$$

Obtém-se então

$$\begin{aligned} \hat{H} &= \int d^3r \sum_{\lambda} a_{\lambda}^{\dagger}(t) \Phi_{\lambda}^*(\mathbf{r}) \mathcal{H} \sum_{\lambda'} a_{\lambda'}(t) \Phi_{\lambda'}(\mathbf{r}) \\ &= \sum_{\lambda, \lambda'} \langle \lambda | \mathcal{H} | \lambda' \rangle a_{\lambda}^{\dagger}(t) a_{\lambda'}(t) \end{aligned} \quad (16.30)$$

em que

$$\langle \lambda | \mathcal{H} | \lambda' \rangle = \int d^3r \Phi_{\lambda}^*(\mathbf{r}) \mathcal{H} \Phi_{\lambda'}(\mathbf{r}) \quad (16.31)$$

No caso de se escolher como base completa de funções as funções próprias do Hamiltoniano

$$\mathcal{H}\Phi_\lambda = \epsilon_\lambda \Phi_\lambda \quad (16.32)$$

o Hamiltoniano toma a forma simplificada

$$\hat{H} = \sum_\lambda \epsilon_\lambda a_\lambda^\dagger a_\lambda \quad (16.33)$$

No caso de uma partícula livre $\lambda = \mathbf{k}$

$$\hat{H} = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \quad (16.34)$$

Podemos então interpretar $a_\lambda^\dagger a_\lambda$ como um operador número de partículas com momento \mathbf{k} . A energia de um sistema livre é a soma para todos os momentos da energia de cada momento vezes o número de partículas com esse momento. Em geral o operador $a_\lambda^\dagger a_\lambda$ conta o número de partículas no estado λ . Este procedimento aplica-se a qualquer operador.

É então natural definir um operador hermitico que nos dá o número de partículas no sistema:

$$N = \int \psi^\dagger \psi \, d^3r \quad (16.35)$$

Provemos agora que o operador assim definido é uma constante do movimento. A sua evolução temporal é dada por:

$$i\hbar \dot{N} = [N, H]$$

O Hamiltoniano é escrito como:

$$H = \int \left(\frac{\hbar^2}{2m} \nabla \psi^\dagger \cdot \nabla \psi + V \psi^\dagger \psi \right) d^3r$$

pelo que a evolução temporal é dada por:

$$i\hbar \dot{N} = \left[\int \psi_1^\dagger \psi_1 \, d^3r_1, \int \left(\frac{\hbar^2}{2m} \nabla_2 \psi_2^\dagger \cdot \nabla_2 \psi_2 + V_2 \psi_2^\dagger \psi_2 \right) d^3r_2 \right] \quad (16.36)$$

em que $V_2 = V(\mathbf{r}_2)$ é uma função numérica e $\psi_i = \psi(\mathbf{r}_i)$.

Para calcular o comutador da Equação (16.36) começamos pelo termo relativo ao potencial:

$$\iint V_2(\psi_1^\dagger \psi_1 \psi_2^\dagger \psi_2 - \psi_2^\dagger \psi_2 \psi_1^\dagger \psi_1) d^3 r_1 d^3 r_2 \quad (16.37)$$

Tendo em conta que

$$[\psi_1, \psi_2^\dagger] = \delta^3(\mathbf{r}_1 - \mathbf{r}_2) \Rightarrow \psi_1 \psi_2^\dagger = \psi_2^\dagger \psi_1 + \delta^3(\mathbf{r}_1 - \mathbf{r}_2) \quad (16.38)$$

a integranda da Equação (16.37) fica:

$$\psi_1^\dagger \psi_1 \psi_2^\dagger \psi_2 - \psi_2^\dagger \psi_2 \psi_1^\dagger \psi_1 = \psi_1^\dagger [\psi_2^\dagger \psi_1 + \delta^3(\mathbf{r}_1 - \mathbf{r}_2)] \psi_2 - \psi_2^\dagger \psi_2 \psi_1^\dagger \psi_1$$

Uma vez que $[\psi_1, \psi_2] = [\psi_1^\dagger, \psi_2^\dagger] = 0$, o termo anterior fica

$$\begin{aligned} & \psi_1^\dagger \psi_2^\dagger \psi_2 \psi_1 + \psi_1^\dagger \psi_2 \delta^3(\mathbf{r}_1 - \mathbf{r}_2) - \psi_2^\dagger \psi_2 \psi_1^\dagger \psi_1 \\ &= \psi_2^\dagger [\psi_2 \psi_1^\dagger - \delta^3(\mathbf{r}_1 - \mathbf{r}_2)] \psi_1 + \psi_1^\dagger \psi_2 \delta^3(\mathbf{r}_1 - \mathbf{r}_2) - \psi_2^\dagger \psi_2 \psi_1^\dagger \psi_1 \\ &= (\psi_2^\dagger \psi_1 + \psi_1^\dagger \psi_2) \delta^3(\mathbf{r}_1 - \mathbf{r}_2) = 0 \end{aligned}$$

De modo análogo, pode mostrar-se que

$$[\psi_1^\dagger \psi_1, \nabla_2 \psi_2^\dagger \cdot \nabla_2 \psi_2] = [\psi_1^\dagger \nabla_2 \psi_2 - (\nabla_2 \psi_2^\dagger) \psi_1] \cdot \nabla_2 \delta^3(\mathbf{r}_1 - \mathbf{r}_2)$$

cujo integral em $d^3 r_1$ e $d^3 r_2$ é nulo.

Assim, concluímos finalmente que $[N, H] = 0$, pelo que N é uma constante do movimento, como queríamos demonstrar.

Procuremos agora uma representação em que N é diagonal. Para tal, exprimimos a função de onda como uma combinação linear de funções que formam uma base ortonormal,

$$\int u_k(\mathbf{r}) u_l(\mathbf{r}) d^3 r = \delta_{kl} \quad (16.39)$$

Os coeficientes são operadores escritos na representação de Heisenberg,

$$\psi(\mathbf{r}, t) = \sum_k a_k(t) u_k(\mathbf{r}) \quad e \quad \psi^\dagger(\mathbf{r}, t) = \sum_k a_k^\dagger(t) u_k^*(\mathbf{r}) \quad (16.40)$$

visto estarmos a considerar operadores com dependência explícita no tempo.

Multiplicando à esquerda por $u_k^*(\mathbf{r})$ e $u_k(\mathbf{r})$, respectivamente, e integrando,

$$a_k(t) = \int u_k^*(\mathbf{r})\psi(\mathbf{r}, t) d^3r \quad e \quad a_k^\dagger(t) = \int u_k(\mathbf{r})\psi^\dagger(\mathbf{r}, t) d^3r$$

O comutador destes dois operadores pode ser facilmente calculado,

$$[a_k(t), a_l^\dagger(t)] = \iint u_k^*(\mathbf{r}_1)u_l(\mathbf{r}_2)(\psi_1\psi_2^\dagger - \psi_2^\dagger\psi_1) d^3r_1 d^3r_2$$

Pela relação (16.38),

$$\psi_1\psi_2^\dagger - \psi_2^\dagger\psi_1 = [\psi_1, \psi_2^\dagger] = \delta^3(\mathbf{r}_1 - \mathbf{r}_2)$$

e ficamos com

$$[a_k(t), a_l^\dagger(t)] = \iint u_k^*(\mathbf{r}_1)u_l(\mathbf{r}_2)\delta^3(\mathbf{r}_1 - \mathbf{r}_2)d^3r_1 d^3r_2 = \delta_{kl} \quad (16.41)$$

onde se usou a ortogonalidade das funções de onda, equação (16.39).

Substituindo as funções de onda assim escritas, a Equação (16.40) na definição de N , e a Equação (16.35):

$$N = \int \sum_k a_k^\dagger(t)u_k^*(\mathbf{r}) \sum_l a_l(t)u_l(\mathbf{r})d^3r = \sum_k a_k^\dagger a_k = \sum_k N_k$$

em que se definiu $N_k = a_k^\dagger a_k$. Interpretando k como o momento de uma partícula, recuperamos a representação introduzida anteriormente.

Os operadores a_k e a_k^\dagger com as relações de comutação da equação (16.41) são de novo semelhantes aos operadores do oscilador harmónico. Tendo em conta as propriedades encontradas para aquele problema, os estados quânticos na representação em que N_k é diagonal são representados por: $|n_1, \dots, n_k, \dots\rangle$, onde n_k são os valores próprios de N_k , inteiros positivos (o valor próprio n_k indica-nos quantas partículas existem no estado k).

A aplicação dos operadores a_k e a_k^\dagger aos estados próprios de N_k é idêntica à encontrada para o oscilador harmónico, i.e.,

$$a_k |n_1, \dots, n_k, \dots\rangle = \sqrt{n_k} |n_1, \dots, n_k - 1, \dots\rangle \quad (16.42)$$

$$a_k^\dagger |n_1, \dots, n_k, \dots\rangle = \sqrt{n_k + 1} |n_1, \dots, n_k + 1, \dots\rangle \quad (16.43)$$

Capítulo 14

Microscopic description of interacting bosonic and fermionic systems

14.1 Second quantization

Let us then consider a system of N identical particles, contained in a box of volume Ω , with periodic boundary conditions. To describe the state of the system we can either use a wave function $\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ or else we can specify the state of the system indicating the number of particles, $n_{\vec{p}}$, in a given one-particle state described by a plane wave of momentum \vec{p} . We can then characterize the state of the system by the vector

$$|n_1, \dots, n_p, \dots\rangle \quad (14.1)$$

This state is given by a symmetrical or antisymmetrical product of plane waves, one for each particle. As said above this is not a very convenient description to carry out calculations.

Let us define an operator, called creation operator, by

$$a_p^\dagger |\dots, n_p, \dots\rangle = \sqrt{n_p + 1} |\dots, n_p + 1, \dots\rangle \quad (14.2)$$

Similarly we can define a destruction operator by

$$a_p |\dots, n_p, \dots\rangle = \sqrt{n_p} |\dots, n_p - 1, \dots\rangle \quad (14.3)$$

Let us consider first the case of bosons. Consider the product of two creation operators acting on the vacuum state (state with zero occupation numbers)

$$a_p^\dagger a_{p'}^\dagger |0, \dots, 0\rangle = |0, \dots, 1, \dots, 1, \dots, 0\rangle \quad (14.4)$$

indicating that a particle was created with momentum \vec{p} and another with momentum \vec{p}' . Since the wave function is symmetrical with respect to the permutation of particles the final state is the same if instead we act with the operators in reverse order like $a_{p'}^\dagger a_p^\dagger |0\rangle$, using a simplified notation to represent the vacuum state. Therefore we conclude that

$$a_p^\dagger a_{p'}^\dagger - a_{p'}^\dagger a_p^\dagger = 0 \quad (14.5)$$

meaning that the operators commute. In the same we may conclude that

$$[a_p, a_{p'}] = [a_p^\dagger, a_{p'}^\dagger] = 0$$

$$[a_p, a_{p'}^\dagger] = \delta_{p,p'}$$

The last result is easily shown:

$$\begin{aligned} a_p a_{p'}^\dagger |\dots, n_p, \dots, n_{p'}, \dots\rangle &= \sqrt{n_{p'} + 1} \sqrt{n_p} |\dots, n_p - 1, \dots, n_{p'} + 1, \dots\rangle \\ a_{p'}^\dagger a_p |\dots, n_p, \dots, n_{p'}, \dots\rangle &= \sqrt{n_p + 1} \sqrt{n_{p'}} |\dots, n_p - 1, \dots, n_{p'} + 1, \dots\rangle \end{aligned} \quad (14.6)$$

Therefore if $p \neq p'$ we get that

$$(a_p a_{p'}^\dagger - a_{p'}^\dagger a_p) |\dots, n_p, \dots, n_{p'}, \dots\rangle = 0 \quad (14.7)$$

If $p = p'$ then

$$\begin{aligned} a_p a_p^\dagger |\dots, n_p, \dots\rangle &= \sqrt{n_p + 1} \sqrt{n_p + 1} |\dots, n_p, \dots\rangle \\ a_p^\dagger a_p |\dots, n_p, \dots\rangle &= \sqrt{n_p} \sqrt{n_p} |\dots, n_p, \dots\rangle \end{aligned} \quad (14.8)$$

and

$$(a_p a_p^\dagger - a_p^\dagger a_p) |\dots, n_p, \dots\rangle = (n_p + 1 - n_p) |\dots, n_p, \dots\rangle \quad (14.9)$$

which proves the above result. The operator $n_p = a_p^\dagger a_p$ is a number operator since the state $|\dots, n_p, \dots\rangle$ is an eigenvector with eigenvalue n_p .

Let us now consider the case of fermions. In the same way we define creation and destruction operators. We have to keep in mind now that the wave function is antisymmetric and that the exchange of two particles involves a minus sign. Then

$$c_p^\dagger c_{p'}^\dagger |\dots\rangle = -c_{p'}^\dagger c_p^\dagger |\dots\rangle \quad (14.10)$$

This implies that these operators anticommute instead of commuting:

$$(c_p^\dagger c_{p'}^\dagger + c_{p'}^\dagger c_p^\dagger) = \{c_p^\dagger, c_{p'}^\dagger\} = 0 \quad (14.11)$$

Also we can show that

$$\begin{aligned} \{c_p, c_{p'}\} &= 0 \\ \{c_p^\dagger, c_{p'}\} &= \delta_{p,p'} \end{aligned} \quad (14.12)$$

In the same way $n_p = c_p^\dagger c_p$ is a number operator. As a consequence $c_p^2 = 0$ and $n_p^2 = n_p$. We have then to define that

$$\begin{aligned} c_p |\dots, n_p, \dots\rangle &= (-1)^{-\sum_{j<p} n_j} n_p |\dots, n_p - 1, \dots\rangle \\ c_p^\dagger |\dots, n_p, \dots\rangle &= (-1)^{-\sum_{j<p} n_j} (1 - n_p) |\dots, n_p + 1, \dots\rangle \end{aligned} \quad (14.13)$$

We have to assign a specified order to the one particle states. Any order is allowed since only relative signs have physical meaning. For instance we may choose to define

$$|\dots, n_p = 1, \dots, n_{p'} = 1, \dots\rangle = c_p^\dagger c_{p'}^\dagger |\dots, 0, \dots, 0, \dots\rangle \quad (14.14)$$

We have overlooked so far that the fermions have half-integer spin. Let us then consider spin-1/2 fermions and add the spin index to the operators. The anticommutation relations are then properly rewritten as

$$\begin{aligned} \{c_{p\sigma}, c_{p'\sigma'}\} &= \{c_{p\sigma}^\dagger, c_{p'\sigma'}^\dagger\} = 0 \\ \{c_{p\sigma}, c_{p'\sigma'}^\dagger\} &= \delta_{\sigma\sigma'} \delta_{pp'} \end{aligned} \quad (14.15)$$

14.1.1 Field operators

A normalized plane wave (eigenfunction of the momentum operator) is written as

$$\phi_p(\vec{r}) = \frac{1}{\sqrt{\Omega}} e^{i\vec{p}\cdot\vec{r}}$$

As a consequence $\phi_p(\vec{r})a_{\vec{p}}$ is the probability amplitude to destroy a boson (or a fermion if we replace the boson destruction operator by the fermion destruction operator) of momentum \vec{p} at the point \vec{r} . Let us then define a field operator $\psi(\vec{r})$ by the Fourier transform of the destruction operator

$$\begin{aligned}\psi(\vec{r}) &= \frac{1}{\sqrt{\Omega}} \sum_{\vec{p}} a_{\vec{p}} e^{i\vec{p} \cdot \vec{r}} \\ a_{\vec{p}} &= \frac{1}{\sqrt{\Omega}} \int d^3r \psi(\vec{r}) e^{-i\vec{p} \cdot \vec{r}}\end{aligned}\quad (14.16)$$

This field operator destroys a particle at site \vec{r} . Clearly the conjugate is an operator that creates a particle at point \vec{r} defined by

$$\psi^\dagger(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{p}} a_{\vec{p}}^\dagger e^{-i\vec{p} \cdot \vec{r}} \quad (14.17)$$

It is easy to show using the commutation and anticommutation relations of the momentum operators that for bosons the field operators obey

$$\begin{aligned}[\psi(\vec{r}), \psi(\vec{r}')] &= [\psi^\dagger(\vec{r}), \psi^\dagger(\vec{r}')] = 0 \\ [\psi(\vec{r}), \psi^\dagger(\vec{r}')] &= \delta(\vec{r} - \vec{r}')$$

and for fermions obey

$$\begin{aligned}\{\psi_\sigma(\vec{r}), \psi_{\sigma'}(\vec{r}')\} &= \{\psi_\sigma^\dagger(\vec{r}), \psi_{\sigma'}^\dagger(\vec{r}')\} = 0 \\ \{\psi_\sigma(\vec{r}), \psi_{\sigma'}^\dagger(\vec{r}')\} &= \delta_{\sigma\sigma'} \delta(\vec{r} - \vec{r}')$$

Let us now consider the fermion gas as an example. The number operators in momentum and in real spaces are defined by

$$\begin{aligned}n_{p\sigma} &= c_{p\sigma}^\dagger c_{p\sigma} \\ \rho(\vec{r}, \sigma) &= \psi_\sigma^\dagger(\vec{r}) \psi_\sigma(\vec{r})\end{aligned}\quad (14.18)$$

respectively. Clearly the total number of particles operator is given by

$$N = \sum_{\vec{p}, \sigma} c_{p\sigma}^\dagger c_{p\sigma} = \sum_{\sigma} \int d^3r \psi_\sigma^\dagger(\vec{r}) \psi_\sigma(\vec{r}) \quad (14.19)$$

Let us now see how the various quantum mechanical operators are written in second quantization. In particular we are interested in representing the Hamiltonian in terms of destruction and creation operators. The kinetic energy is given by

$$\begin{aligned} T &= \sum_i \frac{\vec{p}_i^2}{2m} \\ &= \sum_{\vec{p}, \sigma} \frac{\vec{p}^2}{2m} n_{p\sigma} \end{aligned} \quad (14.20)$$

which is straightforwardly written in second quantization replacing the number of particles in a given momentum and with a given spin projection by the corresponding number operator. Therefore in second quantization the kinetic energy is written as

$$T = \sum_{\vec{p}, \sigma} \frac{\vec{p}^2}{2m} c_{p\sigma}^\dagger c_{p\sigma} \quad (14.21)$$

We can as well recall that the kinetic energy in first quantization is written as

$$T = -\frac{\hbar^2}{2m} \sum_{\sigma} \int d^3r \psi_{\sigma}^*(\vec{r}) \nabla^2 \psi_{\sigma}(\vec{r}) \quad (14.22)$$

where now $\psi_{\sigma}(\vec{r})$ is the wave function and interpret this wave function as a field operator. Then if we Fourier transform it is easy to see that we recover the above expression.

The potential energy is more complicated to handle. Let us consider a two-body factorizable potential energy like

$$V = \frac{1}{2} \sum_{i \neq j} V(\vec{r}_i - \vec{r}_j) \quad (14.23)$$

The particle density is given by

$$\rho(\vec{r}) = \sum_i \delta(\vec{r} - \vec{r}_i) \quad (14.24)$$

Then the potential energy is written as

$$V = \frac{1}{2} \int d^3r \int d^3r' V(r - r') [\rho(r)\rho(r') - \rho(r)\delta(r - r')] \quad (14.25)$$

The last term eliminates the interaction of a particle with itself. Recalling that $\rho(r) = \sum_{\sigma} \rho(r, \sigma)$ and interpreting the density functions as operators we can write the potential energy as

$$\begin{aligned} V &= \frac{1}{2} \int d^3r \int d^3r' V(r - r') \left[\sum_{\sigma, \sigma'} \psi_{\sigma}^{\dagger}(r) \psi_{\sigma}(r) \psi_{\sigma'}^{\dagger}(r') \psi_{\sigma'}(r') \right. \\ &\quad \left. - \sum_{\sigma} \psi_{\sigma}^{\dagger}(r) \psi_{\sigma}(r) \delta(r - r') \right] \end{aligned} \quad (14.26)$$

Using that

$$\begin{aligned} [A, BC] &= [A, B]C - B[C, A] \\ &= \{A, B\}C - B\{C, A\} \end{aligned} \quad (14.27)$$

where A, B, C are operators we can rewrite the potential energy in so-called normal order where all the creation operators are to the left and all the destruction operators are to the right

$$V = \frac{1}{2} \int d^3r \int d^3r' V(r - r') \sum_{\sigma, \sigma'} \psi_{\sigma}^{\dagger}(r) \psi_{\sigma'}^{\dagger}(r') \psi_{\sigma'}(r') \psi_{\sigma}(r) \quad (14.28)$$

In momentum space it can be written as

$$V = \frac{1}{2\Omega} \sum_{\sigma, \sigma'} \sum_{k, p, q} V_q c_{p+q, \sigma}^{\dagger} c_{p-q, \sigma'}^{\dagger} c_{k, \sigma'} c_{p, \sigma} \quad (14.29)$$

Here V_q is the Fourier transform of the two-body potential energy.

Due to the potential there will be scattering of the particles in two-body collisions. One particle with momentum k and spin projection σ' will scatter off a particle with momentum p and spin projection σ . After the collision a particle with momentum $k - q$ and spin projection σ' will emerge and another one with momentum $q + p$ and spin projection σ . Since the particles are indistinguishable any of the two particles will be in either final state leading to two possible outcomes of the collision. The scattering matrix element of a pair of fermions $(p, \sigma; k, \sigma')$ into a new state $(q + p, \sigma; k - q, \sigma')$ is given by

$$\langle \dots; q + p, \sigma; \dots; k - q, \sigma'; \dots | V | \dots; p, \sigma; \dots; k, \sigma'; \dots \rangle \quad (14.30)$$

Replacing the expression for the potential energy we have to evaluate the matrix element

$$\langle p+q, \sigma; k-q, \sigma' | c_{p'+q', \sigma_1}^\dagger c_{k'-q', \sigma_2}^\dagger c_{k', \sigma_2} c_{p', \sigma_1} | p, \sigma; k, \sigma' \rangle$$

and then sum over the spin projections σ_1, σ_2 and over the momenta k', p', q' and multiply by $V_{q'}$. There are two possibilities: in one we act with the destruction operator c_{k', σ_2} in the single particle state $|p, \sigma\rangle$ and the destruction operator c_{p', σ_1} in the single particle state $|p, \sigma\rangle$ or vice-versa (recall that the order of the operators is important when we are dealing with fermions). The matrix element yields for fermions $V_q - V_{p+q-k} \delta_{\sigma, \sigma'}$ and for bosons it yields $V_q + V_{p+q-k}$. The different relative sign between the boson and fermion cases is very important. The first term V_q is called the direct term and the other term is the exchange term. Note that the exchange term in the fermion case only appears if the spins are equal. We will get back to this later.

To finish this section it is easy to see that

$$\begin{aligned} \rho_q &= \int d^3r \rho(r) e^{-iq \cdot r} \\ \rho(q) &= \sum_{p\sigma} c_{p\sigma}^\dagger c_{p+q, \sigma} \end{aligned} \quad (14.31)$$

and that the current operator is given in second quantization by

$$\vec{J}_q = \sum_{p\sigma} \frac{\vec{p}}{m} c_{p\sigma}^\dagger c_{p+q, \sigma} \quad (14.32)$$

14.2 Bose liquid

Let us consider a Bose liquid like for instance He^4 . We are interested in the low energy excitations which we expect are those corresponding to low momenta or equivalently long distances. These are therefore the sound waves of the system. The energy dispersion relation is then of the form $\omega = uk$ or $\epsilon = up$ where u is the sound velocity. This only holds at low energies. As the momentum increases then *epsilon* will no longer be linear in p . In general we have that

$$E(\rho, v) = \frac{1}{2} \int d^3r \rho(r) v^2(r) + E^{(1)}(\rho) \quad (14.33)$$

where the second term is independent of velocity. Let us then expand the density around its mean value

$$\rho(\vec{r}) = \bar{\rho} \delta \rho(\vec{r})$$

Let us consider small fluctuations and small velocities. The average density is given by

$$\bar{\rho} = \frac{1}{V} \int d^3r \rho(\vec{r})$$

and

$$\int d^3r \delta \rho(\vec{r}) = 0$$

We can now expand

$$\begin{aligned} E^{(1)}(\rho) &= E^{(1)}(\bar{\rho}) + \int d^3r \psi(\vec{r}) \delta \rho(\vec{r}) \\ &+ \frac{1}{2} \int d^3r \int d^3r' \varphi(\vec{r}, \vec{r}') \delta \rho(\vec{r}) \delta \rho(\vec{r}') \end{aligned}$$

$\psi(\vec{r})$ and $\varphi(\vec{r}, \vec{r}')$ are determined by the equilibrium properties of the liquid (with no oscillations in the density). In this case they are homogeneous and isotropic and $\psi(\vec{r}) = \psi$ and $\varphi(\vec{r}, \vec{r}') = \varphi(|\vec{r} - \vec{r}'|)$. This implies that the total energy of the liquid is expanded

$$\begin{aligned} E &= E^{(1)}(\bar{\rho}) + \frac{1}{2} \bar{\rho} \int d^3r \vec{v}^2 \\ &+ \frac{1}{2} \int d^3r \int d^3r' \varphi(|\vec{r} - \vec{r}'|) \delta \rho(\vec{r}) \delta \rho(\vec{r}') \end{aligned}$$

The continuity equation is given by

$$\dot{\rho} + \vec{\nabla} \cdot (\rho \vec{v}) = 0$$

which reduces to leading order to

$$\dot{\rho} + \bar{\rho} \vec{\nabla} \cdot \vec{v} = 0$$

let us expand in a Fourier series the density oscillations and the velocity

$$\begin{aligned}\delta\rho(\vec{r}) &= \frac{1}{V} \sum_{\vec{p}} \rho_{\vec{p}} e^{i\vec{p}\cdot\vec{r}} \\ \vec{v}(\vec{r}) &= \frac{1}{V} \sum_{\vec{p}} \vec{v}_{\vec{p}} e^{i\vec{p}\cdot\vec{r}} \\ \varphi(\vec{r}) &= \frac{1}{V} \sum_{\vec{p}} \varphi_{\vec{p}} e^{i\vec{p}\cdot\vec{r}}\end{aligned}$$

The low energy oscillations of the liquid are longitudinal that is $\vec{v}_{\vec{p}} = a_{\vec{p}}\vec{p}$. The continuity equation then leads to

$$\vec{v}_{\vec{p}} = i \frac{\dot{\rho}_{\vec{p}}}{\bar{\rho}} \frac{\vec{p}}{p^2}$$

Then the energy is written as

$$E = E^{(1)}(\bar{\rho}) + \frac{1}{V} \sum_{\vec{p}} \left(\frac{|\dot{\rho}_{\vec{p}}|^2}{2\bar{\rho}p^2} + \frac{1}{2} \varphi_{\vec{p}} |\rho_{\vec{p}}|^2 \right) \quad (14.34)$$

The term corresponding to the fluctuations describes a set of harmonic oscillators of frequencies ω_p given by

$$\omega_p^2 = \bar{\rho} p^2 \varphi_{\vec{p}}$$

This implies that

$$\epsilon_{\vec{p}} = \omega_{\vec{p}} \left(n + \frac{1}{2} \right)$$

The ground state is therefore

$$E_0 = E^{(1)}(\bar{\rho}) + \sum_{\vec{p}} \frac{1}{2} \omega_{\vec{p}}$$

Using the continuity equation we can write that

$$\frac{1}{2} V \omega_p = \frac{1}{2\bar{\rho}p^2} |\dot{\rho}_{\vec{p}}|^2 + \frac{1}{2} \varphi_p |\varphi_p^-|^2 = \varphi_p |\rho_p|^2$$

this implies that

$$\epsilon(p) = \omega_p = \frac{p^2}{2mS(\vec{p})} \quad (14.35)$$

where

$$S(\vec{p}) = \frac{|\rho_p|^2}{Vm\bar{\rho}} \quad (14.36)$$

This is the Fourier transform of $S(\vec{r} - \vec{r}')$ the density correlation function given by

$$S(\vec{r} - \vec{r}') = \frac{[n(\vec{r}) - \bar{n}][n(\vec{r}') - \bar{n}]}{\bar{n}}$$

Here $n(\vec{r}) = \rho(\vec{r})/m$ is the number of particles per unit volume. At small momenta $\epsilon \sim up$ which implies that

$$S \sim \frac{p}{2mu}$$

At large momenta (small distances) we should recover the free particle spectrum where $\epsilon \sim p^2$.

14.3 Interacting bosons

Let us then consider a system of bosons with weak interactions. The system is described by an Hamiltonian of the form

$$\begin{aligned} H &= H_0 + H_{int} \\ H_0 &= \sum_{\vec{k}} \epsilon_{\vec{k}} a_{\vec{k}}^\dagger a_{\vec{k}} \\ H_{int} &= \frac{1}{2} \frac{1}{V^2} \int d^3r_1 \int d^3r_2 V(\vec{r}_1 - \vec{r}_2) a^\dagger(\vec{r}_1) a(\vec{r}_1) a^\dagger(\vec{r}_2) a(\vec{r}_2) \end{aligned} \quad (14.37)$$

Here $\epsilon_{\vec{k}} = \frac{k^2}{2M}$. This model may apply to He^4 a system of weakly coupled bosons where the number of bosons is conserved.

Let us define

$$a(\vec{r}) = \sum_{\vec{k}} e^{i\vec{k} \cdot \vec{r}} a_{\vec{k}}$$

Then we can rewrite the interaction term introducing the Fourier transform of the interaction potential as

$$H_{int} = \frac{1}{2} \sum_{k_1, k'_1, k_2, k'_2} \delta_{k_1+k_2, k'_1+k'_2} V(k'_1 - k_1) a_{k_1}^\dagger a_{k_2}^\dagger a_{k'_2} a_{k'_1} + \frac{1}{2} NV(r=0) \quad (14.38)$$

Here $N = \sum_{\vec{k}} a_k^\dagger a_k$.

If $V = 0$ we have seen that the system will show Bose-Einstein condensation at $T < T_c$ and all the particles will be in the state of $k = 0$ at $T = 0$. If $V \neq 0$ then some particles will occupy excited states ($\vec{k} \neq 0$). There will be N_e in the excited states and N_0 in the ground state with $N_e \ll N_0$.

Let us now look at the structure of the terms of the Hamiltonian. They represented particles in different states and involved in different processes like:

i) Interactions in the ground state: $a_0^\dagger a_0^\dagger a_0 a_0$. This leads to a contribution $\frac{1}{2} V(k=0) N_0^2$.

ii) One particle out of the ground state: $a_0^\dagger a_k^\dagger a_k a_0$ and $a_k^\dagger a_0^\dagger a_0 a_k$. Lead to $N_0 V(k=0) \sum_k' a_k^\dagger a_k$. The $'$ excludes the momentum $k = 0$.

iii) Exchange of one particle in the ground state with another one in a different state: $a_k^\dagger a_0^\dagger a_k a_0$ and $a_0^\dagger a_k^\dagger a_0 a_k$. Lead to $N_0 \sum_k' V(k) a_k^\dagger a_k$.

iv) Both particles in the ground state or both in an excited state: $a_0^\dagger a_0^\dagger a_k a_{-k}$ or $a_k^\dagger a_{-k}^\dagger a_0 a_0$. Lead to $\frac{1}{2} N_0 \sum_k' V(k) (a_k a_{-k} + a_k^\dagger a_{-k}^\dagger)$.

