Modèle de Hartree-Fock-Bogoliubov: une perspective théorique et numérique

Séverine Paul

Université de Cergy-Pontoise

Modèle de Hartree-Fock-Bogoliubov : une perspective théorique et numérique

Thèse de doctorat en Mathématiques

présentée par

Séverine PAUL

JURY

Directeur :	**Mathieu LEWIN** Chargé de recherche CNRS, Université de Cergy-Pontoise
Rapporteurs :	**Éric CANCÈS** Professeur, École Nationale des Ponts et Chaussées
	Christian HAINZL Professeur, Tübingen Universität (Allemagne)
Examinateurs :	**Laurent BRUNEAU** Maître de Conférences, Université de Cergy-Pontoise
	Bernard DUCOMET Ingénieur-chercheur, Commissariat à l'Énergie Atomique
	Éric SÉRÉ Professeur, Université Paris-Dauphine

le 30 octobre 2012

À mon grand-père

Modèle de Hartree-Fock-Bogoliubov : une perspective théorique et numérique.

Résumé : Cette thèse est consacrée à l'étude mathématique et numérique du modèle de Hartree-Fock-Bogoliubov (HFB) pour les systèmes quantiques attractifs, qui est abondamment utilisé en physique nucléaire. Après avoir présenté le modèle et ses principales caractéristiques, nous expliquons comment le discrétiser et nous montrons des résultats de convergence. Nous examinons tout particulièrement l'algorithme de point fixe (parfois appelé Roothaan) et montrons qu'il converge ou alors oscille entre deux états dont aucun n'est solution du problème. Ceci généralise au cadre HFB des résultats de Cancès et Le Bris pour le modèle plus simple de Hartree-Fock dans le cas répulsif. Suivant ces mêmes auteurs, nous proposons un algorithme basé sur la contrainte relâchée et pour lequel la convergence est garantie. Dans la dernière partie de la thèse, nous illustrons le comportement de ces algorithmes par des simulations numériques pour plusieurs modèles. Dans un premier temps nous considérons un système purement gravitationnel où les particules interagissent avec le potentiel de Newton. Nos simulations montrent que la matrice d'appariement est toujours non nulle, un fait qui n'a pas encore pu être démontré rigoureusement. Nous étudions ensuite un modèle très simplifié pour la description de protons et neutrons dans le noyau atomique.

Mots clés : Hartree-Fock-Bogoliubov, appariement, analyse numérique, physique quantique.

Hartree-Fock-Bogoliubov Theory: a Theoretical and Numerical Perspective.

Abstract: This work is devoted to the theoretical and numerical study of Hartree-Fock-Bogoliubov (HFB) theory for attractive quantum systems, which is one of the main methods in nuclear physics. We first present the model and its main properties, and then explain how to discretize it. We prove some convergence results, in particular for the simple fixed point algorithm (sometimes called Roothaan). We show that it converges, or oscillates between two states, none of them being a solution. This generalizes to the HFB case previous results of Cancès and Le Bris for the simpler Hartree-Fock model in the repulsive case. Following these authors, we also propose a relaxed constraint algorithm for which convergence is guaranteed. In the last part of the thesis, we illustrate the behavior of these algorithms by some numerical experiments. We first consider a system where the particles only interact through the Newton potential. Our numerical results show that the pairing matrix never vanishes, a fact that has not yet been proved rigorously. We then study a very simplified model for protons and neutrons in a nucleus.

Keywords: Hartree-Fock-Bogoliubov, pairing, numerical analysis, quantum phy--sics.

Université de Cergy-Pontoise, Laboratoire de Mathématiques UMR CNRS 8088
2, avenue Adolphe Chauvin, 95 302 CERGY-PONTOISE Cedex, FRANCE

Contents

Remerciements

Mes tout premiers remerciements vont à mon directeur de thèse Mathieu. Merci avant tout pour ta patience ! Qualité dont tu as fait preuve au quotidien avec moi pendant ces quatre années, face à mes questions parfois idiotes, parfois les mêmes d'une semaine à l'autre, face aux travaux inachevés, mes hésitations et mes silences traduisant l'incompréhension. Merci pour ta disponibilité que ce soit par ta présence au labo, par mail ou par skype. Merci aussi pour tes encouragements et ton indulgence. Je mesure la chance que j'ai eu de t'avoir comme directeur. Tu es pour moi l'exemple du chercheur passionné et passionnant et je regrette parfois de n'avoir pu profiter plus de ton savoir.

Je remercie très chaleureusement mes rapporteurs Eric Cancès et Christian Hainzl pour avoir accepté de consacrer une partie de leur temps à la lecture et à l'évaluation de mes travaux. Je remercie aussi cordialement Laurent Bruneau, Bernard Ducomet et Eric Séré d'avoir accepté d'être membre de mon jury de thèse.

Un grand merci au laboratoire AGM dans lequel j'ai passé ces quatre années de thèse et à son directeur Vladimir qui a bien voulu être mon directeur officiel pendant ma première année. Merci aussi à tous leurs membres et employés pour l'ambiance de travail chaleureuse qui règne dans ces lieux. Les ANR "NoNAP" et ERC "MNIDS" m'ont permis de visiter les mathématiques et des mathématiciens à différents endroits en Europe et je tiens à en remercier leurs éminents porteurs et tous leurs membres. Merci beaucoup à Christian Hainzl qui m'a permis d'assister à la conférence, très demandée à Vienne en août 2012.

Merci à Julien, Salma, Virginie, Simona, Antoine, Gaspard, Loïc, Jéremy et Mauricio, membres du groupe de travail des thésards "Méthodes variationnelles en Physique Quantique" pour l'ambiance joviale autour de discussions mathématiques.

Ma gratitude s'adresse également à Abil, Azba, Hong Cam, Sixiang, Andrey, Nicolas G., David, Anne-Sophie, Sébastien, Lysianne, Amal, Constanza, Xavier, Christophe, Salahaddine, Davit, Hayk, Jérome, Giona, Nam, Nicolas R. et Rafik, thésards, ex-thésards et post-doc du labo pour les pauses café, les séminaires et les longues discussions soit footbalistiques soit cinématographiques pendant les déjeuners au CROUS. Merci aux maîtres de conf des bureaux voisins, François,

Thierry, Gerasim et maintenant Kenny pour les brefs moments de repos partagés entre deux cours. Merci Patrick pour approvisionner toujours à temps le labo de chocolat.

Un grand merci aux secrétaires Marie, Linda et à Josette, responsable du personnel pour les conversations chaleureuses, les paperasses administratives et leur présence à tout moment face aux soucis de badge, photocopieuse ou imprimante. Merci à Bernard Charles pour son dévouement. Je ne compte pas le nombre de fois où j'ai pu l'embêter pour des problèmes informatiques qui ne concernaient même pas mon travail.

Merci aussi à Sepideh, Rachida, Tina, Mathieu et Giacomo les thésards du bureau 3D18 du Laboratoire Jacques-Louis Lions à chevaleret pour m'avoir accueillie les jours de flemne où Cergy me semblait trop loin.

Merci aux étudiants rencontrés pendant ces quatre années qui m'ont fait apprécier l'enseignement.

Merci de tout coeur à ma maman, mon père et mes trois adorables petites soeurs, toujours une question de math en tête. Merci à mes oncles et tantes. Merci aussi à mes grands-parents. Merci grand-père, pour ta foi en la science, ta curiosité, ton amour des mathématiques et ton émerveillement. Merci pour vos encouragements et votre soutien. Merci à mes amis toujours présents Guylaine, Caroline, Nouri, Alexia, Pakito et Christophe qui m'ont largement soutenue quand le découragement me guettait.

Mes derniers remerciements sont pour Olivier qui a réussit l'expoit de supporter mes jérémiades au quotidien et m'avoir encouragée pendant la rédaction de cette thèse. Merci aussi pour tes bons conseils de programmeur, pour les longues discussions sur nos différents sujets. Merci à mon petit Léon; à tes longues siestes dans le porte-bébé qui m'ont permis de préparer cette présentation, et à ton sourire, tes bavouilles, tes grosses joues et tes Areuh !

Introduction

Cette thèse présente divers résultats théoriques et numériques concernant le modèle de Hartree-Fock-Bogoliubov qui permet la description de systèmes fermioniques attractifs. Il s'agit d'un modèle non linéaire dans lequel l'état du système est décrit par deux inconnues, qui sont des opérateurs compacts sur $L^2(\mathbb{R}^3, \mathbb{C}^2)$:

- γ, la matrice densité à un corps
- α, la matrice d'appariement qui est nulle pour les systèmes répulsifs.

Un problème important est de savoir si la matrice d'appariement est non nulle pour les systèmes attractifs, pour l'état fondamental, lorsque le nombre moyen de particules est fixé. La non-nullité de cette matrice est supposée décrire des phénomènes physiques comme l'apparition des paires de Cooper, que l'on retrouve en physique de la matière condensée et en physique nucléaire, et qui est reliée à la supraconductivité et à la superfluidité. Dans le cas où la matrice d'appariement serait nulle, on retrouve le modèle de Hartree-Fock utilisé dans de nombreuses applications mais qui ne permet pas la description de ces phénomènes.

Nous décrivons ci-après les résultats obtenus dans les divers chapitres de la thèse.

Chapitre 1 : le modèle Hartree-Fock-Bogoliubov

Dans le premier chapitre, nous décrivons le modèle à N-corps non relativiste avec spin. La matière est composée de plusieurs types de particules, qui sont toutes des fermions qui interagissent entre eux à travers différentes forces. Ce sont les bosons qui servent de particules médiatrices décrivant les échanges d'énergie entre les fermions de la matière. Les fermions sont de spin demi-entier, le spin étant une propriété quantique intrinsèque associée à chaque particule. Les seuls connus jusqu'à présent sont de spin $1/2$.

Un système quantique composé de N particules fermioniques dans l'espace $\mathbb{R}^3$ est décrit par une fonction d'onde Ψ à N-corps telle que

$$\Psi(x_1, \sigma_1, ..., x_N, \sigma_N) \in \bigwedge_1^N L^2\big(\mathbb{R}^3 \times \{\pm 1/2\}; \mathbb{C}\big) := \mathfrak{H}_N.$$

Les quantités $\sigma_k, \forall k = 1, ..., N$ représentent le spin de chacune des particules, celui-ci ne pouvant prendre que deux valeurs : $-1/2$ et $1/2$. Par le principe de Pauli,

il ne peut y avoir que 2 fermions de spins différents dans un même état, ce qui se traduit ici mathématiquement par le caractère anti-symétrique de la fonction Ψ. On considère donc N fermions interagissant à travers un potentiel W et sans potentiel exterieur. Il est décrit par le hamiltonien :

$$H(N) = \sum_{i=1}^{N}(-\Delta)_{x_i} + \sum_{1\leq k<\ell\leq N} W(x_k - x_\ell),$$

où $(-\Delta)_{x_i}, \forall i = 1, ..., N$ est l'énergie cinétique de chaque particule et où on suppose W suffisamment régulière et s'annulant à l'infini pour que $H(N)$ soit bien défini.

Notre problème va être la détermination de l'état fondamental, c'est-à-dire de l'état de plus basse énergie du système. Dans le cas du modèle à N-corps ci-dessus, l'énergie est bornée inférieurement, mais ne possède pas de minimiseur. Ceci est dû à l'invariance par translation du hamiltonien $H(N)$ du système. Nous allons considérer une approximation non linéaire du modèle à N-corps, pour laquelle il peut y avoir des minimiseurs à cause des effets non linéaires. Cette approximation consiste à restreindre le calcul de l'énergie aux fonctions qui s'écrivent sous la forme d'un déterminant de Slater, autrement dit d'un produit antisymétrique de fonctions d'onde à une particule :

$$\begin{aligned}\Psi(x_1, \sigma_1, ..., x_N, \sigma_N) =& \varphi_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_N(x_1, \sigma_1, ..., x_N, \sigma_N) \\ :=& \frac{1}{\sqrt{N!}} \det\big(\varphi_i(x_j, \sigma_j)\big),\end{aligned} \tag{0.0.1}$$

où φ_i sont les orbitales de chacune des particules, qui vérifient par hypothèse la contrainte de normalisation

$$\langle \varphi_i, \varphi_j \rangle_{L^2(\mathbb{R}^3\times\{\pm 1/2\};\mathbb{C})} = \delta_{ij}.$$

Nous obtenons le modèle de Hartree-Fock (HF) qui permet le calcul de l'état fondamental approché de le hamiltonien à N-corps. L'énergie du système s'écrit alors

$$\begin{aligned}\langle \Psi, H(N)\Psi \rangle =& \mathcal{E}_{\mathrm{HF}}(\varphi_1, ..., \varphi_N) \\ =& \sum_{i=1}^{N} \int_{\mathbb{R}^3} |\nabla \varphi_i|^2 + \frac{1}{2} \iint_{\mathbb{R}^3\times\mathbb{R}^3} \rho_\Psi(x)\rho_\Psi(y)W(x-y)dxdy \\ & - \frac{1}{2} \iint_{\mathbb{R}^3\times\mathbb{R}^3} |\gamma_\Psi(x,y)|^2 W(x-y)dxdy.\end{aligned} \tag{0.0.2}$$

Ici γ_Ψ est la matrice densité à 1-corps du système définie par

$$\gamma_\Psi = \sum_{i=1}^{N} |\varphi_i\rangle\langle\varphi_i|, \tag{0.0.3}$$

qui vérifie les contraintes $0 \leqslant \gamma_\Psi \leqslant 1$ et $\mathrm{Tr}(\gamma_\Psi) = N$ et dont le noyau intégral est donc

$$\gamma_\Psi(x,y)_{\sigma,\sigma'} = \sum_{i=1}^{N} \varphi_i(x,\sigma)\overline{\varphi_i(y,\sigma')}.$$

D'un autre côté,

$$\begin{aligned}\rho_\Psi(x) =& N \sum_{\sigma_1,\ldots,\sigma_N=\pm 1/2} \int_{\mathbb{R}^3} dx_2 \cdots \int_{\mathbb{R}^3} dx_N |\Psi(x,\sigma_1,x_2,\sigma_2,...,x_N,\sigma_N)|^2 \\ =& \sum_{\sigma=\pm 1/2} \gamma_\Psi(x,x)_{\sigma\sigma} \\ =& \sum_{i=1}^{N} \sum_{\sigma=\pm 1/2} |\varphi_i(x,\sigma)|^2\end{aligned}$$

est la densité de charge associée. Grâce à la nonlinéarité du modèle HF, cette énergie peut admettre un minimiseur, même si le système est invariant par translation.

Le modèle HF décrit un système fermé de N particules. Dans la suite de notre étude, nous introduisons une brisure de symétrie. Le nombre de particules du système N est alors remplacé par un opérateur $\mathcal{N}$ et seul le nombre moyen de particules est fixé. Plus précisément, il est essentiel d'introduire un nouvel espace, appelé *espace de Fock*

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathfrak{H}_N,$$

qui est la somme directe des espaces de Hilbert à N particules avec $N \geq 0$, et dans lequel nous recherchons l'état qui minimise l'énergie du système. L'opérateur de nombre est alors défini par

$$\mathcal{N} = \bigoplus_{N\geq 0} N$$

et ses valeurs propres sont $0, 1, 2...$ De façon similaire, le hamiltonien à N-corps est remplacé par sa seconde quantification

$$\mathbb{H} := 0 \oplus \bigoplus_{n\geq 1} H(n)$$

sur l'espace de Fock $\mathcal{F}$.

Nous n'allons pas minimiser le hamiltonien $\mathbb{H}$ sur tout l'espace de Fock, mais plutôt sur une classe naturelle d'états qui généralisent les états HF de la formule (0.0.1) et qui sont appelés *états quasi-libres* ou *Hartree-Fock-Bogoliubov*. Ces états de $\mathcal{F}$ peuvent être définis comme l'ensemble de tous les états de Gibbs associés à tous les hamiltoniens qui sont quadratiques en les opérateurs de création et d'annihilation. La définition précise est un peu longue à donner dans cette introduction et nous renvoyons pour cela au Chapitre 1. On pourra simplement

retenir que, pour les déterminer, on utilise le formalisme des matrices densités. Les états sont entièrement caractérisés par deux matrices densité γ et α qu'on regroupe généralement dans une grande matrice Γ :

$$\Gamma = \begin{pmatrix} \gamma & \alpha \\ \alpha^* & 1-\overline{\gamma} \end{pmatrix}. \tag{0.0.4}$$

Ici γ est la même matrice densité à 1-corps que précédement, et α est la *matrice d'appariement*. Ces deux matrices sont définies à l'aide des opérateurs de création et d'annihilation par les relations

$$\langle g, \gamma f\rangle = \langle a^*(f)a(g)\rangle \qquad \langle g, \alpha\overline{f}\rangle = \langle a(f)a(g)\rangle.$$

Lorsque $\alpha \neq 0$, l'état correspondant dans l'espace de Fock $\mathcal{F}$ n'est pas un état propre de l'opérateur $\mathcal{N}$ et la symétrie du nombre de particules est brisée. L'opérateur Γ doit vérifier les contraintes

$$0 \leq \Gamma \leq 1 \qquad \text{et} \qquad \alpha^T = -\alpha.$$

Le modèle de Hartree-Fock-Bogoliubov est utilisé pour l'étude de systèmes attractifs à petit ou grand nombre de particules. Il consiste à restreindre le hamiltonien quantifié $\mathbb{H}$ sur l'espace de Fock aux états quasi-libres. L'énergie HFB en fonction des opérateurs γ et α, et décrivant un système fermionique non-relativiste soumis à l'interaction W s'écrit alors

$$\begin{aligned}&\mathcal{E}_{\mathrm{HFB}}(\gamma,\alpha)\\ &\quad= \mathrm{Tr}(-\Delta)\gamma + \frac{1}{2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\Big(\rho_\gamma(x)\rho_\gamma(y) - |\gamma(x,y)|^2 + |\alpha(x,y)|^2\Big)\,dx\,dy.\end{aligned}$$

Cette énergie a une forme très similaire à l'énergie HF vue en (0.0.2), avec laquelle elle coïncide quand $\alpha \equiv 0$ et γ est un projecteur de rang fini. On remarque l'apparition d'un quatrième terme qui représente l'appariement du système.

Après avoir montré que cette énergie était bien définie et bornée inférieurement, dans le chapitre 1 nous définissons deux problèmes de minimisation, $I(N)$ sur l'ensemble des états mixtes ($0 \leq \Gamma \leq 1$) et $J(N)$ sur l'ensemble des états purs ($\Gamma = \Gamma^2$). Le principe variationnel de Lieb dit que si $W > 0$ est repulsif, l'état fondamental sera toujours un état pur HF sans appariement $\alpha = 0$. Ce résultat est très important et a été exploité par É. Cancès et C. Le Bris pour développer des stratégies numériques que nous verrons dans le chapitre 3. Dans le cas purement attractif $W < 0$, un résultat dû à Bach, Fröhlich et Jonsson nous dit que pour des fermions de spin 1/2 et une attraction fortement attractive, l'état HFB est à nouveau pur et les matrices densité et d'appariement sont sous une forme spéciale (cf 1.4.30) et $I(N) = J(N)$. Ce résultat suggère qu'il est possible d'adapter les méthodes numériques utilisées en chimie quantique au cas des systèmes attractifs, ce qui est l'objectif principal de cette thèse.

Par la suite nous énonçons un résultat de E. Lenzmann et M. Lewin qui fournit l'existence de minimiseurs pour le problème $I(N)$ (en fait, leur résultat est valable dans le cas relativiste mais la preuve s'adapte sans problème au cas non relativiste). Ces minimiseurs Γ sont solutions d'une équation non linéaire sous la forme

$$\Gamma = \mathbb{1}_{(-\infty,0)}\big(F_\Gamma - \mu\mathcal{N}\big) + \delta \tag{0.0.5}$$

où $0 \leq \delta \leq \mathbb{1}_{\{0\}}(F_\Gamma - \mu\mathcal{N})$ prend la même forme que Γ,

$$\mathcal{N} := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

et F_Γ est la matrice de Fock (ou matrice de champ moyen) définie par

$$F_\Gamma = \begin{pmatrix} h_\gamma & \pi \\ \pi^* & -\overline{h_\gamma} \end{pmatrix}. \tag{0.0.6}$$

Ici

$$h_\gamma = -\Delta + \rho_\gamma * W - W(x-y)\gamma(x,y)$$

et

$$\pi(x,y) = \alpha(x,y)W(x-y).$$

Dans le cas où $\alpha = 0$ on retombe sur l'équation des minimiseurs HF

$$\gamma = \mathbb{1}_{(-\infty,\mu)}(h_\gamma) + \delta$$

ou, exprimé en fonction des orbitales φ_i,

$$h_\gamma \varphi_i = \lambda_i \, \varphi_i, \qquad i = 1, ..., N$$

quand $\delta \equiv 0$. Dans le problème HFB, la matrice de champ moyen F_Γ est non bornée inférieurement, contrairement à h_γ (son spectre est en fait symétrique par rapport à 0). De plus, $\mathcal{N}$ ne commute pas avec F_Γ, sauf quand $\alpha \equiv 0$, et donc les équations HFB ne peuvent être écrites aussi facilement que pour HF. Le fait que γ ne soit plus un projecteur de rang N, comme pour le problème HF ($\alpha \equiv 0$), mais soit un opérateur de rang infini quand l'appariement $\alpha \neq 0$ change radicalement le comportement du problème.

Dans cette thèse, nous nous intéressons particulièrement au problème de savoir si l'appariement est non nul pour les minimiseurs HFB, dans le cas où W est suffisament attractif. Nous répondrons à cette question par des simulations numériques.

Chapitre 2 : le modèle Hartree-Fock-Bogoliubov discrétisé

Dans le chapitre 2, nous expliquons comment discrétiser le problème de minimisation HFB afin de le résoudre numériquement. La méthode est d'utiliser une approximation de Galerkin. On introduit donc un espace discret et on montre que la solution du problème sur cet espace converge vers la solution du problème en

dimension infinie (Théorème 2.1.1 ci-dessous). On peut donc s'attarder sur l'étude du problème en dimension finie.

On choisit une base $(\chi_i)_{i=1}^{N_b}$ et, après avoir écrit les noyaux de γ et α dans cette base, on trouve pour l'énergie HFB

$$\mathcal{E}(G,A) = \mathrm{Tr}(hG) + \frac{1}{2}\,\mathrm{Tr}(G\,J(G)) - \frac{1}{2}\,\mathrm{Tr}(G\,K(G)) + \frac{1}{2}\,\mathrm{Tr}(A^*\,X(A)),$$

où G et A sont les matrices densité et d'appariement discrétisées, vérifiant $G^* = G$ et $A^T = -A$. Par ailleurs, J, K et X sont des opérateurs linéaires sur l'espace des matrices $N_b \times N_b$. La contrainte s'écrit

$$0 \leq \Upsilon \mathbf{S} \Upsilon \leq \Upsilon$$

où Υ est défini par

$$\Upsilon := \begin{pmatrix} G & A \\ A^* & \overline{\Sigma}^{-1} S \Sigma^{-1} - \overline{G} \end{pmatrix},$$

et $\mathbf{S}$, S et Σ sont des matrices de recouvrement définies en (2.2.7) et (2.2.8) qui dépendent des fonctions de bases choisies.

Le problème HFB présente certaines symétries que l'on choisit d'inclure dans les contraintes pour réduire le coût de calcul. La première est la symétrie de spin car le laplacien et la fonction d'interaction W n'agissent en aucun cas sur la variable de spin. La deuxième est due au fait que le laplacien et l'opérateur W sont réels. On l'appelle symétrie de la conjugaison complexe. Si on choisit d'imposer ces symétries, alors le problème est de trouver (après élimination du spin) G et A, deux matrices réelles symétriques, telles que

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \leq \Upsilon := \begin{pmatrix} G & A \\ A & 1-G \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

et l'énergie s'écrit

$$\mathcal{E}(G,A) = 2\,\mathrm{Tr}(hG) + 2\,\mathrm{Tr}(G\,J(G)) - \mathrm{Tr}(G\,K(G)) + \mathrm{Tr}(A\,X(A)).$$

Le nombre de particule est réduit par 2, c'est-à-dire $\mathrm{Tr}(SG) = N/2$.

La dernière symétrie est celle du groupe des rotations généré par l'opérateur moment angulaire $L = x \times (-i\nabla)$. Si W est une fonction radiale, alors les états du système sont invariants par rotation, et s'écrivent en fonction du moment angulaire total ℓ tel que $0 \leq \ell \leq \ell_{\max}$ et du moment angulaire azimutal m tel que $-\ell \leq m \leq \ell$, ainsi que des harmoniques sphériques Y_ℓ^m. L'énergie discrétisée et la contrainte s'écrivent aussi en fonction du moment angulaire total ℓ, et nous obtenons $\ell_{\max} + 1$ équations non linéaires couplées, partageant le même multiplicateur d'Euler-Lagrange μ.

Chapitre 3 : algorithmes et analyse de leur convergence

Dans le chapitre 3, nous étudions la convergence des algorithmes de Roothaan et ODA, utilisés pour résoudre le problème de minimisation de l'énergie de Hartree-Fock-Bogoliubov. Ils ont largement été étudiés par É. Cancès et C. Le Bris pour la résolution des équations de Hartree-Fock en chimie quantique, mais pas pour les systèmes attractifs comme dans cette thèse.

L'algorithme de Roothaan est un simple algorithme de point fixe sur l'équation non linéaire (0.0.5), qui n'assure *a priori* pas la décroissance de l'énergie. Nous donnons tout d'abord une définition du caractère bien posé de cet algorithme, et démontrons ensuite un théorème concernant ses problèmes de convergence et dont la preuve repose sur de récents travaux d'A. Levitt. En effet il se peut que l'algorithme oscille entre deux états dont aucun des deux n'est solution du problème de minimisation HFB. Pour comprendre ce phénomène, nous introduisons une fonctionnelle d'énergie à deux variables

$$\tilde{\mathcal{E}}(\Upsilon, \Upsilon') := \mathrm{Tr}(hG) + \mathrm{Tr}(hG') + 2\,\mathrm{Tr}(G\,J(G')) - \mathrm{Tr}(G\,K(G')) + \mathrm{Tr}(A\,K(A')),$$

en suivant É. Cancès et C. Le Bris. Il se trouve qu'appliquer l'algorithme de point fixe est équivalent à minimiser la fonctionnelle $\tilde{\mathcal{E}}$ alternativement par rapport à ses deux variables. Nous avons alors démontré le théorème suivant :

Théorème (Convergence de l'algorithme de Roothaan). *Soit Υ_0 un état initial HFB tel que la suite (Υ_n) générée par l'algorithme de Roothaan est* uniformement bien posée*, c'est-à-dire*

$$\forall n, \qquad \left|\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\right| \geq \eta > 0,$$

où $\mathbf{F}_{\tilde{\Upsilon}_n}$ est la version discrétisée de l'opérateur de champ moyen (0.0.6)*. Alors*

- *la suite $\tilde{\mathcal{E}}(\Upsilon_{2n}, \Upsilon_{2n+1})$ décroit vers une valeur critique $\tilde{\mathcal{E}}$;*
- *la suite $(\Upsilon_{2n}, \Upsilon_{2n+1})$ converge vers un point critique (Υ, Υ') de $\tilde{\mathcal{E}}$;*
- *si $\Upsilon = \Upsilon'$, alors cet état est une solution des équations HFB, mais si $\Upsilon \neq \Upsilon'$, alors aucun de ces deux états n'est solution.*

Par la suite, pour contourner les difficultés de convergence de l'algorithme de Roothaan, nous introduisons le *Optimal Damping Algorithm* (ODA), qui a été proposé en premier par É. Cancès et C. Le Bris pour les modèles de chimie quantique. Il consiste à relâcher la contrainte $(\Upsilon_n)^2 = \Upsilon_n$ en $0 \leq (\Upsilon_n)^2 \leq \Upsilon_n$. Cela signifie qu'on ne cherche plus à minimiser l'énergie sur l'ensemble des états purs, mais sur les états mixtes du système. L'algorithme ODA engendre donc deux suites Υ_n pour les états purs et $\tilde{\Upsilon}_n$ pour les états mixtes. Le théorème qui suit prouve sa convergence :

Théorème (Convergence d'ODA). *Soit $\Upsilon_0 = \tilde{\Upsilon}_0$ un état initial HFB tel que l'algorithme ODA soit* uniformément bien posé*, c'est-à-dire*

$$\forall n, \qquad \left|\mathbf{F}_{\tilde{\Upsilon}_n} - \mu_{n+1}\mathbf{N}\right| \geq \eta > 0,$$

où $\mathbf{F}_{\tilde{\Upsilon}_n}$ *est la version discrétisée de l'opérateur de champ moyen* (0.0.6). *Alors*

- *la suite* $\mathcal{E}(\tilde{\Upsilon}_n)$ *décroît vers une valeur critique de* $\mathcal{E}$*;*
- *la suite* Υ_n *converge* numériquement *vers un point critique* Υ *de* $\mathcal{E}$*, au sens où* $\Upsilon_{n+1} - \Upsilon_n \to 0$, $\Upsilon_{n+1} - \tilde{\Upsilon}_n \to 0$. *Par ailleurs tous les points d'accumulation* Υ *de la suite* (Υ_n) *vérifient l'équation HFB* $\Upsilon = \mathbb{1}_{(-\infty,0)}(\mathbf{F}_\Upsilon - \mu\mathbf{N})$.

Chapitres 4–5 : simulations numériques

Dans les chapitres 4 et 5, nous appliquons les deux algorithmes étudiés précédemment à deux systèmes différents pour résoudre numériquement le problème HFB.

Le premier, étudié au chapitre 4, est un modèle gravitationnel qui peut être utilisé pour décrire les étoiles à neutrons ou les naines blanches quand $N >> 1$ (mais dans nos simulations, N n'est pas très grand). Nous considérons donc un système de N particules non relativistes de spin 1/2 interagissant à travers le potentiel Newtonien

$$W(x) = -\frac{g}{|x|}, \qquad g > 0,$$

qui est très attractif à courte distance et qui décroit suffisamment lentement pour que deux particules très éloignées s'attirent toujours.

Le second est un modèle inspiré de ceux utilisés en physique nucléaire. Il modélise l'interaction entre nucléons qui est répulsive à courte distance et attractive à moyenne distance. Le potentiel qui décrit cette interaction est

$$W(x) = \frac{\kappa}{|x|} - a_1\, e^{-b_1|x|^2} + a_2\, e^{-b_2|x|^2}, \qquad \text{avec } a_2, a_1 > 0,\ b_1 < b_2.$$

La constante κ vaut 1 pour les protons et 0 pour les neutrons.

Pour simuler ces deux systèmes physiques, nous utilisons le logiciel libre Scilab. Tout d'abord nous comparons les valeurs de l'énergie obtenues pour les modèles HF et HFB. Les résultats obtenus montrent qu'il y a tout le temps de l'appariement (il est malgré tout nécessaire de bien choisir l'espace de discrétisation).

Aussi nous comparons le comportement des algorithmes de Roothaan et ODA. Dans certains cas, nous observons que l'algorithme de Roothaan converge très lentement par rapport à ODA, ou alors présente des oscillations entre deux états, dont aucun des deux n'est solution du problème. L'algorithme de Roothaan est connu et beaucoup utilisé par les physiciens. Pour contrer les problèmes qu'il présente, ils utilisent un paramètre de mélange t fixe et non optimisé comme dans ODA. Nous espérons donc que nos travaux permettront d'améliorer les algorithmes existants.

L'observation numérique de l'existence de l'appariement dans tous les cas pose évidemment une question théorique très intéressante. Nous espérons que les résultats de cette thèse stimuleront l'obtention de nouveaux théorèmes dans ce sens.

Part I

Theory

Chapter 1

Hartree-Fock-Bogoliubov model for attractive systems

In this chapter we quickly present the Hartree-Fock-Bogoliubov (HFB) model and we review some known mathematical results.

1.1 Linear N-body Schrödinger model

In this section, we introduce the many-body problem for N nonrelativistic particles interacting with a (local) potential W. We start by recalling the basics of quantum mechanics.

1.1.1 Many-body wavefunctions

The state of one quantum particle in $\mathbb{R}^3$

The quantum state of one particule in the 3 dimensional space and with q internal degrees of freedom is described by a wavefunction

$$\psi \in L^2(\mathbb{R}^3 \times \{1, ..., q\}; \mathbb{C}),$$

such that

$$\sum_{\sigma=1}^{q} \int_{\mathbb{R}^3} |\psi(x, \sigma)|^2 dx = 1.$$

The internal variable σ depends on the studied particle and it is a feature of it. For electrons, we have $q = 2$ and σ describes the spin which is an intrinsic quantum property. The value $|\psi(x, \sigma)|^2$ is the density of probability to find the particle at x in $\mathbb{R}^3$ with a spin $\sigma \in \{1, ..., q\}$. Similarly $|\widehat{\psi}(p, \sigma)|^2$ represents the density of probability to find the particle with a momentum p and a spin $\sigma \in \{1, ..., q\}$. Here and everywhere the Fourier transform is defined by

$$\hat{\psi}(p, \sigma) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} \psi(x, \sigma) e^{-ip.x} dx.$$

It is an isometry in $L^2(\mathbb{R}^3, \mathbb{C}^q)$:

$$\sum_{\sigma=1}^{q} \int_{\mathbb{R}^3} |\psi(x,\sigma)|^2 dx = \sum_{\sigma=1}^{q} \int_{\mathbb{R}^3} |\widehat{\psi}(p,\sigma)|^2 dp = 1.$$

The space $L^2(\mathbb{R}^3 \times \{1,...,q\}; \mathbb{C})$ is equipped with the following norm

$$\|\psi\|^2_{L^2(\mathbb{R}^3\times\{1,...,q\};\mathbb{C})} = \sum_{\sigma=1}^{q} \int_{\mathbb{R}^3} |\psi(x,\sigma)|^2 dx. \tag{1.1.1}$$

Note that we can equivalently see $\psi(x,\sigma)$ as a vector with q coordinates. In other words $L^2(\mathbb{R}^3 \times \{1,...,q\}, \mathbb{C})$ is isometric to $L^2(\mathbb{R}^3, \mathbb{C}^q)$. Using the convention that

$$|\psi(x)|^2 := \sum_{\sigma=1}^{q} |\psi(x,\sigma)|^2$$

is the usual norm in $\mathbb{C}^q$, we then get

$$\int_{\mathbb{R}^3} |\psi(x)|^2 \, dx = \sum_{\sigma=1}^{q} \int_{\mathbb{R}^3} |\psi(x,\sigma)|^2 dx.$$

Example 1.1.1 (Hydrogen atom)**.** *The hydrogen atom is the simplest atomic system and it is a good starting point for presenting the formalism of quantum mechanics. It is composed of a nucleus and just one electron. The nucleus is assumed to be fixed, classical and pointlike (Born-Oppenheimer approximation). This approximation can be justified by the fact that the nucleus has a mass which is much larger than that of the electron. The electron is submitted to the Coulomb potential*

$$V(|x|) = -\frac{e^2}{4\pi\epsilon_0|x|}$$

of the nucleus. Here e is the charge of the electron, m_e is its mass and ϵ_0 the permittivity of the vacuum. The classical energy of the electron located in $x \in \mathbb{R}^3$ and with a momentum $p \in \mathbb{R}^3$ is

$$\frac{|p|^2}{2m} - \frac{e^2}{4\pi\epsilon_0|x|}.$$

In a convenient choice of units we can take $2m = e^2/(4\pi\epsilon_0) = 1$. The quantum energy of the electron is then

$$\int_{\mathbb{R}^3} |p|^2 |\widehat{\psi}(p)|^2 \, dp - \int_{\mathbb{R}^3} \frac{|\psi(x)|^2}{|x|} \, dx = \left\langle \psi, \left(-\Delta - \frac{1}{|x|} \right) \psi \right\rangle$$

The Hamiltonian

$$H_1 = -\Delta - \frac{1}{|x|}$$

appearing in the energy does not act on the spin variable. To apply H_1 on a wavefunction $\psi \in L^2(\mathbb{R}^3;\mathbb{C}^2)$, we just have to apply H_1 to each component of the wavefunction ψ. This means H_1 can be written with respect to the spin variable as a diagonal matrix.

$$\begin{pmatrix} -\Delta - \frac{1}{|x|} & 0 \\ 0 & -\Delta - \frac{1}{|x|} \end{pmatrix}.$$

So when we take care of the spin, the spectrum of the hydrogen atom is the same as that of the hydrogen atom without spin, but each eigenvalue has its multiplicity multiplied by two. This means that in every state of the spectrum without spin, we can put one electron with the spin $\uparrow$, and one with the spin $\downarrow$.

The spectrum of H_1 on $L^2(\mathbb{R}^3;\mathbb{C})$ can be computed explicitly. The first eigenvalue is

$$\lambda_1 = -1$$

with corresponding eigenfunction

$$\psi_1(x) = \frac{1}{2^{3/2}\sqrt{\pi}} e^{-|x|/2}.$$

The next eigenvalues are given by the formula

$$\lambda_n = -\frac{1}{n^2}, \qquad n \geq 1.$$

The states of N quantum particles in $\mathbb{R}^3$

Let us consider a system composed of N identical and indistinguishable particles. The state of the system is now represented by a normalized wavefunction

$$\Psi(x_1,\sigma_1,...,x_N,\sigma_N) \in L^2((\mathbb{R}^3 \times \{1,...,q\})^N;\mathbb{C})$$

where $|\Psi(x_1,\sigma_1,...,x_N,\sigma_N)|^2$ is the density of probability that particle number 1 be in $x_1 \in \mathbb{R}^3$ with 'spin' $\sigma_1 \in \{1,...,q\}$, particle number 2 be in $x_2 \in \mathbb{R}^3$ with 'spin' $\sigma_2 \in \{1,...,q\}$, etc ... Since we have assumed that the particles are indistinguishable, the probability to find particle number 1 in x_1 with 'spin' σ_1 and particle number 2 in x_2 with 'spin' σ_2, is the same as the one to find particle number 1 in x_2 with 'spin' σ_2 and particle number 2 in x_1 with 'spin' σ_1. This is equivalent to say that the function

$$(x_1,\sigma_1,...,x_N,\sigma_N) \mapsto |\Psi(x_1,\sigma_1,...,x_N,\sigma_N)|^2$$

must be symmetric with respect to exchanges of its variables $(x_i,\sigma_i) \in \mathbb{R}^3 \times \{1,...,q\}$. Then there are two possibilities:

- Either $(x_1,\sigma_1,...,x_N,\sigma_N) \mapsto \Psi(x_1,\sigma_1,...,x_N,\sigma_N)$ is symmetric, which means

  $$\Psi(x_{\tau(1)},\sigma_{\tau(1)},...,x_{\tau(N)},\sigma_{\tau(N)}) = \Psi(x_1,\sigma_1,...,x_N,\sigma_N)$$

  for all permutations $\tau \in \mathcal{S}_N$. In this case Ψ describes particles called *bosons*. We will denote by $L^2_s((\mathbb{R}^3 \times \{1,...,q\})^N,\mathbb{C})$ the subspace of such symmetric functions.

- Or $(x_1, \sigma_1, ..., x_N, \sigma_N) \mapsto \Psi(x_1, \sigma_1, ..x_N, \sigma_N)$ is antisymmetric, which means

$$\Psi(x_{\tau(1)}, \sigma_{\tau(1)}, ..., x_{\tau(N)}, \sigma_{\tau(N)}) = \epsilon(\tau)\Psi(x_1, \sigma_1, ..x_N, \sigma_N)$$

  for all permutations $\tau \in \mathcal{S}_N$, where $\epsilon(\tau) \in \pm 1$ is the signature of the permutation τ. In this case Ψ describes particles called *fermions*. We will denote by $L^2_a((\mathbb{R}^3 \times \{1, ..., q\})^N, \mathbb{C})$ the subspace of the antisymmetric functions.

In this thesis, we will only consider fermions which means that we have to work with antisymmetric wavefunctions.

Remark 1.1.1. *Exchanging two space variables x_i without exchanging the corresponding σ_i has no meaning.*

Example 1.1.2 (Atoms and molecules)**.** *One famous example is that of the electrons in a molecule, with the nuclei treated as classical fixed pointlike particles. The Hamiltonian for the N electrons is the operator [31]*

$$H_N = \sum_{j=1}^{N} (-\Delta)_{x_j} + V(x_j) + \sum_{1 \le k < \ell \le N} \frac{1}{|x_k - x_\ell|}$$

where V is the external potential induced by the nuclei:

$$V(x) = -\sum_{m=1}^{M} \frac{z_m}{|x - R_m|}.$$

The ambient Hilbert space is the fermionic space $L^2_a((\mathbb{R}^3 \times \{\uparrow, \downarrow\})^N, \mathbb{C})$ and the energy is $\langle \Psi, H_N \Psi \rangle$. In this thesis we will be mainly interested in attractive *systems which are not submitted to any external potential V.*

1.1.2 Basis of the N-body spaces, tensor products

Tensor product of Hilbert spaces

We fix as space for one quantum particle

$$\mathfrak{H} := L^2(\mathbb{R}^3, \mathbb{C}^q) = L^2(\mathbb{R}^3 \times \{1, ..., q\}, \mathbb{C}).$$

We define the tensor product of two functions $\varphi_1, \varphi_2 \in \mathfrak{H}$ by

$$(\varphi_1 \otimes \varphi_2)(x_1, \sigma_1; x_2, \sigma_2) = \varphi_1(x_1, \sigma_1)\varphi_2(x_2, \sigma_2),$$

which is a function of $L^2(\mathbb{R}^3 \times \{1, ..., q\} \times \mathbb{R}^3 \times \{1, ..., q\}, \mathbb{C})$ by Fubini's theorem. We know that

$$L^2((\mathbb{R}^3, \{1, ..., q\}) \times (\mathbb{R}^3, \{1, ..., q\})) = L^2(\mathbb{R}^3, \{1, ..., q\}, \mathbb{C}) \otimes L^2(\mathbb{R}^3, \{1, ..., q\}, \mathbb{C})$$

which means that $L^2((\mathbb{R}^3,\{1,...,q\})\times(\mathbb{R}^3,\{1,...,q\}))$ is the closure of the vector space spaned by all the finite linear combinations of tensor product of the form $\varphi_1\otimes\varphi_2$. More precisely, if $(\varphi_i)_{i\geq 1}$ is an orthonormal basis of $\mathfrak{H}$, then $(\varphi_i\otimes\varphi_j)_{i,j\geq 1}$ is an orthonormal basis of $\mathfrak{H}\otimes\mathfrak{H}$.

