

## First order correction $\hat{H}_1$

Now, let us evaluate the first order correction term. Clearly, in the perturbative approach, the **first order correction** term is obtained just by taking the matrix element of the potential energy term between the non-interacting Fermi ground state:

$$E^{(1)} = \langle F | \hat{H}_1 | F \rangle = \frac{e^2}{2V} \sum_{\vec{k}\vec{p}\vec{q}}' \sum_{\lambda\mu} \frac{4\pi}{q^2} \langle F | a_{\vec{k}+\vec{q},\lambda}^\dagger a_{\vec{p}-\vec{q},\mu}^\dagger a_{\vec{p}\mu} a_{\vec{k}\lambda} | F \rangle$$

where we have

$$a_{\vec{p}\mu} a_{\vec{k}\lambda} | F \rangle \neq 0 \quad \text{only if } |\vec{k}| \leq k_F, |\vec{p}| \leq k_F$$

and

$$\langle F | a_{\vec{k}+\vec{q},\lambda}^\dagger a_{\vec{p}-\vec{q},\mu}^\dagger = \left( a_{\vec{p}-\vec{q},\mu} a_{\vec{k}+\vec{q},\lambda} | F \rangle \right)^\dagger \neq 0 \quad \text{only if } |\vec{k} + \vec{q}| \leq k_F, |\vec{p} - \vec{q}| \leq k_F$$

where the condition is such that the wave vectors are inside the Fermi sphere. Hence the matrix element  $\langle F | \hat{H}_1 | F \rangle$  can be different from zero only if certain conditions apply. In particular, the first two destruction operators  $a_{\vec{p}\mu} a_{\vec{k}\lambda}$  destroy two particles while the other two creation operators  $a_{\vec{k}+\vec{q},\lambda}^\dagger a_{\vec{p}-\vec{q},\mu}^\dagger$  create two particles and, in order that the action of these last two creation operators gives a result different from zero, they should act on an empty state. More specifically, the matrix element can be different from zero only if the two creation operators exactly fill up the holes made by the two destruction operators, namely two particles are destroyed and then two additional particles are created in exactly the same state inside the Fermi sphere. By considering the wave vector index of these operators, there are only two possibilities:

$$\begin{cases} \vec{k} + \vec{q} = \vec{k} & (\lambda = \lambda) \\ \vec{p} - \vec{q} = \vec{p} & (\mu = \mu) \end{cases} \quad \begin{cases} \vec{k} + \vec{q} = \vec{p} & (\lambda = \mu) \\ \vec{p} - \vec{q} = \vec{k} & (\mu = \lambda) \end{cases}$$

where the contribution on the left is called **direct term** while the one on the right **exchange term**. The direct term realizes only if  $\vec{q} = 0$ , but in the sum  $\sum'$  the  $\vec{q} = 0$  term is not considered. Hence, in our specific case, the condition in the direct term is not possible. On the contrary, the condition of the exchange term can be realized. It is called exchange term, because if we interpreted  $\vec{q}$  as a wave vector, or momentum, there is momentum transferred from one particle to another one. Moreover, this case is applicable if we consider interactions only between electrons with equal spins,  $\delta_{\lambda,\mu}$ . In practice, this means that the only possibility is:

$$\begin{aligned} E^{(1)} &= \frac{e^2}{2V} \sum_{\vec{k}\vec{q},\lambda}' \frac{4\pi}{q^2} \langle F | a_{\vec{k}+\vec{q},\lambda}^\dagger a_{\vec{k},\lambda}^\dagger a_{\vec{k}+\vec{q},\lambda} a_{\vec{k}\lambda} | F \rangle \\ &= \frac{e^2}{2V} \sum_{\vec{k}\vec{q},\lambda}' \frac{4\pi}{q^2} \langle F | a_{\vec{k}+\vec{q},\lambda}^\dagger \left( -a_{\vec{k}+\vec{q},\lambda} a_{\vec{k}\lambda}^\dagger + \underbrace{\delta_{\vec{k},\vec{k}+\vec{q}}}_{=0 \text{ } (\vec{q} \neq 0)} \right) a_{\vec{k}\lambda} | F \rangle \end{aligned}$$

where the green transforms due to the anticommutation rule and where the  $\delta$  is different from zero only if  $\vec{q} = 0$ . Hence, since in our case the  $\vec{q} = 0$  term is not present, the  $\delta$  gives no contribution (in practice, we can interchange these two