We get then that the Hamiltonian is written as

$$\begin{aligned} H = & \sum_k \epsilon_k a_k^\dagger a_k + \frac{1}{2} N_0^2 V_0 + N_0 V_0 \sum_k' a_k^\dagger a_k \\ & + N_0 \sum_k' V(k) a_k^\dagger a_k + \frac{1}{2} \sum_k' V(k) (a_k a_{-k} + a_k^\dagger a_{-k}^\dagger) + \dots \end{aligned} \quad (14.39)$$

Using that $N = N_0 + \sum_k' a_k^\dagger a_k$ we can rewrite the Hamiltonian as

$$H = \frac{1}{2} N_0^2 V_0 + \sum_k' (\epsilon_k + N V_k) a_k^\dagger a_k + \frac{1}{2} \sum_k' V(k) (a_k a_{-k} + a_k^\dagger a_{-k}^\dagger) + \dots \quad (14.40)$$

14.3.1 Diagonalization of interacting Hamiltonian

Let us diagonalize the system using two possible methods.

a) Equations of motion

The equations of motion of the operators

$$\begin{aligned} i\frac{\partial}{\partial t}a_k^\dagger &= [a_k^\dagger, H] = -(\epsilon_k + NV_k)a_k^\dagger - NV_ka_{-k} \\ i\frac{\partial}{\partial t}a_{-k} &= [a_{-k}, H] = (\epsilon_k + NV_k)a_{-k} + NV_ka_k^\dagger \end{aligned} \quad (14.41)$$

Taking the Fourier transform we get that

$$\begin{aligned} [-\omega + \epsilon_k + NV_k]a_k^\dagger + NV_ka_{-k} &= 0 \\ -NV_ka_k^\dagger + [-\omega - \epsilon_k - NV_k]a_{-k} &= 0 \end{aligned} \quad (14.42)$$

Taking the determinant to be zero we get that

$$\omega^2 - (\epsilon_k + NV_k)^2 + (NV_k)^2 = 0 \quad (14.43)$$

The solution is therefore

$$\omega = \sqrt{\epsilon_k^2 + 2NV_k\epsilon_k} \quad (14.44)$$

In the limit when $k \rightarrow 0$ we get that

$$\omega_{k \rightarrow 0} = \sqrt{\frac{NV_0}{M}}k$$

In the limit when $k \gg 1$ we get that

$$\omega_{k \gg 1} = \epsilon_k \left(1 + \frac{NV_k}{\epsilon_k} \right) \sim \epsilon_k + NV_k$$

Therefore when $V = 0$ the dispersion at small momenta is quadratic in k . However when there is an interaction between the bosons $V \neq 0$ the dispersion is linear in the momentum (note that $V_0 > 0$ meaning a repulsive potential). Any infinitesimal potential changes qualitatively the nature of the spectrum.

b) Bogoliubov-Valatin transformation

The equations of motion couple a_k^\dagger and a_{-k} . Let us then define new operators α_k^\dagger and α_k such that

$$H = \sum_k \lambda_k \alpha_k^\dagger \alpha_k + \text{const} \quad (14.45)$$

With this choice we see that the

$$[\alpha_k^\dagger, H] = -\lambda_k \alpha_k^\dagger$$

$$[\alpha_k, H] = \lambda_k \alpha_k$$

and since the new operators are also bosons they obey that

$$[\alpha_k, \alpha_{k'}^\dagger] = \delta_{k,k'}$$

The Hamiltonian may be written in a convenient way like

$$\begin{aligned} H &= \frac{1}{2} N^2 V_0 + \sum_k' H_k \\ H_k &= \omega_0 (a_k^\dagger a_k + a_{-k}^\dagger a_{-k}) \\ \omega_0 &= \epsilon_k + N V_k \\ \omega_1 &= N V_k \end{aligned} \quad (14.46)$$

Let us then consider the linear combinations

$$\begin{aligned} \alpha_k &= u_k a_k - v_k a_{-k}^\dagger \\ \alpha_k^\dagger &= u_k a_k^\dagger - v_k a_{-k} \end{aligned} \quad (14.47)$$

The commutation relations between the new operators lead to

$$[\alpha_k, \alpha_k^\dagger] = u_k^2 - v_k^2 = 1$$

We can invert the linear combination and obtain that

$$a_k = u_k \alpha_k + v_k \alpha_{-k}^\dagger$$

Calculating the following commutator

$$\begin{aligned} [\alpha_k^\dagger, H_k] &= u_k (-\omega_0 \alpha_k^\dagger - \omega_1 a_{-k}) - v_k (\omega_0 a_{-k} + \omega_1 \alpha_k^\dagger) \\ &= -\lambda_k (u_k \alpha_k^\dagger - v_k a_{-k}) \end{aligned} \quad (14.48)$$

This implies that

$$\begin{aligned}\omega_0 u_k + \omega_1 v_k &= \lambda_k u_k \\ \omega_1 u_k + \omega_0 v_k &= -\lambda_k v_k\end{aligned}\tag{14.49}$$

These equations have a solution if the determinant of the matrix

$$\begin{pmatrix} \omega_0 - \lambda_k & \omega_1 \\ \omega_1 & \omega_0 + \lambda_k \end{pmatrix}$$

The eigenvalues are given by

$$\lambda_k^2 = (\epsilon_k + NV_k)^2 - (NV_k)^2$$

Therefore in the limit $k \rightarrow 0$ we get as before

$$\lambda_{k \rightarrow 0} = \left(\frac{NV_0}{M} \right)^{1/2} k$$

14.3.2 Excitations. Second sound

Let us consider to be specific that the interaction potential between the bosons is of the Lennard-Jones type. The potential and its Fourier transform are shown in the Figures. Using this potential the spectrum of the interacting Bose liquid is shown in the Figure. At low momenta the spectrum is linear like that of phonons in a solid. As the momentum grows there is a local minimum in the dispersion relation which is related to the minimum in the Lennard-Jones potential. At high momenta the energy dispersion converges to that of free bosons, as previously discussed.

In the neighborhood of the local minimum located around $1A^{-1} - 2A^{-1}$ a new type of quasi-particles appears which were called rotons by Landau. The average distance between the atoms in He^4 in the liquid is about $r \sim (2\pi)/(2A^{-1}) \sim 3.1A$. The structure factor of the liquid shows a high probability of finding two atoms at a distance of about $3A$. There is therefore short range order in the liquid.

Let us now consider a particle of mass M and velocity \vec{v} that is moving through the Bose liquid. As it moves through the liquid it will interact with the elementary excitations of the liquid. There is conservation of energy and momentum expressed

by

$$\begin{aligned}\frac{1}{2}Mv^2 &= \frac{1}{2}M(v')^2 + \omega_k \\ M\vec{v} &= M\vec{v}' + \vec{k}\end{aligned}\tag{14.50}$$

where \vec{k} and ω_k are the momentum and energy of the elementary excitation of the Bose liquid, respectively. We can then get that

$$\begin{aligned}(M\vec{v} - \vec{k})^2 &= M^2(v')^2 = M^2v^2 - 2M\vec{v} \cdot \vec{k} + k^2 \\ 0 &= \omega_k - \vec{v} \cdot \vec{k} + \frac{k^2}{2M}\end{aligned}\tag{14.51}$$

The smallest velocity that satisfies this condition is

$$v_c = \min \left(\frac{\omega_k}{k} + \frac{k}{2M} \right) > \min \frac{\omega_k}{k}\tag{14.52}$$

If the potential $V = 0$ the critical velocity vanishes (since the dispersion is quadratic). However if $V \neq 0$ then the ratio of ω_k/k is finite. This implies that the critical velocity is non-zero. If the velocity of the incident particle is $v < v_c$ there is no loss of energy to the medium and the particle moves without losing energy. The system becomes a superfluid. Note that if the system of Bosons is a gas (non-interacting) it can not be superfluid. Note however that there are other excitations which we have not considered which have a different nature from the elementary excitations we have paid attention up to now, called vortices, that decrease the value of the critical velocity in about an order of magnitude. The vortices require a lower threshold to be excited but still below the renormalized critical velocity there will be superfluidity.

The coefficient of the dispersion relation with respect to the momentum is the usual sound (called first sound) such that $\omega = c_1 k$ where c_1 is the velocity of sound. There are however other excitations which are constituted of waves of energy density that appear that are of the form $\omega = c_2 k$ which are called second sound. Let us determine the velocity of propagation of these waves.

Let us assume there is no Umklapp and that the mean free path of the phonons is smaller than the wave length of the second sound. Let us define

$$f(k, x) d^3 k d^3 x$$

the number of phonons in an elementary volume in phase space. Then the rate of change of the density of phonons is given by the Boltzmann like equation

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \vec{v} \cdot \vec{\nabla} f = \Delta_c f$$

Which is the rate of change of f due to the collisions. Now

$$\vec{v} = c_1 \frac{\vec{k}}{|\vec{k}|}$$

Using this we get that

$$\frac{\partial f}{\partial t} + c_1 \frac{\vec{k}}{|\vec{k}|} \cdot \vec{\nabla} f = \Delta_c f$$

The total momentum and the total energy of the liquid are conserved. They are given by

$$\begin{aligned} \vec{P}(\vec{r}) &= \int d^3k \vec{k} f \\ U(\vec{r}) &= \int d^3k c_1 k f \end{aligned} \quad (14.53)$$

Since these quantities are conserved

$$\begin{aligned} \frac{d\vec{P}}{dt} &= 0 = \int d^3k \vec{k} \Delta_c f \\ \frac{dU}{dt} &= 0 = \int d^3k c_1 k \Delta_c f \end{aligned} \quad (14.54)$$

Multiplying by k_μ and integrating over the momentum we get that the equation of motion for f gives

$$\frac{\partial}{\partial t} \int d^3k k_\mu f + \int d^3k \frac{c_1}{k} k_\mu \sum_\nu k_\nu \frac{\partial f}{\partial x_\nu} = \int d^3k k_\mu \Delta_c f = 0$$

and therefore

$$\frac{\partial}{\partial t} P_\mu + \sum_\nu \frac{\partial}{\partial x_\nu} c_1 \int d^3k \frac{k_\mu k_\nu}{k} f = 0 \quad (14.55)$$

Multiplying now by $c_1 k$ and integrating over the momentum we get that

$$\frac{\partial}{\partial t} \int d^3 k c_1 k f + \int d^3 k c_1^2 \sum_{\nu} k_{\nu} \frac{\partial f}{\partial x_{\nu}} = \int d^3 k c_1 \Delta_c f = 0$$

and therefore

$$\frac{\partial U}{\partial t} + c_1^2 \sum_{\nu} \frac{\partial}{\partial x_{\nu}} P_{\nu} = 0 \quad (14.56)$$

If we now consider a situation close to equilibrium $f \sim f^0$ we can simplify

$$c_1 \int d^3 k \frac{k_{\mu} k_{\nu}}{k} f \sim \delta_{\mu\nu} c_1 \int d^3 k \frac{k_{\mu}^2}{k} f$$

as the dominant contribution. This is approximately given by

$$\frac{1}{3} \delta_{\mu\nu} c_1 \int d^3 k \frac{k^2}{k} f \sim \frac{1}{3} \delta_{\mu\nu} U$$

Then we get that

$$\frac{\partial}{\partial t} P_{\mu} + \frac{1}{3} \frac{\partial U}{\partial x_{\mu}} = 0 \quad (14.57)$$

Taking now the time derivative in the equation () we get that

$$\frac{\partial^2 U}{\partial t^2} - \frac{c_1^2}{3} \sum_{\nu} \frac{\partial^2 U}{\partial x_{\nu}^2} = 0 \quad (14.58)$$

We obtain therefore an equation for the energy that is a wave equation with a velocity of propagation $c_2 = c_1/\sqrt{3}$. These waves which are energy density waves are called second sound.

14.4 Solid state physics and the fermi liquid

The electrons of the atoms in a solid have orbitals that overlap. The electrons are therefore shared between the ions some of them becoming strongly delocalized to become conduction electrons. The microscopic Hamiltonian of the system is quite complex due to the strong electromagnetic interactions between the charges of the system. Clearly some simplification has to be done to obtain the energy states.

The Hamiltonian is in general of the form

$$\begin{aligned}
H &= H_{el} + H_{ion} + H_{el-ion} \\
H_{el} &= H_{el,kin} + H_{el-el} = \sum_k \frac{p_k^2}{2m} + \frac{1}{8\pi\epsilon_0} \sum'_{k,k'} \frac{e^2}{|\vec{r}_k - \vec{r}_{k'}|} \\
H_{ion} &= H_{ion,kin} + H_{ion-ion} = \sum_i \frac{p_i^2}{2M_i} + \frac{1}{2} \sum'_{i,i'} V_{ion}(\vec{R}_i - \vec{R}_{i'}) \\
H_{el-ion} &= \sum_{k,i} V_{el-ion}(\vec{r}_k - \vec{R}_i) \\
H_{ion,ion} &= H_{ion-ion}^0 + H_{ph} \\
H_{el,ion} &= H_{el-ion}^0 + H_{el-ph}
\end{aligned} \tag{14.59}$$

Here H^0 respects to the equilibrium positions of the ions and the label ph is the part due to the phonons. We can make approximations. Consider first the case of the electrons. As a first step we can write that

$$H_{el} = - \sum_k \frac{\hbar^2}{2m} \nabla_k^2 + \frac{1}{8\pi\epsilon_0} \sum'_{k,k'} \frac{e^2}{|\vec{r}_k - \vec{r}_{k'}|} + H_+ \tag{14.60}$$

This Hamiltonian respects only to the electrons. The positive charges of the ions are replaced by a positively charged background (required to maintain charge neutrality and stabilize the system). This is the so-called jellium model. It can be improved if instead of considering a uniform background we implement the symmetries of the lattice of ions. This amounts, as reviewed before to replace the free electron spectrum by the band structure due to the periodicity of the lattice. Clearly the important new step is now the interactions between the electrons in comparison with the nearly free electron model previously discussed.

We can as well make some approximations for the Hamiltonian of the ions. Take as Hamiltonian

$$H_{ion} = - \sum_i \frac{\hbar^2}{2M_i} \nabla_i^2 + \frac{1}{2} \sum'_{i,i'} V_{ion}(\vec{R}_i - \vec{R}_{i'}) + H_- \tag{14.61}$$

In the same way we replace the interaction of the ions with the electrons by an effective negative medium. The solution of this Hamiltonian gives the dynamics of the lattice. For consistency $H_+ + H_- = 0$.

We have effectively separated the dynamics of the electrons and of the ions in a first approximation. This constitutes the basis of the adiabatic approximation of Born-Oppenheimer. The electrons have masses which are far smaller than those of the ions. Therefore since the strength of the forces is similar the electrons move much faster than the ions. The ions respond slowly to any change of configuration of the electrons but the electrons follow any changes of the configuration of the ions. The ions respond to an average of the configurations of the electrons and the wave functions of the electrons can be parametrized by the distribution of the ions. This implies that we can write that

$$(H_{el} + H_{el-ion})\psi = E_{el}\psi \quad (14.62)$$

The overall wave function Ψ is such that

$$\Psi = \psi(\vec{r}_1, \dots, \vec{r}_N; \vec{R}_1, \dots, \vec{R}_N) \varphi(\vec{R}_1, \dots, \vec{R}_N) \quad (14.63)$$

In the part of the wave function respecting to the electrons the coordinates of the ions are just parameters. We can write that

$$\begin{aligned} H\Psi &= (H_{el} + H_{ion} + H_{el-ion})\psi\varphi \\ &= \psi(H_{ion} + E_{el})\varphi - \sum_i \frac{\hbar^2}{2M_i} (\varphi \nabla_i^2 \psi + 2\nabla_i \varphi \cdot \nabla_i \psi) \end{aligned} \quad (14.64)$$

If we neglect the second term we get that

$$(H_{ion} + E_{el})\varphi = E\varphi \quad (14.65)$$

where E_{el} is obtained from the solution for the electrons.

Let us then focus our attention in the electron problem. This is given by the solution of

$$(H_{el} + H_{el-ion}^0)\psi = E_{el}\psi \quad (14.66)$$

where we have neglected the interactions of the electrons with the phonons. Usually in solid state physics the interactions between the conduction electrons are neglected and only the free electron problem in the presence of a periodic potential is solved. However the interactions between the electrons may be important. However, the restrictions on the phase space due to the exclusion principle have to be taken into

account. Let us then consider that the interactions between the electrons are weak enough that they do not destroy the concept of the Fermi surface even there may be (and surely are) changes of the energy levels. Also there may be some changes to the equilibrium Fermi-Dirac distribution due to the interactions. This is the approximation taken by Landau, that there remains the concept of a Fermi surface, that will have to be justified later on. The Fermi liquid is then assumed to have overall characteristics that are similar to those of the Fermi gas, with excitations that are fermions even though with different energy levels.

Let us then assume that the excitations of the Fermi liquid are of the same type as those of the Fermi gas. That is the concept of the Fermi surface prevails. Actually we will see later that, even at $T = 0$ the discontinuity of the occupation number at the Fermi surface, even though it remains by hypothesis, will have a different (smaller) amplitude. The discontinuity is smaller because the interactions between the electrons renormalize parameters in the theory. Therefore even at $T = 0$ there is some depletion of the levels close to the Fermi surface.

Consider the case of the Fermi gas at finite temperature. Excitations arise that transfer particles from the levels inside the Fermi surface to states above the Fermi surface. This implies that the excited states (due to the thermal excitation) differ from the equilibrium distribution due to the excitation of particles. In the case of the Fermi liquid there also changes in the Fermi-Dirac distribution that represents the excitations. However due to the interactive character of the particles (and of the excitations themselves) the energy of the excitations involves interactive terms in the density of the particles. Due to the interactions the excitations have now a finite lifetime since they may scatter off other excitations. Therefore they are actually quasi-particles. Only very close to the Fermi surface the lifetime of the excitations tends to infinite. In a Fermi liquid the excitations close to the Fermi surface are well defined and have a character similar to the excitations of the Fermi gas. The theory is therefore limited to energies close to the Fermi surface.

14.5 Hartree-Fock approximation

Let us consider again the electron Hamiltonian in the jellium model

$$\begin{aligned}
 H &= -\sum_k \frac{\hbar^2}{2m} \nabla_k^2 + \frac{1}{8\pi\epsilon_0} \sum'_{k,k'} \frac{e^2}{|\vec{r}_k - \vec{r}_{k'}|} + H_+ \\
 H &= \sum_k H_k + \sum'_{k,k'} H_{kk'}
 \end{aligned} \tag{14.67}$$

We want now to focus our attention on the term of the interactions and not on the details of the energy bands. Eventually we may later on replace the plane waves by the energy levels in the presence of a periodic potential. For the moment we will use the plane waves as a starting point to describe free electrons that interact among themselves using perturbation theory.

If there were no interactions between the electrons the problem would be separable in the problem of one electron (independent electron approximation). Let us now attempt to obtain the equation of motion of one electron taking into account the interactions with the other electrons (perhaps in mean field?). This approximation in leading order is called the Hartree-Fock approximation.

If $H_{kk'} = 0$ the solution is trivial

$$\sum_k H_k \phi = E \phi$$

The solution is of the type

$$\phi(\vec{r}_1, \dots, \vec{r}_N) = \varphi_1(\vec{r}_1) \varphi_2(\vec{r}_2) \cdots \varphi_N(\vec{r}_N)$$

which then we have to antisymmetrize. The energy of the independent electrons is just additive $E = \sum_k E_k$, where E_k are the eigenvalues of the single particle states

$$H_k \varphi(\vec{r}_k) = E_k \varphi(\vec{r}_k)$$

Let us insert the solution of the problem with no interactions in the Schroedinger equation. Let us calculate

$$\langle \phi | H | \phi \rangle = \sum_k \langle \varphi_k | H_k | \varphi_k \rangle + \frac{e^2}{8\pi\epsilon_0} \sum'_{k,k'} \langle \varphi_k \varphi_{k'} | \frac{1}{|\vec{r}_k - \vec{r}_{k'}|} | \varphi_k \varphi_{k'} \rangle \tag{14.68}$$

Let us now use the variational principle and consider the states φ_k as variational functions. The states have a well defined norm. Let us add Lagrange multipliers, E_k , and let us call $E = \langle \phi | H | \phi \rangle$. Then

$$\delta[E - \sum_k E_k (\langle \varphi_k | \varphi_k \rangle - 1)] = 0$$

We obtain that

$$\langle \delta \varphi_j | H_j | \varphi_j \rangle + \frac{e^2}{4\pi\epsilon_0} \sum_{k \neq j} \langle \delta \varphi_j \varphi_k | \frac{1}{|\vec{r}_j - \vec{r}_k|} | \varphi_j \varphi_k \rangle - E_j \langle \delta \varphi_j | \varphi_j \rangle = 0$$

Since the variations are independent this leads to the Hartree equation

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V(r) + \frac{e^2}{4\pi\epsilon_0} \sum_{k \neq j} \int dV' \frac{|\varphi_k(\vec{r}')|^2}{|\vec{r} - \vec{r}'|} \right) \varphi_j(\vec{r}) = E_j \varphi_j(\vec{r}) \quad (14.69)$$

This is not quite correct because we have not antisymmetrized the single particle wavefunctions. We should have used the Slater determinant and we should have considered that each particle has a spin projection σ . The wave function is actually $q = (\vec{r}; \sigma)$

$$\phi = \frac{1}{\sqrt{N!}} \begin{pmatrix} \varphi_1(q_1) & \cdots & \varphi_N(q_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(q_N) & \cdots & \varphi_N(q_N) \end{pmatrix}$$

Let us now recalculate the average value of the energy using the Slater determinant.

$$\begin{aligned} E &= \langle \phi | H | \phi \rangle \\ E &= \sum_k \int dV_1 \varphi_k^*(q_1) H_k \varphi_k(q_1) + \frac{e^2}{8\pi\epsilon_0} \sum_{k,k'} \int dV_1 \int dV_2 \frac{|\varphi_k(q_1)|^2 |\varphi_{k'}(q_2)|^2}{|\vec{r}_1 - \vec{r}_2|} \\ &\quad - \frac{e^2}{8\pi\epsilon_0} \sum_{k,k'} \int dV_1 \int dV_2 \frac{\varphi_k^*(q_1) \varphi_k(q_2) \varphi_{k'}^*(q_2) \varphi_{k'}(q_1)}{|\vec{r}_1 - \vec{r}_2|} \end{aligned} \quad (14.70)$$

(The integrations include a sum over the spin components). Using again the variational method

$$\delta[E - \sum_{k,k'} \lambda_{kk'} (\langle \varphi_k | \varphi_{k'} \rangle - \delta_{kk'})] = 0$$

This leads to

$$\begin{aligned} & \left[-\frac{\hbar^2}{2m} \nabla_1^2 + V(\vec{r}_1) \right] \varphi_k(q_1) + \frac{e^2}{4\pi\epsilon_0} \sum_{k'} \int dV_2 \frac{|\varphi_{k'}(q_2)|^2}{|\vec{r}_1 - \vec{r}_2|} \varphi_k(q_1) \\ & - \frac{e^2}{4\pi\epsilon_0} \sum_{k'} \int dV_2 \frac{\varphi_{k'}^*(q_2) \varphi_k(q_2)}{|\vec{r}_1 - \vec{r}_2|} \varphi_{k'}(q_1) = \sum_{k'} \lambda_{kk'} \varphi_{k'}(q_1) \end{aligned} \quad (14.71)$$

Note that in the last term due to orthogonality the spins must be equal. This then leads to the Hartree-Fock equation

$$\begin{aligned} & \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r}) \right] \varphi_j(\vec{r}) + \frac{e^2}{4\pi\epsilon_0} \sum_{k \neq j} \int dV' \frac{|\varphi_k(\vec{r}')|^2}{|\vec{r} - \vec{r}'|} \varphi_j(\vec{r}) \\ & - \frac{e^2}{4\pi\epsilon_0} \sum_{k \neq j; \text{spins} =} \int dV' \frac{\varphi_k^*(\vec{r}') \varphi_j(\vec{r}')}{|\vec{r} - \vec{r}'|} \varphi_k(\vec{r}) = E_j \varphi_j(\vec{r}) \end{aligned} \quad (14.72)$$

We can diagonalize $\lambda_{kk'}$ to obtain its eigenvalues E_k and make a rotation on the wave functions.

Let us now rederive the same results using second quantization. The interacting term of the Hamiltonian is written as

$$H_{int} = \frac{1}{2} \sum'_{ij} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} c_i^\dagger c_i c_j^\dagger c_j \quad (14.73)$$

Introducing Fourier components

$$c_i = \frac{1}{\sqrt{N}} \sum_k e^{ik \cdot r_i} c_k$$

we get that the Hamiltonian can be written as

$$H = \sum_{k,\sigma} \epsilon_k c_{k,\sigma}^\dagger c_{k,\sigma} + \sum_q \frac{2\pi}{q^2} e^2 \sum_{kk'} \sum_{\sigma\sigma'} c_{k+q,\sigma}^\dagger c_{k-q,\sigma'}^\dagger c_{k',\sigma'} c_{k,\sigma} \quad (14.74)$$

We have introduced the Fourier transform of the Coulomb potential

$$\int d^3x e^{-i\vec{k} \cdot \vec{r}} \frac{e^{-\alpha|\vec{r}|}}{|\vec{r}|} = \frac{4\pi}{k^2 + \alpha^2}$$

We have introduced, for regularization a small constant α , which at the end is taken to zero to calculate the Fourier transform of the Coulomb potential. Note that the Fourier transform diverges as $q \rightarrow 0$ due to the long range nature of the potential. This divergence has important consequences as we will see.

The Hamiltonian can now be diagonalized if we take a mean field approach (we will see that the mean field approach leads to the Hartree-Fock approximation). It is actually convenient to use the real space description. In terms of the field operators the Hamiltonian is written as

$$H = \int d^3x \psi^\dagger(x) \left[\frac{p^2}{2m} + v(x) \right] \psi(x) + \frac{1}{2} \int d^3x \int d^3y \psi^\dagger(x) \psi^\dagger(y) V(x-y) \psi(y) \psi(x) \quad (14.75)$$

Let us expand the field operators in terms of the wave functions of the free problem

$$\psi(x) = \sum_j c_j \varphi_j(x)$$

where the wave functions are the solutions of

$$\left[\frac{p^2}{2m} + v(x) \right] \varphi_j(x) = \epsilon_j \varphi_j(x)$$

Using that the field operators anticommute

$$\{\psi(x), \psi^\dagger(x')\} = \delta(x - x')$$

we conclude using that $\{c_j, c_k^\dagger\} = \delta_{j,k}$

$$\sum_j \varphi_j(x) \varphi_j^*(x') = \delta(x - x')$$

If the potential is constant then it is clear that the wave functions are plane waves and we recover the previous case.

Let us consider the equation of motion of the field operator

$$\begin{aligned} i \frac{\partial \psi(x)}{\partial t} &= [\psi(x), H] \\ &= \left(\frac{p^2}{2m} + v(x) \right) \psi(x) + \frac{1}{2} \int d^3y \psi^\dagger(y) V(x-y) \psi(y) \psi(x) + \frac{1}{2} \int d^3y \psi^\dagger(y) V(y-x) \psi(y) \psi(x) \\ &= \left(\frac{p^2}{2m} + v(x) \right) \psi(x) + \int d^3y \psi^\dagger(y) V(x-y) \psi(y) \psi(x) \end{aligned} \quad (14.76)$$

if the potential is symmetric. Let us use now a mean field approximation

$$i\frac{\partial\psi(x)}{\partial t} \sim [\frac{p^2}{2m} + v(x) + \int d^3y V(x-y) \langle \psi^\dagger(y)\psi(y) \rangle] \psi(x) - \int d^3y V(x-y) \langle \psi^\dagger\psi(x) \rangle \psi(y) \quad (14.77)$$

We have replaced

$$\psi^\dagger(y)\psi(y)\psi(x) \rightarrow \langle \psi^\dagger(y)\psi(y) \rangle \psi(x) - \langle \psi^\dagger\psi(x) \rangle \psi(y)$$

To clarify the procedure let us consider a variable A that fluctuates around an average value $\langle A \rangle$. Let us consider the fluctuations $\delta A = A - \langle A \rangle$ and take them to be small. Then we can write that

$$A = \langle A \rangle + (A - \langle A \rangle) = \langle A \rangle + \delta A$$

Let us then consider the product of two fluctuating variables. Using the same decomposition we can write that

$$\begin{aligned} AB &= (\langle A \rangle + \delta A)(\langle B \rangle + \delta B) \\ &= \langle A \rangle \langle B \rangle + \langle A \rangle \delta B + \delta A \langle B \rangle + \delta A \delta B \\ &\sim \langle A \rangle \langle B \rangle + \langle A \rangle (B - \langle B \rangle) + (A - \langle A \rangle) \langle B \rangle \\ AB &\sim \langle A \rangle B + A \langle B \rangle - \langle A \rangle \langle B \rangle \end{aligned}$$

Going back to the equation of motion for the field operator we note that the average values of the type $\langle \psi \rangle$ or $\langle \psi^\dagger \rangle$ vanish because they do not conserve the number of particles. The minus sign of the second term appears because of the anticommutation relations of the field operators

$$\psi(y)\psi(x) = -\psi(x)\psi(y)$$

It is easy to see that in the decoupled equation of motion the first term is the direct Hartree term and the second term is the exchange or Fock term. Expanding the field operators in terms of the eigenfunctions of the free problem we get

$$\begin{aligned} i\frac{\partial\psi(x)}{\partial t} &= \sum_j \epsilon_j c_j \varphi_j(x) \\ &= [\frac{p^2}{2m} + v(x) + \int d^3y V(x-y) \sum_i \varphi_i^*(y) \varphi_i(y) \langle c_i^\dagger c_i \rangle] \sum_j c_j \varphi_j(x) \\ &\quad - \int d^3y V(x-y) \sum_i \varphi_i^*(y) \varphi_i(x) \langle c_i^\dagger c_i \rangle \sum_j c_j \varphi_j(y) \end{aligned} \quad (14.78)$$

We have used that $\langle c_i^\dagger c_j \rangle = \delta_{ij} \langle c_i^\dagger c_i \rangle$. The sum $\sum_i \langle c_i^\dagger c_i \rangle$ is a sum over the occupied sites at zero temperature. Comparing the coefficients of the operator c_j we get that

$$\begin{aligned} \epsilon_j \varphi_j(x) &= \left[\frac{p^2}{2m} + v(x) + \sum_{i, \text{occupied}} \int d^3 y V(x-y) \varphi_i^*(y) \varphi_i(y) \right] \varphi_j(x) \\ &- \sum_{i, \text{occupied}, = \text{spin}} \int d^3 y V(x-y) \varphi_i^*(y) \varphi_i(x) \varphi_j(x) \end{aligned} \quad (14.79)$$

which are again the Hartree-Fock equations. As before we note that only parallel spins are correlated in this approximation.

Let us now look at the consequences of the Hartree-Fock approximation.

14.5.1 Hartree and Fock terms

Let us specify the jellium model. The potential is then

$$v(x) = - \int d^3 y |e| \frac{\rho_0^+(y)}{|x-y|}$$

Here the charge density of the positive background is uniform and given by

$$\rho_0^+(y) = \frac{N|e|}{V}$$

where as before N is the number of electrons. Since we have a uniform background charge the wave functions are plane waves

$$\varphi_k(x) = \frac{1}{\sqrt{V}} e^{ik \cdot x}$$

Then the Hartree equation reduces to

$$\left[\frac{k^2}{2m} - \frac{Ne^2}{V} \int d^3 y \frac{1}{|x-y|} + \frac{(N-1)e^2}{V} \int d^3 y \frac{1}{|x-y|} \right] \frac{1}{\sqrt{V}} e^{ik \cdot x} = \epsilon_k \frac{1}{\sqrt{V}} e^{ik \cdot x} \quad (14.80)$$

In the thermodynamic limit the free particle energy is a solution and $\epsilon_k = k^2/(2m)$.

Consider now the Fock term. Let us consider again plane waves. The Fock term gives

$$- \sum_{l, occupied} \int d^3y \frac{1}{V} e^{i(k_j - k_l) \cdot y} V(x - y) \frac{1}{\sqrt{V}} e^{ik_l \cdot x}$$

It is a solution of the equation

$$\begin{aligned} \epsilon_{k_j} &= \frac{k_j^2}{2m} - \frac{1}{V} \sum_{k_l} \int d^3y e^{i(k_j - k_l) \cdot (y - x)} V(x - y) \\ &= \frac{k_j^2}{2m} - \frac{1}{V} \sum'_{l, occupied} G(k_j - k_l) \end{aligned}$$

where

$$G(k) = \frac{4\pi e^2}{k^2}$$

Defining the z axis parallel to \vec{k}_j and using spherical coordinates we can calculate the integral and obtain the exchange energy due to the Fock term as

$$E_{exchange} = -\frac{e^2}{\pi} \left(\frac{k_F^2 - k_j^2}{2k_j} \ln \left| \frac{k_F + k_j}{k_F - k_j} \right| + k_F \right) \quad (14.81)$$

In the figure we show the exchange energy as a function of momentum. We see that at the Fermi momentum the derivative is infinite. This means that the density of states represented in the figure vanishes at the Fermi energy. This must be wrong. There are two reasons: the first is that plane waves are not a good solution of the interacting problem and second in this approximation the particles interact via the Coulomb potential while in reality there is screening of the potential between two charges due to the effect of the other charges.

We leave as an exercise to show that the energy of the system per atom is given within the Hartree-Fock approximation by

$$\epsilon = \frac{2.21}{r_s^2} - \frac{0.916}{r_s} \quad (14.82)$$

where $r_s = r_0/r_H$ (r_0 is the average distance between the atoms and r_H is the Bohr radius). We see therefore that within the Hartree-Fock approximation there is binding between the electrons. There are bound states as shown in the figure.