This is true for more than two Hilbert spaces: For all N we have

$$\mathfrak{H}_N := \bigotimes_1^N \mathfrak{H} = L^2((\mathbb{R}^3\times\{1,...,q\})^N,\mathbb{C}),$$

the N-fold tensor product.

Now we look at the symmetric subspaces of $\mathfrak{H}_N$.

Fermionic case

We consider N functions $\varphi_1,...,\varphi_N$ in $\mathfrak{H}_N$ and define their antisymmetric tensor product by

$$(\varphi_1\wedge...\wedge\varphi_N)(x_1,\sigma_1;...;x_N,\sigma_N)=\frac{1}{\sqrt{N!}}\sum_{\pi\in\mathfrak{S}_N}\epsilon(\pi)\varphi_{\pi(1)}(x_1,\sigma_1)...\varphi_{\pi(N)}(x_N,\sigma_N).$$

This is an antisymmetric function with respect to exchange of the variables (x_i,σ_i). Hence we have

$$\varphi_1\wedge...\wedge\varphi_N\in L^2_a((\mathbb{R}^3)^N,\mathbb{C}^q).$$

This tensor product can be written as a determinant

$$\begin{aligned}(\varphi_1\wedge...\wedge\varphi_N)(x_1,...,x_N)&=\frac{1}{\sqrt{N!}}\det(\varphi_i(x_j,\sigma_j))\\&=\frac{1}{\sqrt{N!}}\det\begin{pmatrix}\varphi_1(x_1,\sigma_1) & ... & \varphi_N(x_1,\sigma_1)\\ \varphi_1(x_2,\sigma_2) & ... & \varphi_N(x_2,\sigma_2)\\ . & ... & .\\ . & ... & .\\ . & ... & .\\ \varphi_1(x_N,\sigma_N) & ... & \varphi_N(x_N,\sigma_N)\end{pmatrix}\end{aligned}$$

which is called a *Slater determinant* in the Physics and Chemistry literature.

Similarly to the usual tensor product $\otimes$, it can be shown that $(\varphi_{j_1}\wedge\varphi_{j_2}\wedge...\wedge\varphi_{j_N})$ is an orthonormal basis of $L^2_a((\mathbb{R}^3\times\{1,...,q\})^N,\mathbb{C})$, when $(\varphi_j)_{j\geq 1}$ is an orthonormal basis of $\mathfrak{H}$. We interpret this by saying that

$$L^2_a((\mathbb{R}^3\times\{1,...,q\})^N,\mathbb{C})=\bigwedge_1^N L^2(\mathbb{R}^3\times\{1,...,q\},\mathbb{C}).$$

As we have seen, a system of N fermions is described by a normalized function $\Psi\in L^2_a((\mathbb{R}^3\times\{1,...,q\})^N,\mathbb{C})$. It can therefore be written

$$\Psi=\sum_{i_1<..<i_N}c_{i_1..i_N}\,\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N}$$

with

$$\sum_{i_1<...<i_N} |c_{i_1,...,i_N}|^2 = 1$$

when $(\varphi_i)_{i\geq 1}$ is any orthonormal basis of $L^2(\mathbb{R}^3, \mathbb{C}^q)$.

Even if we will not need it later, we mention the similar properties of the bosonic space.

Bosonic case

We define the symmetric tensor product of N functions $\varphi_i \in L^2(\mathbb{R}^3 \times \{1, ..., q\}, \mathbb{C})$ by

$$(\varphi_1 \vee ... \vee \varphi_N)(x_1, \sigma_1; ...; x_N, \sigma_N) = \frac{1}{N!} \sum_{\pi \in \mathfrak{S}_N} \varphi_{\pi(1)}(x_1, \sigma_1)...\varphi_{\pi(N)}(x_N, \sigma_N),$$

which is a symmetric function with respect to exchange between (x_i, σ_i). The normalization with $N!$ is chosen to ensure that when $\varphi_j = \varphi$ for $j = 1, ..., N$, this special state is normalized (this corresponds to a perfectly condensed system where all the particles are in the same state φ). Similarly it can be shown that $(\varphi_1 \vee \varphi_2 \vee ... \vee \varphi_N)$ is a basis of $L^2_s((\mathbb{R}^3 \times \{1, ..., q\})^N, \mathbb{C})$, which we interpret by saying that

$$L^2_s((\mathbb{R}^3 \times \{1, ..., q\})^N, \mathbb{C}) = \bigvee_1^N L^2(\mathbb{R}^3 \times \{1, ..., q\}, \mathbb{C}).$$

Note that the basis is not orthonormal, but only orthogonal, however.

1.1.3 The N-body Hamiltonian

We now consider a system of N nonrelativistic fermions in $\mathbb{R}^3$, interacting with a potential W. We assume that our particles are not submitted to any external potential. Hence the system is invariant under translations. To describe this N-body system, we use the Schrödinger operator $H(N)$ defined for $N \geqslant 2$ by

$$H(N) = \sum_{i=1}^{N} (-\Delta)_{x_i} + \sum_{1\leq k<\ell\leq N} W(x_k - x_\ell) \tag{1.1.2}$$

The Hamiltonian $H(N)$ acts on the fermionic subspace $\bigwedge_1^N L^2(\mathbb{R}^3 \times \{1, ..., q\}; \mathbb{C})$ of $L^2((\mathbb{R}^3 \times \{1, ..., q\})^N; \mathbb{C})$. The particles have q internal degrees of freedom ($q = 2$ for spin-1/2 particles like electrons).

Most of what follows is valid in an abstract setting. However, for the sake of simplicity, in most of the thesis we will restrict ourselves to the special case mentioned before.

The first term in the definition (1.1.2) of $H(N)$ is a *one-body-operator* (it involves only one x_i at a time) which is just the sum of the nonrelativistic kinetic

energies of each particle in the system. Here $-\Delta_{x_i}$ denotes the Laplacian relatively to the $x_i \in \mathbb{R}^3$ variable. The second term involving W in (1.1.2) is a two-body term (it involves two particles x_k and x_ℓ at a time) which describes the interaction between the fermions. It depends on the relative spatial positions of the particles. In principle $W(x_k - x_\ell)$ could also be a function of the two internal variables $\sigma_k, \sigma_\ell \in \{1, ..., q\}$ of the particles k and ℓ. Again for simplicity, we will assume that W only depends on the space variable $x_k - x_\ell$. Finally, we make the assumption that W is smooth and decays fast enough at infinity to ensure that H is bounded from below. To make this more explicit, we assume everywhere that

$$\boxed{W = W_1 + W_2 \in L^p(\mathbb{R}^3) + L^q(\mathbb{R}^3) \quad \text{for some } 2 \leq p \leq q < \infty.} \tag{1.1.3}$$

Sometimes we will make more precise assumptions on W.

The lowest (ground state) energy of the system is

$$\inf\{\langle \Psi, H(N)\Psi\rangle, \Psi \in \mathfrak{H}_N, \|\Psi\|_{\mathfrak{H}_N} = 1\} = \inf \mathrm{Spec}\big(H(N)\big).$$

Even if W is attractive (that is, negative somewhere), the previous minimization problem never has any minimizer Ψ. This is due to the translation-invariance of the Hamiltonian $H(N)$ and this is best understood by removing the center of mass. This means we make the change of variables [42]

$$x'_0 = \frac{1}{N}\sum_{j=1}^{N} x_j, \quad x'_1 = x_2 - x_1, ..., x'_{N-1} = x_N - x_1.$$

A computation shows that the original Hamiltonian $H(N)$ can be rewritten as

$$\begin{aligned} H(N) &= \frac{|p'_0|^2}{2N} + \left(\sum_{j=1}^{N-1} \frac{|p'_j|^2}{2} + \frac{1}{2}\left|\sum_{j=1}^{N-1} p'_j\right|^2 + \sum_{j=1}^{N-1} W(x'_j) + \sum_{1\leq k<\ell\leq N-1} W(x'_k - x'_\ell)\right) \\ &:= \frac{|p'_0|^2}{2N} + H'(N-1). \end{aligned}$$

Here, as usual $p'_j = -i\nabla_{x'_j}$. The bottom of the spectrum of $H(N)$ is also the bottom of the spectrum of $H'(N-1)$ and, because of the kinetic energy $|p'_0|^2/(2N)$ of the center of mass, there cannot be a ground state for $H(N)$.

To account for the original statistics of our particles, the latter Hamiltonian $H'(N-1)$ is restricted to $(N-1)$–body functions Φ that are symmetric (bosons) or antisymmetric (fermions), and additionally satisfy the following relation

$$\Phi(-x'_1, x'_2 - x'_1, \cdots, x'_{N-1} - x'_1) = \tau\, \Phi(x'_1, x'_2, \cdots, x'_{N-1})$$

with $\tau = 1$ for bosons and $\tau = -1$ for fermions.

If we forget the center of mass variable x'_0, then the Hamiltonian $H'(N-1)$ can itself have a ground state. Now the remaining $(N-1)$ particles feel an effective external potential W as well as an interaction potential (also given by the function W). If $W \geq 0$, then we have not gained anything, however, and the Hamiltonian $H'(N-1)$ also has no bound state. The potential W must be sufficiently negative somewhere to ensure that the bottom of the spectrum of $H'(N-1)$ is an eigenvalue.

1.2 Hartree-Fock approximation

We have seen that the Hamiltonian $H(N)$ never has any bound state, but that it may have some, after we have removed the center of mass. We now consider a nonlinear approximation of the many-body problem which can have bound states even without removing the center of mass. This is then a purely non linear effect.

The Hartree-Fock model is a non linear theory used to approximate the ground state energy of the N-body system. It consists in restricting the N-body energy $\langle H(N)\Psi, \Psi\rangle$ to the class of the functions Ψ which can be written as a Slater determinant (antisymmetrized product) defined as:

$$\Psi = \varphi_1 \wedge \varphi_2 \wedge ... \wedge \varphi_N, \qquad \langle \varphi_i, \varphi_j \rangle_{L^2(\mathbb{R}^3;\mathbb{C}^q)} = \delta_{ij},$$

where $\varphi_1, ..., \varphi_N$ are called orbitals. More precisely the wavefunction which describes the system can be written as

$$\Psi(x_1, \sigma_1, ...x_N, \sigma_N) = \frac{1}{\sqrt{N!}} \det(\varphi_i(x_j, \sigma_j)).$$

The constraint that the $(\varphi)_i$ must be orthonormal can be written in terms of the Gram matrix as

$$\operatorname{Gram} \Phi = I_N$$

where

$$(\operatorname{Gram} \Phi)_{ij} = (\langle \varphi_i, \varphi_j \rangle)_{ij} = \int_{\mathbb{R}^3} \varphi_i^* \varphi_j.$$

Here I_N the identity matrix of $\mathbb{C}^N$.

For any wavefunction $\Psi \in \bigwedge_1^N L^2(\mathbb{R}^3; \mathbb{C}^q)$, it is convenient to define the one-body density matrix γ_Ψ associated with Ψ. This is a self-adjoint trace-class operator acting on $L^2(\mathbb{R}^3; \mathbb{C}^q)$, such that

1. $0 \leqslant \gamma_\Psi \leqslant 1$;
2. $\operatorname{Tr}(\gamma_\Psi) = N$, the number of particules in the system;
3. its kernel is defined by

$$\gamma_\Psi(x, \sigma, y, \sigma') = N \sum_{\sigma_2=1}^{q} \cdots \sum_{\sigma_N=1}^{q} \int_{\mathbb{R}^3} \cdots \int_{\mathbb{R}^3} \Psi(x, \sigma, x_2, \sigma_2 ..., x_N, \sigma_N) \times \\ \times \overline{\Psi(y, \sigma', x_2, \sigma_2, ..., x_N, \sigma_N)} dx_2 \cdots dx_N.$$

The associated density of charge is defined by

$$\begin{aligned} \rho_\Psi(x) &= \sum_{\sigma=1}^{q} \gamma_\Psi(x, \sigma, x, \sigma) \\ &= N \sum_{\sigma_1=1}^{q} \cdots \sum_{\sigma_N=1}^{q} \int_{\mathbb{R}^3} \cdots \int_{\mathbb{R}^3} |\Psi(x, \sigma_1, x_2, \sigma_2, ..., x_N \sigma_N)|^2 dx_2 \cdots dx_N. \end{aligned}$$

For a Hartree-Fock state, one finds that γ_Ψ is precisely the orthogonal projector on the N-dimensional space spanned by $(\varphi_1, ..., \varphi_N)$:

$$\gamma_\Psi = \sum_{i=1}^{N} |\varphi_i\rangle\langle\varphi_i|. \tag{1.2.4}$$

This operator can be written as

$$\gamma_\Psi(x, \sigma, y, \sigma') = \sum_{i=1}^{N} \varphi_i(x, \sigma)\overline{\varphi_i(y, \sigma')}. \tag{1.2.5}$$

When we interpret $L^2(\mathbb{R}^3 \times \{1, ..., q\}, \mathbb{C}) \simeq L^2(\mathbb{R}^3, \mathbb{C}^q)$, this means that we can think about the kernel of γ_Ψ as a $q \times q$ matrix. Equivalently

$$\gamma_\Psi(x, y) = \sum_{i=1}^{N} \varphi_i(x)\varphi_i(y)^* \tag{1.2.6}$$

where $v^* = \overline{v}^T$ is the row vector which contains the complex components of the column vector v. We then get for the density of charge

$$\rho_\Psi(x) = \mathrm{Tr}_{\mathbb{C}^q}\, \gamma_\Psi(x, x) = \sum_{i=1}^{N} |\varphi_i(x)|^2$$

The ground state Hartree-Fock energy for the N-body system is

$$I_{HF}(N) = \inf\left\{\langle\Psi, H^V(N)\Psi\rangle, \Psi \in \mathcal{S}_N, \|\Psi\|_{L^2(\mathbb{R}^3)} = 1\right\} \tag{1.2.7}$$

where $\mathcal{S}_N$ is the submanifold of $\mathfrak{H}_N$ composed of all the functions which can be written as a Slater determinant, and which have a finite kinetic energy:

$$\mathcal{S}_N = \left\{\Psi = \frac{1}{\sqrt{N!}} \det\left(\varphi_i(x_j, \sigma_j)\right),\ \Phi \in H^1(\mathbb{R}^3, \mathbb{C}^q)^N,\ \mathrm{Gram}\,\Phi = I_N\right\}.$$

It is possible to compute the energy of any HF state of this set and one finds

$$\begin{aligned}\mathcal{E}_{HF}(\varphi_1, ..., \varphi_N) &= \langle\Psi, H^N\Psi\rangle \\ &= \sum_{i=1}^{N} \int_{\mathbb{R}^3} |\nabla\varphi_i|^2 + \frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} \rho_\Psi(x)\rho_\Psi(y)W(x-y)dxdy \\ &\qquad - \frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} |\gamma_\Psi(x, y)|^2 W(x-y)dxdy. \end{aligned}\tag{1.2.8}$$

Here we recall that W is the interaction between the particles. In the HF energy (1.2.8), the first term is nothing but the total kinetic energy. The second one is called the *direct term* and it is the classical electrostatic energy of the charge

distribution ρ_Ψ. The third term is purely quantum and it is called the *exchange term.*

When minimizers for (1.2.7) exist, they satisfy the Euler-Lagrange equations

$$\left\{ \begin{array}{rcl} F_\Phi \varphi_i & = & \lambda_i \varphi_i \\ \langle \varphi_i, \varphi_j \rangle & = & \delta_{ij} \end{array} \right.$$

where F_Φ is the *Fock operator*

$$F_\Phi = -\Delta + \left(\sum_{j=1}^N |\varphi_j|^2 \star W(x) \right) - \sum_{j=1}^N \big((\cdot\, \varphi_j) \star W \big) \varphi_j$$

and the λ_i are the Lagrange multipliers associated with the orthonormaly constraints. Let us remark that minimizers (when they exist) are never unique. By translation invariance, we always have a whole manifold of minimizers obtained by translating the system in space.

The fundamental works concerning the existence of HF minimizers are due to Lieb and Simon [32] and to Lions [37] who also constructed excited states. Both papers deal with atoms and molecules. In [15], Friesecke was the first to consider the Hartree-Fock model as a many-body quantum system, again for atom and molecules. The first results for the translation-invariant Hartree-Fock theory mentioned in this Section, are due to Gogny and Lions [18] for models in nuclear physics, to Lenzmann and Lewin in [26] for a pseudo-relativistic gravitational model, and to Lewin in [28] in a more general (non-relativistic) situation. The following is taken from [28]:

Theorem 1.2.1 (Translation-invariant Hartree-Fock theory)**.** *Assume that*

$$W \in L^p(\mathbb{R}^3) + L^q(\mathbb{R}^3)$$

for some $2 \leq p \leq q < \infty$ *and that* $N \geq 2$*. Then the following assertions are equivalent*

1. *All the minimizing sequences* (Ψ_n) *for the HF minimization problem* (1.2.7) *are precompact in* $\bigwedge_1^N H^1(\mathbb{R}^3, \mathbb{C}^q)$*;*

2. *The* binding inequalities

$$I_{\rm HF}(N) < I_{\rm HF}(N-k) + I_{\rm HF}(k), \quad \text{for all } k = 1, ..., N-1. \tag{1.2.9}$$

are satisfied.

Furthermore, if W *is Newtonian at infinity, that is*

$$W(x) \leq -\frac{a}{|x|} \quad \text{for } a > 0 \text{ and } |x| \geq R, \tag{1.2.10}$$

then the previous two equivalent conditions are verified for all $N \geq 2$*.*

It is very important that the binding inequalities in (1.2.9) are quantized (they involve the integer k). This is *not* true anymore in the Hartree-Fock-Bogoliubov model which we present in the next section, see (1.4.31) below.

As we have seen in Section 1.1.3, by translation invariance the Hamiltonian $H(N)$ has no ground state (that is, the bottom of its spectrum cannot be an eigenvalue). But it may have one once the center of mass is removed, if W is sufficiently negative. In Hartree-Fock theory, because of the nonlinearity there can be a ground state, even if the system is translation-invariant. Of course, translation invariance is not lost and there are then infinitely many ground states, obtained by translating the system arbitrarily.

1.3 Second Quantization

In order to present a more precise nonlinear approximation of the original N-body problem, we go back to the N-body operator

$$H(N) = \sum_{i=1}^{N} (-\Delta)_i + \sum_{1 \leq i < j \leq N} W_{ij}$$

defined on the Hilbert space $\mathfrak{H}_N = \bigwedge_1^N \mathfrak{H}$. The number of particles N is fixed, which is called the *canonical picture*. For attractive systems, it is more useful not to fix the particule number at the outset. So we allow for one symmetry breaking namely that of particle number. We replace the fixed number N by an operator $\mathcal{N}$ whose eigenvalues are $0, 1, 2, ...$ Only the average number of particle makes sense for a quantum state. The convenient way of formulating this is to use the formalism of second quantization, which we summarize in this section.

1.3.1 Fock spaces

We recall that the one-particle Hilbert space is

$$\mathfrak{H} = L^2(\mathbb{R}^3 \times \{1, ..., q\}, \mathbb{C}),$$

(but we can take any abstract separable Hilbert space instead). We define like in the previous section the N-fermionic space to be the antisymmetric tensor product

$$\mathfrak{H}_N = \underbrace{\mathfrak{H} \wedge ... \wedge \mathfrak{H}}_{N \text{times}} \qquad \forall N = 1, 2, ...$$

with the identification $\mathfrak{H}_1 = \mathfrak{H}$ and $\mathfrak{H}_0 = \mathbb{C}$. The vector $1 \in \mathbb{C} = \mathfrak{H}_0$ is often denoted as $1 = |0\rangle$ and it is called the vacuum. It is useful to consider all particle numbers at the same time. To do this we introduce the Fock space

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathfrak{H}_N = \mathfrak{H}_0 \oplus \mathfrak{H}_1 \oplus \mathfrak{H}_2 ... \tag{1.3.11}$$

When all particles numbers are considered at the same time, we call this situation the grand canonical picture. In this thesis, we do not study bosonic systems but for completeness we give the definition of the bosonic Fock space

$$\mathcal{F}^B = \bigoplus_{N=0}^{\infty} \bigvee_{1}^{N} \mathfrak{H}.$$

1.3.2 Creation and annihilation operators

Let us introduce operators in the Fock space, which are important tools in studying many body problems. For any vector f in the one-particle Hilbert space $\mathfrak{H}$, we introduce the creation operator $a^\dagger(f) : \mathcal{F} \to \mathcal{F}$ defined on the fermionic Fock space $\mathcal{F}$ by the following action

$$a^\dagger(f)(f_1 \wedge ... \wedge f_N) = f \wedge f_1 \wedge ... \wedge f_N.$$

The annihilation operator $a(f) : \mathcal{F} \to \mathcal{F}$ is by definition the adjoint of the creation operator

$$a(f) = \left(a^\dagger(f)\right)^*.$$

In particular we obtain

$$a(f_1)(f_1 \wedge ... \wedge f_N) = f_2 \wedge ... \wedge f_N,$$

$$a(f_{N+1})(f_1 \wedge ... \wedge f_N) = 0,$$

where the functions $f_1, ..., f_{N+1}$ are assumed to be orthonormal. On the vacuum, they act as follows $a(f)\Omega = 0$ and $a^\dagger(f)\Omega = f$. The physical interpretation of these operators is that $a^\dagger(f)$ creates a fermion in the single particle state f, since it raises the number of particles from N to $N + 1$. The creation and annihilation operators satisfy the Canonical Anticommutation Relations (CAR)

1. $\{a(f), a(g)\} = a(f)a(g) + a(g)a(f) = 0$,
2. $\{a(f), a^\dagger(g)\} = a(f)a^\dagger(g) + a^\dagger(g)a(f) = \langle f, g\rangle_{\mathfrak{H}} I$.

1.3.3 Observables

In this section we explain how one and two-body operators in Fock space can be written using creation and annihilation operators.

1-body operators

Consider a self-adjoint operator A acting on the one-body space $\mathfrak{H}$. We denote by

$$\mathbb{A} = \bigoplus_{N \geq 0} \sum_{i=1}^{N} (A)_{x_i},$$

the operator on the Fock space $\mathcal{F}$ which, in each N-body space $\mathfrak{H}_N$, acts as a sum of operators acting on each variable separately. By convention this operator is 0 on $\mathfrak{H}_0$.

The operator $\mathbb{A}$ can be expressed with the creation and annihilation operators. Let $\{f_i\}_{i\geq 1}$ be an orthonormal basis of $\mathfrak{H}$, such that $f_i \in D(A)$ for all i. A computation shows that

$$\mathbb{A} = \sum_{i,j\geq 1} A_{ij} a^\dagger(f_i) a(f_j), \quad \text{with} \quad A_{ij} = \langle f_i, A f_j \rangle_{\mathfrak{H}}.$$

This can be proved by looking at the action of this operator on the orthonormal basis $(f_{i_1} \wedge \cdots \wedge f_{i_N})_{i_1<\cdots<i_N}$ of $\mathfrak{H}_N$.

Example 1.3.1 (The number operator)**.** *The number operator is defined by*

$$\mathcal{N} = \bigoplus_{N\geq 0} N$$

on the Fock space $\mathcal{F}$. It is just equal to N (times the identity) on any $\mathfrak{H}_N$. In particular, the average number of particles of a state $\Psi \in \mathcal{F}$ is given by the formula

$$\langle \Psi, \mathcal{N}\Psi \rangle = \sum_{n\geq 1} n \|\psi_n\|^2. \tag{1.3.12}$$

In the second quantization formalism, the particle number operator can be written as

$$\mathcal{N} = \sum_{i\geq 1} a^\dagger(f_i) a(f_i).$$

It corresponds to taking $A = 1$.

2-body operators

Consider a self-adjoint operator W acting on the two-body space $\mathfrak{H}_2$. We denote by

$$\mathbb{W} = \bigoplus_{N\geq 0} \sum_{1\leq k<\ell\leq N} W_{x_k,x_\ell},$$

the operator on the Fock space $\mathcal{F}$ which, in each N-body space $\mathfrak{H}_N$, acts as a sum of operators acting on two variables at a time. By convention this is 0 on $\mathfrak{H}_0$ and $\mathfrak{H}_1$.

The operator can also be written in terms of creation and annihilation operators as follows:

$$\mathbb{W} = \sum_{\substack{1\leq i<j \\ 1\leq k<\ell}} W_{ijk\ell}\, a^\dagger(f_i)\, a^\dagger(f_j)\, a(f_\ell)\, a(f_k)$$

where

$$W_{ijk\ell} = \langle f_i \wedge f_j, W f_k \wedge f_\ell \rangle_{\mathfrak{H}_2}.$$

Hamiltonian

Similarly, our N-body Hamiltonian $H(N)$ can also be lifted to Fock space as follows:

$$\mathbb{H} = \bigoplus_{N\geq 0} H(N)$$

which is just the diagonal operator

$$\begin{pmatrix} H(0) & 0 & ... & ... \\ 0 & H(1) & 0 & ... \\ \vdots & 0 & H(2) & 0 \\ \vdots & \vdots & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

It can be written in second quantization form as

$$\mathbb{H} = \sum_{i,j\geq 1} T_{ij}\, a^\dagger(f_i)a(f_j) + \sum_{\substack{1\leq i<j \\ 1\leq k<\ell}} W_{ijk\ell}\, a^\dagger(f_i)\, a^\dagger(f_j)\, a(f_\ell)\, a(f_k) \tag{1.3.13}$$

with

$$T_{ij} = \int_{\mathbb{R}^3} \nabla\overline{f_i}\nabla f_j$$

and

$$W_{ijk\ell} = \sum_{\sigma=1}^{q}\sum_{\sigma'=1}^{q} \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \overline{\big(f_i \wedge f_j\big)(x,\sigma,y,\sigma')} W(x-y) \big(f_k \wedge f_\ell\big)(x,\sigma,y,\sigma')\, dx\, dy.$$

1.3.4 Quasi-free states

A quasi-free state is a special state in Fock space for which it is possible to compute the expectation value of any observable, only in terms of the (generalized) one-particle density matrix.

In order to define the one-particle density matrix properly, let us consider an abstract state ω acting on $\mathcal{F}$, that is, a linear form acting on the space of bounded operators $\mathcal{B}(\mathcal{F})$ such that $\omega(1) = 1$ and $\omega(A^*A) \geqslant 0$ for every A in $\mathcal{B}(\mathcal{F})$. We define the two one-particle matrices $\gamma : \mathfrak{H} \to \mathfrak{H}$ and $\alpha : \mathfrak{H} \to \mathfrak{H}$ by:

$$\langle f, \gamma g\rangle = \omega(a^\dagger(g)a(f)), \qquad \langle f, \alpha\overline{g}\rangle = \omega(a(g)a(f)). \tag{1.3.14}$$

The operator γ is usually called the *one-body density matrix* and α is called the *pairing density matrix* of ω. They satisfy the constraints

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \leq \Gamma = \begin{pmatrix} \gamma & \alpha \\ \alpha^* & 1-\overline{\gamma} \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

where Γ is an operator acting on $\mathfrak{H} \oplus \mathfrak{H}$. Also we have $\alpha^T = -\alpha$ where $\alpha^T = \overline{\alpha^*}$. Note that

$$0 \leq \gamma = \gamma^* \leq 1.$$

Also, writing that $\Gamma^2 \leq \Gamma$, we find that

$$\alpha\alpha^* \leq \gamma - \gamma^2.$$

The kernel $\alpha(x, \sigma; x', \sigma')$ of the operator α can be interpreted as a two-body fermionic wavefunction of Cooper pairs.

Now to any such pair (γ, α), there is a unique state (called quasi-free) in Fock space which is such that the expectation value of any polynomial in the creation and annihilation operators can be computed only in terms of γ and α by Wick formula [3]

$$\omega(e_1, e_2, .., e_{2N-1}) = 0 \tag{1.3.15}$$

and

$$\omega(e_1, e_2, .., e_{2N}) = \sum_{\pi} (-)^{\pi} \omega(e_{\pi(1)} e_{\pi(2)}) ... \omega(e_{\pi(2N-1)} e_{\pi(2N)}) \tag{1.3.16}$$

where $\sum_\pi$ is the sum over permutations π that satisfy $\pi(1) < \pi(3) < ... < \pi(2N-1)$ and $\pi(2j-1) < \pi(2j)$ for all $1 \leq j \leq N$, and the operators $e_1, e_2, ..., e_{2N}$ are each either a $a^\dagger$ or a a. The following is taken from [3]:

Theorem 1.3.1 (Quasi-free states)**.** *Let $0 \leq \Gamma \leq 1$ be an operator on $\mathfrak{H} \oplus \mathfrak{H}$ of the form*

$$\Gamma = \begin{pmatrix} \gamma & \alpha \\ \alpha^* & 1 - \overline{\gamma} \end{pmatrix},$$

with $\gamma^* = \gamma$ and $\alpha^T = -\alpha$, and assume furthermore that $\mathrm{Tr}(\gamma) < \infty$. Then there exists a unique quasi-free state ω with finite particle number such that Γ is the generalized one-body density matrix of ω.

This state ω is a pure state $\omega(A) = \langle \Psi, A\Psi \rangle$ if and only if the 1-pdm Γ with $\mathrm{Tr}(\gamma) < \infty$ is a projection on $\mathfrak{H} \oplus \mathfrak{H}$, i.e. $\Gamma^2 = \Gamma$.

When they have a finite average number of particles, quasi-free states are sometimes called Hartree-Fock-Bogoliubov states in the literature, which is a convention that we will also use. When $\alpha \equiv 0$ and $\gamma^2 = \gamma$ is a projection, then the unique state of Theorem 1.3.1 is nothing else but the usual Hartree-Fock state

$$\Psi = \varphi_1 \wedge \cdots \wedge \varphi_N$$

with $(\varphi_1, ..., \varphi_N)$ an orthonormal basis of the range of γ. Therefore the manifold of HFB states contains the usual HF states.

1.4 Hartree-Fock-Bogoliubov theory

Hartree-Fock-Bogoliubov theory consists in restricting the many-body Hamiltonian $\mathbb{H}$ on the Fock space $\mathcal{F}$ to the special class of quasi-free (HFB) states. This set contains the Slater determinants but it is larger. So HFB theory is a generalization of HF theory presented in Section 1.2.

1.4.1 Hartree-Fock-Bogoliubov states and their energy

We have seen that translation-invariance can be broken in a nonlinear approximation of the linear Schrödinger model, such as Hartree-Fock. When the interaction potential W is attractive ($W \leq 0$), or at least partially attractive ($W \leq 0$ on a set of measure non zero), it is often convenient to allow for another symmetry breaking, namely that of *particle number*. This means that the fixed particle number N is replaced by the number operator $\mathcal{N}$ in Fock space, whose eigenvalues are $0, 1, 2, ...$ Only the *average particle number* is well defined for a quantum state. Allowing to have $\psi_n \neq 0$ for $n \neq N$ is useful to describe some physical properties of attractive systems. In most practical cases it is expected that the variance $\sum_{n\geq 0}(n-N)^2 \left\|\psi_n\right\|^2_{\mathfrak{H}^n}$ will be quite small, i.e. that Ψ will live in a neighborhood of $\mathfrak{H}^N$.

The Hartree-Fock-Bogoliubov (HFB) model generalizes the well-known Hartree-Fock (HF) method and it allows for breaking of particle number in a very simple fashion. The method consists in restricting the many-body Hamiltonian $\mathbb{H}$ on $\mathcal{F}$ to the special class of states called *quasi-free states* (or *Hartree-Fock-Bogoliubov states*), which we have introduced in the previous section.

For the present work, we will only need the formula of the total energy, in terms of γ and α:

$$\begin{aligned}
&\omega(\mathbb{H}) \\
&\quad = \sum_{n\geq 0} \langle \psi_n, H(n)\psi_n \rangle_{\mathfrak{H}^n} \\
&\quad = \mathrm{Tr}(-\Delta)\gamma + \frac{1}{2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\Big(\rho_\gamma(x)\rho_\gamma(y) - |\gamma(x,y)|^2 + |\alpha(x,y)|^2\Big)\,dx\,dy \\
&\quad := \mathcal{E}(\gamma,\alpha)
\end{aligned} \tag{1.4.17}$$

where $\rho_\gamma(x) = \mathrm{Tr}_{\mathbb{C}^q}(\gamma(x,x))$ is the density of particles in the system. This expression is valid whether ω is a pure state or not. The terms in the double integral are respectively called the *direct*, *exchange* and *pairing* terms. Taking $\alpha \equiv 0$ one recovers the usual Hartree-Fock energy which has been introduced in Section 1.2. The value of the HFB energy can be recovered by using the second-quantization formula (1.3.13) of $\mathbb{H}$ as well as the definition of the density matrices γ and α.

Our main goal here is to study the minimization of the nonlinear functional $\mathcal{E}(\gamma,\alpha)$, when γ and α are submitted to the above constraints, and its numerical implementation. We will show below that, under our assumption (1.1.3) on W, the energy $\mathcal{E}$ is well defined in an appropriate function space.

Note that, for a pure state, the variance of the particle number for a HFB state Ψ in Fock space can be expressed only in terms of α by

$$\Big\langle \Psi, \big(\mathcal{N} - \langle \Psi, \mathcal{N}\Psi\rangle_{\mathcal{F}}\big)^2 \Psi \Big\rangle_{\mathcal{F}} = \sum_{n\geq 0}(n-N)^2 \left\|\psi_n\right\|^2_{\mathfrak{H}^n} = 2\ \mathrm{Tr}_{\mathfrak{H}}(\alpha\alpha^*),$$

see Lemma 2.7 in [3]. The spreading of the HFB state over the different spaces $\mathfrak{H}^n$ is therefore determined by the Hilbert-Schmidt norm of the pairing matrix α. We

recover the fact that a mixed HFB state has a given particle number if and only if its pairing matrix α vanishes.

1.4.2 Rigorous definition of the HFB energy

Recall our assumption that $\mathrm{Tr}(\gamma) = N < \infty$. Since $\gamma \geq 0$, this means that γ must be a trace-class operator. Because $\alpha\alpha^* \leq \gamma$, we see that α must be a Hilbert-Schmidt operator. In the following we denote by $\mathfrak{S}_p$ the Schatten spaces with $1 \leq p < \infty$ consisting of the bounded operators A such that

$$\|A\|_{\mathfrak{S}_p} = \mathrm{Tr}(|A|^p)^{1/p} < \infty.$$

The space $\mathfrak{S}_\infty$ denotes the space of compact operators on $L^2(\mathbb{R}^3, \mathbb{C}^q)$, with the operator norm $\|\cdot\|$. So we have $\gamma \in \mathfrak{S}_1$ and $\alpha \in \mathfrak{S}_2$.

The trace-class operator $\gamma \in \mathfrak{S}_1$ has a kernel $\gamma(x, y)$ which can be viewed as a $q \times q$ hermitian matrix for all $x, y \in \mathbb{R}^3$, where we recall that q is an internal degree of freedom like spin. The density ρ_γ is the unique nonnegative function in $L^1(\mathbb{R}^3)$ such that

$$\mathrm{Tr}(\gamma V) = \int_{\mathbb{R}^3} \rho_\gamma(x) V(x) dx$$

for any bounded function V. The pairing density matrix α is Hilbert-Schmidt and it also has a kernel which is a $q \times q$ matrix. Because of the constraint that $\alpha^T = -\alpha$, it must be antisymmetric: $\alpha(x, y)^T = -\alpha(y, x)$, where ^{T} is the transposition of matrices. Recall that we also have the condition

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \leq \Gamma := \begin{pmatrix} \gamma & \alpha \\ \alpha^* & 1 - \overline{\gamma} \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.4.18}$$

The set of one-particle density matrices (γ, α) of mixed HFB states is therefore a convex set, whose extremal points correspond to pure state, $\Gamma^2 = \Gamma$. The natural question arises whether a minimizer, when it exists, is automatically a pure state. The answer to this question is positive in many situations, as we will see below.

Before turning to the comparison between the minimization among pure and mixed states, we first introduce the variational sets on which the energy is well defined. The sets of all pure and mixed HFB states with finite kinetic energy are respectively given by

$$\mathcal{P} = \{(\gamma, \alpha) \in \mathfrak{S}_1(\mathfrak{H}) \times \mathfrak{S}_2(\mathfrak{H}) : \alpha^T = -\alpha, \Gamma = \Gamma^* = \Gamma^2, \mathrm{Tr}(-\Delta)\gamma < \infty\} \tag{1.4.19}$$

and the closed and convex subset

$$\mathcal{K} = \{(\gamma, \alpha) \in \mathfrak{S}_1(\mathfrak{H}) \times \mathfrak{S}_2(\mathfrak{H}) : \alpha^T = -\alpha, 0 \leq \Gamma = \Gamma^* \leq 1, \mathrm{Tr}(-\Delta)\gamma < \infty\}, \tag{1.4.20}$$

The expression $\mathrm{Tr}(-\Delta)\gamma$ is to be understood in the sense of quadratic forms, that is

$$\mathrm{Tr}(-\Delta)\gamma = \sum_{k=1}^{3} \mathrm{Tr}(p_k \gamma p_k) \in [0, +\infty], \qquad \text{with} \quad p_k = -i\partial_{x_k}.$$

In practice, we want to fix the average number of particles. To do that, we add the assumption $\mathrm{Tr}_{\mathfrak{H}}\gamma = N$ and introduce the corresponding sets

$$\mathcal{P}_N = \{(\gamma,\alpha) \in \mathcal{P} : \mathrm{Tr}_{\mathfrak{H}}\gamma = N\} \tag{1.4.21}$$

and

$$\mathcal{K}_N = \{(\gamma,\alpha) \in \mathcal{K} : \mathrm{Tr}_{\mathfrak{H}}\gamma = N\}, \tag{1.4.22}$$

of pure and mixed HFB with average particle number N.

The following lemma says that the energy is a well-defined functional on the largest of the above sets $\mathcal{K}$, and that it is bounded from below on $\mathcal{K}(N)$ for any $N \geq 0$.

Lemma 1.4.1 (The HFB energy is bounded-below on $\mathcal{K}(N)$)**.** *When $W = W_1 + W_2 \in L^p(\mathbb{R}^3) + L^q(\mathbb{R}^3)$ with $2 \leq p \leq q < \infty$, then $\mathcal{E}(\gamma,\alpha)$ is well defined for any $(\gamma,\alpha) \in \mathcal{K}$. It also satisfies a bound of the form*

$$\forall (\gamma,\alpha) \in \mathcal{K}, \qquad \mathcal{E}(\gamma,\alpha) \geq \frac{1}{2}\,\mathrm{Tr}(-\Delta)\gamma - C(N) \tag{1.4.23}$$

for some constant $C(N)$ depending only on $N = \mathrm{Tr}(\gamma)$.

