

Simulation numérique des modèles

Ce chapitre est une introduction à la résolution numérique des modèles issus de la chimie quantique qui ont été décrits d'un point de vue théorique au cours des chapitres précédents.

La section 6.2, qui concerne le problème électronique, en constitue le cœur. Les méthodes qui y sont exposées (développement sur une base d'orbitales atomiques, résolution des équations d'Euler-Lagrange par un algorithme SCF) seront examinées plus en détail au sein des Chapitres 7 et 8 dans le cadre du modèle Hartree-Fock.

La section 6.3 aborde ensuite brièvement la résolution numérique du problème d'optimisation de géométrie. On fera à cette occasion quelques rappels sur les techniques standard de l'optimisation numérique sans contraintes.

La section 6.1 présente quant à elle le seul exemple (ou presque) de système moléculaire pour lequel on dispose d'une expression analytique du fondamental et des états excités. Il s'agit de l'ion hydrogénoïde.

6.1 Prolégomène : étude de l'ion hydrogénoïde

Un ion hydrogénoïde est par définition un système moléculaire comportant un seul noyau, de charge Z , et un seul électron. Un cas particulier d'ion hydrogénoïde est l'atome d'Hydrogène lui-même, qui correspond à $Z = 1$.

L'exemple de l'ion hydrogénoïde est intéressant à plusieurs égards. D'un point de vue historique, c'est sur le calcul du spectre de l'hydrogène atomique que Schrödinger a montré la validité de l'équation qui porte son nom (c'était en 1926). Par ailleurs, l'étude de l'ion hydrogénoïde permet non seulement de comprendre en profondeur la structure de l'atome d'hydrogène mais aussi d'obtenir des informations qualitatives sur des systèmes voisins, en particulier

sur les électrons de cœur des atomes lourds¹ et sur l'électron de valence des alcalins (lithium, sodium, potassium, ...).

L'hamiltonien de l'ion hydrogénoïde s'écrit sous l'approximation de Born-Oppenheimer

$$H = -\frac{1}{2}\Delta - \frac{Z}{|x|}. \quad (6.1)$$

C'est un opérateur auto-adjoint sur $L^2(\mathbb{R}^3)$ de domaine $D(H) = H^2(\mathbb{R}^3)$. Les théorèmes généraux de la théorie spectrale (cf. annexe B) nous fournissent des renseignements qualitatifs sur le spectre de cet opérateur :

1. le potentiel $-\frac{Z}{|x|}$ étant une perturbation compacte de l'opérateur $H_0 = -\frac{1}{2}\Delta$, le spectre essentiel de H est celui de H_0 : $\sigma_{ess}(H) = \sigma_{ess}(H_0) = [0, +\infty[$;
2. H n'a pas de valeurs propres positives ou nulles ;
3. H possède une infinité de valeurs propres strictement négatives, de multiplicité finie, qui forment une suite croissante tendant vers 0.

Comme nous allons le voir dans les deux sections suivantes, il est possible d'aller plus loin et d'explicitier les valeurs propres et les vecteurs propres de H .

6.1.1 Identification du fondamental

Le fondamental normalisé de H est relativement facile à déterminer dès qu'on a prouvé son unicité (à une phase globale près).

Considérons à cette fin un fondamental normalisé à valeurs réelles ψ_0 de H dont la valeur en un certain point $x_0 \neq 0$ est strictement positive. Remarquons maintenant que ψ_0 est solution du problème de minimisation

$$\inf \left\{ \langle \psi, H\psi \rangle = \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \psi|^2 - Z \int_{\mathbb{R}^3} \frac{|\psi|^2}{|x|}, \quad \psi \in H^1(\mathbb{R}^3), \quad \|\psi\|_{L^2} = 1 \right\}$$

et que toute solution de ce problème de minimisation est un fondamental normalisé. On voit donc en particulier que $|\psi_0|$ est un fondamental normalisé (car si $\psi \in H^1(\mathbb{R}^3)$, $|\psi| \in H^1(\mathbb{R}^3)$ et vérifie $|\nabla |\psi|| = |\nabla \psi|$ presque partout). Il en résulte que $|\psi_0|$ vérifie l'équation

$$-\frac{1}{2}\Delta |\psi_0| - \frac{Z}{|x|} |\psi_0| = E_0 |\psi_0|.$$

¹ Pas trop lourds quand même ! Lorsque Z est grand, l'énergie cinétique d'un électron de cœur n'est plus négligeable devant l'énergie de masse $E = mc^2$; on pourra faire le calcul à titre d'exercice pour l'Uranium ($Z = 92$). Pour calculer la dynamique des électrons de cœur des atomes lourds, il faut utiliser un modèle relativiste comme celui de Dirac-Fock.

On en conclut par régularité elliptique et à l'aide de l'inégalité de Harnack que $|\psi_0|$ est continue et strictement positive. On a donc nécessairement $|\psi_0| = \psi_0 > 0$ sur \mathbb{R}^3 . Supposons maintenant que la valeur propre fondamentale E_0 soit de multiplicité supérieure ou égale à 2. Si c'était le cas, on pourrait trouver un autre fondamental réel normalisé de H , noté $\tilde{\psi}_0$, orthogonal à ψ_0 . En reprenant le raisonnement ci-dessus, on pourrait écrire $\tilde{\psi}_0 = \pm \phi$, ϕ désignant un fondamental normalisé de H réel strictement positif sur \mathbb{R}^3 . On aurait alors

$$0 = \pm \int_{\mathbb{R}^3} \psi_0 \tilde{\psi}_0 = \int_{\mathbb{R}^3} \psi_0 \phi > 0,$$

ce qui met en évidence une contradiction. Ceci prouve l'unicité du fondamental normalisé (au signe près ou à une phase globale près si on considère des vecteurs propres à valeurs complexes).

Remarque 6.1 *Ce raisonnement classique s'étend à tout hamiltonien de la forme $H = -\frac{1}{2}\Delta + V$, V désignant un opérateur de multiplication suffisamment régulier ($V \in L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ convient) : si H admet un fondamental, alors ce fondamental est unique, à une constante multiplicative près, qui peut être choisie telle que le fondamental soit strictement positif presque partout.*

Remarquons maintenant que pour toute rotation R de l'espace, $\psi_0(R \cdot x)$ est aussi un fondamental normalisé strictement positif. L'argument d'unicité exposé ci-dessus permet de conclure que ψ_0 est une fonction radiale : $\psi_0(x) = f(|x|)$, f désignant la solution strictement positive normalisée (dans $L^2(\mathbb{R}^+, 4\pi r^2 dr)$) de l'équation aux valeurs propres posée sur \mathbb{R}^{+*}

$$-\frac{1}{2}f''(r) - \frac{f'(r)}{r} - \frac{Zf(r)}{r} = E_0 f(r).$$

On voit facilement que $f(r) = \exp(-Zr)$ est une solution strictement positive de cette équation associée à la valeur propre $E_0 = -Z^2/2$.

Le fondamental normalisé de H est donc donné par

$$\psi_0(x) = \frac{Z^{3/2}}{\sqrt{\pi}} e^{-Z|x|}$$

et la valeur propre fondamentale de H vaut $E_0 = -Z^2/2$.

Remarque 6.2 *On peut déterminer analytiquement l'énergie fondamentale de l'ion hydrogénoïde sans avoir recours à l'approximation de Born-Oppenheimer. Comme en mécanique classique, un problème de mécanique quantique à deux corps peut être en effet ramené à l'étude d'un problème à un corps. Écrivons pour cela l'hamiltonien complet du système*

$$H = -\frac{1}{2M}\Delta_{\bar{x}} - \frac{1}{2m}\Delta_x - \frac{Z}{|x - \bar{x}|}$$

en fonction des variables

$$x_G = \frac{M\bar{x} + mx}{M+m} \quad \text{et} \quad x_r = x - \bar{x},$$

qui correspondent respectivement à la position du centre de masse et à la position relative de l'électron par rapport au noyau. En notant $\mu = mM/(m+M)$ la masse réduite du système, on obtient ainsi l'hamiltonien

$$H = -\frac{1}{2(m+M)}\Delta_{x_G} - \frac{1}{2\mu}\Delta_{x_r} - \frac{Z}{|x_r|}$$

dans lequel mouvement du centre de masse et mouvement relatif des deux particules sont dissociés. L'énergie fondamentale de cet hamiltonien est donc obtenue en minimisant séparément l'énergie E_G de translation du centre de masse et l'énergie du mouvement relatif. On obtient facilement $E_G = 0$ (construire à titre d'exercice une suite minimisante) et on déduit de ce qui précède $E_r = -\mu Z^2/2$. Sous l'approximation de Born-Oppenheimer ($M = +\infty$, et donc $\mu = 1$), on retrouve bien l'expression (6.1) de l'hamiltonien et l'énergie fondamentale $E_0 = -Z^2/2$.

6.1.2 Description complète du spectre discret

On cherche maintenant à exhiber toute la suite des valeurs propres (et des vecteurs propres) de H . Afin de rendre la lecture plus digeste, on se contente dans cette section de donner le principe et les résultats de cette recherche. Le lecteur intéressé par les détails techniques est invité à se plonger dans les problèmes 6.1 et 6.2.

Remarquons tout d'abord qu'en coordonnées sphériques, l'hamiltonien H s'écrit

$$H = -\frac{1}{2r}\frac{\partial^2}{\partial r^2}r - \frac{1}{2r^2}\Delta_S - \frac{Z}{r},$$

Δ_S désigne l'opérateur de Laplace-Beltrami. Cet opérateur, qui est l'analogue du laplacien sur la variété S^2 , n'agit que sur les variables θ et ϕ :

$$\Delta_S = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}.$$

Les vecteurs propres de Δ_S sur la sphère S^2 sont connus (voir le rappel ci-dessous) : il s'agit des harmoniques sphériques Y_l^m . On remarque ensuite que H et Δ_S (Δ_S est vu désormais comme un opérateur agissant sur \mathbb{R}^3) commutent. On peut donc construire une base de vecteurs propres de H en imposant que ces vecteurs propres soient aussi vecteurs propres de Δ_S . Ces fonctions s'écrivent donc en coordonnées sphériques sous la forme

$$\psi(r, \theta, \phi) = f(r) Y_l^m(\theta, \phi) \tag{6.2}$$

où f est une fonction appropriée de $L^2(\mathbb{R}^+, 4\pi r^2 dr)$.

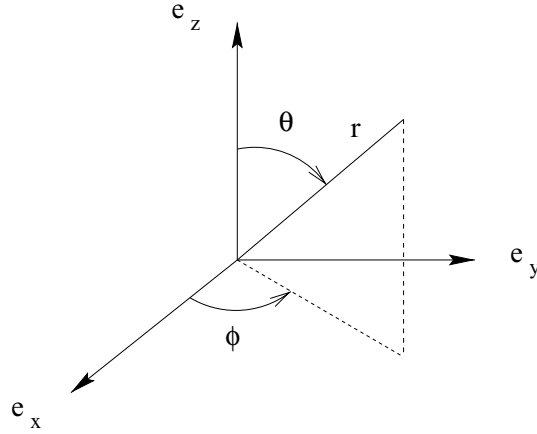


Fig. 6.1. Coordonnées sphériques.

Notons par ailleurs que $\Delta_S = -L^2$ où $L = -ix \times \nabla$ désigne l'opérateur de moment cinétique (cf. Annexe A), et que les fonctions de la forme (6.2) pour f fonction quelconque de $L^2(\mathbb{R}^+, 4\pi r^2 dr)$, sont exactement les fonctions propres communes aux opérateurs L^2 et $L_z = L \cdot e_z = -i \frac{\partial}{\partial \phi}$. Nous sommes donc en train de chercher une base de vecteurs propres de H en imposant que ces vecteurs propres soient aussi vecteurs propres de L^2 et de L_z . Notons enfin que H s'écrit aussi

$$H = -\frac{1}{2r} \frac{\partial^2}{\partial r^2} r + \frac{1}{2r^2} L^2 - \frac{Z}{r}.$$

Il n'est peut-être pas inutile à ce stade de faire un bref

Rappel. Les harmoniques sphériques $(Y_l^m)_{l \geq 0, -l \leq m \leq l}$ forment une base hilbertienne de $L^2(S^2)$:

- on a pour tout (l, m) , $l \geq 0$, $-l \leq m \leq l$ et tout (l', m') , $l' \geq 0$, $-l' \leq m' \leq l'$

$$(Y_l^m, Y_{l'}^{m'})_{L^2(S^2)} = \int_0^\pi \int_0^{2\pi} Y_l^m(\theta, \phi)^* Y_{l'}^{m'}(\theta, \phi) \sin \theta d\theta d\phi = \delta_{ll'} \delta_{mm'} ;$$

- toute fonction $f \in L^2(S^2)$ se décompose de manière unique sur la base des harmoniques sphériques :

$$f = \sum_{l=0}^{+\infty} \sum_{m=-l}^l (Y_l^m, f)_{L^2(S^2)} Y_l^m,$$

la convergence de la série s'entendant au sens de la convergence L^2 .