14.5.2 Fermi hole

Let us now consider the correlations between two electrons where one is located at site x and the other at site y . If they have opposite spins they are distinguishable and the wave function of the pair is simply

$$\phi(x, y) = e^{ik_1 \cdot x + ik_2 \cdot y}$$

Therefore the probability density is given by

$$g(x, y) = |\phi(x, y)|^2 = 1$$

However if they have the same spin then we have to antisymmetrize the spatial part of the wave function of the pair. We get that

$$\phi(x, y) = \frac{1}{\sqrt{2}} \left(e^{ik_1 \cdot x + ik_2 \cdot y} - e^{ik_2 \cdot x + ik_1 \cdot y} \right)$$

and

$$g(x, y) = 1 - \cos[(k_1 - k_2) \cdot (x - y)]$$

Then

$$\begin{aligned} \langle g(r) \rangle &= \frac{1}{N^2} \sum_{k_1, k_2} (1 - \cos[(k_1 - k_2) \cdot r]) \\ &= \frac{1}{N^2} \Re \sum_{k_1, k_2} \left(1 - e^{i(k_1 - k_2) \cdot r} \right) \\ &= 1 - \Re \frac{1}{N^2} \frac{V^2}{(2\pi)^6} \int d^3 k_1 \int d^3 k_2 e^{i(k_1 - k_2) \cdot r} \end{aligned} \quad (14.83)$$

Defining

$$F(k_F r) = \frac{V}{N_p (2\pi)^3} \int_{k < k_F} d^3 k e^{ik \cdot r} \quad (14.84)$$

(where N_p is the number of parallel spins), we see that

$$\langle g(r) \rangle = 1 - F^2(k_F r)$$

Calculating the integral and using that $V/N_p = 6\pi^2/k_F^3$ we get that

$$F(k_F r) = \frac{3}{(k_F r)^3} (\sin(k_F r) - k_F r \cos(k_F r)) \quad (14.85)$$

This function goes to zero at infinity and it tends to one at the origin. In the figure we show the function $\langle g(r) \rangle$. Integrating over space

$$\frac{1}{V} \int d^3r F^2(k_F r) = \frac{1}{N}$$

Therefore the area missing is just one state: this is called the Fermi hole. Due to the exclusion principle around each electron there is a depletion around it of exactly one state.

Capítulo 13

Electromagnetic response

13.1 Dielectrics and metals

A simplistic description of the difference between insulators and metals is associated with two possible extreme cases where the electrons are either in bound states or free to move around the solid, respectively. A obvious insulating limit is obtained if we consider a gas of neutral atoms, far from each other such that all electrons are bound to their nuclei. As the atoms approach each other one may expect that the wave functions start to overlap and the electrons may be transferred from one atom to another leading to possible extended states, assuming the atoms have some regular distribution in a solid structure. We have seen before that the periodicity of the atoms distribution leads to the possibility of bands and gaps between them originating in the Bragg reflections. A consequence of the gap opening due to the band structure provides a separation between metals and insulators as the bands are partially filled or completely filled, respectively.

In order to obtain information about any system properties we must interact with it and study its response. The behavior of the system is affected by changing various external parameters such as temperature, pressure, chemical potential, magnetic field or electrical field. In this chapter we are interested on the effects of applying an electromagnetic field on a solid. The effects of a static magnetic field were considered before.

To start with let us apply a static electric field along a given direction. If the electrons are free they will accelerate and lead to current transport that increases

indefinitely as time goes by (in the ideal case of no collisions) and gives rise to a conductivity, σ , that goes to infinity as $t \rightarrow \infty$. If the electrons are bound, the electric field will displace the electrons with respect to the position of the nuclei and will lead to a finite polarization and a dielectric response, characterized by the dielectric constant ϵ . If instead of a static electric field we apply an electric field that is oscillating in time, the electrons will oscillate back and forth and the difference between the insulator and the conductor become less sharp. Both the conductivity and the dielectric constant are now a function of the frequency of oscillation of the electric field and are not fully independent. We may rewrite the conductivity as a function $\sigma(\omega)$ and the dielectric constant as $\epsilon(\omega)$.

Let us recall Maxwell's equations

$$\begin{aligned}\vec{\nabla} \cdot \vec{D} &= 4\pi\rho \\ \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{H} &= \frac{4\pi}{c}\vec{J} + \frac{1}{c}\frac{\partial \vec{D}}{\partial t} \\ \vec{\nabla} \times \vec{E} &= -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}\end{aligned}\tag{13.1}$$

These equations are supplemented by the relations

$$\begin{aligned}\vec{D} &= \epsilon\vec{E} \\ \vec{B} &= \mu\vec{H}\end{aligned}\tag{13.2}$$

between the electric displacement vector, \vec{D} , and the electric field, the magnetic induction, \vec{B} , and the magnetic field, \vec{H} , where μ is the magnetic permeability. The displacement vector is related to the polarization of the medium as

$$\vec{D} = \vec{E} + 4\pi\vec{P}\tag{13.3}$$

which implies that

$$\epsilon = 1 + 4\pi\chi_e\tag{13.4}$$

where we have defined the dielectric susceptibility, χ_e , as $\vec{P} = \chi_e\vec{E}$. In addition, we define the current, \vec{J} , through the equation that reflects the charge conservation

$$\vec{\nabla} \cdot \vec{J} + \frac{\partial \rho}{\partial t} = 0\tag{13.5}$$

and in a normal metal we have Ohm's law $\vec{J} = \sigma \vec{E}$.

A plane wave is a solution of Maxwell's equations. Choosing $\vec{E} = \vec{E}_0 e^{i\vec{k}\cdot\vec{r} - i\omega t}$ we can write that

$$\vec{\nabla} \times \vec{H} = \left(\frac{4\pi}{c} \sigma - \frac{i\omega}{c} \right) \vec{E} = -\frac{i\omega}{c} \left(1 + i \frac{4\pi\sigma}{\omega} \right) \vec{E} \quad (13.6)$$

In this expression we have neglected the polarization part by assuming that $\chi_e \sim 0$. Therefore, in a simplified way we may consider that the system is a dielectric with a complex frequency dependent dielectric constant, written as

$$\epsilon(\omega) = 1 + i \frac{4\pi\sigma}{\omega} \quad (13.7)$$

where the imaginary part is proportional to the conductivity (considered here frequency independent) or, equivalently, we may consider the system as a conductor with a complex frequency dependent conductivity given by

$$\sigma(\omega) = -\frac{i\omega}{4\pi} (\epsilon(\omega) - 1) \quad (13.8)$$

Recall now that the index of refraction is defined as $n = c/v$, where v is the light velocity in a given medium and c the velocity in vacuum. Using that $v = \omega/k$ we obtain that

$$n(\omega) = \frac{ck}{\omega} \quad (13.9)$$

Recalling that $n = \sqrt{\epsilon\mu}$ we get that the momentum is given by

$$k^2 = \frac{\mu\omega^2}{c^2} \epsilon(\omega) = \frac{\mu\omega^2}{c^2} \left(1 + i \frac{4\pi\sigma}{\omega} \right) \quad (13.10)$$

This implies that the momentum has a real and an imaginary part $k = \alpha + i\beta$ and that in general any wave will be attenuated in the material.

Using that $k^2 = \omega^2 \mu \epsilon(\omega) / c^2$ we may conclude the following:

- If ϵ is real, (which happens in general if $\sigma = 0$) and if $\epsilon > 0$, then k is real and a wave propagates inside the material. If $\epsilon < 0$ then k is imaginary and the wave is attenuated exponentially away from the surface.
- If ϵ is complex, in general waves are attenuated.

- If $\epsilon = 0$, for which $\chi_e = -1/(4\pi)$, the effective potential due to a charge diverges. Recall that inside a dielectric an external potential, for instance due to a test charge, is renormalized as V_{ext}/ϵ . Therefore, if ϵ vanishes the effective potential diverges. The physical consequence is that, even if the potential is vanishingly small, it is possible to have a finite effective potential. This means an instability of the system and leads to normal oscillation modes.
- If $\epsilon \rightarrow \infty$ the effective potential vanishes. Considering the potential on one charge as due to neighboring charges (such as on one atom due to the neighboring atoms) the forces between the atoms become vanishingly small. This implies the possibility of a structural instability.

A simple microscopic model used to calculate the dielectric constant consists in regarding the electrons as particles of mass m and charge $-e$ as forced damped harmonic oscillators. The classical equation of motion for each electron is written as

$$m \left(\frac{d^2}{dt^2} + \lambda \frac{d}{dt} + \omega_0^2 \right) x = F \quad (13.11)$$

Here λ is the damping coefficient, ω_0 is the frequency of the oscillator, x is the electron displacement from its equilibrium position, and F the external force. Looking for a solution of the type $e^{-i\omega t}$ we find that

$$m (-\omega^2 - i\lambda\omega + \omega_0^2) x_\omega = F_\omega \quad (13.12)$$

Writing that $x_\omega = \alpha(\omega)F_\omega$, where α is the polarizability, we get that

$$\alpha(\omega) = \frac{1}{m (\omega_0^2 - \omega^2 - i\omega\lambda)} \quad (13.13)$$

Using this result we may obtain an expression for the dielectric constant:

$$\epsilon(\omega) = 1 + 4\pi \frac{P_\omega}{E_\omega} = -\frac{4\pi n e x_\omega}{E_\omega} = 4\pi n e^2 \alpha(\omega) \quad (13.14)$$

Therefore

$$\epsilon(\omega) = 1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\omega\lambda} \quad (13.15)$$

where we have defined

$$\omega_p^2 = \frac{4\pi n e^2}{m} \quad (13.16)$$

is the plasma frequency, as we will explain later.

Separating the dielectric constant in real and imaginary parts $\epsilon(\omega) = \epsilon'(\omega) + i\epsilon''(\omega)$ we get that

$$\begin{aligned} \epsilon'(\omega) &= 1 + \omega_p^2 \frac{(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \omega^2 \lambda^2} \\ \epsilon''(\omega) &= \omega_p^2 \frac{\omega \lambda}{(\omega_0^2 - \omega^2)^2 + \omega^2 \lambda^2} \end{aligned} \quad (13.17)$$

Using the simplified relation between the conductivity and the dielectric constant we obtain that

$$\begin{aligned} \sigma(\omega) &= -i \frac{\omega}{4\pi} (\epsilon(\omega) - 1) \\ &= -i \frac{\omega}{4\pi} \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\omega\lambda} \end{aligned} \quad (13.18)$$

Separating the conductivity in the real and imaginary parts $\sigma(\omega) = \sigma'(\omega) + i\sigma''(\omega)$ we obtain that

$$\begin{aligned} \sigma'(\omega) &= \frac{\omega^2 \lambda}{4\pi} \frac{\omega_p^2}{(\omega_0^2 - \omega^2)^2 + \omega^2 \lambda^2} \\ \sigma''(\omega) &= -\frac{\omega \omega_p^2}{4\pi} \frac{(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \omega^2 \lambda^2} \end{aligned} \quad (13.19)$$

Let us now consider that the electrons are free. This can be particularized taking

$\omega_0 = 0$ (no harmonic force on the electron). In this case we get that

$$\begin{aligned}
 \alpha(\omega) &= \frac{1}{m\omega} \frac{-1}{\omega + i\lambda} \\
 \epsilon'(\omega) &= 1 - \omega_p^2 \frac{1}{\omega^2 + \lambda^2} \\
 \epsilon''(\omega) &= \frac{\omega_p^2}{\omega} \frac{\lambda}{\omega^2 + \lambda^2} \\
 \sigma'(\omega) &= \frac{\lambda}{4\pi} \frac{\omega_p^2}{\omega^2 + \lambda^2} \\
 \sigma''(\omega) &= \frac{\omega\omega_p^2}{4\pi} \frac{1}{\omega^2 + \lambda^2}
 \end{aligned} \tag{13.20}$$

Note that λ has dimensions of frequency. Therefore, it may also be written as the inverse of a time scale as $\lambda = 1/\tau$. This time scale has the interpretation of an average time between collisions that lead to damping of the electrons.

An alternative way to present the expression for the conductivity in terms of the scattering time between collisions is

$$\begin{aligned}
 \sigma'(\omega) &= \sigma_0 \frac{1}{1 + (\omega\tau)^2} \\
 \sigma''(\omega) &= \sigma_0 \frac{\omega\tau}{1 + (\omega\tau)^2}
 \end{aligned} \tag{13.21}$$

where

$$\sigma_0 = \frac{ne^2\tau}{m} \tag{13.22}$$

is the d.c. (zero frequency) conductivity of a gas of non-interacting free electrons subject to a drag force due to collisions with imperfections, phonons, etc.

In general, in a dielectric there are different types of oscillatory motion, with different natural frequencies, ω_j . Each oscillator contributes additively to the polarization of the system leading straightforwardly to a dielectric function of the form

$$\epsilon(\omega) = 1 + \frac{4\pi ne^2}{m} \sum_j \frac{f_j}{\omega_j^2 - \omega^2 - i\omega\lambda_j} \tag{13.23}$$

where λ_j is the damping coefficient of oscillator j , and f_j is the number of electrons with natural frequency, ω_j . Considering that there are z electrons per molecule (and n is the number of molecules per unit volume) there is a sum rule $\sum_j f_j = z$.

13.1.1 Ideal conductor: Drude peak and plasma frequency

A very pure material is a system for which the friction coefficient is very small (small λ) or the time between collisions is very long (large τ). A good conductor is therefore approximated if we consider $\lambda \rightarrow 0$ (or $\tau \rightarrow \infty$). It is convenient to consider this limit in the expression of the polarizability. Using that

$$\lim_{\lambda \rightarrow 0} \frac{1}{\omega + i\lambda} = P\left(\frac{1}{\omega}\right) - i\pi\delta(\omega) \quad (13.24)$$

we obtain that the polarizability is given by

$$\alpha(\omega) = -\frac{1}{m\omega} \left(P\left(\frac{1}{\omega}\right) - i\pi\delta(\omega) \right) \quad (13.25)$$

The conductivity is given by

$$\begin{aligned} \sigma'(\omega) &= \frac{1}{4}\omega_p^2\delta(\omega) \\ \sigma''(\omega) &= \frac{1}{4\pi}\frac{\omega_p^2}{\omega} \end{aligned} \quad (13.26)$$

and the dielectric constant by

$$\begin{aligned} \epsilon'(\omega) &= 1 - \frac{\omega_p^2}{\omega^2} \\ \epsilon''(\omega) &= \frac{\pi\omega_p^2}{\omega}\delta(\omega) \end{aligned} \quad (13.27)$$

Several comments should be made at this point:

- Consider first the real part of the conductivity. Eq. (13.20) shows that if the damping term is finite (λ finite) the real part has a Lorentzian shape centered at zero frequency and with a width given by λ . This is called the Drude peak of the conductivity and is a characteristic of a conductor. As the material becomes a better metal and $\lambda \rightarrow 0$ the peak sharpens around zero frequency and in the ideal limit becomes a Dirac- $\delta(\omega)$ infinite d.c. conductivity as expected of a pure conductor.

- In the same limit of a very pure conductor the real part of the dielectric function eq. (13.27) has a singular behavior. Near zero frequency it diverges but it is negative. As it turns out it is negative for frequencies $\omega < \omega_p$. Since $\epsilon < 0$ this implies that any wave of this frequency range when incident on the system, is totally reflected (meaning that it is attenuated exponentially inside the material). In the frequency range $\omega > \omega_p$ the dielectric function is positive, and since in the limit $\lambda \rightarrow 0$ the imaginary part vanishes, waves propagate in the material. The frequency point $\omega = \omega_p$ is special since $\epsilon = 0$. This corresponds to oscillatory normal modes called plasma oscillations. Considering that the nuclei are fixed (due to their large mass and interactions that lead to the regular structure of the solid, only the electrons oscillate around their average positions. This displacement leads to the polarization that ultimately leads to a non-trivial dielectric function. As a result, we have a relative motion of the negative charges with respect to the positive charges and is the origin of the plasma oscillations.

13.1.2 Atomic contribution to dielectric function

If there are several oscillator modes the results obtained are easily generalized. Reintroducing the dielectric limit eq. (13.23) we may consider that in general the damping coefficients are small $\lambda_j \sim 0$. Therefore, $\epsilon(\omega)$ has a small imaginary part. Neglecting the imaginary part if the frequency is not very close to one of the resonant natural frequencies, the fraction

$$\frac{1}{\omega_j^2 - \omega^2}$$

is positive if $\omega < \omega_j$ and is negative if $\omega > \omega_j$. Therefore, close to the resonance, $\omega = \omega_j$, the fraction changes sign from positive to negative as the frequency increases. For very small frequency, smaller than any of the natural frequencies of the oscillators, $\epsilon(\omega) > 1$. As the frequency increases, the frequency ω will cross each natural frequency in turn. Each time a crossing occurs negative contributions appear that decrease the dielectric function. When $\omega > \omega_j$ for any j the sum is necessarily negative and $\epsilon(\omega) < 1$. For large enough frequency we get that

$$\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2} \quad (13.28)$$

where

$$\omega_p^2 = \frac{4\pi nze^2}{m} \quad (13.29)$$

is the plasma frequency. Using that the refraction index $n(\omega) = \sqrt{\epsilon(\omega)}$, where we have considered that $\mu \sim 1$, and replacing eq. (13.28) we can write that

$$\omega n(\omega) = \sqrt{\omega^2 - \omega_p^2} \quad (13.30)$$

Since $n(\omega) = ck/\omega$ we get the dispersion relation

$$\omega^2 = \omega_p^2 + c^2k^2 \quad (13.31)$$

Since ω_p is finite, this dispersion relation has a (large) gap. These excitations require a large energy to be excited of the order of several electron Volt.

If the frequency is very close to one of the natural frequencies of the oscillators, the damping can not be neglected. At the resonance energy the contribution to the dielectric function is purely imaginary and the imaginary part has a peak. This signals energy dissipation and since the imaginary part is positive it signals resonant absorption. Clearly, if one of the natural frequencies is vanishingly small we recover the Drude peak found previously.

In general the dielectric function is obtained as the sum of various contributions. The electronic, ionic and dipolar. The electronic and ionic contributions are of similar type. The ionic contribution arises since the inner electronic orbitals are also polarizable and the ions are not strictly static. The dipolar contribution has its origin in the possible existence of permanent dipolar moments in an atom or molecule. Their motion in the presence of a perturbation is different if we assume that the dipolar moment has a fixed magnitude and at most rotates rigidly in the presence of an orienting electric field. In this case its contribution to the real part of the dielectric function has a maximal value at zero frequency and then it decreases as frequency increases. There is no plasma oscillation mode in this case, and therefore the characteristic resonance through the natural oscillator frequency is absent.

13.2 Transverse and longitudinal modes

An electromagnetic wave is a transverse mode of propagation. From the wave equation we have that

$$\begin{aligned}\frac{\partial^2 \vec{D}}{\partial t^2} &= c^2 \nabla^2 \vec{E} \\ \vec{E} &\sim e^{-i\omega t} e^{i\vec{k} \cdot \vec{r}} \\ \vec{D} &= \epsilon(\omega, \vec{k}) \vec{E}\end{aligned}\tag{13.32}$$

This leads to

$$c^2 k^2 = \epsilon(\omega, \vec{k}) \omega^2 \tag{13.33}$$

This is equivalent to the previous relation $\omega n(\omega) = ck$ (since $n(\omega) = \sqrt{\epsilon(\omega)}$).

Let us consider the plasma of the electrons and ions (considered static) in a solid. The ions contribute to the dielectric constant and change the vacuum contribution ($\epsilon = 1$) to a different value, call it $\epsilon(\infty)$. The dielectric function of the electrons in the presence of the ions is therefore corrected to

$$\epsilon(\omega) = \epsilon(\infty) - \frac{4\pi n e^2}{m\omega^2} \tag{13.34}$$

Therefore, we get that

$$\epsilon(\omega) = \epsilon(\infty) \left(1 - \frac{\tilde{\omega}_p^2}{\omega^2} \right) \tag{13.35}$$

where the plasma frequency is changed to

$$\tilde{\omega}_p^2 = \frac{4\pi n e^2}{\epsilon(\infty) m} \tag{13.36}$$

As before, if the frequency is smaller than the plasma frequency the wave is fully reflected since the momentum becomes purely imaginary ($k^2 < 0$). The dispersion relation of the transverse wave is given by

$$\omega^2 = \tilde{\omega}_p^2 + \frac{c^2 k^2}{\epsilon(\infty)} \tag{13.37}$$

13.2.1 Plasmons

The normal longitudinal modes are given by the zeros of the dielectric function. Since they are spontaneous oscillations do not require the stimulus of an external field and are the result of mutual interactions between the plasma constituents. In this context this is achieved at the frequency $\omega_L = \tilde{\omega}_p$. A straightforward picture is obtained considering a solid where the ions are static at given positions (such as the crystal structure of the material) and where, at an initial stage the electrons' average positions correspond to centers of local neutrality. Assume some internal perturbation induces a displacement of the electrons from their equilibrium average positions. To simplify we neglect here spatial modulations and assume a zero momentum, uniform displacement of the electrons. We can view the system as two slabs one charged positively with the ions and one charged negatively with the electrons, that are displaced with respect to the another. It is convenient to think of both slabs as with finite thickness. The relative displacement induces a surface charge, for instance in the ions since they are assumed static. The surface charges have different signs and are given by $\sigma = \pm neu$, where u is the displacement of the electron slab with respect to the ion slab. As a consequence, an electric field arises inside the slab $E = 4\pi neu$. This electric field acts on the electrons with a force that tends to restore equilibrium and leads to an equation of motion

$$nm \frac{d^2 u}{dt^2} = -neE = -4\pi n^2 e^2 u \quad (13.38)$$

This can be rewritten as

$$\frac{d^2 u}{dt^2} + \omega_p^2 u = 0 \quad (13.39)$$

where ω_p is the plasma frequency. Clearly the solution is an oscillatory motion around the equilibrium position. Also, it is a longitudinal wave. It is also clear that, since $\epsilon = 0$ this implies $\vec{D} = 0$ and a field that opposes the polarization $\vec{E} = -4\pi\vec{E}$, as expected.

These modes may be quantized and are called plasmons. Their dispersion relation will be obtained later on.

13.3 Screening

If we insert a test charge in a collection of electric charges, the system responds to this perturbation through a polarization that creates a field that adds up to the external field. The total field acts locally to create the polarization and, therefore, the solution of the problem is only obtained self-consistently: the system of charges is polarized due to the total local potential which is the sum of the external field and the field due to the polarization. The local potential is given by the external field divided by the dielectric function.

Consider an external potential of the type $\delta V_{ext}(\vec{q}) = V e^{i\vec{q}\cdot\vec{r}}$. We consider a static perturbation (zero frequency) or a perturbation that acts very slowly on the system of charges. A similar form is assumed for the polarization $\delta\Phi(\vec{q}) = \Phi e^{i\vec{q}\cdot\vec{r}}$ and the total potential $\delta U(\vec{q}) = u e^{i\vec{q}\cdot\vec{r}}$. We can write that

$$\delta U(\vec{q}) = \frac{\delta V_{ext}(\vec{q})}{\epsilon(\vec{q}, \omega = 0)} \quad (13.40)$$

13.3.1 Thomas-Fermi approximation

We further simplify the perturbation assuming that in a small region the potential energy is changed with respect to neighboring regions. Therefore, the chemical potential is changed locally. Assume this change is positive. Since the perturbation is very slow there is a flow of charge from this region to the neighboring regions to equilibrate the chemical potential. As a consequence there is an electron charge depletion at the same region. The depletion in the electron density is given approximately by

$$\delta n(\vec{r}) = -D(\epsilon_F) \delta U(\vec{r}) \quad (13.41)$$

where $D(\epsilon_F)$ is the density of states per volume of the electronic system, at the Fermi level. This is the Thomas-Fermi approximation. A change in the electronic density induces a change in the potential energy via Poisson's equation

$$\nabla^2 (\delta\Phi) = -4\pi(-e)\delta\rho \quad (13.42)$$

where $\delta\rho = -e\delta n(\vec{r})$. Therefore

$$\nabla^2 (\delta\Phi) = -4\pi e^2 \delta n(\vec{r}) = 4\pi e^2 D(\epsilon_F) \delta U(\vec{r}) = \lambda^2 \delta U(\vec{r}) \quad (13.43)$$

Using the plane wave solution we get that

$$\begin{aligned}
 -q^2 \delta\Phi &= \lambda^2 \delta U(\vec{r}) \\
 \delta\Phi &= -\frac{\lambda^2}{q^2} \delta U \\
 \delta U &= V + \delta\Phi = V - \frac{\lambda^2}{q^2} \delta U \\
 \delta U \left(1 + \frac{\lambda^2}{q^2}\right) &= V \\
 \delta U &= \frac{V}{1 + \frac{\lambda^2}{q^2}}
 \end{aligned} \tag{13.44}$$

This final result implies that the dielectric function is given, in this limit, by

$$\epsilon(q) = 1 + \frac{\lambda^2}{q^2} \tag{13.45}$$

where $\lambda^2 = 4\pi e^2 D(\epsilon_F)$.

Consider now a test charge, such as an electron, at a given point in space. The potential exerted on any other electron at a distance r from the test charge, is given by

$$\delta V = \frac{e^2}{r} \tag{13.46}$$

Taking the Fourier transform gives the potential

$$\delta V(\vec{q}) = \frac{4\pi e^2}{q^2} \tag{13.47}$$

The total potential is given by

$$\delta U(\vec{q}) = \frac{\delta V(\vec{q})}{\epsilon(\vec{q}, 0)} = \frac{4\pi e^2}{q^2} \frac{1}{1 + \frac{\lambda^2}{q^2}} \tag{13.48}$$

This leads to the results

$$\begin{aligned}
 \delta U(\vec{q}) &= \frac{4\pi e^2}{q^2 + \lambda^2} \\
 \delta U(\vec{r}) &= \frac{e^2}{r} e^{-\lambda r}
 \end{aligned} \tag{13.49}$$

The effective potential is therefore not Coulombic. It does not decrease as the distance inverse but much faster, exponentially. $1/\lambda$ gives the distance scale over which the potential decays. This is the screening effect one observes in a conductor. It turns out that in an insulator the behavior is different, as will be shown later. In an insulator the decay is Coulombic.

13.3.2 Mott transition

Consider a system of protons and an equal number of electrons and single out one of these electrons. The potential energy between each proton and this electron is screened and of the type

$$U(r) = -\frac{e^2}{r}e^{-\lambda r} \quad (13.50)$$

where r is the distance between the electron and a given proton. Recall that $\lambda^2 = 4\pi e^2 D(\epsilon_F) = 3.939 n_0^{1/3}/a_0$, where n_0 is the electronic density and a_0 is the Bohr radius. Note that as the density increases λ also increases. If λ is large, the potential energy is small unless r is small. Therefore, the potential is short ranged. As a consequence even though attractive it may not lead to a bound state between the electron and the proton. Therefore the electron may be delocalized since the proton is not able to bind the electron. In three dimensions a minimum attractive potential is needed to create a bound state. It can be shown that there is a bound state if $\lambda < 1.19/a_0$. If the bound state occurs, then the system is expected to be an insulator: the electrons will be bound to their respective protons. A sufficiently small λ occurs if the density is small enough, or the atoms are sufficiently far apart. Clearly, if the atoms are very far from each other the system is insulator. As the atoms approach, their relative distances decrease, the density increases and, assuming a simple cubic structure ($n_0 = 1/a^3$), the system will have a transition from insulator to a metal when the distance between the atoms (lattice constant) is $a < a_c = 2.78a_0$. A more detailed calculation has lead Mott to the prediction that $a_c = 4.5a_0$.

13.3.3 Screening and phonons

Phonons were introduced before as the vibrational modes of the atoms of a lattice structure that result from harmonic motions around equilibrium positions. The model used considered that the atoms are connected by springs with given force constants that are adjusted phenomenologically, for instance measuring the sound

velocity, or different types of deformations of the solid. Ultimately, the interactions between the atoms have an electromagnetic origin. We will next try to establish a connection between the dielectric response of a solid and its vibrations.

Consider, to simplify, longitudinal acoustic phonons. As discussed above, the longitudinal normal modes are obtained from the zeros of the real part of the dielectric function $\epsilon(\omega_L) = 0$. We will consider the total dielectric function due to the sum of the electronic and ionic contributions. We will consider the adiabatic approximation since the ions move slowly with respect to the electrons. The interactions between the charges of the ions are screened by the electrons. We will use the Thomas-Fermi approximation for the dielectric function of the electrons

$$\epsilon_{(e)}(\omega, \vec{k}) = 1 + \frac{\lambda^2}{k^2} \quad (13.51)$$

We consider that the ions are far from each other and that move independently from each other. We use $\epsilon(\omega, 0)$ with the mass of the ions, M , leading to the total dielectric function

$$\epsilon(\omega, k) = 1 - \frac{4\pi ne^2}{M\omega^2} + \frac{\lambda^2}{k^2} \quad (13.52)$$

We are interested in studying the regime where both the frequency and momenta are small, where the dispersion relation of the phonons is linear with a slope given by the sound velocity. In the regime where both k and ω are small, we get that the normal modes are obtained imposing

$$\epsilon(\omega, k) \sim -\frac{4\pi ne^2}{M\omega^2} + \frac{\lambda^2}{k^2} = 0 \quad (13.53)$$

This gives the dispersion relation

$$\omega^2 = \frac{4\pi ne^2}{M\lambda^2} k^2 = \frac{m}{3M} v_F^2 k^2 \quad (13.54)$$

Therefore, the sound velocity is given by

$$v_s = \left(\frac{m}{3M} \right)^{1/2} v_F \quad (13.55)$$

There is a reasonably good agreement with the experimental result for the sound velocity of simple metals like the alkalis. For instance for potassium the calculated value is about $v_s = 1.8 \times 10^5$ cm/s while the experimental value is 2.2×10^5 cm/s.

At high frequencies the Thomas-Fermi approximation for the electrons is not good. It is a better approximation to write in the zero momentum limit

$$\epsilon(\omega, 0) = 1 - \frac{4\pi ne^2}{M\omega^2} - \frac{4\pi ne^2}{m\omega^2} \quad (13.56)$$

The eigenmodes are then just the plasma frequency where the electronic mass is replaced by the reduced mass between the electron and the ion.

13.4 Polaritons

As we have seen previously, if the unit cell of a lattice has more than one atom, in general there are, in addition to the acoustic phonons, higher energy branches called optical phonons. Considering two atoms per cell, while the acoustic phonons describe in phase motion of the atoms of the unit cell, the optical phonons describe motion where the atoms move in opposition of phase. Both the acoustical and the optical phonons in general have longitudinal and transverse modes.

If electromagnetic waves are incident in the material one expects that they may polarize the ions and excite transverse oscillation modes. Since the speed of light is very high and much higher than the speed of sound, the dispersion relation $\omega = ck$ will not intersect the acoustical branch. We expect, therefore, that the photon will interact mainly with the optical phonon since it will become degenerate with this mode at very small momentum but finite energy. This will lead to some perturbation of the motion of the ions, that requires degenerate perturbation theory involving two modes, and leads to the opening of a gap, that is not related to the periodicity of the lattice. The eigenmodes that result from the coupling of the photons and the phonons are called polaritons.

It is actually easier to consider this problem classically, in the sense that electromagnetism contains the root for this coupling. Recall that we can write that

$$c^2 k^2 E = \omega^2 (E + 4\pi P) \quad (13.57)$$

Let us call ω_T the zero momentum energy of the transverse acoustical phonon. In general the dispersion relation near zero momentum is flat and, since the momentum region where the coupling occurs is very narrow, we may consider that the frequency is a constant and independent of momentum, k . If some force is applied on the

atoms, these will oscillate around their equilibrium positions giving rise to some polarization, $P = nqu$, where u is the displacement. The equation of motion for the polarization is of the type

$$\frac{d^2 P}{dt^2} + \omega_T^2 P = F$$

(where F is the force) which leads to

$$-\omega^2 P + \omega_T^2 P = \frac{nq^2}{M} E \quad (13.58)$$

Here n is the number of pairs of ions of effective charge q and reduced mass, M , per unit volume. Let us neglect the electronic contribution to the polarization and focus on the ionic contribution. Putting together Maxwell's equation for the polarization and the forced harmonic oscillator equation we get that

$$\begin{aligned} -(c^2 k^2 - \omega^2) E + 4\pi\omega^2 P &= 0 \\ \frac{nq^2}{M} E - (\omega_T^2 - \omega^2) P &= 0 \end{aligned} \quad (13.59)$$

A non-trivial solution is obtained if the determinant of the coefficients of the electric field and the polarization vanishes. (We may think of the electric field as the "photon" and the polarization as the "phonon"). The determinant is given by

$$\det \begin{pmatrix} \omega^2 - c^2 k^2 & 4\pi\omega^2 \\ \frac{nq^2}{M} & \omega^2 - \omega_T^2 \end{pmatrix} = 0$$

The solution leads to two modes

$$\omega_{\pm}^2 = \frac{1}{2} \left(c^2 k^2 + \omega_T^2 + \frac{4\pi nq^2}{M} \pm \sqrt{\left(c^2 k^2 + \omega_T^2 + \frac{4\pi nq^2}{M} \right)^2 - 4c^2 k^2 \omega_T^2} \right) \quad (13.60)$$

At zero momentum, $k = 0$, the two branches give

$$\begin{aligned} \omega_+^2 &= \omega_T^2 + \frac{4\pi nq^2}{M} \\ \omega_-^2 &= 0 \end{aligned} \quad (13.61)$$

On the other hand, in the limit when the momentum is large the lower mode has the limit $\omega_- = \omega_T$. Therefore, a gap appears in the spectrum. At low momenta the lower branch is very similar to a photon while the upper branch is similar to an acoustical phonon. As the momentum increases the two branches result from a strong mixing of the photon and phonon and at larger momenta the lower branch is similar to the optical phonon and the upper branch is similar to the photon mode.