Proof. The assumption that $W = W_1 + W_2 \in L^p(\mathbb{R}^3) + L^q(\mathbb{R}^3)$ with $2 \leq p \leq q < \infty$ implies that W is relatively form-bounded with respect to the Laplacian, with relative bound as small as we want [9]. This means $|W| \leq \epsilon(-\Delta) + C_\epsilon$ in the sense of quadratic forms, for all $\epsilon > 0$ and for some constant C_ϵ. This can now be used to verify that the energy is well defined under the assumption that $\mathrm{Tr}(-\Delta)\gamma < \infty$. First, we have for the direct term

$$\begin{aligned}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |W(x-y)|\,\rho_\gamma(x)\rho_\gamma(y)\,dx\,dy &\leq \epsilon N \int_{\mathbb{R}^3} |\nabla\sqrt{\rho_\gamma}|^2 + C_\epsilon\, N^2 \\ &\leq \epsilon N\,\mathrm{Tr}(-\Delta)\gamma + C_\epsilon\, N^2,\end{aligned}$$

where in the last line we have used the Hoffmann-Ostenhof inequality [23],

$$\int_{\mathbb{R}^3} |\nabla\sqrt{\rho_\gamma}|^2 \leq \mathrm{Tr}(-\Delta)\gamma. \tag{1.4.24}$$

The exchange term is bounded similarly by applying the inequality $|W| \leq \epsilon(-\Delta) + C_\epsilon$ in x with y fixed:

$$\begin{aligned}&\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |W(x-y)|\,|\gamma(x,y)|^2\,dx\,dy \\ &\qquad\leq \epsilon \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |\nabla_x\gamma(x,y)|^2\,dx\,dy + C_\epsilon \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |\gamma(x,y)|^2\,dx\,dy \\ &\qquad= \epsilon\,\mathrm{Tr}(-\Delta)\gamma^2 + C_\epsilon\,\mathrm{Tr}\,\gamma^2 \leq \epsilon\,\mathrm{Tr}(-\Delta)\gamma + C_\epsilon N,\end{aligned}$$

since $\gamma^2 \leq \gamma$. Similarly we have, since $\alpha\alpha^* \leq \gamma - \gamma^2 \leq \gamma$,

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |W(x-y)|\,|\alpha(x,y)|^2\,dx\,dy \leq \operatorname{Tr}\big(\epsilon(-\Delta)+C_\epsilon\big)\alpha\alpha^* \leq \epsilon\operatorname{Tr}(-\Delta)\gamma + C_\epsilon N.$$

All this shows that all the terms in the energy are well defined when $(\gamma,\alpha)\in\mathcal{K}(N)$. Also, we have

$$\mathcal{E}(\gamma,\alpha) \geq \big(1-\epsilon-\epsilon N/2\big)\operatorname{Tr}(-\Delta)\gamma - C_\epsilon N - C_\epsilon N^2/2. \tag{1.4.25}$$

Taking $\epsilon = 1/(2+N)$ finishes the proof. □

Lemma 1.4.1 allows us to define the minimization problems for pure and mixed states as follows:

$$I(N) := \inf_{(\gamma,\alpha)\in\mathcal{K}(N)} \mathcal{E}(\gamma,\alpha), \tag{1.4.26}$$

$$J(N) := \inf_{(\gamma,\alpha)\in\mathcal{P}(N)} \mathcal{E}(\gamma,\alpha). \tag{1.4.27}$$

Of course we have $J(N)\geq I(N)$ since $\mathcal{P}(N)\subset\mathcal{K}(N)$. In many cases, we have that $I(N)=J(N)$ and that any minimizer, when it exists, is automatically a pure HFB state. We give two results in the literature going in this direction. The first deals with *purely repulsive interactions* and it is Lieb's famous variational principle [29] (see also Thm. 2.11 in [3]).

Theorem 1.4.1 (Lieb's HF Variational Principle [29])**.** *Assume that*

$$W \geq 0$$

and let N be an integer. Then for any $(\gamma,\alpha)\in\mathcal{K}(N)$, there exists $(\gamma',0)\in\mathcal{P}(N)$ such that

$$\mathcal{E}(\gamma,\alpha) \geq \mathcal{E}(\gamma',0).$$

In particular, we have $I(N)=J(N)$.

If $W>0$ a.e., then any minimizer for $I(N)$, when it exists, is necessarily of the form $(\gamma',0)$ with $(\gamma')^2=\gamma'$.

We see that for repulsive interactions, $W>0$, there is never pairing ($\alpha\equiv 0$) and the ground state is always a pure HF state, that is, a Slater determinant. The fact that, in HF theory, one can minimize over mixed states and get the same ground state energy is very important from a numerical point of view. This was used by Cancès and Le Bris [6, 7] to derive well-behaved numerical strategies, to which we will come back later in Section 3.2.

The second result with the contraint imposed is the famous theorem of Bach, Fröhlich and Jonsson, for purely attractive interactions:

Theorem 1.4.2 (HFB Constrained Variational Principle)**.** *Assume that the number of spin states is $q = 2$ (spin-1/2 fermions), and that W can be decomposed in the form*

$$W(x-y) = -\int_\Omega d\mu(\omega) f_\omega(x)\, f_\omega(y) \tag{1.4.28}$$

on a given measure space (Ω, μ), with $(f_\omega)_{\omega\in\Omega}$ a family of bounded real-valued functions on $\mathbb{R}^3$. Then for any $(\gamma, \alpha) \in \mathcal{K}_N$ (with N a given integer), we have

$$\mathcal{E}(\gamma, \alpha) \geq \mathcal{E}(\gamma', \alpha'), \tag{1.4.29}$$

with

$$\gamma' = g \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \alpha' = \sqrt{g(1-g)} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{1.4.30}$$

(the second matrices act on the spin variables), and

$$g = g^T = \overline{g} = \frac{\gamma_{\uparrow\uparrow} + \gamma_{\downarrow\downarrow} + \overline{\gamma_{\uparrow\uparrow}} + \overline{\gamma_{\downarrow\downarrow}}}{4}.$$

This HFB state is pure: $(\gamma', \alpha') \in \mathcal{P}(N)$. In particular, we have $I(N) = J(N)$. Furthermore, if $W < 0$ a.e., then any ground state is necessarily of the previous form.

Remark 1.4.1. *Note that N does not have to be an even integer in this result. Since $\mathrm{Tr}_{L^2(\mathbb{R}^3)}(g) = N/2$, the operator g must have one eigenvalue different from 1 when N is odd, and it follows that $\alpha \neq 0$ in this special case.*

In Theorem 1.4.2, we have decomposed the operator γ acting on $L^2(\mathbb{R}^3 \times \{\uparrow, \downarrow\}, \mathbb{C})$ according to the spin variables as follows:

$$\gamma = \begin{pmatrix} \gamma_{\uparrow\uparrow} & \gamma_{\downarrow\uparrow} \\ \gamma_{\uparrow\downarrow} & \gamma_{\downarrow\downarrow} \end{pmatrix}.$$

Theorem 1.4.2 says that when W satisfies (1.4.28), one can minimize over states which are pure, real, and have a simple spin symmetry. The antisymmetry of α is only contained in the spin variables, hence the Cooper pairs are automatically in a singlet state. Of course, one can express the total energy only in terms of the real operator g, as follows

$$\begin{aligned}\mathcal{E}(\gamma', \alpha') = 2\ \mathrm{Tr}_{L^2(\mathbb{R}^3)}(-\Delta)g \\ + \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\Big(2\rho_g(x)\rho_g(y) - |g(x,y)|^2 + |\sqrt{g(1-g)}(x,y)|^2\Big).\end{aligned}$$

In practice it will be more convenient to keep a pairing term $a(x, y)$ not *a priori* related to g and to optimize over both g and a, that is, to consider mixed states. When W satisfies the assumptions of the theorem, any ground state will automatically lead to $a = \pm\sqrt{g(1-g)}$.

Let us conclude our comments on Theorem 1.4.2, by noticing that several simple attractive potentials W can be written in the form (1.4.28). For instance the Fefferman-de la Llave formula [12]

$$\frac{1}{|x-y|} = \frac{1}{\pi}\int_0^\infty \frac{dr}{r^5}\int_{\mathbb{R}^3} dz\ \mathbb{1}_{B(z,r)}(x)\,\mathbb{1}_{B(z,r)}(y)$$

shows that a simple Newtonian interaction $W(x-y) = -|x-y|^{-1}$ is covered (here $\mathbb{1}_{B(z,r)}$ is the characteristic function of the ball centered at z, of radius r). Hainzl and Seiringer showed in [21] that any smooth enough radial function W can be written in the form

$$W(x-y) = \int_0^\infty dr\,\tilde{g}(r)\int_{\mathbb{R}^3} dz\ \mathbb{1}_{B(z,r)}(x)\,\mathbb{1}_{B(z,r)}(y)$$

for some explicit function $\tilde{g}$, whose sign can easily be studied.

Remark 1.4.2. *The total HFB energies $I(N)$ is an upper bound to the full grand canonical Schrödinger energy in Fock space*

$$J_{\rm Sch}(N) := \inf\{\omega(\mathbb{H}),\ \omega(\mathcal{N}) = N\} = \inf_{\substack{\sum_{n\geq 0}\alpha_n = 1\\ \sum_{n\geq 0} n\alpha_n = N}} \sum_{n\geq 0}\alpha_n \inf \mathrm{Spec}(H(n)).$$

Without appropriate stability assumptions on W, the latter can be equal to $-\infty$. In general, it is much lower than the canonical energy $\inf \mathrm{Spec}(H(N))$.

1.4.3 Existence results and properties of minimizers

The HFB minimization problem for mixed states is

$$I(N) = \inf\{\mathcal{E}(\gamma,\alpha) : (\gamma,\alpha)\in\mathcal{K}_N\}.$$

The following deals with the existence of minimizers and it is the equivalent of Theorem 1.2.1.

Theorem 1.4.3 (Existence of minimizers and compactness of minimizing sequences)**.** *We assume as before that $W = W_1 + W_2 \in L^p(\mathbb{R}^3) + L^q(\mathbb{R}^3)$ with $2 \leq p \leq q < \infty$. Let $\lambda > 0$. Then the following assertions are equivalent:*

1. *All the minimizing sequences $(\gamma_n,\alpha_n) \subset \mathcal{K}(\lambda)$ for $I(\lambda)$ are precompact up to translations, that is there exists a sequence $(x_k)\subset\mathbb{R}^3$ and $(\gamma,\alpha)\in\mathcal{K}(\lambda)$ such that, for a subsequence,*

$$\lim_{k\to\infty} \|(1-\Delta)^{1/2}(\tau_{x_k}\gamma_{n_k}\tau_{-x_k}-\gamma)(1-\Delta)^{1/2}\|_{\mathfrak{S}_1} = \lim_{k\to\infty}\|(1-\Delta)^{1/2}(\tau_{x_k}\alpha_{n_k}\tau_{-x_k}-\alpha)\|_{\mathfrak{S}_2} = 0.$$

In particular (γ,α) is a minimizer for $I(\lambda)$.

2. The binding inequalities

$$I(\lambda) < I(\lambda - \mu) + I(\mu) \qquad \text{for all } 0 < \mu < \lambda \tag{1.4.31}$$

are satisfied.

Furthermore, if W is Newtonian at infinity, that is

$$W(x) \leq -\frac{a}{|x|} \quad \text{for } a > 0 \text{ and } |x| \geq R, \tag{1.4.32}$$

then the previous two equivalent conditions are verified.

The proof of Theorem 1.4.3 follows along the lines of the paper [26] of Lenzmann and Lewin, in which the Laplacian $-\Delta$ was replaced by the more complicated operator $\sqrt{1-\Delta}-1$. It will not be detailed here.

The assumption that the interaction is Newtonian at infinity is a big simplification, as it means that two subsystems receding from each other always attract at large distances. One can expect that minimizers exist even if the potential is not attractive at infinity, as soon as it has a sufficiently large negative component. A typical effective potential $W(x)$ used in nuclear physics is nonnegative for small and large $|x|$, and has a negative well at intermediate distances [39]. At infinity it typically decays like $+\kappa|x|^{-1}$ for two protons, and exponentially fast when one of the two particles is a neutron. Even in HF theory, we are not aware of any existence result dealing with such potentials, however.

Note the difference between the unquantized binding condition (1.4.31) in HFB theory and the quantized condition (1.2.9) in HF theory.

The form of the nonlinear equation solved by minimizers is well-known in the physics literature, and it was re-explained in [3]. The following result summarizes some known properties.

Theorem 1.4.4 (HFB equation and properties of minimizers [3, 26])**.** *A HFB minimizer on $\mathcal{K}(N)$ solves the nonlinear equation*

$$\Gamma = \mathbb{1}_{(-\infty,0)}\big(F_\Gamma - \mu\mathcal{N}\big) + \delta \tag{1.4.33}$$

where $0 \leq \delta \leq \mathbb{1}_{\{0\}}(F_\Gamma - \mu\mathcal{N})$ has the same form as Γ, and where

$$\mathcal{N} := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad F_\Gamma = \begin{pmatrix} h_\gamma & \pi \\ \pi^* & -\overline{h_\gamma} \end{pmatrix} \tag{1.4.34}$$

with $h_\gamma = -\Delta + \rho_\gamma * W - W(x-y)\gamma(x,y)$ and $\pi(x,y) = \alpha(x,y)W(x-y)$.

If $W(x-y) = -\kappa|x-y|^{-1}$ (Newtonian interaction) and N is an integer, then all the minimizers are of the special form (1.4.30)*. In this case, we have either $\alpha \equiv 0$ and γ is a projector of rank N, or $\alpha \neq 0$ and γ has an infinite rank.*

The nonlinear equation (1.4.33) is in principle similar to the usual equation obtained in HF theory,

$$\gamma = \mathbb{1}_{(-\infty,\mu)}(h_\gamma) + \delta \tag{1.4.35}$$

Indeed, (1.4.33) reduces to (1.4.35) when $\alpha \equiv 0$. Let us however emphasize that the mean-field operator F_Γ has a spectrum which is symmetric with respect to 0. Hence F_Γ is usually not even semi-bounded, on the contrary to h_γ which is always bounded from below. Furthermore, the operator $\mathcal{N}$ does *not* commute with F_Γ (except when $\alpha \equiv 0$) and the equation cannot be written in a simple form as in HF theory. This will cause several difficulties to which we will come back at length later.

The fact that γ has an infinite rank when there is pairing, $\alpha \neq 0$, is a dramatic change of behavior compared to the simple HF case. However, no information on the decay of the eigenvalues of γ seems to be known.

An important open question is to show that minimizers actually exhibit non-vanishing pairing $\alpha \neq 0$, at least for a sufficiently strong attractive potential W. On heuristic grounds, one expects such a phenomenon of *"Cooper pair formation"* to be energetically favorable due to the attractive interaction among particles. However, it seems to be a formidable task to find mathematical proof for this claim. The existence of pairing is known in some particular situations (when N is odd and W is Newtonian, see Remark 1.4.1, for the Hubbard model [3], or in BCS theory [13, 14, 19, 22]), but for the model presented here, we are not aware of any result of this sort. One of the purposes of this thesis is to investigate this question numerically.

Chapter 2

Discretization

In this chapter we explain how to discretize the HFB energy in order to handle it on a computer. First in Section 2.1 we show that, under suitable conditions on the discretization space V_h, the minimum energy and any sequence of ground states converge to the solution of the problem in the infinite dimensional case. Then, in Section 2.2 we give the form of the discretized HFB energy and equations. Finally, in Section 2.3 we will discuss the treatment of symmetries in the system.

2.1 Convergence analysis

In this section we show that the HFB ground state energy in a finite basis converges to the true HFB ground state energy when the size of the basis grows. We consider a sequence of finite-dimensional spaces $V_h \subset H^1(\mathbb{R}^3, \mathbb{C}^q)$ for $h \to 0$. We assume that any function $f \in H^1(\mathbb{R}^3, \mathbb{C}^q)$ can be approximated by function in V_h:

$$\forall f \in H^1(\mathbb{R}^3, \mathbb{C}^q), \quad \exists f_h \in V_h \quad \text{such that} \quad \|f - f_h\|_{H^1} \underset{h\to 0}{\longrightarrow} 0. \tag{2.1.1}$$

We typically think of a sequence V_h given by the Finite Elements Method [1]. Let π_h denote the orthogonal projection on V_h in $L^2(\mathbb{R}^3, \mathbb{C}^q)$[1]. We define the set of density matrices living on V_h (with average particle number N) as follows:

$$\mathcal{K}_h(N) = \left\{(\gamma, \alpha) \in \mathcal{K}(N) \ : \ \pi_h \gamma \pi_h = \gamma, \ \pi_h \alpha \, \overline{\pi_h} = \alpha\right\}, \tag{2.1.2}$$

where $\mathcal{K}(N)$ is defined in Chapter 1 by (1.4.22). All the states in $\mathcal{K}_h(N)$ have a finite rank and can be expanded in a basis of V_h. The corresponding minimization problem is now

$$I_h(N) = \inf_{(\gamma,\alpha)\in\mathcal{K}_h(N)} \mathcal{E}(\gamma, \alpha). \tag{2.1.3}$$

Since $\mathcal{K}_h(N) \subset \mathcal{K}(N)$ by definition, it is obvious that $I_h(N) \geq I(N)$. The following result is a simple consequence of Theorem 1.4.3.

[1] Note that $\pi_h f \to f$ when $h \to 0$ for any fixed $f \in L^2$. This follows from the density of H^1 is L^2: we fix $\epsilon > 0$ and take $g \in H^1$ such that $\|f - g\|_{L^2} \leq \epsilon$. Then $\|(1 - \pi_h)g\|_{L^2} \leq \|g - g_h\|_{L^2} \leq \|g - g_h\|_{H^1} \to 0$ and the statement follows.

Theorem 2.1.1 (Convergence of the approximate HFB problem)**.** *When* $W = W_1 + W_2 \in L^p(\mathbb{R}^3) + L^q(\mathbb{R}^3)$ *with* $2 \leq p \leq q < \infty$ *and under Assumption* (2.1.1) *on the sequence* (V_h)*, we have*

$$\lim_{h\to 0} I_h(N) = I(N). \tag{2.1.4}$$

If the binding inequality (1.4.31) *is satisfied, then any sequence of minimizers* $(\gamma_h, \alpha_h) \in \mathcal{K}_h(N)$ *for* $I_h(N)$ *converges, up to a subsequence and up to a translation, to a minimizer* $(\gamma, \alpha) \in \mathcal{K}(N)$ *of* $I(N)$*, in the sense that*

$$\lim_{h_k\to 0} \left\|(1-\Delta)^{1/2}(\tau_{x_k}\gamma_{h_k}\tau_{-x_k} - \gamma)(1-\Delta)^{1/2}\right\|_{\mathfrak{S}_1} = \lim_{h_k\to 0} \left\|(1-\Delta)^{1/2}(\tau_{x_k}\alpha_{h_k}\tau_{-x_k} - \alpha)\right\|_{\mathfrak{S}_2} = 0. \tag{2.1.5}$$

Proof. We only have to show that $I_h(N) \to I(N)$ as $h \to 0$. Then, any sequence of exact minimizers (γ_h, α_h) for $I_h(N)$ is also a minimizing sequence for $I(N)$. Applying Theorem 1.4.3 concludes the proof. We need the following lemma, whose proof is given below.

Lemma 2.1.1. *The finite-rank operators are dense in* $\mathcal{K}(N)$*.*

Let $(\gamma, \alpha) \in \mathcal{K}(N)$ be any such finite rank operator. Let $(f_i)_{i=1}^K$ be an orthonormal basis of the range of γ. By Löwdin's theorem (Lemma 13 in [28]), we know that the two-body wavefunction α can be expanded in the same basis $(f_1, ..., f_K)$. Now, for every $i = 1, .., K$, we apply (2.1.1) and take a sequence $f_i^h \in V_h$ be such that $f_i^h \to f_i$ in $H^1(\mathbb{R}^3)$. The system $(f_i^h)_{i=1}^k$ is not necessarily orthonormal but we have $\langle f_i^h, f_j^h \rangle \to \langle f_i, f_j \rangle = \delta_{ij}$. Applying the Gram-Schmidt procedure, we can therefore replace the $(f_i^h)_{i=1}^K$ by an orthonormal set $(g_i^h)_{i=1}^K \subset V_h$ having the same properties. An equivalent procedure is to take $g_i^h = \sum_{j=1}^K (S_h^{-1/2})_{ji} f_j^h$ where S_h is the Gram matrix $(\langle f_i^h, f_j^h \rangle)_{i,j}$. Let now U_h be any unitary operator on $L^2(\mathbb{R}^3)$ which is such that $U_h f_i = g_i^h$ for all $i = 1, ..., K$. We then take $\gamma_h := U_h \gamma U_h^*$ and $\alpha_h := U_h \alpha U_h^T$. In words, we just replace f_i by g_i^h in the decomposition of γ and α. To see that $(\gamma_h, \alpha_h) \in \mathcal{K}(N)$, we just notice that

$$\begin{pmatrix} \gamma_h & \alpha_h \\ \alpha_h^* & 1 - \overline{\gamma_h} \end{pmatrix} = \begin{pmatrix} U_h & 0 \\ 0 & \overline{U_h} \end{pmatrix} \begin{pmatrix} \gamma & \alpha \\ \alpha^* & 1 - \overline{\gamma} \end{pmatrix} \begin{pmatrix} U_h^* & 0 \\ 0 & \overline{U_h}^* \end{pmatrix}.$$

Note also that $\mathrm{Tr}(\gamma_h) = \mathrm{Tr}(\gamma) = N$ since U_h is unitary. Now γ_h and α_h belong to $\mathcal{K}_h(N)$ by definition, hence we have that $\mathcal{E}(\gamma_h, \alpha_h) \geq I_h(N)$. On the other hand, by the convergence of f_i^h (hence of g_i^h) towards f_i in $H^1(\mathbb{R}^3)$, we easily see that

$$\lim_{h\to 0} \mathcal{E}(\gamma_h, \alpha_h) = \mathcal{E}(\gamma, \alpha),$$

by continuity of $\mathcal{E}$. Therefore we have proved that

$$\limsup_{h\to 0} I_h(N) \leq \mathcal{E}(\gamma, \alpha).$$

This is valid for all finite-rank $(\gamma,\alpha)\in\mathcal{K}(N)$, hence we deduce that

$$\limsup_{h\to 0} I_h(N) \leq \inf_{(\gamma,\alpha)\in\mathcal{K}(N)} \mathcal{E}(\gamma,\alpha) = I(N).$$

On the other hand the inequality $I_h(N)\geq I(N)$ is always satisfied, hence we have proved the claimed convergence $I_h(N)\to I(N)$. □

Proof of Lemma 2.1.1. Let $(\psi_i)_i$ be a Hilbert basis of $L^2(\mathbb{R}^3)$ such that $\psi_i\in H^1$ for all i. We define the projector over the m first basis functions as $P_m=\sum_{i=1}^m |\psi_i\rangle\langle\psi_i|$. Considere the two operators γ_m and α_m defined by $\gamma_m = P_m\,\gamma\,P_m$ and $\alpha_m = P_m\,\alpha\,P_m$. The matrice Γ_m verifies the constraint (1.4.18):

$$0\leq \begin{pmatrix} P_m & 0\\ 0 & P_m\end{pmatrix}\begin{pmatrix}\gamma & \alpha\\ \alpha^* & 1-\overline{\gamma}\end{pmatrix}\begin{pmatrix} P_m & 0\\ 0 & P_m\end{pmatrix}\leq 1.$$

Of course γ_m and α_m are finite-rank operators since P_m has rank m. The trace is not bigger than m. Also we have $\gamma_m\to\gamma$ when $m\to\infty$ in the sense that $\|(1-\Delta)^{1/4}(\gamma-\gamma_m)(1-\Delta)^{1/4}\|_{\mathfrak{S}_1}\to 0$ and $(1-\Delta)^{1/2}(\alpha-\alpha_m)_{\mathfrak{S}_2}\to 0$. In particular, $\gamma_m\to\gamma$ in $\mathfrak{S}_1$ hence $\mathrm{Tr}(\gamma_m)\to\mathrm{Tr}(\gamma)$. Now we define

$$\tilde{\gamma}_m = \gamma_m \,+\, (\mathrm{Tr}(\gamma)-\mathrm{Tr}(\gamma_m))\,|\psi_{m+1}\rangle\langle\psi_{m+1}|.$$

So the finite-rank operator $\tilde{\gamma}_m$ verify the constraint $\mathrm{Tr}\tilde{\gamma}_m=\mathrm{Tr}\gamma$ and the result is proved. □

2.2 Discretization of the HFB model

In this section we compute the HFB energy $\mathcal{E}(\gamma,\alpha)$ of a discretized state $(\gamma,\alpha)\in\mathcal{K}_h(N)$ and we write the corresponding self-consistent equation. We fix once and for all the approximation space V_h, and we consider a basis set $(\chi_i)_{i=1}^{N_b}$ of V_h, which is not necessarily orthonormal. We will assume that V_h is stable under complex conjugation, which means that $f\in V_h\Rightarrow\overline{f}\in V_h$ (this amounts to replacing V_h by $\mathrm{Span}(V_h,\overline{V_h})$).

Since $\pi_h\gamma\pi_h=\gamma$ and $\pi_h\alpha\overline{\pi_h}=\alpha$, we can write the kernels of γ and α as follows

$$\gamma(x,y)_{\sigma,\sigma'}=\sum_{i,j=1}^{N_b} G_{ij}\,\chi_i(x)_\sigma\overline{\chi_j(y)_{\sigma'}},\qquad \alpha(x,y)_{\sigma,\sigma'}=\sum_{i,j=1}^{N_b} A_{ij}\,\chi_i(x)_\sigma\chi_j(y)_{\sigma'}. \tag{2.2.6}$$

In the same way, the density can be expressed as

$$\rho_\gamma(x)=\sum_{\sigma=1}^{q}\gamma(x,\sigma;x,\sigma)=\sum_{\sigma=1}^{q}\sum_{i,j=1}^{N_b} G_{ij}\chi_i(x)_\sigma\overline{\chi_j(x)_\sigma}=\sum_{i,j=1}^{N_b} G_{ij}\chi_j(x)^*\chi_i(x),$$

where $\chi_j(x)^*\chi_i(x)=\sum_{\sigma=1}^q\chi_i(x)_\sigma\overline{\chi_j(x)_\sigma}$.

The matrices G and A (defined by the previous relation) satisfy the constraints $G^* = G$ and $A^T = -A$. Note that since A is antisymmetric, we can also write

$$\begin{aligned}\alpha(x,y)_{\sigma,\sigma'} &= \sum_{1\le i<j\le N_b} A_{ij}\left(\chi_i(x)_\sigma\chi_j(y)_{\sigma'} - \chi_j(x)_\sigma\chi_i(y)_{\sigma'}\right)\\ &= \sqrt{2}\sum_{1\le i<j\le N_b} A_{ij}\,\chi_i\wedge\chi_j(x,\sigma,y,\sigma'),\end{aligned}$$

where

$$(\chi_i\wedge\chi_j)(x,\sigma;y,\sigma') := \left(\chi_i(x)_\sigma\chi_j(y)_{\sigma'} - \chi_i(y)_{\sigma'}\chi_j(x)_\sigma\right)/\sqrt{2}$$

is a two-body Slater determinant. Let us also remark that (2.2.6) can be written in the operator form

$$\gamma = \sum_{i,j=1}^{N_b} G_{ij}\,|\chi_i\rangle\langle\chi_j|, \qquad \alpha = \sum_{i,j=1}^{N_b} A_{ij}\,|\chi_i\rangle\langle\overline{\chi_j}|.$$

Let us define the so-called *overlap matrices* S and Σ by

$$S_{ij} := \langle\chi_i,\chi_j\rangle_{\mathfrak{H}} = \sum_{\sigma=1}^{q}\int_{\mathbb{R}^3}\overline{\chi_i(x)_\sigma}\chi_j(x)_\sigma\,dx \tag{2.2.7}$$

and

$$\Sigma_{ij} := \langle\overline{\chi_i},\chi_j\rangle_{\mathfrak{H}} = \sum_{\sigma=1}^{q}\int_{\mathbb{R}^3}\chi_i(x)_\sigma\chi_j(x)_\sigma\,dx. \tag{2.2.8}$$

Lemma 2.2.1. *The constraint*

$$\begin{pmatrix}0&0\\0&0\end{pmatrix} \le \Gamma := \begin{pmatrix}\gamma & \alpha\\ \alpha^* & 1-\overline{\gamma}\end{pmatrix} \le \begin{pmatrix}1&0\\0&1\end{pmatrix} \tag{2.2.9}$$

is equivalently written for the matrices G and A in the form

$$\begin{pmatrix}0&0\\0&0\end{pmatrix} \le \begin{pmatrix}SGS & SA\Sigma\\ \overline{\Sigma}A^*S & S-\overline{\Sigma}\,\overline{G}\Sigma\end{pmatrix} \le \begin{pmatrix}S&0\\0&S\end{pmatrix}, \tag{2.2.10}$$

or

$$\begin{pmatrix}0&0\\0&0\end{pmatrix} \le \begin{pmatrix}S&0\\0&\overline{\Sigma}\end{pmatrix}\begin{pmatrix}G & A\\ A^* & \overline{\Sigma}^{-1}S\Sigma^{-1}-\overline{G}\end{pmatrix}\begin{pmatrix}S&0\\0&\Sigma\end{pmatrix} \le \begin{pmatrix}S&0\\0&S\end{pmatrix}, \tag{2.2.11}$$

or

$$0 \le \Upsilon\mathbf{S}\Upsilon \le \Upsilon \tag{2.2.12}$$

where Υ and $\mathbf{S}$ are defined by

$$\Upsilon := \begin{pmatrix}G & A\\ A^* & \overline{\Sigma}^{-1}S\Sigma^{-1}-\overline{G}\end{pmatrix} \quad \textit{and} \quad \mathbf{S} := \begin{pmatrix}S&0\\0&\Sigma S^{-1}\overline{\Sigma}\end{pmatrix}. \tag{2.2.13}$$

Note that another way to write the constraint (2.2.12) is $0 \leq \mathbf{S}^{1/2}\Upsilon\mathbf{S}^{1/2} \leq 1$.

Proof. Let f and g be two functions in V_h, which are decomposed in the basis (χ_i) as follows:

$$f(x) = \sum_{i=1}^{N_b} x_i\chi_i(x), \qquad g(x) = \sum_{i=1}^{N_b} y_i\chi_i(x).$$

The constraint (2.2.9) can be written as

$$0 \leq \left\langle \begin{pmatrix} f \\ g \end{pmatrix}, \Gamma \begin{pmatrix} f \\ g \end{pmatrix} \right\rangle \leq \|f\|^2 + \|g\|^2. \tag{2.2.14}$$

That implies

$$0 \leq \langle f, \gamma f\rangle + \langle f, \alpha g\rangle + \langle g, \alpha^* f\rangle + \|g\|^2 - \langle g, \overline{\gamma} g\rangle \leq \|f\|^2 + \|g\|^2. \tag{2.2.15}$$

We compute every term of (2.2.15) and obtain

$$\begin{aligned} \langle f, \gamma f\rangle &= \sum_{ij} \overline{x_i} x_j \langle \chi_i, \gamma\chi_j\rangle \\ &= \sum_{ijk\ell} \overline{x_i} x_j G_{k\ell}\langle \chi_i, \chi_k\rangle\langle \chi_\ell, \chi_j\rangle \\ &= x^*(SGS)x, \end{aligned} \qquad \begin{aligned} \langle f, \alpha g\rangle &= \sum_{ij} \overline{x_i} y_j \langle \chi_i, \alpha\chi_j\rangle \\ &= \sum_{ijk\ell} \overline{x_i} y_j A_{k\ell}\langle \chi_i, \chi_k\rangle\langle \overline{\chi_\ell}, \chi_j\rangle \\ &= x^*(SA\Sigma)y, \end{aligned}$$

$$\begin{aligned} \langle g, \alpha^* f\rangle &= \sum_{ij} \overline{y_i} x_j \langle \chi_i, \alpha^*\chi_j\rangle, \\ &= \sum_{ijk\ell} \overline{y_i} x_j \overline{A_{k\ell}}\langle \chi_i, \overline{\chi_k}\rangle\langle \chi_\ell, \chi_j\rangle \\ &= y^*(\overline{\Sigma} A^* S)x, \end{aligned} \qquad \begin{aligned} \langle g, \overline{\gamma} g\rangle &= \sum_{ij} \overline{y_i} y_j \langle \chi_i, \overline{\gamma}\chi_j\rangle \\ &= \sum_{ijk\ell} \overline{y_i} y_j \overline{G_{k\ell}}\langle \chi_i, \overline{\chi_k}\rangle\langle \overline{\chi_\ell}, \chi_j\rangle \\ &= y^*(\overline{\Sigma}\,\overline{G}\Sigma)y. \end{aligned}$$

So we obtain the inequality

$$0 \leq x^*(SGS)x + x^*(SA\Sigma)y + y^*(\overline{\Sigma} A^* S)x + y^* S y - y^*(\overline{\Sigma}\ \overline{G}\Sigma)y \leq x^* S x + y^* S y.$$

The latter can be written in an equivalent way as follows:

$$0 \leq \begin{pmatrix} x \\ y \end{pmatrix}^* \begin{pmatrix} SGS & SA\Sigma \\ \overline{\Sigma} A^* S & S - \overline{\Sigma} G \Sigma \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} x \\ y \end{pmatrix}^* \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

This concludes the proof of the lemma. □

For the sake of simplicity, we will assume in the following that the basis (χ_i) is real, $\overline{\chi_i} = \chi_i \quad \forall i = 1, .., N_b$. Since $\overline{V_h} = V_h$, this is always possible. We then have

$$S = \overline{S} = \Sigma = \overline{\Sigma}$$

and

$$\Upsilon := \begin{pmatrix} G & A \\ A^* & S^{-1} - \overline{G} \end{pmatrix} \quad , \quad \mathbf{S} := \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix}. \tag{2.2.16}$$

We now compute the HFB energy.

Lemma 2.2.2. *We have*

$$\mathcal{E}(\gamma,\alpha) = \mathrm{Tr}(hG) + \frac{1}{2}\,\mathrm{Tr}(G\,J(G)) - \frac{1}{2}\,\mathrm{Tr}(G\,K(G)) + \frac{1}{2}\,\mathrm{Tr}(A^*\,X(A)). \quad (2.2.17)$$

where

$$h_{ij} = \langle \chi_i, (-\Delta)\chi_j\rangle = \sum_{\sigma=1}^{q} \int_{\mathbb{R}^3} \nabla\chi_i(x)_\sigma \cdot \nabla\chi_j(x)_\sigma\, dx,$$

$$J(G)_{ij} = \sum_{k,\ell=1}^{N_b} (ij|\ell k)\, G_{k\ell}, \quad K(G)_{ij} = \sum_{k,\ell=1}^{N_b} (ik|\ell j)\, G_{k\ell}, \quad X(A)_{ij} = \sum_{k,\ell=1}^{N_b} (ik|j\ell)\, A_{k\ell}, \quad (2.2.18)$$

and

$$(ij|k\ell) := \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\,\chi_i(x)^*\,\chi_j(x)\,\chi_k(y)^*\,\chi_\ell(y)\,dx\,dy. \quad (2.2.19)$$

We use here the notation $a^*b = \sum_{\sigma=1}^{q} \overline{a}_\sigma\, b_\sigma$. *Similarly, we have*

$$\mathrm{Tr}(\gamma) = \mathrm{Tr}(SG). \quad (2.2.20)$$

We recall that S is given by

$$S_{ij} = \langle \chi_i, \chi_j\rangle_{\mathfrak{h}}.$$

Remark 2.2.1. *The trace here is the usual one for* $N_b \times N_b$ *matrices. As we think that there is no possible confusion with* $\mathcal{E}(\gamma,\alpha)$, *we will also denote by* $\mathcal{E}(G,A)$ *this discretized energy functional.*

Remark 2.2.2. *Recall we use that* $\overline{\chi_i} = \chi_i$ *is real. We define the discretized number operator as*

$$\mathbf{N} = \begin{pmatrix} S & 0 \\ 0 & -S \end{pmatrix}. \quad (2.2.21)$$

The constraint $\mathrm{Tr}(\gamma) = N$ *can be written equivalently as*

$$\mathrm{Tr}(\mathbf{N}\Upsilon) = 2N - N_b.$$

Remark 2.2.3. *The formulas are the same when the basis* (χ_i) *is not real.*

Proof. The discretized kinetic energy is

$$\mathrm{Tr}(T\gamma) = \sum_{i,j=1}^{N_b} G_{ij}\mathrm{Tr}(T|\chi_i\rangle\langle\chi_j|) = \sum_{i,j=1}^{N_b} G_{ij}\,\langle\chi_j, T\chi_i\rangle = \mathrm{Tr}(G\,h),$$

where $h_{ji} := \langle \chi_j, T\chi_i\rangle = \displaystyle\int_{\mathbb{R}^3} \nabla\chi_j \cdot \nabla\chi_i$.

The direct term is expressed as

$$\begin{aligned}
&\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\rho_\gamma(x)\rho_\gamma(y)dxdy \\
&\quad= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\sum_{i,j=1}^{N_b}\sum_{\sigma=1}^{q} G_{ij}\overline{\chi_j(x)}_\sigma\chi_i(x)_\sigma \sum_{k,\ell=1}^{N_b}\sum_{\sigma'=1}^{q} G_{k\ell}\overline{\chi_\ell(y)}_{\sigma'}\chi_k(y)_{\sigma'}\,dx\,dy \\
&\quad= \sum_{i,j,k,\ell=1}^{N_b} G_{ij}G_{k\ell}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\chi_j(x)^*\chi_i(x)\chi_\ell(y)^*\chi_k(y)dxdy \\
&\quad= \sum_{i,j,k,\ell=1}^{N_b} G_{ij}\,G_{k\ell}\,(ji|\ell k),
\end{aligned}$$

where $(ji|k\ell) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\chi_j(x)^*\chi_i(x)\chi_k(y)^*\chi_\ell(y)dxdy$.

Recall that χ_i is a spinor that is a function of $L^2(\mathbb{R}^3,\mathbb{C}^q)$. Note that

$$\begin{aligned}
|\gamma(x,y)|^2 &= \mathrm{Tr}_{\mathbb{C}^q}\gamma(x,y)^*\gamma(x,y) \\
&= \sum_{\sigma,\sigma'=1}^{q} \overline{\gamma(x,y)}_{\sigma',\sigma}\gamma(x,y)_{\sigma',\sigma} \\
&= \sum_{\sigma,\sigma'=1}^{q} |\gamma(x,y)_{\sigma,\sigma'}|^2 \\
&= \sum_{\sigma,\sigma'=1}^{q}\sum_{i,j,k,\ell=1}^{N_b} \overline{G_{ij}}\,\overline{\chi_i(x)}_{\sigma'}\chi_j(y)_\sigma G_{k\ell}\chi_k(x)_{\sigma'}\overline{\chi_\ell(y)}_\sigma \\
&= \sum_{ijk\ell} \overline{G_{ij}}G_{k\ell}\chi_i(x)^*\chi_k(x)\chi_\ell(y)^*\chi_j(y).
\end{aligned}$$

So for the exchange term we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)|\gamma(x,y)|^2dxdy \\
&\qquad= \sum_{ijk\ell} \overline{G_{ij}}G_{k\ell}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\chi_i(x)^*\chi_k(x)\chi_\ell(y)^*\chi_j(y)dxdy \\
&\qquad= \sum_{ijk\ell} \overline{G_{ij}}\,G_{k\ell}\,(ik|\ell j).
\end{aligned}$$

The same computation for the pairing term gives

$$\begin{aligned}|\alpha(x,y)|^2 &= \sum_{\sigma,\sigma'=1}^{q} |\alpha(x,y)_{\sigma,\sigma'}|^2 \\ &= \sum_{\sigma,\sigma'=1}^{q} \sum_{i,j,k,\ell=1}^{N_b} \overline{A_{ij}}\, \overline{\chi_i(x)}_\sigma\, \overline{\chi_j(y)}_{\sigma'} A_{k\ell}\chi_k(x)_\sigma \chi_\ell(y)_{\sigma'} \\ &= \sum_{i,j,k,\ell=1}^{N_b} \overline{A_{ij}}\, A_{k\ell}\, \chi_i(x)^*\, \chi_k(x)\, \chi_j(y)^*\, \chi_\ell(y).\end{aligned}$$

In conclusion we obtain

$$\begin{aligned}&\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)|\alpha(x,y)|^2 dxdy \\ &\qquad = \sum_{ijk\ell} \overline{A_{ij}} A_{k\ell} \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} W(x-y)\chi_i(x)^*\, \chi_k(x)\, \chi_j(y)^*\, \chi_\ell(y) dxdy \\ &\qquad = \sum_{ijk\ell} \overline{A_{ij}} A_{k\ell} (ik|j\ell).\end{aligned}$$

This ends the proof of Lemma 2.2.2. □

We deduce from the previous result that the variational problem $I_h(N)$ can be written in finite dimension in a real basis $\overline{\chi_i} = \chi_i$ as

$$I_h(N) = \min\Big\{\mathcal{E}(G,A) \ :\ 0 \le \Upsilon\mathbf{S}\Upsilon \le \Upsilon,\ \mathrm{Tr}(\mathbf{N}\Upsilon) = 2N - N_b\Big\}, \qquad (2.2.22)$$

where we recall that Υ and $\mathbf{S}$ have been defined in (2.2.13) and that $\mathcal{E}(G,A)$ is given by the formula (2.2.17). Here the infimum is always attained because the problem is finite dimensional. More precisely, the energy is clearly continuous with respect to (G,A), and the minimization set is bounded and closed, hence compact.

In this form, the discretized problem is very similar to the usual discretized Hartree-Fock problem [5, 25], in dimension $2N_b$ instead of N_b. There is a big difference, however. In HF theory the constraint involves the matrix $\mathbf{S}$ instead of $\mathbf{N}$. This difference will cause several difficulties. To understand the problem, let us introduce a new variable $\Upsilon' = \mathbf{S}^{1/2}\Upsilon\mathbf{S}^{1/2}$ (which is the same as orthonormalizing the basis (χ_i)). Then the constraint $0 \le \Upsilon\mathbf{S}\Upsilon \le \Upsilon$ is transformed into $0 \le \Upsilon' \le 1$. However, the constraint on the number of particles becomes $\mathrm{Tr}(\mathbf{S}^{-1/2}\mathbf{N}\mathbf{S}^{-1/2}\Upsilon') = 2N - N_b$. In usual Hartree-Fock theory, the matrix $\mathbf{S}^{-1/2}\mathbf{N}\mathbf{S}^{-1/2}$ is replaced by the identity. The fact that this matrix then commutes with the Fock Hamiltonian (defined below) simplifies dramatically the self-consistent equations. Here we get

$$\mathbf{S}^{-1/2}\mathbf{N}\mathbf{S}^{-1/2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which commutes with the Fock Hamiltonian if and only if $A \equiv 0$.

