# THE HARTREE-FOCK MODEL

## 1. Introduction

*System of interacting electrons*. We now study the system of $N$ electrons in the presence of an external potential (due to static nuclei) and with a Coulomb interaction. The corresponding Hamiltonian (in appropriate units) is what we called[1] *atomic* or *molecular* Hamiltonians:

$$H_N = \sum_{j=1}^{N} \big( -\Delta_{x_j} + V(x_j) \big) + \sum_{1 \le i < j \le N} \frac{1}{|x_i - x_j|}, \tag{1}$$

where $V$ takes the form of:

$$V(x) = -\sum_{m=1}^{M} \frac{Z_m}{|x - R_m|}, \ Z_m > 0, \ R_m \in \mathbb{R}^3. \tag{2}$$

In the appropriate units, we must think of $V$ as ($-1$ times) the electrostatic potential created by $M$ positive charges $Z_m > 0$ located at $R_m \in \mathbb{R}^3$. We write $Z := \sum_{j=1}^{N} Z_m$ the total charge of the nuclei.

Several remarks must be made. First, so far we have been considering systems of *distinguishable* particles of the same kind. This lead to considering tensor products of the Hilbert space of one particle. Electrons however are indistinguishable particles that must satisfy the Pauli principle. That is two electrons cannot be in the same state. It is implemented analytically by considering the Hilbert space of antisymmetrical functions, or equivalently, by considering *wedge* products of the Hilbert space of one particle.

Secondly, and as we did throughout the lecture, we consider for simplicity spinless particles, and we only take into account electrostatic interaction, hence the form of the Hamiltonian $H_N$, and the Hilbert space of one particle $L^2(\mathbb{R}^3)$.

*Approximations*. We thus consider such a system of interacting electrons. A relevant problem for instance is to get an estimate on the ground state energy, or to study properties of the ground state. Still, even for $N$ of order – say – 20, this is a difficult problem (from a computational point of view). Approximation models have been introduced to obtain partial answers.

In this part, we will consider the Hartree-Fock approximation which consists in restricting states to the set $\mathcal{T}_N$ of so-called Slater determinants. We will then study the corresponding minimization problem:

$$E_{HF}^N := \inf_{\substack{\psi \in \mathcal{T}_N \cap \mathrm{dom}(H_N), \\ \|\psi\| = 1}} \langle \psi, H_N \psi \rangle. \tag{3}$$

Having the min-max principle in mind, the Hartree-Fock energy $E_{HF}^N$ is larger than the ground state energy. We will prove a theorem due to Lieb and Simon [2], which states[2] that (3) admits a minimizer when $N < Z + 1$.

[1] atomic for $M = 1$.

[2] The same method presented during the seminar ensures us that if $N \ge 2Z + 1$, then (3) has no minimizers.

From a mathematical point of view, the minimization problem has the following features.

(1) It is non-linear: the energy functional is not linear w.r.t. the variable (here the wave function).
(2) It is a problem under constraint (the restriction to Slater determinants).
(3) It is *locally compact* on $\mathbb{R}^3$: minimizing sequences are bounded in the $H^1$-norm (see below for the reason of the name).

To solve it, we will use techniques of calculus of variations. We will also uses three results which are interesting by themselves.

(1) The fact that $H^1(\mathbb{R}^3)$ is compactly injected into $L^2_{\text{loc}}(\mathbb{R}^3)$ (special case of the theorem of Rellich-Kondrachov).
(2) Newton's theorem which states that for a radially symmetric Borel measure $\mu$ in $\mathbb{R}^3$ we have:

$$\begin{aligned}\mu * \frac{1}{|\cdot|}(x) &:= \int_{\mathbb{R}^3} \frac{\mu(\mathrm{d}y)}{|x-y|},\\ &= \frac{1}{|x|}\int_{y:|y|\le|x|} \mu(\mathrm{d}y) + \int_{y:|y|>|x|} \frac{\mu(\mathrm{d}y)}{|y|}.\end{aligned}$$

(3) For a Borel signed Borel measure of finite total variation $\mu$ (that is $\mu = \mu_+ - \mu_-$ with $\mu_+, \mu_-$ finite Borel measures), if $\mu(\mathbb{R}^3) = \mu_+(\mathbb{R}^3) - \mu_-(\mathbb{R}^3) < 0$, then the self-adjoint operator $-\Delta + < \mu * \frac{1}{|\cdot|}$ has infinite discrete spectrum below its essential spectrum $[0, +\infty)$.

We first properly state the problem (we define antisymmetrical functions and Slater determinants and the Hartree-Fock functional), and then show the announced theorem.

## 2. Antisymmetrical functions and Slater determinants

2.1. **The space $L^2_a(\mathbb{R}^{3N})$.** Consider the space $L^2(\mathbb{R}^3 N) \simeq L^2(\mathbb{R}^3)^N$. Recall that we write $\underline{x} = (x_1, \ldots, x_N) \in \mathbb{R}^{3N}$ where $x_j \in \mathbb{R}^3$, $1 \le j \le N$. The space $L^2_a(\mathbb{R}^{3N})$ of antisymmetric function is the closed[3] subspace of $L^2(\mathbb{R}^{3N})$ defined by:

$$\begin{aligned}L^2_a(\mathbb{R}^{3N}) := \big\{\psi \in L^2(\mathbb{R}^{3N}),\ \forall\, 1 \le i < j \le N, \text{ there holds } a.e.\\ \psi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_N) = -\psi(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_N)\big\}.\end{aligned}$$

We will prove below that subspace $L^2_a(\mathbb{R}^{3N})$ is isometric to $\wedge_{1\le i\le N} L^2(\mathbb{R}^3)$.

A legitimate question is whether $H_N$ of (1) is well-defined on $L^2_a(\mathbb{R}^{3N})$ as an unbounded self-adjoint operator. The answer is yes, and this is due to the fact that $H_N$ acts symmetrically on the $N$ variables $x_1, \ldots, x_N \in \mathbb{R}^3$.

Hence $H_N$ with domain $H^2(\mathbb{R}^{3N}) \cap L^2_a(\mathbb{R}^{3N}) =: H^2_a(\mathbb{R}^{3N})$ is a well-defined self-adjoint operator on $L^2_a(\mathbb{R}^{3N})$: if $\psi \in H^2_a(\mathbb{R}^{3N})$, then $H_N\psi$ is also in $L^2_a(\mathbb{R}^{3N})$.

By our study on Slater determinants below, it will be obvious that $H^2_a(\mathbb{R}^{3N})$ is dense in $L^2_a(\mathbb{R}^{3N})$.

Similarly its form domain is $H^1_a(\mathbb{R}^{3N}) := H^1(\mathbb{R}^{3N}) \cap L^2_a(\mathbb{R}^{3N})$.

[3]the antisymmetry condition is continuous w.r.t. the $L^2$-norm.

2.2. **Slater determinants.** There is a special class of vectors in $L^2_a(\mathbb{R}^{3N})$ called Slater determinants, which come directly from the construction of the wedge product.

Consider a family $(\psi_j)_{1\leq j\leq N}$. The wedge product $\psi_1 \wedge \cdots \wedge \psi_N \in L^2_a(\mathbb{R}^{3N})$ of the $\psi_j$'s is defined as follows,

$$\begin{aligned}\psi_1 \wedge \cdots \wedge \psi_N(\underline{x}) &:= \frac{1}{\sqrt{N!}}\det\big((\psi_i(x_j))_{1\leq i,j\leq N}\big),\\ &= \frac{1}{\sqrt{N!}}\sum_{\sigma\in\mathcal{S}_N}\varepsilon(\sigma)\prod_{1\leq i\leq N}\psi_{\sigma(i)}(x_i),\end{aligned}$$

where $\mathcal{S}_N$ denotes the set of permutations of the $N$ first integers. From this definition, it is clear that wedge products lie in $L^2_a(\mathbb{R}^{3N})$. Non-zero wedge products are called *Slater determinants.*

For instance, when $N = 2$, we have:

$$\psi_1 \wedge \psi_2(x,y) := \frac{1}{\sqrt{2}}\big(\psi_1(x)\psi_2(y) - \psi_2(x)\psi_1(y)\big).$$

**Remark 1.**

- *If the family $(\psi_j)_{1\leq j\leq N}$ is linearly dependent, then the wedge product identically vanishes as we can see from the determinant formula.*
- *Observe that the determinant formula also gives:*

  $$\psi_1 \wedge \cdots \wedge \psi_i \wedge \cdots \wedge \psi_j \wedge \cdots \wedge \psi_N = -\psi_1 \wedge \cdots \wedge \psi_j \wedge \cdots \wedge \psi_i \wedge \cdots \wedge \psi_N$$

- *We emphasize that the factorial term $\frac{1}{\sqrt{N!}}$ is a convention, due to our considering wedge products of Hilbert spaces. It is motivated by normalization issues: when $(\psi_j)_{1\leq j\leq N}$ is an orthonormal family, then the corresponding wedge product has unit norm[4] in $L^2_a(\mathbb{R}^{3N})$.*

  *In geometry the factorial term is absent for exterior products of vectors (Grassmann algebra) or wedge products of differential forms.*

It is easy to see that Slater determinants span the space of antisymmetrical functions. Indeed, consider an ONB $(\varphi_i)_{i\in\mathbb{N}}$ of $L^2(\mathbb{R}^3)$. Then we know[5] that the family $(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_N})_{i_j\in\mathbb{N}}$ is an ONB of $L^2(\mathbb{R}^3)^N$.

Then pick $\varphi \in L^2_a(\mathbb{R}^{3N})$. Decomposing $\varphi$ w.r.t. the above basis, and using its antisymmetry we get the following. For $(i_1,\dots,i_N) \in \mathbb{N}^N$, and $\sigma \in \mathcal{S}_N$, we have:

$$\langle \varphi_{\sigma(i_1)} \otimes \cdots \otimes \varphi_{\sigma(i_N)}, \varphi\rangle = \varepsilon(\sigma)\langle \varphi_{i_1} \otimes \cdots \otimes \varphi_{i_N}, \varphi\rangle.$$

It suffices to decompose $\sigma$ into a product of transpositions: the signature $\varepsilon(\sigma)$ corresponds to the parity of the number of transpositions of such a decomposition.

We thus obtain that $\varphi$ can be decomposed as a sum of Slater determinants. We leave as an exercise, the fact that a Slater determinant obtained from an orthonormal family has unit norm.

Consider a normalized Slater determinant $\psi_1 \wedge \cdots \wedge \psi_N$, with $\|\psi_j\|_{L^2} = 1$. If we think of it as describing the system of $N$ particles, the $N$ one-particle states $\psi_j$ are the states occupied by the electrons and are called orbitals.

We recover the fact that the restriction to $L^2_a(\mathbb{R}^{3N})$ implements the Pauli principle: the orbitals are orthonormal to each other, and Slater determinants span $L^2_a(\mathbb{R}^{3N})$.

---

[4] Prove it!

[5] This follows from two things. First continuous functions with compact support are dense in $L^2(\mathbb{R}^3)$. Secondly, fix a compact support $K \subset \mathbb{R}^3$, $K = [-M,M]^3$ for instance. Then by Stone-Weierstrass theorem, continuous functions $f$ of the form $f(x) = f_1(x_1)f_2(x_2)f_3(x_3)$ with $f_i$ continuous span $C(K)$ with the sup-norm.