The dielectric function is given by

$$\epsilon(\omega) = 1 + \frac{4\pi P}{E} = 1 + \frac{4\pi n q^2 / M}{\omega_T^2 - \omega^2} \quad (13.62)$$

Including the core electronic contribution it is renormalized to

$$\epsilon(\omega) = \epsilon(\infty) + \frac{4\pi n q^2 / M}{\omega_T^2 - \omega^2} \quad (13.63)$$

Consider now the case when $\omega = 0$. We get that

$$\epsilon(0) = \epsilon(\infty) + \frac{4\pi n q^2}{M \omega_T^2} \quad (13.64)$$

This means we can rewrite

$$\begin{aligned} \epsilon(\omega) &= \epsilon(\infty) + (\epsilon(0) - \epsilon(\infty)) \frac{\omega_T^2}{\omega_T^2 - \omega^2} \\ &= \frac{\omega_T^2 \epsilon(0) - \omega^2 \epsilon(\infty)}{\omega_T^2 - \omega^2} \\ &= \epsilon(\infty) \left(\frac{\omega_L^2 - \omega^2}{\omega_T^2 - \omega^2} \right) \end{aligned} \quad (13.65)$$

The dielectric function diverges when $\omega = \omega_T$. This means a structural instability resulting from a vanishing force between the system charges. It is a signature of some sort of order in the distribution of the electric charges such as ferroelectric order. The dielectric function vanishes when $\omega = \omega_L$. When the frequency is in the range $\omega_T < \omega < \omega_L$ the dielectric function is negative, the waves are attenuated which means there are no propagating modes, which is consistent with the existence of a

gap in the spectrum. In this interval the reflection coefficient approaches one. When the frequency vanishes we get the Lyddane-Sachs-Teller (LST) relation

$$\frac{\omega_L^2}{\omega_T^2} = \frac{\epsilon(0)}{\epsilon(\infty)} \quad (13.66)$$

13.5 Optical processes

The relation between the dielectric function and the index of refraction allows experimental access using optical methods. As we have shown before, in general the index of refraction has a real and an imaginary part

$$n(\omega) = \sqrt{\epsilon + i \frac{4\pi\sigma}{\omega}} = n_r + in_i \quad (13.67)$$

Recall that the current is given by

$$\frac{4\pi}{c} \vec{J} = -i \frac{\omega}{c} n^2(\omega) \vec{E} \quad (13.68)$$

The power dissipated by Joule's effect is given by

$$\frac{1}{2} \text{Re}(\vec{J} \cdot \vec{E}) = \frac{1}{2} \text{Re}\left(-i \frac{\omega}{c} n^2(\omega) \frac{c}{4\pi}\right) |\vec{E}|^2 \quad (13.69)$$

Using that

$$n^2(\omega) = (n_r + in_i)^2 = n_r^2 - n_i^2 + 2in_r n_i$$

and defining the absorption coefficient, η , by the ratio between the dissipated power and the power transmitted we get that

$$\eta = \frac{\frac{1}{2} \text{Re}(\vec{J} \cdot \vec{E})}{\frac{c}{8\pi} n_r |\vec{E}|^2} = \frac{2n_i \omega}{c} \quad (13.70)$$

A measurement of the absorption coefficient directly reads the imaginary part of the index of refraction.

On the other hand we may also measure the reflection part. Considering normal incidence on the material the reflected part over the incident part, R , is given by

$$R = \left| \frac{1 - n}{1 + n} \right|^2 = \frac{(n_r - 1)^2 + n_i^2}{(n_r + 1)^2 + n_i^2} \quad (13.71)$$

A combined measurement of the absorption and reflection coefficients gives information about the real and imaginary parts of the index of refraction and the dielectric function. Indeed

$$\begin{aligned}\epsilon'(\omega) &= n_r^2 - n_i^2 \\ \epsilon''(\omega) &= 2n_r n_i\end{aligned}\tag{13.72}$$

However, the real and imaginary parts of the dielectric function are not independent and do not need to be both measured. The reason is that the dielectric function is a response function and does satisfy Kramers-Kronig relations which imply that the imaginary and real parts are related, if we know the whole frequency dependency of one of the functions.

Recall that $\epsilon - 1 = 4\pi\chi_e$, where χ_e is a susceptibility that relates the polarization with a given electric field. Therefore, it is a response function that satisfies Kramers-Kronig relations.

$$\chi(\omega) = -iP \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\omega'}{\omega' - \omega}\tag{13.73}$$

Separating the dielectric function in real and imaginary parts we get that

$$\begin{aligned}\epsilon'(\omega) - 1 &= P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\epsilon''(\omega')}{\omega' - \omega} \\ \epsilon''(\omega) &= -P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\epsilon'(\omega') - 1}{\omega' - \omega}\end{aligned}\tag{13.74}$$

These can be rewritten as

$$\begin{aligned}\epsilon'(\omega) - 1 &= \frac{2}{\pi} P \int_0^{\infty} d\omega' \frac{\omega' \epsilon''(\omega')}{(\omega')^2 - \omega^2} \\ \epsilon''(\omega) &= -\frac{2\omega}{\pi} P \int_0^{\infty} d\omega' \frac{\epsilon'(\omega')}{(\omega')^2 - \omega^2}\end{aligned}\tag{13.75}$$

or

$$\begin{aligned}n_r^2(\omega) - n_i^2(\omega) &= \frac{2}{\pi} P \int_0^{\infty} d\omega' \frac{\omega' 2n_r(\omega')n_i(\omega')}{(\omega')^2 - \omega^2} \\ 2n_r(\omega)n_i(\omega) &= -\frac{2\omega}{\pi} P \int_0^{\infty} d\omega' \frac{n_r^2(\omega') - n_i^2(\omega')}{(\omega')^2 - \omega^2}\end{aligned}\tag{13.76}$$

In addition to the "classical" electromagnetic response of incident radiation on a material there are also quantum effects due to the interaction of the photons with the energy bands of the electrons. In general, the electronic levels are organized in bands, separated by finite gaps. Since the velocity of photons are very high the momentum transfer between a photon and an electron is rather small, for a given finite energy transfer. Therefore, with good approximation we may consider that these processes involve no momentum transfer. They are however responsible for transitions between bands and only occur if the energy of an absorbed photon equals the energy difference between different electronic energy bands. Excitations of electrons within the same band can have arbitrarily small energies and small or finite momenta and do not involve absorption of a single photon.

13.5.1 Excitons

Assume that we supply a finite energy and finite momentum so that an electron gains energy and momentum and is able to occupy an empty state leaving behind an empty state. This is a particle-hole excitation. This is a picture of the process in momentum (energy) space. Consider however a similar process in real space. Assume that by sending energy into the system an electron is excited from a given state leaving behind a hole. We may assume that the electron absorbs a photon and is promoted to an excited band. Since the hole and the electron have opposite charges there will be some attraction between the two and they will eventually tend to recombine to form some sort of boundstate called exciton. Consider the possibility that this process occurs but that the electron has not returned to its initial state. If the binding is weak the pair will be loosely bound and the electron and the hole will move together in a pair that extends over space in a way similar to a hydrogen-like atom. This is called a Mott-Wannier exciton and is characterized by an extended wave function. Since this is expected to be a boundstate its energy is negative with respect to the lower edge of the band to which the electron was promoted by the absorption of the photon. We expect therefore energy levels close to the lower edge of the excited band, due to the presence of these excitons. Another possibility is that the system strongly binds the excited electron to the atom it originated from. This happens if the system has a strong ionic character and each electron is strongly bound to atomic orbitals which implies an insulator to begin with. Bound states between the electron and the hole it left behind are also possible but are typically strongly bound. These are called Frenkel excitons and have functions that are rather localized.

These two limiting cases of very weakly bound and strongly bound excitons, lead to a wide range of binding energies that can be as small as 1meV or as a strong as 1 eV.

The Mott-Wannier excitons can be well modelled by a simple attractive potential between the members of the pair as

$$U(r) = -\frac{e^2}{\epsilon r} \quad (13.77)$$

The problem has a straightforward solution as the hydrogen atom and leads to energy levels as

$$E_n = E_g - \frac{\mu e^4}{2\hbar^2 \epsilon^2} \frac{1}{n^2} \quad (13.78)$$

with respect to the energy gap to the excited band and where μ is the reduced mass of the electron and hole. For instance, for Cu_2O we have that $\epsilon = 10, \mu = 0.7m$ which gives a binding energy of about 95meV as compared to the gap to the excited band which is about 2.17 eV.

The strongly coupled nature of the Frenkel excitons suggests that the electrons and their pair could move between nearest atoms, as in the tight-binding model, giving origin to some narrow band below the excited band. Consider a crystal with N atoms and to simplify, consider a one-dimensional arrangement of atoms. Denote the groundstate of the system as $\psi_g = u_1 u_2 \cdots u_N$, where u_i are one-particle Wannier wave functions localized at atom i . Consider now that at site j the electron is in an excited state v_j . The wave function is changed to $\phi_j = u_1 u_2 \cdots u_{j-1} v_j u_{j+1} \cdots u_N$. Consider now an Hamiltonian that essentially is kinetic energy. It will be possible for the electron to stay at the same atom or to change its location either to the left or the right of site j . We consider then that the action of the Hamiltonian on the wave function ϕ_j is of the type

$$H\phi_j = \epsilon\phi_j + T(\phi_{j-1} + \phi_{j+1}) \quad (13.79)$$

where T is the probability amplitude that the electron, under the action of the Hamiltonian, changes its location. Clearly, the wave function ϕ_j is not an eigenfunction of the Hamiltonian since the electron in the excited state v_j may delocalize. In order to diagonalize the Hamiltonian we may choose instead a wave function of the Bloch form

$$\psi_k = \sum_j e^{ijk a} \phi_j \quad (13.80)$$

Applying the Hamiltonian to this trial wave function we get that

$$\begin{aligned}
 H\psi_K &= \sum_j e^{ijka} H\phi_j \\
 &= \sum_j e^{ijka} (\epsilon\phi_j + T\phi_{j-1} + T\phi_{j+1}) \\
 &= \sum_j e^{ijka} (\epsilon + Te^{ika} + Te^{-ika}) \phi_j \\
 &= (\epsilon + 2T \cos(ka)) \psi_k
 \end{aligned} \tag{13.81}$$

which is therefore an eigenfunction of the Hamiltonian with eigenvalue

$$E_k = \epsilon + 2T \cos(ka) \tag{13.82}$$

The bandwidth is determined by T , which is small if the electron-hole pair is strongly bound. An example of this type of exciton is found in ionic crystals such as KBr .

14.7 Coulomb interaction. Lindhardt function

Let us now consider second order perturbation theory to see the effect of the Coulomb potential on the electron gas in any number of dimensions. Consider the correction to the energy of two particles in states 1 and 2 in second order

$$\epsilon_{1,2}^{(2)} = - \sum_{3,4} 2m \frac{\langle 1, 2 | V | 3, 4 \rangle \langle 3, 4 | V | 1, 2 \rangle}{(k_3^2 + k_4^2) - (k_1^2 + k_2^2)} \quad (14.97)$$

where

$$\langle 1, 2 | V | 3, 4 \rangle = \frac{4\pi e^2}{\Omega q^2} \delta_{\vec{k}_1 + \vec{k}_2, \vec{k}_3 + \vec{k}_4} \delta_{\vec{q} = \vec{k}_1 - \vec{k}_3 = \vec{k}_4 - \vec{k}_2}$$

and Ω is the volume. We can expand the above expression to obtain that

$$\begin{aligned} \epsilon_{1,2}^{(2)} &= -2m \left(\frac{4\pi e^2}{\Omega q^2} \right)^2 \sum_{\vec{q}} \frac{1}{q^4} \frac{1}{(\vec{k}_1 - \vec{q})^2 + (\vec{k}_2 + \vec{q})^2 - \vec{k}_1^2 - \vec{k}_2^2} \\ &= -m \left(\frac{4\pi e^2}{\Omega q^2} \right)^2 \sum_{\vec{q}} \frac{1}{q^4} \frac{1}{q^2 + \vec{q} \cdot (\vec{k}_2 - \vec{k}_1)} \\ &= -m \left(\frac{4\pi e^2}{\Omega q^2} \right)^2 \frac{\Omega}{(2\pi)^3} 2\pi \int_0^\infty dq q^2 \int_{-1}^1 d\mu \frac{1}{q^4} \frac{1}{q^2} \frac{1}{1 + \mu \frac{|\vec{k}_1 - \vec{k}_2|}{q}} \\ &= -m \left(\frac{4\pi e^2}{\Omega q^2} \right)^2 \frac{\Omega}{(2\pi)^2} \frac{1}{|\vec{k}_1 - \vec{k}_2|} \int_0^\infty dq \frac{q}{q^4} \ln \left| \frac{q + |\vec{k}_1 - \vec{k}_2|}{q - |\vec{k}_1 - \vec{k}_2|} \right| \quad (14.98) \end{aligned}$$

Now the integrating function behaves as q^{-2} and therefore diverges as $q \rightarrow 0$. Each order in the perturbation expansion is infinite. The cause is the long range nature of the Coulomb potential ($q \rightarrow 0$ is problematic). Actually what happens is that the electronic system screens the interaction. The interaction between two electrons far apart is screened by the other electrons. The Coulomb potential is altered to a Yukawa like potential of the form

$$\frac{e^2}{r} \rightarrow \frac{e^2}{r} e^{-\frac{r}{l_s}}$$

where we have introduced a length l_s which is the screening length (we will see later that $l_s \sim v_F \sqrt{\frac{m}{n e^2}}$). Once again this result is non-analytical which can not be obtained in perturbation theory. We must use another method.

Let us consider a time dependent perturbation (recall the linear response theory earlier discussed). Consider then

$$\delta U(\vec{r}, t) = u e^{-i\vec{q}\cdot\vec{r} + i\omega t} e^{\alpha t}$$

Here $\alpha = 0^+$. The states are perturbed to

$$\psi_{\vec{k}} = |\vec{k}\rangle + b_{\vec{k}-\vec{q}}(t) |\vec{k} - \vec{q}\rangle$$

where

$$|\vec{k}\rangle = e^{i\vec{q}\cdot\vec{r} - i\epsilon(\vec{k})t/\hbar}$$

is a plane wave. The coefficients in the expansion on other states are given by

$$b_{\vec{k}-\vec{q}}(t) = \frac{\langle \vec{k} - \vec{q} | \delta U | \vec{k} \rangle}{\epsilon(\vec{k}) - \epsilon(\vec{k} - \vec{q}) + \hbar\omega - i\hbar\alpha} = \frac{u e^{i\omega t} e^{\alpha t}}{\epsilon(\vec{k}) - \epsilon(\vec{k} - \vec{q}) + \hbar\omega - i\hbar\alpha}$$

The change in the electronic density due to the perturbation is

$$\begin{aligned} \delta\rho(\vec{r}, t) &= e \frac{1}{\Omega} \sum_{\vec{k}} (|\psi_{\vec{k}}(\vec{r}, t)|^2 - 1) \\ &= e \frac{1}{\Omega} \sum_{\vec{k}} \left(b_{\vec{k}-\vec{q}}^*(t) e^{i\vec{q}\cdot\vec{r}} + e^{-i\vec{q}\cdot\vec{r}} b_{\vec{k}-\vec{q}}(t) \right) \\ &= e \frac{1}{\Omega} \sum_{\vec{k}} \left(\frac{u e^{-i\omega t} e^{\alpha t} e^{i\vec{q}\cdot\vec{r}}}{\epsilon(\vec{k}) - \epsilon(\vec{k} - \vec{q}) - \hbar\omega - i\hbar\alpha} + \frac{u e^{i\omega t} e^{\alpha t} e^{-i\vec{q}\cdot\vec{r}}}{\epsilon(\vec{k}) - \epsilon(\vec{k} - \vec{q}) - \hbar\omega + i\hbar\alpha} \right) \end{aligned} \quad (14.99)$$

Here Ω is the volume of the system.

If it is a real perturbation instead of a complex one then we have to sum the terms with $\omega \rightarrow \omega$ and $\vec{q} \rightarrow -\vec{q}$. It gives

$$\begin{aligned} \delta\rho(\vec{r}, t) &= e u \frac{1}{\Omega} \sum_{\vec{k}} \left(\frac{e^{-i\omega t} e^{\alpha t} e^{i\vec{q}\cdot\vec{r}}}{\epsilon(\vec{k}) - \epsilon(\vec{k} - \vec{q}) - \hbar\omega - i\hbar\alpha} + \frac{e^{i\omega t} e^{\alpha t} e^{-i\vec{q}\cdot\vec{r}}}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q}) + \hbar\omega - i\hbar\alpha} + c.c. \right) \\ &= e e^{\alpha t} u e^{i\omega t - i\vec{q}\cdot\vec{r}} \\ &\quad \frac{1}{\Omega} \sum_{\vec{k}, \text{occupied}} \left(\frac{1}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q}) + \hbar\omega - i\hbar\alpha} + \frac{1}{\epsilon(\vec{k}) - \epsilon(\vec{k} - \vec{q}) - \hbar\omega + i\hbar\alpha} \right) \\ &+ c.c. \end{aligned} \quad (14.100)$$

Adding the Fermi functions to give a thermodynamic weight to the occupied states we get that

$$\delta\rho(\vec{r}, t) = ee^{\alpha t} u \frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon(\vec{k})) - f(\epsilon(\vec{k} + \vec{q}))}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q}) + \hbar\omega - i\hbar\alpha} e^{i\omega t - i\vec{q} \cdot \vec{r}} + c.c. \quad (14.101)$$

Recall now that Poisson's equation implies that

$$\nabla^2(\delta\phi) = -4\pi e\delta\rho \quad (14.102)$$

Take

$$\delta\phi = \phi e^{i\omega t - i\vec{q} \cdot \vec{r}} e^{\alpha t}$$

Then we get that

$$\phi = u \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon(\vec{k})) - f(\epsilon(\vec{k} + \vec{q}))}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q}) + \hbar\omega - i\hbar\alpha} \quad (14.103)$$

We should now add ϕ to u in δU . In this way the calculation becomes self-consistent. We have that the external potential plus the polarization of the medium gives the total potential. The total charge density is the external charge term plus the induced charge. We really have

$$i\vec{q} \cdot \vec{D}(\vec{q}, \omega) = 4\pi\rho_{ext}$$

and that the induced charge is related to the polarization through

$$-i\vec{q} \cdot \vec{P} = \rho_{induced}$$

Defining the dielectric constant in the usual way we get that

$$\vec{D}(\vec{q}, \omega) = \vec{E}(\vec{q}, \omega) + 4\pi\vec{P}(\vec{q}, \omega) = \epsilon(\vec{q}, \omega)\vec{E}(\vec{q}, \omega)$$

Then writing the external potential as

$$\delta V = V e^{i\omega t - i\vec{q} \cdot \vec{r}} e^{\alpha t} + c.c.$$

we get that

$$u = \frac{V}{\epsilon(\vec{q}, \omega)}$$

and the dielectric constant is obtained through the self-consistent solution as

$$u = V + u \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon(\vec{k})) - f(\epsilon(\vec{k} + \vec{q}))}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q}) + \hbar\omega - i\hbar\alpha}$$

and the dielectric constant is given by

$$\epsilon(\vec{q}, \omega) = 1 - \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon(\vec{k})) - f(\epsilon(\vec{k} + \vec{q}))}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q}) + \hbar\omega - i\hbar\alpha} \quad (14.104)$$

This is the so-called Lindhard function.

14.7.1 Static properties

Consider the static uniform case. Take $\omega = 0$ and $\vec{q} \rightarrow 0$. In this limit

$$\epsilon(\vec{q} \rightarrow 0, \omega = 0) = 1 - \lim_{\vec{q} \rightarrow 0} \frac{1}{\Omega} \sum_{\vec{k}} \frac{4\pi e^2}{q^2} \frac{f(\epsilon(\vec{k})) - f(\epsilon(\vec{k} + \vec{q}))}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q}) + -i\hbar\alpha}$$

We can expand $\epsilon(\vec{k} + \vec{q}) = \epsilon(\vec{k}) + \vec{q} \cdot \vec{\nabla}_{\vec{k}} \epsilon(\vec{k})$. Also using that

$$\frac{1}{x - i\delta} = P \frac{1}{x} + i\pi\delta(x)$$

($\delta \rightarrow 0$) we conclude that the imaginary part is zero and in this case we can take $\alpha = 0$. Using that

$$f(\epsilon(\vec{k} + \vec{q})) = f(\epsilon(\vec{k})) + \vec{q} \cdot \frac{\partial f}{\partial \epsilon} \cdot \vec{\nabla}_{\vec{k}} \epsilon(\vec{k})$$

we obtain that the dielectric function is given by

$$\epsilon(\vec{q} \rightarrow 0, \omega = 0) = 1 - \lim_{\vec{q} \rightarrow 0} \frac{1}{\Omega} \sum_{\vec{k}} \frac{-\frac{\partial f}{\partial \epsilon} \vec{q} \cdot \vec{\nabla}_{\vec{k}} \epsilon(\vec{k})}{-\vec{q} \cdot \vec{\nabla}_{\vec{k}} \epsilon(\vec{k})} = \lim_{\vec{q} \rightarrow 0} \frac{1}{\Omega} \sum_{\vec{k}} \frac{\partial f}{\partial \epsilon_{\vec{k}}} \quad (14.105)$$

The sum over the momentum can be converted into an integral over the energies. Also at low enough temperatures $(\partial f)/(\partial \epsilon) \sim -\delta(\epsilon - \epsilon_F)$. We get then that

$$\epsilon(\vec{q} \rightarrow 0, \omega = 0) = 1 + \frac{4\pi e^2}{q^2} N(\epsilon_F) \quad (14.106)$$

where $N(\epsilon_F)$ is the density of states per volume at the Fermi level. The dielectric constant diverges at zero momentum.

Physically we can understand the form of the dielectric constant in a simple way. We can imagine the electrons as filling the Fermi sea at each site in the lattice. In equilibrium the chemical potential (Fermi level) is homogeneous. Imagine now that at a certain site \vec{r} we apply a perturbation that increases the energy locally. Then at that point the chemical potential is higher. There will be a flow of the electrons at that site to the neighboring sites to decrease the density at that site and try to homogenize the system. Since the electrons flow from that site there is an electronic deficiency at that point given by $\delta n(\vec{r}) = -N(\epsilon_F)\delta u(\vec{r})$, where $\delta u(\vec{r})$ is the increment in the energy at that site. Then again using Poisson's equation we get that

$$\nabla^2(\delta\phi) = -4\pi e^2 \delta n(\vec{r}) = 4\pi e^2 N(\epsilon_F) \delta u(\vec{r}) = \lambda^2 \delta u(\vec{r})$$

This leads to

$$-q^2 \delta\phi = \lambda^2 \delta u(\vec{r})$$

This then implies that

$$\epsilon = 1 + \frac{\lambda^2}{q^2}$$

where $\lambda^2 = 4\pi e^2 N(\epsilon_F)$. This is the so-called Thomas-Fermi approximation.

Consider now the regime of very high temperatures $T \gg T_F$. A classical treatment is therefore justified. Then

$$\delta n = n_0 e^{-\frac{\delta U(\vec{r})}{k_B T}} - n_0 = -n_0 \frac{\delta U(\vec{r})}{k_B T}$$

Using again Poisson's equation we obtain that

$$\nabla^2(\delta\phi) = \frac{4\pi e^2 n_0}{k_B T} \delta U(\vec{r})$$

which implies that in this limit

$$\lambda^2 = \frac{4\pi e^2 n_0}{k_B T}$$

This is the Debye-Huckel approximation.

The screening can be seen in the following way. Let us then consider a site with a positive charge. Then there is a change in the potential energy given by $\delta V = e^2/r$. The Fourier transform is $\delta V(\vec{q}) = 4\pi e^2/q^2$. The change in the potential in the medium is

$$\delta U(\vec{q}) = \frac{\delta V(\vec{q})}{\epsilon(\vec{q}, 0)} = \frac{4\pi e^2}{q^2} \frac{1}{1 + \frac{\lambda^2}{q^2}} = \frac{4\pi e^2}{q^2 + \lambda^2}$$

Inverting the Fourier transform back to real space we get that

$$\delta U(\vec{r}) = \frac{e^2}{r} e^{-\frac{r}{\lambda}} \quad (14.107)$$

There is therefore screening in the action of the charge far from its location and the screening length is λ .

Let us return to the Lindhardt function.

$$\begin{aligned} \epsilon(\vec{q}, \omega) &= 1 - \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon_{\vec{k}}) - f(\epsilon_{\vec{k}+\vec{q}})}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q}) + \hbar\omega - i\hbar\alpha} \\ \epsilon(\vec{q}, 0) &= 1 - \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon_{\vec{k}}) - f(\epsilon_{\vec{k}+\vec{q}})}{\epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{q})} \end{aligned}$$

Take the case of three dimensions and $\omega = 0$. The difference in the momentum summation can be simplified. Make a translation in the second term $\vec{k} \rightarrow \vec{k} - \vec{q}$ and assuming isotropy for free electrons we can write that $\epsilon(\vec{k} - \vec{q}) = \epsilon(\vec{k} + \vec{q})$. Then we can write that

$$\epsilon(\vec{q}, 0) = 1 - \frac{4\pi e^2}{q^2} 2 \frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon_{\vec{k}})}{\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}}}$$

The sum can be carried out

$$\begin{aligned} 2 \frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon_{\vec{k}})}{\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}}} &= \frac{4m}{(2\pi)^3} \int d^3k \frac{f(\epsilon_{\vec{k}})}{k^2 - (\vec{k} + \vec{q})^2} \\ &= -\frac{4m}{(2\pi)^3} 2\pi \int_0^{k_F} dk k^2 \int_{-1}^1 d\mu \frac{1}{2kq\mu + q^2} \end{aligned}$$

The integrations can be performed and yield for the dielectric constant using that $N(\epsilon_F) = (m)/(2\pi^2)$

$$\epsilon^{(3d)}(\vec{q}, 0) = 1 + \frac{4\pi e^2}{q^2} N(\epsilon_F) k_F \left(\frac{1}{2} + \frac{1}{2qk_F} (k_F^2 - \frac{q^2}{4}) \ln \left| \frac{2k_F + q}{2k_F - q} \right| \right) \quad (14.108)$$

Consider now the two-dimensional case. Here we mean a three dimensional system but arranged in layers. The dispersion relation of the free electrons is taken as

$$\epsilon_{\vec{k}} = \frac{1}{2m}(k_x^2 + k_y^2)$$

The vector \vec{q} is also contained in the $x - y$ plane. This is valid for $|k_z| < \frac{1}{2}K$ where K is some cutoff that arises since the spacing in the z direction is very small and therefore the quantization along the z direction is such that only very high momenta states are allowed. The motion along z is frozen.

In this case we can write that

$$\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}} = -\frac{1}{2m}(2\vec{k} \cdot \vec{q} + q^2)$$

Then we can write that

$$\begin{aligned} 2\frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon_{\vec{k}})}{\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}}} &= -\frac{4m}{(2\pi)^3} K \int_0^{k_F} dk k \int_0^{2\pi} d\theta \frac{1}{2kq \cos \theta + q^2} \\ &= -\frac{mK}{2\pi^3} \frac{1}{2q} \int_0^{k_F} dk \frac{2\pi}{\sqrt{\left(\frac{q}{2k}\right)^2 - 1}} \theta(q - 2k) \end{aligned}$$

Using that the density of states is $N(\epsilon) = \frac{mK}{(2\pi)^2}$ we can obtain that

$$\begin{aligned} \epsilon^{(2d)}(\vec{q}, 0) &= 1 + \frac{4\pi e^2}{q^2} N(\epsilon_F), q < 2k_F \\ &= 1 + \frac{4\pi e^2}{q^2} N(\epsilon_F) \left(1 - \frac{1}{q} \sqrt{q^2 - 4k_F^2}\right), q > 2k_F \end{aligned} \quad (14.109)$$

Finally let us consider the one-dimensional case. In this case

$$\epsilon_k = \frac{k_z^2}{2m}$$

Again we consider a system quasi-one-dimensional where $|k_x|, |k_y| < \frac{1}{2}K$. Like before we analyse the sum over the momenta

$$\begin{aligned} 2\frac{1}{\Omega} \sum_{\vec{k}} \frac{f(\epsilon_{\vec{k}})}{\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}}} &= 2\frac{1}{(2\pi)^3} K^2 \int_{-k_F}^{k_F} dk_z \frac{2m}{k_z^2 - (k_z + q)^2} \\ &= -4m \frac{K^2}{(2\pi)^3 2q} \ln \left| \frac{2k_F + q}{2k_F - q} \right| \end{aligned}$$

Using that the density of states at the Fermi level is

$$N(\epsilon_F) = \frac{2K^2 m}{(2\pi)^3 k_F}$$

we get that

$$\epsilon^{(1d)}(\vec{q}, 0) = 1 + \frac{4\pi e^2}{q^2} N(\epsilon_F) \frac{k_F}{q} \ln \left| \frac{2k_F + q}{2k_F - q} \right| \quad (14.110)$$

In the three cases of $1d$, $2d$ and $3d$ the dielectric constant can be written as

$$\begin{aligned} \epsilon(\vec{q}, 0) &= 1 + \frac{4\pi e^2}{q^2} \chi(q) \\ \chi(0) &= N(\epsilon_F) \\ \chi(q) &\geq 0 \end{aligned} \quad (14.111)$$

In the figure we show the susceptibility $\chi(q)$ for the three cases as a function of the momentum. The point $q = 2k_F$ is special. In $1d$ the susceptibility diverges at this point indicative of the instability previously mentioned. In $2d$ the susceptibility is continuous but its derivative is discontinuous. Only in three dimensions the behavior of the susceptibility is smooth.

It is easy to understand why $q = 2k_F$ is a special point. As a perturbation is applied there is an exchange of momentum q between the electrons. We may imagine that the perturbation shifts the Fermi sphere of two electrons by an amount which is the momentum transferred by the external perturbation. At small momentum many of the electrons do not contribute to the process because due to exclusion principle there is a large overlap of the Fermi surfaces and most states are occupied. As the momentum increases the volume that contributes increases and when the special point $q = 2k_F$ is reached all electrons contribute. After this point no qualitative change occurs and the susceptibility decreases since the momentum exchanged keeps increasing.

Let us then consider a phonon with momentum \vec{q} . The electrons follow the motion of the ions and therefore the interactions between the ions (positive charges immersed in the negative medium provided by the electrons) is screened by the electrons. As we have seen the screening is proportional to $1/\epsilon$. Therefore the forces between the ions are screened by the electrons to F/ϵ . The electronic medium changes the coupling constants between the ions. Therefore the phonon spectrum is changed

in an important way when $q = 2k_F$. This is the so-called Kohn effect. In the special case of $d = 1$ the dielectric constant diverges and therefore the effective force between the ions vanishes. The forces are completely screened when $q = 2k_F$. We expect therefore a strong instability in the lattice structure and the system may have a structural phase transition because the ions are no longer bound. The outcome is that the lattice becomes dimerized: this is the so-called Peierls instability. If we distort the crystal to form dimers (boundstates of two ions) the periodicity doubles, as discussed above, the lattice constant is now $2a$ this implies that the Brillouin zone halves the dispersion relation is changed and $q \neq 2k_F$. The force between the ions is restored since the screening is not fully effective anymore and this leads in general to the stabilization of the lattice. There is in general the so-called Peierls distortion (not necessarily dimerization).