## Lecture 6.

Friday 27<sup>th</sup> March, 2020.

Compiled:

Wednesday 22<sup>nd</sup>

April, 2020.

First order

correction

Direct and exchange term

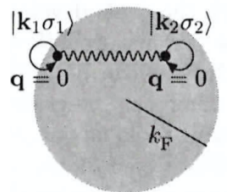


Figure 1: Direct interaction.

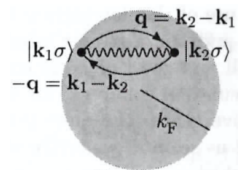


Figure 2: Exchange interaction.

operators by adding a minus sign). By replacing the destruction and creation operator with the number ones:

$$E^{(1)} = -\frac{e^2}{2V} \sum_{\substack{\vec{k}, \vec{q}, \lambda \\ |\vec{k}| \leq k_F, |\vec{k} + \vec{q}| \leq k_F}}' \frac{4\pi}{q^2} \langle F | \hat{n}_{\vec{k} + \vec{q}, \lambda} \hat{n}_{\vec{k}, \lambda} | F \rangle < 0!$$

this exchange term results negative even though the electron-electron interaction is repulsing. It is called **Hartree-Fock** approximation, for a homogeneous system. Now, we should try to really evaluate this expression. First of all, we observe that:

*Hartree-Fock  
approximation*

$$\begin{cases} \hat{n}_{\vec{k}, \lambda} = |F\rangle & \text{if } |\vec{k}| \leq k_F \\ \hat{n}_{\vec{k} + \vec{q}, \lambda} = |F\rangle & \text{if } |\vec{k} + \vec{q}| \leq k_F \end{cases}$$

because by construction  $\vec{k}$  and  $\vec{k} + \vec{q}$  are in the Fermi sphere. Therefore, we get:

$$E^{(1)} = -\frac{e^2}{2V} \sum_{\substack{\vec{k}, \vec{q}, \lambda \\ |\vec{k}| \leq k_F, |\vec{k} + \vec{q}| \leq k_F}}' \frac{4\pi}{q^2} \xrightarrow{L \rightarrow \infty} -\frac{2e^2}{2V} \frac{V^2}{(2\pi)^6} \int d\vec{k} \int d\vec{q} \frac{4\pi}{q^2} \Theta(k_F - k) \Theta(k_F - |\vec{k} + \vec{q}|)$$

where by considering the thermodynamic limit we compute integrals instead of sums (it is simpler), going from a double sum to a double integral. In particular, we get the yellow term "2" by summing over  $\lambda$  and the  $\theta$  functions are introduced to enforce the conditions in the sums. Moreover,  $\vec{q} \neq 0$  can be omitted since it effects the integrand at only a single point, so it is negligible compared to the rest of the contribution.

Now it is convenient to make the expression more symmetric. For that purpose, we introduce new variables:

$$\vec{p} \equiv \vec{k} + \frac{\vec{q}}{2} \Rightarrow \vec{k} = \vec{p} - \frac{\vec{q}}{2}$$