**Remark 2.** *As a word of caution: the orbitals $\psi_j$ are not uniquely determined. The only well-defined geometrical object is the $N$-dimensional space spanned by the $\psi_j$'s.*

*It is not surprising. If we have the Grassmann algebra in mind, a normalized Slater determinant is a $N$-vector. As a state is defined up to a phase, from a geometrical point of view, it characterizes the $N$-dimensional space.*

*We will recover this fact when considering the reduced one-particle density matrices.*

2.3. **Reduced one-particle density matrices.** For a normalized $\psi \in L^2_a(\mathbb{R}^{3N})$, its reduced one-particle density matrix $\gamma_\psi$ is the bounded operator on $L^2(\mathbb{R}^3)$, whose integral kernel is:

$$\gamma_\psi(x,y) := N \int_{\underline{x}' \in \mathbb{R}^{3(N-1)}} \psi(x,\underline{x}')\overline{\psi(y,\underline{x}')}\mathrm{d}\underline{x}'. \tag{4}$$

From (4), we can show:

$$0 \le \gamma_\psi \le 1. \tag{5}$$

This result is proved below in Section 5.1.

For a (normalized) Slater determinant $\psi \in \mathcal{T}_N$, $\gamma_\Psi$ is the orthogonal projection onto the vector space spanned by its orbitals[6]:

$$\gamma_\psi(x,y) = \sum_{j=1}^N \psi_j(x)\overline{\psi_j(y)}, \text{ or equivalently } \gamma_\psi = \sum_{j=1}^N |\psi_j\rangle\langle\psi_j|.$$

The diagonal of $\gamma_\psi$ is called the density of $\psi$ and written $\rho_\psi$:

$$\begin{array}{rcl} \rho_\psi(x) & := & \gamma_\psi(x,x), \\ & \underset{\psi \text{ Slater}}{=} & \sum_{j=1}^N |\psi_j(x)|^2. \end{array}$$

## 3. The Hartree-Fock model

As said in the introduction, we restrict the set of admissible states and only consider Slater determinants. As we are mainly interested in states our convention will be that whenever we write $\psi \in \mathcal{T}_N$, it is understood that $\|\psi\|_{L^2} = 1$ and that $\psi = \psi_1 \wedge \cdots \wedge \psi_N$ where $\psi_j$ are its (normalized) orbitals.

### 3.1. **The Hartree-Fock functional.**

3.1.1. *Form of the energy functional.* The energy of a Slater determinant $\psi$ can be written in terms of the orbitals $\psi_j$, or more compactly in terms of the reduced density matrix $\gamma_\psi$. Below, for $\psi \in H^2_a(\mathbb{R}^{3N})$, we write $R_{\gamma_\psi}$ the operator defined by the integral kernel:

$$R_{\gamma_\psi}(x,y) = \frac{\gamma_\psi(x,y)}{|x-y|}.$$

For such a $\psi$, Hardy's inequality ensures that the operator $R_{\gamma_\psi}$ is Hilbert-Schmidt.

**Lemma 3.** *Let $\psi \in \mathscr{T}_N \cap H^2_a(\mathbb{R}^{3N})$ be a Slater determinant. Its energy is:*

$$\langle \psi, H_N\psi\rangle_{L^2_a(\mathbb{R}^{3N})} = \sum_{j=1}^N \int_{\mathbb{R}^3} (|\nabla\psi_j|^2 + V|\psi_j|^2) + \frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{\rho_\psi(x)\rho_\psi(y) - |\gamma_\psi(x,y)|^2}{|x-y|}\mathrm{d}x\mathrm{d}y.$$

*More compactly there holds:*

$$\langle \psi, H_N\psi\rangle_{L^2_a(\mathbb{R}^{3N})} = \mathrm{tr}((-\Delta + V)\gamma_\psi) + D(\rho_\psi) - X(\gamma_\psi), \tag{6}$$

[6] This is left as an exercise.

*where $D(\rho_\psi)$ and $X(\gamma_\psi)$ denote:*

$$D(\rho_\psi) = \frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{\rho_\psi(x)\rho_\psi(y)}{|x-y|}\mathrm{d}x\mathrm{d}y = \frac{1}{2}\,\mathrm{tr}(\rho_\psi * \tfrac{1}{|\cdot|}\gamma_\psi),$$

$$X(\gamma_\psi) = \frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{|\gamma_\psi(x,y)|^2}{|x-y|}\mathrm{d}x\mathrm{d}y = \frac{1}{2}\,\mathrm{tr}(R_{\gamma_\psi}\gamma_\psi).$$

**Remark 4.** *The term $D(\rho_\psi)$ is called the* direct term *and corresponds to the electrostatic potential of the density $\rho_\psi$. The term $X(\gamma_\psi)$ is called the exchange term and is a purely quantic term.*

The Lemma is proven below in Section 5.2.

3.1.2. *Properties of the functional.* As we have seen in the part on quadratic forms, we can define the energy on the form domain of $H_N$. We recover this fact in (6): the functional energy is well-defined if the orbitals $\psi_j$'s only have a $H^1$-regularity. It is convenient to introduce the functional energy as a function of the orbitals.

**Definition 1.** *Let $\mathcal{E} : H^1(\mathbb{R}^3)^N \to \mathbb{R}$ be the functional defined as follows. For $\underline{\psi} = (\psi_1, \dots, \psi_N) \in H^1(\mathbb{R}^3)^N$, we define*

$$\mathcal{E}(\underline{\psi}) := \sum_{j=1}^N \int_{\mathbb{R}^3} (|\nabla\psi_j|^2 + V|\psi_j|^2) + D(\rho_{\underline{\psi}}) - X(\gamma_{\underline{\psi}}),$$

*where $\gamma_{\underline{\psi}}$ denotes the compact operator $\sum_{j=1}^N |\psi_j\rangle\langle\psi_j|$ and $\rho_{\underline{\psi}}(x) := \sum_j |\psi_j|^2$.*

**Remark 5.** *Observe that for $\psi \in H^1(\mathbb{R}^3)^N$ we have:*

$$\begin{aligned}|\gamma_\psi(x,y)|^2 &= \left|\sum_i \psi_i(x)\overline{\psi_i(y)}\right|^2, \\ &\le \sum_i |\psi_i(x)|^2 \sum_i |\psi_i(y)|^2 = \rho_\psi(x)\rho_\psi(y).\end{aligned} \tag{7}$$

*In particular we have $D(\rho_\psi) - X(\gamma_\psi) \ge 0$.*

For $\underline{\psi} \in H^1(\mathbb{R}^3)^N$, we also defines its gram matrix as the matrix of the inner products of its elements[7]:

$$\mathrm{Gram}(\underline{\psi}) := \big(\langle\psi_i, \psi_j\rangle\big)_{1\le i,j\le N} \in \mathbb{C}^{N\times N}$$

The Hartree-Fock functional $\mathcal{E}_{HF}$ on $H^1(\mathbb{R}^3)^N$ corresponds to the restriction of $\mathcal{E}$ to the orthonormal families $\underline{\psi} \in H^1(\mathbb{R}^3)^N$, or equivalently to the families which satisfy:

$$\mathrm{Gram}(\underline{\psi}) = \mathbb{1}_{\mathbb{C}^N}. \tag{8}$$

We thus introduce the sets:

$$\begin{aligned}\mathcal{M} &:= \{\underline{\psi} \in H^1(\mathbb{R}^3)^N,\ \mathrm{Gram}(\underline{\psi}) = \mathbb{1}_{\mathbb{C}^N}\}, \\ \mathcal{K} &:= \{\underline{\psi} \in H^1(\mathbb{R}^3)^N,\ 0 \le \mathrm{Gram}(\underline{\psi}) \le \mathbb{1}_{\mathbb{C}^N}\}.\end{aligned} \tag{9}$$

We have indeed the following.

**Lemma 6.** *The set $\mathcal{K}$ is the closure of $\mathcal{M}$ under the weak $H^1$-topology.*

[7]From a geometrical point of view, the Gram matrix corresponds to the pullback of the $L^2$-metric through the submersion $\underline{z} := (z_1, \dots z_N) \mapsto \sum_{j=1}^N z_j\psi_j$ from $\mathbb{C}^N$ to $\mathrm{span}\{\psi_j,\ 1 \le j \le N\}$. This simply means that for $\underline{z}, \underline{w} \in \mathbb{C}^N$ we have:

$$\Big\langle \sum_i z_i\psi_i, \sum_j w_j\psi_j \Big\rangle_{L^2(\mathbb{R}^3)} = \langle \underline{z}, \mathrm{Gram}(\underline{\psi})\underline{w}\rangle_{\mathbb{C}^N} = \underline{z}^* \,\mathrm{Gram}(\underline{\psi})\underline{w}.$$

**Remark 7.** *For $\mathcal{M}$ and $\mathcal{K}$, we can replace the $H^1$-norm by the $L^2$-norm in both the definition and the lemma: the result remains.*

Lemma 6 is to be related to the following proposition (proven in Section 5.3).

**Proposition 8.** *[Properties of $\mathcal{E}$] The functional $\mathcal{E} : H^1(\mathbb{R}^3)^N \to \mathbb{R}$ is well-defined and satisfies the following properties.*

(1) *It is $\|\cdot\|_{H^1}$-continuous.*
(2) *It is $H^1$-weakly lower semi-continuous. That is, if $(\underline{\psi}^{(n)})_n$ is a sequence of $H^1(\mathbb{R}^3)^N$ which converges weakly to $\underline{\psi}$, then we have:*

$$\mathcal{E}(\underline{\psi}) \leq \liminf_{n\to+\infty} \mathcal{E}(\underline{\psi}^{(n)}).$$

(3) *The functional $\mathcal{E}_{|\mathcal{K}}$ is bounded from below, that is $E_{\mathcal{K}} := \inf_{\psi\in\mathcal{K}} \mathcal{E}(\psi) \in \mathbb{R}$. Also there exists $C_Z, C_2 > 0$ such that:*

$$\forall\, \psi \in \mathcal{K},\ \|\psi\|_{H^1}^2 \leq C_2\mathcal{E}(\psi) + NC_Z.$$

Furthermore, the functional $\mathcal{E}$ is invariant under rotation in the following sense. The group of unitary matrices $\mathcal{U}(\mathbb{C}^N)$ acts on $H^1(\mathbb{R}^3)^N$ as follows: for $U \in \mathcal{U}(\mathbb{C}^N)$ we define

$$U \cdot \underline{\psi} := \Big(\sum_{j=1}^{N} U_{ij}\psi_j\Big)_{1\leq i\leq N}.$$

Under this action, the energy remains unchanged.

**Proposition 9.** *For $U \in \mathcal{U}(\mathbb{C}^N)$ and $\underline{\psi} \in H^1(\mathbb{R}^3)^N$, there holds $\mathcal{E}(U \cdot \underline{\psi}) = \mathcal{E}(\underline{\psi})$.*

The proof of Proposition 8 uses the following result, which is important by itself. This is a special case of the theorem of Rellich-Kondrachov (see [1]).