The case of an insulator is slightly different. The treatment is also different from the one we followed for the metal because now we can not simply take the conduction electrons to be free. We want to consider the case when the insulator is originated in the band structure which implies taking into account the periodic potential (even though it is also possible to originate an insulator solely due to the effect of interactions leading to the so-called Mott insulator; we will return to this point later). Let us then consider that the gap arises due to the Bragg reflections (clearly the result depends now on the particular lattice structure). The dielectric constant is written as

$$\epsilon(\vec{q}, \omega) = 1 - \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_{\vec{k}, \vec{g}} \frac{| \langle \vec{k} | e^{i\vec{q} \cdot \vec{r}} | \vec{k} + \vec{q} + \vec{g} \rangle |^2 \left(f(\epsilon_{\vec{k}}) - f(\epsilon_{\vec{k} + \vec{q} + \vec{g}}) \right)}{\epsilon_{\vec{k}} - \epsilon_{\vec{k} + \vec{q} + \vec{g}} + \hbar\omega - i\hbar\alpha} \quad (14.112)$$

Consider now the static case $\omega = 0$. The term $\vec{g} = 0$ corresponds to two energies in the same band. Since we are considering an insulator either both bands are empty or both are full. Therefore it does not contribute. Among the reciprocal vector values $\vec{g} \neq 0$ the most relevant is the one such that

$$\epsilon_{\vec{k} + \vec{q} + \vec{g}} - \epsilon_{\vec{k}} \sim E_{gap}$$

Let us consider small momentum \vec{q} (the case of $q = 0$ does not contribute). The leading term is proportional to q^2 . This term cancels the same factor in the denominator and therefore the divergence disappears. We get that

$$\epsilon(\vec{q} \rightarrow 0, 0) = 1 + \frac{4\pi n e^2 \hbar^2}{m(E_{gap})^2} \quad (14.113)$$

It is finite: an insulator can not screen completely. Note that the dielectric constant can be written as

$$\epsilon(\vec{q} \rightarrow 0, 0) = 1 + \left(\frac{\hbar\omega_p}{E_{gap}} \right)^2 \quad (14.114)$$

Here

$$\omega_p = \sqrt{\frac{4\pi n e^2}{m}}$$

is called the plasma frequency for reasons that will be clear in a moment.

14.7.2 Dynamic properties

Consider the momentum and frequency dependent dielectric constant. In general it has a real and an imaginary part

$$\epsilon(q, \omega) = \epsilon_1(q, \omega) + i\epsilon_2(q, \omega)$$

Using the expression obtained before for the dielectric constant and using that

$$\frac{1}{\epsilon_k - \epsilon_{k+q} + \hbar\omega - i\hbar\alpha} = P \frac{1}{\epsilon_k - \epsilon_{k+q} + \hbar\omega} + i\pi\delta(\epsilon_k - \epsilon_{k+q} + \hbar\omega)$$

we obtain that the real part is given at $T = 0$ by

$$\begin{aligned} \epsilon_1(q, \omega) = & 1 + \frac{q_{TF}^2}{q^2} \left\{ \frac{1}{2} + \frac{k_F}{4q} \left[\left(\frac{(\omega + \frac{q^2}{2m})^2}{(qv_F)^2} - 1 \right) \ln \left| \frac{\omega - v_F q + \frac{q^2}{2m}}{\omega + v_F q + \frac{q^2}{2m}} \right| \right. \right. \\ & \left. \left. - \left(\frac{(\omega + \frac{q^2}{2m})^2}{(qv_F)^2} - 1 \right) \ln \left| \frac{\omega - v_F q + \frac{q^2}{2m}}{\omega + v_F q + \frac{q^2}{2m}} \right| \right] \right\} \end{aligned} \quad (14.115)$$

where the Thomas-Fermi momentum is given by

$$q_{TF}^2 = \frac{3\omega_p^2}{v_F^2} = 3 \frac{4\pi n e^2}{m v_F^2}$$

The imaginary part is given by

$$\begin{aligned}
 \epsilon_2(q, \omega) &= \frac{\pi}{2} \frac{q_{TF}^2}{q^2} \frac{\omega}{qv_F}, 0 \leq v_F q - \frac{q^2}{2m} \\
 &= \frac{\pi}{4} \frac{q_{TF}^2}{q^2} \frac{k_F}{q} \left(1 - \frac{(\omega - \frac{q^2}{2m})^2}{(qv_F)^2} \right), qv_F - \frac{q^2}{2m} \leq \omega \leq qv_F + \frac{q^2}{2m} \\
 &= 0, \omega \geq qv_F + \frac{q^2}{2m}
 \end{aligned} \tag{14.116}$$

Note however that this is a continuous function of the frequency.

In the limit $q \rightarrow 0$ the real part of the dielectric constant diverges. For small values of the momentum q we show in the figure the real and imaginary parts of the dielectric constant as a function of the frequency ($q \ll q_{TF}, k_F$). The imaginary part vanishes at zero frequency as we saw before. The real part is finite in the static case then it decreases goes through zero becomes negative and at high frequencies it becomes positive again crossing the real axis at the plasma frequency ω_p . When the real part becomes negative the imaginary part is finite. Then the imaginary part decreases to zero and at the plasma frequency both the real and the imaginary parts vanish. Therefore the dielectric constant vanishes at the plasma frequency. Therefore the electric field is finite even if the applied external field is zero. There is an excitation of the system which are called plasmons where an internal electric field is spontaneously created by any infinitesimal fluctuation in the system.

Since we are in a regime where the imaginary part is zero we can take $\alpha = 0$ in the expression for the dielectric constant. We obtain then that

$$\epsilon(q, \omega) = 1 - \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_k f(\epsilon_k) \left(\frac{1}{\epsilon_k - \epsilon_{k+q} + \hbar\omega} + \frac{1}{\epsilon_k - \epsilon_{k-q} - \hbar\omega} \right)$$

Let us consider small q and large ω . Then

$$\epsilon(q, \omega) \sim 1 - \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_k f(\epsilon_k) (\epsilon_{k+q} - \epsilon_k + \epsilon_{k-q} - \epsilon_k) \sim 1 - \frac{4\pi e^2}{q^2} \frac{q^2}{m} \frac{1}{\Omega} \sum_k f(\epsilon_k)$$

We finally get that

$$\epsilon(q, \omega) \sim 1 - \frac{\omega_p^2}{\omega^2} \tag{14.117}$$

where ω_p is the plasma frequency. Clearly when $\omega = \omega_p$ we obtain that $\epsilon(q, \omega) = 0$.

The meaning of the plasma frequency is easy to understand physically. Define the dipolar moment in the usual way

$$\vec{P} = ne\vec{x}$$

If there are no external fields the displacement vector vanishes

$$\vec{D} = \vec{E} + 4\pi\vec{P} = 0$$

So we get that

$$\vec{E} = -4\pi\vec{P} = -4\pi ne\vec{x}$$

The equation of motion of the electrons is then given by

$$m \frac{d^2 \vec{x}}{dt^2} = e\vec{E} = -4\pi ne^2 \vec{x}$$

The eigenmodes have a frequency which is the plasma frequency, ω_p . This is an energy of $10 - 20 eV$. It is therefore a very high energy to be excited by thermal excitation. These mode of oscillation of the charges are therefore usually frozen (the Coulomb forces are very strong).

Let us take again the dielectric constant given by

$$\begin{aligned} \epsilon(q, \omega) &= 1 - \frac{4\pi e^2}{q^2} \frac{1}{\Omega} \sum_k f(\epsilon_k) \frac{2(\epsilon_k - \epsilon_{k+q})}{(\epsilon_k - \epsilon_{k+q})^2 - \omega^2} \\ &= 1 - \frac{\omega_p^2}{\omega^2} + \frac{4\pi e^2}{q^2 \omega^4} \frac{1}{\Omega} \sum_k f(\epsilon_k) 2(\epsilon_k - \epsilon_{k+q})^3 + O\left(\frac{1}{\omega^6}\right) \end{aligned}$$

Considering small momenta converting the sum over the momenta as an integral and performing the integral we can obtain that

$$\epsilon(q, \omega) = 1 - \frac{\omega_p^2}{\omega^2} \left(1 + \frac{1}{\omega^2} \frac{3}{5} (qv_F)^2 \right)$$

The zeros of the dielectric constant are obtained from the equation $\epsilon(q, \omega_p(q)) = 0$ and are given by

$$\omega_p^4(q) - \omega_p^2(q) \omega_p^2 - \frac{3}{5} \omega_p^2 (qv_F)^2 = 0$$

The solutions are

$$\omega^2 = \frac{\omega_p^2}{2} \left(1 + \sqrt{1 + \frac{12}{5} \frac{(qv_F)^2}{\omega_p^2}} \right) \quad (14.118)$$

Expanding we get that for small momenta

$$\omega = \omega_p + \frac{3}{10} \frac{(qv_F)^2}{\omega_p}$$

At zero momentum the frequency is just the plasma frequency. As the momentum increases the eigenmodes increase quadratically with the momentum. These excitations are called plasmons. They are well defined very high energy excitations of the system. They are stable for small momentum values. However there are other excitations in the system since there are particle-hole excitations that are allowed. Since these are two-particle excitations they form a continuum. As shown in the figure they have very low energy at small momentum and are therefore favorable energetically while the plasmons remain frozen. When the dispersion relation of the plasmons meets the Stoner continuum of excitations the plasmons may decay into particle-hole excitations and are no longer stable excitations.

To finish this section let us summarize some of the results we reached. The susceptibility that we have defined is actually the change in the charge density caused by the external perturbation. Therefore

$$\delta\rho = -e\chi(q, \omega)\delta u = e \frac{1}{\Omega} \sum_k \frac{f(\epsilon_k) - f(\epsilon_{k+q})}{\epsilon_k - \epsilon_{k+q} + \omega - i\alpha} \delta u$$

If we do not consider the self-consistent nature of the problem we get that

$$\delta\rho_{H-F} = -e\chi(q, \omega)\delta V_{ext}$$

and we do not get screening. If however we consider the perturbation in a self-consistent way this leads to the so-called Random Phase Approximation like

$$\delta\rho_{RPA} = -e \frac{\chi(q, \omega)}{1 + \frac{4\pi e^2}{q^2} \chi(q, \omega)} \delta V_{ext}$$

In this way we get the proper screening property of the system of charges.

15.6 ϕ^4 theory

15.6.1 Representation of the Ising model by functional integrals

Let us recall the gaussian integral

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}}$$

If we make a translation $x \rightarrow x - b$ we obtain the integral

$$\int_{-\infty}^{\infty} dx e^{-\frac{1}{4a^2}x^2 + sx} = \sqrt{4\pi a^2} e^{a^2 s^2}$$

This expression can be generalized

$$\int_{-\infty}^{\infty} \prod_{i=1}^N dx_i e^{-\frac{1}{4}x_i V_{ij}^{-1}x_j + S_i x_i} = C e^{S_i V_{ij} S_j} \quad (15.46)$$

where C is a constant.

Let us consider again the Ising model. The Hamiltonian is of the type

$$H = - \sum_{ij} J_{ij} S_i S_j - \sum_i h_i S_i$$

The partition function is

$$\begin{aligned} Z\{H_i\} &= \sum_{\{S_i\}} e^{-\beta H\{S_i\}} \\ &= \sum_{\{S_i\}} e^{\sum_{ij} K_{ij} S_i S_j + \sum_i H_i S_i} \\ &= \sum_{\{S_i\}} \int_{-\infty}^{\infty} \prod_{i=1}^N d\phi_i e^{-\frac{1}{2}\phi_i K_{ij}^{-1}\phi_j + (\phi_i + H_i)S_i} \\ &= \int_{-\infty}^{\infty} \prod_{i=1}^N d\phi_i e^{-\frac{1}{4}(\phi_i - H_i)K_{ij}^{-1}(\phi_j - H_j)} \sum_{\{S_i\}} e^{\phi_i S_i} \end{aligned} \quad (15.47)$$

The final sum is easy to carry out:

$$\sum_{\{S_i\}} e^{\phi_i S_i} = \prod_i (2 \cosh \phi_i) = C e^{\sum_i \ln(\cosh \phi_i)}$$

Therefore the partition function can be rewritten as

$$Z\{H_i\} = C \int_{-\infty}^{\infty} \prod_{i=1}^N d\phi_i e^{-\frac{1}{4}(\phi_i - H_i) K_{ij}^{-1} (\phi_j - H_j)} e^{\sum_i \ln(\cosh \phi_i)} \quad (15.48)$$

Defining a new field like

$$\psi_i = \frac{1}{2} K_{ij}^{-1} \phi_j$$

we obtain that

$$Z\{H_i\} \sim e^{-\frac{1}{4} H_i K_{ij}^{-1} H_j} \int \mathcal{D}\psi e^{-\psi_i K_{ij} \psi_j + H_i \psi_i} e^{\sum_i \ln(\cosh(2K_{ij} \psi_j))} \quad (15.49)$$

Here we have written that $\mathcal{D}\psi = \prod_{i=1}^N d\psi_i$. Expanding

$$\ln \cosh x = \frac{1}{2} x^2 - \frac{1}{12} x^4 + \dots$$

we get that the partition function takes the form

$$Z\{H_i\} \sim e^{-\frac{1}{4} H_i K_{ij}^{-1} H_j} \int \mathcal{D}\psi e^{-\psi_i K_{ij} \psi_j + H_i \psi_i} e^{2 \sum_i (K_{ij} \psi_j)^2 - \frac{16}{12} \sum_i (K_{ij} \psi_j)^4 + \dots} \quad (15.50)$$

We have replaced the partition function as a sum over spin configurations by a functional integral over bosonic fields with an effective action that is given by an expansion in powers of this bosonic field.

15.6.2 Landau's mean-field theory

The partition function is then of the type

$$Z \sim \int \mathcal{D}\phi e^{-\frac{H(\phi)}{T}}$$

Let us then consider that in general

$$\frac{H(\phi)}{T} = \int d^d x [a_0 - B(x)\phi(x) + R(T)\phi^2(x) + u(T)\phi^4(x) + S(T) \left(\vec{\nabla}\phi(x) \right)^2 + \dots] \quad (15.51)$$

where we have added a gradient term to allow for non-homogeneous field configurations. To obtain the thermodynamics we have to calculate the partition function. In general this is very complicated. If $u(T) = 0$ and all higher order terms vanish as well we may solve exactly the problem: this is the gaussian model. However this is not in good agreement with the experimental results. We may therefore develop several approximation methods. In this case we are interested in looking for solutions that correspond to values of $\langle \phi \rangle \neq 0$ that is to phase transitions.

The term $-B(x)\phi(x)$ corresponds to the action of an external field (if $\phi = M$, B is the magnetic field). Since the system has to be invariant if we make the transformation $\phi \rightarrow -\phi$ and at the same time $B \rightarrow -B$ there can be no odd terms like ϕ^3, ϕ^5, \dots . The meaning of $R(T)$ will become clear momentarily. The coefficient $u(T)$ is the amplitude of the interaction part. Finally $S(T)$ is the coefficient that results of the possible spatial variation of the field. We assume that the variations are slow and therefore we neglect higher order derivatives. In the case of uniform solutions this term vanishes. If a uniform solution is preferred physically then the presence of this term means that it costs energy to introduce a spatial modulation if $S(T) > 0$. We will take $S(T) = 1$.

To calculate the partition function let us expand the action around its minimum (actually saddle-point)

$$H(\phi) = H(\bar{\phi}) + \frac{1}{2} \left(\frac{\partial^2 H}{\partial \phi^2} \right)_{\bar{\phi}} (\phi - \bar{\phi})^2 + \frac{1}{3!} \left(\frac{\partial^3 H}{\partial \phi^3} \right)_{\bar{\phi}} (\phi - \bar{\phi})^3 + \dots$$

where

$$\left(\frac{\partial H}{\partial \phi} \right)_{\bar{\phi}} = 0 \quad (15.52)$$

Then we obtain that

$$\begin{aligned}
Z &\sim e^{-\frac{H(\bar{\phi})}{T}} \int \mathcal{D}(\phi - \bar{\phi}) e^{-\frac{1}{T} \frac{1}{2} \left(\frac{\partial^2 H}{\partial \phi^2} \right)_{\bar{\phi}} (\phi - \bar{\phi})^2 - \frac{1}{T} \frac{1}{3!} \left(\frac{\partial^3 H}{\partial \phi^3} \right)_{\bar{\phi}} (\phi - \bar{\phi})^3 - \frac{1}{T} \frac{1}{4!} \left(\frac{\partial^4 H}{\partial \phi^4} \right)_{\bar{\phi}} (\phi - \bar{\phi})^4 + \dots} \\
&\sim e^{-\frac{H(\bar{\phi})}{T}} \int \mathcal{D}x e^{-\frac{1}{2T} H'' x^2} \left[1 - \frac{x^3}{3!T} H''' + \frac{1}{2!} \left(\frac{x^3}{3!T} H''' \right)^2 + \dots \right] \\
&\quad \left[1 - \frac{1}{4!T} x^4 H^{iv} + \frac{1}{2!} \left(\frac{x^4}{4!T} H^{iv} \right)^2 + \dots \right] [1 + \dots]
\end{aligned} \tag{15.53}$$

The leading order constitutes the mean-field approximation. At this order we get that

$$Z^{(0)} \sim e^{-\frac{H(\bar{\phi})}{T}} \tag{15.54}$$

The next order includes also the gaussian fluctuations around the saddle-point solution. It is given by

$$Z^{(1)} \sim e^{-\frac{H(\bar{\phi})}{T}} \int \mathcal{D}x e^{-\frac{1}{2T} H'' x^2} \tag{15.55}$$

For the moment let us focus our attention on the mean-field approximation. The equation of state is written as

$$\int d^3x [-B(x) + 2R(T)\bar{\phi}(x) + 4u(T)\bar{\phi}^3(x) - 2\nabla^2\bar{\phi}(x)]\delta\phi = 0 \tag{15.56}$$

where we have varied the action with respect to a variation of the field and where we have integrated by parts the gradient term like

$$\delta \left(\vec{\nabla} \phi \cdot \vec{\nabla} \phi \right) = 2\vec{\nabla} \phi \cdot \delta(\vec{\nabla} \phi) = -2\nabla^2 \phi \delta \phi$$

The equation of state reduces therefore to

$$2R(T)\bar{\phi}(x) + 4u(T)\bar{\phi}^3(x) - 2\nabla^2\bar{\phi}(x) = B(x) \tag{15.57}$$

Let us take, for simplicity, a uniform field $B(x) = B$. Then we expect that the field is uniform $\bar{\phi}(x) = \bar{\phi}$. The equation of state reduces in this case to

$$2R(T)\bar{\phi} + 4u(T)\bar{\phi}^3 = B$$

If $B = 0$ there are two solutions. The trivial one $\bar{\phi} = 0$ and a second solution

$$\bar{\phi} = \pm \left(\frac{-R(T)}{2u(T)} \right)^{1/2}$$

Clearly this implies that this second solution only exists if $R(T) < 0$. If $T > T_c$ the trivial solution is the chosen one. Therefore we expect that we can write $R(T) = a(T - T_c)$. If $T < T_c$ the coefficient $R(T) < 0$ and the two solutions are possible. If $T > T_c$ the coefficient $R(T) > 0$ and only the trivial solution is possible. We have to determine when $T < T_c$ which of the two solutions is the most stable. It is easy to see that the non-trivial solutions are the most stable. Close to T_c the coefficient is a constant $u(T_c)$. Therefore close to the critical temperature we have that

$$\bar{\phi} \sim (T_c - T)^{1/2} \quad (15.58)$$

defining a critical exponent β as

$$\bar{\phi} \sim (T_c - T)^\beta$$

we see that in mean field this exponent is $\beta = 1/2$. At the critical point $R(T_c) = 0$ we get from the equation of state that $4u(T_c)\bar{\phi}^3 = B$. Defining another critical exponent like

$$\bar{\phi} \sim B^{1/\delta}$$

we get in mean field theory that $\delta = 3$.

We can also study the behavior of the susceptibility around the critical point. The susceptibility is defined by $\chi_T = (\partial\bar{\phi})/(\partial B)_T$. Using the equation of state we get that

$$[2R(T) + 12U(T)\bar{\phi}^2] \left(\frac{\partial\bar{\phi}}{\partial B} \right)_T = 1$$

Considering $T > T_c$ ($\bar{\phi} = 0$) we get that

$$\chi_T = (\partial\bar{\phi})/(\partial B)_T = \frac{1}{2R(T)} \sim \frac{1}{T - T_c} \sim (T - T_c)^{-\gamma} \quad (15.59)$$

which means that the critical exponent $\gamma = 1$ in mean field. If we consider a temperature below the critical temperature $T < T_c$ and using the non trivial solution for $\bar{\phi}$ we get the same exponent.

It turns out however that the exponents observed experimentally are not in agreement with the mean field predictions. The reason is that the fluctuations that are neglected in mean field theory are important and not negligible near the critical point. A proper treatment of the fluctuations is necessary.

15.6.3 Ginzburg criterion

The functional is then of the type

$$\frac{1}{2}|\nabla\phi|^2 + \frac{1}{2}a\phi^2 + \frac{b}{4!}\phi^4 + \dots$$

where $a \sim (T - T_c)$. Varying with respect to ϕ and taking a uniform solution we get as before $a\phi_0 + b/6\phi_0^3 = 0$. The solution is $\phi_0^2 = -6a/b$. The fluctuations are given by

$$\langle \phi(x)\phi(x') \rangle = \int \frac{d^d k}{(2\pi)^d} \frac{e^{ik(x-x')}}{k^2 + a}$$

The local field fluctuation is given by

$$\langle \phi^2(x) \rangle = \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 + a} \sim a^{\frac{d-2}{2}} \sim (T - T_c)^{\frac{d-2}{2}} \quad (15.60)$$

Then the relative fluctuation is given by

$$\frac{\langle \phi^2 \rangle}{\langle \phi \rangle^2} \sim \frac{(T - T_c)^{\frac{d}{2}-1}}{(T - T_c)} \sim (T - T_c)^{\frac{d}{2}-2} \sim (T - T_c)^{\frac{d-4}{2}} \quad (15.61)$$

If the space dimension is $d > d_c$ where $d_c = 4$ the fluctuations are not important. If on the other hand $d < d_c$ the fluctuations can not be neglected. If $T - T_c$ is not too small then the quotient is finite. Only when $T \rightarrow T_c$ there are problems. The Ginzburg criterion establishes therefore a region around the critical point where fluctuations are important if $d < d_c$.

15.8 Mermin-Wagner theorem

Let us consider once again the classical Heisenberg model

$$H = -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j$$

and let us consider in general the spin vector as having n components like

$$\vec{S} = (S_1, S_2, \dots, S_n)$$

with $|\vec{S}| = 1$. As before for $n = 1$ we have the Ising model, for $n = 2$ the $X - Y$ model, for $n = 3$ the Heisenberg model and for $n = \infty$ the spherical model. The partition function is obtained integrating

$$Z = \int d^N \Omega e^{-\frac{H}{k_B T}}$$

Consider now the continuum limit and rewrite the partition function as

$$Z = \int D\vec{S} e^{-\int d^d x [(\nabla \vec{S})^2 + u S^2 + v (S^2)^2]}$$

Note that $(\nabla \vec{S})^2 = (\vec{\nabla} S_1)^2 + \dots + (\vec{\nabla} S_n)^2$. For $n = 2, 3, \dots$ we have excitations of the spin-wave type in the low temperature phase where $|\vec{S}| = \text{const}$ but $\vec{\nabla} \vec{S} \neq 0$. The Hamiltonian for the spin-waves (this will be treated again later on in detail) is

$$\begin{aligned} H_{\text{spin-waves}} &= J' \int d^d x (\nabla S)^2 \\ &= J'' \sum_{\vec{q}} q^2 |S_{\vec{q}}|^2 \end{aligned} \quad (15.71)$$

Using the Equipartition Theorem we can establish that

$$J'' q^2 \langle |S_{\vec{q}}|^2 \rangle = \frac{1}{2}$$

Therefore

$$\langle (\delta \vec{S}(\vec{r}))^2 \rangle \sim \int d^d q \langle |S_{\vec{q}}|^2 \rangle \sim \int_0^\Lambda dq q^{d-1} \frac{1}{q^2}$$

This integral diverges when $q \sim 0$ if $d \leq 2$. If $d > 4$ the mean field solution should work well. This is the upper critical dimension for the phase transition to an ordered phase. However if $d \leq 2$ the fluctuations should destroy the ordered phase since the fluctuations diverge, by the simple above argument. Therefore for $d \leq 2$ there should not be a phase transition. If $n > 2$ then $d = 2$ is the lower critical dimension. But as we will see later if $n = 2$ and $d = 2$ there is a transition of another type in which there is no spontaneous magnetization in the low temperature phase but there is long range order of topological defects (vortices). Let us now show that the above argument is correct and that indeed $d = 2$ is the lower critical dimension: that is for dimensions equal or lower there is no phase transition at finite temperature. This is the essence of the so-called Mermin-Wagner which we will next prove.

Let us first consider the so-called Bogoliubov inequality.

$$\frac{1}{2} \langle \{A, A^\dagger\} \rangle \langle [C, H], C^\dagger \rangle \geq k_B T \langle [C, A] \rangle^2$$

Let us prove this inequality. First define

$$(A, B) = \sum_{i,j} \langle i|A|j \rangle^* \langle i|B|j \rangle \frac{W_i - W_j}{E_j - E_i}$$

Here

$$W_i = \frac{e^{-\beta E_i}}{\text{Tr} e^{-\beta H}}$$

It is easy to show that

$$\frac{W_i - W_j}{E_j - E_i} > 0$$

Also

$$\begin{aligned} (A, A) &\leq \frac{1}{2} \beta \langle \{A, A^\dagger\} \rangle \\ (A, A)(B, B) &\geq |(A, B)|^2 \end{aligned} \tag{15.72}$$

Choosing $B = [C^\dagger, H]$ then

$$\begin{aligned} (A, B) &= \langle [C^\dagger, A^\dagger] \rangle \\ (B, B) &= \langle [C^\dagger, [H, C]] \rangle \end{aligned} \tag{15.73}$$

This shows the inequality.

Let us now apply it to the Heisenberg model

$$H = - \sum_{\vec{R}, \vec{R}'} J(\vec{R} - \vec{R}') \vec{S}(\vec{R}) \cdot \vec{S}(\vec{R}') - h \sum_{\vec{R}} S_z(\vec{R}) e^{-i\vec{q} \cdot \vec{R}}$$

Here we are taking that $J(0) = 0$, $J(\vec{R}) = J(-\vec{R})$. Also the interaction j has a finite range so that

$$\sum_{\vec{R}} R^2 |J(\vec{R})|$$

is convergent. Let us choose

$$C = S_+(\vec{k}), A = S_-(-\vec{k} - \vec{q})$$

Then the Bogoliubov inequality leads to

$$\begin{aligned} & \langle \{S_+(\vec{k} + \vec{q}), S_-(-\vec{k} - \vec{q})\} \rangle \leq N^2 k_B T S_z^2 \\ & \frac{1}{\frac{1}{N} \sum_{\vec{k}'} [J(\vec{k}' - \vec{q}) - J(\vec{k}' - \vec{q})] \langle S_z(-\vec{k}') S_z(\vec{k}') \rangle + \frac{1}{4} \{S_+(\vec{k}'), S_-(-\vec{k}')\} \rangle} > + \frac{N}{2} h S_z \end{aligned} \quad (15.74)$$

where

$$S_z = \frac{1}{N} \sum_{\vec{R}} \langle S_z(\vec{R}) e^{i\vec{q} \cdot \vec{R}} \rangle$$

Using that

$$\sum_{\vec{k}'} \langle S_i(-\vec{k}') S_i(\vec{k}') \rangle = N \sum_{\vec{R}} \langle S_i(\vec{R}) S_i(\vec{R}) \rangle = N^2 \langle S_i(\vec{R}_0) S_i(\vec{R}_0) \rangle$$

and summing equation (A) over \vec{q} we obtain that

$$S(S+1) > 2k_B T S_z^2 \frac{1}{N} \sum_{\vec{k}} \frac{1}{h S_z + S(S+1) \sum_{\vec{R}} R^2 J(\vec{R}) k^2} \quad (15.75)$$

When the volume of the system goes to infinity we get that the sum over the momenta is replaced by an integral like

$$S(S+1) > \frac{2k_B T S_z^2}{\rho} \int_{1st\,zone} \frac{d^d k}{(2\pi)^d} \frac{1}{S(S+1) \sum_{\vec{R}} R^2 J(\vec{R}) k^2 + h S_z} \quad (15.76)$$

Here $1/\rho$ is the volume per spin. Denoting $\bar{\omega} = S(S+1) \sum_{\vec{R}} R^2 J(\vec{R})$ we can write that

$$S_z^2 < \frac{\rho S(S+1)}{2k_B T} \frac{1}{\int_{1st zone} \frac{d^d k}{(2\pi)^d} \frac{1}{\bar{\omega} k^2 + h S_z}} \quad (15.77)$$

The first Brillouin zone can be taken as a sphere of radius k_0 . Using that

$$\begin{aligned} \frac{1}{\pi} \int_0^{k_0} dk \frac{1}{k^2 + a} &= \frac{1}{\pi} \frac{1}{\sqrt{a}} \arctan \frac{k_0}{\sqrt{a}}, d = 1 \\ \frac{1}{2\pi} \int_0^{k_0} dk \frac{k}{k^2 + a} &= \frac{1}{4\pi} \ln \frac{k_0^2 + a}{a}, d = 2 \\ \frac{1}{2\pi^2} \int_0^{k_0} dk \frac{k^2}{k^2 + a} &= \frac{1}{2\pi^2} \{k_0 - \sqrt{a} \arctan \frac{k_0}{\sqrt{a}}\}, d = 3 \end{aligned} \quad (15.78)$$

We have therefore that in the limit that $h \rightarrow 0$

$$\begin{aligned} S_z &< \text{const} \frac{h^{1/3}}{T^{2/3}}, d = 1 \\ S_z &< \frac{\text{const}}{T^{1/2}} \frac{1}{|\ln h|^{1/2}}, d = 2 \\ S_z^2 &< \frac{\rho S(S+1)}{2k_B T} \frac{2\pi^2 \bar{\omega}}{k_0}, d = 3 \end{aligned} \quad (15.79)$$

In the first two cases $S_z \rightarrow 0$ as $h \rightarrow 0$. In the case of $d = 3$ the magnetization is independent of h and nothing is concluded. The conclusion proves the theorem: in $d = 1, 2$ there is no spontaneous magnetization.

Capítulo 16

Magnetism

Spin systems are described in terms of spin operators which have commutation relations that are neither bosons nor fermions. However as we will now see in some circumstances their excitations are boson like and later on we will see that for instance in one dimension their spectrum has a fermion like dispersion (or more precisely they are hard core bosons). One of the reasons for this duality is that one can represent spin operators by bilinear representations in terms of either bosons or fermions. We will not elaborate at this point on the various possible representations of the spin operators but will throughout this book refer to different representations, as needed.

At low temperatures spin systems (meaning here localized spins interacting by some exchange overlap whose origin is not specified but that is eventually originated and in the electron orbitals that may overlap between the atoms where the spins are defined or else via the extended conduction electron orbitals, an example of which we will see later) may order magnetically in different manners. We will here focus only on the ferromagnetic and the antiferromagnetic orders since these are the most commonly treated ones.

16.1 Ferromagnetism: mean-field theory

A simple model that describes exchange interactions between localized magnetic momenta is the Heisenberg model that can be written as

$$H = \frac{1}{2}J \sum_{i,j} \vec{S}_i \cdot \vec{S}_j \quad (16.1)$$

The sum extends over nearest-neighbor spins located on a discrete lattice. The coupling between the spins is parametrized by the exchange interaction J . This model describes magnetism in an insulating system. The description of magnetism in a conducting system is studied later.