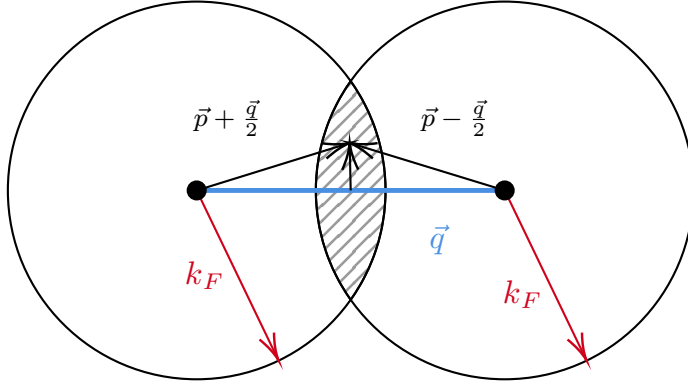
Thus we obtain:

$$E^{(1)} = -\frac{4\pi e^2 V}{(2\pi)^6} \int d\vec{q} \frac{1}{q^2} \int d\vec{p} \Theta\left(k_F - \left|\vec{p} - \frac{\vec{q}}{2}\right|\right) \Theta\left(k_F - \left|\vec{p} + \frac{\vec{q}}{2}\right|\right)$$

and considering that  $\left|\vec{p} \pm \frac{\vec{q}}{2}\right| \leq k_F$ , the region of integration over  $\vec{p}$  is the intersection between 2 spheres of radius  $k_F$ . Indeed, it is easy to demonstrate that you can have a non vanishing contribution only if  $|\vec{q}| \leq 2k_F$ . Hence, this integral is equal to compute the volume of the intersection between two spheres of radius  $k_F$  (see Fig.3) that are located in a such a way that the distance between their centers is exactly given by  $\vec{q}$ . Thus we have reduced an integral to a geometrical problem by evaluating the intersection between two spheres at distance  $\vec{q}$ .

It is convenient to go to dimensionless variables:

$$x \equiv \frac{q}{2k_F}$$



**Figure 3:** Intersection between two spheres of radius  $k_F$  at distance  $|\vec{q}| \leq 2k_F$ .

Hence:

$$\begin{aligned}
 \int \frac{d\vec{q}}{q^2} \int d\vec{p} \Theta\left(k_F - \left|\vec{p} - \frac{\vec{q}}{2}\right|\right) \Theta\left(k_F - \left|\vec{p} + \frac{\vec{q}}{2}\right|\right) &= \int \frac{d\vec{q}}{q^2} \frac{4\pi k_F^3}{3} \left(1 - \frac{3}{2}x + \frac{x^3}{2}\right) \Theta(1-x) \\
 &= 4\pi \int_0^\infty dq \frac{q^2}{q^2} \left[ \frac{4\pi k_F^3}{3} \left(1 - \frac{3}{2}x + \frac{x^3}{2}\right) \Theta(1-x) \right] \\
 &= 4\pi 2k_F \int_0^\infty dq \frac{1}{2k_F} \left[ \frac{4\pi k_F^3}{3} \left(1 - \frac{3}{2}x + \frac{x^3}{2}\right) \Theta(1-x) \right] \\
 &= 4\pi 2k_F \int_0^\infty dx \left[ \frac{4\pi k_F^3}{3} \left(1 - \frac{3}{2}x + \frac{x^3}{2}\right) \right] \Theta(1-x) \\
 &= 4\pi 2k_F \int_0^1 dx \left[ \frac{4\pi k_F^3}{3} \left(1 - \frac{3}{2}x + \frac{x^3}{2}\right) \right]
 \end{aligned}$$

The first correction term results:

$$\begin{aligned}
 E^{(1)} &= -\frac{4\pi e^2 V}{(2\pi)^6} (4\pi 2k_F) \frac{4\pi k_F^3}{3} \underbrace{\left[ \int_0^1 dx \left(1 - \frac{3}{2}x + \frac{x^3}{2}\right) \right]}_{=3/8} \\
 &= \frac{2}{3} \frac{e^2 V}{\pi^3} k_F^3 k_F \left(\frac{3}{8}\right) \stackrel{(a)}{=} \frac{2e^2 N}{\pi} \left(\frac{9\pi}{4}\right)^{1/3} \frac{1}{r_s a_0} \left(\frac{3}{8}\right) \\
 &= -\left(\frac{e^2}{2a_0}\right) \frac{N}{r_s} \underbrace{\left[ \left(\frac{9\pi}{4}\right)^{1/3} \frac{3}{2\pi} \right]}_{=0.916}
 \end{aligned}$$

where in the step (a) we have made the substitutions:

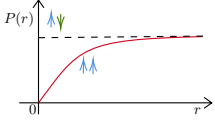
$$k_F^3 = 3\pi^2 \frac{N}{V} \quad k_F = \left(\frac{9\pi}{4}\right)^{1/3} \frac{1}{r_s a_0}$$

By substituting the numerical result we have:

$$E^{(1)} = -\underbrace{\left(\frac{e^2}{2a_0}\right)}_{=1\text{Ryd}} \frac{N}{r_s} \cdot 0.916 < 0 \quad (1)$$

so the first order correction (exchange or Hartree-Fock term) is negative as anticipated.

Now, the question is how we can interpret this contro-intuitive fact from a physical point of view, namely the reason why the electron-electron interaction gives a negative



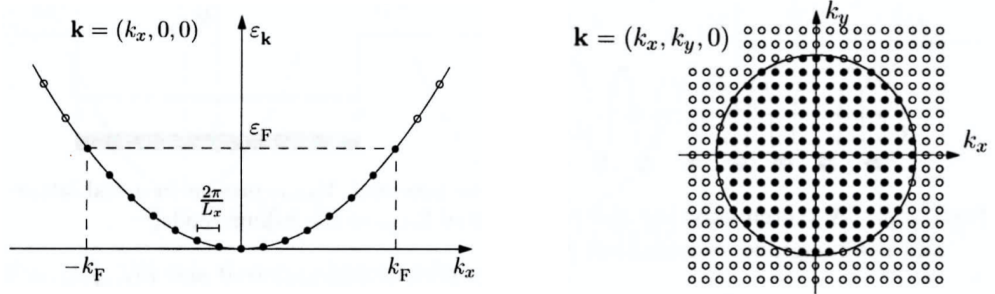
**Figure 4:** Probability of finding an electron at the distance  $r$ , for parallel and antiparallel spins.

### Exchange-hole

contribution to the energy. Since mathematically the minus sign rise due to the antisymmetry of the wave functions, the physical origin is the Pauli exclusion principle, which prevents two electrons with parallel spin from getting too close. More specifically, we can visualize the situation as follow (see Fig.4): suppose that we put an electron in the origin and we want to look at the probability of finding another electron at a certain distance  $r$ . If we consider an independent particle approach and suppose that electrons have parallel spins, since there is no interaction term the probability of finding the electron at the distance  $r$  is uniform (i.e. the same at all the distances). However, we are considering electrons with parallel spins hence, even if there is no direct Coulomb interaction acting on the system, in any case there is the Pauli exclusion principle: we cannot have the two electrons with parallel spins occupy the same position or even very close to each other. This produced the so called **exchange-hole** concept, which means that the probability of finding another electron with same spin very close to the given electron is reduced with respect to the situation where the Pauli principle where not active.

In other words, the electrons are forced somehow to avoid each other since only one electron at a time can be at a given point in space: the "direct" classical Coulomb interaction ( $\vec{q} = 0$ ) does not take this into account (it was just used to cancel divergences) and overestimates the energy while the "exchange" part correct for this by being negative. In other words, while a direct classical term does not take Pauli exclusion principle into account and tend to overestimate the energy, the exchange part corrects for this and gives a negative contribution. This is the reason why we really get a minus sign for this contribution to the energy.

In Fig.5 are shown some pictures to have a better idea of what is happening. In particular, on the right there is the Fermi sphere, on the left it is projected just along one dimension and the states are occupied up to the Fermi level.