**Theorem 10.** *In any dimension $d \in \mathbb{N}$ the following holds.*
*Let $(\psi^{(n)})_n$ be a $H^1$-bounded sequence which converges $H^1$-weakly to $\psi \in H^1(\mathbb{R}^d)$.*
*Then there exists a subsequence $(\psi_{n_k})_{k\geq 0}$ such that for all compact set $K \subset \mathbb{R}^d$, the sequence $((\psi_{n_k})_{|K})_{k\geq 0}$ converges to $\psi_{|K}$ in $L^2(K)$.*
*We say that the set $H^1(\mathbb{R}^d)$ is compactly injected in $L^2_{\mathrm{loc}}(\mathbb{R}^d)$.*

We prove this theorem in Section 5.4.

*Proof of Lemma 6.* We first check that $\mathcal{K}$ is weakly closed. Consider a sequence $\underline{\psi}^{(n)} \overset{H^1}{\rightharpoonup} \underline{\psi}$ in $\mathcal{K}$. Then for $\underline{z} \in \mathbb{C}^N$, we have[8]:

$$\begin{aligned}\langle \underline{z}, \mathrm{Gram}(\underline{\psi})\underline{z}\rangle_{\mathbb{C}^N} &= \|\sum_j z_j\psi_j\|_{L^2(\mathbb{R}^3)}^2,\\ &\leq \liminf_{n\to+\infty} \|\sum_j z_j\psi_j^{(n)}\|_{L^2(\mathbb{R}^3)}^2,\\ &\leq \liminf_{n\to+\infty} \langle \underline{z}, \mathrm{Gram}(\underline{\psi}^{(n)})\underline{z}\rangle_{\mathbb{C}^N} \leq \|\underline{z}\|_{\mathbb{C}^N}^2.\end{aligned}$$

As $\mathcal{M} \subset \mathcal{K}$, this also shows that the weak closure of $\mathcal{M}$ is included in $\mathcal{K}$.

Conversely let $\underline{\psi} \in \mathcal{K}$. We show that $\underline{\psi}$ is a weak limit of a sequence in $\mathcal{M}$.

[8] by the uniform boundedness principle.

Let $U \in \mathcal{U}(\mathbb{C}^N)$ such that $U\,\mathrm{Gram}(\underline{\psi})U^* = \mathrm{diag}(\varepsilon_1, \dots, \varepsilon_N)$ with $0 \le \varepsilon_i \le 1$. Observe that $U^*\,\mathrm{Gram}(\underline{\psi})U$ is the Gram matrix of $U \cdot \underline{\psi}$ as we have:

$$\begin{aligned}
\big\langle U^*\underline{z}, \mathrm{Gram}(\underline{\psi})U^*\underline{z}\big\rangle^2_{\mathbb{C}^N} &= \big\| \sum_i (U^*\underline{z})_i \psi_i \big\|^2_{\mathbb{C}^N}, \\
&= \big\| \sum_i \sum_j U^*_{ij} z_j \psi_i \big\|^2_{\mathbb{C}^N}, \\
&= \big\| \sum_j \big( \sum_i U_{ji}\psi_i \big) z_j \big\|^2_{\mathbb{C}^N}.
\end{aligned}$$

So w.l.o.g. we can assume that the Gram matrix is diagonal, or in other words that the family $(\psi_i)_{1\le i\le N}$ is orthogonal with $\|\psi_i\|^2_{L^2} = \varepsilon_i$.

Let $\chi \in C_0^\infty(\mathbb{R}^3)$ with $\|\chi\|^2_{L^2} = 1$ and let $n_i \in \mathbb{S}^2$, $1 \le i \le N$ be $N$ *different* directions. For every $R > 0$ and $i$, we define $\chi_{i,R}$ by:

$$\chi_{i,R}(x) := \psi_i(x) + \sqrt{1-\varepsilon_i}\,\chi(x - Rn_i),$$

and we define $(\psi_{i,R})_{1\le i\le N}$ as the family obtained by applying the Gram-Schmidt procedure to $(\chi_{i,R})_{1\le i\le N}$: in particular $\mathrm{Gram}((\psi_{i,R})_{1\le i\le N}) = \mathbb{1}_{\mathbb{C}^N}$.

Then, as $R \to 0$, it is straightforward[9] to see that $(\psi_{i,R})_{1\le i\le N}$ weakly converges to $\underline{\psi}$ as $R$ tends to $+\infty$.

□

*Proof of Prop. 9.* It follows from the fact that $\gamma_{U\cdot\underline{\psi}} = \gamma_{\underline{\psi}}$. Indeed, there holds:

$$\begin{aligned}
\gamma_{U\cdot\underline{\psi}}(x,y) &= \sum_i ((U\cdot\underline{\psi})_i(x))\overline{(U\cdot\underline{\psi})_i(y)}, \\
&= \langle U\cdot\underline{\psi}(y), U\cdot\underline{\psi}(x)\rangle_{\mathbb{C}^N}, \\
&= \langle \underline{\psi}(y), \underline{\psi}(x)\rangle_{\mathbb{C}^N} = \gamma_{\underline{\psi}}(x,y).
\end{aligned}$$

hence we also have $\rho_{U\cdot\underline{\psi}} = \rho_{\underline{\psi}}$, showing the equalities of the two direct terms, resp. exchange terms.

If we do not use the trace formula for the kinetic energy and the interaction energy with the nuclei, observe:

$$\begin{aligned}
\sum_i \int |\nabla(U\cdot\underline{\psi})_i|^2 &= \sum_k \int_{\mathbb{R}^3} \langle \partial_k(U\cdot\underline{\psi}), \partial_k(U\cdot\underline{\psi})\rangle_{\mathbb{C}^N}, \\
&= \sum_k \int_{\mathbb{R}^3} \langle U\cdot(\partial_k\underline{\psi}), U\cdot(\partial_k\underline{\psi})\rangle_{\mathbb{C}^N}, \\
&= \sum_k \|\partial_k\underline{\psi}\|^2_{L^2(\mathbb{R}^3)^N} = \sum_i \int \|\nabla\psi_i\|^2_{L^2(\mathbb{R}^3)}.
\end{aligned}$$

Similarly:

$$\begin{aligned}
\sum_i \int_{\mathbb{R}^3} |(U\cdot\underline{\psi})_i|^2 V &= \int_{\mathbb{R}^3} \|U\cdot\underline{\psi}\|^2_{\mathbb{C}^N} V, \\
&= \int_{\mathbb{R}^3} \|\underline{\psi})\|^2_{\mathbb{C}^N} V = \sum_i \int_{\mathbb{R}^3} |\psi_i|^2 V.
\end{aligned}$$

□

### 3.2. **Existence of a minimizer.**

[9]We leave the details as an exercise.

3.2.1. *Statement of the main theorems.*

**Theorem 11.** *[Lieb and Simon] If $N < Z+1$, then there exists a minimizer to* (3)*. That is, there exists a Slater determinant $\psi \in \mathcal{T}_N \cap H^2_a(\mathbb{R}^{3N})$ with $\|\psi\|_{L^2} = 1$ such that:*

$$\langle \psi, H_N \psi \rangle = \inf_{\substack{\varphi \in \mathcal{T}_N \cap H^2_a(\mathbb{R}^{3N}) \\ \|\varphi\|_{L^2}=1}} \langle \varphi, H_N \varphi \rangle =: E^N_{HF}.$$

To prove this theorem, we need *a priori* information on minima.

From a mathematical point of view, the minimization problem is under constraint, hence we expect putative minimizers to satisfy some *Euler-Lagrange* equations. Here it turns out that the E.L. equations involve the so-called mean-field operator.

**Definition 1** (Mean-field operator)**.** *For $\psi \in \mathcal{T}_N \cap H^1_a(\mathbb{R}^{3N})$, the mean-field operator associated to $\psi$ is the self-adjoint operator on $L^2(\mathbb{R}^3)$:*

$$F_\psi := -\Delta + V + \rho_{\gamma_\psi} * \tfrac{1}{|\cdot|} - R_{\gamma_\psi}. \tag{10}$$

*This definition naturally extends to $H^1(\mathbb{R}^3)^N$.*

We recall the important feature of this operator in the following lemma.

**Lemma 12.** *Let $\underline{\psi} \in H^1(\mathbb{R}^3)^N$. Then its mean-field operator $F_{\underline{\psi}}$ is self-adjoint on $L^2(\mathbb{R}^3)$ with domain $H^2(\mathbb{R}^3)$. Furthermore its essential spectrum is $[0, +\infty)$.*

**Theorem 13.** *[Properties of minima] Assume that there exists a minimizer $\psi \in \mathcal{T}_N$ to $E^N_{HF}$ and let $F_\psi$ be its mean-field operator.*

*Then up to replacing $\psi$ by $U \cdot \psi$ for some $U \in \mathcal{U}(\mathbb{C}^N)$, the $\psi_i$'s are the eigenfunctions of $F_\psi$ corresponding to the $N$ lowest eigenvalues $\lambda_1 \leq \cdots \leq \lambda_N$ of $F_\psi$.*

*If the $N+1$-th min-max coefficient $\mu_{N+1}(F_\psi)$ of $F_\psi$ is an eigenvalue, then we have:*

$$\lambda_N = \mu_N(F_\psi) < \mu_{N+1}(F_\psi) \leq 0 \qquad \text{(Aufbau principle)}.$$

*In that case, we have:*

$$\gamma_\psi = \mathbb{1}_{(-\infty, \lambda_N]}(F_\psi).$$

3.2.2. *Proof of Lemma 12.* The key-result is Hardy's inequality. To simplify notations, we write $\psi$ instead of $\underline{\psi}$.

We have seen that $V \in L^2 + L^\infty$. Furthermore, Hardy's inequality gives that $\rho_\psi * \frac{1}{|\cdot|}$ is bounded. Indeed for all $x \in \mathbb{R}^3$ we have:

$$\begin{aligned}\left[\int \frac{\rho_\psi(y)}{|x-y|}\mathrm{d}y\right]^2 &\leq \int \rho_\psi(y)\mathrm{d}y \int \frac{\rho_\psi(y)}{|x-y|^2}\mathrm{d}y, \\ &\leq 4\|\psi\|^2_{L^2(\mathbb{R}^3)^N}\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)^N}.\end{aligned}$$

Similarly Hardy's inequality implies that $\gamma_\psi$ is Hilbert-Schmidt with:

$$\iint \frac{|\gamma_\psi(x,y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y \leq 4\,\mathrm{tr}(-\Delta\gamma_\psi) = 4\|\nabla\psi\|^2_{L^2}.$$

Hence by Kato's theorem, we get that $F_\psi$ is self-adjoint with domain $H^2(\mathbb{R}^3)$.

We now use Weyl's theorem for the invariance of the essential spectrum and show that $V + \rho_\psi * \frac{1}{|\cdot|} - R_{\gamma_\psi}$ is $(-\Delta)$-compact.

First, $R_{\gamma_\psi}$ is already compact.

For $v_\rho := \rho_\psi * \frac{1}{|\cdot|}$ which is $L^\infty$, it suffices to show that it converges to 0 at infinity to get that $v_\rho(-\Delta+1)^{-1}$ is compact as $p \mapsto (|p|^2+1)$ is $L^\infty$ and tends to 0 at infinity.