Cette base est caractérisée par les relations

$$\Delta_S \cdot Y_l^m = -l(l+1)Y_l^m \quad \text{et} \quad -i \frac{\partial}{\partial \phi} Y_l^m(\theta, \phi) = m Y_l^m(\theta, \phi). \quad (6.3)$$

L'expression des harmoniques sphériques en fonction de θ et ϕ est donnée par

$$Y_l^m(\theta, \phi) = \left(\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!} \right)^{1/2} P_l^m(\cos \theta) e^{im\phi}$$

où les $P_l^m(\cos \theta)$ désignent les fonctions de Legendre. Celles-ci sont obtenues à partir des polynômes de Legendre² par dérivation :

$$\begin{cases} P_l^m(\cos \theta) = (\sin \theta)^m P_l^{(m)}(\cos \theta) & \text{pour } 0 \leq m \leq l, \\ P_l^{-m}(\cos \theta) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(\cos \theta) & \text{pour } 0 \leq m \leq l, \end{cases}$$

la notation $P_l^{(m)}$ désignant la dérivée m -ième du polynôme de Legendre P_l . Les premiers représentants de la famille des harmoniques sphériques sont :

$$\begin{aligned} Y_0^0(\theta, \phi) &= \frac{1}{\sqrt{4\pi}} \\ Y_1^0(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \cos \theta, \quad Y_1^{\pm 1}(\theta, \phi) = \pm \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \\ Y_2^0(\theta, \phi) &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1), \quad Y_2^{\pm 1}(\theta, \phi) = \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi}, \\ Y_2^{\pm 2}(\theta, \phi) &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}. \end{aligned}$$

Ce rappel étant fait, revenons à notre problème et cherchons donc les états propres de H sous la forme

$$\psi(r, \theta, \phi) = R(r) Y_l^m(\theta, \phi).$$

En reportant dans l'équation aux valeurs propres $H\psi = E\psi$ et en utilisant (6.3), on obtient l'équation radiale

$$-\frac{1}{2r} \frac{d^2}{dr^2} (r R(r)) + \frac{l(l+1)}{2r^2} R(r) - \frac{Z}{r} R(r) = E R(r). \quad (6.4)$$

A l fixé, l'équation (6.4) possède une infinité dénombrable de solutions normalisées (i.e. telle que $\int_0^{+\infty} r^2 |R(r)|^2 dr = 1$). Chacune d'elle peut être caractérisée (à une phase globale près) par un entier $n > l$; la solution $R_{nl}(r)$ est donnée par

² La base des polynômes de Legendre $P_l(x)$ est l'orthonormalisée de Gramm-Schmidt de la base des monômes $(1, x, x^2, x^3, \dots)$ pour le produit scalaire usuel de $L^2(-1, 1)$. En fait, l'usage est plutôt d'utiliser le procédé de normalisation consistant à imposer $P_l(1) = 1$. On a ainsi $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{1}{2}(3x^2 - 1)$, et plus généralement (formule de Rodrigues) $P_l(x) = \frac{(-1)^l}{2^l l!} \frac{d^l}{dx^l} (1 - x^2)^l$.

$$R_{nl}(r) = Q_{nl}(Zr) e^{-Zr}$$

$Q_{nl}(x) = x^l(C_0 + C_1x + \cdots + C_{n-l-1}x^{n-l-1})$ désignant le polynôme dont les coefficients sont définis par la relation de récurrence

$$C_i = -\frac{2}{n} \frac{n-l-i}{i(2l+i+1)} C_{i-1},$$

C_0 étant déterminé à une phase près par la relation de normalisation. Il lui correspond la valeur propre $E_{nl} = -Z^2/2n^2$ qui est indépendante de l . Soulignons que cette situation est exceptionnelle : si on remplace le potentiel coulombien par un autre potentiel central possédant même décroissance à l'infini et même singularité à l'origine, on obtient qualitativement des résultats similaires, mais l'énergie E_{nl} dépend cette fois-ci effectivement de n et de l .

En tenant compte de cette dégénérescence "accidentelle", on peut reformuler plus clairement les résultats ci-dessus de la façon suivante :

- les valeurs propres de H forment une suite $(E_n)_{n \geq 1}$ définie par

$$E_n = -\frac{Z^2}{2n^2};$$

- la valeur propre E_n est de multiplicité n^2 ;
- les n^2 vecteurs

$$\psi_{nlm}(r, \theta, \phi) = Q_{nl}(Zr) e^{-Zr} Y_l^m(\theta, \phi), \quad 0 \leq l \leq n-1, \quad -l \leq m \leq l,$$

définissent une base du sous-espace propre associé à la valeur propre E_n formée de vecteurs propres des opérateurs L^2 et L_z ;

- on a

$$H\psi_{nlm} = E_n\psi_{nlm}, \quad L^2\psi_{nlm} = l(l+1)\psi_{nlm}, \quad L_z\psi_{nlm} = m\psi_{nlm}$$

et ψ_{nlm} est l'unique vecteur (à une constante multiplicative près) satisfaisant simultanément ces trois relations : le triplet d'observables (H, L^2, L_z) forme un Ensemble Complet d'Observables qui Commutent (cf. section A.2.2.4).

Les entiers n , l et m sont appelés respectivement *nombre quantique principal*, *nombre quantique azimutal* et *nombre quantique magnétique*. Il est d'usage en chimie de caractériser les fonctions propres ψ_{nlm} par un entier (le nombre quantique principal n) suivi d'une lettre identifiant le nombre quantique azimutal l selon la règle de correspondance

$$l = 0 \longrightarrow s, \quad l = 1 \longrightarrow p, \quad l = 2 \longrightarrow d, \quad l = 3 \longrightarrow f.$$

L'état ψ_{100} est ainsi dénommé 1s, les trois états ψ_{21m} 2p, les cinq états ψ_{32m} 3d, etc.

6.2 Résolution numérique du problème électronique

Dans les sections 6.2.1 à 6.2.5, on raisonne pour fixer les idées sur le modèle de Hartree-Fock sans spin dont on rappelle ici la forme :

$$\inf \{ E^{HF}(\Phi), \quad \Phi \in \mathcal{W}_N \}, \quad (6.5)$$

avec

$$\mathcal{W}_N = \left\{ \Phi = \{ \phi_i \}_{1 \leq i \leq N}, \quad \phi_i \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}, \quad 1 \leq i, j \leq N \right\},$$

$$\begin{aligned} E^{HF}(\Phi) = & \sum_{i=1}^N \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \phi_i|^2 + \int_{\mathbb{R}^3} \rho_\Phi V \\ & + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\Phi(x) \rho_\Phi(x')}{|x - x'|} dx dx' \\ & - \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|\tau_\Phi(x, x')|^2}{|x - x'|} dx dx', \end{aligned}$$

$$\tau_\Phi(x, x') = \sum_{i=1}^N \phi_i(x) \phi_i(x') \quad \text{et} \quad \rho_\Phi(x) = \tau_\Phi(x, x) = \sum_{i=1}^N |\phi_i(x)|^2. \quad (6.6)$$

Soulignons cependant que les mêmes techniques s'appliquent pour les modèles RHF et UHF utilisés en pratique³. La résolution numérique des modèles de type Kohn-Sham issus de la théorie de la fonctionnelle de la densité, est discutée brièvement dans la section 6.2.6.

6.2.1 Approximation de Galerkin

Dans toute la suite, on note $\mathcal{M}(m, n)$ l'espace vectoriel des matrices $m \times n$ à coefficients réels. et $\mathcal{M}_S(n)$ l'espace vectoriel des matrices $n \times n$ symétriques à coefficients réels ;

Pour résoudre numériquement un problème de Hartree-Fock, la méthode la plus efficace consiste à utiliser une approximation de Galerkin⁴. Plus précisément, on approche le problème (6.5) par

$$\inf \{ E^{HF}(\Phi), \quad \Phi \in \mathcal{W}_N(\mathcal{V}) \} \quad (6.7)$$

³ Pour le modèle RHF, le formalisme est similaire ; pour le modèle UHF, il faut considérer deux matrices de coefficients C_α et C_β rassemblant les coefficients dans la base d'orbitales atomiques des spin-orbitales de type α et β respectivement.

⁴ Voir cependant [140] concernant l'utilisation de méthodes de grilles (différences finies, multigrilles).

avec

$$\mathcal{W}_N(\mathcal{V}) = \left\{ \Phi = \{\phi_i\}_{1 \leq i \leq N}, \quad \phi_i \in \mathcal{V}, \quad \int_{\mathbf{R}^3} \phi_i \phi_j = \delta_{ij}, \quad 1 \leq i, j \leq N \right\},$$

où \mathcal{V} est un sous-espace $H^1(\mathbf{R}^3)$ de dimension finie N_b . Soit $\{\chi_\mu\}_{1 \leq \mu \leq N_b}$ une base de \mathcal{V} et $S \in \mathcal{M}_S(N_b)$ la matrice de recouvrement définie par

$$S_{\mu\nu} = \int_{\mathbf{R}^3} \chi_\mu \chi_\nu.$$

À $\Phi = \{\phi_i\} \in \mathcal{W}_N(\mathcal{V})$, on associe la matrice $C \in \mathcal{M}(N_b, N)$ rassemblant les coefficients des ϕ_i dans la base $\{\chi_\mu\}$:

$$\forall 1 \leq i \leq N, \quad \phi_i(x) = \sum_{\mu=1}^{N_b} C_{\mu i} \chi_\mu(x).$$

Les contraintes $\int_{\mathbf{R}^3} \phi_i \phi_j = \delta_{ij}$, s'écrivent alors

$$\delta_{ij} = \int_{\mathbf{R}^3} \phi_i \phi_j = \int_{\mathbf{R}^3} \left(\sum_{\nu=1}^{N_b} C_{\nu i} \chi_\nu \right) \left(\sum_{\mu=1}^{N_b} C_{\mu j} \chi_\mu \right) = \sum_{\mu=1}^{N_b} \sum_{\nu=1}^{N_b} C_{\mu j} S_{\mu\nu} C_{\nu i},$$

soit encore sous forme matricielle

$$C^T S C = I_N,$$

I_N désignant la matrice identité de rang N . Par ailleurs,

$$\begin{aligned} \sum_{i=1}^N \frac{1}{2} \int_{\mathbf{R}^3} |\nabla \phi_i|^2 + \int_{\mathbf{R}^3} \rho_\Phi V &= \sum_{i=1}^N \left(\frac{1}{2} \int_{\mathbf{R}^3} |\nabla \phi_i|^2 + \int_{\mathbf{R}^3} V |\phi_i|^2 \right) \\ &= \sum_{i=1}^N \left(\frac{1}{2} \int_{\mathbf{R}^3} \left| \nabla \sum_{\mu=1}^{N_b} C_{\mu i} \chi_\mu \right|^2 + \int_{\mathbf{R}^3} V \left| \sum_{\mu=1}^{N_b} C_{\mu i} \chi_\mu \right|^2 \right) \\ &= \sum_{i=1}^N \sum_{\mu=1}^{N_b} \sum_{\nu=1}^{N_b} h_{\mu\nu} C_{\nu i} C_{\mu i} \\ &= \text{Tr} (h C C^T), \end{aligned}$$

où h désigne la matrice de l'hamiltonien de cœur $-\frac{1}{2}\Delta + V$ dans la base $\{\chi_\mu\}$:

$$h_{\mu\nu} = \frac{1}{2} \int_{\mathbf{R}^3} \nabla \chi_\mu \cdot \nabla \chi_\nu + \int_{\mathbf{R}^3} V \chi_\mu \chi_\nu.$$

En notant

$$(\mu\nu|\kappa\lambda) = \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{\chi_\mu(x)\chi_\nu(x)\chi_\kappa(x')\chi_\lambda(x')}{|x-x'|} dx dx',$$

et pour toute matrice $D \in \mathcal{M}(N_b, N_b)$,

$$[J(D)]_{\mu\nu} = \sum_{\kappa, \lambda=1}^{N_b} (\mu\nu|\kappa\lambda) D_{\kappa\lambda}, \quad [K(D)]_{\mu\nu} = \sum_{\kappa, \lambda=1}^{N_b} (\mu\lambda|\kappa\nu) D_{\kappa\lambda},$$

on obtient de même l'expression des termes de répulsion interélectroniques en fonction de C

$$\begin{aligned} \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{\rho_\Phi(x)\rho_\Phi(x')}{|x-x'|} dx dx' &= \sum_{\mu, \nu, \kappa, \lambda=1}^{N_b} \sum_{i, j=1}^N (\mu\nu|\kappa\lambda) C_{\mu i} C_{\nu i} C_{\kappa j} C_{\lambda j} \\ &= \text{Tr} (J(CC^T)CC^T), \end{aligned}$$

$$\begin{aligned} \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{|\tau_\Phi(x, x')|^2}{|x-x'|} dx dx' &= \sum_{\mu, \nu, \kappa, \lambda=1}^{N_b} \sum_{i, j=1}^N (\mu\lambda|\kappa\nu) C_{\mu i} C_{\nu i} C_{\kappa j} C_{\lambda j} \\ &= \text{Tr} (K(CC^T)CC^T). \end{aligned}$$

On aboutit donc au problème de minimisation

$$\inf \{ E^{HF}(CC^T), \quad C \in \mathcal{W}_N \} \quad (6.8)$$

où

$$E^{HF}(D) = \text{Tr}(hD) + \frac{1}{2}\text{Tr}(J(D)D) - \frac{1}{2}\text{Tr}(K(D)D),$$

et

$$\mathcal{W}_N = \{ C \in \mathcal{M}(N_b, N), \quad C^T S C = I_N \}.$$

Aucune confusion avec le problème de départ (6.5) n'étant à craindre, nous avons conservé par commodité les mêmes notations E^{HF} et \mathcal{W}_N pour désigner l'énergie de Hartree-Fock et l'ensemble de minimisation dans l'approximation de Galerkin (6.8). En posant

$$G(D) = J(D) - K(D),$$

on peut écrire l'énergie de Hartree-Fock sous la forme plus concise

$$E^{HF}(D) = \text{Tr}(hD) + \frac{1}{2}\text{Tr}(G(D)D).$$

Remarquons qu'en raison des symétries du tenseur à quatre indices $(\mu\nu|\kappa\lambda)$, on a pour toutes matrices D et D' de $\mathcal{M}(N_b, N_b)$,

$$\begin{aligned}\mathrm{Tr}(J(D)D') &= \mathrm{Tr}(J(D')D), & \mathrm{Tr}(K(D)D') &= \mathrm{Tr}(K(D')D), \\ \mathrm{Tr}(G(D)D') &= \mathrm{Tr}(G(D')D).\end{aligned}$$