**Figure 5:** Two aspects of  $|F\rangle$  in  $\vec{k}$ -space. To the left the dispersion relation  $\varepsilon_{\vec{k}}$  is plotted along the line  $\vec{k} = (k_x, 0, 0)$  and  $\varepsilon_F$  and  $k_F$  are indicated. To the right the occupation of the states is shown in the plane  $\vec{k} = (k_x, k_y, 0)$ . The Fermi sphere is shown as a circle with radius  $k_F$ . Filled and empty circles represent occupied and unoccupied states, respectively.

### Jellium model: final results and comments

To summarize, we have seen that in the high density limit ( $r_s \rightarrow 0$ ), the total energy of the Jellium system is given by:

$$E = E_0 + E^{(1)} + \dots$$

Thus, the energy per particle is:

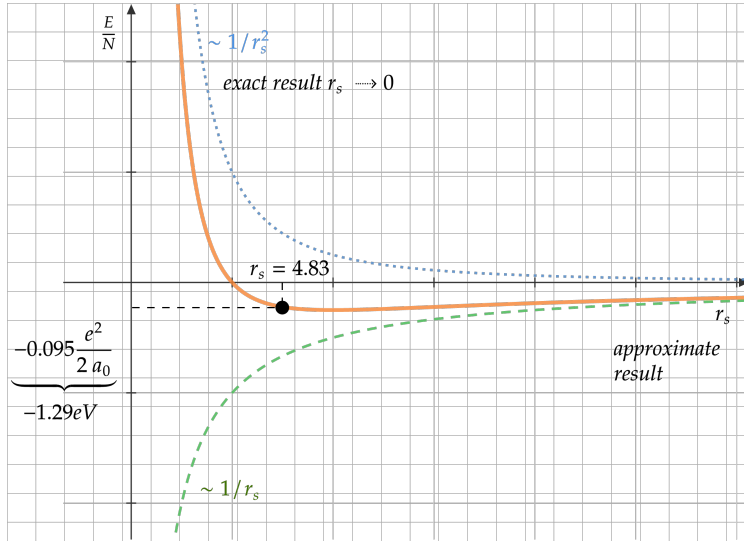
$$\frac{E}{N} \underset{r_s \rightarrow 0}{=} \left( \frac{e^2}{2a_0} \right) \left[ \frac{2.21}{r_s^2} - \frac{0.916}{r_s} + \dots \right]$$

where the yellow term correspond to the kinetic energy term, while the green one the Hartree-Fock exchange energy term. Of course, in principle, we have an infinite series of terms; all the remaining terms are denoted by the definition of Wigner (1933) as **correlation energy**. This historical definition is misleading, because in principle also the Hartree-Fock exchange energy has to do with some sort of correlations between the electrons. Another definition is given by Feymann (1972) that called all the remaining terms **stupidity energy**, because of the difficulty to go beyond the Hartree-Fock term in order of expansion.

Let us consider only the first two terms; hence, the energy per particle at the first order in the high density limit results:

$$\frac{E}{N} \underset{r_s \rightarrow 0}{\simeq} \left( \frac{e^2}{2a_0} \right) \left[ \frac{2.21}{r_s^2} - \frac{0.916}{r_s} \right]$$

It is interesting to consider  $E/N$  as a function of  $r_s$  as in Fig.6. We can note that the dominant term as  $r_s \rightarrow 0$  is the kinetic energy term, while the exchange term is the dominant one if  $r_s$  is large. The combination of the two contribution is the solid line in orange. In particular, this curve has a *minimum* for a negative value of the energy and for a given value of  $r_s$ . Physically, it means that the system is stable system and bounded which explains why the metals are stable. One should remember that while in the vanishing  $r_s \rightarrow 0$  region the exact solution is the one founded, we cannot say the same for relatively large value of the  $r_s$  parameter. Indeed, even around the minimum the result is just an approximation because we stopped at the first order term.