For $x \in \mathbb{R}^3$ with $|x| \geq 2$, we have:

$$v_\rho(x) = \int_{y:|x-y|\leq\frac{|x|}{2}} \frac{\rho_\psi(y)}{|x-y|}\mathrm{d}y + \int_{y:|x-y|\geq\frac{|x|}{2}} \frac{\rho_\psi(y)}{|x-y|}\mathrm{d}y.$$

The last integral is smaller than $\frac{2}{|x|}\|\psi\|_{L^2}^2$. By Cauchy-Schwarz's inequality and Hardy's inequality, the square of the first integral is smaller than

$$4\|\nabla\psi\|_{L^2}^2 \int_{y:|x-y|\leq\frac{|x|}{2}} \rho_\psi(y)\mathrm{d}y \leq 4\|\nabla\psi\|_{L^2}^2 \int_{y:|y|\geq\frac{|x|}{2}} \rho_\psi(y)\mathrm{d}y,$$

which tends to 0 as $x \to +\infty$ by monotone convergence.

As last, as we have seen in earlier lecture, we have $V \in L^2 + L^\infty$ where the $L^\infty$ part tends to 0 at infinity. Hence, as $p \in \mathbb{R}^3 \mapsto (p^2+1)^{-1}$ is also in $L^2(\mathbb{R}^3)$, we get that $V(-\Delta+1)^{-1}$ is compact.

3.2.3. *Proof of Theorem 13.* We derive the Euler-Lagrange equations from the minimizing property.

We start with the following observation: each orbital $\psi_i$, $1 \leq i \leq N$ is the minimizer of the minimization problem:

$$\begin{aligned} E_i(\psi) := \inf\{&\mathcal{E}(\psi_1,\ldots,\psi_{i-1},\varphi,\psi_{i+1},\ldots,\psi_N), \\ &\varphi \in H^2(\mathbb{R}^3),\ \|\varphi\|_{L^2(\mathbb{R}^3)} = 1 \ \&\ \langle\psi_j,\varphi\rangle_{L^2(\mathbb{R}^3)} = 0,\ j \neq i\}. \end{aligned}$$

This corresponds to freezing all the orbitals but $\psi_i$ and tuning $\psi_i$.

A computation yields[10]:

$$\mathcal{E}(\psi_1,\ldots,\psi_{i-1},\varphi,\psi_{i+1},\ldots,\psi_N) = \mathcal{E}(\psi_1,\ldots,\psi_{i-1},0,\psi_{i+1},\ldots,\psi_N) + \langle F_\psi^{(i)}\varphi,\varphi\rangle_{L^2(\mathbb{R}^3)}, \tag{11}$$

where $F_\psi^{(i)}$ denotes the mean-field operator with the orbital $\psi_i$ removed:

$$F_\psi^{(i)} := -\Delta + V + (\rho_\psi - |\psi_i|^2) * \frac{1}{|\cdot|} - \left(\frac{\gamma_\psi(x,y)}{|x-y|} - \frac{\psi_i(x)\overline{\psi_i(y)}}{|x-y|}\right).$$

We use the minimization property: let $t \in \mathbb{C}$ and $\varphi \in H^2(\mathbb{R}^3)$ with $\langle\psi_j,\varphi\rangle_{L^2(\mathbb{R}^3)} = 0$, $1 \leq j \leq N$. We have:

$$\mathcal{E}\Big(\psi_1,\ldots,\psi_{i-1},\frac{\psi_i+t\varphi}{\sqrt{1+|t|^2}},\psi_{i+1},\ldots,\psi_N\Big) \geq E_{HF}^N = \mathcal{E}(\psi_1,\ldots,\psi_N).$$

Equivalently, we have

$$\Big\langle F_\psi^{(i)}\frac{\psi_i+t\varphi}{\sqrt{1+|t|^2}},\ \frac{\psi_i+t\varphi}{\sqrt{1+|t|^2}}\Big\rangle_{L^2(\mathbb{R}^3)} - \langle F_\psi^{(i)}\psi_i,\psi_i\rangle_{L^2(\mathbb{R}^3)} \geq 0,$$

which we can rewrite as

$$\frac{2}{1+|t|^2}\mathrm{Re}\,\langle F_\psi^{(i)}\varphi_i, t\varphi\rangle_{L^2(\mathbb{R}^3)} + \frac{|t|^2}{1+|t|^2}\Big(\langle F_\psi^{(i)}\varphi,\varphi\rangle_{L^2(\mathbb{R}^3)} - \langle F_\psi^{(i)}\psi_i,\psi_i\rangle_{L^2(\mathbb{R}^3)}\Big) \geq 0. \tag{12}$$

*Variation at first order*. The first order term (in $t$) in (12) implies that $F_\psi^{(i)}\psi_i$ is orthogonal to $\varphi$ for every $\varphi \in H^2(\mathbb{R}^3)$ orthogonal to the $\psi_j$'s.

In other words: $F_\psi^{(i)}\psi_i \in \mathrm{span}(\psi_j)_{1\leq j\leq N}$. However for $1 \leq i \leq N$ we have:

$$F_\psi^{(i)}\psi_i = F_\psi\psi_i, \tag{13}$$

because[11]

$$|\psi_i|^2 * \frac{1}{|\cdot|}\psi_i - \frac{\psi_i(x)\overline{\psi_i(y)}}{|x-y|}\psi_i = 0.$$

[10]The computation is left as an exercise.

[11]one sometimes say that an electron does not see its own field.

So we obtain the Euler-Lagrange equations:

$$F_\psi \psi_i = \sum_j \lambda_{ij} \psi_j,$$

where $(\lambda_{ij})_{1\le i,j\le N}$ are the Lagrange multipliers. As the operator $F_\psi$ is self-adjoint, we have $\lambda_{ji} = \overline{\lambda_{ij}}$ and the matrix $(\lambda_{ij})_{1\le i,j\le N}$ is Hermitian. So up to unitary $U \in \mathcal{U}(\mathbb{C}^N)$, the Euler-Lagrange equations read:

$$F_\psi \psi_i = \lambda_i \psi_i,$$

with $\lambda_1 \le \cdots \le \lambda_N$. In other words, the E.L. equations say that the $\psi_i$'s are (up to a rotation) eigenfunctions of the mean-field operator.

**Remark 14.** *An electron does not see its own field, but in general for $\varphi \in H^2(\mathbb{R}^3)$, a computation yields:*

$$\begin{aligned}\langle F_\psi^{(i)} \varphi, \varphi\rangle_{L^2(\mathbb{R}^3)} &= \langle F_\psi \varphi, \varphi\rangle_{L^2(\mathbb{R}^3)} - \iint \frac{|\psi_i(x)|^2 |\varphi(x)|^2}{|x-y|} \mathrm{d}x\mathrm{d}y + \iint \frac{(\psi_i\overline{\varphi})(x)(\overline{\psi_i}\varphi)(y)}{|x-y|}\mathrm{d}x\mathrm{d}y,\\ &= \langle F_\psi \varphi, \varphi\rangle_{L^2(\mathbb{R}^3)} - \frac{1}{2}\iint \frac{|\varphi\wedge\psi_i(x,y)|^2}{|x-y|}\mathrm{d}x\mathrm{d}y.\end{aligned}$$

*Variation at second order*. The second order term (in $t$) in (12) implies that for all normalized $\varphi \in H^2(\mathbb{R}^3)$ orthogonal to the $\psi_j$'s we have:

$$\langle F_\psi^{(i)} \varphi, \varphi\rangle_{L^2(\mathbb{R}^3)} \ge \langle F_\psi^{(i)} \psi_i, \psi_i\rangle_{L^2(\mathbb{R}^3)},$$

which can be rewritten thanks to Remark 14 as:

$$\begin{aligned}\langle F_\psi \varphi, \varphi\rangle_{L^2(\mathbb{R}^3)} &\ge \langle F_\psi^{(i)} \psi_i, \psi_i\rangle_{L^2(\mathbb{R}^3)} + \frac{1}{2}\iint \frac{|\varphi\wedge\psi_i(x,y)|^2}{|x-y|}\mathrm{d}x\mathrm{d}y,\\ &\ge \langle F_\psi \psi_i, \psi_i\rangle_{L^2(\mathbb{R}^3)} + \frac{1}{2}\iint \frac{|\varphi\wedge\psi_i(x,y)|^2}{|x-y|}\mathrm{d}x\mathrm{d}y.\end{aligned}$$

This holds for all $1 \le i \le N$, hence by the min-max principle we get:

$$\mu_{N+1}(F_\psi) \ge \lambda_{N+1},$$

which ends the proof. Indeed, if $F_\psi$ has at least $N+1$ eigenvalues below 0, the condition above with $\varphi$ the corresponding $N+1$-th eigenfunction gives $\lambda_{N+1} > \lambda_N$, as $\varphi\wedge\psi_N \ne 0$.

3.2.4. *Proof of Theorem 11.*

We are now ready to prove the existence of a minimizer in the case $N < Z+1$.

The method goes as follows. We will relax the constraint $\mathrm{Gram}(\psi) = \mathbb{1}_{\mathbb{C}^N}$ and show that there exists a minimizer $\psi$ of $\mathcal{E}$ on $\mathcal{K}$ (see (9) for the definition of $\mathcal{K}$).

Then, we will show that when $N < Z+1$, the minimizer $\psi$ is in fact in $\mathcal{M}$, that is that it satisfies $\mathrm{Gram}(\psi) = \mathbb{1}_{\mathbb{C}^N}$. In particular it gives rise to a minimizer $\psi_1\wedge\cdots\wedge\psi_N$ of $E_{HF}^N$.

*Existence of a minimizer of $E_\mathcal{K} = \inf_{\psi\in\mathcal{K}} \mathcal{E}(\psi)$.* We use the direct method of calculus of variations. Let $(\psi^{(n)})_n$ be a minimizing sequence for $E_\mathcal{K}$, that is $\psi^{(n)} \in \mathcal{K}$ and:

$$\lim_{n\to+\infty} \mathcal{E}(\psi^{(n)}) = E_\mathcal{K}.$$

By Proposition 8, the sequence $(\psi^{(n)})_n$ is $H^1$-bounded. By the theorem of Banach-Alaoglu, up to extracting a subsequence, we can assume that it converges $H^1$-weakly to $\psi \in \mathcal{K}$. And Proposition 8 ensures us that $\psi$ is a minimizer as $\mathcal{E}$ is $H^1$-weakly lower semi-continuous on $\mathcal{K}$.

*End of the proof*. Consider the minimizer $\psi \in \mathcal{K}$ for $E_{\mathcal{K}}$. We know that $E_{\mathcal{K}} \leq 0$ as $0 \in \mathcal{K}$.

Up to applying a unitary $U \in \mathcal{U}(\mathbb{C}^N)$ to $\psi$, we can assume that $\mathrm{Gram}(\psi) = \mathrm{diag}(\varepsilon_1, \dots, \varepsilon_N)$ with $0 \leq \varepsilon_i \leq 1$. In other words, we can assume that for all $i \neq j$ there holds $\langle \psi_i, \psi_j \rangle_{L^2(\mathbb{R}^3)} = 0$.