Remarquons également que la matrice $D = CC^T$ intervenant dans l'expression (6.8) n'est autre que la représentation dans la base $\{\chi_\mu\}$ de la matrice densité réduite $\tau_\Phi(x, x')$ définie par (6.6)⁵. En effet,

$$\begin{aligned}\tau_\Phi(x, x') &= \sum_{i=1}^N \phi_i(x) \phi_i(x') = \sum_{i=1}^N \left(\sum_{\mu=1}^{N_b} C_{\mu i} \chi_\mu(x) \right) \left(\sum_{\nu=1}^{N_b} C_{\nu i} \chi_\nu(x') \right) \\ &= \sum_{\mu, \nu=1}^{N_b} \left(\sum_{i=1}^N C_{\mu i} C_{\nu i} \right) \chi_\mu(x) \chi_\nu(x') = \sum_{\mu, \nu=1}^{N_b} D_{\mu\nu} \chi_\mu(x) \chi_\nu(x').\end{aligned}$$

Pour cette raison, la matrice $D = CC^T$ est appelée *matrice densité*. Notons enfin qu'il est possible de reformuler le problème de Hartree-Fock discrétisé (6.8) en un problème de minimisation ne faisant intervenir que la matrice densité $D = CC^T$. On montre en fait facilement (exercice 6.2) que

$$\{D = CC^T, \quad C \in \mathcal{W}_N\} = \{D \in \mathcal{M}_S(N_b), \quad DSD = D, \quad \mathrm{Tr}(SD) = N\}, \quad (6.9)$$

$\mathcal{M}_S(N_b)$ désignant l'espace vectoriel des matrices symétriques de taille $N_b \times N_b$. On notera désormais :

$$\mathcal{P}_N = \{D \in \mathcal{M}_S(N_b), \quad DSD = D, \quad \mathrm{Tr}(SD) = N\}.$$

On voit immédiatement que si $S = I_{N_b}$, autrement dit si la base $\{\chi_\mu\}$ est orthonormale, \mathcal{P}_N n'est autre que l'ensemble des projecteurs orthogonaux de rang N sur \mathbb{R}^{N_b} ; il s'agit d'une variété grassmannienne, objet bien connu en géométrie différentielle [83]. Le problème de Hartree-Fock discrétisé peut donc s'écrire

$$\inf \{E^{HF}(D), \quad D \in \mathcal{P}_N\}. \quad (6.10)$$

On se trouve finalement confronté à un problème de minimisation sous contraintes, (6.8) ou (6.10), posé sur un espace de dimension finie ; il existe pour résoudre ce problème, toute une batterie de méthodes plus ou moins efficaces, nous en reparlerons à la section 6.2.5.

6.2.2 Bases d'orbitales atomiques

Le choix de l'espace \mathcal{V} va bien évidemment conditionner la qualité du résultat du calcul. Un choix judicieux consiste à prendre un espace \mathcal{V} engendré par une base d'orbitales atomiques (OA). On parle alors d'approximation LCAO (*linear combination of atomic orbitals*).

⁵ Notons cependant que la matrice dans la base $\{\chi_\mu\}$ de la restriction de l'opérateur densité \mathcal{D}_Φ à l'espace \mathcal{V} est DS (et non D).

Les bases d'orbitales atomiques fonctionnent selon le principe suivant :

1. A chaque élément chimique A du tableau périodique, on associe une collection $\{\xi_\mu^A\}_{1 \leq \mu \leq n_A}$ de n_A fonctions de $H^1(\mathbb{R}^3)$ linéairement indépendantes : ce sont les orbitales atomiques relatives à l'élément chimique A.
2. Pour effectuer un calcul sur un système moléculaire donné, on construit la base $\{\chi_\mu\}$ en regroupant toutes les orbitales atomiques relatives à tous les atomes du système. Ainsi par exemple, pour résoudre le problème électronique associé à une molécule d'eau H_2O , on prend

$$\{\chi_\mu\} = \{\xi_1^H(x - \bar{x}_{H_1}), \dots, \xi_{n_H}^H(x - \bar{x}_{H_1}); \xi_1^H(x - \bar{x}_{H_2}), \dots, \xi_{n_H}^H(x - \bar{x}_{H_2}); \xi_1^O(x - \bar{x}_O), \dots, \xi_{n_O}^O(x - \bar{x}_O)\},$$

où on a noté \bar{x}_{H_1} , \bar{x}_{H_2} et \bar{x}_O les positions respectives dans \mathbb{R}^3 des deux noyaux d'Hydrogène et du noyau d'Oxygène. On a donc ainsi pour cet exemple $n = 2n_H + n_O$ fonctions de base.

Une base d'orbitales atomiques correspond à la donnée des $\{\xi_\mu^A\}_{1 \leq \mu \leq n_A}$ pour tous les éléments du tableau périodique⁶. L'ensemble des OA relatives à l'atome A est optimisé de façon à engendrer de bonnes approximations du fondamental et des premiers états excités de l'atome isolé et de petites molécules simples contenant A. Ce qu'il y a de remarquable avec ces bases, c'est qu'un petit nombre d'OA suffit pour obtenir *in fine* une approximation très précise du fondamental Hartree-Fock pour un système moléculaire quelconque. Ainsi la base d'orbitales atomiques gaussiennes 6-311G++(3df,3pd) ne comporte "que" 18 OA pour l'atome d'hydrogène (1 électron), 39 OA pour l'atome de carbone (6 électrons), 47 OA pour l'atome de magnésium (12 électrons). Elle est pourtant considérée comme une base de grande taille à laquelle on n'a recours que si l'on veut un résultat d'une grande précision.

Il est *a priori* tentant de prendre comme OA des orbitales de Slater, c'est-à-dire des fonctions de la forme

$$\xi(r, \theta, \phi) = P(r)Y_l^m(\theta, \phi)e^{-\alpha r} \quad (6.11)$$

où P est un polynôme, $\alpha > 0$ et Y_l^m une harmonique sphérique, puisque les états propres du seul système moléculaire pour lequel on dispose d'expressions analytiques, à savoir l'ion hydrogénoïde, sont précisément de cette forme (cf. section 6.1). Ce sont d'ailleurs des orbitales de Slater qui ont servi aux premiers calculs de chimie quantique.

Il s'avère cependant plus efficace de prendre comme OA des "gaussiennes contractées" soit en d'autres termes des combinaisons linéaires finies de gaussiennes-polynômes :

⁶ On ne trouve cependant dans les bases de données des codes standard de chimie quantique que les orbitales atomiques correspondant aux éléments chimiques usuels : celui qui veut faire un calcul sur l'Holmium ($Z = 67$) doit lui-même rentrer les caractéristiques des OA relatives à cet atome.

$$\xi(x) = \sum_{k=1}^d c_k x_1^{n_{1,k}} x_2^{n_{2,k}} x_3^{n_{3,k}} e^{-\alpha_k |x|^2}$$

où les $n_{j,k}$ sont des entiers positifs et les α_k des réels positifs optimisés pour minimiser, avec un certain empirisme, le nombre de degré de liberté dans un calcul LCAO. L'intérêt de telles fonctions est qu'elles se prêtent facilement au calcul des N_b^4 intégrales biélectroniques

$$(\mu\nu|\kappa\lambda) = \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{\chi_\mu(x) \chi_\nu(x) \chi_\kappa(x') \chi_\lambda(x')}{|x - x'|} dx dx', \quad (6.12)$$

qui est l'étape limitante de la méthode Hartree-Fock en termes de temps de calcul. Plus précisément,

1. les quantités (6.12) qui s'expriment *a priori* sous la forme d'intégrales sur \mathbf{R}^6 peuvent en fait se ramener à des intégrales sur $[0, 1]$ du type

$$F(w) = \int_0^1 e^{-w s^2} ds$$

lorsque les $\chi_\mu(x) = \xi(x - \bar{x}_k)$ sont des gaussiennes ; ceci fait l'objet de l'exercice 6.3 ;

2. en remarquant que les gaussiennes-polynômes sont engendrées par dérivation à partir des gaussiennes, on peut par intégration par parties établir des relations de récurrence permettant de calculer les intégrales biélectroniques pour des gaussiennes-polynômes de degré d à partir de celles de degré inférieur.

La découverte de cette propriété par Boys [45] a marqué une avancée considérable de la chimie quantique.

La complétude asymptotique des bases d'orbitales atomiques gaussiennes est prouvée au Chapitre 7. Cela signifie qu'on peut approcher l'énergie de Hartree-Fock d'aussi près que l'on veut (par valeur supérieure puisque l'approximation de Galerkin consiste à minimiser la fonctionnelle de Hartree-Fock sur un sous-espace de \mathcal{W}_N) en utilisant une base d'orbitales atomiques gaussiennes.

En pratique, plusieurs bases d'orbitales atomiques gaussiennes sont implémentées dans les codes et il appartient à l'utilisateur de sélectionner l'une d'elle lorsqu'il lance un calcul. Le choix de la base sur laquelle on réalise les calculs est une étape clé : il s'agit de trouver un compromis entre temps de calcul et qualité des résultats. A l'heure actuelle, seuls l'intuition et le savoir-faire du chimiste guident ce choix, qui est donc en quelque sorte une concession de l'*ab initio* à l'empirisme. Pour valider le choix effectué, il faudrait pouvoir répondre à la question suivante : une fois calculée l'énergie de Hartree-Fock dans une base donnée, peut-on savoir si on a obtenu une bonne approximation de l'énergie exacte de Hartree-Fock ? Autrement dit, peut-on estimer l'erreur due à l'incomplétude de la base ? La méthode des estimations *a posteriori* développée au Chapitre 7, fournit un majorant de cette erreur.

Remarque 6.3 *Les bases d'OA ne sont pas les seules utilisées en pratique. On utilise aussi dans certains cas des bases d'ondes planes. Ces bases sont en fait adaptées au cas cristallin, pour lequel on travaille sur un borné, la maille cristalline, avec conditions aux bords périodiques (cf. section 9.1). Pour calculer un système moléculaire non cristallin avec des ondes planes, on périodise le système : on place le système physique dans une “maille élémentaire” rectangulaire suffisamment grande pour que les interactions inter-maillages soient négligeables et on construit un cristal fictif en répliquant indéfiniment la maille élémentaire dans les trois dimensions de l'espace.*

6.2.3 Equations de Hartree-Fock discrétisées

Le modèle de Hartree-Fock discrétisé (6.8) revêt la forme d'un problème de minimisation sous contraintes : il s'agit de trouver l'infimum de la fonctionnelle $E : C \mapsto E^{HF}(CC^T)$ sur l'ensemble des $C \in \mathcal{M}(N_b, N)$ qui vérifient les contraintes égalités $J(C) = C^T SC - I_N = 0$ (J est à valeurs dans $\mathcal{M}_S(N)$). Pour établir les équations d'Euler-Lagrange de ce problème, il faut calculer $\nabla E(C)$ et $J'(C)^T$. Un calcul simple (laissé en exercice) montre que pour tout $\Sigma \in \mathcal{M}_S(N)$,

$$J'(C)^T \cdot \Sigma = 2SC\Sigma,$$

et que pour tout $C \in \mathcal{M}(N_b, N)$

$$\nabla E(C) = F(CC^T)C,$$

où pour toute matrice D

$$F(D) = h + G(D)$$

désigne la *matrice de Fock* associée à D . Les équations d'Euler-Lagrange du problème (6.8) s'écrivent donc

$$\begin{cases} F(D)C = SCA \\ C^T SC = I_N \\ D = CC^T. \end{cases} \quad (6.13)$$

La matrice $A \in \mathcal{M}_S(N)$ joue le rôle du multiplicateur de Lagrange de la contrainte $C^T SC = I_N$.

Remarquons maintenant que si C est un point critique de (6.8) et U une matrice orthogonale de $\mathcal{M}(N)$, alors CU est aussi un point critique de (6.8). On peut donc utiliser cette propriété pour diagonaliser la matrice A dans (6.13) et en déduire que l'ensemble des points critiques de (6.8) peut être engendré en résolvant le système

$$\begin{cases} F(D)C = SCE \\ C^T SC = I_N \\ D = CC^T \end{cases} \quad (6.14)$$

où $E = \text{Diag}(\epsilon_1, \dots, \epsilon_N)$ désigne une matrice diagonale. En notant Φ_k la k -ième colonne de C on voit que

$$F(D)\Phi_k = \epsilon_k S\Phi_k, \quad (6.15)$$

et que (6.14) est donc un problème aux valeurs propres (généralisé) non linéaire. Les équations (6.14) sont parfois appelées *équations de Hartree-Fock-Roothaan-Hall*. Notons bien que résoudre (6.14) fournit seulement *certaines* solutions de (6.8) (ainsi que les minima locaux non globaux, des points selles et des maxima locaux), mais que toutes les solutions de (6.8) peuvent être obtenues à partir des solutions de (6.14) en faisant agir sur ces solutions le groupe des matrices orthogonales. Un autre point de vue consiste à dire que parmi les matrices D solutions de (6.14), il y a *toutes* les solutions de (6.10).

6.2.4 Principe *Aufbau*

Nous avons vu dans le chapitre précédent que pour tout minimiseur du problème de Hartree-Fock non discrétisé (5.7), les valeurs propres ϵ_i intervenant dans les équations de Hartree-Fock (5.30) étaient en fait les N plus petites valeurs propres de l'opérateur de Fock. Cette propriété survit à l'approximation LCAO : si C est un minimiseur de (6.8) vérifiant la condition (6.14), les coefficients $(\epsilon_1, \epsilon_2, \dots, \epsilon_N)$ de la matrice diagonale E sont les N plus petites valeurs propres du problème aux valeurs propres généralisé (6.15). On dit alors que les matrices C et D correspondantes vérifient le principe *Aufbau*. Comme nous le verrons section suivante, cette propriété est largement exploitée dans la construction d'algorithmes SCF. Démontrons-là : soit $C = (\Phi_1|\Phi_2|\dots|\Phi_{N-1}|\Phi_N)$ un minimiseur de (6.8) vérifiant (6.14) et considérons $C' = (\Phi_1|\Phi_2|\dots|\Phi_{N-1}|\Phi_a)$ avec $a \in [N+1, N_b]$ l'état obtenu en remplaçant une orbitale de C par une autre fonction propre du problème (6.15). Une simple manipulation algébrique conduit à

$$E^{HF}(C'C'^T) - E^{HF}(CC^T) = \epsilon_a - \epsilon_N - \frac{1}{2} \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{|\phi_N(x)\phi_a(y) - \phi_N(y)\phi_a(x)|^2}{|x-y|} dx dy$$

où

$$\phi_N(x) = \sum_{\mu=1}^{N_b} [\Phi_N]_{\mu} \chi_{\mu}(x), \quad \phi_a(x) = \sum_{\mu=1}^{N_b} [\Phi_a]_{\mu} \chi_{\mu}(x).$$

Si C est un minimiseur de l'énergie de Hartree-Fock,

$$E^{HF}(C'C'^T) - E^{HF}(CC^T) \geq 0.$$

D'où

$$\epsilon_a - \epsilon_N \geq \frac{1}{2} \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{|\phi_N(x)\phi_a(y) - \phi_N(y)\phi_a(x)|^2}{|x-y|} dx dy > 0.$$

On en conclut

- que $(\epsilon_1, \dots, \epsilon_N)$ sont les N plus petites valeurs propres du problème spectral généralisé (6.15) ;
- qu'il existe un *gap* strictement positif entre la plus haute orbitale moléculaire occupée (*Highest Occupied Molecular Orbital*, HOMO) et la plus basse orbitale moléculaire virtuelle (*Lowest Unoccupied Molecular Orbital*, LUMO).