The self-consistent equation is obtained like in [3]. The result is as follows.

Lemma 2.2.3 (Discretized HFB equation)**.** *Assume that the basis is real, $\overline{\chi_i} = \chi_i$. Let Υ be a minimizer for the discretized variational problem $I_h(N)$. Then there exists $\mu \in \mathbb{R}$ such that Υ solves the linear problem*

$$\min\Big\{ \mathrm{Tr}(\mathbf{F}_\Upsilon - \mu\mathbf{N})\tilde{\Upsilon} \ : \ 0 \leq \tilde{\Upsilon}\mathbf{S}\tilde{\Upsilon} \leq \tilde{\Upsilon}\Big\}, \tag{2.2.23}$$

where

$$\mathbf{F}_\Upsilon := \begin{pmatrix} h_G & X(A) \\ X(A)^* & -\overline{h_G} \end{pmatrix}, \qquad h_G := h + J(G) - K(G). \tag{2.2.24}$$

The solution Υ can be written in the form

$$\Upsilon = \mathbf{S}^{-1/2}\Big(\chi_{(-\infty,0)}\Big(\mathbf{S}^{-1/2}(\mathbf{F}_\Upsilon - \mu\mathbf{N})\mathbf{S}^{-1/2}\Big) + \delta\Big)\mathbf{S}^{-1/2} \tag{2.2.25}$$

where $0 \leq \delta \leq 1$ lives only in the kernel of $\mathbf{S}^{-1/2}(\mathbf{F}_\Upsilon - \mu\mathbf{N})\mathbf{S}^{-1/2}$.

The solution Υ of the self-consistent equation (2.2.25) may be equivalently written by considering the generalized eigenvalue problem

$$\big(\mathbf{F}_\Upsilon - \mu\mathbf{N}\big) f_i = \epsilon_i \, \mathbf{S} \, f_i, \qquad \langle f_i, \mathbf{S} f_j \rangle = \delta_{ij}. \tag{2.2.26}$$

Then we have simply (assuming $\epsilon_i \neq 0$ for all $i = 1, ..., 2N_b$)

$$\Upsilon = \sum_{\epsilon_i < 0} f_i \, f_i^*.$$

Again, this is similar to the Hartree-Fock solution [5,25] except that μ is unknown and $\mathbf{N}$ does not always commute with $\mathbf{F}_\Upsilon$.

Remark 2.2.4. *Let us recall that the Hartree-Fock discretized problem [8] is*

$$\inf\{\mathcal{E}_{HF}(G), 0 \leq GSG \leq G, \ \mathrm{Tr}SG = N\}$$

where the Hartree-Fock energy is

$$\mathcal{E}_{HF}(G) = \mathcal{E}(G, 0) = \mathrm{Tr}(hG) + \frac{1}{2}\mathrm{Tr}(J(G)G) - \frac{1}{2}\mathrm{Tr}(X(G)G).$$

Here S, h J, X and G are exactly as before. The solution can be written in the form

$$G = C\,C^* = \sum_{i=1}^{N} C_i \, C_i^*,$$

where $C = (C_1, .., C_N)$ and $\left\{ \begin{array}{rcl} h_G\, C & = & S\, C\, \Lambda \\ C^*\, S\, C & = & I_N \\ G & = & C\, C^* \end{array} \right.$

with Λ, the matrix of Lagrange multipliers can be diagonalized and the composants of Λ are the lowest N eigenvalues of the Fock operator h_G:

$$h_G\, C_i = \lambda_i \, S\, C_i.$$

Remark that although the basis functions χ_i are real, the density matrix Υ is not necessarily real. In the next section, we will restrict ourselves to real-valued density matrices and impose some spin symmetry.

2.3 Using symmetries

The HFB energy $\mathcal{E}(\Gamma)$ has some natural symmetry invariances which we describe in detail in this section. Recall that since $\mathcal{E}$ is a *nonlinear* functional, it cannot be guaranteed that the HFB minimizers will all have the same symmetries as the HFB energy. The set of all minimizers will be invariant under the action of the symmetry group, but each minimizer alone does not have to be invariant.

We have already allowed for the breaking of particle-number symmetry and we hope to find an HFB ground state. It will then automatically break the translational invariance of the system. There are three other symmetries (spin, complex conjugation and rotations) in the system, which are of interest to us. We have the choice of imposing these symmetries by adding appropriate constraints, or not. Because this reduces the computational cost, it will be convenient to impose them.

2.3.1 Time-reversal symmetry

Let us now assume that $q = 2$, which means that our fermions are spin-1/2 particles. Since the Laplacian and the interaction function W do not act on the spin variable, the HFB energy has some *spin symmetry*, which can be written for $q = 2$ as

$$\boxed{\forall k = 1, 2, 3, \qquad \mathcal{E}(\Sigma_k \Gamma \Sigma_k^*) = \mathcal{E}(\Gamma)}$$

where

$$\Sigma_k := \begin{pmatrix} i\sigma_k & 0 \\ 0 & \overline{i\sigma_k} \end{pmatrix},$$

with $\sigma_1, \sigma_2, \sigma_3$ being the usual Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.3.27}$$

Note that Σ_k has the form of a Bogoliubov transformation, hence $\Sigma_k \Gamma \Sigma_k^*$ is also an HFB state. The number operator is also invariant which means that

$$\Sigma_k \mathcal{N} \Sigma_k^* = \mathcal{N}$$

for all $k = 1, 2, 3$. Thus, the contraint $\mathrm{Tr}(\gamma) = N$ is conserved and we have $\Sigma_k \Gamma \Sigma_k^* \in \mathcal{K}(N)$ when $\Gamma \in \mathcal{K}(N)$.

Another important symmetry is that of *complex conjugation* which means this time that

$$\boxed{\mathcal{E}(\overline{\Gamma}) = \mathcal{E}(\Gamma)}$$

and which is based on the fact that the Laplacian and W are real operators. Again we have $\mathrm{Tr}(\overline{\gamma}) = \mathrm{Tr}(\gamma)$ hence $\mathcal{K}(N)$ is invariant under complex conjugation.

As was explained in [20] (see Remark 5 page 1032), the density matrices (γ, α) can be written in the special form

$$\gamma' = g \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \alpha' = a \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \text{with } g = g^T = \overline{g} \text{ and } a = a^T = \overline{a} \tag{2.3.28}$$

(the 2×2 matrix refers to the spin variables), if and only if

$$\begin{cases} \Sigma_k \Gamma \Sigma_k^* = \Gamma, & \text{for } k = 1, 2, \text{ and} \\ \overline{\Gamma} = \Gamma. \end{cases} \tag{2.3.29}$$

In other words, Γ is invariant under the action of the group generated by Σ_1, Σ_2 and complex conjugation. This invariance is sometimes called the *time-reversal symmetry*. As remarked in [20], imposing $\Sigma_k \Gamma \Sigma_k^* = \Gamma$ for all $k = 1, 2, 3$ (instead of just $k = 1, 2$) implies $\alpha \equiv 0$ which is not interesting for us.

When W can be written in the form

$$W(x - y) = -\int_\Omega d\mu(\omega) f_\omega(x)\, f_\omega(y), \tag{2.3.30}$$

the Theorem 1.4.2 of Bach, Fröhlich and Jonsson [2], tells us that there is no breaking of the time-reversal symmetry for minimizers. That is, we can always minimize over such special states. For other interactions W this is not necessarily true but it is often convenient to impose this symmetry anyhow.

Because it holds

$$F_{\Sigma_k \Gamma \Sigma_k^*} = \Sigma_k F_\Gamma \Sigma_k^*, \quad F_{\overline{\Gamma}} = \overline{F_\Gamma},$$

it can then be verified that minimizers under the additional symmetry constraint, satisfy the same self-consistent equation as when no constraint is imposed.

When we discretize the problem by imposing time-reversal symmetry, we use two real symmetric matrices G and A, related through the constraint that

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \leq \Upsilon := \begin{pmatrix} G & A \\ A & 1 - G \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.3.31}$$

The energy becomes

$$\mathcal{E}(G, A) = 2\,\mathrm{Tr}(hG) + 2\,\mathrm{Tr}(G\,J(G)) - \mathrm{Tr}(G\,K(G)) + \mathrm{Tr}(A\,X(A)) \tag{2.3.32}$$

and the associated particle number constraint is $\mathrm{Tr}(SG) = N/2$. In practice we always assume that N is even for simplicity. The basis (χ_i) is now composed of (real-valued) functions in $H^1(\mathbb{R}^3, \mathbb{R})$, instead of functions in $H^1(\mathbb{R}^3, \mathbb{C}^2)$ as before, and the formulas for S, h, J, K and X are the same as before.

2.3.2 Rotational symmetry

The group $SO(3)$ of rotations in $\mathbb{R}^3$ also acts on HFB states and it leaves the energy invariant when the interaction W is a radial function. In this section we always assume that the spin variable has already been removed according to the previous section and we denote by g and a the corresponding (real symmetric) density matrices. If the spin were still present, rotations would act on it as well.

To any rotation $R \in SO(3)$ we can associate a unitary operator on $L^2(\mathbb{R}^3)$, denoted also by R and defined by $(R\phi)(x) = \phi\big(R^{-1}x\big)$. Now we say that an HFB state Γ with density matrices (γ, α) is invariant under rotations when it satisfies

$$\boxed{\mathscr{R}\,\Gamma\mathscr{R}^* = \Gamma, \qquad \text{where} \quad \mathscr{R} = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix} \qquad \forall R \in SO(3).}$$

Note that $\mathscr{R}$ is a Bogoliubov rotation since $\overline{R} = R$. The density matrices of an invariant state satisfy

$$g(Rx, Ry) = g(x, y), \qquad a(Rx, Ry) = a(x, y)$$

for all $x, y \in \mathbb{R}^3$ and any rotation $R \in SO(3)$.

As the angular momentum $L = x \times (-i\nabla)$ generates the group of rotations, a (smooth enough) HFB state is invariant under rotations if and only if

$$\boxed{\mathscr{L}\Gamma = \Gamma\mathscr{L}, \qquad \text{where} \quad \mathscr{L} = \begin{pmatrix} L & 0 \\ 0 & L \end{pmatrix}.} \tag{2.3.33}$$

Let us now give the special form of rotation-invariant states. We recall that the spherical harmonics $(Y_\ell^m)_{\ell\geq 0, -\ell\leq m\leq \ell}$ is a Hilbert basis of $L^2(\mathbb{S}^2)$, where $\mathbb{S}^2$ is the sphere of radius 1 in dimension 3. We have for all (ℓ, m), such that $\ell \geq 0$, $-\ell \leq m \leq \ell$ and for all (ℓ', m'), with $\ell' \geq 0$, $-\ell' \leq m' \leq \ell'$,

$$(Y_\ell^m, Y_{\ell'}^{m'})_{L^2(\mathbb{S}^2)} = \int_0^\pi \int_0^{2\pi} Y_\ell^m(\theta,\phi)^* Y_{\ell'}^{m'}(\theta,\phi)\sin(\theta) d\theta d\phi = \delta_{\ell\ell'}\delta_{mm'}. \tag{2.3.34}$$

Now we choose a basis of $L^2(\mathbb{R}^3)$ of the form

$$\Psi_{i\ell m} = \chi_i(r) Y_\ell^m(\theta,\phi) \quad \text{ with } i \geq 1, m = -\ell, .., \ell \text{ and } \ell \geq 0.$$

Every basis functions can be decomposed as a radial function and a spherical harmonics. Here (χ_i) is a Hilbert basis of $L^2([0,\infty), r^2\,dr)$. The operators g and a can be written in this basis,

$$g(x,y) = \sum_{i,j\geq 1} \sum_{\ell,\ell'\geq 0} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} g_{ij}^{\ell\ell' mm'} |\Psi_{i\ell m}\rangle\langle\Psi_{j\ell' m'}|,$$

$$a(x,y) = \sum_{i,j\geq 1} \sum_{\ell,\ell'\geq 0} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} a_{ij}^{\ell\ell' mm'} |\Psi_{i\ell m}\rangle\langle\Psi_{j\ell' m'|}.$$

By (2.3.33), we know that

$$L^2 g = g L^2, \qquad L_z g = g L_z \qquad \text{and} \qquad L_\pm g = g L_\mp, \tag{2.3.35}$$

where L is the angular momentum vector such that $L = (L_x, L_y, L_z)$, $L^2 = L_x^2 + L_y^2 + L_z^2$ and $L_\pm = L_x \pm i L_y$.

The spherical harmonics Y_ℓ^m are the eigenfunctions of the operators build upon L and they satisfy the following relations:

$$L^2 Y_\ell^m = \ell(\ell{+}1) Y_\ell^m, \quad L_z Y_\ell^m = m Y_\ell^m \quad \text{and} \quad L_\pm Y_\ell^m = \sqrt{(\ell \mp m)(\ell \pm m + 1)}\, Y_\ell^{m\pm 1}. \tag{2.3.36}$$

Lemma 2.3.1. *We have*

$$g_{ij}^{\ell\ell' mm'} = g_{ij}^{\ell} \qquad \forall i,j \geq 1, \ell, \ell' \geq 0, m = -\ell, .., \ell \text{ and } m' = -\ell', .., \ell'.$$

Proof. Writing that $L^2\, g = g\, L^2$, we obtain

$$\sum_{i,j\geq 1} \sum_{\ell,\ell'\geq 0} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} \ell(\ell+1) g_{ij}^{\ell\ell' mm'} |\Psi_{i\ell m}\rangle\langle \Psi_{j\ell' m'}|$$

$$= \sum_{i,j\geq 1} \sum_{\ell,\ell'\geq 0} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} \ell'(\ell'+1) g_{ij}^{\ell\ell' mm'} |\Psi_{i\ell m}\rangle\langle \Psi_{j\ell' m'}|.$$

By uniqueness of the decomposition, we deduce that

$$\ell(\ell+1) g_{ij}^{\ell\ell' mm'} = \ell'(\ell'+1) g_{ij}^{\ell\ell' mm'} \qquad \forall i,j,\ell,\ell',m,m'.$$

So $g_{ij}^{\ell\ell' mm'} = 0$, when $\ell \neq \ell'$, and therefore g depends only of ℓ.

On the other hand, using that $L_z\, g = g\, L_z$ with the azimutal angular momentum L_z, we obtain:

$$\sum_{i,j\geq 1} \sum_{\ell\geq 0} \sum_{m,m'=-\ell}^{\ell} m g_{ij}^{\ell mm'} |\Psi_{i\ell m}\rangle\langle \Psi_{j\ell m'}| = \sum_{i,j\geq 1} \sum_{\ell\geq 0} \sum_{m,m'=-\ell}^{\ell} m' g_{ij}^{\ell mm'} |\Psi_{i\ell m}\rangle\langle \Psi_{j\ell m'}|,$$

hence

$$m\, g_{ij}^{\ell mm'} = m'\, g_{ij}^{\ell mm'} \qquad \forall i,j,\ell,m,m'.$$

As before, we find that $g_{ij}^{\ell mm'} = 0$, when $m \neq m'$. So g depends only of m.

Finally we apply the property $L_+\, g = g\, L_-$ of the ladder operator L_+ and L_-, and we find the following equation

$$\sum_{i,j\geq 1} \sum_{\ell\geq 0} \sum_{m=-\ell}^{\ell} \sqrt{(\ell-m)(\ell+m+1)}\, g_{ij}^{\ell m+1} |\Psi_{i\ell m+1}\rangle\langle \Psi_{j\ell m}|$$

$$= \sum_{i,j\geq 1} \sum_{\ell\geq 0} \sum_{m=-\ell}^{\ell} \sqrt{(\ell+m)(\ell-m+1)}\, g_{ij}^{\ell m-1} |\Psi_{i\ell m}\rangle\langle \Psi_{j\ell m-1}|,$$

$$= \sum_{i,j\geq 1} \sum_{\ell\geq 0} \sum_{m=-\ell-1}^{\ell-1} \sqrt{(\ell-m)(\ell+m+1)}\, g_{ij}^{\ell m-1} |\Psi_{i\ell m+1}\rangle\langle \Psi_{j\ell m}|.$$

Therefore

$$\sqrt{(\ell-m)(\ell+m+1)}g_{ij}^{\ell m} = \sqrt{(\ell+m)(\ell-m+1)}g_{ij}^{\ell m} \tag{2.3.37}$$

because $gL_+ = L_-g$. We see that $g_{ij}^{\ell m} = 0$ when $m \neq -m$. This means that g is independent of m, and the statement is proved for g. The same computation can be done for a. □

By Lemma (2.3.1), we conclude that the density matrices g and a can then be written in the special form

$$g(x,y) = \sum_{i,j\geq 1}\sum_{\ell\geq 0}\sum_{m=-\ell}^{\ell} g_{ij}^{\ell}\,\chi_i(r)\,\chi_j(r')\,Y_\ell^m(\theta,\phi)\,Y_\ell^m(\theta',\phi'), \tag{2.3.38}$$

$$a(x,y) = \sum_{i,j\geq 1}\sum_{\ell\geq 0}\sum_{m=-\ell}^{\ell} a_{ij}^{\ell}\,\chi_i(r)\,\chi_j(r')\,Y_\ell^m(\theta,\phi)\,Y_\ell^m(\theta',\phi'). \tag{2.3.39}$$

We now use that

$$\sum_{m=-\ell}^{\ell} Y_\ell^m(\theta,\phi)Y_\ell^m(\theta',\phi') = \frac{2\ell+1}{4\pi}P_\ell(\omega\cdot\omega'),$$

where P_ℓ is the Legendre polynomial of degree ℓ, which is such that $P_\ell(1) = 1$. We get

$$\boxed{g(x,y) = \frac{1}{4\pi}\sum_{\ell\geq 0} g^\ell(|x|,|y|)\,(2\ell+1)\,P_\ell\big(\omega_x\cdot\omega_y\big),} \tag{2.3.40}$$

$$\boxed{a(x,y) = \frac{1}{4\pi}\sum_{\ell\geq 0} a^\ell(|x|,|y|)\,(2\ell+1)\,P_\ell\big(\omega_x\cdot\omega_y\big),} \tag{2.3.41}$$

where $g^\ell(|x|,|y|) = g_{ij}^\ell\,\chi_i(|x|)\,\chi_j(|y|)$ and $a^\ell(|x|,|y|) = a_{ij}^\ell\,\chi_i(|x|)\,\chi_j(|y|)$. The constraint on g and a are transfered in each angular momentum sector (labelled by $\ell \geq 0$), leading to

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \leq \begin{pmatrix} g^\ell & a^\ell \\ a^\ell & 1-g^\ell \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

on $L^2([0,\infty), r^2\,dr) \oplus L^2([0,\infty), r^2\,dr)$. However, there is no such constraint between different ℓ's. The total average particle number is given by

$$\mathrm{Tr}(g) = \sum_{\ell\geq 0}(2\ell+1)\,\mathrm{Tr}(g^\ell) = N/2.$$

This is the only constraint which mixes the different angular momentum density matrices.

2.3.3 Discretization of the radial HFB problem

Now we can discretize the radial HFB problem. We choose a finite-dimensional subspace $V_{\rm rad}$ in $L^2([0,\infty), r^2dr)$ with basis $(\chi_1, ..., \chi_{N_b})$, which we use to expand the density matrices g^ℓ and a^ℓ. This is the same as truncating the sums in the previous calculations. Then we fix a maximal angular momentum $\ell_{\max}$ and we truncate the series in (2.3.40) and (2.3.41). This is the same as taking as discretization space

$$V = \Big\{ f(|x|)\, Y_\ell^m(\omega_x) \ : \ f \in V_{\rm rad},\ 0 \le \ell \le \ell_{\max},\ -\ell \le m \le \ell \Big\} \subset L^2(\mathbb{R}^3, \mathbb{R})$$

where Y_ℓ^m is the spherical harmonics of total angular momentum ℓ and azimuthal angular momentum m. We then assume that g and a are radial and live in this space. The matrices G^ℓ and A^ℓ of g^ℓ and a^ℓ in the basis (χ_i) are defined similarly as before by

$$g^\ell(r,r') = \sum_{i,j=1}^{N_b} G^\ell_{ij}\, \chi_i(r)\chi_j(r'), \qquad a^\ell(r,r') = \sum_{i,j=1}^{N_b} A^\ell_{ij}\, \chi_i(r)\chi_j(r'). \tag{2.3.42}$$

The constraints on the matrices G^ℓ and A^ℓ are

$$0 \le \Upsilon^\ell \mathbf{S} \Upsilon^\ell \le \Upsilon^\ell, \quad \text{with} \quad \Upsilon^\ell := \begin{pmatrix} G^\ell & A^\ell \\ A^\ell & S^{-1} - G^\ell \end{pmatrix} \quad \text{and} \quad \mathbf{S} = \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix} \tag{2.3.43}$$

with

$$S_{ij} = \int_0^\infty \chi_i(r)\, \chi_j(r)\, r^2\, dr$$

and

$$\boxed{\sum_{\ell=0}^{\ell_{\max}} (2\ell+1)\, \mathrm{Tr}(S G^\ell) = N/2.} \tag{2.3.44}$$

Lemma 2.3.2 (Rotation-invariant discretized HFB energy)**.** *The total HFB energy is*

$$\mathcal{E}(G^0, ..., G^{\ell_{\max}}, A^0, ..., A^{\ell_{\max}}) = 2 \sum_{\ell=0}^{\ell_{\max}} (2\ell+1)\, \mathrm{Tr}(h^\ell G^\ell)$$
$$+ \sum_{\ell,\ell'=0}^{\ell_{\max}} (2\ell+1)(2\ell'+1) \Big(2\, \mathrm{Tr}(G^\ell\, J(G^{\ell'})) - \mathrm{Tr}(G^\ell\, K^{\ell\ell'}(G^{\ell'})) + \mathrm{Tr}(A^\ell\, K^{\ell\ell'}(A^{\ell'})) \Big), \tag{2.3.45}$$

where

$$h^\ell_{ij} = \int_0^\infty \chi_i'(r)\chi_j'(r) r^2 dr + \ell(\ell+1) \int_0^\infty \chi_i(r)\chi_j(r) dr,$$

$$J(G^{\ell'})_{ij} := \sum_{m,n=0}^{N_b} (ij|nm)_{0,0}\, G^{\ell'}_{mn}, \qquad K^{\ell\ell'}(G^{\ell'})_{ij} := \sum_{m,n=0}^{N_b} (im|jn)_{\ell,\ell'}\, G^{\ell'}_{mn},$$

$$(ij|mn)_{\ell,\ell'} := \int_0^\infty r^2\, dr \int_0^\infty s^2\, ds\; \chi_i(r)\, \chi_j(r)\, \chi_m(s)\, \chi_n(s)\, w_{\ell,\ell'}(r,s)$$

and

$$w_{\ell,\ell'}(r,s) := \frac{1}{2}\int_{-1}^{1} W\left(\sqrt{r^2+s^2-2rst}\right)\, P_\ell(t)\, P_{\ell'}(t)\, dt.$$

Any minimizer $(G^0, ..., G^{\ell_{\max}}, A^0, ..., A^{\ell_{\max}})$ of $\mathcal{E}$ under the constraints (2.3.43) and (2.3.44) is of the form

$$\Upsilon^\ell = \sum_{\epsilon_i^\ell<0} f_i^\ell \big(f_i^\ell\big)^T, \qquad 0 \le \ell \le \ell_{\max},$$

where the f_i^ℓ solve the generalized eigenvalue problem

$$\big(\mathbf{F}^\ell - \mu \mathbf{N}\big) f_i^\ell = \epsilon_i^\ell\, \mathbf{S}\, f_i^\ell, \qquad \langle f_i, \mathbf{S} f_j\rangle = \delta_{ij} \tag{2.3.46}$$

with

$$\mathbf{F}^\ell := \begin{pmatrix} h^\ell & 0 \\ 0 & -h^\ell \end{pmatrix} + \sum_{\ell'=0}^{\ell_{\max}} \begin{pmatrix} 2J\big(G^{\ell'}\big) - K^{\ell\ell'}\big(G^{\ell'}\big) & K^{\ell\ell'}\big(A^{\ell'}\big) \\ K^{\ell\ell'}\big(A^{\ell'}\big) & -2J\big(G^{\ell'}\big) + K^{\ell\ell'}\big(G^{\ell'}\big) \end{pmatrix}.$$

The Euler-Lagrange multipler μ appearing in (2.3.46) is common to all the different angular momentum sectors and it is chosen to ensure the validity of the constraint (2.3.44).

Proof. The kinetic energy can be computed as follows

$$\begin{aligned}
&\mathrm{Tr}(hg) = \mathrm{Tr}\Big((-\Delta)\sum_{\ell=0}^{\ell_{\max}} g_\ell\Big) = \sum_{\ell=0}^{\ell_{\max}} \mathrm{Tr}(-\Delta g_\ell) \\
&= \sum_{ij=1}^{N_b}\sum_{\ell=0}^{\ell_{\max}}\sum_{m=-\ell}^{\ell} g_{ij}^\ell \mathrm{Tr}(-\Delta|\Psi_{i\ell m}\rangle\langle\Psi_{j\ell m}|) \\
&= \sum_{ij=1}^{N_b}\sum_{\ell=0}^{\ell_{\max}}\sum_{m=-\ell}^{\ell} g_{ij}^\ell \mathrm{Tr}(-\Delta|\chi_i Y_\ell^m\rangle\langle\chi_i Y_\ell^m|) \\
&= \sum_{ij=1}^{N_b}\sum_{\ell=0}^{\ell_{\max}}\sum_{m=-\ell}^{\ell} g_{ij}^\ell \mathrm{Tr}((-\frac{1}{r^2}\partial_r(r^2\partial_r) + \frac{L^2}{r^2})|\chi_i Y_\ell^m\rangle\langle\chi_i Y_\ell^m|) \\
&= \sum_{ij=1}^{N_b}\sum_{\ell=0}^{\ell_{\max}}\sum_{m=-\ell}^{\ell} g_{ij}^\ell \left(\int_0^\infty -\frac{1}{r^2}\left(\frac{\partial}{\partial r}\left(r^2\frac{\partial\chi_i}{\partial r}\right)\right)\chi_j r^2 dr + \ell(\ell+1)\int_0^\infty \chi_i(r)\chi_j(r)dr\right) \\
&= \sum_{ij=1}^{N_b}\sum_{\ell=0}^{\ell_{\max}}(2\ell+1) g_{ij}^\ell \left(\int_0^\infty \chi_i'(r)\chi_j'(r) r^2 dr + \ell(\ell+1)\int_0^\infty \chi_i(r)\chi_j(r)dr\right).
\end{aligned}$$

We recall that the spherical harmonics are normalized on the sphere S^2 as follows

$$\int_{0,0}^{\pi,2\pi} |Y_\ell^m(\theta,\varphi)|^2 \sin\theta d\theta d\varphi = 1.$$

Furthermore, we have used that the spherical harmonics Y_ℓ^m are the eigenfunctions of $|L|^2$, see (2.3.36). If we introduce the kinetic energy matrix

$$h_{ij}^\ell = \int_0^\infty \chi_i'(r)\chi_j'(r)r^2dr + \ell(\ell+1)\int_0^\infty \chi_i(r)\chi_j(r)dr,$$

then we can express the kinetic energy as

$$\mathrm{Tr}(-\Delta)\gamma = \sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\sum_{ij=1}^{Nb} h_{ij}^\ell g_{ij}^\ell = \sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\mathrm{Tr}(h^\ell G^\ell).$$

As for the direct term, we compute

$$\begin{aligned}
&D(\rho_g,\rho_g)\\
&\quad= \iint_{\mathbb{R}^3\times\mathbb{R}^3} \rho_g(x)\rho_g(y)W(x-y)dxdy\\
&\quad= \iint_{\mathbb{R}^3\times\mathbb{R}^3} \sum_{\ell,\ell'=0}^{\ell_{\max}} g^\ell(x,x)\, g^{\ell'}(y,y)W(x-y)dxdy\\
&\quad= \sum_{i,j,m,n=1}^{N_b}\sum_{\ell,\ell'=0}^{\ell_{\max}} \frac{2\ell+1}{4\pi}\frac{2\ell'+1}{4\pi} g_{ij}^\ell g_{mn}^{\ell'} \times\\
&\qquad\qquad \times \iint_{\mathbb{R}^3\times\mathbb{R}^3} \chi_i(|x|)\chi_j(|x|)P_\ell(\omega\omega)\chi_m(|y|)\chi_n(|y|)P_{\ell'}(\omega'\omega')W(|x-y|)dxdy.
\end{aligned}$$

and we know that $P_\ell(\omega\omega) = P_\ell(1) = 1$. Similarly, $P_{\ell'}(\omega'\omega') = 1$. So we obtain

$$\begin{aligned}
D(\rho_g,\rho_g) &= \sum_{ijmn=1}^{N_b}\sum_{\ell,\ell'=0}^{\ell_{\max}} \frac{2\ell+1}{4\pi}\frac{2\ell'+1}{4\pi} g_{ij}^\ell g_{mn}^{\ell'} \times\\
&\qquad \times \iint_{\mathbb{R}^3\times\mathbb{R}^3} \chi_i(|x|)\chi_j(|x|)\chi_m(|y|)\chi_n(|y|)W(|x-y|)dxdy\\
&= \sum_{\ell,\ell'=0}^{\ell_{\max}}(2\ell+1)(2\ell'+1)\sum_{ijmn=1}^{N_b} G_{ij}^\ell G_{mn}^{\ell'}(ij,nm)_{00},
\end{aligned}$$

where

$$(ij,nm)_{00} := \frac{1}{(4\pi)^2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} \chi_i(|x|)\chi_j(|x|)\chi_m(|y|)\chi_n(|y|)W(|x-y|)dxdy.$$

Changing variables

$$\begin{aligned}|x-y|^2 &= |x|^2 + |y|^2 - 2x\cdot y \\ &= r^2 + r'^2 - 2rr'\omega\omega',\end{aligned}$$

we find for the double integral

$$\begin{aligned}&(ij,nm)_{00} \\ &= \frac{1}{(4\pi)^2}\int_0^\infty r^2dr\int_0^\infty s^2ds\int_{\mathbb{S}^2}d\omega\int_{\mathbb{S}^2}d\omega' W(\sqrt{r^2+s^2-2rs\omega\omega'})\chi_i(r)\chi_j(r)\chi_m(s)\chi_n(s).\end{aligned}$$

We let $t=\omega\cdot\omega'$ and we finally obtain

$$(ij,nm)_{00} = \frac{1}{2}\int_0^\infty r^2dr\int_0^\infty s^2ds\chi_i(r)\chi_j(r)\chi_m(s)\chi_n(s)\int_{-1}^1 W(\sqrt{r^2+s^2-2rst})dt,$$

where we have introduced

$$w_{00}(r,s) = \frac{1}{2}\int_{-1}^1 W(\sqrt{r^2+s^2-2rst})dt.$$

We have used that

$$\iint_{\mathbb{S}^2\times\mathbb{S}^2} d\omega d\omega' f(\omega\cdot\omega') = \frac{(4\pi)^2}{2}\int_{-1}^1 f(t)$$

for any integrable function f.

For the exchange term, we have

$$\begin{aligned}&\iint |g(x,y)|^2 W(x-y)dxdy \\ &= \iint_{\mathbb{R}^3\times\mathbb{R}^3}\sum_{\ell=0}^{\ell_{\max}} g^\ell(x,y)\sum_{\ell'=0}^{\ell'_{\max}} g^{\ell'}(x,y)W(x-y)dxdy \\ &= \sum_{i,j,m,n=1}^{N_b}\sum_{\ell,\ell'=0}^{\ell_{\max}}\frac{2\ell+1}{4\pi}\frac{2\ell'+1}{4\pi}g^\ell_{ij}g^{\ell'}_{mn}\times \\ &\qquad\times\iint_{\mathbb{R}^3\times\mathbb{R}^3}\chi_i(|x|)\chi_j(|y|)P_\ell(\omega\omega')\chi_m(|x|)\chi_n(|y|)P_{\ell'}(\omega\omega')W(|x-y|)dxdy \\ &= \sum_{\ell,\ell'=0}^{\ell_{\max}}(2\ell+1)(2\ell'+1)\sum_{i,j,m,n=1}^{N_b}G^\ell_{ij}G^{\ell'}_{mn}(im,jn)_{\ell\ell'},\end{aligned}$$

where we have introduced

$$\begin{aligned}&(im,jn)_{\ell\ell'} \\ &= \frac{1}{(4\pi)^2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\chi_i(|x|)\chi_j(|y|)\chi_m(|x|)\chi_n(|y|)W(|x-y|)P_\ell(\omega\omega')P_{\ell'}(\omega\omega')dxdy.\end{aligned}$$

Changing variables again, we find for the double integral

$$\begin{aligned}
&(im, jn)_{\ell\ell'} \\
&= \frac{1}{(4\pi)^2}\int_0^\infty r^2 dr \int_0^\infty s^2 ds \int_{\mathbb{S}^2} d\omega \int_{\mathbb{S}^2} d\omega' \times \\
&\qquad\qquad \times W(\sqrt{r^2+s^2-2rs\omega\omega'}) P_\ell(\omega\omega') P_{\ell'}(\omega\omega') \chi_i(r)\chi_j(s)\chi_m(r)\chi_n(s) \\
&= \frac{1}{2}\int_0^\infty r^2 dr \int_0^\infty s^2 ds\, \chi_i(r)\chi_j(s)\chi_m(r)\chi_n(s) w_{\ell\ell'}(r,s)
\end{aligned}$$

where we note

$$w_{\ell\ell'}(r,s) = \frac{1}{2}\int_{-1}^1 W(\sqrt{r^2+s^2-2rst}) P_\ell(t) P_{\ell'}(t) dt.$$

The computation is the same for the pairing term. □

Chapter 3

Algorithmic strategies and convergence analysis

In this section we study the convergence of two algorithms which can be used in practice to solve the HFB minimization problem (1.4.27). In order to simplify our presentation, we restrict ourselves to the finite-dimensional case, that is, to the discretized problem (2.2.22). We also assume that the discretization basis (χ_j) is orthonormal, such that

$$\mathbf{S} = I_{2N_b},$$

the $(2N_b) \times (2N_b)$ identity matrix. Finally, we only consider states which are invariant under time-reversal symmetry like in Section 2.3.1. This means that the HFB state is described by real and symmetric matrices G and A such that

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \leq \Upsilon := \begin{pmatrix} G & A \\ A & 1-G \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.0.1}$$

The energy is given by (2.3.32),

$$\mathcal{E}(\Upsilon) = 2\,\mathrm{Tr}(hG) + 2\,\mathrm{Tr}(G\,J(G)) - \mathrm{Tr}(G\,K(G)) + \mathrm{Tr}(A\,K(A)). \tag{3.0.2}$$

The extension to more general situations is straightforward.

The energy $\mathcal{E}$ is continuous (it is indeed real-analytic) with respect to Υ. Also the set $\mathcal{K}$ of density matrices Υ of the form (3.0.1) is compact in finite dimension. Hence minimizers always exist and, as we have seen, they solve the nonlinear equation

$$\Upsilon = \mathbb{1}_{(-\infty,0)}\Big(\mathbf{F}_\Upsilon - \mu\mathbf{N}\Big) + \delta, \tag{3.0.3}$$

where μ is a Lagrange multiplier chosen to ensure the constraint that $\mathrm{Tr}(G) = N/2$. Of course we must have $N/2 \leq N_b$, the dimension of the (no-spin) discretization space V_h, otherwise the minimization set is always empty.

3.1 Roothaan Algorithm

The most natural technique used in practice to solve the equation (3.0.3) is a simple fixed point method [10, 38]. This is usually reffered to as the *Roothaan algorithm* in the chemistry literature. The iteration scheme is the following

$$\Upsilon_{n+1} = \mathbb{1}_{(-\infty,0)}\Big(\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\Big) + \delta_{n+1} \tag{3.1.4}$$

where $\mathrm{Tr}(\mathbf{N}\Upsilon_{n+1}) = \dfrac{N}{2} - N_b$. At each step, one has to determine μ_{n+1} such as to satisfy the constraint $\mathrm{Tr}(G_{n+1}) = N/2$. If the operator $\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}$ has a trivial kernel, then $\delta_{n+1} \equiv 0$. This is the usual situation encountered in practice. In the iteration (3.1.4), the state is assumed to be pure at each step (Υ_n is an orthogonal projection). Recall that by Theorem 1.4.2 we know that minimizers of $\mathcal{E}$ under a particle number constraint are always pure, under suitable assumptions on the interaction potential W. The algorithm is stopped when the commutator

$$\big\|\big[\Upsilon_n, \mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\big]\big\|$$

and/or when the variation of the HFB state

$$\|\Upsilon_{n+1} - \Upsilon_n\|$$

are smaller than a prescribed ε. Here the norm is a classical Frobenius norm.

Our purpose in this section is to study the behavior of the Roothaan algorithm (3.1.4). In the Hartree-Fock case, it was shown in a fundamental work of Cancès and Le Bris [6, 7], that the algorithm converges or oscillates between two points, none of them being a solution to the equation (3.0.3). This result was recently improved by Levitt in [27]. We will explain that the results of Cancès-Le Bris and Levitt can be generalized to the HFB model. Actually, in a discretization space of dimension N_b, HFB is equivalent to a Hartree-Fock-like minimization problem in dimension $2N_b$, with additional constraints; see the remark 3.3.3. The adaptation of the previously cited works in the HF case reduces to handling these contraints.

In order to avoid the convergence problems of the Roothaan algorithm, Cancès and Le Bris have proposed the *Optimal Damping Algorithm* (ODA). We will study the equivalent of this algorithm in HFB theory in Section 3.2.

To start with, we show that the Roothaan algorithm is well defined, in the sense that for any HFB state Υ_n, there exists $(\Upsilon_{n+1}, \mu_{n+1}, \delta_{n+1})$ solving (3.1.4). To this end, we follow [6, 7] and introduce the auxillary functional

$$\tilde{\mathcal{E}}(\Upsilon, \Upsilon') := \mathrm{Tr}(hG) + \mathrm{Tr}(hG') + 2\,\mathrm{Tr}(G\,J(G')) - \mathrm{Tr}(G\,K(G')) + \mathrm{Tr}(A\,K(A')) \tag{3.1.5}$$

as well as the variational problem

$$I_\Upsilon(\lambda) := \min_{\Upsilon'}\left\{\tilde{\mathcal{E}}(\Upsilon, \Upsilon') \;:\; \mathrm{Tr}(G') = \frac{N}{2}\right\} \tag{3.1.6}$$

which consists in minimizing over Υ' with Υ fixed. The matrix Υ' must be an admissible HFB state which, in our context, means that

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \leq \Upsilon' := \begin{pmatrix} G' & A' \\ A' & 1 - G' \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, (G')^T = \overline{G'} = G', (A')^T = \overline{A'} = A'. \tag{3.1.7}$$

Recall that we have chosen an orthonormal basis and that the spin has been eliminated. It is clear that (3.1.6) admits at least one solution Υ', as soon as $0 \leq N/2 \leq N_b$, where we recall that N_b is the dimension of the discretization space V_h. The following says that these solutions are exactly those solving the equation of the Roothaan method.

Lemma 3.1.1 (The Roothaan algorithm is well defined)**.** *The function $N/2 \in [0, N_b] \mapsto I_\Upsilon(N/2)$ is convex, hence left and right differentiable. For any $N/2 \in [0, N_b]$, the minimizers Υ' of $I_\Upsilon(N/2)$ are exactly the states of the form*

$$\Upsilon' = \mathbb{1}_{(-\infty,0)}\big(\mathbf{F}_\Upsilon - \mu'\mathbf{N}\big) + \delta' \tag{3.1.8}$$

where $\mu' \in [\partial_- I_\Upsilon(N/2), \partial_+ I_\Upsilon(N/2)]$ and $0 \leq \delta' \leq \mathbb{1}_{\{0\}}(\mathbf{F}_\Upsilon - \mu'\mathbf{N})$. If $0 \notin \sigma(\mathbf{F}_\Upsilon - \mu'\mathbf{N})$, then $\delta' \equiv 0$ and Υ' is unique, for any such $\mu' \in [\partial_- I_\Upsilon(N/2), \partial_+ I_\Upsilon(N/2)]$.