To simplify we begin by considering a simplified model such that the interaction is limited to one spin component such as the z -th component

$$H = \frac{1}{2}J \sum_{i,j} S_i^z S_j^z - g\mu_B \sum_i S_i^z \quad (16.2)$$

where we have added the effect of an external magnetic field, B , along the z -th direction. Here, g is the Landé factor and μ_B is the Bohr magneton. This model can be treated in a mean-field approximation where fluctuations are neglected. Specifically, like before we rewrite the exchange part of the Hamiltonian as

$$\frac{1}{2}J \sum_{i,j} [(S_i^z - \langle S_i^z \rangle)(S_j^z - \langle S_j^z \rangle) + \langle S_j^z \rangle S_i^z + \langle S_i^z \rangle S_j^z - \langle S_i^z \rangle \langle S_j^z \rangle] \quad (16.3)$$

In the mean-field approximation the first term is neglected since it is proportional to the square of the fluctuations around the mean value of the spin at a given point in space. Within this approximation the effect of the interactions between the spins can be understood as an effective magnetic field on each spin. In addition to the external magnetic field there is an exchange term that is the result of the neighboring spins. The Hamiltonian may then be written as

$$H = - \left(g\mu_B B - J \sum_{j=1}^z \langle S_j^z \rangle \right) \sum_i S_i^z - \frac{J}{2} z \sum_i \langle S_i^z \rangle^2 \quad (16.4)$$

where z is the coordination number (number of first neighbors). Equivalently we may write

$$H = -g\mu_B B_{eff} \sum_i S_i^z + C \quad (16.5)$$

where C is a constant. Note that it is straightforward to generalize these expressions to a rotationally symmetric expression.

Consider now, to simplify, that the system is composed of spins $S = 1/2$. Recalling the results obtained for a paramagnetic system in the presence of an external field the magnetization can be written as

$$N = N\mu \tanh\left(\frac{\mu B_{eff}}{k_B T}\right) \quad (16.6)$$

N is the number of spins, $\mu = g\mu_B S$. Consider now that the external field vanishes, $B = 0$. The magnetization satisfies the equation of state

$$M = N\mu \tanh\left(\frac{\mu\lambda M}{k_B T}\right) \quad (16.7)$$

Here $\lambda = B_{eff}/M$. Considering a ferromagnetic system, for which $J = -|J|$ and using the translational invariance of the system, $M_j = M/N$, where M_j is the magnetization at site j , we see that

$$B_{eff} = \frac{|J|z}{g\mu_B} \frac{M}{N}, \lambda = \frac{|J|z}{g\mu_B N} \quad (16.8)$$

The equation of state has a trivial solution, $M = 0$, that corresponds to the paramagnetic phase. There is also a non-trivial solution with finite M . At high temperatures the only solution is the trivial one since the thermal energy completely misaligns the spins. At zero temperature the spins are all parallel and the magnetization is finite (and maximal). Between the two extreme regimes there is a transition from zero magnetization and finite magnetization. This may occur as a crossover or as a transition. It turns out that there is a transition at a finite temperature that is called critical temperature, T_c . Let us assume that the transition is continuous in the sense that the magnetization raises from zero continuously. Therefore near the critical temperature it is small.

We may linearize the equation of state

$$M = N\mu \frac{\mu\lambda M}{k_B T_c} = \frac{N\mu^2\lambda}{k_B T_c} M \quad (16.9)$$

Within mean-field the critical temperature is given by

$$k_B T_c = N\mu^2\lambda = g\mu_B z S^2 |J| \quad (16.10)$$

Define $m = M/(N\mu)$. This leads to

$$m = \tanh\left(\frac{T_c}{T}m\right) \quad (16.11)$$

This equation may be solved by the graphical method. Since the slope at the origin of the \tanh dependence is T_c/T , if $T > T_c$ there is only one solution, $m = 0$. However, if $T_c < T$ there are two solutions. One where $m = 0$ and one for which m is finite, showing that there is a non-trivial phase where there is a spontaneous magnetization that results from the exchange effective field due to the neighboring spins.

In the paramagnetic regime, $T > T_c$, and restoring the external field, B , we can write that

$$M = \chi_p (B + \lambda M) \quad (16.12)$$

where $\chi_p = C/T$ is the standard Curie paramagnetic susceptibility, where C is Curie's constant. The susceptibility in the limit when the external field is small, can be written as

$$\chi = \frac{M}{B} = \frac{\chi_p}{1 - \lambda\chi_p} = \frac{C}{T - T_c} \quad (16.13)$$

where $T_c = \lambda C$. At this point several remarks are in order. The susceptibility diverges at the finite temperature, T_c , instead of at zero temperature as in the paramagnetic phase. The susceptibility may be expanded as

$$\begin{aligned} \chi &= \chi_p (1 + \lambda\chi_p + (\lambda\chi_p)^2 + \cdots) \\ &= \frac{C}{T} + \frac{C}{T}\lambda\frac{C}{T} + \frac{C}{T}\lambda\frac{C}{T}\lambda\frac{C}{T} + \cdots \end{aligned} \quad (16.14)$$

Truncating the sum at any finite order the divergence of the susceptibility occurs at zero temperature. Only summing the infinite series the divergence temperature

shifts to the finite critical temperature, T_c . Recalling that $\lambda \sim |J|$ this shows that the result is non-perturbative in the coupling between the spins.

In the realistic case when the spin coupling is of a vectorial nature it can be easily shown that the critical temperature is given by

$$k_B T_c = \frac{1}{3} g \mu_B z S(S+1) |J| \quad (16.15)$$

16.2 Ferrimagnetism and antiferromagnetism

Some magnetic systems show a zero temperature magnetization that does not reach the saturation value expected using Hund's rules to determine the spin value of the magnetic moment. An example is magnetite, Fe_3O_4 . The iron atoms are involved in two types of bonds FeO and Fe_2O_3 . The electronic configuration of oxygen is $2s^2 2p^4$ and of iron is $3d^6 4s^2$. A s -orbital is full with two electrons, a p -orbital with 6 electrons and a d -orbital with 10 electrons. In a bond FeO the two $4s$ electrons are transferred to the p -orbital of the oxygen atom. The iron atom is therefore in the valence state Fe^{2+} with a configuration $3d^6$. Hund's rules lead to $S = 2$ and $L = 2$ but is expected to be quenched by the crystal field.

In the bonds Fe_2O_3 the four s -electrons of the iron atoms are transferred to the p -orbitals of the oxygens plus one d -electron of each iron atom. The valence state is therefore Fe^{3+} and the electronic configuration of each iron atom is $3d^5$. This leads to $S = 5/2$ and $L = 0$. Therefore we expect a spin $N(1 \times 2 + 2 \times 5/2) = N7$. However, the zero temperature magnetization corresponds to a spin $S = 2.05$. In the unit cell of the lattice structure of magnetite there are two types of lattice sites. There are A sites with tetrahedral symmetry filled with Fe^{3+} ions and B sites with octohedral symmetry filled with Fe^{3+} and Fe^{2+} ions in equal amounts. The interactions between the various spins is such that their exchange integrals favor antiparallel orientations. Specifically, $J_{AA}, J_{AB}, J_{BB} < 0$, are all of an antiferromagnetic origin. The $S = 5/2$ of the A sites are compensated by the $S = 5/2$ spins of the B sites. We are therefore left with the $S = 2$ spins of the remaining B sites explaining the observed experimental value.

It turns out that this explanation lies on the fact that the largest exchange constant is the interaction between the A and B spins: $|J_{AB}| > |J_{AA}|, |J_{BB}|$. It is energetically favorable that at low temperatures the A spins are parallel, the same with the B spins and the A and B spins are antiparallel with respect to each other.

Even though some of the orientations individually do not satisfy the alignment suggested by the exchange couplings (some bonds are frustrated) the overall configuration minimizes the energy.

The exchange fields satisfy

$$\begin{aligned}\vec{B}_A &= -\lambda\vec{M}_A - \mu\vec{M}_B \\ \vec{B}_B &= -\mu\vec{M}_A - \nu\vec{M}_B\end{aligned}\tag{16.16}$$

Where λ, μ, ν are positive constants and the minus signs indicate that each coupling has an antiferromagnetic nature. The energy of the interaction of the magnetic moments with the exchange magnetic fields is given by

$$\begin{aligned}U &= -\frac{1}{2}(\vec{B}_A \cdot \vec{M}_A + \vec{B}_B \cdot \vec{M}_B) \\ &= \frac{1}{2}\lambda M_A^2 + \mu\vec{M}_A \cdot \vec{M}_B + \frac{1}{2}\nu M_B^2\end{aligned}\tag{16.17}$$

The first and third terms are clearly positive. However the second term favors an antiparallel alignment between the A and B spins. There is a trivial solution $M_A = M_B = 0$ that leads to a vanishing energy. A more favorable phase is obtained if the energy is negative. This implies

$$\frac{1}{2}(\lambda M_A^2 + \nu M_B^2) < \mu M_A M_B\tag{16.18}$$

if we take $\vec{M}_A \cdot \vec{M}_B = -M_A M_B$.

To simplify let us take $\lambda = \nu = 0$. In this case we get

$$\begin{aligned}\vec{B}_A &= -\mu\vec{M}_B \\ \vec{B}_B &= -\mu\vec{M}_A\end{aligned}\tag{16.19}$$

Applying an external magnetic field, B_a , and following similar steps to the ferromagnetic case, we can write, in the paramagnetic phase, that

$$\begin{aligned}M_A T &= C_A (B_a - \mu M_B) \\ M_B T &= C_B (B_a - \mu M_A)\end{aligned}\tag{16.20}$$

where C_A, C_B are the Curie constants of the spins of sites A and B , respectively, and T is the temperature.

Consider now a vanishing external field $B_a = 0$. A non-trivial solution is obtained near the critical point if

$$T_c = \mu \sqrt{C_A C_B} \quad (16.21)$$

and we get that

$$M_A = -\frac{\sqrt{C_A}}{\sqrt{C_B}} M_B \quad (16.22)$$

The susceptibility is given by

$$\chi = \frac{M_A + M_B}{B_a} = \frac{(C_A + C_B) T - 2\mu C_A C_B}{T^2 - T_c^2} \quad (16.23)$$

and diverges at the critical temperature T_C , similarly to a ferromagnet.

An interesting case occurs if $C_A = C_B$. In this case $M_A = -M_B$ and is called the antiferromagnetic case. In this case $C = C_A = C_B$ and $T_c = \mu C$ which is called the Néel temperature, T_N . The susceptibility is reduced to

$$\chi = \frac{2C}{T + T_N} \quad (16.24)$$

It is no longer divergent. As one enters the antiferromagnetic phase the susceptibility has a cusp but no divergence. Also, the total magnetization at zero temperature vanishes. The magnetization of each sublattice is finite however, but since the spins are oriented in opposite directions the total magnetization vanishes.

Considering a temperature smaller than the Néel temperature it turns out that the susceptibility depends on the orientation of the applied magnetic field with respect to the direction of the anti-aligned spins. If the magnetic field is applied perpendicular to the direction of the spins both the A and B spins will tend to align along the direction of the applied (small) magnetic field. This costs energy since the exchange energy is no longer minimized. If each magnetization A and B approaches the direction of the applied field by an angle ϕ , the angle between each magnetic moment is 2ϕ . Denoting $M = |M_A| = |M_B|$ we get that the energy has the form

$$\begin{aligned} U &= \mu \vec{M}_A \cdot \vec{M}_B - \vec{B}_a \cdot (\vec{M}_A + \vec{M}_B) \\ &= -\mu M^2 \left(1 - \frac{1}{2}(2\phi)^2\right) - 2B_a M \phi \end{aligned} \quad (16.25)$$

Searching for an equilibrium solution we may find the angle that minimizes the energy. This leads to

$$\frac{dU}{d\phi} = 0 = 4\mu M^2 \phi - 2B_a M \quad (16.26)$$

and therefore an equilibrium value

$$\phi = \frac{B_a}{2\mu M} \quad (16.27)$$

The perpendicular susceptibility is therefore a constant at low temperatures

$$\chi_{\perp} = \frac{2M\phi}{B_a} = \frac{1}{\mu} \quad (16.28)$$

On the other hand the parallel susceptibility obtained applying an external field along the direction of the magnetic moments costs no exchange energy and the susceptibility vanishes. This happens if the magnetic moments of the A and B sites change orientation by the same angle.

16.3 Excitations in ferromagnetic systems

Let us consider once again the Heisenberg model and consider for now the one-dimensional case

$$H = -2J \sum_{i=1}^N \vec{S}_i \cdot \vec{S}_{i+1}$$

Let us consider here the classical model where we take the spins as classical vectors and let us take $J > 0$. The ground state is obtained if all the spins are parallel. In this case $\vec{S}_i \cdot \vec{S}_{i+1} = S^2$ and the energy of the groundstate is given by $E_0 = -2NJS^2$. We want now to study the excitations around this minimum energy state.

i) Consider for instance an excitation in which one spin is turned "down". This state has an energy $E = E_0 + 8JS^2$. This excitation is a single particle (single spin) excitation.

ii) It is also possible to construct an excitation that requires much less energy if we divide the fluctuation over all spins in the system. Let us consider the collective

excitation that is constituted of a wave that corresponds to the change in the orientation of the spins in the plane perpendicular to their preferential direction (which was assumed as pointing "up"). The terms in the energy that involve the spin i are

$$-2J\vec{S}_i \cdot (\vec{S}_{i+1} + \vec{S}_{i-1})$$

The magnetic moment at site i is $\vec{\mu}_i = -g\mu_B\vec{S}_i$. Therefore we can rewrite that the contribution to the energy due to the site i is

$$-\vec{\mu}_i \cdot \left[\left(\frac{-2J}{g\mu_B} \right) (\vec{S}_{i+1} + \vec{S}_{i-1}) \right] \sim -\vec{\mu}_i \cdot \vec{B}$$

where \vec{B} is the effective magnetic field acting on the spin i . From the electromagnetism we know that

$$\hbar \frac{d\vec{S}_i}{dt} = \vec{\mu}_i \times \vec{B}_i$$

we get therefore the equation of motion for the spins in the form

$$\frac{d\vec{S}_i}{dt} = \frac{2J}{\hbar} (\vec{S}_i \times \vec{S}_{i+1} + \vec{S}_i \times \vec{S}_{i-1}) \quad (16.29)$$

If the amplitude of the excitation is small then $S_i^x, S_i^y \ll S$ and we can linearize the equation of motion taking $S_i^z = S$ and neglecting terms of the order of $S^x S^y$

$$\begin{aligned} \frac{dS_i^z}{dt} &= 0 \\ \frac{dS_i^x}{dt} &= \frac{2JS}{\hbar} (2S_i^y - S_{i+1}^y - S_{i-1}^y) \\ \frac{dS_i^y}{dt} &= -\frac{2JS}{\hbar} (2S_i^x - S_{i+1}^x - S_{i-1}^x) \end{aligned} \quad (16.30)$$

Let us look for solutions of the type

$$S_p^x = ue^{i(pka - \omega t)}$$

$$S_p^y = ve^{i(pka - \omega t)}$$

Substituting these solutions in the equations of motion we get that

$$\begin{aligned} -i\omega u &= \frac{2JS}{\hbar} \left(2 - e^{-ika} - e^{ika} \right) v = \frac{4JS}{\hbar} (1 - \cos ka) v \\ -i\omega v &= -\frac{2JS}{\hbar} \left(2 - e^{-ika} - e^{ika} \right) u = -\frac{4JS}{\hbar} (1 - \cos ka) u \end{aligned} \quad (16.31)$$

To get a nontrivial solution the determinant of the matrix

$$\begin{pmatrix} i\omega & \frac{4JS}{\hbar}(1 - \cos ka) \\ -\frac{4JS}{\hbar}(1 - \cos ka) & i\omega \end{pmatrix}$$

must vanish. This leads to the dispersion relation

$$\hbar\omega = 4JS(1 - \cos ka) \quad (16.32)$$

The solution corresponds to $v = -iu$ which is a circular precession around the z axis. At low momenta the dispersion relation is quadratic $\omega \sim k^2$. It saturates at π/a with a zero velocity. To be more precise

$$\hbar\omega \sim (2JSa^2)k^2$$

at small momenta.

Let us now quantize these waves called spin-waves. They are bosonic like (classical) and let us quantize them by using the spectrum of the harmonic oscillator

$$\epsilon_k = \left(n_k + \frac{1}{2}\right)\hbar\omega_k$$

(In this approximation the spin waves do not interact and therefore the linearization leads to independent modes which we have quantized. Later on we will solve the problem with more care). Since these excitations are bosons their thermodynamic occupation number obeys the Bose-Einstein distribution (with a zero chemical potential since their number is not conserved) like

$$\langle n_k \rangle = \frac{1}{e^{\frac{\hbar\omega_k}{k_B T}} - 1}$$

The total number of these quantized spin waves (called magnons) is given by

$$\sum_k n_k = \int d\omega D(\omega) \langle n(\omega) \rangle$$

Here the density of states per unit volume is given by

$$D(\omega) = \left(\frac{1}{2\pi}\right)^3 4\pi k^2 \frac{dk}{d\omega}$$

Using that $\hbar\omega \sim (2JSa^2)k^2$ we obtain that

$$D(\omega) = \frac{1}{4\pi^2} \left(\frac{\hbar}{2JSa^2}\right)^{3/2} \omega^{1/2}$$

The total number of magnons is therefore

$$\begin{aligned} \sum_k n_k &= \frac{1}{4\pi^2} \left(\frac{\hbar}{2JSa^2}\right)^{3/2} \int_0^\infty d\omega \frac{\omega^{1/2}}{e^{\beta\hbar\omega} - 1} \\ &= \frac{1}{4\pi^2} \left(\frac{k_B T}{2JSa^2}\right)^{3/2} \int_0^\infty dx \frac{x^{1/2}}{e^x - 1} \end{aligned}$$

The integral is equal to $4\pi^2 0.0587$. Therefore we Bloch's Law that gives that

$$\frac{\sum_k n_k}{NS} = \frac{\Delta M}{M(0)} \sim \left(\frac{k_B T}{JS}\right)^{3/2} \quad (16.33)$$

for the deviation in the magnetization from the groundstate.

16.3.1 Holstein-Primakoff transformation

Again consider the Heisenberg Hamiltonian

$$H = -J \sum_{\vec{j}, \vec{\delta}} \vec{S}_{\vec{j}} \cdot \vec{S}_{\vec{j}+\vec{\delta}} - 2\mu_0 H_0 \sum_{\vec{j}} S_{\vec{j},z} \quad (16.34)$$

with $\mu_0 = 2\mu_B/2$. The total spin of the system is such that $\vec{S}^2 = (\sum_j \vec{S}_j)^2$ (we have replaced \vec{j} by j to simplify the notation) and $S_z = \sum_j S_{j,z}$. The groundstate of the ferromagnetic system is such that

$$\begin{aligned} \vec{S}^2 |0\rangle &= NS(S+1) |0\rangle \\ S_z |0\rangle &= NS |0\rangle \end{aligned}$$

It is a state completely ordered where all spins are aligned parallel. This is the ground state even if the magnetic field vanishes. Then the direction of alignment is arbitrary and after a direction is fixed this breaks the initial rotational symmetry.

Let us now introduce a bosonic description of the spin operators. First let us define the raising and lowering spin operators like

$$\begin{aligned} S_j^+ &= S_{jx} + iS_{jy} = \sqrt{2S} \sqrt{1 - \frac{a_j^\dagger a_j}{2S}} a_j \\ S_j^- &= S_{jx} - iS_{jy} = \sqrt{2S} a_j^\dagger \sqrt{1 - \frac{a_j^\dagger a_j}{2S}} \end{aligned} \quad (16.35)$$

The operators a_j^\dagger and a_j are bosonic operators satisfying

$$[a_j, a_l^\dagger] = \delta_{jl}$$

Using that $S_z^2 = S(S+1) - S_x^2 - S_y^2$, introducing the above expressions for the ladder operators, using that $[a^\dagger a, a] = -a$ we can show that

$$S_{jz} = S - a_j^\dagger a_j \quad (16.36)$$

To verify that this bosonic representation is accurate we have to show that the operators obey the commutation relations of the spin operators. It is easy to verify that for instance $[S^+, S^-] = 2S^z$ using the above definitions in terms of the bosonic operators.

Let us now consider the Fourier transform of the bosonic operators

$$\begin{aligned} b_k &= \frac{1}{\sqrt{N}} \sum_j e^{ik \cdot x_j} a_j \\ b_k^\dagger &= \frac{1}{\sqrt{N}} \sum_j e^{-ik \cdot x_j} a_j^\dagger \end{aligned}$$

These definitions lead to a commutation relation $[b_k, b_{k'}^\dagger] = \delta_{k,k'}$ which shows that they are bosons as well. We can then write the spin operators in terms of these

momentum space bosonic operators like

$$\begin{aligned}
S_j^+ &= \sqrt{2S} \left[a_j - \frac{1}{4S} a_j^\dagger a_j a_j + \dots \right] \\
&= \sqrt{\frac{2S}{N}} \left[\sum_k e^{-ik \cdot x_j} b_k - \frac{1}{4SN} \sum_{kk'k''} e^{i(k-k'-k'') \cdot x_j} b_k^\dagger b_{k'} b_{k''} + \dots \right] \\
S_j^- &= \sqrt{\frac{2S}{N}} \left[\sum_k e^{ik \cdot x_j} b_k^\dagger - \frac{1}{4SN} \sum_{kk'k''} e^{i(k+k'-k'') \cdot x_j} b_k^\dagger b_{k'}^\dagger b_{k''} + \dots \right] \\
S_j^z &= S - a_j^\dagger a_j = S - \frac{1}{N} \sum_{kk'} e^{i(k-k') \cdot x_j} b_k^\dagger b_{k'}
\end{aligned} \tag{16.37}$$

Also the deviation from the maximal magnetization is given by

$$\delta S_z = NS - \sum_j S_{jz} = \sum_j a_j^\dagger a_j = \sum_k b_k^\dagger b_k \tag{16.38}$$

Let us now rewrite the Hamiltonian in terms of these bosonic operators. It yields considering that there are z nearest neighbors

$$H = -JNzS^2 - 2\mu_0 H_0 NS + H_0 + H_1 \tag{16.39}$$

Here

$$H_0 = -JzS \sum_k \left(\gamma_k b_k b_k^\dagger + \gamma_{-k} b_k^\dagger b_k - 2b_k^\dagger b_k \right) + 2\mu_0 H_0 \sum_k b_k^\dagger b_k \tag{16.40}$$

where

$$\gamma_k = \frac{1}{z} \sum_{\delta} e^{ik \cdot \delta}$$

Note that $\sum_k \gamma_k = 0$. If there is a center of symmetry then $\gamma_k = \gamma_{-k}$. In this case we get that

$$H_0 = \sum_k n_k \omega_k \tag{16.41}$$

where the dispersion relation of the excitations (spin-waves) is

$$\omega_k = 2JSz(1 - \gamma_k) + 2\mu_0 H_0 \tag{16.42}$$

The term H_1 contains all higher order terms in the boson operators like cubic and higher. These terms describe the interactions between the spin waves and can be systematically included. We will not pursue this here.

Writing the dispersion relation as $\omega_k = Dk^2$, where D is called the spin stiffness, we get that, if the magnetic field is zero, $D = 2SJa^2$. The total energy is given by

$$\begin{aligned} U &= \sum_k \langle n_k \rangle \omega_k = \sum_k \omega_k \frac{1}{e^{\beta\omega_k} - 1} = \frac{1}{(2\pi)^3} \int d^3k Dk^2 \frac{1}{e^{\beta Dk^2} - 1} \\ &= \frac{(k_B T)^{5/2}}{4\pi^2 D^{3/2}} \int_0^{x_m} dx x^{3/2} \frac{1}{e^x - 1} \end{aligned} \quad (16.43)$$

This means that $U \sim T^{5/2}$ and that the specific heat is $C_V \sim T^{3/2}$.

16.4 Ferrimagnetism and anti-ferromagnetism

Consider now the case of antiferromagnetism described by the Heisenberg Hamiltonian but now with an opposite coupling constant: it is now favorable classically for the spins to align in opposite directions. There are now two interpenetrating sublattices such that the spins on each lattice are preferably aligned while spins in different lattices are antiparallel. Therefore the nearest neighbors of each spin belong to the other sublattice.

The Hamiltonian can now be written conveniently as

$$H = J \sum_{j\delta} \vec{S}_j \cdot \vec{S}_{j+\delta} - 2\mu_0 H_A \sum_j S_{jz}^a + 2\mu_0 H_A - \sum_j m_j S_{jz}^b \quad (16.44)$$

We have divided the system in sublattices A and B and assigned the labels a and b to the spins belonging to the two sublattices, respectively. Here $J > 0$ to have antiferromagnetism.

Let us now define two sets of bosonic operators one associated to the spins a and

the other set to the spins b . Then we define

$$\begin{aligned}
 S_{aj}^+ &= \sqrt{2S} \left(1 - \frac{1}{2S} a_j^\dagger a_j \right)^{1/2} a_j \\
 S_{aj}^- &= \sqrt{2S} a_j^\dagger \left(1 - \frac{1}{2S} a_j^\dagger a_j \right)^{1/2} \\
 S_{bl}^+ &= \sqrt{2S} b_l^\dagger \left(1 - \frac{1}{2S} b_l^\dagger b_l \right)^{1/2} \\
 S_{bl}^- &= \sqrt{2S} \left(1 - \frac{1}{2S} b_l^\dagger b_l \right)^{1/2} b_l \\
 S_{aj}^z &= S - a_j^\dagger a_j \\
 S_{bl}^z &= -S + b_l^\dagger b_l
 \end{aligned}$$

Note that the spins of the B lattice point in the opposite direction and therefore the bosons describe deviations from the opposite direction. The presence of bosons associated with the A sites describes the lowering of the z -component of the magnetization. Therefore lowering the spin component is associated with creating a boson. On the other hand, for the B sites the presence of bosons is associated with the increase of the spin component: therefore the raising operator is associated with the creation of a boson. Like before we define the Fourier transforms of the bosonic fields

$$\begin{aligned}
 c_k &= \frac{1}{\sqrt{N}} \sum_j^a e^{ik \cdot x_j} a_j \\
 c_k^\dagger &= \frac{1}{\sqrt{N}} \sum_j^a e^{-ik \cdot x_j} a_j^\dagger \\
 d_k &= \frac{1}{\sqrt{N}} \sum_l^b e^{-ik \cdot x_l} b_l \\
 d_k^\dagger &= \frac{1}{\sqrt{N}} \sum_l^b e^{ik \cdot x_l} b_l^\dagger
 \end{aligned} \tag{16.45}$$

Substituting these definitions in the expressions for the spin operators in terms of the bosons and keeping only the quadratic terms in the bosonic operators we get

that as before the Hamiltonian can be written as

$$H = -2NzJS^2 - 4N\mu_0H_AS + H_0 + H_1 \quad (16.46)$$

where as before H_1 is the term that contains the interactions between the magnons and

$$H_0 = 2JzS \sum_k [\gamma_k (c_k^\dagger d_k^\dagger + c_k d_k) + (c_k^\dagger c_k + d_k^\dagger d_k)] + 2\mu_0H_A \sum_k (c_k^\dagger c_k + d_k^\dagger d_k) \quad (16.47)$$

here

$$\gamma_k = \frac{1}{z} \sum_\delta e^{ik \cdot \delta} = \gamma_{-k} \quad (16.48)$$

if there is a center of symmetry, as before.

To diagonalize the Hamiltonian let us make a Bogoliubov-Valatin transformation. Define

$$\begin{aligned} \alpha_k &= u_k c_k - v_k d_k^\dagger \\ \alpha_k^\dagger &= u_k c_k^\dagger - v_k d_k \\ \beta_k &= u_k d_k - v_k c_k^\dagger \\ \beta_k^\dagger &= u_k d_k^\dagger - v_k c_k \end{aligned} \quad (16.49)$$

As before using that the new operators are also bosons we get that $u_k^2 - v_k^2 = 1$. Comparing with our solution of the phonons in the Bose liquid we get that making the replacements

$$a_k \rightarrow c_k; a_k^\dagger \rightarrow c_k^\dagger; a_{-k} \rightarrow d_k; a_{-k}^\dagger \rightarrow d_k^\dagger$$

and making the replacements

$$\omega_0 \rightarrow 2JzS + 2\mu_0H_A; \omega_1 \rightarrow 2JzS\gamma_k$$

This leads then to

$$\omega_k^2 = (\omega_e + \omega_A)^2 - \omega_e^2 \gamma_k^2 \quad (16.50)$$

with $\omega_e = 2JzS$ and $\omega_A = 2\mu_0H_A$. In terms of the new operators we can write the Hamiltonian like

$$H_0 = -N(\omega_e + \omega_A) + \sum_k \omega_k (\alpha_k^\dagger \alpha_k + \beta_k^\dagger \beta_k + 1)$$

and the full Hamiltonian is then written as

$$H = -2NzJS(S+1) - 4\pi\mu_0 H_A(S + \frac{1}{2}) + \sum_k \omega_k(n_k + \frac{1}{2}) + H_1 \quad (16.51)$$

here the sum over the momentum has a weight 2 for each momentum value. If the magnetic field vanishes and we consider small momenta we get that

$$\omega_k \sim (4\sqrt{3}JSa)k \quad (16.52)$$

and therefore the dispersion relation is linear in k .

It is interesting to calculate the zero temperature magnetization of each sublattice. For instance

$$\begin{aligned} S_a^z &= \sum_j S_{aj}^z = NS - \sum_k c_k^\dagger c_k \\ &= NS - \sum_k (u_k^2 \alpha_k^\dagger \alpha_k + v_k^2 \beta_k \beta_k^\dagger + \text{terms not diagonal}) \end{aligned}$$

Consider then $T = 0$ and $n_k = 0$. Then the magnetization deviation is given by

$$\begin{aligned} \Delta S^z &= NS - \langle S^z \rangle = \sum_k v_k^2 \\ &= \frac{1}{2} \sum_k (\cosh 2\chi_k - 1) \\ &= \frac{1}{2} \sum_k \left(\frac{1}{\sqrt{1 - \gamma_k^2}} - 1 \right) \end{aligned}$$

The change of magnetization per unit volume is

$$\Delta S^z = -\frac{N}{2} + \frac{1}{2(2\pi)^3} \int d^3k \frac{1}{\sqrt{1 - \gamma_k^2}} = 0.078N \neq 0 \quad (16.53)$$

This numerical result is for the cubic lattice. The quantum result is different from the classical case since the sublattice magnetization is not fully saturated.

Capítulo 18

Superconductivity

18.1 Zero resistance and Meissner effect

In 1911 Kamerling Onnes found a sudden drop of the resistance of Hg at very low temperatures. At about 4.2 Kelvin the resistance is apparently zero. In 1933 the Meissner effect was discovered: a similar material penetrated by a magnetic field completely expels the magnetic field lines from its interior (the magnetic induction vanishes inside the material $B = 0$) as the temperature is lowered beyond the point where the resistance vanishes.

A set of properties later discovered may shed some light into the mechanism that gives rise to such spectacular phenomenon, called superconductivity due to its infinite conductivity.

i) Above the critical temperature the materials are conductors but the best metals do not become superconductors.

ii) The materials that are good magnets also do not become superconductors.

iii) At low temperatures the specific heat changes from a behavior typical of a conductor with the electronic and phonon contributions to a regime where it becomes exponentially small at very low temperatures. This is indicative of the presence of a gap in energy spectrum, as seen for instance in a two-level system.

iv) The temperatures at which the system becomes a superconductor are typically very small, usually below 20 Kelvin.

v) An important clue was provided by the discovery of the isotope effect. Substituting some atoms by different isotopes their mass is changed. It turns out that

the critical temperature is changed, indicating that somehow the ions are involved in the process.

vi) A large magnetic field destroys the superconducting state.

It should be noted that the above set of properties are common to the so-called conventional superconductors. Other superconductors do not follow those properties but we will focus here in explaining the conventional ones for which a theory has been constructed with great success.

The list of the above properties seems to suggest the following picture: interactions must have an important role since very good conductors do not become superconductors. It is expected that in very good conductors interactions are not important and resistance is low. The existence of a gap suggests some form of bound state between the electrons that must be weak since the critical temperatures are usually rather low. Magnetism and magnetic fields compete with superconductivity. The influence of the lattice suggests that a possible mechanism for the creation of bound states involves the lattice vibrations, that is the phonons. Indeed, in order to create a bound state between the electrons, one expects an attractive interaction between the electrons. The direct Coulomb interaction is repulsive, even though it should be screened. Some other interaction must yield an attractive force. It has been proposed that this mechanism involves the exchange of phonons between the electrons. If boundstates between the electrons exist one expects that these extend in space since their energy should be small. On the other hand, since the electrons are fermions a pair of electrons is like a boson. Therefore, we may expect at low temperatures some sort of boson condensation and some sort of superfluid state, with no resistance. Note however that this picture is somewhat simplified because a pair of electrons only seems like a boson far away from it. This picture is oversimplified since the radius of a pair is relatively large (of the order of a thousand Angstrom).