As in the proof of Theorem 13, each $\psi_i$ is a minimizer of the minimization problem:

$$\inf\{\mathcal{E}(\psi_1, \dots, \psi_{i-1}, \varphi, \psi_{i+1}, \dots, \psi_N), \\ \varphi \in H^1(\mathbb{R}^3),\ \|\varphi\|_{L^2(\mathbb{R}^3)} \leq 1,\ \langle \psi_j, \varphi \rangle_{L^2(\mathbb{R}^3)} = 0,\ j \neq i\}.$$

Equivalently (see (11)), it is a minimizer for:

$$I_i := \inf\{\langle F_\psi^{(i)} \varphi, \varphi \rangle_{L^2(\mathbb{R}^3)},\ \varphi \in H^1(\mathbb{R}^3),\ \|\varphi\|_{L^2(\mathbb{R}^3)} \leq 1,\ \langle \psi_j, \varphi \rangle_{L^2(\mathbb{R}^3)} = 0,\ j \neq i\},$$

where the expectation is to be understood in the quadratic form sense ($H^1(\mathbb{R}^3)$ is the form domain of $F_\psi^{(i)}$). Taking $\varphi = 0$ we have $I_i \leq 0$, hence either $I_i = 0$ or $I_i < 0$.

*Claim:* it suffices to prove that if $N < Z + 1$ then $I_i < 0$.

Indeed, in that case, as $F_\psi^{(i)}$ is bounded from below, and the minimizer $\psi_i$ is non-zero (the essential spectrum of $F_\psi^{(i)}$ is $[0, +\infty)$). Up to considering $\frac{\psi_i}{\|\psi_i\|_{L^2}}$, this minimizer has necessarily norm 1 and as this holds for all $1 \leq i \leq N$, this gives

$$\mathrm{Gram}(\psi) = \mathbb{1}_{\mathbb{C}^N}.$$

**Lemma 15.** *Let $\psi \in \mathcal{K}$. If $N < Z + 1$, then $F_\psi^{(i)}$ has an infinite discrete spectrum below its essential spectrum* $[0, +\infty)$.

*Proof of the Lemma.* First we observe that $R_{\gamma_\psi}$ is a non-negative operator as for all $\varphi \in L^2(\mathbb{R}^3)$ and $1 \leq i \leq N$ we have:

$$\begin{aligned}\langle \varphi, R_{|\psi_i\rangle\langle\psi_i|} \varphi \rangle_{L^2(\mathbb{R}^3)} &:= \iint \frac{\overline{\varphi}(x)\psi_i(x)\overline{\psi_i(y)}\varphi(y)}{|x-y|} \mathrm{d}x\mathrm{d}y, \\ &= 4\pi \int_{\mathbb{R}^3} \frac{|(\overline{\psi}_i \varphi)\hat{\ }(p)|^2}{|p|} \mathrm{d}p \geq 0.\end{aligned}$$

Hence we have the quadratic form inequality:

$$F_\psi^{(i)} \leq -\Delta + V + (\rho_\psi - |\psi_i|^2) * \frac{1}{|\cdot|} =: G_i.$$

By the min-max principle, it suffices to show that $G_i$ has infinite discrete spectrum below $[0, +\infty)$, and to do so we will use the Rayleigh-Ritz method.

The reason is that $G_i$ takes the form $-\Delta + \mu * \frac{1}{|\cdot|}$, where $\mu$ is a finite Borel measure with $\mu(\mathbb{R}^3) \leq Z - (N-1) < 0$.

As we have done in a previous lecture, we introduce a *radial* function $\chi \in C_0^\infty(\mathbb{R}^3, [0, +\infty))$ with norm 1 and whose support lies in the annulus $\{x \in \mathbb{R}^3,\ 1 < |x| < 2\}$. For every $R > 1$, we write:

$$\chi_R(x) := R^{-3/2} \chi(x/R).$$

For a given $R_0 \geq 1$, the family $(\chi_{2^n R_0})_{n \geq 0}$ is an orthonormal family.

For $R > 1$, the following estimates hold.

- First we have

$$\|\nabla \chi_R\|_{L^2}^2 = R^{-2} \|\nabla \chi\|_{L^2}^2.$$

- Then we have:

$$\int |\psi_j|^2 * \frac{1}{|\cdot|}|\chi_R|^2 = \int_x R^{-3}\mathrm{d}x|\chi(R^{-1}x)|^2 \int_y \frac{|\psi_j(y)|^2}{|y-x|}\mathrm{d}y,$$
$$= \frac{1}{R}\int_x \mathrm{d}x|\chi(x)|^2 \int_y \frac{|\psi_j(y)|^2}{|y/R-x|}\mathrm{d}y,$$
$$= \frac{1}{R}\int_y \mathrm{d}y|\psi_j(y)|^2 \int_x \frac{|\chi(x)|^2}{|y/R-x|}\mathrm{d}x.$$

As the function $\chi$ is radial: $\chi(x) = f(|x|)$, a computation gives:

$$\int_x \frac{|\chi(x)|^2}{|y-x|}\mathrm{d}x = 2\pi \int_0^{+\infty} r^2 f(r)^2 \mathrm{d}r \int_{\theta\in[0,\pi]} \frac{\sin(\theta)\mathrm{d}\theta}{\sqrt{r^2+|y|^2-2\cos(\theta)r|y|}},$$
$$= 2\pi \int_0^{+\infty} r^2 f(r)^2 \frac{(r+|y|)-|r-|y||}{r|y|}\mathrm{d}r,$$
$$(\textbf{Newton's formula}) = \frac{1}{|y|}\int_{x:|x|\le|y|} |\chi(x)|^2\mathrm{d}x + \int_{x:|x|\ge|y|} \frac{|\chi(x)|^2}{|x|}\mathrm{d}x,$$

and thus

$$\int |\psi_j|^2 * \frac{1}{|\cdot|}|\chi_R|^2 \le \int_y |\psi_i(y)|^2 \frac{1}{R}\int_x \frac{|\chi(x)|}{|x|}\mathrm{d}x\mathrm{d}y \le \frac{1}{R}\int_x \frac{|\chi(x)|}{|x|}\mathrm{d}x.$$

- Similarly, for $R > 0$ large enough, we have:

$$-z_m \int_x \frac{|\chi_R(x)|^2}{|x-R_m|}\mathrm{d}x = -\frac{z_m}{R}\int \frac{|\chi(x)|^2}{|x-\frac{R_m}{R}|}\mathrm{d}x,$$
$$= -\frac{z_m}{R}\int \frac{|\chi(x)|^2}{|x|}\mathrm{d}x.$$

(We have used Newton's formula and the fact that supp $\chi \subset \{x,\ |x| \ge 1\}$).

So for $n$ large enough (say $n \ge n_0$), we have:

$$\langle \chi_{2^nR_0}, G_i\chi_{2^nR_0}\rangle_{L^2(\mathbb{R}^3)} \le \frac{N-1-Z}{2^nR_0}\int \frac{|\chi(x)|^2}{|x|}\mathrm{d}x + \frac{1}{4^nR_0^2}\|\nabla\chi\|_{L^2}^2 < 0.$$

This ends the proof by the Rayleigh-Ritz principle applied to the orthonormal family $(\chi_{2^nR_0})_{n\ge n_0}$.

□

## 4. The Hartree-Fock functional as a function of the reduced one-particle density matrix

4.1. **Reminder on trace-class operators.** As we have seen, we can see the Hartree-Fock functional as a function of the reduced one-body density matrix only. The formula involves the trace. In this course we did not properly introduce trace-class operators: we refer the reader to [4, Sec. VI.6].

In this part we only states results without proof.

We recall that for a non-negative operator $\gamma \ge 0$, the trace $\mathrm{tr}(\gamma)$ is defined as:

$$\mathrm{tr}(\gamma) := \sum_i \langle \varphi_i, \gamma\varphi_i\rangle,$$

where $(\varphi_i)_i$ is any ONB of the underlying Hilbert space (here $L^2(\mathbb{R}^3)$). First, monotone convergence ensures that this formula is well-defined for a given ONB, and Parseval's identity together with the fact that $\langle \varphi_i, \gamma\varphi_i\rangle = \|\sqrt{\gamma}\varphi_i\|^2$ ensures us that this formula does not depend on the ONB.

Trace-class operators are the bounded operators $\gamma$ for which $|\gamma| = \sqrt{\gamma^*\gamma}$ has finite trace. Such operators are necessarily compact, and the family of their singular values define an element in $\ell^1(\mathbb{N})$. The latter condition characterizes trace-class operators.

We can then define the trace if a trace-class opeartor $\gamma$ by the same formula: the formula does not depend on the ONB and the corresponding series are absolutely convergent.

It can be shown that the set of trace-class operators form a $*$-ideal of bounded operators. Furthermore it is also a Banach space for the norm:

$$\|\gamma\|_{\mathfrak{S}_1} := \mathrm{tr}\ |\gamma|.$$

Using the canonical form of compact operators, we can show that the trace-class operators are precisely the operators that can be written as a product of two Hilbert-Schmift operators. In $L^2(\mathbb{R}^3)$, a trace-class operator $\gamma$ has an integral kernel $\gamma(x, y)$ and we have:

$$\mathrm{tr}(\gamma) = \int \gamma(x, x)\mathrm{d}x.$$

We also emphasize that the space $\mathfrak{S}_1(L^2(\mathbb{R}^3))$ of trace-class opeartors is, through the trace, the dual of compact operators on $L^2(\mathbb{R}^3)$ (with the operator norm), and that its dual is the banach space of bounded operators on $L^2(\mathbb{R}^3)$. This is very similar to the results know for sequences:

$$\ell^1(\mathbb{N}) = c_0(\mathbb{N})' \quad \& \quad (\ell^1(\mathbb{N}))' = \ell^\infty(\mathbb{N}).$$

Here for a compact operator $K$ and a trace-class operator $\gamma$, the duality is given by

$$\langle \gamma, K\rangle_{\mathrm{Comp}'\times\mathrm{Comp}} = \mathrm{tr}(\gamma K).$$

4.2. **Reduced density matrix of Slater determinant.** We have seen that for a normalized $\psi \in L^2_a(\mathbb{R}^{3N})$, we have $0 \le \gamma_\psi \le 1$, and that it is a rank $N$ projector if $\psi$ is a Slater determinant. In fact the converse is also true.

**Lemma 16.** *Let a normalized $\psi \in L^2_a(\mathbb{R}^{3N})$. Then $\gamma_\psi$ is an orthonormal projector* iff *$\psi$ is a Slater determinant*

*Proof.* The proof is similar as the one in Section 5.2. Let $(\varphi_i)_{i\ge i_0}$ be an ONB of $L^2(\mathbb{R}^3)$ such that $(\varphi_i)_{i_0\le i\le 0}$ is an ONB of $\mathrm{span}\,\gamma_\psi$ and $(\varphi_i)_{i\ge 1}$ is an ONB of $\ker\gamma_\psi$.

For a subset $I = \{i_1 < \cdots < i_N\} \subset \mathbb{N} + i_0$ of $N$ elements, we write $\varphi_I$ the Slater determinant $\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}$.

Let us decompose:

$$\psi = \sum_I c_I \varphi_I.$$

A computation yields

$$\begin{aligned}\gamma_\psi(x, y) &= N \int \psi(x, \underline{x}')\overline{\psi(y, \underline{x}')}\mathrm{d}\underline{x}',\\ &= \sum_{\substack{I\subset[|i_0,0|]\\ |I|=N}} |c_I|^2 \gamma_{\varphi_I}.\end{aligned}$$

We recall that $|i_0, 0|]$ denotes $\{i_0, i_0+1, \dots, -1, 0\}$ or $\mathbb{Z}_-$ if $i_0 = -\infty$. From this formula, it is easy to check the equivalence of the two assertions. $\square$

The proof also shows that $\gamma_\psi$ is in the convex hull of reduced one-body density matrices of Slater determinants.