Proof. To see that $I_\Upsilon(N/2)$ is convex, let $0 \leq N_1/2 < N_2/2 \leq N_b$ and let Υ'_i be a minimizer for $I_\Upsilon(N_i/2)$ with $i = 1, 2$. Then $t\Upsilon'_1 + (1-t)\Upsilon'_2$ is a test state for the problem $I_\Upsilon(tN_1/2 + (1-t)N_2/2)$. Therefore it holds

$$\begin{aligned} I_\Upsilon(tN_1/2 + (1-t)N_2/2) \leq \tilde{\mathcal{E}}(\Upsilon, t\Upsilon'_1 + (1-t)\Upsilon'_2) &= t\tilde{\mathcal{E}}(\Upsilon, \Upsilon'_1) + (1-t)\tilde{\mathcal{E}}(\Upsilon, \Upsilon'_2) \\ &= tI_\Upsilon(N_1/2) + (1-t)I_\Upsilon(N_2/2). \end{aligned}$$

Then, by convexity we get that $I_\Upsilon(N'/2) \geq I_\Upsilon(N/2) + \mu(N'/2 - N/2)$ for any $N'/2 \in [0, N_b]$ and any $\mu \in [\partial_- I_\Upsilon(N/2), \partial_+ I_\Upsilon(N/2)]$. Thus

$$\begin{aligned} I_\Upsilon(N/2) - \mu N/2 &= \min\{I_\Upsilon(N'/2) - \mu N'/2 \ : \ 0 \leq N/2 \leq N_b\} \\ &= \min_{\Upsilon'} \left\{\tilde{\mathcal{E}}(\Upsilon, \Upsilon') - \mu \operatorname{Tr}(G')\right\} \\ &= \operatorname{Tr}(hG) + \frac{1}{2} \min_{\Upsilon'} \operatorname{Tr}(\mathbf{F}_\Upsilon - \mu\mathbf{N})\Upsilon'. \end{aligned}$$

In the previous two mins, Υ' is varied over all possible HFB states, without any particle number constraint. It is well known that the minimizers of the problem on the right side are exactly the solutions of the equation (3.1.8). □

Lemma 3.1.1 tells us that for any given Υ_n, there always exists at least one solution $(\Upsilon_{n+1}, \mu_{n+1}, \delta_{n+1})$ of the equation (3.1.4). It is obtained by solving the minimization problem $I_{\Upsilon_n}(N/2)$, and one has to take $\mu_{n+1} \in [\partial_- I_{\Upsilon_n}(N/2), \partial_+ I_{\Upsilon_n}(N/2)]$. We can always take by convention

$$\mu_{n+1} := \frac{\partial_- I_{\Upsilon_n}(N/2) + \partial_+ I_{\Upsilon_n}(N/2)}{2}.$$

However, Υ_{n+1} is not uniquely defined yet because of the possibility of having $\delta_{n+1} \neq 0$. As we have seen it is unique when

$$0 \notin \sigma\big(\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\big). \tag{3.1.9}$$

In this section we always assume that it is possible to find $(\Upsilon_{n+1}, \mu_{n+1}, \delta_{n+1})$ exactly. Later in Section 3.3 we explain how to do this numerically. We will also see that the condition (3.1.9) is "very often" satisfied. This vague statement is made precise in Lemma 3.3.2 below.

In practice, we will only know $(\Upsilon_{n+1}, \mu_{n+1}, \delta_{n+1})$ approximately. But here we assume for simplicity that we know them exactly. Following Cancès and Le Bris, we now introduce the concept of uniform well-posedness.

Definition 3.1.1 (Uniform well-posedness)**.** *We say that for a given initial HFB state Υ_0, the sequence (Υ_n) generated by the Roothaan algorithm is* uniformly well posed *when there exists $\eta > 0$ such that*

$$|\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}| \geq \eta \tag{3.1.10}$$

for all $n \geq 0$, where $\mu_{n+1} = \big(\partial_- I_{\Upsilon_n}(N/2) + \partial_+ I_{\Upsilon_n}(N/2)\big)/2$.

Note that the condition $|\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}| \geq \eta$ is equivalent to $(-\eta, \eta) \cap \sigma(\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}) = \emptyset$. Later in Section 3.3 we will make several comments concerning Assumption (3.1.10).

We have seen that the sequence generated by the Roothaan algorithm can be obtained by solving the minimization problem

$$I_{\Upsilon_n} = \min_{\Upsilon'} \tilde{\mathcal{E}}(\Upsilon_n, \Upsilon').$$

Since $\tilde{\mathcal{E}}(\Upsilon, \Upsilon') = \tilde{\mathcal{E}}(\Upsilon', \Upsilon)$, we conclude that the Roothaan algorithm is the same as minimizing $\tilde{\mathcal{E}}$ with respect to its first and second variables one after another, inductively. This fact allows to prove the following result, which is the HFB equivalent of Theorem 7 in [7] and Theorem 5.1 in [27] in the HF case.

Theorem 3.1.1 (Convergence of the Roothaan algorithm)**.** *Assume that $0 < N/2 < N_b$. Let Υ_0 be an initial HFB state such that the sequence (Υ_n) generated by the Roothaan algorithm is* uniformly well posed*. Then*

- *The sequence $\tilde{\mathcal{E}}(\Upsilon_{2n}, \Upsilon_{2n+1})$ decreases towards a critical value of $\tilde{\mathcal{E}}$;*
- *The sequence $(\Upsilon_{2n}, \Upsilon_{2n+1})$ converges towards a critical point (Υ, Υ') of $\tilde{\mathcal{E}}$;*
- *If $\Upsilon = \Upsilon'$, then this state is a solution of the original HFB equation* (3.0.3)*, but if $\Upsilon \neq \Upsilon'$, then none of these two states is a solution to* (3.0.3)*.*

Theorem 3.1.1 says that (provided it is uniformly well posed) the sequence Υ_n will either converge to a solution of the self-consistent Equation (3.0.3), or oscillate between two points Υ and Υ', none of them being a solution to the desired equation.

Proof. We split the proof into several steps.

Step 1: μ_n ***is uniformly bounded.*** It will be very useful to know that the sequence μ_n is uniformly bounded. The following says that, as soon as we fix $\mathrm{Tr}(G) = N/2$ with $0 < N/2 < N_b$, the chemical potential μ cannot be too negative and too positive.

Lemma 3.1.2 (Bounds on the multiplier μ)**.** *Let Υ' be any fixed HFB state and*

$$\Upsilon_\mu := \mathbb{1}_{(-\infty,0)}\big(\mathbf{F}_{\Upsilon'} - \mu\mathbf{N}\big) = \begin{pmatrix} G_\mu & A_\mu \\ A_\mu & 1 - G_\mu \end{pmatrix}$$

the corresponding HFB ground state at chemical potential μ. There exists a constant C which is independent of Υ' and μ, such that

$$\forall \mu \leq -C, \qquad \mathrm{Tr}\, G_\mu \leq \frac{C}{|\mu|} \tag{3.1.11}$$

and

$$\forall \mu \geq C, \qquad \mathrm{Tr}\, G_\mu \geq N_b - \frac{C}{\mu}. \tag{3.1.12}$$

The lemma says that the average number of particles in $\mathbb{1}_{(-\infty,0)}\big(\mathbf{F}_\Upsilon - \mu\mathbf{N}\big)$ tends to N_b when $\mu \to \infty$ whereas it tends to 0 when $\mu \to -\infty$, this *uniformly with respect to the state* Υ' used to build the mean-field operator $\mathbf{F}_{\Upsilon'}$.

Proof. We first remark that there exists a constant C such that

$$\|F_{\Upsilon'}\| \leq C \tag{3.1.13}$$

for any HFB state Υ'. This follows from the fact that $F_{\Upsilon'}$ is continuous with respect to Υ' and that the latter lives in a compact set since we always have $0 \leq \Upsilon \leq 1$. The chosen norm for $\|F_{\Upsilon'}\|$ does not matter since we are in finite dimension. Now, for μ large enough we can use regular perturbation theory (see the proof of Lemma 3.3.1) and obtain that

$$\begin{aligned} &\big\|\mathbb{1}_{(-\infty,0)}\big(F_{\Upsilon'} - \mu\mathbf{N}\big) - \mathbb{1}_{(-\infty,0)}\big(-\mu\mathbf{N}\big)\big\| \\ &\qquad = \left\|\mathbb{1}_{(-\infty,0)}\left(\frac{F_{\Upsilon'}}{|\mu|} - \frac{\mu}{|\mu|}\mathbf{N}\right) - \mathbb{1}_{(-\infty,0)}\left(-\frac{\mu}{|\mu|}\mathbf{N}\right)\right\| \leq \frac{C}{|\mu|}. \end{aligned}$$

Note that

$$\mathbb{1}_{(-\infty,0)}\left(-\frac{\mu}{|\mu|}\mathbf{N}\right) = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \text{for } \mu > 0, \\ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & \text{for } \mu < 0. \end{cases}$$

Taking the trace against $\mathbf{N}$ gives the result. □

From Lemma 3.1.2 we deduce that our sequence (μ_n) is bounded. Indeed, since we have by construction $\mathrm{Tr}(G_{n+1}) = N/2$ with $0 < N/2 < N_b$, we must have

$$-\max\left(C\,,\,\frac{2C}{N}\right) \leq \mu_{n+1} \leq \max\left(C\,,\,\frac{C}{N_b - N/2}\right) \tag{3.1.14}$$

otherwise $\mathrm{Tr}(G_{n+1})$ would be too small or too large.

Step 2: *convergence of* $\tilde{\mathcal{E}}(\Upsilon_n, \Upsilon_{n+1})$. We now follow [4, 6, 7]. At each step, we know from Lemma 3.1.1 that Υ_{n+1} is a solution of the minimization problem $\min_{\Upsilon'} \tilde{\mathcal{E}}(\Upsilon_n, \Upsilon')$. In particular, we deduce that

$$\tilde{\mathcal{E}}(\Upsilon_n, \Upsilon_{n+1}) \leq \tilde{\mathcal{E}}(\Upsilon_n, \Upsilon_{n-1}) = \tilde{\mathcal{E}}(\Upsilon_{n-1}, \Upsilon_n). \tag{3.1.15}$$

Thus the sequence of real numbers $\tilde{\mathcal{E}}(\Upsilon_n, \Upsilon_{n+1})$ is non-increasing. It is also bounded from below, hence it converges to a limit ℓ. We now use the uniform well-posedness to prove an inequality which is more precise than (3.1.15). We remark that

$$\begin{aligned}
\tilde{\mathcal{E}}(\Upsilon_n, \Upsilon_{n-1}) - \tilde{\mathcal{E}}(\Upsilon_n, \Upsilon_{n+1}) &= \frac{1}{2}\,\mathrm{Tr}\,\mathbf{F}_{\Upsilon_n}\big(\Upsilon_{n-1} - \Upsilon_{n+1}\big) \\
&= \frac{1}{2}\,\mathrm{Tr}\,\big(\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\big)\big(\Upsilon_{n-1} - \Upsilon_{n+1}\big) \\
&\geq \frac{1}{2}\,\mathrm{Tr}\,\big|\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\big|\big(\Upsilon_{n-1} - \Upsilon_{n+1}\big)^2 \\
&\geq \frac{\eta}{2}\,\mathrm{Tr}\,\big(\Upsilon_{n-1} - \Upsilon_{n+1}\big)^2 = \frac{\eta}{2}\,\|\Upsilon_{n-1} - \Upsilon_{n+1}\|^2\,.
\end{aligned} \tag{3.1.16}$$

In the above calculation we have used that $\mathrm{Tr}\,\mathbf{N}\,\Upsilon_{n+1} = \mathrm{Tr}\,\mathbf{N}\,\Upsilon_{n-1} = N - N_b$. We have also used that Υ_{n+1} is the negative spectral projector of $\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}$, such that we can write

$$\big(\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\big) = \big|\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\big|\big(\Upsilon_{n+1}^{\perp} - \Upsilon_{n+1}\big).$$

Finally, we have used that $0 \leq \gamma \leq 1$ is equivalent to $\gamma^2 \leq \gamma$. We can write this as

$$\begin{aligned}
\gamma^2 &= (\gamma - P + P)^2, \\
&= (\gamma - P)^2 + P^2 + (\gamma - P)P + P(\gamma - P), \\
&= (\gamma - P)^2 - P + \gamma P + P\gamma, \\
&\leq \gamma,
\end{aligned}$$

and we obtain

$$\begin{aligned}
(\gamma - P)^2 &\leq \gamma + P - \gamma P - P\gamma, \\
&\leq P^{\perp}\gamma P^{\perp} + P^{\perp}\gamma P + P\gamma P^{\perp} + P\gamma P - P^{\perp}\gamma P - P\gamma P - P\gamma P^{\perp} - P\gamma P + P, \\
&\leq P^{\perp}\gamma P^{\perp} - P\gamma P + P, \\
&\leq P^{\perp}\,(\gamma - P)P^{\perp} - P\,(\gamma - P)P,
\end{aligned}$$

for any orthogonal projector P, thus

$$\Upsilon_{n+1}^{\perp}\big(\Upsilon_{n-1}-\Upsilon_{n+1}\big)\Upsilon_{n+1}^{\perp}-\Upsilon_{n+1}\big(\Upsilon_{n-1}-\Upsilon_{n+1}\big)\Upsilon_{n+1}\geq\big(\Upsilon_{n-1}-\Upsilon_{n+1}\big)^2.$$

Since $\tilde{\mathcal{E}}(\Upsilon_n,\Upsilon_{n+1})$ converges to a limit ℓ, we deduce that

$$\sum_{n\geq 1}\|\Upsilon_{n+1}-\Upsilon_{n-1}\|^2<\infty.$$

In particular, $\|\Upsilon_{n+1}-\Upsilon_{n-1}\|\to 0$ which is called *numerical convergence* in [6, 7].

Step 3: *convergence of* Υ_{2n} *and* Υ_{2n+1}. In order to upgrade the numerical convergence to true convergence, we use a recent method of Levitt [27]. Namely we show that

$$\Big(\tilde{\mathcal{E}}(\Upsilon_n,\Upsilon_{n-1})-\ell\Big)^{\theta}-\Big(\tilde{\mathcal{E}}(\Upsilon_n,\Upsilon_{n+1})-\ell\Big)^{\theta}\geq\frac{\eta'}{2}\|\Upsilon_{n-1}-\Upsilon_{n+1}\| \tag{3.1.17}$$

for a well chosen $0<\theta<1/2$ and $\theta'>0$. Summing over n and using the convergence of $\tilde{\mathcal{E}}(\Upsilon_n,\Upsilon_{n-1})$, hence of $(\tilde{\mathcal{E}}(\Upsilon_n,\Upsilon_{n-1})-\ell)^{\theta}$, then gives the convergence of Υ_{2n} and Υ_{2n+1}.

For the proof of (3.1.17), we argue as follow. Consider a (real, no-spin) pure HFB state Υ. It is possible to parametrize the manifold of pure HFB states around Υ by using Bogoliubov transformations as follows:

$$H\mapsto e^{H}\Upsilon e^{-H}$$

where H is assumed to be of the form

$$\begin{pmatrix}h & p\\ -p & -h\end{pmatrix},\qquad h^T=-\overline{h}=-h,\quad p^T=\overline{p}=p.$$

These constraints ensure that iH is a self-adjoint Hamiltonian such that $e^{H}=e^{-i(iH)}$ is a Bogoliubov rotation. They also ensure that $e^{H}\Upsilon e^{-H}$ stays real. That $H\mapsto e^{H}\Upsilon e^{-H}$ is a local map of the manifold of pure HFB states around Υ follows from the arguments in [3] as well as simple considerations in linear algebra.

Let us now consider the energy $\tilde{\mathcal{E}}$ in a neighborhood of any fixed (Υ,Υ'). The map

$$f:(H,H')\mapsto\tilde{\mathcal{E}}\big(e^{H}\Upsilon e^{-H},e^{H'}\Upsilon' e^{-H'}\big)-\frac{\mu'}{2}\,\mathrm{Tr}\,\mathbf{N}e^{H}\Upsilon e^{-H}-\frac{\mu}{2}\,\mathrm{Tr}\,\mathbf{N}e^{H'}\Upsilon' e^{-H'}$$

is real analytic in a neighborhood of $(0,0)$ for any fixed $\mu,\mu'\in\mathbb{R}$ and any fixed pure HFB states (Υ,Υ'). The Łojasiewicz inequality (Theorem 2.1 in [27]) then tells us that there exist $0<\theta\leq 1/2$ and a constant $\kappa>0$ such that $\|H\|+\|H'\|\leq\kappa$ implies

$$|f(H,H')-f(0)|^{1-\theta}\leq\kappa^{-1}\Big(|\nabla_H f(H,H')|+|\nabla_{H'}f(H,H')|\Big).$$

A simple computation shows that

$$\nabla_H f(H,H') = \frac{1}{2}\big[\mathbf{F}_{\Upsilon'} - \mu'\mathbf{N}\,,\, e^H \Upsilon e^{-H}\big], \quad \nabla_{H'} f(H,H') = \frac{1}{2}\big[\mathbf{F}_{\Upsilon} - \mu\mathbf{N}\,,\, e^{H'} \Upsilon' e^{-H'}\big].$$

If we rephrase all this in our setting, this means that for any fixed pure HFB states $(\Upsilon_1, \Upsilon_1')$ and any $\mu, \mu' \in \mathbb{R}$, there is a $\kappa > 0$ such that for any $(\Upsilon_2, \Upsilon_2')$ another pure HFB state which is at most at a distance κ from $(\Upsilon_1, \Upsilon_1')$, we have

$$\begin{aligned}\Big|\tilde{\mathcal{E}}\big(\Upsilon_1, \Upsilon_1'\big) - \tilde{\mathcal{E}}\big(\Upsilon_2, \Upsilon_2'\big) + \mu' \operatorname{Tr}\mathbf{N}(G_2 - G_1) + \mu \operatorname{Tr}\mathbf{N}(G_2' - G_1')\Big|^{1-\theta} \\ \leq \kappa^{-1}\Big(\big\|\big[\mathbf{F}_{\Upsilon_2'} - \mu'\mathbf{N}\,,\, \Upsilon_2\big]\big\| + \big\|\big[\mathbf{F}_{\Upsilon_2} - \mu\mathbf{N}\,,\, \Upsilon_2'\big]\big\| \Big). \quad (3.1.18)\end{aligned}$$

The constants κ and θ depend on μ, μ' and of the reference point $(\Upsilon_1, \Upsilon_1')$. But they stay positive as soon as μ, μ' and $(\Upsilon_1, \Upsilon_1')$ stay in a compact set. By a simple compactness argument, we therefore deduce that there exists a neighborhood of the compact set

$$\Big\{(\Upsilon, \Upsilon') \;:\; \tilde{\mathcal{E}}(\Upsilon, \Upsilon') = \ell,\ \operatorname{Tr}(G) = \operatorname{Tr}(G') = N/2\Big\}$$

such that for any (Υ, Υ') in this neighborhood and μ, μ' in a compact set in $\mathbb{R}$, we have

$$\begin{aligned}\Big|\tilde{\mathcal{E}}\big(\Upsilon, \Upsilon'\big) - \ell + \mu'\big(N/2 - \operatorname{Tr}\mathbf{N}G\big) + \mu\big(N/2 - \operatorname{Tr}\mathbf{N}G'\big)\Big|^{1-\theta} \\ \leq \kappa^{-1}\Big(\big\|\big[\mathbf{F}_{\Upsilon'} - \mu'\mathbf{N}\,,\, \Upsilon\big]\big\| + \big\|\big[\mathbf{F}_{\Upsilon} - \mu\mathbf{N}\,,\, \Upsilon'\big]\big\| \Big) \quad (3.1.19)\end{aligned}$$

for some $0 < \theta \leq 1/2$ and some $\kappa > 0$. We recall that ℓ is by definition the limit of $\tilde{\mathcal{E}}\big(\Upsilon_n, \Upsilon_{n+1}\big)$.

Recall our inequality (3.1.14) which says that μ_n is uniformly bounded. Also, we know that $\tilde{\mathcal{E}}(\Upsilon_n, \Upsilon_{n+1})$ converges to ℓ so, for n large enough, $(\Upsilon_n, \Upsilon_{n+1})$ must be in the neighborhood of the level set ℓ. Choosing $\mu = \mu_{n+1}$ and $\mu' = \mu_n$ and using that G_n and G_{n+1} have the correct trace, we get the estimate

$$\begin{aligned}\Big(\tilde{\mathcal{E}}\big(\Upsilon_n, \Upsilon_{n-1}\big) - \ell\Big)^{1-\theta} &\leq \kappa^{-1}\left(\big\|\big[\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\,,\, \Upsilon_{n-1}\big]\big\| + \big\|\big[\mathbf{F}_{\Upsilon_{n-1}} - \mu_n\mathbf{N}\,,\ ,\, \Upsilon_n\big]\big\|\right) \\ &= \kappa^{-1}\big\|\big[\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}\,,\, \Upsilon_{n-1} - \Upsilon_{n+1}\big]\big\| \\ &\leq C\,\|\Upsilon_{n-1} - \Upsilon_{n+1}\|\end{aligned}$$

for n large enough. Here we have used that Υ_n commutes with $\mathbf{F}_{\Upsilon_{n-1}} - \mu_n\mathbf{N}$ and that Υ_{n+1} commutes with $\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}$ by construction, and that $\|\mathbf{F}_{\Upsilon_n}\|$ and μ_{n+1} are both uniformly bounded. In order to conclude, we use the concavity of $x \mapsto x^\theta$

and (3.1.16) like in [27] to obtain

$$\begin{aligned}\Big(\tilde{\mathcal{E}}\big(\Upsilon_n,\Upsilon_{n-1}\big)-\ell\Big)^{\theta}-\Big(\tilde{\mathcal{E}}\big(\Upsilon_n,\Upsilon_{n+1}\big)-\ell\Big)^{\theta}&\\ \geq\frac{\theta}{\Big(\tilde{\mathcal{E}}\big(\Upsilon_n,\Upsilon_{n-1}\big)-\ell\Big)^{1-\theta}}\Big(\tilde{\mathcal{E}}\big(\Upsilon_n,\Upsilon_{n-1}\big)-\tilde{\mathcal{E}}\big(\Upsilon_n,\Upsilon_{n+1}\big)\Big)&\\ \geq\frac{\eta\,\theta}{2\Big(\tilde{\mathcal{E}}\big(\Upsilon_n,\Upsilon_{n-1}\big)-\ell\Big)^{1-\theta}}\left\|\Upsilon_{n+1}-\Upsilon_{n-1}\right\|^2&\\ \geq\eta\theta/(2C)\,\left\|\Upsilon_{n+1}-\Upsilon_{n-1}\right\|&\end{aligned}$$

by (3.1.16). This concludes the proof of the inequality (3.1.17), hence the proof of the convergence of $(\Upsilon_{2n},\Upsilon_{2n+1})$, towards some pure HFB states (Υ,Υ').

Step 4: ***the limit*** (Υ,Υ') ***of*** $(\Upsilon_{2n},\Upsilon_{2n+1})$ ***is a critical point of*** $\tilde{\mathcal{E}}$**.** Since we have $\Upsilon_{2n}\to\Upsilon$ and $\Upsilon_{2n+1}\to\Upsilon'$, we deduce that $\mathbf{F}_{\Upsilon_{2n}}\to\mathbf{F}_{\Upsilon}$ and $\mathbf{F}_{\Upsilon_{2n+1}}\to\mathbf{F}_{\Upsilon'}$, by continuity of the map $\Upsilon\mapsto\mathbf{F}_{\Upsilon}$. Extracting a subsequence, we can assume that $\mu_{2n_k}\to\mu'$ and $\mu_{2n_k+1}\to\mu$. We have

$$\Upsilon_{2n_k}=\mathbb{1}_{(-\infty,0)}(\mathbf{F}_{\Upsilon_{2n_k-1}}-\mu_{2n_k}\mathbf{N}),\qquad\Upsilon_{2n_k+1}=\mathbb{1}_{(-\infty,0)}(\mathbf{F}_{\Upsilon_{2n_k}}-\mu_{2n_k+1}\mathbf{N})$$

and, by uniform well-posedness,

$$\left|\mathbf{F}_{\Upsilon_{2n_k-1}}-\mu_{2n_k}\mathbf{N}\right|\geq\eta,\qquad\left|\mathbf{F}_{\Upsilon_{2n_k}}-\mu_{2n_k+1}\mathbf{N}\right|\geq\eta.$$

Passing to the limit $k\to\infty$ we get

$$\Upsilon=\mathbb{1}_{(-\infty,0)}(\mathbf{F}_{\Upsilon'}-\mu'\mathbf{N})\quad\text{and}\quad\Upsilon'=\mathbb{1}_{(-\infty,0)}(\mathbf{F}_{\Upsilon}-\mu\mathbf{N}).$$

This exactly means that (Υ,Υ') is a critical point of $\tilde{\mathcal{E}}$ on $\mathcal{P}_h(N/2)\times\mathcal{P}_h(N/2)$. Note that we have also $\left|\mathbf{F}_{\Upsilon'}-\mu'\mathbf{N}\right|\geq\eta$ and $\left|\mathbf{F}_{\Upsilon}-\mu\mathbf{N}\right|\geq\eta$.

The remaining statements are verified exactly like in the HF case. This concludes the proof of Theorem 3.1.1. □

3.2 Optimal Damping Algorithm

In the previous section we have studied the convergence properties of the Roothaan algorithm, which consists in solving the self-consistent equation by a fixed point method. We have seen that the algorithm can either converge or oscillate between two states, none of them being a solution to the problem.

Examples of such oscillations in quantum chemistry have been exhibited by Cancès and Le Bris [6, 7]. In this case the potential W is repulsive and there is no pairing. In order to cure this problem of oscillations, Cancès and Le Bris proposed in [7] a *relaxed algorithm* called the *Optimal Damping Algorithm* (ODA). This method makes use of the important fact that one can minimize over mixed

states and get the same ground state as when minimizing over pure states only (Theorem 1.4.1).

The same oscillations can *a priori* happen in HFB with an attractive potential W. They are frequently seen with the Roothaan algorithm, at least during the first iterations (examples will be given in Section 4). Even when the sequence (Υ_n) eventually converges towards a single state Υ, these oscillations slow down the convergence considerably. This phenomenon is well known in nuclear physics. Dechargé and Gogny already advocate in [10] the use of a *damping parameter* between two successive iterations, in order to *"slow down the convergence on the density matrix. In this way the average field varies slowly and we can insure the convergence on the pairing tensor step by step"* (see [10] page 1574). Even in the modern computations, this damping parameter is fixed all along the algorithm.

We suggest to transpose the method of Cancès and Le Bris to the HFB setting by using an optimal damping parameter, chosen such as to minimize the energy. This means resorting to mixed states even if the final ground state is always a pure HFB state. This is theoretically justified when the assumptions of the Bach-Fröhlich-Jonsson Theorem 1.4.2 are fulfilled.

The ODA involves two density matrices Υ_n and $\tilde{\Upsilon}_n$. The HFB state Υ_n is always pure but $\tilde{\Upsilon}_n$ can (and will usually) be a mixed HFB state. The starting point $\Upsilon_0 = \tilde{\Upsilon}_0$ being chosen, the sequence is then constructed by induction as follows:

1. One finds $(\Upsilon_{n+1}, \mu_{n+1})$ solving

$$\Upsilon_{n+1} = \mathbb{1}_{(-\infty,0)}\big(\mathbf{F}_{\tilde{\Upsilon}_n} - \mu_{n+1}\mathbf{N}\big) \quad \text{and} \quad \mathrm{Tr}(G_{n+1}) = N/2.$$

   This is always possible, by Lemma 3.1.1 and we can take as before

$$\mu_{n+1} := \frac{\partial_- I_{\tilde{\Upsilon}_n}(N/2) + \partial_+ I_{\tilde{\Upsilon}_n}(N/2)}{2},$$

   in case 0 is in the spectrum of $\mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}$.

2. One lets

$$\tilde{\Upsilon}_{n+1} = t_{n+1}\tilde{\Upsilon}_n + (1 - t_{n+1})\Upsilon_{n+1}$$

   where the damping parameter $t_{n+1} \in [0, 1]$ is chosen such as to minimize the (quadratic) function

$$t \mapsto \mathcal{E}\big(t\tilde{\Upsilon}_n + (1-t)\Upsilon_{n+1}\big).$$

3. The algorithm is stopped when $\|[\Upsilon_n, \mathbf{F}_{\Upsilon_n} - \mu_{n+1}\mathbf{N}]\|$ and/or $\|\Upsilon_{n+1} - \Upsilon_n\|$ are smaller than a prescribed ε.

The general strategy of the ODA is displayed in Figure 3.1. By construction we see that $\mathcal{E}(\tilde{\Upsilon}_n)$ is a non-increasing sequence. This guarantees the convergence of the ODA. The result is the following

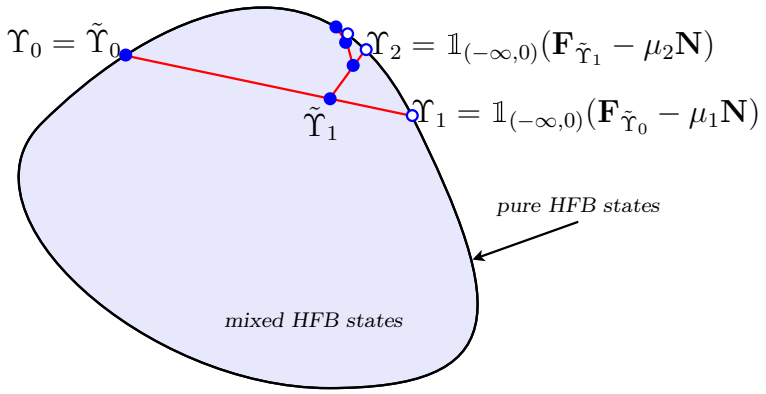

Figure 3.1: Schematic representation of the Optimal Damping Algorithm of Cancès & Le Bris in the HFB case. In reality, pure HFB states are only the extremal points of the set of states.

Theorem 3.2.1 (Convergence of the ODA)**.** *Assume that* $0 < N/2 < N_b$*. Let* $\Upsilon_0 = \tilde{\Upsilon}_0$ *be an initial HFB state such that the sequence* (Υ_n) *generated by the ODA is* uniformly well posed*, that is*

$$\forall n, \qquad \left|\mathbf{F}_{\tilde{\Upsilon}_n} - \mu_{n+1}\mathbf{N}\right| \geq \eta > 0. \tag{3.2.20}$$

Then

- *The sequence* $\mathcal{E}(\tilde{\Upsilon}_n)$ *decreases towards a critical value of* $\mathcal{E}$*;*
- *The sequence* Υ_n numerically *converges towards a critical point* Υ *of* $\mathcal{E}$*, in the sense that* $\Upsilon_{n+1} - \Upsilon_n \to 0$*,* $\Upsilon_{n+1} - \tilde{\Upsilon}_n \to 0$ *and that all the limit points* Υ *of subsequences of* (Υ_n) *solve* $\Upsilon = \mathbb{1}_{(-\infty,0)}(\mathbf{F}_\Upsilon - \mu\mathbf{N})$*.*

Proof. The proof is exactly the same as in the Hartree-Fock case [4,8] and we only sketch it. First we have by definition $\mathcal{E}(\tilde{\Upsilon}_{n+1}) \leq \mathcal{E}(\tilde{\Upsilon}_n)$, so $\mathcal{E}(\tilde{\Upsilon}_n)$ must converge to a limit ℓ. Now we have

$$\mathcal{E}(\tilde{\Upsilon}_{n+1}) = \mathcal{E}\big((1-t_{n+1})\tilde{\Upsilon}_n + t_{n+1}\Upsilon_{n+1}\big) = \mathcal{E}(\tilde{\Upsilon}_n) - t_{n+1}a_{n+1} + t_{n+1}^2 b_{n+1}$$

where

$$t_{n+1} = \operatorname{argmin}_{t\in[0,1]}\big(-ta_{n+1} + t^2 b_{n+1}\big)$$

and with

$$a_{n+1} := \operatorname{Tr}\mathbf{F}_{\tilde{\Upsilon}_n}\big(\tilde{\Upsilon}_n - \Upsilon_{n+1}\big) = \operatorname{Tr}\left|\mathbf{F}_{\tilde{\Upsilon}_n} - \mu\mathbf{N}\right|\big(\tilde{\Upsilon}_n - \Upsilon_{n+1}\big)^2 \geq \eta\left\|\tilde{\Upsilon}_n - \Upsilon_{n+1}\right\|^2$$

with $\mathbf{F}_\Upsilon = \mathbf{h} + \mathcal{L}(\Upsilon - \mathbf{P})$,

$$b_{n+1} = \operatorname{Tr}((\tilde{\Upsilon}_{n+1} - \Upsilon_n)\mathcal{L}(\tilde{\Upsilon}_{n+1} - \Upsilon_n)).$$

In finite dimension we have $|b_{n+1}| \leq C \left\| \tilde{\Upsilon}_{n+1} - \Upsilon_n \right\|^2 \leq (C/\eta) a_{n+1}$. This can be used to prove that

$$-t_{n+1} a_{n+1} + t_{n+1}^2 b_{n+1} \leq -\epsilon\, a_{n+1} \leq -\epsilon\, \eta \left\| \tilde{\Upsilon}_n - \Upsilon_{n+1} \right\|^2$$

for some $\epsilon > 0$ independent of n. This now proves that

$$\sum_n \left\| \tilde{\Upsilon}_n - \Upsilon_{n+1} \right\|^2 < \infty,$$

hence that $\tilde{\Upsilon}_n - \Upsilon_{n+1} \to 0$. In order to conclude the proof, we notice that

$$\Upsilon_{n+1} - \Upsilon_n = \Upsilon_{n+1} - \tilde{\Upsilon}_n + (1 - t_n)\big(\tilde{\Upsilon}_{n-1} - \Upsilon_n\big)$$

which finally implies

$$\sum_n \|\Upsilon_n - \Upsilon_{n+1}\|^2 < \infty \quad \text{and} \quad \sum_n \left\| \tilde{\Upsilon}_n - \tilde{\Upsilon}_{n+1} \right\|^2 < \infty.$$

Since $\Upsilon_{n+1} = \mathbb{1}_{(-\infty,0)}(\mathbf{F}_{\tilde{\Upsilon}_n} - \mu_{n+1}\mathbf{N})$ by definition, the proof that any limit Υ of a subsequence of (Υ_n) satisfies the self-consistent equation is elementary. □

3.3 Handling constraints

Both the Roothaan algorithm and the ODA are based on Lemma 3.1.1 which says that for any given $\mathbf{F}_\Upsilon$, there exist μ', δ' and Υ' such that

$$\begin{cases} \Upsilon' = \mathbb{1}_{(-\infty,0)}\big(\mathbf{F}_\Upsilon - \mu'\mathbf{N}\big) + \delta', \\ \operatorname{Tr} \mathbf{N}\Upsilon' = N - N_b. \end{cases} \tag{3.3.21}$$

The purpose of this section is to explain how to solve this problem numerically. To simplify our notation, we consider in this section a generic matrix

$$\mathbf{F} = \begin{pmatrix} h & p \\ p & -h \end{pmatrix}, \qquad \text{with } p = \overline{p} = p^T \text{ and } h = \overline{h} = h^T \tag{3.3.22}$$

and we study the problem consisting in finding Υ, μ and δ such that

$$\begin{cases} \Upsilon = \mathbb{1}_{(-\infty,0)}\big(\mathbf{F} - \mu\mathbf{N}\big) + \delta, \\ \operatorname{Tr} \mathbf{N}\Upsilon = N - N_b. \end{cases} \tag{3.3.23}$$

Assume first that $p \equiv 0$ (Hartree-Fock case). Then we have

$$\mathbf{F} = \begin{pmatrix} h & 0 \\ 0 & -h \end{pmatrix}$$

which commutes with $\mathbf{N}$. The solution of (3.3.23) is then given by the *Aufbau principle*,

$$\Upsilon = \begin{pmatrix} G & 0 \\ 0 & 1-G \end{pmatrix}, \qquad G = \mathbb{1}_{(-\infty,\mu)}(h) + \delta$$

where μ is the $(N/2)$th eigenvalue of h, counted with multiplicity and δ lives in the corresponding eigenspace. Equivalently,

$$G = \sum_{i=1}^{K} v_i v_i^T + \sum_{i=K+1}^{K'} n_i \, v_i v_i^T$$

where the v_i's solve the eigenvalue equation

$$h \, v_i = \epsilon_i \, v_i,$$

where ϵ_i are the eigenvalues of h sorted in ascending order, and v_i the orthonormal eigenfunctions. $K = \operatorname{Tr} \mathbb{1}_{(-\infty,\epsilon_{N/2})}(h)$ is the dimension of the direct sum of all the eigenspaces corresponding to the eigenvalues $< \epsilon_{N/2}$ and $K' = \operatorname{Tr} \mathbb{1}_{(-\infty,\epsilon_{N/2}]}(h)$ is the dimension of the direct sum of all the eigenspaces corresponding to the eigenvalues $\leq \epsilon_{N/2}$. The n_i's are chosen such that

$$0 \leq n_i \leq 1, \qquad K + \sum_{i=K+1}^{K'} n_i = \frac{N}{2}.$$

Therefore, finding Υ, μ and δ in the Hartree-Fock case only requires to diagonalize h once.

In the Hartree-Fock-Bogoliubov case ($p \neq 0$), the situation is more complicated since $\mathbf{N}$ does *not* commute with $\mathbf{F}$. Let us consider the real function

$$\boxed{\nu_{\mathbf{F}} : \mu \mapsto \nu_{\mathbf{F}}(\mu) = \frac{\operatorname{Tr} \mathbf{N} \, \mathbb{1}_{(-\infty,0)}\big(\mathbf{F} - \mu \mathbf{N}\big) + N_b}{2}.} \tag{3.3.24}$$

We are interested in solving the equation

$$\nu_{\mathbf{F}}(\mu) = N/2.$$

In the Hartree-Fock case, $\nu_{\mathbf{F}}$ is a non-decreasing piecewise constant function. There is a solution μ to $\nu_{\mathbf{F}}(\mu') = N/2$ when $N/2$ belong to the range of $\nu_{\mathbf{F}}$. Otherwise, one has to partially fill a shell using the matrix δ.

In the Hartree-Fock-Bogoliubov case, $\nu_{\mathbf{F}}$ is also non-decreasing and in general it is much smoother when $p \neq 0$. The following lemma summarizes some important properties of $\nu_{\mathbf{F}}$ in both the HF and HFB cases.

Lemma 3.3.1 (Elementary properties of $\nu_{\mathbf{F}}$)**.** *Let* $\mathbf{F}$ *be as in* (3.3.22)*. Then the function* $\nu_{\mathbf{F}}$ *defined in* (3.3.24) *is increasing with respect to* μ*. It can only have finitely many jumps. It satisfies for some constant* C *depending only on* N_b

- $\nu_{\mathbf{F}}(\mu) \leq C/\mu$ *for* $\mu \leq -C$;
- $\nu_{\mathbf{F}}(\mu) \geq N_b - C/\mu$ *for* $\mu \geq C$.

If $0 \notin \sigma\big(\mathbf{F} - \mu\mathbf{N}\big)$, *then*

$$\frac{d\nu_{\mathbf{F}}}{d\mu}(\mu) = 2 \sum_{\substack{\epsilon_i<0 \\ \epsilon_j>0}} \frac{|\langle v_j, \mathbf{N} v_i\rangle|^2}{\epsilon_j - \epsilon_i} \geq 0 \tag{3.3.25}$$

where $(\mathbf{F} - \mu\mathbf{N})v_i = \epsilon_i\, v_i$.

Proof. The behavior of $\nu_{\mathbf{F}}$ for $|\mu| \gg 1$ was already studied in Lemma 3.1.2.

The matrix $\mathbf{F} - \mu\mathbf{N}$ is a linear function of $\mu \in \mathbb{R}$, hence by [24], we know that its eigenvalues form a set of real analytic functions. They cannot be constant because the matrix $\mathbf{N}$ does not vanish. The eigenvalues of $\mathbf{F} - \mu\mathbf{N}$ all behave like $\pm\mu$ for large μ, by perturbation theory. We conclude that 0 can be an eigenvalue of $\mathbf{F} - \mu\mathbf{N}$ for a finite number of μ's, say $\mu_1 < \cdots < \mu_K$. On the other hand, the map $\mu \mapsto \mathbb{1}_{(-\infty,0)}\big(\mathbf{F} - \mu\mathbf{N}\big)$ is real-analytic outside of the μ_k's (see [24]). So $\nu_{\mathbf{F}}$ is itself real-analytic outside of this set and it can have at most a finite number of jumps.

Once we know that the derivative of $\nu_{\mathbf{F}}$ is given by (3.3.25), the fact that $d\nu_{\mathbf{F}}/d\mu \geq 0$ proves that $\nu_{\mathbf{F}}$ is increasing with respect to μ, in between the μ_k's. That the jumps are all positive can be easily seen by an approximation argument using Lemma 3.3.2 below. We skip the details. This concludes the proof of Lemma 3.3.1.