Remarquons que les équations (6.14) peuvent s'écrire en prenant comme inconnue principale la matrice D , sous la forme

$$[F(D), D] = 0 \quad (6.16)$$

où la notation $[\cdot, \cdot]$ désigne le “commutateur” défini par $[A, B] = ABS - SBA$.

Nous donnons pour finir une caractérisation utile des matrices D vérifiant les équations (6.14) complétées par le principe *Aufbau* ; cette caractérisation s'écrit (voir exercice 6.4)

$$D = \operatorname{arginf} \{ \operatorname{Tr}(F(D)D'), \quad D' \in \mathcal{P}_N \}. \quad (6.17)$$

où, comme précédemment,

$$\mathcal{P}_N = \{ D \in \mathcal{M}_S(N_b), \quad DSD = D, \quad \operatorname{Tr}(SD) = N \}.$$

6.2.5 Algorithmes SCF

Les conditions (6.16) et (6.17) incitent à définir deux critères de convergence adaptés à la résolution numérique des équations de Hartree-Fock :

- on dira que la suite $(D_n)_{n \in \mathbb{N}}$ converge numériquement vers une solution des équations de Hartree-Fock si
 - (1) $D_{n+1} - D_n \longrightarrow 0$
 - (2) $[F(D_n), D_n] \longrightarrow 0$;
- on dira que la suite $(D_n)_{n \in \mathbb{N}}$ converge numériquement vers une solution des équations de Hartree-Fock vérifiant le principe *Aufbau* si
 - (1) $D_{n+1} - D_n \longrightarrow 0$
 - (2') $\operatorname{Tr}(F(D_n)D_n) - \inf \{ \operatorname{Tr}(F(D)D), \quad D \in \mathcal{P}_N \} \longrightarrow 0$.

Algorithme de Roothaan et procédures similaires

L'algorithme “naturel” de résolution de (6.14) est l'algorithme de point fixe suivant : choisir C_0 vérifiant $C_0^T S C_0 = I_N$, poser $D_0 = C_0 C_0^T$ et construire par récurrence une suite $(C_n)_{n \in \mathbb{N}}$ vérifiant

$$\begin{cases} F(D_{n-1})C_n = S C_n E_n, & \text{avec} \quad E_n = \operatorname{Diag}(\epsilon_1^n, \dots, \epsilon_N^n) \\ C_n^T S C_n = I_N \\ D_n = C_n C_n^T \end{cases}$$

où $\epsilon_1^n \leq \epsilon_2^n \leq \dots \leq \epsilon_N^n$ sont les N plus petites valeurs propres de

$$F(D_{n-1})\Phi = \epsilon S\Phi$$

et où les colonnes de C_n sont des vecteurs propres S -orthogonaux associés. Cet algorithme est connu sous le nom d'algorithme de Roothaan [196]. Notons qu'il s'écrit aussi sous la forme compacte (voir exercice 6.4)

$$D_n = \operatorname{arginf} \{ \operatorname{Tr}(F(D_{n-1})D), \quad D \in \mathcal{P}_N \}.$$

Cet algorithme fonctionne correctement sur des systèmes chimiques simples (composés de trois ou quatre atomes de la première ligne du tableau périodique : Hydrogène, Carbone, Azote, Oxygène, ...) à condition d'utiliser des bases de petite taille. Toutes les simulations effectuées dans les premiers temps de la chimie computationnelle s'effectuaient dans ce cadre, et c'est la raison pour laquelle l'algorithme de Roothaan a été utilisé pendant une vingtaine d'années sans que cela pose trop de problèmes. Quand les ordinateurs ont permis de faire tourner des calculs plus importants, l'algorithme de Roothaan s'est révélé insuffisant : en général, il ne converge pas. Plusieurs algorithmes ont été proposés pour tenter de remédier à cette situation, qu'on peut classer en deux catégories :

1. les algorithmes consistant, comme celui de Roothaan, à résoudre les équations d'Euler-Lagrange (6.14) par une méthode de point fixe ; c'est la stratégie employée dans la quasi-totalité des codes antérieurs à ce jour ;
2. les algorithmes consistant à résoudre directement le problème d'optimisation sous contrainte (6.8).

Dans la première catégorie, mentionnons l'algorithme de *level-shifting* [201] (voir ci-dessous), l'algorithme de *damping* [233] et surtout l'algorithme DIIS [184], qui est encore à l'heure actuelle le plus utilisé. L'algorithme DIIS converge vite quand il converge mais ne converge pas toujours et converge dans de nombreux cas vers un minimum local non global. Dans la seconde catégorie, citons l'algorithme de *steepest descent* [166] (il s'agit d'une méthode de gradient projeté), l'algorithme de Bacskey [10] (réécriture du problème (6.8) sous la forme d'un problème d'optimisation sans contrainte et mise en œuvre de l'algorithme de Newton), ainsi que ses variantes de type quasi-Newton [69, 87], et l'algorithme de Shepard [209] (méthode de Hessian réduit pour résoudre (6.8)). Les algorithmes de la seconde catégorie convergent toujours vers un minimum local mais ce minimum local est rarement le minimum global pour les systèmes et/ou les bases de grande taille. Tous ces algorithmes sont décrits en détail dans [50], sections 29 et 30.

Au Chapitre 8, nous étudions en détail les propriétés de convergence des algorithmes de Roothaan et de *level shifting*, ce dernier étant défini par

$$\begin{cases} (F(D_{n-1}) - bD_{n-1})C_n = SC_nE_n & E_n = \operatorname{Diag}(\epsilon_1^n, \dots, \epsilon_N^n) \\ C_n^T SC_n = I_N \\ D_n = C_n C_n^T \end{cases}$$

où b est un paramètre strictement positif fixé, où $\epsilon_1^n \leq \epsilon_2^n \leq \dots \leq \epsilon_N^n$ sont les N plus petites valeurs propres de

$$(F(D_{n-1}) - bD_{n-1})\Phi = \epsilon S\Phi,$$

et où les colonnes de C_n sont des vecteurs propres S -orthogonaux associés. Remarquons que l'algorithme de *level shifting* coïncide avec l'algorithme de Roothaan dans le cas limite où $b = 0$.

Comme nous le verrons au Chapitre 8, les performances des algorithmes de Roothaan et de *level shifting* sont loin d'être satisfaisantes.

Algorithmes à contraintes relâchées

La méthode de relaxation des contraintes introduite dans [60, 53] permet de construire une nouvelle catégorie d'algorithmes de résolution du problème de Hartree-Fock. Elle est fondée sur l'observation suivante : le problème

$$\inf \left\{ E^{HF}(D), \quad D \in \tilde{\mathcal{P}}_N \right\} \quad (6.18)$$

avec

$$\tilde{\mathcal{P}}_N = \{D \in \mathcal{M}_S(N_b), \quad DSD \leq D, \quad \text{Tr}(SD) = N\} \quad (6.19)$$

obtenu à partir du problème (6.10) en remplaçant la contrainte $DSD = D$ par la contrainte relâchée $DSD \leq D$ a les mêmes minima locaux que le problème (6.10). Cette propriété, spécifique à l'énergie de Hartree-Fock, se prouve de la façon suivante : on commence par montrer (le faire en exercice) qu'un point critique D du problème

$$\inf \left\{ E^{HF}(D), \quad D \in \tilde{\mathcal{P}}_N \right\}. \quad (6.20)$$

vérifie des équations d'Euler-Lagrange de la forme

$$\begin{cases} F(D)C = SCE, & E = \text{Diag}(\epsilon_1, \dots, \epsilon_{N_b}) \\ C^T SC = I_{N_b} \\ D = C\Delta C^T, & \Delta = \text{Diag}(n_1, \dots, n_{N_b}) \\ n_i = 0 & \text{si } \epsilon_i > \epsilon_F \\ n_i = 1 & \text{si } \epsilon_i < \epsilon_F \\ 0 \leq n_i \leq 1 & \text{si } \epsilon_i = \epsilon_F \\ \sum_{i=1}^{N_b} n_i = N \end{cases}$$

où ϵ_F est un paramètre réel appelé *niveau de Fermi*, que l'on peut voir comme le multiplicateur de Lagrange de la contrainte $\text{Tr}(SD) = N$. D'un point de vue physique, n_i représente le *nombre d'occupation* de l'orbitale moléculaire Φ_i , Φ_i désignant le i -ème vecteur colonne de la matrice C (qui est ici de taille $N_b \times N_b$ et forme donc une base S -orthonormée de l'espace de discrétisation).

L'état électronique défini par la matrice D comprend donc potentiellement des orbitales totalement occupées ($n_i = 1$), des orbitales partiellement occupées ($0 < n_i < 1$), et des orbitales virtuelles ($n_i = 0$).

Supposons maintenant que le point critique D n'appartienne pas à \mathcal{P}_N , autrement dit qu'il ne vérifie pas les contraintes $DS D = D$. Cela se traduit par le fait qu'il existe au moins deux orbitales moléculaires $\Phi \in \mathbb{R}^{N_b}$ et $\Phi' \in \mathbb{R}^{N_b}$ de même énergie ϵ_F , partiellement occupées ; on note n et n' leurs nombres d'occupations (n et n' sont dans l'intervalle ouvert $]0, 1[$). Il est alors facile de voir que si l'on transfère δn électrons de Φ à Φ' , ce qui revient à remplacer D par la matrice

$$D' = D + \delta n \left(\Phi' \Phi'^T - \Phi \Phi^T \right),$$

qui appartient bien à $\tilde{\mathcal{P}}_N$ si $|\delta n|$ est suffisamment petit, on obtient la variation d'énergie

$$\Delta E = E(D') - E(D) = -\frac{\delta n^2}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|\phi(x) \phi'(y) - \phi(y) \phi'(x)|^2}{|x - y|} dx dy < 0,$$

où $\phi(x) = \sum_{\mu=1}^{N_b} \Phi_{\mu} \chi_{\mu}(x)$ et $\phi'(x) = \sum_{\mu=1}^{N_b} \Phi'_{\mu} \chi_{\mu}(x)$. Il en résulte donc que tout minimiseur local du problème (6.18) appartient en fait à \mathcal{P}_N et est donc un minimiseur local du problème de Hartree-Fock (6.10).

La supériorité de la formulation (6.18) sur la formulation (6.10) est que l'ensemble $\tilde{\mathcal{P}}_N$ sur lequel on minimise est convexe.

Les algorithmes cherchant à résoudre (6.18) plutôt que (6.8) ou (6.10) sont regroupés sous le vocable *Relaxed Constrained Algorithms* (RCA). L'*Optimal Damping Algorithm* (ODA), introduit dans [60], est l'algorithme RCA le plus simple ; il s'apparente à une méthode de gradient à pas optimal et s'écrit

$$\begin{cases} F(\tilde{D}_n) C_{n+1} = S C_{n+1} E_{n+1} \\ C_{n+1}^T S C_{n+1} = I_N \\ D_{n+1} = C_{n+1} C_{n+1}^T \\ \tilde{D}_{n+1} = \arg \inf \left\{ E^{HF}(\tilde{D}), \quad \tilde{D} \in \text{Seg}[\tilde{D}_n, D_{n+1}] \right\}. \end{cases} \quad (6.21)$$

où $\text{Seg}[\tilde{D}_n, D_{n+1}] = \left\{ (1 - \lambda) \tilde{D}_n + \lambda D_{n+1}, \lambda \in [0, 1] \right\}$. Comme E^{HF} comporte deux termes dont l'un est linéaire et l'autre quadratique en la matrice densité, la dernière ligne de (6.21) consiste simplement à minimiser un polynôme du second ordre en λ sur $[0, 1]$, ce qui peut bien sûr être fait analytiquement à moindre coût. L'algorithme est initialisé en choisissant un point de départ $D_0 \in \mathcal{P}_N$ et en posant $\tilde{D}_0 = D_0$.

L'algorithme ODA engendre donc deux suites de matrices :

- la suite principale $(D_n)_{n \in \mathbb{N}}$ dont les éléments sont des matrices densité admissibles (appartenant à \mathcal{P}_N) ;
- une suite secondaire $(\tilde{D}_n)_{n \geq 1}$ de pseudo-matrices densité dont les éléments sont dans $\tilde{\mathcal{P}}_N$ (de par la convexité de cet ensemble).

Notons que pour tout $D' \in \mathcal{P}_N$, et pour tout $\lambda \in [0, 1]$,

$$\begin{aligned} E^{HF}(\tilde{D}_n + \lambda(D' - \tilde{D}_n)) &= E^{HF}(\tilde{D}_n) + \lambda \text{Tr}(F(\tilde{D}_n) \cdot (D' - \tilde{D}_n)) \\ &\quad + \frac{\lambda^2}{2} \text{Tr}(G(D' - \tilde{D}_n) \cdot (D' - \tilde{D}_n)). \end{aligned}$$

La direction de plus forte pente correspond au minimiseur de

$$\inf \left\{ \text{Tr}(F(\tilde{D}_n)(D' - \tilde{D}_n)), D' \in \mathcal{P}_N \right\},$$

autrement dit de

$$\inf \left\{ \text{Tr}(F(\tilde{D}_n)D'), D' \in \mathcal{P}_N \right\}.$$

Ce minimiseur est précisément la matrice densité D_{n+1} définie par (6.21). C'est en ce sens que l'algorithme ODA peut s'interpréter comme un algorithme de gradient à pas optimal.