More generally the following holds.

**Lemma 17.** *The closed convex hull of*

$$\mathcal{P}_N := \{\gamma \text{ rank } N \text{orthogonal projector in } L^2(\mathbb{R}^3)\}$$

*under the trace-norm is the convex set:*

$$\mathcal{K}_{\gamma,N} := \{\gamma \in \mathcal{L}(L^2(\mathbb{R}^3)),\ 0 \leq \gamma \leq 1,\ \mathrm{tr}(\gamma) = N\}.$$

*The set $\mathcal{K}_{\gamma,N}$ is also the closure of $\mathcal{P}_N$ under the $*$-weak topology of the space of trace-class operators.*

*Proof.* The proof of the first assertion goes as follows. The condition $0 \leq \gamma \leq 1$ is continuous in the $*$-weak topology of $\mathfrak{S}_1$ as we have:

$$\langle \psi, \gamma\psi \rangle = \mathrm{tr}(\gamma|\psi\rangle\langle\psi|).$$

By density, it then suffices to show that finite rank elements $\gamma$ of $\mathcal{K}_{\gamma,N}$ can be written as a convex combination of elements in $\mathcal{P}_N$. Using the spectral theorem on $\gamma$, this result can be shown by induction on the rank of $\gamma$.

The proof of the second assertion is very similar to that of Lemma 6 (take the same sequence *mutatis mutandis*) and is left to the reader. (By density it suffices to show that finite rank elements $\gamma \in \mathcal{K}_{\gamma,N}$ are limits of a weakly converging sequence in $\mathcal{P}_N$).

The details are left to the reader. □

4.3. **The Hartree-Fock functional and Lieb's variational principle.** We recall that the Hartree-Fock functional is given formally by:

$$\mathcal{E}_{HF}(\gamma) = \mathrm{tr}((-\Delta + V)\gamma) + D(\rho_\gamma) - X(\gamma),$$

where $\rho_\gamma(x)$ denotes the diagonal $\gamma(x,x)$.

Originally, it was only defined on the one-body density matrices of Slater determinants. The trace $\mathrm{tr}(-\Delta\gamma)$ has to be understood in the quadratic form sense (in the Hilbert space of Hilbert-Schmidt operators) and denotes:

$$\mathrm{tr}(-\Delta\gamma) = \sum_i n_i \|\nabla\psi_i\|_{L^2}^2,$$

where $\gamma = \sum_i n_i |\psi_i\rangle\langle\psi_i|$ is the spectral decomposition of $\gamma$. If we write $\hat{\gamma}(p,q)$ the integral kernel of its Fourier transform $\mathscr{F}\gamma\mathscr{F}^{-1}$, we also have the formula:

$$\mathrm{tr}(-\Delta\gamma) = \int_{\mathbb{R}^3} |p|^2 \hat{\gamma}(p,p).$$

Having (17) in mind, we can relax the constraint and extend $\mathcal{E}_{HF}$ to

$$\mathcal{A}_{\gamma,N} := \big\{\gamma \in \mathcal{K}_{\gamma,N},\ \mathrm{tr}(-\Delta\gamma) < +\infty\big\},$$

and we consider the corresponding variational problem

$$E_{gHF}^N := \inf_{\gamma_0 \in \mathcal{A}_{\gamma,N}} \mathcal{E}_{HF}(\gamma_0).$$

**Lemma 18.** *The Hartree-Fock functional is well-defined and bounded from below on $\mathcal{A}_{\gamma,N}$. Furthermore we have $E_{gHF}^N = E_{HF}^N$, and if there is a minimizer, then it is the reduced density matrix of a Slater determinant.*

The second part of the Lemma follows from Lieb's variationnal principle, which we state for $\mathcal{A}_{\gamma,N}$.

**Lemma 19** (Lieb's variationnal principle)**.** *Let $\gamma \in \mathcal{A}_{\gamma,N}$, then there exists a rank $N$ projector $\gamma_1$ in $\mathcal{A}_{\gamma,N}$ such that $\mathcal{E}_{HF}(\gamma_1) \leq \mathcal{E}_{HF}(\gamma)$.*

*Proof.* The fact that $\mathcal{E}_{HF}$ is well-defined and bounded from below on $\mathcal{A}_{\gamma,N}$ follows from Hardy's inequality and the fact that $\sqrt{\gamma_\psi} \in H^1(\mathbb{R}^3)$. We prove below the last statement and give a partial proof of the first result to the reader.

For $\gamma \in \mathcal{A}_{\gamma,N}$, we consider its spectral decomposition $\gamma = \sum_i n_i |\psi_i\rangle\langle\psi_i|$, where $n_i \geq 0$. We have

$$\|\nabla\sqrt{\rho_\gamma}\|_{L^2}^2 = \int \frac{|\nabla\rho_\gamma|^2}{4\rho_\gamma} \leq \sum_i n_i \|\nabla\psi_i\|_{L^2}^2 = \mathrm{tr}(-\Delta\gamma),$$

where we have used Cauchy-Schwarz's inequality to get

$$\left|2\mathrm{Re}\sum_i n_i\overline{\psi_i}\nabla\psi_i\right|^2 \leq 4\sum_i n_i|\psi_i|^2 \sum_i n_i|\nabla\psi_i|^2.$$

We leave as an exercise the proof of

$$\begin{aligned} X(\gamma_\psi) &\leq 2\|\gamma\|_{\mathfrak{S}_2}\sqrt{\mathrm{tr}(-\Delta\gamma)} \leq 2\sqrt{N\,\mathrm{tr}(-\Delta\gamma)}, \\ D(\rho_\gamma) &\leq 2\|\sqrt{\rho_\gamma}\|_{L^2}\|\nabla\sqrt{\rho_\gamma}\|_{L^2} \leq 2\sqrt{N\,\mathrm{tr}(-\Delta\gamma)}, \end{aligned}$$

and the end of the proof.

*Lieb's variationnal principle.* Lett $\gamma \in \mathcal{A}_{\gamma,N}$, we consider its spectral decomposition $\gamma = \sum_{i\geq 1} n_i|\psi_i\rangle\langle\psi_i|$, where $0 \leq n_i \leq 1$. Assume that there exists $i_1$ with $0 < n_{i_1} < 1$. As $\sum_i n_i = N$, there exists necessarily a second index $i_2$ with $0 < n_{i_2} < 1$. W.l.o.g. we can assume that $i_1 = 1$ and $i_2 = 2$ and write:

$$\gamma = n_1|\psi_1\rangle\langle\psi_1| + n_2|\psi_2\rangle\langle\psi_2| + g.$$

Let $\delta\gamma := |\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2|$. Let $t \in \mathbb{R}$: for $|t|$ small enough we still have $\gamma_t := \gamma + t\delta\gamma \in \mathcal{A}_{\gamma,N}$. A computation yields

$$\begin{aligned} \mathcal{E}_{HF}(\gamma_t) &= \mathcal{E}_{HF}(\gamma) + t\big(\langle F_\gamma\psi_1,\psi_1\rangle_{L^2} - \langle F_\gamma\psi_2,\psi_2\rangle_{L^2}\big) + t^2\big(D(\rho_{\delta\gamma}) - X(\delta\gamma)\big), \\ &= \mathcal{E}_{HF}(\gamma) + t\big(\langle F_\gamma\psi_1,\psi_1\rangle_{L^2} - \langle F_\gamma\psi_2,\psi_2\rangle_{L^2}\big) - t^2\iint \frac{|\psi_1\wedge\psi_2(x,y)|^2}{|x-y|}\mathrm{d}x\mathrm{d}y, \end{aligned}$$

where we recall that $F_\gamma = -\Delta + V + \rho_\gamma * \frac{1}{|\cdot|} - R_\gamma$.

Thus optimizing in $t$, we can choose $t_0 \in \mathbb{R}$ such that $\gamma_{t_0} \in \mathcal{A}_{\gamma,N}$ has a smaller energy and satisfies either $\gamma_{t_0}\psi_1 = \psi_1$ or $\gamma_{t_0}\psi_2 = \psi_2$.

By induction[12] on the indices $i$ such that $0 < n_i < 1$, we find a rank $N$ projector which has a smaller energy.

The same argument shows that a minimizer must necessarily be a projector. □

## 5. Technical proofs

In the first two subsections, the proofs are all about spotting the permutations in $\mathcal{S}_N$ that have a non-trivial contribution to the formula at hand.

### 5.1. Proof of (4).

To show the first inequality observe that for $f \in L^2(\mathbb{R}^3)$, we have:

$$\begin{aligned} \langle f, \gamma_\psi f\rangle_{L^2(\mathbb{R}^3)} &:= N\iint \mathrm{d}x\mathrm{d}y \int_{\underline{x}'} \overline{f(x)}\psi(x,\underline{x}')\overline{\psi(y,\underline{x}')}f(y)\mathrm{d}\underline{x}', \\ &= N\int\left|\int \overline{f(x)}\psi(x,\underline{x}')\mathrm{d}x\right|^2\mathrm{d}\underline{x}'. \end{aligned} \tag{14}$$

[12]The argument as it is only works for finite rank operators. To make the argument rigorous in general, you can for instance choose at each step two of the eigenvalues of $\gamma$ which are the farthest from 0 and 1. By construction, you end up with a sequence which converges in trace-norm to a projector. You then have to use a weak semi lower-continuity argument.

To show that $\gamma_\psi \leq 1$, we can decompose[13] $\psi$ w.r.t. a basis of Slater determinants:

$$\psi = \sum_{\substack{I\subset\mathbb{N}\\|I|=N}} c_I\varphi_I,$$

where $\varphi_I$ denotes $\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N}$ with $\{i_1<i_2<\cdots<i_N\}=I$ and $(\varphi_i)_i$ is an ONB.

W.l.o.g. we can assume that $\|f\|_{L^2}=1$ and $f=\varphi_1$. By density we can assume that the decomposition of $\psi$ is finite.

From the formula of the Slater determinants $\varphi_I$, the expression 14 for $\psi=\varphi_I$ is non-zero if and only if $1\in I$. Recall that we write $I=\{i_1<\cdots<i_N\}$, $\underline{x}=(x_1,\ldots,x_N)$, and that we assume $f=\varphi_1$. We thus have:

$$\begin{aligned}
\sqrt{N!}\int_{x\in\mathbb{R}^3}\overline{f(x_1)}\psi(x_1,\underline{x}')\mathrm{d}x_1 &= \sqrt{N!}\sum_I c_I\int_{x\in\mathbb{R}^3}\overline{f(x_1)}\varphi_I(x_1,\underline{x}')\mathrm{d}x_1,\\
&=\sum_I\frac{c_I}{\sqrt{(N-1)!}}\sum_{\sigma\in\mathcal{S}_N}\varepsilon(\sigma)\overline{f(x_1)}\prod_{j=1}^N\varphi_{i_{\sigma(j)}}(x_j)\mathrm{d}x_1,\\
&=\sum_{I:\,1\in I}\frac{c_I}{\sqrt{(N-1)!}}\sum_{\substack{\sigma\in\mathcal{S}_N\\ \sigma(1)=1}}\varepsilon(\sigma)\overline{f(x_1)}\prod_{j=1}^N\varphi_{i_{\sigma(j)}}(x_j)\mathrm{d}x_1,\\
&=\sum_{I:\,I=\{1\}\cup I'}c_I\varphi_{I'},
\end{aligned}$$

where $\varphi_{I'}$ denotes the Slater determinant $\varphi_{i_2}\wedge\cdots\wedge\varphi_{i_N}$ in $L^2_a(\mathbb{R}^{3(N-1)})$.