The formula (3.3.25) of $\frac{d\nu_{\mathbf{F}}}{d\mu}$ can be obtained by usual perturbation theory, as we now explain in detail. In order to simplify the notation, we introduce $A = \mathbf{F} - \mu\mathbf{N}$ for some μ. We want to expand $\nu_{\mathbf{F}}(\mu + \delta\mu)$ for some $\delta\mu << 1$. Recall that

$$\begin{aligned}\nu_{\mathbf{F}}(\mu + \delta\mu) &= \frac{N_b + \mathrm{Tr}\big(\mathbb{1}_{(-\infty,0)}(\mathbf{F} - (\mu + \delta\mu)\mathbf{N})\mathbf{N}\big)}{2} \\ &= \frac{N_b}{2} + \frac{1}{2}\mathrm{Tr}(\mathbf{N}\mathbb{1}_{(-\infty,0)}(A - \delta\mu\mathbf{N})).\end{aligned}$$

If we diagonalize A in the form

$$A = \sum_{i=1}^{d} \epsilon_i |v_i\rangle\langle v_i|$$

then

$$\mathbb{1}_{(-\infty,0)}(A) = \sum_{\epsilon_i<0} \epsilon_i |v_i\rangle\langle v_i|.$$

Our assumption that $\mu \notin \{\mu_k\}$ means that $0 \notin \sigma(A)$. So $\epsilon_i \neq 0 \quad \forall i$. Let us consider a simple closed curve $\mathcal{C}$ in $\mathbb{C}$, which intersects the real axis and encloses

all the negative eigenvalues of A (andonly the negative ones). We remark that the residuum formula gives

$$\frac{1}{2i\pi}\oint_{\mathcal{C}}\frac{dz}{z-\epsilon}=\begin{cases}1 & \text{if} \quad \epsilon\in\sigma(A)\cap\mathbb{R}^-\\ 0 & \text{if} \quad \epsilon\notin\sigma(A)\cap\mathbb{R}^-\end{cases}$$

From this we deduce that

$$\mathbb{1}_{(-\infty,0)}(A)=\frac{1}{2i\pi}\oint_{\mathcal{C}}dz(z-A)^{-1}.$$

Because

$$\frac{1}{z-A}=\sum_{i=1}^{d}\frac{1}{z-\epsilon_i}|v_i\rangle\langle v_i|,$$

hence

$$\begin{aligned}\frac{1}{2i\pi}\oint_{\mathcal{C}}\frac{dz}{z-A}&=\sum_{i=1}^{d}\left(\frac{1}{2i\pi}\oint_{\mathcal{C}}\frac{dz}{z-\epsilon_i}\right)|v_i\rangle\langle v_i|,\\&=\sum_{\epsilon_i<0}|v_i\rangle\langle v_i|.\end{aligned}$$

We now considere that $B=-\delta\mu\mathbf{N}$. For B small enough (that is, for $\delta\mu$ small), we have similarly

$$\mathbb{1}_{(-\infty,0)}(A+B)=\frac{1}{2i\pi}\oint_{\mathcal{C}}dz(z-A-B)^{-1},$$

and this

$$\mathbb{1}_{(-\infty,0)}(A+B)-\mathbb{1}_{(-\infty,0)}(A)=\frac{1}{2i\pi}\oint_{\mathcal{C}}\left(\frac{1}{z-(A+B)}-\frac{1}{z-A}\right)dz.$$

Now we use that

$$\frac{1}{z-A-B}-\frac{1}{z-A}=\frac{1}{z-A}B\frac{1}{z-A}+\underbrace{\frac{1}{z-A}B\frac{1}{z-A}B\frac{1}{z-A-B}}_{\mathcal{O}(\|B\|^2)}.$$

So,

$$\mathbb{1}_{(-\infty,0)}(A+B)-\mathbb{1}_{(-\infty,0)}(A)=\frac{1}{2i\pi}\oint_{\mathcal{C}}\left(\frac{1}{z-A}B\frac{1}{z-A}+\mathcal{O}(\|B\|^2)\right)dz.$$

To compute the first-order term, we use the decomposition:

$$1=\Pi^+ +\Pi^- \text{ with } \Pi^+(A)=\sum_{a_i>0}|v_i\rangle\langle v_i| \text{ and } \Pi^-(A)=\sum_{a_i<0}|v_i\rangle\langle v_i|$$

and the spectral decomposition of A, and we find

$$\begin{aligned}\frac{1}{z-A} &= \sum_{\epsilon_i>0} \frac{1}{z-\epsilon_i}|v_i\rangle\langle v_i| + \sum_{\epsilon_i<0} \frac{1}{z-\epsilon_i}|v_i\rangle\langle v_i|,\\ B\frac{1}{z-A} &= \sum_{\epsilon_i>0} \frac{1}{z-\epsilon_i}|Bv_i\rangle\langle v_i| + \sum_{\epsilon_i<0} \frac{1}{z-\epsilon_i}|Bv_i\rangle\langle v_i|,\end{aligned}$$

so

$$\begin{aligned}&\frac{1}{z-A}B\frac{1}{z-A}\\ =&\sum_{\substack{\epsilon_i>0\\ \epsilon_j>0}} \frac{1}{(z-\epsilon_i)(z-\epsilon_j)}\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i| + \sum_{\substack{\epsilon_i<0\\ \epsilon_j>0}} \frac{1}{(z-\epsilon_i)(z-\epsilon_j)}\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i|,\\ &+\sum_{\substack{\epsilon_i>0\\ \epsilon_j<0}} \frac{1}{(z-\epsilon_i)(z-\epsilon_j)}\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i| + \sum_{\substack{\epsilon_i<0\\ \epsilon_j<0}} \frac{1}{(z-\epsilon_i)(z-\epsilon_j)}\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i|.\end{aligned}$$

Because

$$\frac{1}{(z-\epsilon_i)(z-\epsilon_j)} = \frac{1}{\epsilon_i-\epsilon_j}\left(\frac{1}{z-\epsilon_i} - \frac{1}{z-\epsilon_j}\right),$$

the residuum formula gives

$$\frac{1}{2i\pi}\oint_{\mathcal{C}} \frac{dz}{(z-\epsilon_i)(z-\epsilon_j)} = \begin{cases} 0 & \text{if} \quad \epsilon_i \in \sigma(A)\cap\mathbb{R}^- \quad \text{and} \quad \epsilon_j \in \sigma(A)\cap\mathbb{R}^- \\ 0 & \text{if} \quad \epsilon_i \notin \sigma(A)\cap\mathbb{R}^- \quad \text{and} \quad \epsilon_j \notin \sigma(A)\cap\mathbb{R}^- \\ \dfrac{1}{\epsilon_i-\epsilon_j} & \text{if} \quad \epsilon_i \in \sigma\cap\mathbb{R}^-(A) \quad \text{and} \quad \epsilon_j \notin \sigma(A)\cap\mathbb{R}^- \\ \dfrac{1}{\epsilon_j-\epsilon_i} & \text{if} \quad \epsilon_i \notin \sigma(A)\cap\mathbb{R}^- \quad \text{and} \quad \epsilon_j \in \sigma(A)\cap\mathbb{R}^- \end{cases}$$

and we find

$$\begin{aligned}&\frac{1}{2i\pi}\oint_{\mathcal{C}}\left(\frac{1}{z-A}B\frac{1}{z-A}\right)dz\\ =&\frac{1}{2i\pi}\oint_{\mathcal{C}}\left(\sum_{\substack{\epsilon_i<0\\ \epsilon_j>0}} \frac{1}{(z-\epsilon_i)(z-\epsilon_j)}\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i| + \sum_{\substack{\epsilon_i>0\\ \epsilon_j<0}} \frac{1}{(z-\epsilon_i)(z-\epsilon_j)}\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i|\right)dz\\ =&\frac{1}{2i\pi}\oint_{\mathcal{C}}\left(\sum_{\substack{\epsilon_i<0\\ \epsilon_j>0}} \frac{1}{\epsilon_i-\epsilon_j}\left(\frac{1}{z-\epsilon_i}-\frac{1}{z-\epsilon_j}\right)\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i|\right.\\ &\left.+\sum_{\substack{\epsilon_i>0\\ \epsilon_j<0}} \frac{1}{\epsilon_i-\epsilon_j}\left(\frac{1}{z-\epsilon_i}-\frac{1}{z-\epsilon_j}\right)\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i|\right)dz.\end{aligned}$$

We obtain

$$\begin{aligned}
&\mathbb{1}_{(-\infty,0)}(A+B)-\mathbb{1}_{(-\infty,0)}(A)\\
&\qquad=\sum_{\substack{\epsilon_i<0\\ \epsilon_j>0}}\frac{1}{\epsilon_i-\epsilon_j}\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i|+\sum_{\substack{\epsilon_i>0\\ \epsilon_j<0}}\frac{-1}{\epsilon_i-\epsilon_j}\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i|+\mathcal{O}(\|B\|^2),\\
&\qquad=\sum_{\substack{\epsilon_i<0\\ \epsilon_j>0}}\frac{1}{\epsilon_i-\epsilon_j}\Big(\langle v_j|Bv_i\rangle|v_j\rangle\langle v_i|+\overline{\langle v_j|Bv_i\rangle}|v_i\rangle\langle v_j|\Big)+\mathcal{O}(\|B\|^2).
\end{aligned}$$

Let us recall that $B=-\delta\mu\mathbf{N}$. So the final answer is

$$\mathrm{Tr}\Big(\mathbf{N}(\mathbb{1}_{(-\infty,0)}(A-\delta\mu\mathbf{N})-\mathbb{1}_{(-\infty,0)}(A))\Big)=2\delta\mu\sum_{\substack{\epsilon_i<0\\ \epsilon_j>0}}\frac{|\langle v_j|\mathbf{N}v_i\rangle|^2}{\epsilon_j-\epsilon_i}+\mathcal{O}(\|B\|^2).$$

Therefore

$$\boxed{\frac{d\nu_{\mathbf{F}}}{d\mu}(\mu)=2\sum_{\substack{\epsilon_i<0\\ \epsilon_j>0}}\frac{\big|\langle v_j,\mathbf{N}v_i\rangle\big|^2}{\epsilon_j-\epsilon_i}.}\tag{3.3.26}$$

□

Remark 3.3.1. *In the case of rotation invariant states, let us recall that the constraint takes the form:*

$$\sum_{\ell=0}^{\ell_{\max}}\mathrm{Tr}S\ G_\ell=\frac{N}{2}.\tag{3.3.27}$$

In this section $S=I$. Then we have to change our definition of ν which becomes:

$$\begin{aligned}
\nu_{\mathbf{F}}(\mu)&=\sum_{\ell=0}^{\ell_{\max}}\nu_{\mathbf{F}^\ell}(\mu)(2\ell+1)\\
&=\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\frac{N_b+\mathrm{Tr}\big(\mathbb{1}_{(-\infty,0)}(\mathbf{F}^\ell-\mu\mathbf{N})\mathbf{N}\big)}{2}.
\end{aligned}$$

Of course we find using Lemma 3.3.1 that

$$\begin{aligned}
\nu'_{\mathbf{F}}(\mu)&=\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\nu'_{\mathbf{F}^\ell}(\mu)\\
&=\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\ 2\sum_{\substack{\epsilon^\ell_i<0\\ \epsilon^\ell_j>0}}\frac{\big|\big\langle v^\ell_j,\mathbf{N}v^\ell_i\big\rangle\big|^2}{\epsilon^\ell_j-\epsilon^\ell_i},
\end{aligned}$$

where $(\mathbf{F}^\ell-\mu\mathbf{N})v^\ell_i=\epsilon^\ell_i v^\ell_i,\quad\forall\ell=0,..\ell_{\max}$.

The shape of the function $\nu_{\mathbf{F}}$ is very different in the HF and HFB cases. For a Hartree-Fock state, the function $\nu_{\mathbf{F}}$ is piecewise constant and it has jumps at the eigenvalues $\epsilon_1 < \cdots < \epsilon_{N_b}$ of h_G. The size of the jumps is equal to the multiplicity of the associated eigenvalue. An HFB state will most always have a very smooth $\nu_{\mathbf{F}}$. Of course, the smaller p in the Hamiltonian $\mathbf{F}_\Upsilon$, the more $\nu_{\mathbf{F}}$ looks like a step function.

In Figure 3.2 below, we show the function $\nu_{\mathbf{F}}$ for different values of the pairing term. More precisely, we have randomly chosen two symmetric real matrices h and p of size $N_b = 5$, and we display the function $\nu_{\mathbf{F}}$ when the pairing is replaced by tp for $t = 0$ (Hartree-Fock case), $t = 0.1$ and $t = 1$. Figure 3.3 is a plot of the eigenvalues of $\mathbf{F} - \mu\mathbf{N}$ for $t = 0.1$, as functions of μ. Note that there are some crossings of eigenvalues above and below the real line (recall that the spectrum is symmetric with respect to 0). But, around 0 the crossings are avoided and there is a gap.

If we repeat the numerical experiment with several random matrices h and p, we *never* see any jump for $\nu_{\mathbf{F}}$. The purpose of the next result is to clarify this observation.

Lemma 3.3.2 (Generic behavior of $\nu_{\mathbf{F}}$)**.** *The Fock matrix $\mathbf{F} - \mu\mathbf{N}$ is invertible if and only if $h \pm ip - \mu$ are invertible. More precisely,*

$$\min \sigma\big(|\mathbf{F} - \mu\mathbf{N}|\big) = \min\left(\left\|(h + ip - \mu)^{-1}\right\|^{-1}, \left\|(h - ip - \mu)^{-1}\right\|^{-1}\right). \quad (3.3.28)$$

The set of real symmetric matrices h and p such that

$$\sigma\big(\mathbf{F} - \mu\mathbf{N}\big) \cap \{0\} = \emptyset \quad \textit{for all } \mu \in \mathbb{R}$$

is open and dense in $\{(h, p) \ : \ h = h^T = \overline{h},\ p = p^T = \overline{p}\}$. For h and p in this set, $\nu_{\mathbf{F}}$ is real-analytic on $\mathbb{R}$.

It is obvious that there are matrices h and p for which $\mathbf{F}-\mu\mathbf{N}$ has 0 as eigenvalue for some $\mu \in \mathbb{R}$. The simplest examples are HF Hamiltonians for which $p \equiv 0$ and $\mathbf{F}-\mu\mathbf{N}$ is not invertible each time μ equals an eigenvalue of h. If p does not vanish but commutes with h, then we have $|h+ip-\mu|^2 = |h-ip-\mu|^2 = (h-\mu)^2+p^2$ and we see that 0 is never in the spectrum of $\mathbf{F} - \mu\mathbf{N}$ when the kernel of p does not contain the eigenvectors of h. However there are counterexamples with p invertible not commuting with h. For instance, $\mathbf{F} + \mathbf{N}$ is not invertible for

$$h = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad p = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}.$$

We now turn to the proof of Lemma 3.3.2.

Proof. The operator $\mathbf{F} - \mu\mathbf{N}$ is unitarily equivalent to

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \big(\mathbf{F} - \mu\mathbf{N}\big) \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = -i \begin{pmatrix} 0 & h + ip - \mu \\ -(h - ip - \mu) & 0 \end{pmatrix}.$$

From this we deduce that $\mathbf{F}-\mu\mathbf{N}$ is invertible if and only if $h+ip-\mu$ and $h-ip-\mu$ are invertible. Then we have

$$\big(\mathbf{F}-\mu\mathbf{N}\big)^{-1} = i\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 0 & -(h-ip-\mu)^{-1} \\ (h+ip-\mu)^{-1} & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

and

$$\Big\|\big(\mathbf{F}-\mu\mathbf{N}\big)^{-1}\Big\| = \max\Big(\,\big\|(h+ip-\mu)^{-1}\big\|\;,\;\big\|(h-ip-\mu)^{-1}\big\|\,\Big).$$

The statement now follows from the fact that, on a dense open set, the spectra of $h \pm ip$ do not intersect the real axis. □

Lemma 3.3.2 is interesting when we apply the Roothaan or the ODA, because it means that, as soon as $p \neq 0$, most often we will have no choice for μ_{n+1} and we will take $\delta_{n+1} = 0$. Saying differently, it is really reasonable to assume that the sequence generated by the Roothaan and the ODA are uniformly well posed (of course when the final state is believed to have a non vanishing pairing), as we did in Theorem 3.1.1 and 3.2.1.

Even if it is in general smooth, the function $\nu_{\mathbf{F}}$ can still vary quickly and this will be the case when the pairing term p is small. The appropriate method to find the solution of $\nu_{\mathbf{F}}(\mu) = N/2$ then depends on the properties of $\nu_{\mathbf{F}}$. If the Hamiltonian $\mathbf{F}$ has a large enough pairing matrix p, then $\nu_{\mathbf{F}}$ is smooth and we can use a Newton-like method to solve the equation $\nu_{\mathbf{F}}(\mu) = N/2$. A trial chemical potential μ^0 being given, we compute the derivative $\partial\nu_{\mathbf{F}}/\partial\mu(\mu^0)$ using Formula (3.3.25) and then let

$$\mu^1 := \mu^0 + \big(N/2 - \nu_{\mathbf{F}}(\mu^0)\big)\left(\frac{\partial\nu_{\mathbf{F}}}{\partial\mu}(\mu^0)\right)^{-1}.$$

The method can be iterated until convergence of μ^n towards the desired μ. The convergence is very fast, as soon as $\nu_{\mathbf{F}}$ is smooth.

If the Hamiltonian $\mathbf{F}$ has a small pairing matrix p, the function $\nu_{\mathbf{F}}$ will be smooth but close to a step function. Its derivative varies very quickly and the previous Newton method is not appropriate. In this case we can use a simple bisection method. The bounds on $\nu_{\mathbf{F}}(\mu)$ for large $|\mu|$ can be used to find a good starting interval $[\mu_l, \mu_r]$ such that $\nu_{\mathbf{F}}(\mu_l) < N/2$ and $\nu_{\mathbf{F}}(\mu_r) > N/2$.

We have to find a new μ_{n+1} at each step of the Roothaan or ODA. It is of course not efficient to find μ_{n+1} with a very high precision all along the algorithm. Dechargé and Gogny advice in Section II.E of [10] to apply the Newton scheme only once at each step. This means that

$$\mu_{n+1} = \mu_n + \big(N/2 - \nu_n(\mu_n)\big)\left(\frac{\partial\nu_n}{\partial\mu}(\mu_n)\right)^{-1}$$

where ν_n is the function ν corresponding to $\mathbf{F} = \mathbf{F}_{\Upsilon_n}$. This is then the same as doing perturbation theory at first order. We use a slightly different strategy which we explain in the next chapter.

Remark 3.3.2. *For rotation-invariant states, the multiplier μ is updated using the formula*

$$\mu' = \frac{\mu + N/2 - \sum_{\ell=0}^{\ell_{\max}} (2\ell+1)\nu_{\mathbf{F}^\ell}(\mu)}{\sum_{\ell=0}^{\ell_{\max}} (2\ell+1)\partial\nu_{\mathbf{F}^\ell}/\partial\mu(\mu)}. \tag{3.3.29}$$

See Remark 3.3.1.

Remark 3.3.3. ***Link with Hartree-Fock theory in dimension*** $2N_b$**.** *In order to emphasize the abstract structure of the HFB model, we write the energy as follows*

$$\mathcal{E}(\Upsilon) = \mathrm{Tr}\left(\mathbf{h}(\Upsilon - \mathbf{P})\right) + \frac{1}{2}\mathrm{Tr}(\Upsilon - \mathbf{P})\mathcal{L}(\Upsilon - \mathbf{P}) \tag{3.3.30}$$

where

$$\mathbf{h} = \begin{pmatrix} h & 0 \\ 0 & -h \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and where $\mathcal{L}$ is a linear operator defined on the space of $(2N_b) \times (2N_b)$ symmetric real matrices, by

$$\mathcal{L}\begin{pmatrix} M_1 & M_2 \\ M_2 & M_3 \end{pmatrix} = \begin{pmatrix} 2J(M_1) - K(M_1) & K(M_2) \\ K(M_2) & 2J(M_3) - K(M_3) \end{pmatrix}.$$

The only important property of $\mathcal{L}$ for the following is the symmetry

$$\mathbf{U}\mathcal{L}(M)\mathbf{U}^{-1} = \mathcal{L}(\mathbf{U}M\mathbf{U}^{-1}), \qquad \textit{where} \quad \mathbf{U} := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.3.31}$$

All the results of this chapter apply exactly the same when h, $\mathcal{L}$, $\mathbf{P}$, $\mathbf{U}$, $\mathbf{N}$ are abstract operators satisfying (3.3.31)*,*

$$\mathbf{U}\mathbf{h}\mathbf{U}^{-1} = -\mathbf{h}$$

and we can assume that $\mathbf{h}$ is any matrix satisfying this property. The constraint that Υ takes the form (3.0.1) *can be written in terms of the matrices $\mathbf{P}$ and $\mathbf{U}$ as*

$$\mathbf{U}\big(\Upsilon - \mathbf{P}\big)\mathbf{U}^{-1} = -\big(\Upsilon - \mathbf{P}\big) \quad \Longleftrightarrow \quad \mathbf{U}\Upsilon\mathbf{U}^{-1} = I_{2N_b} - \Upsilon. \tag{3.3.32}$$

Finally, the particle number constraint $\mathrm{Tr}(G) = N/2$ can also be written as

$$\mathrm{Tr}\left(\mathbf{N}(\Upsilon - \mathbf{P})\right) = N. \tag{3.3.33}$$

As a conclusion, the HFB discretized minimization problem can be rewritten as

$$I_h(N) = \min\Big\{ \mathrm{Tr}\left(\mathbf{h}(\Upsilon - \mathbf{P})\right) + \frac{1}{2}\mathrm{Tr}(\Upsilon - \mathbf{P})\mathcal{L}(\Upsilon - \mathbf{P}) \ : 0 \le \Upsilon \le 1, \\ \mathbf{U}\Upsilon\mathbf{U}^{-1} = I_{2N_b} - \Upsilon, \ \mathrm{Tr}\left(\mathbf{N}\Upsilon\right) = N - N_b \Big\}. \tag{3.3.34}$$

As we have already mentioned, this problem takes the same form as the Hartree-Fock minimization problem [8], with two main differences:

- *we have the symmetry constraint* $\mathbf{U}\Upsilon\mathbf{U}^{-1} = I_{2N_b} - \Upsilon$ *on the density matrix* Υ*;*
- *we have a constraint on* $\mathrm{Tr}(\mathbf{N}\Upsilon)$ *instead of* $\mathrm{Tr}(\Upsilon)$*.*

We can consider the abstract minimization problem of the form (3.3.34)*. We assume that* $\mathbf{P}$ *is a given projection of rank* N_b*, and that* $\mathbf{U}$ *is an isometry such that* $\mathbf{U}\mathbf{P}\mathbf{U}^{-1} = 1 - \mathbf{P}$*. We also assume that*

$$\mathbf{U}\mathbf{h}\mathbf{U}^{-1} = -\mathbf{h}, \qquad \mathbf{U}\mathbf{N}\mathbf{U}^{-1} = -\mathbf{N} \quad \textit{and} \quad \mathbf{U}\mathcal{L}(\cdot)\mathbf{U}^{-1} = \mathcal{L}(\mathbf{U}\cdot\mathbf{U}^{-1}). \tag{3.3.35}$$

The solution of this minimization problem is

$$\Upsilon = \mathbb{1}_{(-\infty,0)}\Big(\mathbf{F}_\Upsilon - \mu\mathbf{N}\Big) + \delta,$$

with $\mathrm{Tr}(\mathbf{N}\Upsilon) = \dfrac{N}{2} - N_b$*.*

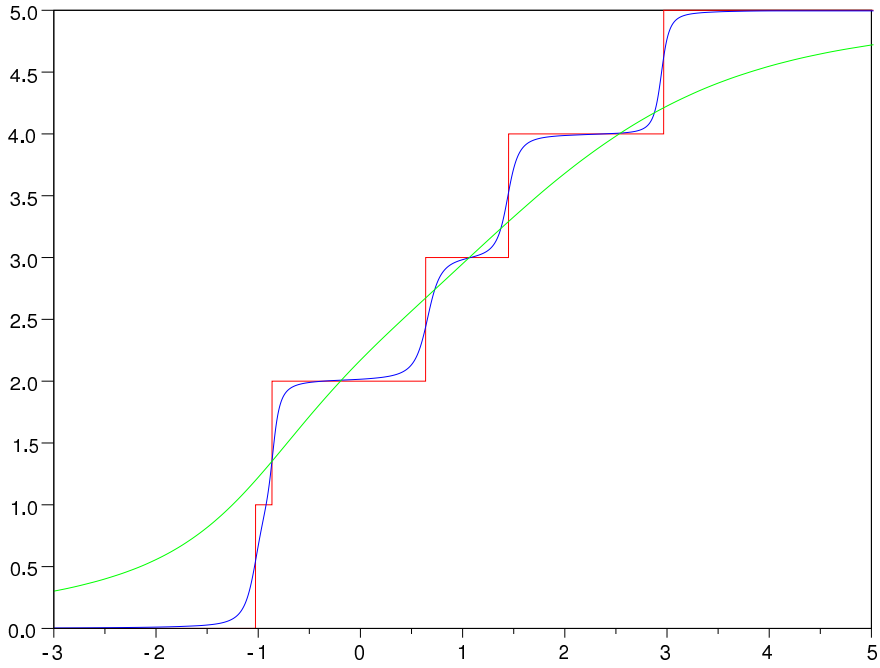

Figure 3.2: The function $\nu_{\mathbf{F}}(\mu)$ which gives the average number of particles in the state $\mathbb{1}_{(-\infty,0)}(\mathbf{F}-\mu\mathbf{N})$, in terms of the chemical potential μ. The pairing term in $\mathbf{F}$ is equal to tp with $t = 0$ (Hartree-Fock case, red curve), $t = 0.1$ (blue curve) and $t = 1$ (green curve).

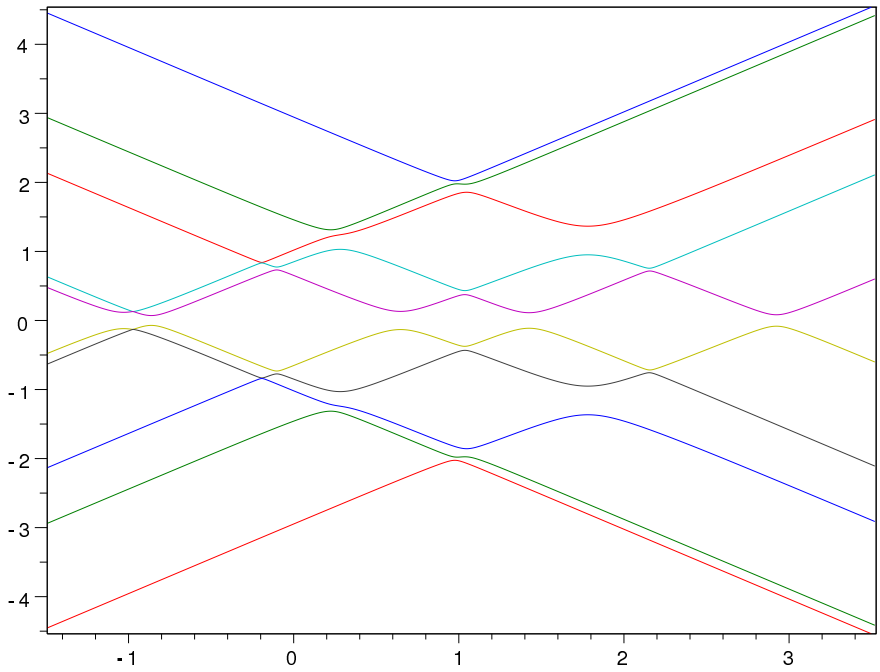

Figure 3.3: The eigenvalues of $\mathbf{F}-\mu\mathbf{N}$ in terms of μ for $t = 0.1$.

Part II

Numerical simulation of some real systems

Chapter 4

Non relativistic gravitational systems

4.1 Model

Here we consider a system of N spin-1/2 neutral non-relativistic particles, only interacting through the Newtonian interaction

$$W(x) = -\frac{g}{|x|}, \qquad g > 0. \tag{4.1.1}$$

This potential is strongly attractive at short distances. Since $1/|x|$ does not decay too fast at infinity, it is also quite attractive at large distances. The kinetic energy does not scale the same as the potential energy. By a simple scaling argument, we can therefore always assume that

$$g \equiv 1.$$

This model can be used to describe neutron stars and white dwarfs when $N \gg 1$. It has been particularly studied from a theoretical point of view in the pseudo-relativistic case where the kinetic energy is given by $T = \sqrt{c^4 m^2 - c^2 \Delta} - mc^2$, see [26, 35, 36]. In our simulations we restrict ourselves to the non-relativistic case of the Laplacian $T = -\Delta/(2m)$ (in units such that $m = 1/2$). It would be interesting to take N large but this is of course much too difficult from a numerical point of view.

As mentioned in the previous chapters, we always impose the spin and time-reversal symmetries, which is perfectly justified for the ground state since the interaction (4.1.1) satisfies the assumption of the Bach-Fröhlich-Jonsson Theorem 1.4.2. We also impose spherical symmetry which, on the contrary, is not known to hold for the true ground state.

One advantage of the Newtonian interaction (4.1.1) is that the matrices h, S, and the operators J and $K^{\ell\ell'}$ can be explicitly computed in the basis of "hat"

functions. We have shown in Lemma 2.3.2 that the energy can be written

$$\mathcal{E}(G^0,...,G^{\ell_{\max}},A^0,...,A^{\ell_{\max}}) = 2\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\,\mathrm{Tr}(h^\ell G^\ell)$$
$$+\sum_{\ell,\ell'=0}^{\ell_{\max}}(2\ell+1)(2\ell'+1)\Big(2\,\mathrm{Tr}(G^\ell\,J(G^{\ell'}))-\mathrm{Tr}(G^\ell\,K^{\ell\ell'}(G^{\ell'}))+\mathrm{Tr}(A^\ell\,K^{\ell\ell'}(A^{\ell'}))\Big), \tag{4.1.2}$$

where

$$h^\ell_{ij}=\int_0^\infty \chi_i'(r)\chi_j'(r)r^2dr+\ell(\ell+1)\int_0^\infty\chi_i(r)\chi_j(r)dr,$$
$$J(G^{\ell'})_{ij}:=\sum_{m,n=0}^{N_b}(ij|nm)_{0,0}\,G^{\ell'}_{mn},\qquad K^{\ell\ell'}(G^{\ell'})_{ij}:=\sum_{m,n=0}^{N_b}(im|jn)_{\ell,\ell'}\,G^{\ell'}_{mn},$$
$$(ij|mn)_{\ell,\ell'}:=\int_0^\infty r^2\,dr\int_0^\infty s^2\,ds\;\chi_i(r)\,\chi_j(r)\,\chi_m(s)\,\chi_n(s)\,w_{\ell,\ell'}(r,s)$$

and

$$w_{\ell,\ell'}(r,s):=\frac{1}{2}\int_{-1}^1 W\left(\sqrt{r^2+s^2-2rst}\right)\,P_\ell(t)\,P_{\ell'}(t)\,dt$$
$$=-\frac{1}{2}\int_{-1}^1\frac{P_\ell(t)\,P_{\ell'}(t)}{\sqrt{r^2+s^2-2rst}}\,dt.$$

Using the well-known formula

$$\frac{1}{\sqrt{r^2+s^2-2rst}}=\sum_{n=0}^\infty\frac{\min(r,s)^n}{\max(r,s)^{n+1}}P_n(t)$$

we deduce that

$$(ij|mn)_{\ell,\ell'}$$
$$=-\frac{1}{2}\sum_{n=0}^\infty\Big(\int_{-1}^1 P_nP_\ell P_{\ell'}\Big)\int_0^\infty r^2\,dr\int_0^\infty s^2\,ds\frac{\min(r,s)^n}{\max(r,s)^{n+1}}\chi_i(r)\,\chi_j(r)\,\chi_m(s)\,\chi_n(s). \tag{4.1.3}$$

The integral over the Legendre polynomials is related to the usual Clebsch-Gordan coefficients as follows

$$\frac{1}{2}\int_{-1}^1 P_n(t)\,P_\ell(t)\,P_{\ell'}(t)\,dt=\begin{pmatrix}\ell&\ell'&n\\0&0&0\end{pmatrix}^2$$

and only a finite number of terms are non zero in the sum over n in (4.1.3).

The previous argument shows that the HFB energy can be written as follows

Corollary 4.1.1 (Gravitational HFB energy for all $\ell_{\max}$)**.** *The gravitational HFB energy functionnal associated with the density and pairing matrices* $(G^\ell, A^\ell)_{0\le\ell\le\ell_{\max}}$ *is*

$$\begin{aligned}
&\mathcal{E}(G^0,...,G^\ell,A^0,...,A^\ell)\\
&\quad= 2\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\mathrm{Tr}(h^\ell G^\ell) - \sum_{\ell,\ell'=0}^{\ell_{\max}}(2\ell+1)(2\ell'+1)\Bigg(2\mathrm{Tr}(G^\ell\mathcal{J}(G^{\ell'}))\\
&\qquad-\sum_{m=0}^{\infty}\begin{pmatrix}\ell & \ell' & m\\ 0 & 0 & 0\end{pmatrix}^2\mathrm{Tr}(G^\ell\mathcal{K}^m(G^{\ell'}))+\sum_{m=0}^{\infty}\begin{pmatrix}\ell & \ell' & m\\ 0 & 0 & 0\end{pmatrix}^2\mathrm{Tr}(A^\ell\mathcal{K}^m(A^{\ell'}))\Bigg),
\end{aligned}$$

where

$$\big(\mathcal{J}(G^{\ell'})\big)_{ij} := \sum_{k,l=1}^{N_b} G^{\ell'}_{kl}\big(ij,kl\big)_0, \tag{4.1.4}$$

$$\Big(\mathcal{K}^m(G^{\ell'})\Big)_{ij} := \sum_{k,l=1}^{N_b} G^{\ell'}_{kl}\big(ik,jl\big)_m, \tag{4.1.5}$$

and

$$\big(ij,kl\big)_m := \int_0^\infty\int_0^\infty \chi_i(r)\chi_j(r)\chi_k(s)\chi_l(s)\frac{\min(r,s)^m}{\max(r,s)^{m+1}}r^2s^2drds. \tag{4.1.6}$$

For the sake of clarity we have defined new operators $\mathcal{J}$ and $\mathcal{K}$ which coincide with J and K up to a minus sign. We thereby explicit the sign of the interaction.

In our computations we have taken $\ell_{\max}=1$. So we give the complete expression of the energy in this special case:

$$\begin{aligned}
&\mathcal{E}(G^0,G^1,A^0,A^1) = 2\mathrm{Tr}(h^0G^0)+6\mathrm{Tr}(h^1G^1)\\
&\quad-2\Big(\mathrm{Tr}(G^0\mathcal{J}(G^0))+6\mathrm{Tr}(G^0\mathcal{J}(G^1))+9\mathrm{Tr}(G^1\mathcal{J}(G^1))\Big)\\
&\quad+\left(\mathrm{Tr}(G^0\mathcal{K}^0(G^0))+2\mathrm{Tr}(G^0\mathcal{K}^1(G^1))+\frac{6}{5}\mathrm{Tr}(G^1\mathcal{K}^2(G^1))+3\mathrm{Tr}(G^1\mathcal{K}^0(G^1))\right)\\
&\quad-\left(\mathrm{Tr}(A^0\mathcal{K}^0(A^0))+2\mathrm{Tr}(A^0\mathcal{K}^1(A^1))+\frac{6}{5}\mathrm{Tr}(A^1\mathcal{K}^2(A^1))+3\mathrm{Tr}(A^1\mathcal{K}^0(A^1))\right).
\end{aligned}$$

This expression is a consequence of the formulas

$$\begin{pmatrix}1 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}^2 = \begin{pmatrix}1 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}^2 = \frac{1}{2}\int_{-1}^1 P_1(t)^2\,dt = \frac{1}{2}\int_{-1}^1 t^2\,dt = \frac{1}{3},$$

$$\begin{pmatrix}1 & 1 & 2\\ 0 & 0 & 0\end{pmatrix}^2 = \frac{1}{2}\int_{-1}^1 P_1(t)^2\,P_2(t)\,dt = \frac{1}{2}\int_{-1}^1 t^2\,\frac{3t^2-1}{2}\,dt = \frac{2}{15}.$$

Any minimizer (G_0,G_1,A_0,A_1) of $\mathcal{E}$ under the constraints

$$0\le\Upsilon^\ell\mathbf{S}\Upsilon^\ell\le\Upsilon^\ell,\ \text{with}\ \Upsilon^\ell := \begin{pmatrix}G^\ell & A^\ell\\ A^\ell & S^{-1}-G^\ell\end{pmatrix}\ \text{and}\ \mathbf{S}=\begin{pmatrix}S & 0\\ 0 & S\end{pmatrix} \tag{4.1.7}$$

for all $\ell = 0, 1$, and

$$\boxed{2\,\mathrm{Tr}(SG^0) + 6\,\mathrm{Tr}(SG^1) = N} \tag{4.1.8}$$

can be written as

$$\begin{cases} \Upsilon^0 = \sum_{\epsilon_i^0<0} f_i^0 \big(f_i^0\big)^T \\ \Upsilon^1 = \sum_{\epsilon_i^1<0} f_i^1 \big(f_i^1\big)^T \end{cases}$$

where the $f_i^{0,1}$ solve the generalized eigenvalue problem

$$\begin{cases} \big(\mathbf{F}^0 - \mu\mathbf{N}\big) f_i^0 = \epsilon_i^0\, \mathbf{S}\, f_i^0, \qquad \big\langle f_i^0, \mathbf{S} f_j^0 \big\rangle = \delta_{ij}, \\ \big(\mathbf{F}^1 - 3\mu\mathbf{N}\big) f_i^1 = \epsilon_i^1\, \mathbf{S}\, f_i^1, \qquad \big\langle f_i^1, \mathbf{S} f_j^1 \big\rangle = \delta_{ij}. \end{cases} \tag{4.1.9}$$

Here $\mathbf{F}^0$ and $\mathbf{F}^1$ are the Fock matrices associated with G^0 and G^1:

$$\mathbf{F}^0 = \begin{pmatrix} H_G^0 & \Pi_A^0 \\ \Pi_A^0 & -H_G^0 \end{pmatrix} \quad \text{and} \quad \mathbf{F}^1 = \begin{pmatrix} H_G^1 & \Pi_A^1 \\ \Pi_A^1 & -H_G^1 \end{pmatrix}$$

where

$$H_G^0 = 2h^0 - 2(2J(G_0) + 6J(G_1)) + 2(K^0(G_0) + K^1(G_1)),$$

$$H_G^1 = 6h^1 - 2(6J(G_0) + 18J(G_1)) + 2\left(K^1(G_0) + \frac{18}{15}K^2(G_1) + 3K^0(G_1)\right),$$

$$\Pi_A^0 = -2(K^0(A_0) + K^1(A_1)),$$

and

$$\Pi_A^1 = -2\left(K^1(A_1) + \frac{18}{15}K^2(A_1) + 3K^0(A_1)\right).$$

4.2 Method

To simulate our physical system, we have used the open source software Scilab [40].

We choose a simple basis set $(\chi_1, ..., \chi_{N_b})$ of $L^2([0,\infty), r^2dr)$, made of "hat functions" associated with a chosen grid

$$0 = r_0 < r_1 < \cdots < r_{N_b} < r_{N_b+1} := r_{\max}.$$

We impose Dirichlet boundary conditions at $r_{\max}$. We have tested different types of grids and there was no important difference between them. The results presented here are all with regular grids. As we will explain later, for a given basis size N_b, the results usually depend a lot on the value of the radius $r_{\max}$ of the ball in which the system is placed.

In order to speed up the computation, we have calculated all the integrals $(ijkl)_m$ using Maple and we have used the explicit formulas in Scilab. These integrals are computed once and for all in the beginning of the algorithm and they are kept in memory.

Our main goal is to investigate the existence of pairing. We therefore always start by doing a precise Hartree-Fock calculation, for which we use the Optimal

Damping Algorithm described in Section 3.2. We take as initial state a simple uniform state

$$G_{\text{init}} = \frac{N}{2\,\text{Tr}(S)}\text{Id}_{N_b} \tag{4.2.10}$$

and we run HF until convergence. We have observed a global stability of the results with respect to initial states, hence the previous simple choice is appropriate (but more clever choices might decrease the total number of iterations). Then, we use the converged HF state G_{opt} as initial datum for the HFB algorithm. Of course we have to perturb it a little bit since any HF solution is also an HFB solution. We proceed as follows. Assuming that the overlap matrix $S = \text{Id}_{N_b}$ and that $\ell_{\max} = 0$ for simplicity, the optimal HF state G_{opt} can be written in the form

$$G_{\text{opt}} = \sum_{k=1}^{N/2} v_k v_k^T,$$

where v_k are the $N/2$ first eigenvectors of the mean-field matrix h,

$$hv_k = \epsilon_k\, v_k.$$

We then choose a number n_v of valence orbitals and a mixing parameter θ, and we perturb G_{opt} as follows

$$G'_{\text{init}} = \sum_{k=1}^{N/2-n_v} v_k v_k^T + \theta \sum_{k=N/2-n_v+1}^{N/2} v_k v_k^T + (1-\theta) \sum_{k=N/2+1}^{N/2+n_v} v_k v_k^T,$$

$$A'_{\text{init}} = \sqrt{\theta(1-\theta)} \sum_{k=N/2-n_v+1}^{N/2+n_v} v_k v_k^T.$$

In most cases, we have observed that $n_v = 1$ and $\theta = 0.95$ works perfectly well, that is, the algorithm escapes from the HF solution G_{opt} and converges towards an optimal HFB state. But other values of n_v and θ seem to work fine also.