18.1.1 London theory

A simple early theory assumes that the superconductor is composed of two fluids a normal and a superfluid with densities n_n and n_s and velocities v_n and v_s , respectively. The electronic density of the material is $n = n_n + n_s$. The current of the normal part, $\vec{J}_n = -en_n\vec{v}_n$, follows the usual Ohm's law

$$\vec{J}_n = \sigma_n \vec{E} \quad (18.1)$$

under the presence of an electric field \vec{E} . This law indicates a dissipative current due to the presence of friction with impurities, imperfections, phonons, etc. The superfluid part does not follow Ohm's law since it has infinite conductivity. The motion of the particles of the superfluid follow Newton's law instead. The particles are accelerated due to the presence of an external force. That is the force is proportional to the acceleration while Ohm's law describes a set of particles whose force is proportional to the velocity. For the superfluid part we postulate that, defining the current as $\vec{J}_s = -en_s\vec{v}_s$,

$$\frac{d\vec{J}_s}{dt} = -en_s\vec{a} = \frac{e^2n_s}{m}\vec{E} \quad (18.2)$$

Using that $\vec{E} \sim \partial\vec{A}/\partial t$ this shows that the current is proportional to the vector potential, \vec{A} , instead of proportional to the electric field.

Maxwell's equation

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t} \quad (18.3)$$

can be written as

$$\vec{\nabla} \times \vec{J}_s = -\frac{n_se^2}{mc} \vec{B} \quad (18.4)$$

Another Maxwell equation is written as

$$\vec{\nabla} \times \vec{B} = \frac{4\pi}{c} \vec{J} = \frac{4\pi}{c} (\vec{J}_n + \vec{J}_s) \quad (18.5)$$

where we have neglected the displacement current. Since we expect the superfluid current to be much larger than the normal current, we also neglect its contribution. Taking the curl of this equation we get that

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{B}) = \frac{4\pi}{c} \vec{\nabla} \times \vec{J}_s \quad (18.6)$$

Using that $\vec{\nabla} \cdot \vec{B} = 0$ we get that

$$\nabla^2 \vec{B} = \frac{4\pi n_se^2}{mc^2} \vec{B} = \frac{1}{\lambda_L^2} \vec{B} \quad (18.7)$$

where, λ_L , the London penetration depth is given by

$$\lambda_L = \left(\frac{mc^2}{4\pi n_se^2} \right)^{1/2} \quad (18.8)$$

To understand why it is called penetration depth let us consider a superconductor in contact with the vacuum. Let us call the direction perpendicular to the surface of the superconductor z and place the surface at $z = 0$. To simplify we assume that the superconductor surface is flat and is contained in the infinite plane $x - y$. Consider that the field in the vacuum is parallel to the surface, $\vec{B} = B\vec{e}_x$, for instance. We have then to solve the set of equations

$$\begin{aligned}\vec{\nabla} \times \vec{B} &= \frac{4\pi}{c} \vec{J}_s \\ \nabla^2 \vec{B} &= \frac{1}{\lambda_L^2} \vec{B} \\ \vec{\nabla} \cdot \vec{B} &= 0\end{aligned}\tag{18.9}$$

near the surface. If the field is tangential and along the x direction then the current is along the y direction $\vec{J}_s = J_s \vec{e}_y$. The first equation then simplifies to

$$\frac{dB}{dz} = \frac{4\pi}{c} J_s\tag{18.10}$$

Also,

$$\begin{aligned}\frac{dJ_s}{dz} &= \frac{n_s e^2}{mc} B \\ \frac{d^2 B}{dz^2} &= \frac{B}{\lambda_L^2}\end{aligned}\tag{18.11}$$

This leads to an exponentially decreasing profile for the amplitude of the magnetic field inside the superconductor with a scale given by the London penetration depth

$$B(z) = B(z=0)e^{-\frac{z}{\lambda_L}}\tag{18.12}$$

Deep inside the superconductor the magnetic field, B , vanishes: this is the result expected from the Meisner effect.

The penetration depth is a function of the superfluid density. At very low temperature the density is basically the superfluid density. For *Al* and *Sn* the electron mass is very close to the free electron mass and $\lambda_L \sim 500$ Angstrom. For transition metals the effective mass is larger and the penetration length increases, $\lambda_L \sim 2000$ Angstrom. Also, for systems with low electronic density the penetration length increases.

18.1.2 Thermodynamics in the presence of external field

Consider a superconductor and a magnet that is brought from infinity into the neighborhood of the superconducting material. Let us call the field created by the magnet on the superconductor, B_a . Assume that this applied field is approximately constant in the region of the superconductor. If the field is sufficiently small, such that the system remains superconducting, Meissner effect implies that the magnetic induction inside the superconductor vanishes. Since $B = H + 4\pi M$, $B = 0$ implies that $M = -B_a/(4\pi)$ (we will use H or B_a interchangeably). This means the superconductor is a perfect diamagnet.

The work done on the superconductor as we bring the magnet from infinity ($B_a = 0$) until a given point where $B_a \neq 0$, is given by

$$W = - \int_0^{B_a} \vec{M} \cdot d\vec{B}_a \quad (18.13)$$

The change in free energy is given by $dF = -\vec{M} \cdot d\vec{B}_a$ which leads to

$$dF_s = \frac{1}{4\pi} B_a dB_a \quad (18.14)$$

This can be integrated easily to yield

$$F_s(B_a) - F_s(0) = \frac{B_a^2}{8\pi} \quad (18.15)$$

If the applied field is large enough the system will suffer a transition from the superconducting phase to the normal phase (conductor). At the transition the free energies of the two phases are equal: $F_s(B_{ac}) = F_N(B_{ac})$, where B_{ac} is the critical field that causes the transition to occur. On the other hand, in a normal metal (non-magnetic) the magnetic susceptibility is small. We may then approximate $F_N(B_{ac}) = F_N(0)$. This holds if the critical field is not large. Therefore,

$$F_N(B_{ac}) = F_s(B_{ac}) = F_s(0) + \frac{B_{ac}^2}{8\pi} = F_N(0) \quad (18.16)$$

This implies that the change in the free energy, *in zero field*, is given by

$$\Delta F = F_N(0) - F_s(0) = \frac{B_{ac}^2}{8\pi} \quad (18.17)$$

As expected the free energy of the normal phase is larger than in the superconducting phase. This difference is called the condensation energy and shows that the system becomes more stable by changing phase into the superconducting regime. Also, it is related to the critical magnetic field showing a connection to the magnetic response of the system.

We may also study the entropy of the two phases. We know that

$$S_N = -\frac{dF_N}{dT}, S_s = -\frac{dF_s}{dT} \quad (18.18)$$

The difference in entropy between the two phases is given by

$$S_N - S_s = -\frac{d}{dT} (F_N - F_s) = -\frac{1}{4\pi} H_c \frac{dH_c}{dT} \quad (18.19)$$

The normal phase is a higher temperature phase and it is expected to have higher entropy. Therefore, the entropy difference should be positive. This implies that the critical magnetic field should decrease with temperature. This is to be expected: at zero field there is a finite critical temperature at which superconductivity emerges. A sufficiently large applied magnetic field destroys superconductivity. Therefore, as temperature increases the superconducting state becomes less robust and a smaller magnetic field is enough to destroy superconductivity.

If there is an entropy difference between the normal and the superconducting phase in the presence of a magnetic field, there is a latent heat at the transition. The latent heat is given by

$$L = T(S_N - S_s) = -\frac{T}{4\pi} H_c \frac{dH_c}{dT} > 0 \quad (18.20)$$

Phenomenologically it is known that, approximately

$$H_c(T) = H_c(0) \left(1 - \frac{T^2}{T_c^2} \right) \quad (18.21)$$

We may also calculate the change in the specific heat at the transition

$$\begin{aligned} C_N - C_s &= T \frac{d}{dT} (S_N - S_s)_{T=T_c} \\ &= -\frac{T}{4\pi} \left(\frac{dH_c}{dT} \right)^2_{T=T_c} \end{aligned} \quad (18.22)$$

There is a discontinuity at the transition. Since there is a latent heat, the transition from the normal phase to the superconducting phase is of first order if the magnetic field is finite. At zero field, the discontinuity vanishes which suggest a second order phase transtion near the critical temperature T_c . (It turns out the transition is actually weakly first order, the discontinuity is very small, and we will assume the transition to be second order).

It is easy to see that

$$\left(\frac{dH_c(T)}{dT} \right)_{T=T_c} = -2 \frac{H_c(0)}{T_c} \quad (18.23)$$

which allows an estimate of the critical field at zero temperature from the slope of dependence of the critical field near the critical point $T = T_c$.

18.2 Cooper pairs

Consider again the Fermi liquid. The Coulomb interactions between the electrons are repulsive in the vacuum. However consider now the electrons in a solid. As we saw the interactions get renormalized in the medium. Let us assume that the interactions between the electrons and the lattice may cause an effective attractive interaction between the electrons. Such is possible through for instance the electron-phonon interaction as we shall see later on. For the moment let us assume that there is an attractive interaction between the electrons. Let us consider two electrons plus the Fermi gas that occupy the Fermi sea.

The electrons have coordinates \vec{r}_1 and \vec{r}_2 and interact via an attractive potential. To simplify let us consider states where the center of mass of the two electrons is at rest. Then the wave function of the pair is

$$\psi(\vec{r}_1, \vec{r}_2) = \psi(\vec{r}_1 - \vec{r}_2) = \sum_{\vec{k}} g(\vec{k}) e^{i\vec{k} \cdot (\vec{r}_1 - \vec{r}_2)}$$

Here $g(\vec{k})$ is the probability amplitude that one electron has momentum \vec{k} (and that the other has momentum $-\vec{k}$). Pauli's principle implies that $g(\vec{k}) = 0$ if $k < k_F$. The Fermi sea is filled and we are considering two electrons above the Fermi sea. The Schroedinger equation is therefore

$$-\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) \psi(\vec{r}_1, \vec{r}_2) + V(\vec{r}_1, \vec{r}_2) \psi(\vec{r}_1, \vec{r}_2) = \left(E + \frac{\hbar^2 k_F^2}{m} \right) \psi(\vec{r}_1, \vec{r}_2)$$

In terms of the momentum amplitude we get that

$$\frac{\hbar^2}{2m}k^2 g(\vec{k}) + \sum_{\vec{k}'} g(\vec{k}') V_{\vec{k},\vec{k}'} = (E + 2E_F)g(\vec{k})$$

where the matrix element is given by

$$V_{\vec{k},\vec{k}'} = \frac{1}{L^3} \int d^3r V(\vec{r}) e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}}$$

($\vec{r} = \vec{r}_1 - \vec{r}_2$).

Consider now that $V < 0$ and let us look for bound states (solutions with negative energy). To be specific let us consider an interaction of the form

$$\begin{aligned} V_{\vec{k},\vec{k}'} &= -\frac{V}{L^3}, \frac{\hbar^2 k^2}{2m} < E_F + \hbar\omega_D, \frac{\hbar^2 (k')^2}{2m} < E_F + \hbar\omega_D \\ &= 0, \text{otherwise} \end{aligned} \quad (18.24)$$

Here ω_D is the Debye frequency of the phonons or any other cutoff frequency representing the maximum exchange energy available. Implicitly $\hbar\omega_D$ is the energy scale exchanged in the scattering of the electrons. Since we are assuming that the mechanism for the attractive interaction is the exchange of phonons between the electrons, $\hbar\omega_D$ is the Debye energy. If the mechanism were another then it will represent the energy exchanged through the excitations responsible for the attractive interaction.

Using the simplified potential we obtain that

$$\left(-\frac{\hbar^2 k^2}{2m} + E + 2E_F \right) g(\vec{k}) = -\frac{V}{L^3} \sum_{\vec{k}'} g(\vec{k}') = C$$

and

$$E_F < \frac{\hbar^2 (k')^2}{2m} < E_F + \hbar\omega_D$$

Summing over \vec{k}

$$\sum_{\vec{k}} g(\vec{k}) = \sum_{\vec{k}} \frac{1}{-E + \frac{\hbar^2 k^2}{2m} - 2E_F} \frac{V}{L^3} \sum_{\vec{k}'} g(\vec{k}')$$

We finally arrive at

$$1 = \frac{V}{L^3} \sum_{\vec{k}} \frac{1}{-E + \frac{\hbar^2 k^2}{2m} - 2E_F} \quad (18.25)$$

Defining $\zeta' = \frac{\hbar^2 (k')^2}{2m} - E_F$ and recalling that the density of states is given by

$$N(\zeta') = \frac{1}{(2\pi)^3} 4\pi (k')^2 \frac{dk'}{d\zeta'}$$

we can rewrite that

$$1 = V \int_0^{\hbar\omega_D} d\zeta' N(\zeta') \frac{1}{2\zeta' - E}$$

Since in general $\hbar\omega_D \ll E_F$ we can approximate the density of states by its value at the Fermi surface $N(0)$. Then the integral gives

$$\begin{aligned} 1 &= VN(0) \int_0^{\hbar\omega_D} d\zeta \frac{1}{2\zeta - E} \\ &= \frac{VN(0)}{2} \ln \frac{E - 2\hbar\omega_D}{E} \end{aligned}$$

This yields in the limit when $VN(0) \ll 1$ (in the weak coupling limit)

$$E \sim -2\hbar\omega_D e^{-\frac{2}{VN(0)}} \quad (18.26)$$

Several points are of relevance. First, as anticipated, the solution is of negative energy and there is a bound state. Second the energy of the boundstate is non-perturbative in the potential. Third whatever the value of the potential there is a bound state. This does not happen in free space as you may recall from elementary quantum mechanics (except in one dimension, which is not our concern here). In free space (3d) there is a minimum value for the potential required for the occurrence of a bound state. Fourth the wave function $\psi(\vec{r}_1, \vec{r}_2)$ has to be anti-symmetric since the electrons are fermions. With our simplified choice $g(\vec{k})$ only depends on $k = |\vec{k}|$. This implies that the spatial part of the wave function is symmetric. This implies on the other hand that the spin part of the wave function must be anti-symmetric. Therefore the two electrons are in a spinsinglet state.

Let us now calculate the mean square radius of a pair of electrons (called Cooper pair). It is given by

$$\rho^2 = \frac{\int d^3r r^2 |\psi(\vec{r}_1 - \vec{r}_2)|^2}{\int d^3r |\psi(\vec{r}_1 - \vec{r}_2)|^2}$$

This expression can be written as

$$\begin{aligned} \rho^2 &= \frac{\sum_{\vec{k}} |\nabla_{\vec{k}} g(\vec{k})|^2}{\sum_{\vec{k}} |g(\vec{k})|^2} \\ &= \frac{N(0) \left(\frac{\partial \zeta}{\partial k} \right)_{\zeta=0}^2 \int_0^\infty d\zeta \left(\frac{\partial g}{\partial \zeta} \right)^2}{N(0) \int_0^\infty d\zeta g^2} \end{aligned}$$

Recall that

$$g = \frac{C}{E - 2\zeta}$$

and that

$$\frac{dk}{d\zeta}_{\zeta=0} = \frac{1}{\hbar v_F}$$

Therefore we get for the mean square radius of the Cooper pair the result

$$\rho = \frac{2}{\sqrt{3}} \frac{\hbar v_F}{E} \quad (18.27)$$

18.3 Electron-phonon interaction and attractive interaction

From the point of view of the electrons the lattice acts as an external potential. The interaction may be written as

$$H_{int} = \int d^3r V(\vec{r}) n(\vec{r}) \quad (18.28)$$

where $V(\vec{r})$ is the potential due to the lattice and $n(\vec{r})$ is the electron density at point \vec{r} . In second quantization and considering that the potential due to the lattice is an infinite sum of contributions from ions located at points \vec{R}_l we can write that

$$H_{int} = \sum_l \int d^3r V(\vec{r} - \vec{R}_l) c^\dagger(\vec{r}) c(\vec{r}) \quad (18.29)$$

(we have omitted the spin dependence, for simplicity). Introducing a Fourier decomposition of the electron operators

$$c(\vec{r}) = \sum_{\vec{k}} e^{i\vec{k} \cdot \vec{r}} c_{\vec{k}}$$

and writing that the positions of the ions are given by a vector that refers to the regular lattice position, \vec{l} , and a small displacement around the equilibrium position, \vec{y}_l , as $\vec{R}_l = \vec{l} + \vec{y}_l$ we can write that

$$H_{int} = \sum_{\vec{k}, \vec{k}', l} c_{\vec{k}}^\dagger \langle \vec{k} | V(\vec{r} - \vec{l} - \vec{y}_l) | \vec{k}' \rangle c_{\vec{k}'} \quad (18.30)$$

The matrix element is the Fourier transform of the potential.

We may then write the full Hamiltonian as $H = H_0 + H_{int}$ where

$$H_0 = \sum_{\vec{k}} \epsilon_{\vec{k}} c_{\vec{k}}^\dagger c_{\vec{k}} + \sum_{\vec{q}, \lambda} \hbar \omega_{\vec{q}\lambda} a_{\vec{q}\lambda}^\dagger a_{\vec{q}\lambda} \quad (18.31)$$

is the sum of a free electron part and a set of harmonic oscillators where λ are the polarizations. Making a shift of variables $\vec{r} \rightarrow \vec{r} + \vec{l} + \vec{y}_l$ we get that

$$\begin{aligned} H_{int} &= \sum_{\vec{k}, \vec{k}', l} \int d^3r e^{-i\vec{k} \cdot \vec{r}} V(\vec{r} - \vec{l} - \vec{y}_l) e^{i\vec{k}' \cdot \vec{r}} c_{\vec{k}}^\dagger c_{\vec{k}'} \\ &= \sum_{\vec{k}, \vec{k}', l} e^{i(\vec{k}' - \vec{k}) \cdot (\vec{l} + \vec{y}_l)} \int d^3r e^{-i(\vec{k} - \vec{k}') \cdot \vec{r}} V(\vec{r}) c_{\vec{k}}^\dagger c_{\vec{k}'} \\ &= \sum_{\vec{k}, \vec{k}', l} e^{i(\vec{k}' - \vec{k}) \cdot (\vec{l} + \vec{y}_l)} V_{\vec{k} - \vec{k}'} c_{\vec{k}}^\dagger c_{\vec{k}'} \end{aligned} \quad (18.32)$$

Since the displacement of the ions from their equilibrium positions is small we may expand

$$\begin{aligned} e^{i(\vec{k}' - \vec{k}) \cdot \vec{y}_l} &= 1 + i(\vec{k}' - \vec{k}) \cdot \vec{y}_l \\ &= 1 + \frac{i}{\sqrt{N}} (\vec{k}' - \vec{k}) \cdot \sum_{\vec{q}} e^{i\vec{q} \cdot \vec{l}} \vec{y}_{\vec{q}} \end{aligned} \quad (18.33)$$

The interaction term may then be split into two terms: one due to the coupling to the regular lattice and one due to the coupling with the phonons. The term due to the regular lattice is written as

$$\begin{aligned} H_{Bloch} &= \sum_{\vec{k}, \vec{k}', l} e^{i(\vec{k}' - \vec{k}) \cdot \vec{l}} V_{\vec{k} - \vec{k}'} c_{\vec{k}}^{\dagger} c_{\vec{k}'} \\ &= N \sum_{\vec{k}, \vec{g}} V_{\vec{g}} c_{\vec{k} + \vec{g}}^{\dagger} c_{\vec{k}} \end{aligned} \quad (18.34)$$

where \vec{g} is a reciprocal lattice vector. This is the usual term that reflects the influence of the lattice on the free electron bands that leads to the band structure of the nearly-free electron model. Far from the borders of the Brillouin zone its influence may be neglected. In other words, if we are far from Bragg reflections we may neglect its effect. The second term is the electron-phonon interaction term that most concerns us here. It may be written as

$$\begin{aligned} H_{e-p} &= \frac{i}{\sqrt{N}} \sum_{\vec{k}, \vec{k}', l, \vec{q}} e^{i(\vec{k}' - \vec{k} + \vec{q}) \cdot \vec{l}} (\vec{k}' - \vec{k}) \cdot \vec{y}_{\vec{q}} V_{\vec{k} - \vec{k}'} c_{\vec{k}}^{\dagger} c_{\vec{k}'} \\ &= i\sqrt{N} \sum_{\vec{k}, \vec{k}'} (\vec{k}' - \vec{k}) \cdot \vec{y}_{\vec{k} - \vec{k}'} V_{\vec{k} - \vec{k}'} c_{\vec{k}}^{\dagger} c_{\vec{k}'} \\ &= i \sum_{\vec{k}, \vec{k}', \vec{\lambda}} \left(\frac{N\hbar}{2M\omega_{\vec{k} - \vec{k}', \vec{\lambda}}} \right)^{1/2} (\vec{k}' - \vec{k}) \cdot \vec{\lambda} V_{\vec{k} - \vec{k}'} \left(a_{\vec{k}' - \vec{k}, \vec{\lambda}}^{\dagger} + a_{\vec{k} - \vec{k}', \vec{\lambda}} \right) c_{\vec{k}}^{\dagger} c_{\vec{k}'} \end{aligned} \quad (18.35)$$

The dot product selects modes that are longitudinal (polarization and momentum difference parallel). Neglecting the Bloch term yields Fröhlich Hamiltonian

$$H = \sum_{\vec{k}, \sigma} \epsilon_{\vec{k}} c_{\vec{k}\sigma}^{\dagger} c_{\vec{k}\sigma} + \sum_{\vec{q}} \hbar\omega_{\vec{q}} a_{\vec{q}}^{\dagger} a_{\vec{q}} + \sum_{\vec{k}, \vec{k}', \sigma} M_{\vec{k}\vec{k}'} \left(a_{-\vec{q}}^{\dagger} + a_{\vec{q}} \right) c_{\vec{k}\sigma}^{\dagger} c_{\vec{k}', \sigma} \quad (18.36)$$

with

$$M_{\vec{k}\vec{k}'} = i\sqrt{\frac{N\hbar}{2M\omega_{\vec{q}}}} |\vec{k}' - \vec{k}| V_{\vec{k} - \vec{k}'} \quad (18.37)$$

Here $\vec{q} = \vec{k} - \vec{k}'$.

In the momentum representation we defined the matrix element $V_{\vec{k}, \vec{k}'}$. This is the matrix element of the electron-electron interaction where in the initial time (*I*) the electrons are in states (\vec{k}, \uparrow) , $(-\vec{k}, \downarrow)$ and in the final time (*II*) the electrons are in states (\vec{k}', \uparrow) , $(-\vec{k}', \downarrow)$. In addition to the Coulomb repulsion term, the matrix element contains two terms resulting from the interaction with the phonons. The interaction with the phonons implies that an electron may create a phonon or absorb a phonon. An effective interaction between two electrons is obtained considering second order terms by considering that one phonon emitted by one electron is absorbed by another electron. This leads to the two terms represented in the figure. The process then involves the exchange of a phonon between two electrons. The matrix element has the structure

$$\begin{aligned} \langle I|V|II \rangle &= \int d^3r_1 \int d^3r_2 e^{-i\vec{k} \cdot (\vec{r}_1 - \vec{r}_2)} V(\vec{r}_1 - \vec{r}_2) e^{i\vec{k}' \cdot (\vec{r}_1 - \vec{r}_2)} \\ &= \int d^3r V(\vec{r}) e^{i\vec{q} \cdot \vec{r}} = V(\vec{q}) \end{aligned}$$

where $\vec{q} = \vec{k}' - \vec{k}$ is the relative momentum. As a function of time two processes may occur: in process (i) the electron with momentum \vec{k} emits a phonon with momentum $-\vec{q}$ and stays in a state with momentum $\vec{k} + \vec{q}$. At a later time the phonon is absorbed by the electron with momentum $-\vec{k}$ which then stays in a state with momentum $-\vec{k} - \vec{q}$. In the other process (ii) first the electron with momentum $-\vec{k}$ emits a phonon with momentum \vec{q} and stays in a state with momentum $-\vec{k} - \vec{q}$. Later the other electron with momentum \vec{k} absorbs the phonon and stays in a state with momentum $\vec{k} + \vec{q}$. The initial and final states of both processes are the same but the intermediate steps are different. The initial state has energy $E_I = 2\zeta_k$ where $\zeta_k = \frac{\hbar^2 k^2}{2m} - E_F$. The final state has energy $E_{II} = 2\zeta_{k'}$. The two intermediate states have also equal energies $E_i = E_{ii} = \zeta_{k'} + \zeta_k + \hbar\omega_q$. Here ω_q is the energy of the phonon of momentum q .

To second order the matrix element of the transition between the two states is given by

$$\begin{aligned} \langle I|H_{el-ph}|II \rangle &= \sum_j \langle I|V|j \rangle \frac{1}{2} \left(\frac{1}{E_{II} - E_j} + \frac{1}{E_I - E_j} \right) \langle j|V|II \rangle \\ &= \frac{|W_q|^2}{\hbar} \left(\frac{1}{\omega - \omega_q} - \frac{1}{\omega + \omega_q} \right) \end{aligned} \quad (18.38)$$

The sum is over the intermediate states we defined $W_q = \langle j|V|II \rangle$ assumed that $W_p^* = \langle I|V|j \rangle$, and defined $\hbar\omega = \zeta_{k'} - \zeta_k$. We obtain finally that

$$\langle I|H|II \rangle = V_q + \frac{2|W_q|^2}{\hbar} \frac{\omega_q}{\omega^2 - \omega_q^2} \quad (18.39)$$

The first term is the direct repulsive Coulomb repulsion between the electrons and the second term is the term due to the mediated electron-electron interaction via the electron-phonon interaction. If $\omega < \omega_q$ it is possible that the total interaction is attractive.

We may as well recall that actually in the solid the interaction between the electrons is screened. Actually, the interaction term should be of the form

$$\langle I|H|II \rangle = \frac{4\pi e^2}{q^2 \epsilon(q, \omega)}$$

Let us retrace steps that we took earlier. Let us assume that small external charges are added to the system with a density

$$\delta\rho(\vec{r}) = \delta\rho \cos(\vec{q} \cdot \vec{r}) e^{i\omega t} + c.c.$$

Then there will appear screening charges

$$\rho(\vec{r}) = \rho \cos(\vec{q} \cdot \vec{r}) e^{i\omega t} + c.c.$$

The dielectric constant is given by the ratio of the external charge by the total charge

$$\epsilon(\vec{q}, \omega) = \frac{\delta\rho}{\rho + \delta\rho}$$

The charge has two contributions one from the electrons and the other from the ions $\rho = \rho_e + \rho_i$. Poisson's equation can be written as

$$\nabla^2 V = -4\pi(\delta\rho + \rho_e + \rho_i)$$

The equation of motion of the ions in the presence of the electric field $\vec{E} = -\vec{\nabla}V$ is given by

$$M \frac{dj_i}{dt} = nZe^2 E$$

where j_i is the ionic current and Ze the charge of each ion. The continuity equation of the ionic charge is

$$\frac{\partial \rho_i}{\partial t} + \vec{\nabla} \cdot \vec{j}_i = 0$$

Assuming small oscillations of the ions we can take $(dj_i)/(dt) \sim (\partial j_i)/(\partial t)$. Then

$$\frac{\partial^2 \rho_i}{\partial t^2} = \frac{nZe^2}{M} \nabla^2 V$$

We conclude therefore that

$$\omega^2 \rho_i(q, \omega) = \omega_i^2 [\rho_i(q, \omega) + \rho_e(q, \omega) + \delta \rho(q, \omega)]$$

Here

$$\omega_i = \sqrt{\frac{4\pi nZe^2}{M}}$$

The two characteristic energies (or frequencies) of the electrons are E_F/\hbar and the plasma frequency given by

$$\omega_p = \left(\frac{4\pi ne^2}{m} \right)^{1/2} \sim \omega_i \left(\frac{M}{m} \right)^{1/2}$$

They are both very high. On the other hand ω is small since is of order ω_i . To calculate ρ_e let us consider $\omega = 0$ (static perturbation limit). Let us use the Thomas-Fermi approximation. In this approximation the electronic density is proportional to $N(0)$ and therefore it is proportional to $(E_F + eV)^{3/2}$. Then its relative variation for small values of V is given by

$$\frac{\rho_e}{-ne} = \frac{3}{2} \frac{eV}{E_F}$$

Then fixing a certain momentum value q and using that

$$\nabla^2 V = -4\pi(\delta \rho + \rho_i + \rho_e)$$

we get that

$$\rho_e = -\frac{k_s^2}{q^2}(\rho + \delta \rho)$$

where we have defined $k_s^2 = \frac{6\pi n e^2}{E_F}$. Then

$$\rho = \rho_i + \rho_e = \left(\frac{\omega_i^2}{\omega^2} - \frac{k_s^2}{q^2} \right) (\rho + \delta\rho)$$

and therefore the dielectric constant is

$$\epsilon(q, \omega) = \frac{\omega^2(k_s^2 + q^2) - \omega_i^2 q^2}{\omega^2 q^2}$$

The spontaneous vibration modes of the system (phonons) can be obtained from $\delta\rho = 0$ or $\epsilon(q, \omega) = 0$. These are

$$\omega_q^2 = \omega_i^2 \frac{q^2}{k_s^2 + q^2}$$

which for $q \ll 1$ leads to the dispersion relation

$$\omega_q \sim v_s q \sim \frac{\omega_i}{k_s} q \quad (18.40)$$

Here $v_s = \omega_i/k_s$ is the sound velocity. Finally we obtain for the dielectric constant

$$\frac{1}{\epsilon(q, \omega)} = \frac{q^2}{k_s^2 + q^2} \left(1 + \frac{\omega_q^2}{\omega^2 - \omega_q^2} \right) \quad (18.41)$$

and therefore we obtain that

$$< I | H | II > = \frac{4\pi e^2}{q^2 \epsilon(q, \omega)} = \frac{4\pi e^2}{q^2 + k_s^2} + \frac{4\pi e^2}{q^2 + k_s^2} \frac{\omega_q^2}{\omega^2 - \omega_q^2} \quad (18.42)$$

Since ω_q is the energy exchanged which is originated in the phonons it depends on the mass of the ions. This leads to the known isotopic effect in superconductivity.

18.4 BCS theory

We have seen that the system is unstable to the creation of pairs of electrons of opposite momenta and opposite spins. Actually we have not considered the case in which the momentum of the center of mass is finite. It can however be shown

that the zero momentum case is particularly favorable requiring a smaller attractive potential.

We can therefore expect that the ground state is composed by pairs of electrons $(\vec{k}, \uparrow), (-\vec{k}, \downarrow)$. We may therefore attempt a wave function for the system of the form

$$|G\rangle = \prod_k \left(u_k + v_k c_{k,\uparrow}^\dagger c_{-k,\downarrow}^\dagger \right) |0\rangle \quad (18.43)$$

Here $|0\rangle$ is the vacuum state and u_k (v_k) are the probability amplitudes of not having (having) one pair of electrons in the state of momentum k . By normalization $u_k^2 + v_k^2 = 1$ (since there is no other possibility).

Let us consider then the Hamiltonian of the system as being

$$H = \sum_{k,\sigma} \epsilon_k c_{k,\sigma}^\dagger c_{k,\sigma} + \sum_{k,k'} V_{kk'} c_{k,\uparrow}^\dagger c_{-k,\downarrow}^\dagger c_{-k',\downarrow} c_{k',\uparrow} \quad (18.44)$$

and let us not consider other terms for not being relevant to the pairing mechanism. Let us then consider that the electrons pair up. This implies that $\langle c_{k,\uparrow}^\dagger c_{-k,\downarrow}^\dagger \rangle$ will have a non-vanishing value in the superconducting phase. In the same way $\langle c_{-k',\downarrow} c_{k',\uparrow} \rangle$ will have a non-vanishing value. Note that the average values do not conserve the number of particles. We may however use the grand canonical ensemble just by making the replacement $\epsilon_k \rightarrow \epsilon_k - \mu$.