Le coût d'une itération ODA est sensiblement équivalent à celui d'une itération de l'algorithme de Roothaan, l'étape de recherche linéaire étant quasiment gratuite. L'analyse de convergence qui fait l'objet de l'exercice 8.3 fournit le résultat suivant :

Théorème 6.4 *Soit $D_0 \in \mathcal{P}_N$ un initial guess tel que la suite $(\tilde{D}_n)_{n \in \mathbb{N}}$ engendrée par l'algorithme ODA soit telle qu'il existe $\gamma > 0$ tel que*

$$\tilde{\epsilon}_{N+1}^n - \tilde{\epsilon}_N^n \geq \gamma \quad (6.22)$$

où $\tilde{\epsilon}_k^n$ désigne la k -ième valeur propre du problème spectral généralisé

$F(\tilde{D}_{n-1})\Phi = \epsilon S\Phi$. Alors

- la suite $\left(E^{HF}(\tilde{D}_n)\right)_{n \in \mathbb{N}}$ décroît vers une valeur stationnaire de l'énergie de Hartree-Fock ;
- la suite $(D_n)_{n \in \mathbb{N}}$ converge numériquement vers une solution des équations de Hartree-Fock vérifiant le principe Aufbau.

Bien que rien ne garantisse sur un plan théorique que le minimum obtenu soit un minimum global (en fait tout minimum local est un point fixe de l'algorithme), les tests numériques montrent que pour les choix usuels de D_0 (obtenus souvent par diagonalisation de l'Hamiltonien de cœur h ou par le calcul de l'état fondamental d'un modèle semi-empirique de type Hückel), l'algorithme ODA converge vers ce qui semble être le minimum global de l'énergie de Hartree-Fock. Le tableau ci-dessous reproduit des résultats numériques

représentatifs : le premier système considéré ($\text{CH}_3\text{-NH-CH=CH-NO}_2$) est une molécule organique standard, le second (Cr_2) et le troisième ($[\text{Fe}(\text{H}_2\text{O})_6]^{2+}$) contiennent des éléments métalliques (Chrome et Fer) pour lesquels les algorithmes SCF rencontrent généralement des difficultés (comme en témoignent les calculs effectués sur ces atomes, voir figure 8.2). La deuxième colonne fournit l'énergie totale du point de départ D_0 , les deuxième et troisième colonnes, celles obtenues à la convergence avec les algorithmes DIIS et ODA respectivement ; ces trois énergies sont en Hartree. La quatrième colonne reproduit la différence d'énergie entre le résultat obtenu avec DIIS (algorithme par défaut dans quasiment tous les codes antérieurs à 2002) et celui obtenu avec ODA en kcal/mol (une liaison covalente CH a une énergie de l'ordre de 100 kcal/mol). On voit que l'erreur commise par DIIS est loin d'être négligeable puisqu'elle est parfois comparable à l'énergie de liaison d'un ou plusieurs atomes.

Système	$E^{HF}(D_0)$	Min DIIS	Min ODA	ΔE (kcal/mol)
$\text{CH}_3\text{-NH-CH=CH-NO}_2$ base 6-31G	-374.0038 -322.2373	-375.3869 Ne CV pas	-375.3869 -375.3869	0 Non défini
Dichrome Cr_2 , base 6-31G	-2069.5400 -2051.4339	-2085.5449 -2085.4042	-2085.8060 -2085.8060	163.71 251.93
Ion $[\text{Fe}(\text{H}_2\text{O})_6]^{2+}$, 178 fonctions de base	-1700.7596 -1538.7283	-1717.8928 -1717.7355	-1718.0151 -1718.0151	76.68 175.31

Comparaison des algorithmes ODA et DIIS. Les énergies du point de départ D_0 (deuxième colonne), du minimum obtenu par l'algorithme DIIS (troisième colonne), et de celui obtenu par ODA (quatrième colonne) sont en unités atomiques. La différence d'énergie entre les deux minima est exprimée en kcal/mol (une liaison covalente : ~ 100 kcal/mol, une liaison hydrogène : ~ 5 kcal/mol).

Pour accélérer l'algorithme ODA, il est naturel de garder en mémoire la suite des itérés D_0, D_1, \dots, D_{n+1} (ou tout au moins une partie d'entre eux), et de remplacer la minimisation sur le segment $\text{Seg}[\tilde{D}_n, D_{n+1}]$ intervenant dans ODA par une minimisation sur le simplexe formé par les $(D_i)_{0 \leq i \leq n+1}$. La procédure ainsi obtenue constitue l'algorithme EDIIS pour *Energy Direct Inversion in the Iterative Subspace* :

$$\begin{cases} F(\tilde{D}_n)C_{n+1} = SC_{n+1}E_{n+1} \\ C_{n+1}^T SC_{n+1} = I_{N_p} \\ D_{n+1} = C_{n+1}C_{n+1}^T \\ \tilde{D}_{n+1} = \arg \inf \left\{ E^{HF}(\tilde{D}), \tilde{D} = \sum_{i=0}^{n+1} c_i D_i, 0 \leq c_i \leq 1, \sum_{i=0}^{n+1} c_i = 1 \right\}. \end{cases} \quad (6.23)$$

Cet algorithme s'inspire de l'algorithme DIIS dont il a été fait mention ci-dessus à ceci près que dans DIIS les coefficients $\{c_i^{opt}\}$ sont obtenus en minimisant le résidu

$$\left\| \sum_{i=0}^{n+1} c_i [F(D_i), D_i] \right\|^2$$

sous la contrainte $\sum_{i=0}^{n+1} c_i = 1$ (dans l'algorithme DIIS, l'extrapolation - i.e. les combinaisons linéaires non convexes des D_i - est théoriquement autorisée mais on constate en pratique que DIIS n'est efficace que lorsque les combinaisons linéaires sont convexes).

Les tests numériques montrent que l'algorithme EDIIS converge plus vite que ODA mais souvent moins vite que DIIS. En revanche, il arrive fréquemment que DIIS converge vers une mauvaise solution (un minimum local non global, voire un point selle) ou ne converge pas, alors que EDIIS converge inconditionnellement vers un minimum local. Notons que l'analyse numérique de l'algorithme DIIS (défaut de convergence dans certains cas, rapidité dans d'autres cas) reste à faire.

Quand on résout numériquement un problème de Hartree-Fock, l'étape limitante réside comme on l'a dit précédemment dans l'assemblage de la matrice de Fock à chaque itération. On peut choisir, ou bien de calculer une fois pour toutes les intégrales biélectroniques (6.12) et de les stocker (sur disque pour les systèmes moléculaires de grande taille), ou bien de ne pas les stocker et de les recalculer chaque fois qu'on en a besoin (on parle dans ce cas de *méthode directe*). Dès que la taille du système est assez importante (typiquement quelques dizaines ou quelques centaines d'atomes selon les machines), on choisit en général la deuxième solution qui s'avère plus économique en termes de temps de calcul : il est en effet plus rapide de calculer une intégrale biélectronique pour des gaussiennes-polynômes que d'effectuer une lecture sur disque. La diagonalisation de la matrice de Fock s'effectue en général par une méthode QR [79] mais des méthodes de Krylov sont implémentées dans certains codes. La complexité algorithmique d'un calcul Hartree-Fock est donc théoriquement en $N_I \times N_b^4$ (N_b désignant le nombre d'OA prises en compte dans le calcul, et N_I le nombre d'itérations SCF, de l'ordre de la dizaine pour les calculs courants). Notons cependant que pour les systèmes de grande taille, la complexité algorithmique de l'assemblage de la matrice de Fock n'est pas en N_b^4 mais plutôt en N_b^2 asymptotiquement : du fait du caractère localisé des OA

(une gaussienne-polynôme est numériquement à support compact) beaucoup d'intégrales biélectroniques $(\mu\nu|\kappa\lambda)$ peuvent être considérées comme nulles.

Remarque 6.5 *Un axe de recherche important consiste à rechercher des méthodes de résolution dont la complexité algorithmique soit en $O(N_b)$, d'une part en assemblant les matrices de Fock par des algorithmes rapides (Fast Multipole Methods [97, 204] par exemple), et d'autre part en utilisant une alternative à la diagonalisation pour calculer directement la matrice densité issue du principe Aufbau sans passer par le calcul des N orbitales moléculaires de plus basse énergie. Ces questions numériques sont évoquées à la section 11.1 (voir aussi la section 6.5).*

Signalons pour terminer cette section que le choix de l'état initial D_0 est une étape importante car elle conditionne la vitesse de convergence voire la convergence elle-même de l'algorithme. Il existe plusieurs techniques pour effectuer ce choix. En général, on utilise une fonction d'onde obtenue par une méthode semi-empirique, ou, au cours d'un processus d'optimisation de géométrie, le résultat final du calcul électronique pour les positions des noyaux de l'itération précédente. Les codes de chimie quantique donnent également la possibilité de démarrer l'algorithme avec $F(D_0) = h$, ce qui revient à prendre comme état initial un état dans lequel les électrons sont dispersés à l'infini.

Le Chapitre 8 consiste en une étude des algorithmes de Roothaan et de *level-shifting* pour la résolution des équations de Hartree-Fock. On y établit notamment que

1. génériquement, l'algorithme de Roothaan a le comportement suivant : ou bien il converge vers une valeur stationnaire de l'énergie ou bien il conduit à une oscillation entre deux états dont aucun n'est solution des équations de Hartree-Fock ;
2. pour un *shift* b assez grand, dépendant de la donnée initiale, l'algorithme de *level-shifting* converge vers une valeur stationnaire de l'énergie.

L'analyse numérique de l'algorithme d'*optimal damping* (ODA) fait l'objet du problème 8.3.

6.2.6 Extension aux modèles DFT

Le modèle de Kohn-Sham (1.17) ayant une parenté certaine (du moins formellement) avec le modèle de Hartree-Fock (6.5), les méthodes numériques utilisées pour résoudre ces deux problèmes sont *grosso modo* les mêmes (approximation de Galerkin par développement sur une base, cf. sections 6.2.1 et 6.2.2, puis mise en oeuvre d'un algorithme SCF, cf. section 6.2.5). Il faut toutefois souligner quelques différences :

- dans le cadre Kohn-Sham, on doit projeter la densité sur une grille pour calculer la composante d'échange-corrélation $E_{xc}(\rho)$ alors que dans le

cadre Hartree-Fock, l'assemblage de la matrice de Fock se fait sans grille, en tirant parti des propriétés des gaussiennes-polynômes ;

- on constate numériquement que l'algorithme de Roothaan se comporte sur le modèle de Kohn-Sham comme il se comporte sur le modèle de Hartree-Fock (convergence ou oscillation entre deux états), mais ceci n'est pas prouvé mathématiquement. De même, l'algorithme de *level shifting* s'avère converger pour des paramètres de *shift* assez grands, mais lentement et souvent vers des points critiques de l'énergie de Kohn-Sham qui ne sont pas des minima locaux. Enfin, la propriété de l'énergie de Hartree-Fock qui fait que tout minimum local de (6.18) est un minimum local de (6.10) n'est pas vérifiée pour l'énergie de Kohn-Sham (avec les choix usuels de fonctionnelles d'échange-corrélation approchées). En revanche, des raisons physiques qu'il serait trop long de décrire ici (voir par exemple [50], Section 15), justifient le fait de remplacer le modèle de Kohn-Sham par un modèle “à contraintes relâchées”, appelé *modèle de Kohn-Sham étendu*, dont les solutions peuvent être calculées par les algorithmes RCA. La complexité algorithmique d'un calcul Kohn-Sham est essentiellement la même que celle d'un calcul Hartree-Fock.

Signalons que parallèlement aux méthodes LCAO ou ondes planes qui sont de loin les plus employées, des méthodes de différences finies utilisant des techniques de type multigrilles ou maillages composites ont été récemment testées avec succès sur la résolution d'équations de Kohn-Sham LDA.

6.2.7 Compléments - Méthodes post Hartree-Fock

Pour calculer une partie de l'énergie de corrélation et améliorer ainsi le résultat obtenu par le modèle de Hartree-Fock, on peut utiliser plusieurs types de méthodes :

- les méthodes de perturbation ;
- les méthodes d'interactions de configurations ;
- les méthodes de multidéterminants.

6.2.7.1 Méthode de perturbation de Møller-Plesset

La méthode de perturbation de Møller-Plesset est une application directe mais originale de la méthode standard des perturbations, qui est un outil incontournable en mécanique quantique et que nous décrivons ici en quelques lignes.

Considérons un système quantique isolé décrit par l'hamiltonien H_0 et ψ_0 un état propre de H_0 d'énergie E_0 . Soumettons ce système à une perturbation extérieure modélisée par un hamiltonien d'interaction noté \mathcal{V} , le système perturbé étant alors décrit par l'hamiltonien $H = H_0 + \mathcal{V}$. On cherche un état propre ψ de H “voisin” de ψ_0 d'énergie E “voisine” de E_0 . Pour l'obtenir, on cherche une solution $(\psi(\lambda), E(\lambda))$ de l'équation

$$(H_0 + \lambda \mathcal{V})\psi(\lambda) = E(\lambda)\psi(\lambda), \quad (6.24)$$

pour $\lambda \in \mathbb{R}$, sous la forme $\psi(\lambda) = \sum_{n=0}^{+\infty} \lambda^n \psi_n$, $E(\lambda) = \sum_{n=0}^{+\infty} \lambda^n E_n$. En reportant ces expressions dans l'équation (6.24) et dans la condition de normalisation $\|\psi(\lambda)\| = 1$, on obtient pour tout $n \in \mathbb{N}^*$

$$(RS_n) \begin{cases} (H_0 - E_0) \cdot \psi_n = E_n \psi_0 + f_n \\ (\psi_0, \psi_n) = \alpha_n \end{cases}$$

où f_n et α_n ne dépendent que des $\{\psi_k\}_{0 \leq k \leq n-1}$ et des $\{E_k\}_{0 \leq k \leq n-1}$. Lorsque le système triangulaire (RS) , défini comme l'union des systèmes (RS_n) , a une solution et une seule, on obtient ainsi deux séries formelles en λ ,

$$\sum_{n=0}^{+\infty} \lambda^n \psi_n \quad \text{et} \quad \sum_{n=0}^{+\infty} \lambda^n E_n$$

appelées *séries de Rayleigh-Schrödinger*. Si ces deux séries ont, par exemple, un rayon de convergence strictement supérieur à 1, il est clair que $\psi = \psi(1) = \sum_{n=0}^{+\infty} \psi_n$ et $E = E(1) = \sum_{n=0}^{+\infty} E_n$ vérifient

$$H\psi = E\psi,$$

c'est-à-dire fournissent un état propre de l'hamiltonien perturbé ainsi que son énergie. En pratique, on ne calcule que les k premiers termes des séries de Rayleigh-Schrödinger ; on dit alors qu'on a mis en oeuvre la méthode des perturbation à l'ordre k .