When the maximum angular momentum $\ell_{\max}$ is larger than 0, we often first run the algorithm with $\ell_{\max} = 0$ for a few iterations before switching to the actual value of $\ell_{\max}$. We stop the algorithm when the commutators $[\mathbf{F}_n^\ell, \Upsilon_n^\ell]$ are smaller than a prescribed error. We know from (2.3.46) and (4.1.9) that these commutators must all vanish for an exact solution of the discretized HFB minimization problem. In terms of the matrix $\mathbf{S}$, the right quantity to look at is

$$\sum_{\ell=0}^{\ell_{\max}} \left\| \mathbf{S}^{-\frac{1}{2}} \big(\mathbf{F}_n^\ell \Upsilon_n^\ell \mathbf{S} - \mathbf{S} \Upsilon_n^\ell \mathbf{F}_n^\ell \big) \mathbf{S}^{-\frac{1}{2}} \right\|$$

where $\|\cdot\|$ is the usual operator norm for $(2N_b) \times (2N_b)$ matrices. There is a similar formula in the HF case [5].

As we have explained, in the HFB case, ensuring the constraint

$$2\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\mathrm{Tr}(SG^{\ell})=N$$

is not as easy as in the HF case. In the beginning of the algorithm, our state Υ is rather close to an HF state by construction. Therefore, the function $\nu_{\mathbf{F}}(\mu)$ defined in Section 3.3 is close to a step function. We choose an error ε and look for the next states Υ_{n+1}^{ℓ} having a total number of particles $\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\,\mathrm{Tr}(SG_{n+1}^{\ell})$ close to $N/2$, within the error ε, using a simple bisection method. We use the bisection for a fixed number of global iterations. Then, when the pairing term is large enough, we switch to a Newton method in order to find the state Υ_{n+1}. We have observed that even if in the beginning several Newton iterations can be employed at each step, usually only one Newton iteration is necessary after a while. To guarantee a good value of the average number of particles in the end, we decrease the error ε on $|\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\,\mathrm{Tr}(SG_{n+1}^{\ell})-N/2|$ along the algorithm.

4.3 Numerical results

In this section we present our numerical results for the purely gravitational system.

4.3.1 Roothaan vs ODA

In the HF case, we have observed that the Roothaan algorithm very often oscillates between two states, none of them being the solution of the problem (as described in Theorem 3.1.1 and in [7]). The Roothaan algorithm seems more well behaved in the HFB case. With the model presented in this section, we never got real oscillations for HFB. Sometimes the convergence is improved by using the ODA, but in most cases the Roothaan algorithm always converges towards the same state as the ODA in the end.

We start by comparing Roothaan and ODA in the HF case. There, oscillations seem to be related to the size of the gap between the largest filled eigenvalue and the smallest unfilled one. Indeed, oscillations in HF seem to only occur when there is pairing in HFB, an effect which is also well-known to be related to the size of the gap (see, e.g., Theorem 5 in [2]). When there is no pairing, the HF Roothaan algorithm always behaves like the ODA. However, the situation is complex and there is no exact rule. Sometimes the Roothaan algorithm does *not* oscillate even when the gap is rather small and there is pairing.

In Figure 4.1 we display the value of the energy obtained along the algorithm for the Roothaan and the ODA, for the following choice of parameters: $N=6$, $N_b=200$, $\ell_{\max}=0$ and $r_{\max}=30$. The ODA converges in about 17 iterations, whereas the Roothaan algorithm oscillates. We also show the value of the norms $\|G_n-G_{n-1}\|$ and $\|G_n-G_{n-2}\|$ along the Roothaan algorithm. The oscillation between two points is clearly demonstrated.

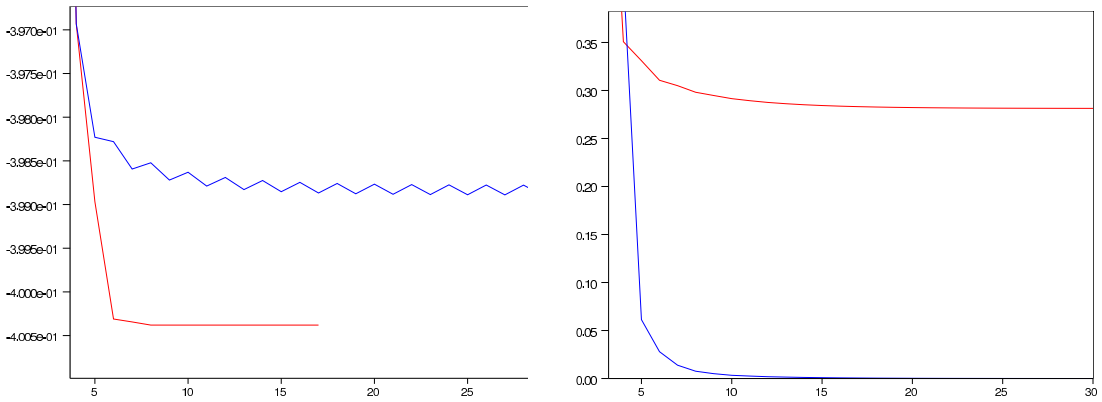

Figure 4.1: *Left:* HF energy along the iterations for the Roothaan Algorithm (blue) and the ODA (red). *Right:* Values of $\|G_n - G_{n-1}\|$ (red) and $\|G_n - G_{n-2}\|$ (blue) along the Roothaan algorithm, showing the oscillations between two states. Here $N = 6$, $N_b = 200$, $\ell_{\max} = 0$ and $r_{\max} = 30$.

When we decrease the parameter $r_{\max}$ but keep $N_b = 200$ constant, the gap is seen to increase slightly and the Roothaan algorithm behaves better. In Table 4.1, we give the numerical value of the last filled eigenvalue and the corresponding gap. The Roothaan algorithm slowly converges for $r_{\max} = 25$ and it coincides with the ODA when $r_{\max} = 20$. The gap for $r_{\max} = 20$ is 2.5 times the one for $r_{\max} = 30$. We will discuss the occurence of pairing in terms of the parameter $r_{\max}$ in the next section.

$r_{\max}$	$\epsilon_{N/2}$	$\epsilon_{N/2+1} - \epsilon_{N/2}$	behavior of HF Roothaan
20	-0.532430	0.159430	fast convergence
25	-0.536706	0.081016	slow convergence
30	-0.529200	0.061928	oscillations
	-0.548554	0.067422	

Table 4.1: Value of the last filled eigenvalue $\epsilon_{N/2}$ and the corresponding gap $\epsilon_{N/2+1} - \epsilon_{N/2}$ in HF, for $N = 6$, $N_b = 200$ and $\ell_{\max} = 0$. For $r_{\max} = 30$ the Roothaan algorithm oscillates and we display the last filled eigenvalue and the gap for the two states.

Lastly, in Figure 4.2 we compare the norm of the commutator $[G_n, h_{G_n}]$ for the Roothaan algorithm and the ODA for $\ell_{\max} = 1$. This is an example of convergence which is improved by using the ODA, even if the Roothaan algorithm does not oscillate.

As we have mentioned the Roothaan algorithm is usually much more well behaved in the HFB case. However, sometimes the convergence can be improved dramatically by using the ODA. In Figure 4.3 we display the energy along the iterations of the algorithm in both the Roothaan and ODA cases, for $N_b = 500$, $N = 16$, $\ell_{\max} = 1$ and $r_{\max} = 10$. In this case the Roothaan algorithm is very

badly behaved. It passes very close to the HF ground state and it takes it a very long time to escape from it. On the other hand, the ODA does not suffer from this problem and it converges much more rapidly.

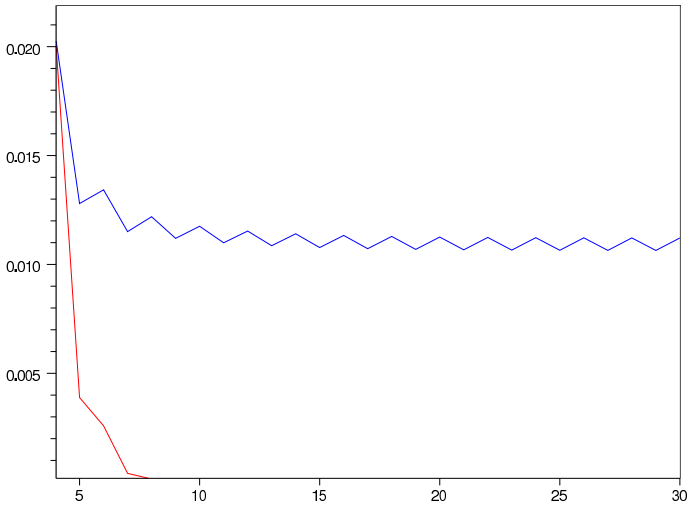

Figure 4.2: Value of $\|[G_n, h_{G_n}]\|$ along the Roothaan (blue) and the ODA (red), for $N = 6$, $N_b = 200$, $\ell_{\max} = 1$ and $r_{\max} = 30$.

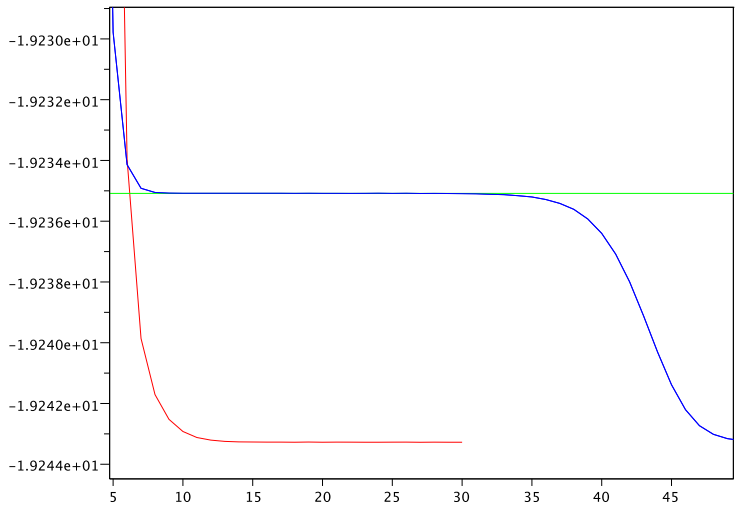

Figure 4.3: HFB energy along the iterations of the Roothaan (blue) and the ODA (red) for $N_b = 500$, $N = 16$, $\ell_{\max} = 1$ and $r_{\max} = 10$. The optimal HF energy is also displayed (green).

4.3.2 Numerical evidence of pairing

Pairing in a finite discretization basis might depend on the properties of the basis. As we have explained, the occurence of pairing is related to the size of the gap in HF theory and this gap varies with the radius $r_{\max}$ in which the system is confined. If $r_{\max}$ is decreased the system is more condensed and the HF gap increases.

In Figure 4.4 we display the HF and HFB ground state energies computed for $N = 16$ in a basis set of size $N_b = 200$, in terms of $r_{\max}$. The HF and HFB curves are distinct for $r_{\max}$ large enough and they merge at $r_{\max} = 6$ approximately. This observation is confirmed by the value of the norm of A plotted on the right of the same figure. The minima of the HF and HFB ground state energies are attained at about $r_{\max} \simeq 10$ in the HF case and $r_{\max} \simeq 10.5$ in the HFB case, which is sufficiently far from the merging point. The minima of these curves correspond to the best possible approximation for a given basis size N_b (here $N_b = 200$) and a given type of grid (here regular). The difference between the corresponding HF and HFB energies is significant. The HF ground state energy at $r_{\max} = 10$ is -19.232176 (in our units in which $m = 1/2$ and $e = 1$), whereas the HFB ground state energy at $r_{\max} = 10.5$ is -19.240176. The norm of the pairing matrix A is rather large at this point: $\|A\| = \sqrt{\mathrm{Tr}(SA_0SA_0) + 3\,\mathrm{Tr}(SA_1SA_1)} \simeq 0.462129$. This goes in favour of the conclusion that pairing really occurs for $N = 16$ in this model. This intuition is confirmed by a more precise calculations with $N_b = 500$ which we discuss below.

The observation of pairing requires to have an appropriate $r_{\max}$ but it does *not* require to have a very large basis set. Even for $N_b = 30$ and $r_{\max} = 10.5$, we already find that the HFB energy is approximately -19.078416 whereas the HF energy is about -19.072954. The corresponding norm of the pairing matrix A is $\|A\| \simeq 0.424124$.

Pairing is a subtle effect which decreases the energy by a small amount (much less than one percent here). Catching this effect requires to be very careful when choosing the radius $r_{\max}$. Taking $r_{\max}$ too small might lead to the conclusion that there is no pairing. In our simulations we have always observed the occurence of pairing, but provided we choose $r_{\max}$ appropriately. The values of $r_{\max}$ at which the HF and HFB energies attain their minimum were always found on the right of the merging point of the two curves. In Table 4.3 below we give our results for $N_b = 200$ and $N = 6, 10, 16$ and 20. The HFB ground state energy is always smaller than the HF energy.

In the paper [26], Lenzmann and Lewin have rigorously studied the gravitational model of this section. They showed the existence of a ground state in both the HF and HFB cases. But, so far, no proof that pairing occurs has been provided. The numerical results of this section tend to show that there is actually always pairing, at least for N not too large. Of course, a more systematic numerical study with different basis sets and explicit convergence rates must be realized in order to definitely conclude that pairing occurs in our model. As the energy shift due to the pairing is always rather small, this would however be a very challenging task.

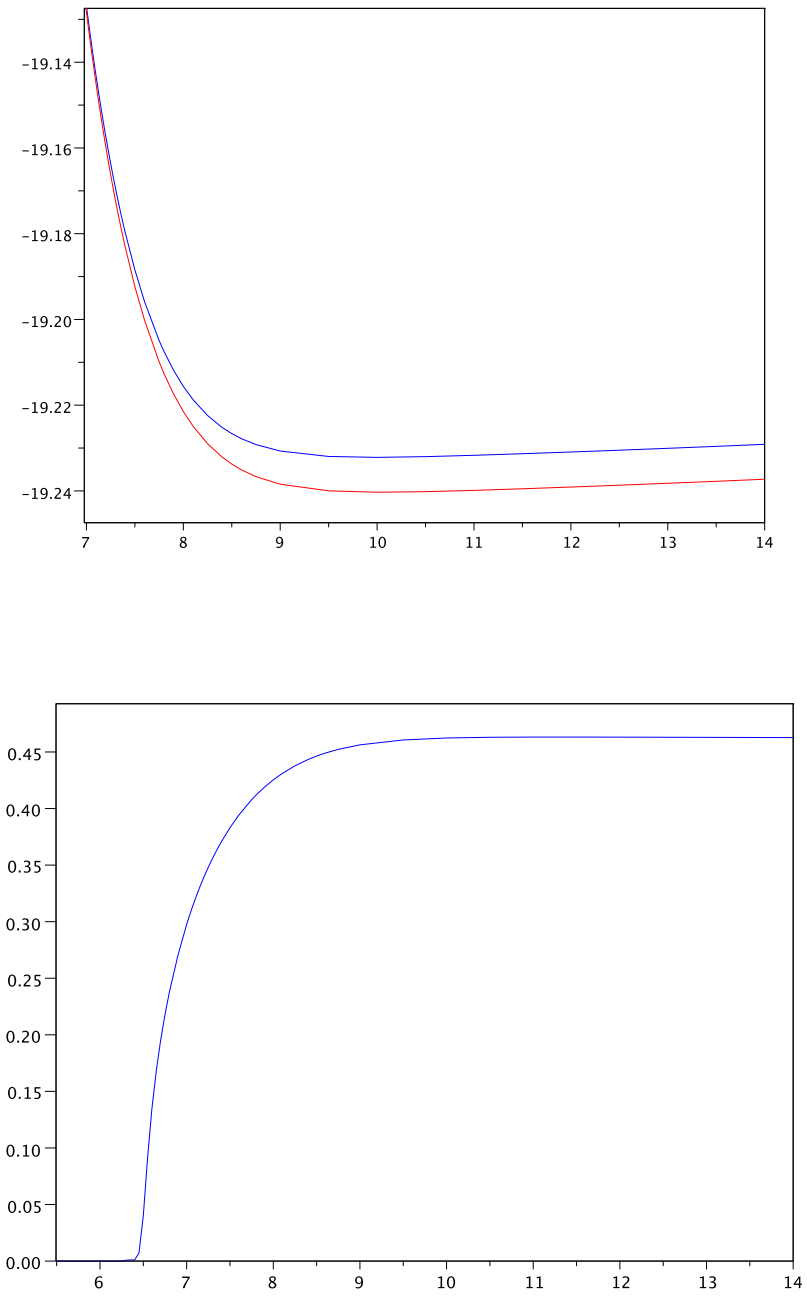

Figure 4.4: Value of the ground state HF and HFB energies (top) and of the norm of the pairing matrix A (bottom), as functions of the radius $r_{\max}$ in which the system is confined, for $N = 16$, $N_b = 200$ and $\ell_{\max} = 1$. The corresponding numerical values are given in Table 4.2 below.

$r_{\max}$	HF energy per spin	HFB energy per spin	$\|A\|$
4.6	-8.151868	-8.151868	
4.8	-8.463037	-8.463037	
5	-8.711979	-8.711979	
5.2	-8.910797	-8.910797	
5.4	-9.069147	-9.069147	
5.6	-9.194796	-9.194796	
5.8	-9.294013	-9.294013	
6	-9.371902	-9.371902	
6.2	-9.432627	-9.432628	
6,25	-9.445539	-9.445539	
6,30	-9.457641	-9.457641	
6,35	-9.468979	-9.468979	
6.4	-9.479593	-9.479593	0.001004
6,45	-9.489525	-9.489525	0.007149
6,50	-9.498811	-9.498811	0.040250
6,55	-9.50749	-9.507494	0.091990
6.6	-9.515596	-9.515626	0.135331
6,65	-9.523161	-9.523233	0.168454
6,70	-9.530218	-9.530348	0.195186
6,75	-9.536795	-9.536997	0.217773
6.8	-9.542921	-9.543206	0.237430
6.9	-9.553930	-9.554411	0.270379
7	-9.563440	-9.564143	0.297453
7.02	-9.565179	-9.565930	0.302304
7.03	-9.566029	-9.566803	0.304666
7.04	-9.566867	-9.567665	0.306988
7.05	-9.567691	-9.568513	0.309274
7.06	-9.568504	-9.569350	0.311518
7.07	-9.569304	-9.570175	0.313726
7.08	-9.570093	-9.570988	0.315899
7.09	-9.570869	-9.571789	0.318039
7.1	-9.571634	-9.572578	0.320144
7.11	-9.572387	-9.573356	0.322217
7.12	-9.573129	-9.574123	0.324255
7.13	-9.573860	-9.574878	0.326262
7.14	-9.574579	-9.575623	0.328237
7.15	-9.575288	-9.576357	0.330182
7.16	-9.575986	-9.577080	0.332098
7.17	-9.576673	-9.577792	0.333984
7.18	-9.577350	-9.578494	0.335842
7.19	-9.578016	-9.579185	0.337673
7.2	-9.578672	-9.579867	0.339475
7.21	-9.579318	-9.580538	0.341252

$r_{\max}$	HF energy per spin	HFB energy per spin	$\|A\|$
7.22	-9.579954	-9.581199	0.343001
7.23	-9.580580	-9.581850	0.344726
7.24	-9.581196	-9.582492	0.346424
7.25	-9.581803	-9.583124	0.348100
7.26	-9.582400	-9.583746	0.349750
7.27	-9.582989	-9.584360	0.351373
7.28	-9.583568	-9.584965	0.352974
7.29	-9.584137	-9.585559	0.354556
7.3	-9.584698	-9.586145	0.356113
7.31	-9.585251	-9.586722	0.357647
7.32	-9.585794	-9.587291	0.359159
7.33	-9.586329	-9.587852	0.360649
7.34	-9.586855	-9.588403	0.362121
7.35	-9.587374	-9.588946	0.363568
7.36	-9.587884	-9.589481	0.364997
7.37	-9.588385	-9.590007	0.366408
7.38	-9.588880	-9.590526	0.367795
7.39	-9.589365	-9.591036	0.369167
7.4	-9.589843	-9.591538	0.370518
7.5	-9.594221	-9.596158	0.383042
7.6	-9.597934	-9.600101	0.393956
7.75	-9.602454	-9.604939	0.407772
7.8	-9.603719	-9.606304	0.411785
7.9	-9.605942	-9.608711	0.419029
8	-9.607802	-9.610740	0.425341
8.1	-9.609353	-9.612444	0.430832
8.25	-9.611202	-9.614493	0.437740
8.4	-9.612592	-9.616050	0.443299
8.5	-9.613316	-9.616869	0.446377
8.6	-9.613908	-9.617544	0.449023
8.75	-9.614597	-9.618336	0.452298
9	-9.615347	-9.619215	0.456284
9.5	-9.615982	-9.619991	0.460618
10	-9.616088	-9.620155	0.462363
10.5	-9.616003	-9.620089	0.462991
11	-9.615847	-9.619938	0.463166
11.5	-9.615662	-9.619752	0.463168
12	-9.615462	-9.619549	0.463105
12.5	-9.615252	-9.619335	0.463016
13	-9.615032	-9.619111	0.462916
13.5	-9.614804	-9.618879	0.462810
14	-9.614567	-9.618637	0.462698

Table 4.2: Numerical values used to get the curves in Figure 4.4, with $N = 16$, $N_b = 200$ and $\ell_{\max} = 1$. The displayed energy is the energy per spin which has to be multiplied by 2 to get the total energy. Each calculation took several hours on a regular PC.

4.3.3 Properties of the HFB ground state

In Table 4.3 below we give our results for $N_b = 200$ and $N = 6, 10, 16$ and 20, for the optimal values of $r_{\max}$. With $\ell_{\max} = 1$ we have observed that the shells are filled alternatively. In HF theory, the cases $N = 10$ and $N = 16$ correspond to closed shells, whereas for $N = 6$ and $N = 20$ the last shell is only partially filled. This is a simple explanation for the fact that the pairing matrix is much bigger in these cases.

N	$r_{\max}$	HF gap	HF energy	HFB energy	$\|A\|$
6	15	0	-1.7327688	-1.9934252	1.0242134
10	11	1.023642	-6.7911634	-6.8148576	0.5871951
16	10	1.404396	-19.232177	-19.2403096	0.4623593
20	9	0	-30.010574	-30.174576	0.8235512

Table 4.3: Results for $N_b = 200$ and $\ell_{\max} = 1$.

In Table 4.4 we display the occupation numbers for the optimal HFB ground state in the closed shell case $N = 16$ and in the open shell case $N = 20$. Because of the spin, these are the eigenvalues of G_0 multiplied by 2 and that of G_1 multiplied by 6. Even in the closed shell case $N = 16$, a rather important pairing effect is observed between the last filled orbital (the second $\ell = 1$ eigenvalue) and the first unfilled one (the third $\ell = 0$ eigenvalue). This results in a decrease of the last occupation number of the HF one-particle density matrix by approximately 0.228.

$N = 16$	
$\ell = 0$	$\ell = 1$
1.9999318	5.9997504
1.9980654	5.7710970
0.2281366	0.0026448
0.0002694	0.0002720
0.0000120	0.0000084
0.0000014	0.0000012
0.0000002	3.316D-07
6.896D-08	1.005D-07
⋮	⋮

$N = 20$	
$\ell = 0$	$\ell = 1$
1.9999832	5.9999490
1.9997458	5.9983572
1.9875134	2.0084946
0.0045222	0.0012960
0.0000690	0.0000552
0.0000058	0.0000066
0.0000010	0.0000002
0.0000002	3.072D-07
⋮	⋮

Table 4.4: Occupation numbers of the HFB minimizer, for $N_b = 200$ and $\ell_{\max} = 1$.

In Figure 4.5 the two (radial) ground state densities of particles in the HF and HFB cases are shown for $N = 16$. We observe that the HFB ground state is slightly more concentrated at the origin.

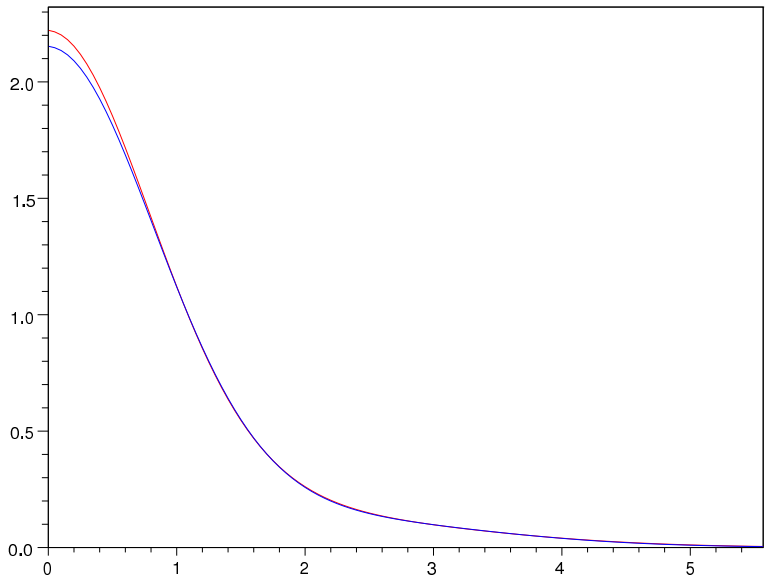

Figure 4.5: The two densities in the HF (blue) and HFB (red) cases, for $N = 16$, $N_b = 200$, $r_{\max} = 10$ and $\ell_{\max} = 1$. The HFB ground state is slightly more concentrated at the origin.

4.3.4 Quality of the approximation in terms of N_b

In Table 4.5 we display the HF and HFB ground state energies for $N = 16$, $\ell_{\max} = 1$ for the the optimal value of $r_{\max}$, in terms of the number of discretization points N_b of the regular grid. The convergence is not very fast, but we see that the difference between the HF and the HFB energy, as well as the norm of the pairing matrix are of the same order for small N_b as they are for larger N_b's. From this observation we can conclude that it is probably not necessary to take N_b very large in order to decide whether pairing occurs or not.

N_b	$r_{\max}$	HF energy	HFB energy	difference	$\|A\|$
30	9	-19.112314	-19.117948	0.005634	0.425604
50	9	-19.189066	-19.196012	0.006946	0.445728
100	9	-19.222300	-19.229872	0.007572	0.454173
150	10	-19.229494	-19.237574	0.008080	0.461725
200	10	-19.232176	-19.240308	0.008132	0.462363
250	10	-19.233420	-19.241576	0.008156	0.462659
300	10	-19.234094	-19.242264	0.008170	0.462821
400	11	-19.234826	-19.243068	0.008242	0.463905
500	11	-19.235206	-19.243456	0.008250	0.463985

Table 4.5: Value of the HF and HFB energies for $N = 16$ and $\ell_{\max} = 1$ and the (approximate) optimal $r_{\max}$.

4.3.5 Dependance on the initial state

We want to prove numerically that the convergence of the ODA algorithm does not depend on the initial state. In all the previous numerical computations, the initial generalized one-body matrix Γ was a multiple of the identity. To show that the convergence does not depend on the initial state, we chose randomly 100 initial matrices Γ and ran the algorithm for each of them. In figure 4.6, each curve represents the relative error between the energy of the state along the iterations of the ODA, and the reference energy (obtained with the identity matrix as initial state). The relative error is plotted in logarithmic scale. Both in the HF and HFB cases, we can see that the energy decreases during the algorithm in the same way whatever the initial random state. We also cheked that the final state was always the same. The relative error which we found for the final state was of the order of 10^{-7} for all the simulations, which was our chosen convergence precision.

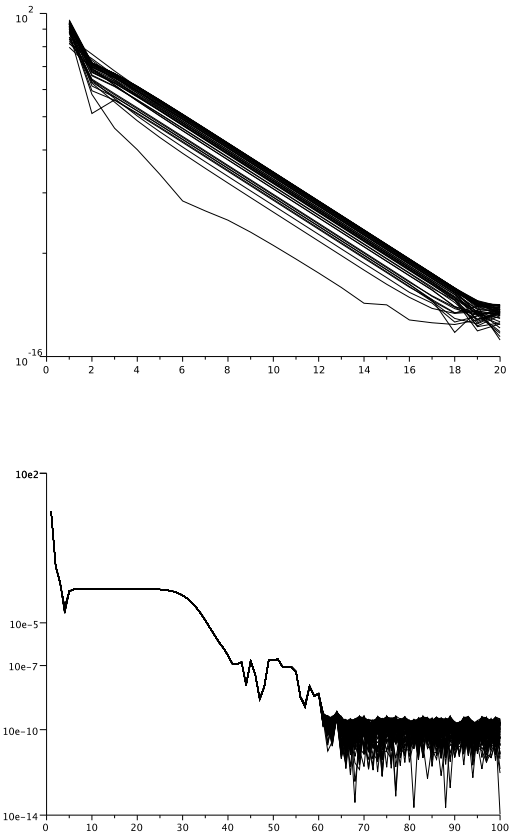

Figure 4.6: Relative error between the energy of a random initial state and the identity state, for $N = 16$, $N_b = 100$, $r_{\max} = 9$ and $\ell_{\max} = 1$ (Top: HF, bottom: HFB).

4.4 Comparison with Thomas-Fermi

In this section we compute the Thomas-Fermi energy which is an exact lower bound to our HF energy, by the Lieb-Thirring inequality [33]. We use this to verify that our computations are of the correct order.

Let us recall that the Hartree-Fock energy with attractive potential $W(x) = -1/|x|$ and with the spin removed is

$$\begin{aligned}\mathcal{E}_{HF}(\gamma) &= 2\left(\mathrm{Tr}(-\Delta\gamma) - \iint \frac{\rho_\gamma(x)\rho_\gamma(y)}{|x-y|}dxdy + \frac{1}{2}\iint \frac{|\gamma(x,y)|^2}{|x-y|}dxdy\right)\\ &\geq 2\left(\mathrm{Tr}(-\Delta\gamma) - \iint \frac{\rho_\gamma(x)\rho_\gamma(y)}{|x-y|}dxdy\right)\end{aligned}$$

The Lieb-Thirring inequality [33, 34] says that

$$\mathrm{Tr}(-\Delta\gamma) \geq c_{LT}\int \rho_\gamma^{5/3}$$

where the best constant known so far [11] is

$$c_{LT} = 0,672 \times \frac{d}{d+2}\left(\frac{d(2\pi)^d}{|S^{d-1}|}\right)^{2/d} = 0,672 \times \frac{3}{5}(6\pi^2)^{2/3}$$

with d the space dimension ($d = 3$ for us) and $|S^{d-1}|$ the surface of the corresponding unit sphere. It is conjectured that the true constant is the semi-classical one:

$$c_{sc} = \frac{3}{5}(6\pi^2)^{2/3}.$$

We get the exact lower bound in terms of the Thomas-Fermi energy

$$\begin{aligned}\mathcal{E}_{HF}(\gamma) &\geq \inf_{\substack{\gamma\\ \mathrm{Tr}(\gamma)=N/2}}\left\{2c_{LT}\int \rho_\gamma^{5/3} - 2D(\rho_\gamma,\rho_\gamma)\right\}\\ &= \inf_{\substack{\rho\geq 0\\ \int_{\mathbb{R}^3}\rho=N/2}}\left\{2c_{LT}\int \rho^{5/3} - 2D(\rho,\rho)\right\}\end{aligned}$$

where we have used the famous representability theorem [31] which says that the set of all the densities ρ_γ is simply $\{\rho \geq 0 \ : \ \int_{\mathbb{R}^3}\rho = N/2\}$.

We now use the scaling $\rho(x) = (N/2)\lambda^3\rho'(\lambda x)$ with the constraint $\int \rho'(x)dx = 1$. We have

$$\int \rho^{5/3} = (N/2)^{5/3}\lambda^5\int (\rho'(\lambda x))^{5/3}dx = (N/2)^{5/3}\lambda^2\int (\rho')^{5/3}$$

and

$$\iint \frac{\rho(x)\rho(y)}{|x-y|}dxdy = (N/2)^2\lambda^6\iint \frac{\rho'(\lambda x)\rho(\lambda y)}{|x-y|}dxdy = (N/2)^2\lambda D(\rho',\rho').$$

The corresponding Thomas-Fermi energy is

$$\inf_{\substack{\rho'\geq 0\\ \int\rho'=1}}\left\{2c_{LT}(N/2)^{5/3}\lambda^2\int(\rho')^{5/3}-2(N/2)^2\lambda D(\rho',\rho')\right\}$$

and we choose λ such that $c_{LT}(N/2)^{5/3}\lambda^2=(N/2)^2\lambda$, that is, $\lambda=\dfrac{(N/2)^{1/3}}{c_{LT}}$. We obtain

$$\mathcal{E}_{HF}(\gamma)\geq\frac{2(N/2)^{7/3}}{c_{LT}}I_{\mathrm{TF}},\tag{4.4.11}$$

where

$$I_{\mathrm{TF}}:=\inf_{\substack{\rho\geq 0\\ \int_{\mathbb{R}^3}\rho=1}}\left\{\int\rho^{5/3}-D(\rho,\rho)\right\}.\tag{4.4.12}$$

4.4.1 Estimate of I_{TF} using the Hardy-Littlewood-Sobolev inequality

The Hardy-Littlewood inequality tells us that

$$D(\rho,\rho)\leq c_{HLS}\|\rho\|^2_{L^{6/5}}$$

where

$$c_{HLS}=\pi^{1/2}\frac{\Gamma(1)}{\Gamma(5/2)}\left(\frac{\Gamma(3/2)}{\Gamma(3)}\right)^{-2/3}=\frac{4^{5/3}\pi^{-1/3}}{3},$$

see [30].

We interpolate $L^{6/5}$ between L^1 and $L^{5/3}$, that is, $5/6=\theta/1+(1-\theta)3/5$ with $\theta=7/12$, and we get

$$\|\rho\|_{L^{6/5}}\leq\|\rho\|_{L^1}^{7/12}\|\rho\|_{L^{5/3}}^{5/12}.$$

We deduce that

$$\begin{aligned}D(\rho,\rho)\leq c_{HLS}\|\rho\|^2_{L^{6/5}}&\leq c_{HLS}\left(\int\rho\right)^{7/6}\|\rho\|^{5/6}_{L^{5/3}}\\&\leq c_{HLS}\left(\int\rho\right)^{7/6}\left(\int\rho^{5/3}\right)^{1/2},\end{aligned}$$

and therefore

$$\begin{aligned}\int\rho^{5/3}-D(\rho,\rho)&\geq\int\rho^{5/3}-c_{HLS}\left(\int\rho^{5/3}\right)^{1/2}\\&\geq X^2-c_{HLS}X\end{aligned}$$

with $X=\left(\int\rho^{5/3}\right)^{1/2}$. The minimum is reached at $X=\dfrac{c_{HLS}}{2}$, so

$$\int\rho^{5/3}-D(\rho,\rho)\geq-\frac{(c_{HLS})^2}{4}.$$

Finally,

$$I_{\mathrm{TF}} \geq -\frac{2(N/2)^{7/3}}{c_{LT}} \frac{(c_{HLS})^2}{4}.$$

Then

$$\mathcal{E}_{HF}(\gamma) \geq -0.085232\, N^{7/3}.$$

Remark 4.4.1. *With the Lieb-Thirring conjecture we find*

$$\mathcal{E}_{HF}(\gamma) \geq -0.057277\, N^{7/3}.$$

4.4.2 Numerical calculation of I_{TF}

We want to calculate

$$I_{\mathrm{TF}} = \inf_{\substack{\rho \geq 0 \\ \int \rho = 1}} \left\{ \int \rho^{5/3} - D(\rho, \rho) \right\}$$

We know that there is a unique minimizer (up to translations), which is radial and decreasing [30]. The Euler-Lagrange equation is

$$\frac{5}{3}\rho^{2/3} = 2\left(\rho \star \frac{1}{|x|} - \mu\right)_+$$

where μ is the Lagrange multiplier. Denoting

$$\theta = \rho \star \frac{1}{|x|} - \mu, \qquad V_\rho = \rho \star \frac{1}{|x|},$$

we find

$$\rho^{2/3} = \frac{6}{5}\left(V_\rho - \mu\right)_+$$

or equivalently

$$\rho = \left(\frac{6}{5}\right)^{3/2} \left(V_\rho - \mu\right)_+^{3/2}$$

We arrive at the Lane-Emden equation, which describes the density of a fluid subject to its own gravitational potential:

$$-\Delta\theta = -\Delta V_\rho = 4\pi\rho = 4\pi \left(\frac{6}{5}\right)^{3/2} \theta_+^{3/2}$$

Since θ is radial, the equation can be written as

$$-\theta'' - \frac{2}{r}\theta' = 4\pi \left(\frac{6}{5}\right)^{3/2} \theta^{3/2}$$

or, equivalently,

$$r\theta'' + 2\theta' + rK\theta^{3/2} = 0,$$

where

$$K = 4\pi \left(\frac{6}{5}\right)^{3/2}.$$

Because θ is radial, the initial conditions are $\theta(0) = a$, $\theta'(0) = 0$ and we have to find a such that

$$\int_{\mathbb{R}^3} \rho = \left(\frac{6}{5}\right)^{3/2} \int_{\mathbb{R}^3} \theta_+^{3/2} = 4\pi \left(\frac{6}{5}\right)^{3/2} \int_0^\infty \theta(r)_+^{3/2} r^2 \, dr = 1.$$

Using a simple bisection method on Mathematica, we find

$$a \simeq 1.713285.$$

In order to find the corresponding Lagrange multiplier μ, we use that

$$\theta(0) = V(0) - \mu = 4\pi \int_0^\infty \rho(r)\, r \, dr - \mu,$$

since, by Newton's theorem,

$$V(0) = \int_{\mathbb{R}^3} \frac{\rho(|y|)}{\max(0, |y|)} \, dy = \int_{\mathbb{R}^3} \frac{\rho(|y|)}{|y|} \, dy.$$

So

$$\mu = 4\pi \int_0^\infty \rho(r)\, r \, dr - \theta(0) = 4\pi \left(\frac{6}{5}\right)^{3/2} \int_0^\infty \theta(r)_+^{3/2} r \, dr - \theta(0).$$

Numerically we find

$$\mu \simeq 1.27265.$$

Then we can compute

$$I_{\mathrm{TF}} = 4\pi \int_0^\infty \big(\rho(r)^{5/3} - \theta(r)\rho(r)\big) r^2 \, dr - \mu \simeq -1.09084$$

and we deduce that

$$\boxed{\mathcal{E}_{\mathrm{HF}}(\gamma) \geq -0.0706699\, N^{7/3}.}$$

This is not much better than the estimate given by the Hardy-Littlewood-Sobolev inequality.

Remark 4.4.2. *With the Lieb-Thirring conjecture we find*

$$\boxed{\mathcal{E}_{\mathrm{HF}}(\gamma) \geq -0.0474902\, N^{7/3}.}$$

For comparison we provide the Thomas-Fermi energy for different values of N in Table 4.6 and reproduce the values of Table 4.3. The energy is always smaller than our computed HF and HFB discretized ground state energy, as expected.

N	TF with c_{LT}	TF with c_{sc}	HF energy	HFB energy
6	-4.62296	-3.10663	-1.7327688	-1.9934252
10	-15.2254	-10.2314	-6.7911634	-6.8148576
16	-45.5877	-30.6349	-19.232177	-19.2403096
20	-76.7310	-51.5632	-30.010574	-30.174576

Table 4.6: The Thomas-Fermi lower bound for different values of N.

Chapter 5

A simplified model for protons and neutrons

In this section we report on our numerical results concerning a simple model inspired of nuclear physics. The interaction between protons and neutrons is not a fundamental law of nature because these are composite particles made of quarks, which interact through strong and electrostatic forces. A common procedure used in nuclear physics is to use *empirical forces* [39] which involve a small number of parameters which are fitted to experiment or to the known behavior of the model in some limits. The most common forces used in practice are the so-called Skyrme [41] and Gogny [10,16,17] forces and they depend nonlinearly on the state itself. Here we consider an effective force which is fixed and does not depend on the quantum state. We also take it spin-independent and isospin-independent. Our goal is to test some simple ideas and not to do a real nuclear physics calculation.

5.1 Model

The nucleon-nucleon potential has been observed to be repulsive at short distances and attractive at medium distances due to the strong forces. More precisely, the potential starts to be attractive at about 0.7 fm and it becomes minimal at about 0.9 fm. It then decays exponentially and becomes minimal at about 2 fm. At distances of about 1.7 fm, it is stronger than the Coulomb interaction but it becomes weaker at about 2 fm. See Figure 5.1 for the typical shape of the nuclear interaction.

A simple choice to describe this potential is to take

$$W(x) = \frac{\kappa}{|x|} - a_1\, e^{-b_1|x|^2} + a_2\, e^{-b_2|x|^2}, \tag{5.1.1}$$

with $a_2, a_1 > 0$, $b_1 < b_2$. The constant κ is 1 for the proton-proton interaction and 0 for the proton-neutron and the neutron-neutron interaction. The other constants usually also depend on the isospin (the quantum variable which determines whether a nucleon is a neutron or a proton), but not too much. The nuclear force is nearly

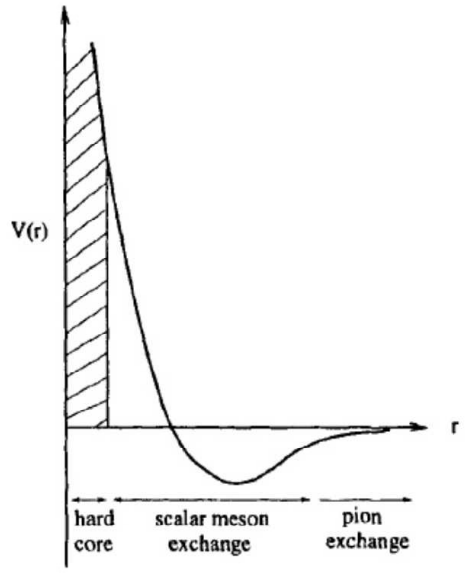

Figure 3-7: Schematic diagram showing the different parts of a nucleon-nucleon potential as a function of distance r between two nucleons. The hard core radius is around 0.4 fm and it takes energy >1 GeV to bring two nucleons closer than (twice) this distance. The main part of the attraction lies at intermediate ranges, at radius ~1 fm, and is believed to be dominated by the exchange of scalar mesons. The long-range part, starting at around 2 fm, is due to the single-pion exchange.