We thus can rewrite $\langle f,\gamma_\psi f\rangle$ as follows:

$$\begin{aligned}
\langle\varphi_1,\gamma_\psi\varphi_1\rangle_{L^2(\mathbb{R}^3)} &= \sum_{I:\,I=\{1\}\cup I'}\ \sum_{J:\,J=\{1\}\cup J'}\langle c_I\varphi_{I'},c_J\varphi_{J'}\rangle_{L^2_a(\mathbb{R}^{3(N-1)})},\\
&=\sum_{I:\,I=\{1\}\cup I'}|c_I|^2\leq 1.
\end{aligned}$$

5.2. **Proof of Lemma 3.** Let $\psi=\psi_1\wedge\cdots\wedge\psi_n$ be a Slater determinant.

We first compute the expectation of $A:=\sum_i(-\Delta_{x_i}+V(x_i))$, and write $A_i:=-\Delta_{x_i}+V(x_i)$.

We then compute the expectation of $B:=\sum_{i<j}\frac{1}{|x_i-x_j|}$.

5.2.1. *The term A.* We have:

$$\begin{aligned}
\sum_i\langle A\psi,\psi\rangle_{L^2_a(\mathbb{R}^{3N})} &= \sum_i\sum_{\sigma,\tau\in\mathcal{S}_N}\frac{\varepsilon(\sigma)\varepsilon(\tau)}{N!}\langle A_i\prod_k\psi_{\sigma(k)}(x_k),\prod_\ell\psi_{\tau(\ell)}(x_\ell)\rangle_{L^2_a(\mathbb{R}^{3N})},\\
&=\sum_i\sum_{\sigma,\tau\in\mathcal{S}_N}\frac{\varepsilon(\sigma)\varepsilon(\tau)}{N!}\langle A_i\psi_{\sigma(i)}(x_i),\psi_{\tau(i)}(x_i)\rangle_{L^2(\mathbb{R}^3)}\prod_{k\neq i}\langle\psi_{\sigma(k)}(x_k),\psi_{\tau(k)}(x_k)\rangle_{L^2(\mathbb{R}^3)},\\
&=\sum_i\sum_{\sigma,\tau\in\mathcal{S}_N}\frac{\varepsilon(\sigma)\varepsilon(\tau)}{N!}\langle A_i\psi_{\sigma(i)}(x_i),\psi_{\tau(i)}(x_i)\rangle_{L^2(\mathbb{R}^3)}\prod_{k\neq i}\delta_{\sigma(k)\,\tau(k)}.
\end{aligned}$$

Each term in the above sum is non-zero *iff* for all $k\neq i$ we have $\sigma(k)=\tau(k)$, that is *iff* $\sigma=\tau$.

[13] there is a very short proof using the CAR and the fact that $\langle f,\gamma_\psi g\rangle_{L^2(\mathbb{R}^3)}=\langle a(g)\psi,a(f)\psi\rangle_{L^2_a(\mathbb{R}^{3N})}$.

Then, given $1 \le i, j \le N$, there exist $(N-1)!$ permutations of $\mathbb{S}_N$ mapping $i$ to $j$. Hence we get:

$$\sum_i \langle A\psi, \psi\rangle_{L^2_a(\mathbb{R}^{3N})} = \frac{1}{N}\sum_i\sum_j \langle A_i\psi_j(x_i), \psi_j(x_i)\rangle_{L^2(\mathbb{R}^3)},$$
$$= \frac{1}{N}\sum_i\sum_j \langle A_i\psi_j, \psi_j\rangle_{L^2(\mathbb{R}^3)} = \sum_j \langle(-\Delta + V)\psi_j, \psi_j\rangle_{L^2(\mathbb{R}^3)}.$$

5.2.2. *The term B.* For the expectation of $B$, we use the antisymmetry and consider the change of variables which switch $x_1$ and $x_i$, and $x_2$ and $x_j$. This gives:

$$\langle B\psi, \psi\rangle_{L^2_a(\mathbb{R}^{3N})} = \sum_{1\le i<j\le N} \int \frac{|\psi(\underline{x})|^2}{|x_i - x_j|}\mathrm{d}\underline{x} = \frac{N(N-1)}{2}\int \frac{|\psi(\underline{x})|^2}{|x_1 - x_2|}\mathrm{d}\underline{x},$$
$$= \frac{N(N-1)}{2}\sum_{\sigma,\tau\in\mathcal{S}_N}\frac{\varepsilon(\sigma)\varepsilon(\tau)}{N!}\int_{x_1,x_2}\frac{\mathrm{d}x_1\mathrm{d}x_2}{|x_1 - x_2|}\int_{\underline{x}'}\prod_k \overline{\psi_{\sigma(k)}(x_k)}\prod_\ell \psi_{\sigma(\ell)}(x_\ell)\mathrm{d}\underline{x}',$$
$$= \frac{1}{2\,(N-2)!}\sum_{\sigma,\tau\in\mathcal{S}_N}\varepsilon(\sigma)\varepsilon(\tau)\int_{x_1,x_2}\frac{\mathrm{d}x_1\mathrm{d}x_2}{|x_1 - x_2|}\prod_{k=1}^2 \overline{\psi_{\sigma(k)}(x_k)}\psi_{\tau(k)}(x_k)\prod_{3\le k'\le N}\langle\psi_{\sigma(k')}, \psi_{\tau(k')}\rangle_{L^2(\mathbb{R}^3)},$$
$$= \frac{1}{2\,(N-2)!}\sum_{\sigma,\tau\in\mathcal{S}_N}\varepsilon(\sigma)\varepsilon(\tau)\int_{x_1,x_2}\frac{\mathrm{d}x_1\mathrm{d}x_2}{|x_1 - x_2|}\prod_{k=1}^2 \overline{\psi_{\sigma(k)}(x_k)}\psi_{\sigma(k)}(x_k)\prod_{3\le k'\le N}\delta_{\sigma(k')\,\tau(k')}.$$

The terms of the above sum are non-zero *iff* for all $3 \le k \le N$ there holds $\sigma(k) = \tau(k)$. Hence either $\tau = \sigma$ or $\tau = (1\,2)\circ\sigma$, in which case $\varepsilon(\sigma)\varepsilon(\tau)$ is equal to 1 resp. $-1$.

Thus

$$\langle B\psi, \psi\rangle_{L^2_a(\mathbb{R}^{3N})} = \frac{1}{2\,(N-2)!}\sum_{\sigma\in\mathcal{S}_N}\frac{\mathrm{d}x_1\mathrm{d}x_2}{|x_1 - x_2|}\big(|\psi_{\sigma(1)}(x_1)|^2|\psi_{\sigma(2)}(x_2)|^2 - [\psi_{\sigma(1)}\overline{\psi_{\sigma(2)}}](x_1)[\psi_{\sigma(2)}\overline{\psi_{\sigma(1)}}](x_2)\big).$$

For given $i \ne j$, there are $(N-2)!$ permutations of $\mathcal{S}_N$ with $\sigma(1) = i$ and $\sigma(2) = j$. Hence, there holds:

$$\langle B\psi, \psi\rangle_{L^2_a(\mathbb{R}^{3N})} = \frac{1}{2}\sum_{i\ne j}\int_{x_1,x_2}\frac{\mathrm{d}x_1\mathrm{d}x_2}{|x_1 - x_2|}\big(|\psi_i(x_1)|^2|\psi_j(x_2)|^2 - [\psi_i\overline{\psi_j}](x_1)[\psi_j\overline{\psi_i}](x_2)\big),$$
$$= \frac{1}{2}\sum_{1\le i,j\le N}\int_{x_1,x_2}\frac{\mathrm{d}x_1\mathrm{d}x_2}{|x_1 - x_2|}\big(|\psi_i(x_1)|^2|\psi_j(x_2)|^2 - [\psi_i\overline{\psi_j}](x_1)[\psi_j\overline{\psi_i}](x_2)\big),$$
$$= \frac{1}{2}\int_{x_1,x_2}\frac{\rho_\psi(x_1)\rho_\psi(x_2) - |\gamma_\psi(x_1, x_2)|^2}{|x_1 - x_2|}\mathrm{d}x_1\mathrm{d}x_2.$$

This ends the proof.

## 5.3. **Proof of Prop. 8.**

5.3.1. *Strong continuity.* The fact that $\mathcal{E}$ is $\|\cdot\|_{H^1}$-continuous follows from Hardy's inequality. Let us deal for instance with the exchange term.

The function $\psi \mapsto X(\gamma_\psi)$ is quartic in $\psi$ (that is homogeneous of degree 4). By Hardy's inequality, the corresponding quadrilinear function is continuous from $(H^1(\mathbb{R}^3)^N)^4$ to $\mathbb{R}$. Indeed, introducing:

$$X_p \quad \begin{array}{ccc} (H^1(\mathbb{R}^3)^N)^4 & \longrightarrow & \mathbb{R}, \\ (\tau, \varphi, \chi, \psi) & \mapsto & \iint \frac{\tau(x)\varphi(x)\chi(y)\psi(y)}{|x-y|}\mathrm{d}x\mathrm{d}y, \end{array}$$

we have:

$$\begin{aligned}|X_p(\tau,\varphi,\chi,\psi)| &\le \iint \frac{|\tau(x)\varphi(x)\chi(y)\psi(y)|}{|x-y|}\mathrm{d}x\mathrm{d}y,\\ &\le \sqrt{\iint \frac{|\tau(x)|^2|\chi(y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y \iint |\varphi(x)|^2|\psi(y)|^2\mathrm{d}x\mathrm{d}y},\\ &\le 2\|\nabla\tau\|_{L^2}\|\varphi\|_{L^2}\|\chi\|_{L^2}\|\psi\|_{L^2}.\end{aligned}$$

The other terms are dealt with in a similar manner.

5.3.2. *Weak lower semi-continuity.* Let us show that $\mathcal{E}$ is weakly lower semi-continuous. Let $\psi^{(n)} \rightharpoonup \psi$ in $H^1(\mathbb{R}^3)^N$. Up to extracting a subsequence, we can assume that

$$\liminf_{n\to+\infty} \mathcal{E}(\psi^{(n)}) = \lim_{n\to+\infty} \mathcal{E}(\psi^{(n)}).$$

By Theorem 10, up to extracting a subsequence, we can assume that the sequence $(\psi^{(n)})_n$ converges in $L^2_{\mathrm{loc}}(\mathbb{R}^3)^N$.

Similarly, up to extracting a subsequence[14], we can assume that $(\psi^{(n)})_n$ converges almost everywhere in $\mathbb{R}^3$.

By the uniform boundedness principle[15] for all $1\le i\le n$, we have:

$$\|\nabla\psi_i\|^2_{L^2} \le \liminf_{n\to+\infty}\|\nabla\psi_i^{(n)}\|^2_{L^2}.$$

Similarly it gives: $\sup_{n\to+\infty}\|\psi_i^{(n)}\|_{H^1}\le C$.