Les fondements mathématiques de la théorie des perturbations des opérateurs linéaires se trouvent notamment dans l'ouvrage de référence de Kato [116].

La méthode de Møller-Plesset consiste à appliquer la technique ci-dessus à la recherche de l'énergie fondamentale de l'hamiltonien électronique H_e (cf. Chapitre 1 pour les notations) à partir de l'estimation obtenue par la méthode de Hartree-Fock. Plus précisément, on se place dans l'espace \mathcal{H}_e et on pose

$$H := H_e \quad \text{et} \quad H_0 := \sum_{i=1}^N \mathcal{F}(\mathcal{D})_{x_i},$$

où $\mathcal{F}(\mathcal{D})$ est l'opérateur de Fock associé au fondamental Hartree-Fock \mathcal{D} . La perturbation \mathcal{V} est donc ici définie par

$$\mathcal{V} = \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} - \sum_{i=1}^N \mathcal{G}(\mathcal{D})_{x_i}.$$

Soit maintenant ψ_0 le déterminant de Slater associé à l'opérateur densité \mathcal{D} . On a

$$H_0\psi_0 = E_0\psi_0$$

avec $E_0 = -\sum_{i=1}^N \epsilon_i$, les $-\epsilon_i$ désignant les N plus petites valeurs propres de l'opérateur de Fock $\mathcal{F}(\mathcal{D})$ en tenant compte des multiplicités (cf. section 6.2.5) : ψ_0 est donc le fondamental dans \mathcal{H}_e de l'opérateur H_0 . Tous les éléments sont réunis pour mettre en oeuvre la méthode des perturbations, qui, on peut l'espérer, fournira le fondamental de l'équation de Schrödinger électronique. En pratique, on ne calcule que la suite des énergies et on définit la méthode de Møller-Plesset à l'ordre n , notée MPn, comme étant la méthode de calcul de l'énergie

$$E^{MPn} = \sum_{k=0}^n E_k.$$

On retrouve à l'ordre 1 l'énergie de Hartree-Fock. En effet

$$E^{MP1} = E_0 + E_1 = E_0 + \langle \psi_0, \mathcal{V}\psi_0 \rangle = -\sum_{i=1}^N \epsilon_i + \frac{1}{2} \text{Tr}(\mathcal{G}(\mathcal{D})\mathcal{D}) = E^{HF}.$$

A l'ordre 2, on obtient

$$E^{MP2} = E_0 + E_1 + E_2 = E^{HF} - \langle \psi_0, \mathcal{V}(H_0 - E_0)^{-1} Q \mathcal{V}\psi_0 \rangle,$$

Q désignant le projecteur $I - |\psi_0\rangle\langle\psi_0|$.

Lors de la résolution du problème de Hartree-Fock sous l'approximation LCAO, on obtient en diagonalisant la matrice de Fock $F(D) \in \mathcal{M}(N_b, N_b)$ en le fondamental D , N spin-orbitales *occupées* (en ce sens que la solution est obtenue en calculant le déterminant de Slater de ces N spin-orbitales) et $N_b - N$ spin-orbitales *virtuelles*. De plus, si l'algorithme de résolution utilise le principe *aufbau* (cf. section 6.2.5), les N spin-orbitales occupées sont celles de plus basses énergies. Un déterminant de Slater construit à l'aide de k orbitales virtuelles et $N - k$ orbitales occupées est appelé une *configuration excitée* ou plus spécifiquement *k-excitation* du fondamental Hartree-Fock.

Le terme E_2 s'évalue facilement dans l'approximation LCAO par *sum over states*, car parmi les états propres de H_0 , qui sont les déterminants de Slater construits à partir des fonctions propres de $F(D)$, seuls ceux correspondant à une diexcitation du fondamental Hartree-Fock ont une contribution non nulle. On obtient plus précisément après calculs [106, 167, 195]

$$E_2 = \sum_{i,j=1}^N \sum_{a,b=1}^{N_b-N} \frac{|\langle \psi_{i,j \rightarrow a,b}, \mathcal{V}\psi_0 \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} = \sum_{i,j=1}^N \sum_{a,b=1}^{N_b-N} \frac{(ij||ab)}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b},$$

où on a noté $\psi_{i,j \rightarrow a,b}$ la diexcitation obtenue en remplaçant dans le fondamental Hartree-Fock les spin-orbitales occupées ϕ_i et ϕ_j par les spin-orbitales virtuelles ϕ_a et ϕ_b . La définition des intégrales biélectroniques $(ij||ab)$ dépend

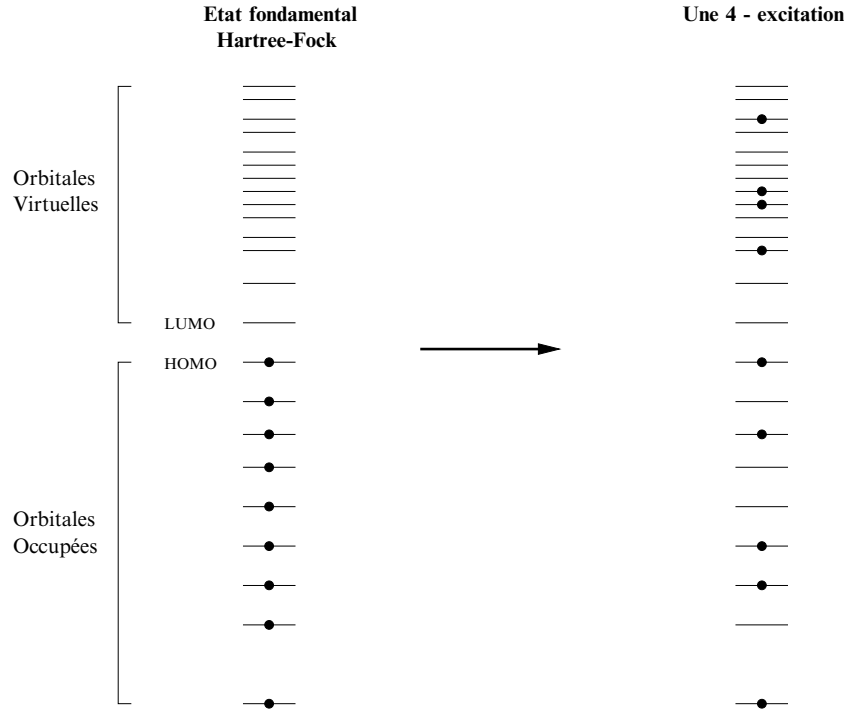


Fig. 6.2. Représentation schématique d'une 4-excitation d'un fondamental Hartree-Fock sans spin.

du modèle de Hartree-Fock considéré. Pour le modèle de Hartree-Fock sans spin, on a par exemple

$$(ij||ab) = (ij|ab) - (ib|aj) = \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{\phi_i(x)\phi_j(x)\phi_a(x')\phi_b(x')}{|x-x'|} dx dx' - \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{\phi_i(x)\phi_b(x)\phi_a(x')\phi_j(x')}{|x-x'|} dx dx'.$$

On peut écrire de même les formules de Møller-Plesset à un ordre quelconque, mais leur mise en oeuvre nécessite des calculs de plus en plus lourds. En pratique, on utilise surtout les approximations MP2, MP3 et MP4 (Møller-Plesset d'ordre 2, 3 et 4), dont la complexité algorithmique est respectivement $O(N_b^5)$, $O(N_b^6)$ et $O(N_b^7)$.

6.2.7.2 Méthode des interactions de configurations

Partant du fondamental du problème de Hartree-Fock sous l'approximation LCAO, on cherche une meilleure fonction d'onde sous la forme d'une combinaison linéaire du fondamental Hartree-Fock ψ_0 et des déterminants de Slater

$\psi_s, \psi_d, \psi_t, \psi_q, \dots$ obtenus par mono-, di-, tri-, quadri-, ... excitation de ψ_0 :

$$\psi_e = c_0\psi_0 + \sum_s c_s\psi_s + \sum_d c_d\psi_d + \sum_t c_t\psi_t + \sum_q c_q\psi_q + \dots$$

et en optimisant les coefficients c_* par méthode variationnelle. Quand on prend effectivement en compte toutes les configurations excitées (il y en a $\frac{N_b!}{N!(N_b-N)!}$), on parle de *full CI*. Le résultat numérique obtenu par *full CI* est le meilleur résultat (i.e. celui qui donne la plus basse énergie) qu'on puisse obtenir avec la base d'OA choisie. En pratique un calcul *full CI* est souvent trop lourd et on se limite à considérer les $N(N_b - N)$ mono-excitations (méthode *CIS*), ou les $N_b(N_b - N)$ mono- et les $\frac{N(N-1)(N_b-N)(N_b-N-1)}{4}$ di-excitations (méthode *CISD*), parfois les $N_b(N_b - N)$ mono-, les $\frac{N(N-1)(N_b-N)(N_b-N-1)}{4}$ di- et les $\frac{N(N-1)(N-2)(N_b-N)(N_b-N-1)(N_b-N-2)}{36}$ tri-excitations (méthode *CISDT*).

La méthode de *coupled cluster* constitue une variante de la méthode d'interaction de configurations permettant de déterminer les coefficients c_* par une technique alternative à la méthode variationnelle, basée sur le développement en série de l'exponentielle dans un formalisme de seconde quantification (voir [108] pour plus de détails). Contrairement à la méthode d'interaction de configurations, la méthode de *coupled cluster* ne fournit pas nécessairement une approximation par excès de l'énergie exacte, mais possède en revanche une propriété d'extensivité : l'énergie d'un système AB formé de deux sous-systèmes A et B *non interagissant* doit être égale à la somme des énergies des sous-systèmes A et B ; c'est le cas pour l'approximation *coupled cluster* mais pas pour l'énergie obtenue par une méthode interaction de configuration tronquée *CIS*, *CISD*, *CISDT*, ... (cette propriété n'est recouverte que pour l'énergie *full CI*).

Etant donné qu'un calcul Hartree-Fock (un seul déterminant) permet déjà d'obtenir entre 90 et 95% du résultat exact, on pourrait penser qu'on améliore fortement le résultat par interactions de configurations. Or pour des systèmes de grande taille (plusieurs centaines d'électrons), il faut souvent prendre en compte plusieurs millions de configurations excitées pour améliorer de 1% le résultat Hartree-Fock.

6.2.7.3 Méthode des multidéterminants

La méthode des multidéterminant, plus connue en chimie sous le nom de méthode MCSCF (*multiconfiguration self-consistent field*) [107, 208, 225], consiste à minimiser l'énergie électronique sur les fonctions d'onde ψ_e qui s'écrivent comme somme de d déterminants de Slater (d étant un entier donné). Autrement dit, la méthode à d déterminants s'écrit (dans un formalisme sans spin)

$$\inf \left\{ \langle \psi_e, H_e \psi_e \rangle, \quad \psi_e = \sum_{i=1}^d c_k \det(\phi_i^k(x_j)), \quad \Phi^k = \{\phi_i^k\} \in \mathcal{W}_N, \quad \|\psi_e\| = 1 \right\}. \quad (6.25)$$

On retrouve Hartree-Fock pour $d = 1$. Par ailleurs la méthode des multidéterminants “converge” vers la solution exacte du problème électronique (1.3) car toute fonction d’onde $\psi_e \in \mathcal{H}_e$ peut s’écrire sous la forme

$$\psi_e(x_1, \dots, x_N) = \sum_{k=1}^{+\infty} c_k \det(\phi_i^k(x_j)),$$

où pour tout $k \in \mathbb{N}^*$, $\Phi^k := \{\phi_i^k\} \in \mathcal{W}_N$.

La méthode des multidéterminants a été étudiée d’un point de vue mathématique par Lewin [143] (suite à des travaux de Le Bris [136] et Friesecke [90]).

Contrairement aux méthodes d’interactions de configurations, dans lesquelles on cherche d’abord les Φ^k par un calcul Hartree-Fock, puis les c_k par méthode variationnelle ou par méthode de *coupled cluster*, on optimise simultanément les Φ^k et les c_k dans un calcul MCSCF. La résolution numérique du problème (6.25) se fait généralement par boucles imbriquées, la boucle interne portant sur l’optimisation des Φ^k , la boucle externe sur celle des c_k , mais on peut aussi optimiser simultanément les Φ^k et les c_k [227, 209].

6.3 Optimisation de géométrie

Le problème de l’optimisation de géométrie (cf. section 1.2) est un problème de minimisation dans \mathbb{R}^{3M} qui s’écrit sous la forme générale

$$\inf \{ W(\bar{x}_1, \dots, \bar{x}_M), \quad (\bar{x}_1, \dots, \bar{x}_M) \in \mathbb{R}^{3M} \}, \quad (6.26)$$

où $W(\bar{x}_1, \dots, \bar{x}_M)$ désigne la somme de l’énergie de répulsion coulombienne entre les noyaux $\sum_{1 \leq k < l \leq M} \frac{z_k z_l}{|\bar{x}_k - \bar{x}_l|}$ et de l’énergie électronique fondamentale $U(\bar{x}_1, \dots, \bar{x}_M)$ correspondant aux positions $(\bar{x}_1, \dots, \bar{x}_M)$ des noyaux, définie par (1.2) et évaluée par l’une des méthodes d’approximation présentées précédemment.