**Figure 6:** Energy per particle at the first order.

For resuming, as  $r_s \rightarrow 0$ , high density limit, this is the exact solution, while for larger values of  $r_s$  the solution is only approximate. However, by using the **variational principle**:

$$\langle F | \hat{H} | F \rangle = \langle F | \hat{H}_0 + \hat{H}_1 | F \rangle \geq \langle O | \hat{H} | O \rangle$$

it states that if we compute the Hamiltonian in the non-interacting ground state, by stopping at the first term of the perturbative expansion, the energy that we find is in any case higher or at most equal to the true ground state. Hence, our solution at the first order is an upperbound of the true solution! Thus since already our solution gives a bound state, the true solution is even more bounded.

The exact solution certainly represents a bound system with energy lying below the approximated curve. As said, this explains why a metal is stable and this very simple model is able to explain the largest part of the binding energy of metals in a semi-quantitative way. For instance, let us consider Table 1 in which we compare our results with the one obtained experimentally for the coesive energy of sodium. The correspondence is quite nice and the error is really small. Therefore, the very important result is that we have a semi-quantitative agreement with experiments.

	$r_s$	$E/N$ (eV)
Na	3.96	-1.13
Jellium model	4.83	-1.29

**Table 1:** Experimental results vs Jellium model.

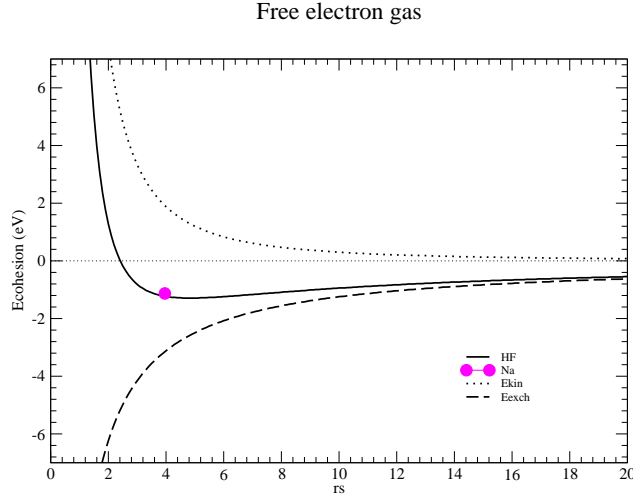
However, one should make some observation that in some sense reduce this good agreement:

- First of all, in our model there is nothing that has to do with the crystal lattice, so with the specific metallic system. One can show that for instance by comparing the results with experimental values for other metals the agreement is less good, so sodium is the best case that we can compare with.
- Despite of the variational principle our estimated binding energy is lower than the experimental one  $E_{\text{Na}} > E_{\text{jellium}}$ , due to the fact that we are neglecting a positive contribution to the kinetic energy due to the localized ions in real metals.

To make the situation clear, let us see the picture in Fig.7 where the dashed line is the exchange contribution, the dotted line the kinetic contribution and the solid line the combination of the two. We can see that the binding energy of sodium is higher than our estimate, while from the variational principle we have expected the opposite. Of course we are neglecting some crucial features of real metals, that are characterized by not a smeared positive background, but by distribution of localized ions. Since this ions are surrounded by core electrons, the valence electrons, that form the dynamical particles of the Jellium model, cannot go too close to these ions, so this means that they have roughly speaking less space to move in the system. Since this means that they should be a little more localized than in the Jellium model, more localization due to uncertainty principle means higher momentum. In conclusion, in our model we are neglecting a positive contribution to the kinetic energy, which would make our solution higher in energy than our estimate.

Despite these observations, the one obtained is a very nice conclusion! Our very simple model is able to explain the most important contributions to the binding energy of real metals.