Let us then use mean field theory and write

$$c_{-k,\downarrow} c_{k,\uparrow} = \langle c_{-k,\downarrow} c_{k,\uparrow} \rangle + (c_{-k,\downarrow} c_{k,\uparrow} - \langle c_{-k,\downarrow} c_{k,\uparrow} \rangle)$$

and the corresponding for the complex conjugate. Then we may write the Hamiltonian as

$$\begin{aligned} H = & \sum_{k,\sigma} \epsilon_k c_{k,\sigma}^\dagger c_{k,\sigma} + \sum_{k,k'} V_{kk'} \left(\langle c_{k,\uparrow}^\dagger c_{-k,\downarrow}^\dagger \rangle + (c_{k,\uparrow}^\dagger c_{-k,\downarrow}^\dagger - \langle c_{k,\uparrow}^\dagger c_{-k,\downarrow}^\dagger \rangle) \right) \\ & \left(\langle c_{-k',\downarrow} c_{k',\uparrow} \rangle + (c_{-k',\downarrow} c_{k',\uparrow} - \langle c_{-k',\downarrow} c_{k',\uparrow} \rangle) \right) \end{aligned}$$

Let us now call

$$\Delta_k = - \sum_{k'} V_{k'k} \langle c_{-k',\downarrow} c_{k',\uparrow} \rangle$$

Then the mean field Hamiltonian is given by

$$H_{MF} = \sum_{k,\sigma} (\epsilon_k - \mu) c_{k,\sigma}^\dagger c_{k,\sigma} - \sum_k \left(\Delta_k c_{k,\uparrow}^\dagger c_{-k,\downarrow}^\dagger + \Delta_k^* c_{-k,\downarrow} c_{k,\uparrow} - \Delta_k^* \langle c_{-k,\downarrow} c_{k,\uparrow} \rangle \right) \quad (18.45)$$

Let us consider now, for instance, the equation of motion of $c_{k,\uparrow}$

$$\begin{aligned} i\hbar \frac{\partial c_{k,\uparrow}}{\partial t} &= [c_{k,\uparrow}, H] \\ &= \sum_{k',\sigma} (\epsilon_{k',\sigma} - \mu) [c_{k,\uparrow}, c_{k',\sigma}^\dagger c_{k',\sigma}] - \sum_{k'} \Delta_{k'} [c_{k,\uparrow}, c_{k',\uparrow}^\dagger c_{-k,\downarrow}^\dagger] - \sum_{k'} \Delta_{k'}^* [c_{k,\uparrow}, c_{-k,\downarrow} c_{k',\downarrow}] \end{aligned}$$

Using that $[A, BC] = \{A, B\}C - B\{C, A\}$ we get

$$i\hbar \frac{\partial c_{k,\uparrow}}{\partial t} = (\epsilon_{k,\uparrow} - \mu) c_{k,\uparrow} - \Delta_k c_{-k,\downarrow}^\dagger$$

Therefore the operators $c_{k,\uparrow}$ and $c_{-k,\downarrow}^\dagger$ are coupled. It makes sense physically: a pair of momenta $(k, \uparrow), (-k, \downarrow)$ is broken either annihilating one electron (k, \uparrow) or creating an electron $(-k, \downarrow)$.

The mean field Hamiltonian is quadratic and therefore may be diagonalized performing a Bogoliubov-Valatin transformation. Let us then define the linear combination

$$\begin{aligned} \alpha_k &= u_k c_{k,\uparrow} - v_k^* c_{-k,\downarrow}^\dagger \\ \beta_{-k} &= u_k c_{-k,\downarrow} + v_k^* c_{k,\uparrow}^\dagger \end{aligned}$$

and their hermitian conjugates. It is easy to show that

$$\{\alpha_k, \alpha_{k'}^\dagger\} = \{\beta_k, \beta_{k'}^\dagger\} = \delta_{k,k'}$$

and that all other anti-commutators vanish. The previous expression implies that

$$|u_k|^2 + |v_k|^2 = 1$$

The equations for the new fermionic operators may be inverted to yield

$$\begin{aligned} c_{k,\uparrow} &= u_k \alpha_k + v_k^* \beta_{-k}^\dagger \\ c_{-k,\downarrow} &= u_k \beta_{-k} - v_k^* \alpha_k^\dagger \end{aligned}$$

Let us substitute this transformation in the effective mean field Hamiltonian. Define ($\tilde{\epsilon}_k = \epsilon_{k,\sigma} - \mu$). Keeping only terms of the type $\alpha_k^\dagger \alpha_k$, $\beta_{-k}^\dagger \beta_{-k}$ we get that

$$\begin{aligned}
 H_{MF} &= \sum_k \tilde{\epsilon}_k (|u_k|^2 - |v_k|^2) (\alpha_k^\dagger \alpha_k + \beta_{-k}^\dagger \beta_{-k}) \\
 &+ \sum_k \Delta_k (\alpha_k^\dagger \alpha_k + \beta_{-k}^\dagger \beta_{-k}) u_k^* v_k \\
 &+ \sum_k \Delta_k^* (\alpha_k^\dagger \alpha_k + \beta_{-k}^\dagger \beta_{-k}) u_k v_k^* \\
 &+ \sum_k \Delta_k^* u_k v_k^* (< \alpha_k^\dagger \alpha_k > + < \beta_{-k}^\dagger \beta_{-k} >) \\
 &+ \sum_k (\tilde{\epsilon}_k 2|v_k|^2 - \Delta_k v_k u_k^* - \Delta_k^* u_k v_k^* - \Delta_k^* u_k v_k^*)
 \end{aligned}$$

We have chosen

$$2u_k v_k \tilde{\epsilon}_k - \Delta_k u_k^2 + \Delta_k^* v_k^2 = 0$$

to annihilate the unwanted crossed terms. Also recall that $|u_k|^2 + |v_k|^2 = 1$. We can rewrite the mean field Hamiltonian as

$$\begin{aligned}
 H_{MF} &= \sum_k \alpha_k^\dagger \alpha_k (\tilde{\epsilon}_k (|u_k|^2 - |v_k|^2) + \Delta_k u_k^* v_k + \Delta_k^* u_k v_k^*) \\
 &+ \sum_k \beta_{-k}^\dagger \beta_{-k} (\tilde{\epsilon}_k (|u_k|^2 - |v_k|^2) + \Delta_k u_k^* v_k + \Delta_k^* u_k v_k^*) \\
 &+ H_0
 \end{aligned} \tag{18.46}$$

where the constant term is given by

$$H_0 = \sum_k \Delta_k^* u_k v_k^* (< \alpha_k^\dagger \alpha_k > + < \beta_{-k}^\dagger \beta_{-k} >) + \sum_k (2|v_k|^2 \tilde{\epsilon}_k - \Delta_k v_k u_k^* - 2\Delta_k^* u_k v_k^*)$$

Recall now that

$$\begin{aligned}
 \Delta_k &= - \sum_{k'} V_{kk'} < c_{-k', \downarrow} c_{k', \uparrow} > \\
 &= - \sum_{k'} V_{kk'} < (u_{k'} \beta_{-k'} - v_{k'}^* \alpha_{k'}^\dagger) (u_{k'} \alpha_{k'} + v_{k'}^* \beta_{-k'}^\dagger) > \\
 &= - \sum_{k'} V_{kk'} < (u_{k'} v_{k'}^* \beta_{-k'} \beta_{-k'}^\dagger - v_{k'}^* u_{k'} \alpha_{k'}^\dagger \alpha_{k'}) >
 \end{aligned} \tag{18.47}$$

We get therefore that

$$\Delta_k = - \sum_{k'} V_{kk'} u_{k'} v_{k'}^* [1 - 2f(E_{k'})] \quad (18.48)$$

where the energy dispersion is obtained from the diagonalized Hamiltonian as

$$E_k = \tilde{\epsilon}_k(|u_k|^2 - |v_k|^2) + \Delta_k u_k^* v_k + \Delta_k^* u_k v_k^* \quad (18.49)$$

and f is the Fermi function. A simplification occurs because u_k and v_k may be chosen as real (this is not possible if an external magnetic field is applied, as we will see later). Then

$$u_k^2 + v_k^2 = 1$$

which implies that we may choose $u_k = \cos \chi_k$ and $v_k = \sin \chi_k$. Then the condition for the amplitudes u and v can be rewritten as

$$\begin{aligned} 2u_k v_k \tilde{\epsilon}_k - \Delta_k u_k^2 + \Delta_k v_k^2 &= 0 \\ \tan 2\chi_k &= \frac{\Delta_k}{\tilde{\epsilon}_k} \end{aligned} \quad (18.50)$$

This implies that

$$\begin{aligned} \sin 2\chi_k &= \pm \frac{\Delta_k}{E_k} = 2u_k v_k \\ \cos 2\chi_k &= \pm \frac{\tilde{\epsilon}_k}{E_k} = u_k^2 - v_k^2 \end{aligned}$$

where $E_k = (\tilde{\epsilon}_k^2 + \Delta_k^2)^{1/2}$. We can also obtain that

$$\begin{aligned} u_k^2 &= \frac{1}{2} \left(1 + \frac{\tilde{\epsilon}_k}{E_k} \right) \\ v_k^2 &= \frac{1}{2} \left(1 - \frac{\tilde{\epsilon}_k}{E_k} \right) \end{aligned}$$

Clearly the Hamiltonian may be written as

$$H_{MF} = \sum_k E_k (\alpha_k^\dagger \alpha_k + \beta_{-k}^\dagger \beta_{-k}) + H_0 \quad (18.51)$$

If $\tilde{\epsilon}_k = 0$ then $E_k = \Delta_k$ and there is a gap in the excitation spectrum. Substituting the previous expressions in the equation for the gap function we get that

$$\Delta_k = - \sum_{k'} \frac{V_{kk'} \Delta_{k'}}{2E_{k'}} (1 - 2f(E_{k'})) \quad (18.52)$$

This is called the gap equation.

18.5 Critical temperature

Let us take the gap near zero $\Delta_k \rightarrow 0$ and therefore $E_k = \tilde{\epsilon}_k$. Then the gap equation becomes

$$\Delta_k = - \sum_{k'} \frac{V_{kk'} \Delta_{k'}}{2|\tilde{\epsilon}_{k'}|} (1 - 2f(|\tilde{\epsilon}_{k'}|))$$

Let us choose a simple form for the interaction like $V_{kk'} = -V$ if $|\tilde{\epsilon}_k|, |\tilde{\epsilon}_{k'}| < \hbar\omega_D$, and $V_{kk'} = 0$, otherwise. Then the gap equation is rewritten as

$$\Delta_k = V \sum_{k', |\tilde{\epsilon}_{k'}| < \hbar\omega_D} \Delta_{k'} \frac{1 - 2f(|\tilde{\epsilon}_{k'}|)}{2|\tilde{\epsilon}_{k'}|} \quad (18.53)$$

If $\Delta_k \rightarrow 0$ then it is independent of k and gives

$$1 = VN(0) \int_0^{\hbar\omega_D} d\epsilon \frac{1 - 2f(\epsilon)}{\epsilon}$$

The Fermi function is in this case calculated at the critical temperature $T = T_c$. Defining $x = \epsilon/(k_B T_c)$ and integrating by parts we obtain that

$$\frac{1}{VN(0)} = \ln \frac{\hbar\omega_D}{k_B T_c} (1 - 2f(\hbar\omega_D)) + 2 \int_0^{\frac{\hbar\omega_D}{k_B T_c}} dx \ln x \frac{d}{dx} \left(\frac{1}{e^x + 1} \right)$$

Neglecting $f(\hbar\omega_D) \ll 1$ we get that

$$\frac{1}{VN(0)} = \ln \left(A \frac{\hbar\omega_D}{k_B T_c} \right)$$

where

$$\ln A = 2 \int_0^\infty dx \ln x \frac{d}{dx} \left(\frac{1}{e^x + 1} \right)$$

($A = \frac{2\gamma}{\pi} \sim 1.13$ where γ is Euler's constant). We obtain finally that

$$k_B T_c = 1.13 \hbar \omega_D e^{-\frac{1}{V N(0)}} \quad (18.54)$$

Once again let us emphasize that the result is non-perturbative in the potential.

18.6 Gap

Consider once again the gap equation

$$\Delta_k = - \sum_{k'} V_{kk'} \frac{\Delta_{k'}}{2E_{k'}} [1 - 2f(E_{k'})]$$

and the same form for the potential as above. Then the gap is independent of the momentum and we get

$$\begin{aligned} 1 &= N(0)V \int_0^{\hbar\omega_D} \frac{d\epsilon}{E} [1 - 2f(E)] \\ E &= (\epsilon^2 + \Delta^2)^{1/2} \end{aligned}$$

Consider now that $\Delta, k_B T_c \ll \hbar\omega_D$. Then we obtain that

$$\frac{1}{N(0)V} = \sinh^{-1} \frac{\hbar\omega_D}{\Delta} + 2 \int_0^\infty d\epsilon \frac{\partial f(E)}{\partial \epsilon} \sinh^{-1} \frac{\epsilon}{\Delta}$$

Integrating by parts we get as before

$$\frac{1}{N(0)V} \sim \ln \frac{2\hbar\omega_D}{\Delta} + F\left(\frac{\Delta}{k_B T}\right)$$

Using the expression for the critical temperature we may write that

$$\Delta(T) = \frac{2}{1.13} k_B T_c e^{F\left(\frac{\Delta}{k_B T}\right)} \quad (18.55)$$

If Δ is finite then when $T \rightarrow 0$ the Fermi-Dirac function tends to zero and $F \rightarrow 0$. Then

$$\Delta(0) = 2\hbar\omega_D e^{-\frac{1}{N(0)V}}$$

and

$$\frac{\Delta(T)}{\Delta(0)} = e^{F\left(\frac{\Delta}{k_B T}\right)}$$

The ratio starts at one at zero temperature and decreases to zero at the critical temperature. Near T_c we get

$$\frac{\Delta(T)}{\Delta(0)} = 1.74\left(1 - \frac{T}{T_c}\right)^{1/2}$$

In our notation of the critical exponents we get $\beta = 1/2$ for the dependence of the order parameter, here taken as the gap function, as a function of the temperature difference to the critical temperature. This is obviously the mean field result as expected since the whole treatment was based upon a mean field approximation of the Hamiltonian. At the same time this result suggests a description of the superconducting phase transition by a functional in which the free energy is expanded in powers of the order parameter. This is actually possible, as we shall see later, and it is called the Ginzburg-Landau functional. Anticipating the later treatment it can be obtained expanding the gap function in powers of Δ and equalizing it to the stationary condition for the free energy.

18.7 Thermodynamics

The Hamiltonian was obtained as

$$H = \sum_{k,\alpha} E_k \gamma_{k,\alpha}^\dagger \gamma_{k,\alpha} + H_0$$

calling now $\gamma_{k,+} = \alpha_k$ and $\gamma_{k,-} = \beta_{-k}$. The average energy is given by

$$U = \langle H \rangle = \sum_{k,\alpha} E_k f(E_k) + H_0 = 2 \sum_k E_k f(E_k) + H_0$$

The specific heat per unit volume is then obtained from

$$c_V = \frac{2}{V} \sum_k E_k \left(\frac{\partial f(E_k)}{\partial T} \right)_V$$

Using that

$$\frac{\partial f(E)}{\partial T} = \frac{\partial f(E)}{\partial E} \frac{\partial E}{\partial T} + \frac{\partial f}{\partial T}$$

we obtain that

$$c_V = -\frac{2}{VT} N(0) \int_{-\infty}^{\infty} d\epsilon (E^2 + \frac{\beta}{2} \frac{\partial \Delta^2}{\partial \beta}) \frac{\partial f(E)}{\partial E} \quad (18.56)$$

Using that

$$\frac{\partial f(E)}{\partial E} = -\frac{1}{k_B T} f(E)(1 - f(E))$$

the specific heat is given by

$$c_V = 2k_B \beta^2 N(0) \int_0^{\infty} d\epsilon f(E)(1 - f(E)) [E^2 - T \Delta \frac{d\Delta}{dT}]$$

At low temperatures $\Delta \frac{d\Delta}{dT}$ is negligible and the specific heat is obtained from

$$\begin{aligned} c_V &\sim 2\beta^2 k_B N(0) \Delta^2 \int_0^{\infty} d\epsilon e^{-\beta \sqrt{\epsilon^2 + \Delta^2}} \\ &\sim 2\beta^2 k_B N(0) \Delta^2 e^{-\beta \Delta} \int_0^{\infty} d\epsilon e^{-\beta \frac{\epsilon^2}{2\Delta}} \end{aligned}$$

we see that at low temperatures the specific heat is exponentially small due to the presence of the gap in the spectrum. The difference between the superconducting and the normal phase specific heats is

$$c_s - c_n = k_B N(0) \beta_c^2 \left(\frac{d\Delta^2}{d\beta} \right)_{T_c} = 10.2 k_B^2 T_c N(0)$$

It is therefore discontinuous at the critical temperature as shown in the figure.

18.8 Bogoliubov-de Gennes equations

Let us consider now an alternative way to solve the problem that is easily generalizable to more complex situations (like the inclusion of magnetic fields, the addition of disorder effects, etc.). Let us consider a description in the real space, and let us

consider creation and destruction operators in the real space. The Hamiltonian of the electrons with an interaction may now be written as

$$\begin{aligned}
 H &= \int d^3r \sum_{\alpha} \psi^{\dagger}(\vec{r}, \alpha) \left(\frac{(\vec{p} - \frac{e}{c} \vec{A})^2}{2m} + U_0(\vec{r}) \right) \psi(\vec{r}, \alpha) \\
 &- \frac{1}{2} V \int d^3r \sum_{\alpha, \beta} \psi^{\dagger}(\vec{r}, \alpha) \psi^{\dagger}(\vec{r}, \beta) \psi(\vec{r}, \beta) \psi(\vec{r}, \alpha)
 \end{aligned} \tag{18.57}$$

Considering a local interaction between the electrons and considering a local potential $U_0(\vec{r})$ and the eventual action of an external magnetic field (\vec{A} is the vector potential). Let us then consider a local pairing of electrons with opposite spins (in this case we are considering a local attractive interaction that corresponds to a Fourier transform independent of the momentum, as before). Then the mean field Hamiltonian is given by

$$\begin{aligned}
 H_{MF} &= \int d^3r \left[\sum_{\alpha} \psi^{\dagger}(\vec{r}, \alpha) H_e(\vec{r}) \psi(\vec{r}, \alpha) + U(\vec{r}) \psi^{\dagger}(\vec{r}, \alpha) \psi(\vec{r}, \alpha) \right. \\
 &\quad \left. + \Delta(\vec{r}) \psi^{\dagger}(\vec{r}, \uparrow) \psi^{\dagger}(\vec{r}, \downarrow) + \Delta^*(\vec{r}) \psi(\vec{r}, \downarrow) \psi(\vec{r}, \uparrow) \right]
 \end{aligned} \tag{18.58}$$

where

$$\begin{aligned}
 \Delta(\vec{r}) &= -V \langle \psi(\vec{r}, \downarrow) \psi(\vec{r}, \uparrow) \rangle = V \langle \psi(\vec{r}, \uparrow) \psi(\vec{r}, \downarrow) \rangle \\
 U(\vec{r}) &= -V \langle \psi^{\dagger}(\vec{r}, \uparrow) \psi(\vec{r}, \uparrow) \rangle = -V \langle \psi^{\dagger}(\vec{r}, \downarrow) \psi(\vec{r}, \downarrow) \rangle \\
 H_e(\vec{r}) &= \frac{1}{2m} (-i\hbar \vec{\nabla} - \frac{e}{c} \vec{A})^2 + U_0(\vec{r}) - E_F
 \end{aligned} \tag{18.59}$$

This Hamiltonian can be diagonalized. Let us then write

$$\begin{aligned}
 \psi(\vec{r}, \uparrow) &= \sum_n \left(\gamma_{n, \uparrow} u_n(\vec{r}) - \gamma_{n, \downarrow}^{\dagger} v_n^*(\vec{r}) \right) \\
 \psi(\vec{r}, \downarrow) &= \sum_n \left(\gamma_{n, \downarrow} u_n(\vec{r}) + \gamma_{n, \uparrow}^{\dagger} v_n^*(\vec{r}) \right)
 \end{aligned}$$

The new operators are fermionic as well. Since this transformation diagonalizes the Hamiltonian we can write it as

$$H_{MF} = E_g + \sum_{n, \alpha} \epsilon_n \gamma_{n, \alpha}^{\dagger} \gamma_{n, \alpha} \tag{18.60}$$

where E_g is the energy of the ground state and (n, α) are the excitations of the system with energies ϵ_n (we are here neglecting the Zeeman term). It is then obvious that the following relations hold

$$[H_{MF}, \gamma_{n,\alpha}] = -\epsilon_n \gamma_{n,\alpha}$$

$$[H_{MF}, \gamma_{n,\alpha}^\dagger] = \epsilon_n \gamma_{n,\alpha}^\dagger$$

Let us then calculate the commutators

$$[\psi(\vec{r}, \uparrow), H_{MF}] = [H_e + U]\psi(\vec{r}, \uparrow) + \Delta(\vec{r})\psi^\dagger(\vec{r}, \downarrow)$$

$$[\psi(\vec{r}, \downarrow), H_{MF}] = [H_e + U]\psi(\vec{r}, \downarrow) - \Delta^*(\vec{r})\psi^\dagger(\vec{r}, \uparrow)$$

Let us introduce the transformation of the fields $\psi(\vec{r}, \alpha)$ in terms of the operators of the excitations of the system $\gamma_{n,\alpha}$ and hermitian conjugate. We obtain comparing the coefficients of the same operators the Bogoliubov-de Gennes equations

$$\begin{aligned} [H_e + U]u(\vec{r}) + \Delta(\vec{r})v(\vec{r}) &= \epsilon u(\vec{r}) \\ -[H_e^* + U]v(\vec{r}) + \Delta^*(\vec{r})u(\vec{r}) &= \epsilon v(\vec{r}) \end{aligned} \quad (18.61)$$

The functions $u(\vec{r})$ and $v(\vec{r})$ are the eigenvectors of these equations and ϵ are the eigenvalues. We have reduced the problem to the solution of a system of eigenvalues. This way to solve the problem is quite useful in many situations. The solution of these equations allows the calculation of all physical quantities, even in complex situations. In general these equations will have to be solved numerically.

18.9 Flux quantization

Os supercondutores são caracterizados por várias propriedades. Uma consequência da estrutura do espectro de excitações é a robustez do sistema a estímulos exteriores de baixa energia, menor do que o gap. De certa forma, as quasipartículas estão protegidas e formam estados estacionários de "partículas" sem interações que se propagam no sistema. São por isso imunes a pequenas perturbações como impurezas, etc. Na aplicação de um campo eléctrico exterior movem-se sem resistência, e por isso a designação de supercondutores: experimentalmente foi determinado que ao se baixar a temperatura estes sistemas transitam para um estado de resistência nula,

tal que correntes estabelecidas num circuito fechado permanecem sem decaimento aparente se o sistema transitar para a fase supercondutora.

Uma solução possível do gap corresponde a uma solução uniforme no sistema e portanto a coerência da fase a longas distâncias (recorde que em geral a função do gap é complexa). Pensando num campo que representa a densidade de electrões emparelhados nestes sistemas, a coerência de fase implica um campo bem definido no espaço.

Outra propriedade importante destes sistemas é que na presença de um campo magnético este é expelido do material (efeito de Meissner) e portanto os supercondutores constituem em muitos casos diamagnetos perfeitos. Alguns supercondutores (designados de tipo-II) permitem a penetração de linhas de campo magnético no seu interior, originando correntes que circulam em torno das linhas de campo magnético (vórtices) para criarem campos magnéticos interiores que blindam o campo magnético aplicado. A coerência de fase implica que quaisquer linhas de campo magnético que penetrem o material devem ter o fluxo magnético quantificado. Na verdade se recordarmos o efeito de Aharonov-Bohm e considerarmos uma geometria em que se tem, por exemplo um anel supercondutor e no interior do orifício do anel fizermos passar linhas de campo magnético, como a função de onda dos pares de electrões (designados de pares de Cooper) é uma função bem definida, se fizermos uma circulação de um par ao longo do anel a função de onda deve ter o mesmo valor. Mas uma circulação de um par de carga $q = 2e$ ao longo de um dado caminho envolve uma fase do tipo

$$\frac{q}{\hbar c} \Phi \quad (18.62)$$

em que Φ é o fluxo do campo magnético que atravessa a área definida pela trajectória. Para que a função de onda seja invariante é necessário que

$$\frac{q}{\hbar c} \Phi = 2\pi n$$

Então o fluxo magnético deve ser dado por

$$\Phi = \frac{2\pi\hbar c}{q} n \quad (18.63)$$

em que n é um inteiro. O fluxo fica quantificado em unidades do quantum de fluxo

$$\Phi = \Phi_0 n \quad (18.64)$$

em que $\Phi_0 = hc/2e$.

18.10 Tunneling in macroscopic systems: Josephson effect

O efeito de túnel foi discutido anteriormente no contexto de sistemas microscópicos. A probabilidade de efeito de túnel decresce exponencialmente com a extensão espacial da barreira e depende essencialmente da diferença de energia para o topo da barreira de potencial. Depende também da massa da partícula e tirando sistemas de massas muito reduzidas tal probabilidade é muito pequena. Na verdade parece ser um fenómeno associado a sistemas microscópicos e que não parece ser possível observar em sistemas macroscópicos.

Os supercondutores são sistemas com várias propriedades singulares algumas das quais foram abordadas na secção anterior. Estes materiais apresentam a possibilidade de um sistema macroscópico ter um comportamento que se pode interpretar como efeito de túnel através de uma barreira, tal como uma partícula na presença de uma barreira de potencial. Na verdade, também é possível observar num supercondutor fenómenos de interferência como em sistemas microscópicos.

Como vimos é possível que em determinadas condições os electrões possam organizar-se em pares e criar um estado colectivo. Visto que um electrão é um fermião, dois electrões têm propriedades semelhantes às de um bóson. Como mencionámos anteriormente, os bósons podem condensar num estado em que existe uma ocupação macroscópica do estado de energia mínima e portanto podemos associar à ocupação deste estado uma grandeza de natureza macroscópica. Ou seja, devemos poder associar ao número de pares no supercondutor uma grandeza de natureza macroscópica. Na verdade, um método alternativo de descrever o estado de um supercondutor consiste em usar uma função de onda variacional da forma

$$|\psi\rangle \sim \prod_{\mathbf{k}} \left(u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k},\uparrow}^{\dagger} c_{-\mathbf{k},\downarrow}^{\dagger} \right) |0\rangle \quad (18.65)$$

em que se assume o caso favorável em que os electrões se emparelham preferencialmente em estados de momentos e spins opostos (designados supercondutores convencionais) e em que se pode interpretar os coeficientes $u_{\mathbf{k}}$ e $v_{\mathbf{k}}$ como parâmetros variacionais. O estado $|0\rangle$ é o estado de vácuo em que não existem electrões no sistema. O estado fundamental corresponde a uma ocupação macroscópica dos estados em que os electrões estão emparelhados.

Consideremos agora um supercondutor dividido em duas regiões separadas por um

isolante de espessura pequena. Pelas considerações anteriores será de esperar que possa haver efeito de túnel dos electrões não-emparelhados através do isolante. Será também de esperar que possa ocorrer efeito de túnel dos pares de electrões através do isolante, visto continuarmos nas condições de um sistema microscópico. Este efeito designa-se de efeito de Josephson. Designemos de ψ_1 a amplitude de probabilidade dos pares de electrões num dos lados da junção e de ψ_2 do outro lado. A concentração de pares $n = \psi^* \psi$ é assumida como constante. Podemos então escrever que

$$\psi = n^{\frac{1}{2}} e^{i\theta} \quad (18.66)$$

visto que a densidade n corresponde aos pares. Vamos agora considerar que

$$\begin{aligned} \psi_1 &= n_1^{\frac{1}{2}} e^{i\theta_1} \\ \psi_2 &= n_2^{\frac{1}{2}} e^{i\theta_2} \end{aligned} \quad (18.67)$$

em que em geral a fase pode ser diferente através da junção isolante. Considerando que pode haver efeito de túnel através da junção podemos estabelecer duas equações de Schrödinger para as amplitudes na forma

$$\begin{aligned} i\hbar \frac{\partial \psi_1}{\partial t} &= \hbar T \psi_2 \\ i\hbar \frac{\partial \psi_2}{\partial t} &= \hbar T \psi_1 \end{aligned} \quad (18.68)$$

em que $\hbar T$ resulta da transferência de pares de electrões através da junção (considera-se a situação em que todos os electrões estão emparelhados correspondendo ao estado fundamental do sistema). Definindo $\delta = \theta_2 - \theta_1$ a diferença de fase entre os dois supercondutores, obtém-se que

$$\begin{aligned} \frac{\partial n_1}{\partial t} &= 2T (n_1 n_2)^{\frac{1}{2}} \sin \delta \\ \frac{\partial n_2}{\partial t} &= -2T (n_1 n_2)^{\frac{1}{2}} \sin \delta \\ \frac{\partial \theta_1}{\partial t} &= -T \left(\frac{n_2}{n_1} \right)^{\frac{1}{2}} \cos \delta \\ \frac{\partial \theta_2}{\partial t} &= -T \left(\frac{n_1}{n_2} \right)^{\frac{1}{2}} \cos \delta \end{aligned} \quad (18.69)$$

Como a corrente é proporcional à variação da densidade temos que

$$J = J_0 \sin \delta \quad (18.70)$$

e portanto é função da diferença de fase entre as amplitudes de probabilidade de pares nos dois supercondutores. Depende portanto de um parâmetro associado a uma grandeza macroscópica, visto descrever um estado de muitas partículas.

É possível, em determinadas circunstâncias, construir um sistema em que a energia do sistema depende do parâmetro δ de uma forma que se assemelha a uma barreira de potencial. Ou seja, existem mínimos locais da energia em função de δ separados por máximos locais. Em função de δ a energia tem um mínimo local, tal que é favorável ao sistema assumir uma configuração em que a diferença de fases entre os dois supercondutores toma o valor correspondente ao mínimo local desta energia potencial (δ_0). Variando o parâmetro δ de uma grandeza infinitesimal a energia potencial aumenta e portanto o sistema deverá assumir o valor δ_0 . No entanto para variações finitas de δ será possível que o sistema adopte outros valores de δ por lhes corresponder uma energia mais favorável. Para que tal aconteça é necessário ocorrer efeito de túnel entre as duas configurações. Note no entanto que este efeito de túnel corresponde a um sistema macroscópico e não a um sistema microscópico: a variável δ (e as variáveis θ_1 e θ_2) correspondem a um parâmetro que descreve um estado colectivo (coerente) de muitas partículas. Este efeito de túnel foi observado experimentalmente.

Podemos agora considerar um sistema numa geometria mais complexa em que se tem duas junções entre um supercondutor e um isolante como indicado na figura. Considera-se um fio supercondutor onde se introduzem dois isolantes e portanto construímos duas junções de Josephson. Para além disso consideremos que no interior do circuito definido pelas duas junções se aplica um campo magnético perpendicular, para simplificar. Mesmo na ausência das junções surge o efeito de Aharonov-Bohm, como vimos anteriormente. Neste caso a circulação da fase está relacionada com o fluxo do campo magnético. Como vimos antes, devemos ter que

$$\hbar c \nabla \theta = q \mathbf{A} \quad (18.71)$$

e portanto a presença do campo magnético implica uma fase que depende do espaço. Neste caso a carga das partículas é dada por $q = -2e$, visto as partículas transportadas ao longo do circuito que inclui o fluxo magnético serem os pares de Cooper. Mas como a amplitude de probabilidade é uma grandeza macroscópica mensurável,

deve ser bem definida e por isso a diferença de fase ao longo da trajectória deve ser um múltiplo de 2π . Ou seja,

$$\int_C \nabla\theta \cdot d\mathbf{l} = 2\pi m \quad (18.72)$$

Portanto o fluxo deve ser quantificado na forma

$$\Phi = \int_C \mathbf{A} \cdot d\mathbf{l} = \frac{\hbar c}{2e} m \quad (18.73)$$

em que C é um contorno fechado e m é inteiro.

Regressando às duas junções de Josephson obtemos que a circulação da fase ao longo do circuito fechado deve ser dada por

$$\int_C \nabla\theta \cdot d\mathbf{l} = \frac{2e}{\hbar c} \Phi \quad (18.74)$$

Chamemos de δ_a a diferença de fase entre os dois supercondutores ao longo de uma junção e δ_b ao longo da outra junção. Temos então que

$$\delta_b - \delta_a = \frac{2e}{\hbar c} \Phi \quad (18.75)$$

e portanto podemos escrever que

$$\begin{aligned} \delta_a &= \delta_0 - \frac{e}{\hbar c} \Phi \\ \delta_b &= \delta_0 + \frac{e}{\hbar c} \Phi \end{aligned} \quad (18.76)$$

A corrente total é dada pela soma das correntes através de cada junção. Portanto

$$J = J_0 \sin\left(\delta_0 - \frac{e}{\hbar c} \Phi\right) + J_0 \sin\left(\delta_0 + \frac{e}{\hbar c} \Phi\right) \quad (18.77)$$

A corrente é então função do fluxo exterior e oscila em função do campo magnético

$$J = 2J_0 \sin \delta_0 \cos \frac{e\Phi}{\hbar c} \quad (18.78)$$

Estas oscilações resultam da interferência entre as duas junções e portanto resultam de uma interferência entre dois sistemas macroscópicos.

18.11 Interface and surface effects