D’un point de vue théorique, le problème de l’existence d’un minimum pour le problème (6.26) a été résolu pour quelques modèles simples d’énergie électronique. Le premier résultat concerne le modèle de Thomas-Fermi et est connu sous le nom de *no-binding theorem* [218, 154]. Il énonce qu’on ne peut pas lier les atomes entre eux, ni donc en particulier former des molécules stables, avec le modèle de Thomas-Fermi. Catto et Lions ont montré par la suite l’existence d’un minimum pour les systèmes neutres (i.e. la “stabilité” de la matière) pour les modèles de Thomas-Fermi-von Weizsäcker, de Thomas-Fermi-Dirac-von

Weiszäcker et de Hartree [61]. La question de l'existence d'un minimum pour les modèles de Hartree-Fock et de Kohn-Sham demeure toujours en suspens, même pour des systèmes très simples.

D'un point de vue numérique, le problème de l'optimisation de géométrie est difficile car d'une part l'espace sur lequel on minimise est de grande dimension (M peut atteindre plusieurs centaines pour les modèles *ab initio*, plusieurs centaines de milliers pour les modèles empiriques), et car d'autre part l'évaluation numérique de la fonction $W(\bar{x}_1, \dots, \bar{x}_M)$ en un point donné de \mathbb{R}^{3M} est déjà en elle-même difficile et coûteuse (du moins pour les modèles *ab initio* et semi-empiriques). Il faut donc faire appel le moins souvent possible à la procédure de calcul de la fonction W au cours de la résolution numérique. Dans le cas où la molécule présente des symétries, on résout en pratique le problème d'optimisation de géométrie en imposant des contraintes pour maintenir la symétrie. Cela permet d'une part de diminuer le nombre des paramètres par rapport auxquels on optimise, et d'autre part d'être assuré de ne pas converger vers un minimum local qui n'a pas la symétrie requise.

6.3.1 Rappels sur les méthodes standard de l'optimisation numérique

Cette section rassemble les idées de base de l'optimisation numérique sans contraintes. Elle a surtout pour but de mettre en évidence la raison d'être de la section 6.3.2 consacrée à la recherche d'expressions analytiques (donc économiques à évaluer numériquement) du gradient et du hessien de l'énergie potentielle W . Pour plus de précisions sur l'optimisation numérique, on recommande vivement au lecteur l'excellent ouvrage (en français!) de Bonnans, Gilbert, Lemaréchal et Sagastizábal [40], dont cette section s'inspire largement.

6.3.1.1 Principe général

On décide ici de noter

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

la fonction à minimiser et on suppose désormais que f est de classe C^2 . On pourra s'amuser à vérifier à titre d'exercice que l'énergie potentielle W est bien de classe C^2 sur l'ouvert

$$\Omega = \{(\bar{x}_1, \dots, \bar{x}_M) \in \mathbb{R}^{3M}, \quad \bar{x}_i \neq \bar{x}_j \text{ pour } i \neq j\}.$$

Toutes les méthodes (déterministes) d'optimisation sans contraintes fonctionnent sur le même principe : ce sont des méthodes itératives visant à construire une suite $(x_k)_{k \geq 0}$ dans \mathbb{R}^n qui converge vers un minimum de f , le passage de x_k à x_{k+1} se faisant en choisissant

1. une direction de descente $d_k \in \mathbb{R}^n$, c'est-à-dire une direction dans laquelle f décroît au moins localement ; une telle direction est caractérisée par $(d_k, g_k) \leq 0$ où l'on a noté $g_k = \nabla f(x_k)$ le gradient de f en x_k ,
2. une longueur de descente $t_k \in \mathbb{R}^+$,

et en posant

$$x_{k+1} = x_k + t_k d_k.$$

Les constructions de d_k et t_k résultent du choix d'un *modèle local* pour f . Typiquement, on considère que lorsqu'on est au point x_k , minimiser f revient à minimiser son développement de Taylor en x_k à l'ordre 1 (méthode de gradient) ou à l'ordre 2 (méthode de Newton).

Dans les méthodes exposées ci-dessous, on détermine *d'abord* la direction de descente d_k , *puis* la longueur de descente t_k en minimisant (de façon approximative) sur \mathbb{R}^+ la fonction

$$q(t) = f(x_k + t d_k).$$

Cette deuxième étape porte le nom de recherche linéaire. Notons cependant qu'il existe d'autres méthodes (par exemple celles dites des régions de confiance) pour lesquelles détermination de la direction et de la longueur de descente sont effectuées conjointement.

Les sections 6.3.1.2 à 6.3.1.5 présentent quatre méthodes plus ou moins performantes de recherche d'une direction de descente : la méthode de plus grande pente, la méthode du gradient conjugué non linéaire, la méthode de Newton, et la (famille) de méthode(s) de quasi-Newton. La section 6.3.1.6 aborde succinctement le problème de la recherche linéaire.

Remarque 6.6 *On obtient donc ainsi une suite $(x_k)_{k \in \mathbb{N}}$ dont on peut espérer qu'elle converge vers un minimum. Comme on ne dispose pas d'un temps infini pour effectuer la minimisation, il faut arrêter l'algorithme dès qu'on estime être suffisamment proche d'un minimum. Le critère d'arrêt généralement retenu porte sur la norme du gradient de la fonction à minimiser : on choisit $\epsilon > 0$ suffisamment petit et on arrête l'algorithme dès que*

$$|g_k| \leq \epsilon.$$

Cette condition ne garantit évidemment pas en toute généralité qu'on soit arrivé au voisinage d'un minimum local, mais elle fonctionne plutôt bien en pratique.

6.3.1.2 Méthode de plus forte pente

La première idée qu'on peut avoir pour déterminer une direction de descente est d'utiliser le modèle au premier ordre

$$f(x) \simeq \tilde{f}(x) = f(x_k) + (\nabla f(x_k), (x - x_k)).$$

La direction de plus forte pente, celle qui pour un pas $|x - x_k| = h$ donné ($|\cdot|$ désigne la norme euclidienne) induit la plus forte décroissance de \tilde{f} , est la direction opposée au gradient. Cela amène à poser

$$d_k = -\nabla f(x_k).$$

Cette méthode élémentaire fonctionne sur le papier : on prouve que si f est minorée, l'algorithme converge toujours en ce sens que pour tout $\epsilon > 0$, il existe k tel que $|g_k| \leq \epsilon$. Cela dit elle est peu efficace en pratique car si l'opposé du gradient est une bonne direction de descente (en fait la meilleure) pour un pas infinitésimal, ce n'est pas forcément le cas pour un pas donné.

La méthode de plus forte pente (*steepest descent*) est donc à proscrire car beaucoup trop lente.

6.3.1.3 Méthode du gradient conjugué non linéaire

La méthode de plus forte pente a un caractère "markovien" : le choix de la direction d_k se fait en oubliant totalement la façon dont on est arrivé en x_k . C'est pour cela que cet algorithme est si peu performant.

L'algorithme du gradient conjugué garde en un certain sens la mémoire des itérations précédentes. C'est à l'origine un algorithme de résolution des systèmes linéaires symétriques du type

$$Ax = b, \tag{6.27}$$

A désignant une matrice carrée symétrique définie positive. Ce problème s'écrivant également sous la forme

$$\inf_{x \in \mathbf{R}^n} f(x), \quad \text{avec} \quad f(x) = \frac{1}{2} x^T A x - b^T x.$$

le gradient conjugué est donc de façon équivalente un algorithme de minimisation d'une fonction quadratique fortement convexe. Son principe consiste à minimiser f à l'itération $k+1$, non pas comme dans la méthode de plus forte pente sur la demi-droite $x_k - tg_k$, mais sur le sous-espace affine contenant x_k et engendré par tous les $(g_l)_{0 \leq l \leq k}$

$$K_k = x_k + \text{Vect}(g_0, \dots, g_k).$$

La puissance de cet algorithme réside dans les deux propriétés suivantes :

1. Le minimum x_{k+1} de f sur le sous-espace K_k est donné par les formules de récurrence

$$\begin{aligned}
x_{k+1} &= x_k + t_k d_k \\
d_k &= -g_k + c_{k-1} d_{k-1} \\
c_{k-1} &= \frac{|g_k|^2}{|g_{k-1}|^2} \\
t_k &= -\frac{(g_k, d_k)}{(d_k, Ad_k)}.
\end{aligned} \tag{6.28}$$

Pour calculer x_{k+1} , il suffit donc de connaître le point x_k , le gradient g_k et la dernière direction de descente d_{k-1} , ce qui représente un coût de stockage très faible. Par ailleurs, le calcul de x_{k+1} ne requiert que l'évaluation de produits matrice-vecteur et de produits scalaires.

2. Les gradients g_k sont deux à deux orthogonaux et les directions de descente d_k sont A -conjuguées, c'est-à-dire telles que $(Ad_i, d_j) = 0$, pour tout couple $i \neq j$. Il en résulte que les vecteurs (g_0, \dots, g_k) forment une famille libre jusqu'à l'itération k pour laquelle $g_k = 0$ (condition de convergence de l'algorithme). Comme une famille libre dans un espace de dimension n ne peut pas contenir plus de n éléments, cela signifie que l'algorithme du gradient conjugué converge en au plus n itérations (fort heureusement, la convergence à ϵ près a lieu en beaucoup moins de n itérations pour un problème bien conditionné).

La généralisation à une fonctionnelle non quadratique de l'algorithme du gradient conjugué porte le nom de gradient conjugué non linéaire. En s'inspirant de (6.29) la direction de descente est prise de la forme

$$d_k = -g_k + c_{k-1} d_{k-1}. \tag{6.29}$$

Contrairement au cas quadratique où l'on pouvait déterminer analytiquement la valeur optimale du coefficient c_{k-1} , il faut ici se donner un critère de choix de ce coefficient. On peut par exemple choisir comme ci-dessus

$$c_{k-1} = \frac{|g_k|^2}{|g_{k-1}|^2}, \tag{6.30}$$

ce qui donne la méthode de Fletcher-Reeves. Un autre choix possible consiste à prendre

$$c_{k-1} = \frac{((g_k - g_{k-1}), g_k)}{|g_{k-1}|^2}. \tag{6.31}$$

C'est la méthode de Polak-Ribière. Notons que les formules (6.30) et (6.31) sont équivalentes dans le cas quadratique (car g_{k-1} et g_k sont alors orthogonaux), mais pas dans le cas général.

On dispose d'une preuve de convergence de l'algorithme de Fletcher-Reeves mais personne ne l'utilise plus car il est trop lent. La situation est inverse pour Polak-Ribière : on sait que cet algorithme ne converge pas à tous les coups, mais on l'utilise beaucoup car c'est une méthode performante la plupart du temps.

Remarque 6.7 *Pour une recherche linéaire exacte, une direction de la forme (6.29) est toujours une direction de descente puisque*

$$(d_k, g_k) = -|g_k|^2 + c_{k-1} (d_{k-1}, g_k) = -|g_k|^2.$$

En revanche, du fait que les recherches linéaires ne sont en pratique qu'approchées, il peut arriver que la direction calculée par la formule (6.29) et l'une des deux règles de Fletcher-Reeves ou de Polak-Ribière ne soit pas une direction de descente ($d_k \cdot g_k \geq 0$). Si cela se produit, on peut par exemple réinitialiser l'algorithme en posant $d_k = -g_k$.

L'algorithme du gradient conjugué non linéaire se met donc sous la forme :

1. Initialisation : choix de $x_0 \in \mathbb{R}^n$ et de $\epsilon > 0$; $k = 0$.
2. Test d'arrêt : calcul de $g_k = \nabla f(x_k)$, et arrêt de l'algorithme si $|g_k| \leq \epsilon$.
3. Calcul de la direction de descente : si $k = 0$, $d_0 = -g_0$, sinon $d_k = -g_k + c_{k-1}d_{k-1}$ avec c_{k-1} donné par (6.31). Si $(d_k, g_k) \geq 0$ (d_k est une direction de "remontée"), poser $d_k = -g_k$.
4. Recherche linéaire le long de d_k pour obtenir t_k .
5. Poser $x_{k+1} = x_k + t_k d_k$, remplacer k par $k + 1$ et aller en 2.

6.3.1.4 Méthode de Newton

La méthode de Newton consiste à considérer un modèle au second ordre de f :

$$f(x) \simeq \tilde{f}(x) = f(x_k) + (\nabla f(x_k), (x - x_k)) + \frac{1}{2} (x - x_k, f''(x_k) \cdot (x - x_k)).$$

Le minimum x_{min} de \tilde{f} est solution de

$$f''(x_k) \cdot (x_{min} - x_k) + \nabla f(x_k) = 0.$$

On prendra donc comme direction de descente la solution d_k de

$$f''(x_k) \cdot d_k + g_k = 0, \tag{6.32}$$

où comme ci-dessus $g_k = \nabla f(x_k)$.

Lorsque la fonction f est quadratique $f = \tilde{f}$, l'algorithme de Newton converge du premier coup si on prend une longueur de descente égale à 1 (on parle de pas de Newton). Dans le cas général, la convergence de la méthode de Newton avec recherche linéaire converge superlinéairement (i.e. pour tout $\epsilon > 0$, $|x_{k+1} - x_{min}| \leq \epsilon |x_k - x_{min}|$ lorsque k est assez grand) et la convergence devient quadratique si f est de classe C^3 au voisinage d'un minimum elliptique pour un pas de descente constant égal à un.

La méthode de Newton est donc très performante mais coûte cher : il faut à chaque pas calculer la matrice hessienne et résoudre le système linéaire (6.32).