One could also think to go further improving the first order result by addition terms in the perturbative approach. Unfortunately, if one tries to improve the 1<sup>st</sup> order result by going to, for instance, 2<sup>nd</sup> order perturbation theory, the result is disastrous: it diverges. It happens similarly for additionally terms. Hence, it is necessary some sort of **regularization** of the divergent behavior by taking higher order perturbation terms into account; the wrong idea is that by taking into account many terms in the perturbative approach, we are able to regularize these divergences. Actually, it is not enough even take a large number of perturbative terms and one has to consider perturbation theory to **infinite order** (i.e. an infinite number of terms),



**Figure 7:** Hartree-Fock energy.

which is not possible using standard perturbative approach, but we have to exploit all the possibilities of the quantum field theory. First of all, we will need to introduce some technical tools to be able to afford such a problem.

### Wigner solid

Up to now we have estimated the basic properties, as equilibrium density and energy, of the Jellium model considering in particular the high density limit, where the kinetic energy is the dominant term. One can also wonder what happens considering the opposite limit: the low density case  $r_s \rightarrow \infty$ . One can show (Wigner, 1938) that a lower energy (than that of the Jellium model) can be obtained by allowing the electrons to "crystallize" in a so called **Wigner solid**. Hence, instead of being free to move above the homogeneous background, electrons are localized in a crystal of negative charges.

*Low density limit*

In fact, if the density is low there is a lot of space available for electrons and the zero-point kinetic energy associated with localizing the electrons becomes negligible in comparison with the electrostatic energy of a classical lattice of charges. Since essentially the electrons are very far from each other and are localized around lattice site, actually the quantum mechanical description is no longer essential because we can treat the particles as classical charged particles in a lattice. Then, for instance the energy can be evaluated, despite of the long-range nature of the Coulomb interaction, by using methods used in solid state theory as for instance the Madelung-Ewald method in ionic solid.

*Remark.* Since from the variational principle we know that the true energy has a energy lower than the Jellium model energy and since in this case we get for large  $r_s$  a solution which is lower in energy than that obtained for the Jellium model, the Wigner solution is better than the Jellium one.

The details of the derivation has been done by Wigner and we illustrate only the final conclusion. Thus, in the limit of low density, the energy per particle is given by:

$$\frac{E}{N} \underset{r_s \rightarrow \infty}{=} \left( \frac{e^2}{2a_0} \right) \left[ -\frac{1.79}{r_s} + \frac{2.66}{r_s^{3/2}} + \dots \right]$$

where the yellow term is obtained as the difference between the potential of the

electrons on fixed lattice sites and the exchange energy<sup>1</sup>, while the green term is the zero-point oscillation of electrons about their equilibrium positions (i.e. electrons are forced to remain close to their lattice positions, but they can oscillate around these positions).

Summarizing, we have obtained two estimates of energy per particle in such a model, one for high density and one for low density. To be more precise:

- in the Jellium model the kinetic term dominates and it is valid for very high density:  $r_s < 1$ .
- in the Wigner-solid model the electrostatic potential energy term dominates and it is valid for very low density:  $r_s > 20$ .

Unfortunately, the region which is more interesting from a physical point of view is the one which corresponds to the densities of real metals and it is intermediate between these two cases:  $2 < r_s < 6$ . In particular, in this region the kinetic and potential energy are comparable. This is clear from Fig.7, in which in the region from 2 to 6 both the kinetic energy and the exchange energy really are comparable.

In practice, it means that, as we will see better in the future, one needs to interpolate between high and low density to try to reproduce the behavior of real metals.

In conclusion, despite of the simplicity of the Jellium model, this is an interesting example in which we can apply the second quantization formalism and it is also interesting from a physical point of view for describing real metals.

Now, we skip to chapter 4 of the [?] to try to develop the important formalism and tools that are required to go further in the perturbative approach. Hence, we do not deal with chapter 2, which is a summary of the basic notions of quantum statistical mechanics but we strongly suggest to read this chapter, because we will need some basic concepts of quantum statistical mechanics in the future.

---

<sup>1</sup>Remember that also for the Jellium model the exchange energy goes as the inverse of  $r_s$ .