Secondly, for every $A>0$, we have:

$$\begin{aligned}\int \frac{|\psi_i^{(n)}(x)|^2}{|x-R_m|}\mathrm{d}x &= \int_{x:|x-R_m|\le A}\frac{|\psi_i^{(n)}(x)|^2}{|x-R_m|}\mathrm{d}x + \int_{x:|x-R_m|>A}\frac{|\psi_i^{(n)}(x)|^2}{|x-R_m|}\mathrm{d}x,\\ &= \int_{x:|x-R_m|\le A}\frac{|\psi_i^{(n)}(x)|^2}{|x-R_m|}\mathrm{d}x + \underset{A\to+\infty}{\mathcal{O}}\left(\frac{C}{A}\right).\end{aligned}$$

By the $L^2_{\mathrm{loc}}$-convergence, and using the fact that

$$\int_{x:|x-R_m|\le A}\frac{\tau(x)\varphi(x)}{|x-R_m|}\mathrm{d}x \le 2\|\nabla\tau\|_{L^2(\mathbb{R}^3)}\|\varphi\|_{L^2(B(R_m,A))},$$

we get that:

$$\lim_{n\to+\infty}\int_{x:|x-R_m|\le A}\frac{|\psi_i^{(n)}(x)|^2}{|x-R_m|}\mathrm{d}x = \int_{x:|x-R_m|\le A}\frac{|\psi_i(x)|^2}{|x-R_m|}\mathrm{d}x.$$

Hence:

$$\lim_{n\to+\infty}\int_{\mathbb{R}^3} V\|\psi^{(n)}\|^2_{\mathbb{C}^N} = \int_{\mathbb{R}^3} V\|\psi\|^2_{\mathbb{C}^N}.$$

At last we deal with the direct and the exchange terms altogether. Recall (7): for $\varphi\in H^1(\mathbb{R}^3)^N$ there holds:

$$\rho_\psi(x)\rho_\psi(y) - |\gamma_\psi(x,y)|^2 \ge 0.$$

By Fatou's lemma we thus have:

$$\iint \frac{\rho_\psi(x)\rho_\psi(y)-|\gamma_\psi(x,y)|^2}{|x-y|}\mathrm{d}x\mathrm{d}y \le \liminf_{n\to+\infty}\iint \frac{\rho_{\psi^{(n)}}(x)\rho_{\psi^{(n)}}(y)-|\gamma_{\psi^{(n)}}(x,y)|^2}{|x-y|}\mathrm{d}x\mathrm{d}y.$$

[14] Note that we have to use a diagonal extraction here.

[15] why?

We thus have proved:

$$\mathcal{E}(\psi) \leq \liminf_{n\to+\infty} \mathcal{E}(\psi^{(n)}).$$

5.3.3. *Estimate.* Let $\psi \in \mathcal{K}$. The constraint gives:

$$\|\psi\|_{L^2}^2 = \int \rho_\psi \leq N.$$

The fact that it is bounded from below follows from Hardy's inequality.

First, by (7) we have $D(\rho_\psi) - X(\gamma_\psi) \geq 0$, and for all $1 \leq i \leq N$ and $\varepsilon > 0$, we have:

$$z_m \int \frac{|\psi_i(x)|^2}{|x - R_m|} \leq 2\varepsilon z_m \|\nabla\psi_i\|_{L^2}^2 + \frac{z_m}{2\varepsilon} \int |\psi_i(x)|^2 \mathrm{d}x.$$

Taking $\varepsilon = (4Z)^{-1}$ for instance yields:

$$\mathcal{E}(\psi) \geq 1/2\|\nabla\psi\|_{L^2} - 2Z^2\|\psi\|_{L^2}^2.$$

For $\psi \in \mathcal{K}$, we thus get the (crude) estimate:

$$\|\psi\|_{H^1}^2 \leq 2\mathcal{E}(\psi) + (4Z^2 + 1)N.$$

5.4. **Proof of Thm 10.** The usual proof of the theorem of Rellich-Kondrachov uses the $L^p(\mathbb{R}^d)$-version of Ascoli's theorem.

Here as we deal with $L^2$-functions, we will use an "elementary" tools to prove the theorem, namely the Fourier transform on a torus $\mathbb{R}^d/(L\mathbb{Z}^d)$ and the diagonal extraction.

Let $(\psi_n)_n$ be a sequence of functions in $H^1(\mathbb{R}^d)$ which is $H^1$-bounded. By banach-Alaoglu, up to an extraction, we can assume that it converges $H^1$-weakly to $\psi \in H^1(\mathbb{R}^d)$. Furthermore, by the unifom boundedness principle, we have:

$$\sup_n \|\psi_n\|_{H^1} = C_0 < +\infty.$$

Let us show that there exists a subsequence of $(\psi_n)$ which converges in $L^2_{\mathrm{loc}}(\mathbb{R}^d)$.

Let $\chi \in C_0^\infty(\mathbb{R}^d, [0,1])$ such that supp $\chi \subset \{x : |x| \leq 2\}$ and $\chi(x) = 1$ for $|x| \leq 1$. For $R \geq 1$, we introduce the function $\chi_R(x) := \chi(x/R)$ and $T_{2R}$ the hypercube $[-2R, 2R]^d$.

*Aim:* For every $L \in \mathbb{N}$, we show that the sequence $(\chi_L\psi_n)_n$ converges in $L^2(T_{2L})$ up to the extraction of a subsequence.

By a diagonal extraction argument, this will provide us with a subsequence $(\psi_{n_k})_k$ such that for all $n$, the sequence $(\chi_L\psi_{n_k})_k$ converges in $L^2(T_{2L})$, which will end the proof.

The key idea is that we can inject *isometrically* the Hilbert space $H_0^1(T_{2L})$ into the Hilbert space $H^1_{\mathrm{per}}(T_{2L})$ of periodic functions in $L^2((\mathbb{R}^d/(4L\mathbb{Z}^d)))$ with an $H^1$-regularity. Thus we can see $(\chi_L\psi_n)_n$ as an $H^1$-bounded sequence in $H^1_{\mathrm{per}}(T_{2L})$. It is bounded because $[\nabla(\chi_L\psi_n)](x) = L^{-1}\nabla\chi(x/L)\psi_n + \chi_L\nabla\psi_n$. The same argument gives also that the $\chi_L\psi_n$'s are uniformly $H^1$-bounded in $n$ and $L \geq 1$, say by $C_1^2$.

We write $T_{2L}^* := \frac{\pi}{2L}\mathbb{Z}^d$ the dual of $\mathbb{R}^d/(2L\mathbb{Z}^d)$. The Fourier transform defines an isometry of $L^2_{\mathrm{per}}(T_{2L}) = L^2((\mathbb{R}^d/(4L\mathbb{Z}^d)))$ onto $\ell^2(T_{2L}^*)$. The $\ell^2$-element of an $L^2$-function is the collection of its Fourier modes:

$$a_{\underline{k}}(f) := (4L)^{-d/2} \int_{T_{2L}} f(x) e^{-i\underline{k}\cdot x} \mathrm{d}x, \ \underline{k} \in T_{2L}^*.$$

We thus have:

$$\|f\|_{L^2(T_{2L})}^2 = \sum_{\underline{k} \in T_{2L}^*} |a_{\underline{k}}(f)|^2.$$

Similarly, we know that[16]

$$\|\nabla f\|_{L^2(T_{2L})}^2 = \sum_{\underline{k}\in T_{2L}^*} |\underline{k}|^2 |a_{\underline{k}}(f)|^2.$$

Here, we consider the Fourier modes of the functions $\chi_L\psi_n$, giving a family $(a_{\underline{k},n})_{\underline{k},n}$. Fixing $\underline{k}$, the sequence $(a_{\underline{k},n})_{n\geq 0}$ is bounded in $\mathbb{C}$ as we have:

$$|a_{\underline{k},n}|^2(1+|\underline{k}|^2) \leq \|\chi_L\psi_n\|_{H^1}^2 \leq C_1^2.$$

Thus for every $\underline{k}$, we can extract a converging subsequence $(a_{\underline{k},n_\ell})_{\ell\geq 0}$.

By a *diagonal extraction*, we can extract a subsequence $(\chi_L\psi_{n_\ell})_\ell$ such that for all $\underline{k} \in T_{2L}^*$, the sequence $(a_{\underline{k},n_\ell})_{\ell\geq 0}$ is convergent (say to $a_{\underline{k}} \in \mathbb{C}$). By the uniform boundedness principle we also have:

$$\sum_{\underline{k}} (1+|\underline{k}|^2)|a_{\underline{k}}|^2 < +\infty.$$

Note that the $a_{\underline{k}}$'s are the Fourier modes of $\chi_L\psi$.

We show that the sequence $\big((a_{\underline{k},n_\ell})_{\underline{k}\in T_{2L}^*}\big)_{\ell\geq 0}$ is convergent in $\ell^2(T_{2L}^*)$, which will end the proof.

For $A \geq 1$, we have:

$$\begin{aligned}\sum_{\underline{k}} |a_{\underline{k},n_\ell} - a_{\underline{k}}|^2 &= \sum_{\substack{\underline{k},\\ |\underline{k}|\leq A}} |a_{\underline{k},n_\ell} - a_{\underline{k}}|^2 + \sum_{\substack{\underline{k},\\ |\underline{k}|> A}} |a_{\underline{k},n_\ell} - a_{\underline{k}}|^2,\\ &= \sum_{\substack{\underline{k},\\ |\underline{k}|\leq A}} |a_{\underline{k},n_\ell} - a_{\underline{k}}|^2 + \sum_{\substack{\underline{k},\\ |\underline{k}|> A}} \frac{1+|\underline{k}|^2}{1+|\underline{k}|^2}|a_{\underline{k},n_\ell} - a_{\underline{k}}|^2,\\ &\leq \sum_{\substack{\underline{k},\\ |\underline{k}|\leq A}} |a_{\underline{k},n_\ell} - a_{\underline{k}}|^2 + \frac{2}{1+A^2}(\|\chi_L\psi_n\|_{H^1}^2 + \|\chi_L\psi\|_{H^1}^2).\end{aligned}$$

taking the limsup yields:

$$\limsup_{\ell\to+\infty} \sum_{\underline{k}} |a_{\underline{k},n_\ell} - a_{\underline{k}}|^2 \leq \frac{4C_1^2}{1+A^2}.$$

As this holds for every $A > 0$ we obtain the convergence in $\ell^2(T_{2L}^*)$, and thus by isometry the convergence in $L^2(T_{2L})$.

## References

[1] Lawrence C. Evans, *Partial differential equations*, 2nd ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR2597943

[2] Elliott H. Lieb and Barry Simon, *The Hartree-Fock theory for Coulomb systems*, Comm. Math. Phys. **53** (1977), no. 3, 185–194. MR0452286

[3] P.-L. Lions, *Solutions of Hartree-Fock equations for Coulomb systems*, Comm. Math. Phys. **109** (1987), no. 1, 33–97. MR879032

[4] Michael Reed and Barry Simon, *Methods of modern mathematical physics. I. Functional analysis*, Academic Press, New York-London, 1972.

[16]or we do not know but we can compute and show that