Figure 5.1: Extract from [43] showing the typical nuclear force.

independent of whether the nucleons are neutrons or protons. For simplicity we work here with particles having a definite isospin. This means that we assume to have either only protons or only neutrons. In particular we want to ask for which intensity of the effective force it becomes possible for the protons to overcome their Coulomb repulsion and form a bound state. In reality a nucleus is made of a certain number of protons and neutrons and one has to use a different HFB state for each species.

In our applications we have chosen for simplicity $b_1 = 1$, $b_2 = 4$, $a_1 = a = 2\,a_2/3$. This means that the effective force takes the form

$$W(x) = \frac{\kappa}{|x|} + a\left(\frac{3}{2}\,e^{-4|x|^2} - e^{-|x|^2}\right). \tag{5.1.2}$$

When $\kappa = 1$, this force is purely repulsive for $a \leq 2.87$ and it becomes attractive at intermediate distances for larger a's. The corresponding force is displayed in Figure 5.2 for $a = 1$ and $\kappa = 0$.

One can ask several questions concerning the model considered in this chapter:

1. For which value of a does a system of N identical nucleons bind in Hartree-Fock theory?

2. Is there always pairing when there is binding?

3. Can pairing effects allow for binding with a smaller a than in HF theory?

These questions are mostly of academic nature for the very simplified model considered in this section. But investigating the same problems with more realistic

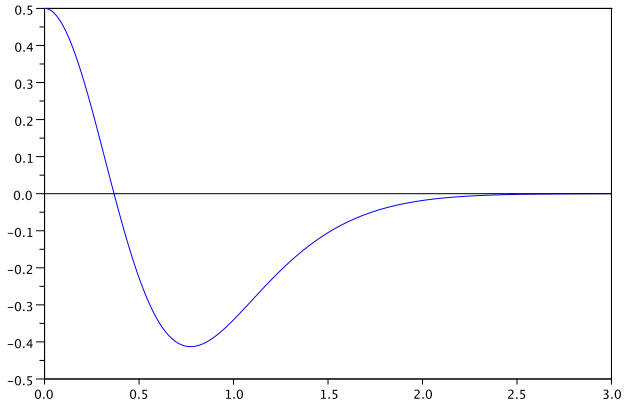

Figure 5.2: The effective force $-e^{-|x|^2} + 3\,e^{-4|x|^2}/2$ used in our calculation of Section 5. The (repulsive) Coulomb potential must be added for protons.

forces is very important from a physical point of view. From a mathematical point of view, nothing seems to be known for simple models of the same form as in this section. It is not even known that binding always occurs for a large enough with the previous interaction. We hope that our calculations will stimulate some further mathematical studies.

5.2 Some computational details

We always minimize over states having the spin, time-reversal and rotation symmetries. The Bach-Fröhlich-Jonsson Theorem 1.4.2 does not apply to the model of this section, hence we are making a further approximation here.

For such symmetric states we have shown in Section 2.3.2 that the energy can be expressed in terms of

$$(ij|kl)_{\ell,\ell'} = \int_0^\infty r^2\,dr \int_0^\infty s^2\,ds\;\chi_i(r)\,\chi_j(r)\,\chi_k(s)\,\chi_l(s)\,w_{\ell,\ell'}(r,s)$$

where, for the model considered in this section,

$$\begin{aligned} w_{\ell,\ell'}(r,s) &= \frac{1}{2}\int_{-1}^{1} W\left(\sqrt{r^2+s^2-2rst}\right)\, P_\ell(t)\, P_{\ell'}(t)\, dt \\ &= \frac{1}{2}\int_{-1}^{1} P_\ell(t)\, P_{\ell'}(t)\Bigg(\frac{\kappa}{\sqrt{r^2+s^2-2rst}} \\ &\qquad - a_1\, e^{-b_1(r^2+s^2-2rst)} + a_2\, e^{-b_2(r^2+s^2-2rst)}\Bigg)\, dt \\ &= \int_{-1}^{1} P_\ell(t)\, P_{\ell'}(t)\Bigg(\kappa \sum_{m=0}^{\infty} \begin{pmatrix} \ell & \ell' & m \\ 0 & 0 & 0 \end{pmatrix}^2 \frac{\min(r,s)^m}{\max(r,s)^{m+1}} \\ &\qquad - \frac{a_1}{2}\, e^{-b_1(r^2+s^2-2rst)} + \frac{a_2}{2}\, e^{-b_2(r^2+s^2-2rst)}\Bigg)\, dt \end{aligned}$$

For $0 \le \ell, \ell' \le \ell_{\max}$ with $\ell_{\max}$ not too large, the Gaussian integrals can be computed exactly and it is possible to find the exact expression of $w_{\ell,\ell'}(r,s)$. A simple but tedious calculation shows that

$$\begin{aligned} w_{0,0}(r,s) = \frac{\kappa}{\max(r,s)} &- \frac{a_1}{4b_1 rs}\, 2\sinh(2b_1 rs)\, e^{-b_1(r^2+s^2)} \\ &+ \frac{a_2}{4b_2 rs}\, 2\sinh(2b_2 rs)\, e^{-b_2(r^2+s^2)}, \end{aligned}$$

$$\begin{aligned} w_{0,1}(r,s) = +\frac{\kappa \min(r,s)}{3\max(r,s)^2} &- \frac{a_1}{4b_1 rs}\left(2\cosh(2b_1 rs) - \frac{2\sinh(2b_1 rs)}{2b_1 rs}\right) e^{-b_1(r^2+s^2)} \\ &+ \frac{a_2}{4b_2 rs}\left(2\cosh(2b_2 rs) - \frac{2\sinh(2b_2 rs)}{2b_2 rs}\right) e^{-b_2(r^2+s^2)}, \end{aligned}$$

and

$$\begin{aligned} w_{1,1}(r,s) = &\frac{2\kappa \min(r,s)^2}{15\max(r,s)^3} \\ &- \frac{a_1}{4b_1 rs}\left(2\sinh(2b_1 rs) - \frac{2\cosh(2b_1 rs)}{b_1 rs} + \frac{2\sinh(2b_1 rs)}{2(b_1 rs)^2}\right) e^{-b_1(r^2+s^2)} \\ &+ \frac{a_2}{4b_2 rs}\left(2\sinh(2b_2 rs) - \frac{2\cosh(2b_2 rs)}{b_2 rs} + \frac{2\sinh(2b_2 rs)}{2(b_2 rs)^2}\right) e^{-b_2(r^2+s^2)}. \end{aligned}$$

The computation of the integral $(ij|mn)_{\ell,\ell'}$ against hat functions is much more tedious, however. It is easy to find an exact expression for the Coulomb part, but not so simple for the Gaussian part. So we have performed a numerical calculation of these integrals. Since we have of the order of $(N_b)^4$ integrals, we could not take N_b too large. The results of the previous chapter indicated that the existence of

pairing effects does not depend very much on the size of the basis, and it seems reasonable to expect the same for the model of this chapter.

Using the calculations of Chapter 2, we deduce that the nuclear HFB energy associated with the density and pairing matrices $(G^\ell, A^\ell)_{0\le\ell\le\ell_{\max}}$ is

$$\begin{aligned}\mathcal{E}(G^0,...,G^\ell,A^0,...,A^\ell)&\\ = 2\sum_{\ell=0}^{\ell_{\max}}(2\ell+1)\mathrm{Tr}(h^\ell G^\ell) + \sum_{\ell,\ell'=0}^{\ell_{\max}}(2\ell+1)(2\ell'+1)\Big(&2\mathrm{Tr}(G^\ell\mathcal{J}(G^{\ell'}))\\ -\mathrm{Tr}(G^\ell\mathcal{K}_{\ell\ell'}(G^{\ell'})) + \mathrm{Tr}(A^\ell\mathcal{K}_{\ell\ell'}(A^{\ell'}))\Big),&\end{aligned}$$

where

$$\big(\mathcal{J}(G^{\ell'})\big)_{ij} := \sum_{k,l=1}^{N_b} G^{\ell'}_{kl}\big(ij,kl\big)_{0,0}, \tag{5.2.3}$$

$$\Big(\mathcal{K}_{\ell\ell'}(G^{\ell'})\Big)_{ij} := \sum_{k,l=1}^{N_b} G^{\ell'}_{kl}\big(ik,jl\big)_{\ell,\ell'}, \tag{5.2.4}$$

and

$$(ij|kl)_{\ell,\ell'} = \int_0^\infty r^2\,dr\int_0^\infty s^2\,ds\;\chi_i(r)\,\chi_j(r)\,\chi_k(s)\,\chi_l(s)\,w_{\ell,\ell'}(r,s), \tag{5.2.5}$$

$$\begin{aligned}w_{\ell,\ell'}(r,s) &= \frac{1}{2}\int_{-1}^1 W\left(\sqrt{r^2+s^2-2rst}\right)\,P_\ell(t)\,P_{\ell'}(t)\,dt\\ &= \frac{1}{2}\int_{-1}^1 P_\ell(t)\,P_{\ell'}(t)\Big(\frac{\kappa}{\sqrt{r^2+s^2-2rst}}\\ &\qquad - a_1\,e^{-b_1(r^2+s^2-2rst)} + a_2\,e^{-b_2(r^2+s^2-2rst)}\Big)\,dt.\end{aligned}$$

In our computations we have taken $\ell_{\max}=1$ and the energy is, in this special case,

$$\begin{aligned}\mathcal{E}(G^0,G^1,A^0,A^1) &= 2\mathrm{Tr}(h^0G^0) + 6\mathrm{Tr}(h^1G^1)\\ &+ 2\Big(\mathrm{Tr}(G^0\mathcal{J}(G^0)) + 6\mathrm{Tr}(G^0\mathcal{J}(G^1)) + 9\mathrm{Tr}(G^1\mathcal{J}(G^1))\Big)\\ &- \big(\mathrm{Tr}(G^0\mathcal{K}_{00}(G^0)) + 6\mathrm{Tr}(G^0\mathcal{K}_{01}(G^1)) + 9\mathrm{Tr}(G^1\mathcal{K}_{11}(G^1))\big)\\ &+ \big(\mathrm{Tr}(A^0\mathcal{K}_{00}(A^0)) + 6\mathrm{Tr}(A^0\mathcal{K}_{01}(A^1)) + 9\mathrm{Tr}(A^1\mathcal{K}_{11}(A^1))\big).\end{aligned}$$

Any minimizer (G_0,G_1,A_0,A_1) of $\mathcal{E}$ under the constraints

$$0\le\Upsilon^\ell\mathbf{S}\Upsilon^\ell\le\Upsilon^\ell,\ \text{with}\ \Upsilon^\ell := \begin{pmatrix}G^\ell & A^\ell\\ A^\ell & S^{-1}-G^\ell\end{pmatrix}\ \text{and}\ \mathbf{S}=\begin{pmatrix}S&0\\0&S\end{pmatrix}\quad\forall\ell=0,1 \tag{5.2.6}$$

and

$$\boxed{2\,\mathrm{Tr}(SG^0) + 6\,\mathrm{Tr}(SG^1) = N} \tag{5.2.7}$$

can be written as

$$\begin{cases} \Upsilon^0 = \sum_{\epsilon_i^0<0} f_i^0 \big(f_i^0\big)^T \\ \Upsilon^1 = \sum_{\epsilon_i^1<0} f_i^1 \big(f_i^1\big)^T \end{cases}$$

where the f_i^0 and f_i^1 solve the generalized eigenvalue problem

$$\begin{cases} \big(\mathbf{F}^0 - \mu\mathbf{N}\big) f_i^0 = \epsilon_i^0\, \mathbf{S}\, f_i^0, & \big\langle f_i^0, \mathbf{S} f_j^0 \big\rangle = \delta_{ij}, \\ \big(\mathbf{F}^1 - 3\mu\mathbf{N}\big) f_i^1 = \epsilon_i^1\, \mathbf{S}\, f_i^1, & \big\langle f_i^1, \mathbf{S} f_j^1 \big\rangle = \delta_{ij}. \end{cases} \tag{5.2.8}$$

Here $\mathbf{F}^0$ and $\mathbf{F}^1$ are the Fock matrices associated with G^0 and G^1:

$$\mathbf{F}^0 = \begin{pmatrix} H_G^0 & \Pi_A^0 \\ \Pi_A^0 & -H_G^0 \end{pmatrix} \quad \text{and} \quad \mathbf{F}^1 = \begin{pmatrix} H_G^1 & \Pi_A^1 \\ \Pi_A^1 & -H_G^1 \end{pmatrix}$$

where

$$H_G^0 = 2h^0 - 4\mathcal{J}(G_0) + 12\mathcal{J}(G_1) + 2\mathcal{K}_{00}(G_0) + 6\mathcal{K}_{10}(G_1),$$
$$H_G^1 = 6h^1 - 12\mathcal{J}(G_0) + 36\mathcal{J}(G_1) + 6\mathcal{K}_{01}(G_0) + 18\mathcal{K}_{11}(G_1),$$
$$\Pi_A^0 = -2\mathcal{K}_{00}(A_0) - 6\mathcal{K}_{01}(G_1),$$

and

$$\Pi_A^1 = -6\mathcal{K}_{01}(A_0) - 18\mathcal{K}_{11}(A_1).$$

5.3 Slow convergence and oscillations of Roothaan

We have observed that the Roothaan algorithm *almost always oscillates*, even in the HFB case (see some examples in Figures 5.3, 5.4 and 5.5). This is in stark contrast with the results of the previous section where the Roothaan algorithm was almost always converging. Sometimes it very slowly converges in the HF case (see, e.g., Figure 5.4). However we have always obtained convergence for the HF Roothaan algorithm when a is small enough, that is, when it is expected that there is actually no binding. For the case displayed in Figure 5.4 we have $a = 20$ but the critical a is about $\simeq 25$.

We conclude that using the ODA is very important for such attractive potentials. The same might be true with the more involved forces used in nuclear physics.

The number of iterations which are necessary for the algorithm to converge is sometimes higher for the model of this section than it was for the purely Newtonian attraction (but not always, this seems to be worse for small N_b's). This is essentially a problem of choosing a good initial trial state. In order to be able to bind, the particles have to be arranged in a certain fashion and getting close to this state is not so easy. In Figure 5.6 we display the energy along the iterations

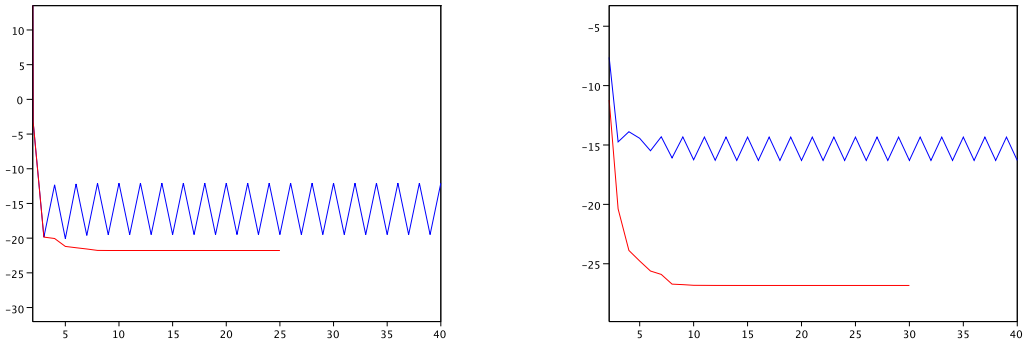

Figure 5.3: Energy along the iterations in the HF (left) and HFB (right) cases, for the Roothaan (blue) and the ODA (red), with $N = 4$, $N_b = 20$, $\ell_{\max} = 1$, $r_{\max} = 3$, $a = 35$ and $\kappa = 1$ (proton-proton case).

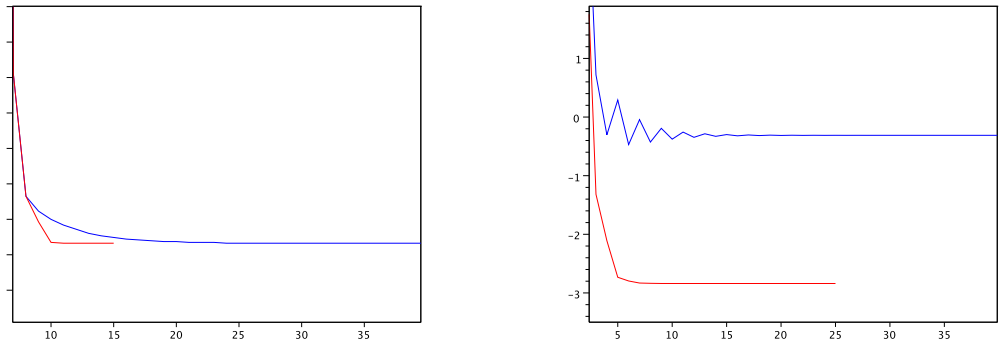

Figure 5.4: Same calculation with $a = 20$ and $\kappa = 1$ (proton-proton case). The Roothaan algorithm slowly converges in the HF case and it oscillates in the HFB case but the two values are very close.

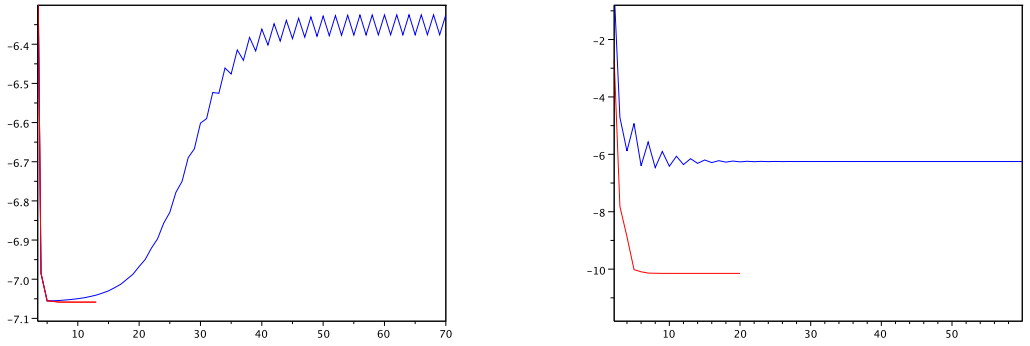

Figure 5.5: Same calculation with $a = 20$ and $\kappa = 0$ (neutron-neutron case). The Roothaan algorithm oscillates in the HFB case, but the two values are very close.

for $N_b = 10$, in both the HF and HFB cases. When we start with the simple initial state given by

$$G_{\text{init}} = \frac{N}{2\,\text{Tr}(S)}\text{Id}_{N_b}$$

(as described in the previous chapter), it takes almost 30 iterations for the HF algorithm to getting close to the ground state. The algorithm seems to be trapped for a while in the neighborhood of an other almost solution, which might be related to a bound state with less particles. The situation is better in the HFB case since we use the HF converged state as initial datum. It is expected that a better choice of an initial state will highly improve the convergence in the HF case. In practice, we always vary the parameters in the model by small amounts and we use the converged ground state of the previous calculation. So only the first calculation might be an issue. Furthermore, we have not encountered this behavior when $N_b = 50$ and so we expect it to be specific to smaller N_b's.

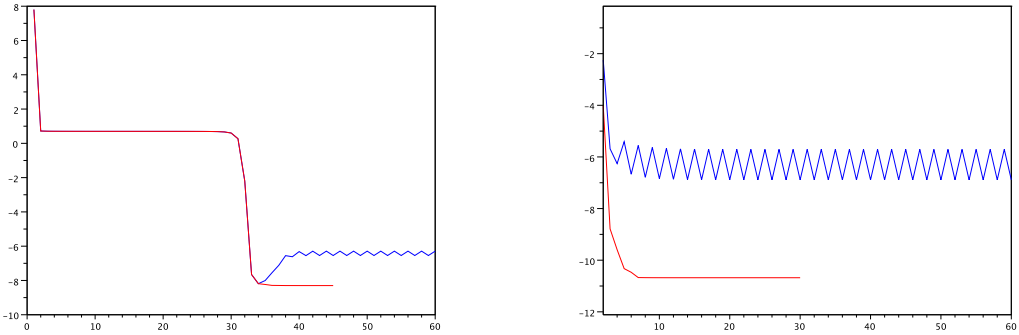

Figure 5.6: Energy along the iterations in the HF (left) and HFB (right) cases, for the Roothaan (blue) and the ODA (red), with $N = 4$, $N_b = 10$, $\ell_{\max} = 1$, $r_{\max} = 9$, $a = 35$ and $\kappa = 1$ (proton-proton case).

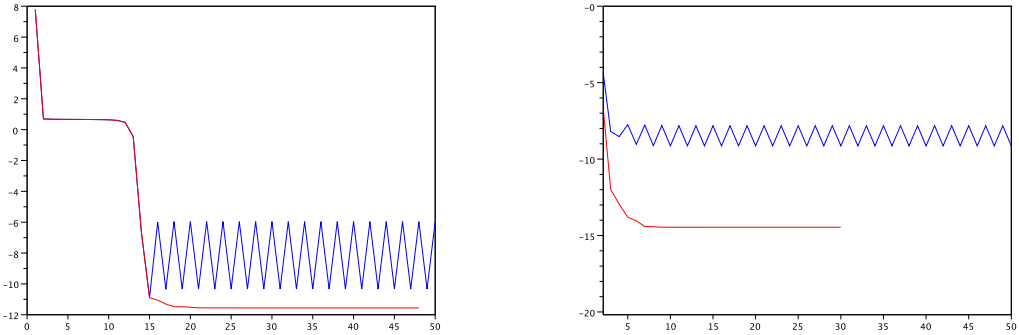

Figure 5.7: Same calculation with $a = 40$.

5.4 The critical intensity

In finite dimension there is always a minimizer. Saying differently, since the particles are trapped in a ball, they always bind. Furthermore, here we work with rotation-invariant states. So, for the true model in infinite dimension, the particles escaping to infinity cannot form a bound state of the same kind. In this special case they will spread out and vanish. We conclude that we can detect the loss of binding by looking at the last filled HF eigenvalue. When it crosses 0, this corresponds to the last particle becoming a scattering state. So we can choose as definition for the critical intensity a, the value at which this eigenvalue is 0. In HFB the definition of the critical a is less clear in finite dimension.

We have made some calculations for $N_b = 20$ and $N = 4$ and the critical intensity is about $a_c \simeq 25$ in the proton-proton case and $a_c \simeq 17.5$ in the neutron-neutron case. In Figure 5.8 we display the HF and HFB energies as functions of the parameter a, for $N = 4$ and $\kappa = 1$ (proton-proton case). Figure 5.9 is the equivalent result for $\kappa = 0$ (neutron-neutron case). The corresponding numbers are given in Tables 5.1 and 5.2.

For these calculations we have chosen $r_{\max} = 3$ which is the optimal choice for a in a neighborhood of the critical value. Like in the previous section the results depend on the radius of the ball in which the system is confined. We see that there is always pairing, the HFB curve is way below the HF curve, especially in the neutron-neutron case for which the potential is much more attractive than for protons, which repel with the Coulomb potential.

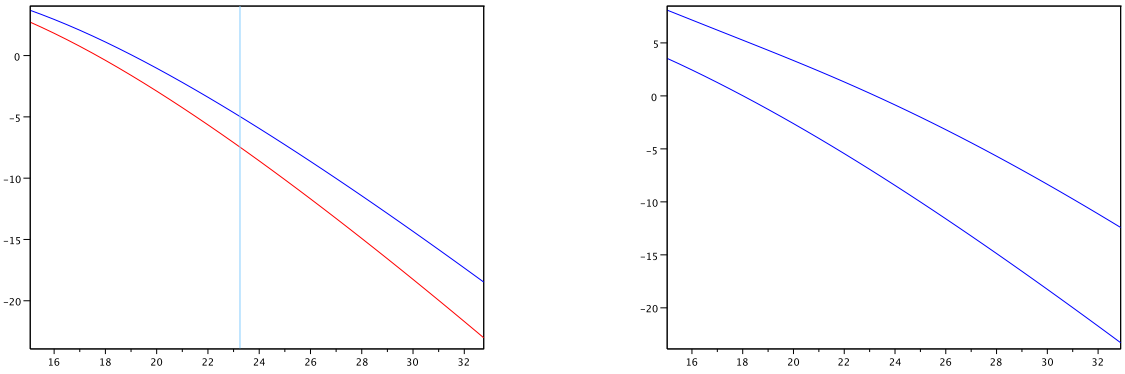

Figure 5.8: Left: Values of the HF (blue) and HFB (red) ground state energies as functions of a, with $N = 4$, $N_b = 20$, $\ell_{\max} = 1$, $r_{\max} = 3$ and $\kappa = 1$ (proton-proton case). The vertical line is the value of a for which the last filled eigenvalue vanishes. Right: Values of the two filled HF eigenvalues for the same a.

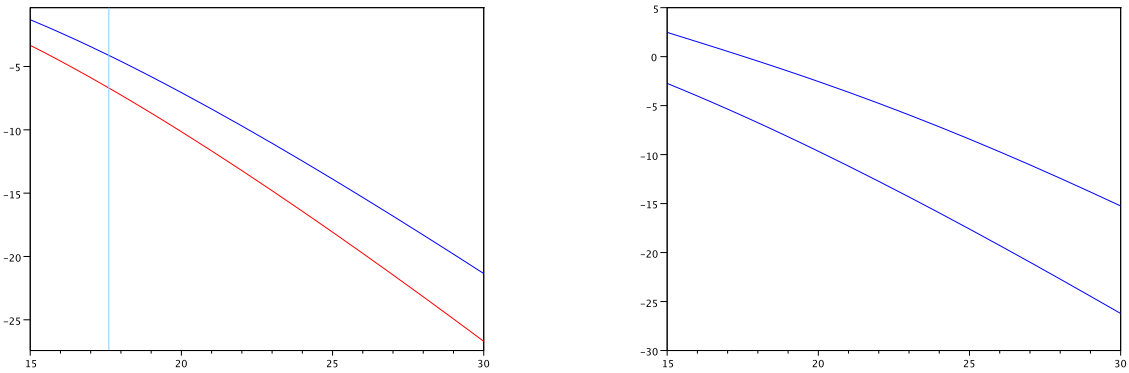

Figure 5.9: Same calculations for $\kappa = 0$ (neutron-neutron case).

a	E_{HF}	$\lambda_{\ell=0}$	$\lambda_{\ell=1}$	gap $\ell=0$	gap $\ell=1$	E_{HFB}	$\|A\|$
15	1.871103	1.782600	4.057559	4.919820	6.015268	1.393393	0.888044
15.5	1.678453	1.507395	3.818953	4.927410	5.981189	1.157264	0.892955
16	1.474612	1.223210	3.580138	4.947227	5.948734	0.909128	0.897492
16.5	1.259743	0.930901	3.341832	4.980112	5.918921	0.649371	0.901695
17	1.034184	0.631111	3.104219	5.026452	5.892746	0.378487	0.905607
17.5	0.798385	0.324276	2.867074	5.086259	5.871136	0.097020	0.909269
18	0.552865	0.010678	2.629913	5.159269	5.854923	-0.194463	0.912716
18.5	0.298169	-0.309495	2.392112	5.245018	5.844840	-0.495396	0.915975
19	0.034841	-0.636106	2.152998	5.342909	5.841517	-0.805236	0.919070
19.5	-0.236589	-0.969031	1.911912	5.452255	5.845490	-1.123463	0.922018
20	-0.515621	-1.308148	1.668236	5.572315	5.857201	-1.449593	0.924835
20.5	-0.801784	-1.653317	1.421413	5.702318	5.877009	-1.783171	0.927534
21	-1.094643	-2.004388	1.170958	5.841479	5.905186	-2.123779	0.930126
21.5	-1.393791	-2.361193	0.916454	5.989017	5.941928	-2.471027	0.932621
22	-1.698858	-2.723553	0.657555	6.144162	5.987354	-2.824557	0.935028
22.5	-2.009500	-3.091281	0.393979	6.306166	6.041514	-3.184038	0.937355
23	-2.325401	-3.464185	0.125505	6.474305	6.104388	-3.549166	0.939611
23.5	-2.646270	-3.842072	-0.148033	6.647886	6.175895	-3.919658	0.941802
24	-2.971839	-4.224749	-0.426750	6.826250	6.255898	-4.295254	0.943938
24.5	-3.301859	-4.612024	-0.710717	7.008772	6.344210	-4.675713	0.946024
25	-3.636102	-5.003712	-0.999967	7.194860	6.440595	-5.060812	0.948068
25.5	-3.974356	-5.399633	-1.294499	7.383964	6.544785	-5.450345	0.950078
26	-4.316423	-5.799613	-1.594282	7.575563	6.656473	-5.844121	0.952059
26.5	-4.662122	-6.203486	-1.899264	7.769178	6.775331	-6.241962	0.954019
27	-5.011283	-6.611092	-2.209371	7.964360	6.901008	-6.643702	0.955966
27.5	-5.363746	-7.022278	-2.524512	8.160695	7.033138	-7.049192	0.957903
28	-5.719363	-7.436900	-2.844585	8.357805	7.171345	-7.458288	0.959841
28.5	-6.077998	-7.854820	-3.169479	8.555341	7.315245	-7.870850	0.961785
29	-6.439519	-8.275908	-3.499074	8.752985	7.464455	-8.286763	0.963734
29.5	-6.803806	-8.700039	-3.833248	8.950449	7.618589	-8.705909	0.965698
30	-7.170745	-9.127095	-4.171872	9.147476	7.777268	-9.128184	0.967683
30.5	-7.540228	-9.556964	-4.514821	9.343831	7.940117	-9.553484	0.969691
31	-7.912155	-9.989541	-4.861964	9.539310	8.106767	-9.981715	0.971718
31.5	-8.286431	-10.424725	-5.213177	9.733729	8.276864	-10.412788	0.973779
32	-8.662965	-10.862420	-5.568333	9.926929	8.450058	-10.846616	0.975862
32.5	-9.041673	-11.302536	-5.927310	10.118771	8.626016	-11.283135	0.977958
33	-9.422475	-11.744987	-6.289988	10.309139	8.804414	-11.722244	0.980076

Table 5.1: Numerical results corresponding to Figure 5.8 ($N = 4$, $N_b = 50$, $r_{\max} = 3$, $\ell_{\max} = 1$, $\kappa = 1$). The displayed energy and eigenvalues are *per spin* (they have to be multiplied by 2).

a	E_{HF}	$\lambda_{\ell=0}$	$\lambda_{\ell=1}$	gap $\ell=0$	gap $\ell=1$	E_{HFB}	$\|A\|$
15	-0.673100	-1.426011	1.183665	5.154380	5.871451	-1.686141	0.932000
15.5	-0.928048	-1.748976	0.940899	5.234162	5.858572	-1.994296	0.933751
16	-1.192004	-2.078660	0.697103	5.326575	5.852222	-2.311473	0.935437
16.5	-1.464437	-2.414895	0.451651	5.430999	5.852996	-2.637147	0.937072
17	-1.744833	-2.757524	0.203929	5.546727	5.861398	-2.970818	0.938662
17.5	-2.032704	-3.106389	-0.046637	5.672999	5.877836	-3.312016	0.940215
18	-2.327587	-3.461323	-0.300566	5.809022	5.902634	-3.660305	0.941733
18.5	-2.629056	-3.822147	-0.558309	5.953996	5.936028	-4.015277	0.943221
19	-2.936711	-4.188674	-0.820251	6.107122	5.978177	-4.376559	0.944679
19.5	-3.250186	-4.560705	-1.086711	6.267618	6.029162	-4.743802	0.946112
20	-3.569139	-4.938040	-1.357944	6.434724	6.088992	-5.116688	0.947522
20.5	-3.893258	-5.320475	-1.634148	6.607708	6.157611	-5.494920	0.948911
21	-4.222254	-5.707807	-1.915468	6.785874	6.234901	-5.878226	0.950284
21.5	-4.555859	-6.099835	-2.202000	6.968556	6.320688	-6.266354	0.951642
22	-4.893826	-6.496364	-2.493797	7.155130	6.414750	-6.659071	0.952992
22.5	-5.235928	-6.897205	-2.790876	7.345008	6.516825	-7.056163	0.954337
23	-5.581952	-7.302173	-3.093222	7.537638	6.626611	-7.457430	0.955683
23.5	-5.931703	-7.711093	-3.400793	7.732509	6.743780	-7.862690	0.957035
24	-6.284997	-8.123798	-3.713525	7.929145	6.867979	-8.271771	0.958400
24.5	-6.641665	-8.540127	-4.031335	8.127107	6.998840	-8.684518	0.959785
25	-7.001548	-8.959928	-4.354125	8.325991	7.135979	-9.100785	0.961200
25.5	-7.364500	-9.383055	-4.681787	8.525426	7.279007	-9.520438	0.962651
26	-7.730382	-9.809372	-5.014204	8.725076	7.427530	-9.943354	0.964148
26.5	-8.099065	-10.238748	-5.351253	8.924634	7.581156	-10.369418	0.965699
27	-8.470427	-10.671060	-5.692808	9.123827	7.739493	-10.798527	0.967314

Table 5.2: Numerical results corresponding to Figure 5.9 ($N = 4$, $N_b = 50$, $r_{\max} = 3$, $\ell_{\max} = 1$, $\kappa = 0$). The displayed energy and eigenvalues are *per spin* (they have to be multiplied by 2).

Bibliography

[1] G. Allaire, *Analyse numérique et optimisation*, École Polytechnique, 2005.

[2] V. Bach, J. Fröhlich, and L. Jonsson, *Bogolubov-Hartree-Fock mean field theory for neutron stars and other systems with attractive interactions*, J. Math. Phys., 50 (2009), pp. 102102, 22.

[3] V. Bach, E. H. Lieb, and J. P. Solovej, *Generalized Hartree-Fock theory and the Hubbard model*, J. Statist. Phys., 76 (1994), pp. 3–89.

[4] É. Cancès, *SCF algorithms for HF electronic calculations*, in Mathematical models and methods for ab initio quantum chemistry, vol. 74 of Lecture Notes in Chem., Springer, Berlin, 2000, ch. 2, pp. 17–43.

[5] É. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris, and Y. Maday, *Computational quantum chemistry: a primer*, in Handbook of numerical analysis, Vol. X, Handb. Numer. Anal., X, North-Holland, Amsterdam, 2003, pp. 3–270.

[6] É. Cancès and C. Le Bris, *Can we outperform the DIIS approach for electronic structure calculations?*, Int. J. Quantum Chem., 79 (2000), pp. 82–90.

[7] ——, *On the convergence of SCF algorithms for the Hartree-Fock equations*, M2AN Math. Model. Numer. Anal., 34 (2000), pp. 749–774.

[8] É. Cancès, C. Le Bris, and Y. Maday, *Méthodes mathématiques en chimie quantique. Une introduction*, vol. 53 of Collection Mathématiques et Applications, Springer, 2006.

[9] E. Davies, *Spectral theory and differential operators*, vol. 42 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1995.

[10] J. Dechargé and D. Gogny, *Hartree-Fock-Bogolyubov calculations with the $D1$ effective interaction on spherical nuclei*, Phys. Rev. C, 21 (1980), pp. 1568–1593.

[11] J. Dolbeault, A. Laptev, and M. Loss, *Lieb-Thirring inequalities with improved constants*, J. Eur. Math. Soc. (JEMS), 10 (2008), pp. 1121–1126.

[12] C. Fefferman and R. de la Llave, *Relativistic stability of matter. I*, Rev. Mat. Iberoamericana, 2 (1986), pp. 119–213.

[13] R. L. Frank, C. Hainzl, S. Naboko, and R. Seiringer, *The critical temperature for the BCS equation at weak coupling*, J. Geom. Anal., 17 (2007), pp. 559–567.

[14] R. L. Frank, C. Hainzl, R. Seiringer, and J. P. Solovej, *Microscopic Derivation of Ginzburg-Landau Theory*, J. Amer. Math. Soc., in press (2012).

[15] G. Friesecke, *The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions*, Arch. Ration. Mech. Anal., 169 (2003), pp. 35–71.

[16] D. Gogny, in Proceedings of the International Conference on Nuclear Physics, J. de Boer and H. Mang, eds., 1973, p. 48.

[17] ——, in Proceedings of the International Conference on Nuclear Self-Consistent Fields, M. Porneuf and G. Ripka, eds., Trieste, 1975, p. 333.

[18] D. Gogny and P.-L. Lions, *Hartree-Fock theory in nuclear physics*, RAIRO Modél. Math. Anal. Numér., 20 (1986), pp. 571——637.

[19] C. Hainzl, E. Hamza, R. Seiringer, and J. P. Solovej, *The BCS functional for general pair interactions*, Comm. Math. Phys., 281 (2008), pp. 349–367.

[20] C. Hainzl, E. Lenzmann, M. Lewin, and B. Schlein, *On blowup for time-dependent generalized Hartree-Fock equations*, Ann. Henri Poincaré, 11 (2010), pp. 1023–1052.

[21] C. Hainzl and R. Seiringer, *General decomposition of radial functions on $\mathbb{R}^n$ and applications to N-body quantum systems*, Lett. Math. Phys., 61 (2002), pp. 75–84.

[22] ——, *The BCS critical temperature for potentials with negative scattering length*, Lett. Math. Phys., 84 (2008), pp. 99–107.

[23] M. Hoffmann-Ostenhof and T. Hoffmann-Ostenhof, *Schrödinger inequalities and asymptotic behavior of the electron density of atoms and molecules*, Phys. Rev. A, 16 (1977), pp. 1782–1785.

[24] T. Kato, *Perturbation theory for linear operators*, Springer, second ed., 1995.

[25] C. Le Bris, *Computational chemistry from the perspective of numerical analysis*, Acta Numerica, 14 (2005), pp. 363–444.

[26] E. Lenzmann and M. Lewin, *Minimizers for the Hartree-Fock-Bogoliubov theory of neutron stars and white dwarfs*, Duke Math. J., 152 (2010), pp. 257–315.

[27] A. Levitt, *Convergence of gradient-based algorithms for the Hartree-Fock equations*, ESAIM: Mathematical Modelling and Numerical Analysis, 46 (2012), pp. 1321–1336.

[28] M. Lewin, *Geometric methods for nonlinear many-body quantum systems*, J. Funct. Anal., 260 (2011), pp. 3535–3595.

[29] E. H. Lieb, *Variational principle for many-fermion systems*, Phys. Rev. Lett., 46 (1981), pp. 457–459.

[30] E. H. Lieb and M. Loss, *Analysis*, vol. 14 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2001.

[31] E. H. Lieb and R. Seiringer, *The Stability of Matter in Quantum Mechanics*, Cambridge Univ. Press, 2010.

[32] E. H. Lieb and B. Simon, *The Hartree-Fock theory for Coulomb systems*, Commun. Math. Phys., 53 (1977), pp. 185–194.

[33] E. H. Lieb and W. E. Thirring, *Bound on kinetic energy of fermions which proves stability of matter*, Phys. Rev. Lett., 35 (1975), pp. 687–689.

[34] ——, *Inequalities for the moments of the eigenvalues of the Schrödinger hamiltonian and their relation to Sobolev inequalities*, Studies in Mathematical Physics, Princeton University Press, 1976, pp. 269–303.

[35] ——, *Gravitational collapse in quantum mechanics with relativistic kinetic energy*, Ann. Physics, 155 (1984), pp. 494–512.

[36] E. H. Lieb and H.-T. Yau, *The Chandrasekhar theory of stellar collapse as the limit of quantum mechanics*, Commun. Math. Phys., 112 (1987), pp. 147–174.

[37] P.-L. Lions, *Solutions of Hartree-Fock equations for Coulomb systems*, Commun. Math. Phys., 109 (1987), pp. 33–97.

[38] P. Quentin and H. Flocard, *Self-Consistent Calculations of Nuclear Properties with Phenomenological Effective Forces*, Ann. Rev. Nucl. Part. Sci., 28 (1978), pp. 523–594.

[39] P. Ring and P. Schuck, *The nuclear many-body problem*, vol. Texts and Monographs in Physics, Springer Verlag, New York, 1980.

[40] Scilab Consortium, *Scilab: The free software for numerical computation*, Scilab Consortium, Digiteo, Paris, France, 2011.

[41] T. Skyrme, *The effective nuclear potential*, Nuclear Physics, 9 (1959), pp. 615–634.

[42] W. E. Thirring, *Quantum Mathematical Physics*, vol. Atoms, Molecules and Large Systems, Springer, Second Edition 2002.

[43] S. S. M. Wong, *Introductory Nuclear Physics*, Wiley-VCH Verlag GmbH, 2007.