6.3.1.5 Méthode de quasi-Newton

Le principe des méthodes de quasi-Newton est d'accélérer la méthode de Newton en court-circuitant les étapes les plus coûteuses du calcul, à savoir l'assemblage de la hessienne $f''(x_k)$ et/ou la résolution du système linéaire (6.32). Pour cela on remplace le système (6.32)

– par le système

$$H_k d_k + g_k = 0.$$

H_k désignant une approximation de $f''(x_k)$ facile à assembler (ce qui évite le calcul du hessien),

– ou mieux par le système

$$d_{k+1} = -B_k g_k$$

où B_k désigne une approximation de $f''(x_k)^{-1}$ facile à assembler (ce qui évite à la fois le calcul du hessien et la résolution du système linéaire).

La question qui se pose alors est de savoir comment construire à peu de frais une approximation B_k de $f''(x_k)^{-1}$. En pratique on démarre souvent l'algorithme avec $B_0 = I$ (mais on peut faire mieux en étudiant la spécificité du problème étudié) et on met à jour la matrice B à chaque itération de façon à satisfaire les deux critères suivants :

1. B_{k+1} est définie positive
2. B_{k+1} vérifie l'équation de Newton $B_{k+1}(g_{k+1} - g_k) = x_{k+1} - x_k$.

Ces deux conditions laissent une vaste gamme de choix. La solution généralement retenue consiste à poser

$$B_{k+1} = B_k - \frac{s_k y_k^T B_k + B_k y_k s_k^T}{(y_k, s_k)} + \left[1 + \frac{(y_k, B_k \cdot y_k)}{(y_k, s_k)} \right] \frac{s_k s_k^T}{(y_k, s_k)} \quad (6.33)$$

où $s_k = x_{k+1} - x_k$ et $y_k = g_{k+1} - g_k$. La méthode ainsi obtenue est désignée par l'acronyme BFGS (pour Broyden, Fletcher, Goldfarb, Shanno).

L'algorithme de quasi-Newton BFGS se résume donc ainsi :

1. Initialisation : choix de $x_0 \in \mathbb{R}^n$ et de $\epsilon > 0$; $k = 0$, $B_0 = I$ (par exemple).
Calcul de $g_0 = \nabla f(x_0)$
2. Test d'arrêt : arrêt de l'algorithme si $|g_k| \leq \epsilon$.
3. Calcul de la direction de descente : $d_k = -B_k g_k$
4. Recherche linéaire le long de d_k pour obtenir t_k .
5. Poser $x_{k+1} = x_k + t_k d_k$ et calculer $g_{k+1} = \nabla f(x_{k+1})$.
6. Mise à jour de la matrice B par la formule BFGS (6.33).
7. Remplacer k par $k + 1$ et aller en 2.

6.3.1.6 Un mot sur la recherche linéaire

Le principal message à retenir de cette section tient en ceci : comme aux itérations intermédiaires, la direction de descente n'est pas celle qui conduit au minimum, il ne sert à rien de mener la recherche linéaire jusqu'à son terme à chaque itération ; il faut s'arrêter dès qu'on estime avoir "suffisamment progressé". La règle d'arrêt considérée à l'heure actuelle comme la plus efficace est la règle de Wolfe ; elle consiste à choisir deux réels $0 < m_1 < m_2 < 1$ et à imposer l'arrêt dès que les deux conditions

$$q(t) \leq q(0) + m_1 t q'(0) \quad (6.34)$$

et

$$q'(t) \geq m_2 q'(0). \quad (6.35)$$

sont satisfaites. La première condition impose que le pas ne soit pas trop grand : si elle n'est pas satisfaite, cela veut dire en effet qu'on se trouve "sur l'autre versant de la cuvette", i.e. qu'on a dépassé le minimum. La deuxième condition exprime que la dérivée a suffisamment décru, ce qui signifie que le pas n'est pas trop petit (on n'est pas resté dans l'immédiat voisinage de x_k).

Le principe de la recherche linéaire est ensuite de chercher une longueur de descente t à l'intérieur d'un intervalle de confiance $]t_g, t_d[$ que l'on réduit au cours des itérations. A l'origine de la recherche on prend $t_g = 0$ et $t_d = +\infty$ et on se donne un pas initial $t > 0$. Trois cas peuvent se produire

1. t vérifie les conditions d'arrêt (i.e. les deux conditions (6.34) et (6.35) pour la règle de Wolfe), auquel cas on a terminé ;
2. t est trop grand (i.e. t ne vérifie pas (6.34)), auquel cas on pose $t_d = t$ et on cherche un nouveau $t \in]t_g, t_d[$ par interpolation,
3. t est trop petit (i.e. t ne vérifie pas (6.35)), auquel cas on pose $t_g = t$ et on cherche un nouveau $t \in]t_g, t_d[$ par interpolation si $t_d < +\infty$ et par extrapolation si $t_d = +\infty$.

Les techniques les plus simples consistent à prendre $t = (t_g + t_d)/2$ pour l'interpolation, $t_{\text{nouveau}} = a t_{\text{ancien}}$ avec $a > 1$ pour l'extrapolation. On utilise plutôt en pratique, pour l'interpolation comme pour l'extrapolation, un ajustement cubique (ou par un polynôme de degré 5 si les dérivées secondes de q sont accessibles) consistant à prendre pour nouveau t le minimum du polynôme interpolant les valeurs de q et de ses dérivées (ainsi que de ses dérivées secondes pour l'ajustement par un polynôme de degré 5) en les deux dernières valeurs de t . On renvoie à [40] pour les détails techniques.

Reste maintenant à choisir le pas initial. Dans la méthode de quasi-Newton, un choix naturel ($t = 1$) est fourni par l'algorithme. Pour la méthode du gradient conjugué non linéaire on peut utiliser l'initialisation de Fletcher $t = -2(f(x_{k-1}) - f(x_k))/(g_k \cdot d_k)$.

6.3.2 Dérivées analytiques

Comme on l'a vu dans la section précédente, les méthodes d'optimisation usuelles utilisent le gradient, et parfois la hessienne, de la fonction à optimiser.

Dans les cadres Hartree-Fock, post Hartree-Fock et Kohn-Sham, et sous l'approximation LCAO, on dispose d'expressions analytiques du gradient du potentiel W qui donnent lieu à un calcul très économique de ce vecteur. Considérons par exemple le cadre Hartree-Fock (sans spin) sous l'approximation LCAO et dérivons le potentiel

$$W = \text{Tr}(hD) + \frac{1}{2}\text{Tr}(G(D)D) + V_{nuc} = h : D + \frac{1}{2}D : A : D + V_{nuc}$$

par rapport à un paramètre λ qui peut être, comme c'est le cas dans la procédure d'optimisation de géométrie, une coordonnée nucléaire. Dans l'expression ci-dessus $V_{nuc} = \sum_{1 \leq k < l \leq M} \frac{z_k z_l}{|\bar{x}_k - \bar{x}_l|}$ désigne le potentiel de répulsion internu-

cléaire, h la matrice de l'hamiltonien de cœur, $A_{\mu\nu\kappa\lambda} = (\mu\nu|\kappa\lambda) - (\mu\lambda|\kappa\nu)$ le tenseur à quatre indices des intégrales biélectroniques et D la matrice densité du fondamental électronique dans la base d'OA choisie. On obtient

$$\frac{\partial W}{\partial \lambda} = \frac{\partial h}{\partial \lambda} : D + \frac{1}{2}D : \frac{\partial A}{\partial \lambda} : D + F(D) : \frac{\partial D}{\partial \lambda} + \frac{\partial V_{nuc}}{\partial \lambda}.$$

Mais on a par ailleurs en utilisant les équations de Hartree-Fock (6.14)

$$\begin{aligned} F(D) : \frac{\partial D}{\partial \lambda} &= \text{Tr} \left(F(D) \frac{\partial}{\partial \lambda} (CC^T) \right) \\ &= \text{Tr} \left(F(D) \frac{\partial C}{\partial \lambda} C^T \right) + \text{Tr} \left(F(D) C \frac{\partial C^T}{\partial \lambda} \right) \\ &= \text{Tr} \left(C^T F(D) \frac{\partial C}{\partial \lambda} \right) + \text{Tr} \left(F(D) C \frac{\partial C^T}{\partial \lambda} \right) \\ &= \text{Tr} \left(EC^T S \frac{\partial C}{\partial \lambda} \right) + \text{Tr} \left(SCE \frac{\partial C^T}{\partial \lambda} \right) \\ &= \text{Tr} \left(E(C^T S \frac{\partial C}{\partial \lambda} + \frac{\partial C^T}{\partial \lambda} SC) \right). \end{aligned}$$

En reportant dans cette expression l'égalité

$$C^T S \frac{\partial C}{\partial \lambda} + \frac{\partial C^T}{\partial \lambda} SC = -C^T \frac{\partial S}{\partial \lambda} C$$

obtenue en dérivant par rapport à λ la condition d'orthonormalité $C^T SC = I_N$, on parvient à l'expression

$$\frac{\partial W}{\partial \lambda} = \frac{\partial h}{\partial \lambda} : D + \frac{1}{2}D : \frac{\partial A}{\partial \lambda} : D - \text{Tr} \left(D_E \frac{\partial S}{\partial \lambda} \right) + \frac{\partial V_{nuc}}{\partial \lambda} \quad (6.36)$$

où $D_E = CEC^T = S^{-1}F(D)D$ désigne la matrice densité pondérée par les énergies (*energy weighted density matrix*). Les orbitales atomiques étant solidaires des noyaux, la base d'OA se déplace lors de l'optimisation de géométrie et les quantités $\frac{\partial h}{\partial \lambda}$, $\frac{\partial A}{\partial \lambda}$ et $\frac{\partial S}{\partial \lambda}$ ne sont donc pas nulles. Lorsque les OA sont des gaussiennes-polynômes, ces quantités sont faciles à calculer, la dérivée d'une gaussienne-polynôme étant encore une gaussienne-polynôme. Cela fournit un argument de plus en faveur de l'utilisation de bases gaussiennes.

Pour résoudre numériquement le problème de l'optimisation de géométrie, on peut donc mettre en oeuvre à moindre coût une méthode de gradient ou de quasi-Newton et assurer ainsi la convergence vers un minimum local.

Les algorithmes les plus utilisés actuellement (tout au moins pour les molécules de taille "raisonnable", quelques dizaines d'atomes), sont des méthodes de quasi-Newton : à la première itération, on calcule une estimation du hessien en utilisant un modèle de champ de forces empirique et on inverse cette matrice, opération qu'on peut se permettre en raison de la "petite" taille de celle-ci. Cette estimation de l'inverse du hessien est ensuite mise à jour par BFGS en utilisant des gradients analytiques (formule (6.36)).

6.3.3 Convergence vers un minimum global

Les méthodes déterministes décrites ci-dessus fonctionnent bien sur des molécules de petite taille (quelques atomes) à condition de partir d'une configuration "raisonnable" des noyaux, que les chimistes savent exhiber. En revanche, pour les molécules de grande taille, il existe un grand nombre de minima locaux, et il faut avoir recours à une méthode d'optimisation probabiliste (recuit simulé, algorithmes génétiques, *Gradient biased Monte-Carlo*) pour espérer la convergence vers un minimum global. De nombreuses références concernant l'application de ces diverses techniques à l'optimisation de géométrie figurent dans [172]. Il faut cependant préciser que les méthodes probabilistes demandent un grand nombre d'évaluations de la fonction à optimiser. Comme le calcul de la fonction W par une méthode *ab initio* est généralement un calcul lourd, les méthodes probabilistes ne sont généralement pas utilisées dans le cadre de la chimie quantique : leur utilisation est limitée au cadre des modèles empiriques dans lequel la fonction W a une expression analytique.

6.4 Résumé

Un problème de Hartree-Fock ou de Kohn-Sham s'exprime sous la forme d'un problème de minimisation sous contraintes dont les équations d'Euler-Lagrange ont la forme d'un problème aux valeurs propres non linéaire. Il s'agit plus précisément de minimiser la fonctionnelle d'énergie électronique correspondant au modèle sur l'ensemble

$$\mathcal{W}_N = \left\{ \Phi = \{\phi_i\}_{1 \leq i \leq N}, \quad \phi_i \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}, \quad 1 \leq i, j \leq N \right\}.$$

Pour approcher ce problème par un problème posé sur un espace de dimension finie, on utilise une approximation de Galerkin consistant généralement à développer les orbitales moléculaires ϕ_i sur une base d'orbitales atomiques gaussiennes.

Pour résoudre numériquement le problème de dimension finie ainsi obtenu (il s'agit toujours d'un problème de minimisation sous contraintes égalités d'une fonctionnelle non quadratique), on peut *a priori* opter pour l'un des deux choix suivants : minimiser directement la fonctionnelle d'énergie ou résoudre les équations d'Euler-Lagrange du problème, qui s'exprime sous la forme d'un système couplé de N équations aux dérivées partielles elliptiques non linéaires, non locales et posées sur \mathbb{R}^3 . C'est en général la deuxième solution qui est retenue, car les temps de calcul sont sensiblement plus courts. Les principaux algorithmes utilisés sont présentés section 6.2.5 dans le cadre Hartree-Fock, et section 6.2.6 dans le cadre Kohn-Sham.

Les problèmes de Hartree-Fock ou de Kohn-Sham apparaissent souvent dans la boucle interne d'un problème d'optimisation de géométrie, consistant à minimiser sur \mathbb{R}^{3M} la fonction d'énergie potentielle des noyaux, notée W . Pour minimiser W , on utilise des méthodes standard d'optimisation de type quasi-Newton, dont on rappelle le principe section 6.3.1. Ces méthodes sont efficaces car on dispose de formules analytiques des dérivées de l'énergie potentielle W par rapport aux coordonnées nucléaires, ce qui permet d'évaluer ces quantités en un temps raisonnable (section 6.3.2).

6.5 Pour en savoir plus

Plusieurs points évoqués brièvement dans ce chapitre seront repris plus en détail dans les chapitres suivants, notamment ceux qui concernent les bases gaussiennes (Chapitre 7), les bases d'ondes planes (section 9.1) et les algorithmes SCF (Chapitre 8).

Nous n'avons pas insisté, faute de place, sur les algorithmes rapides de résolution du problème électronique, ayant une complexité asymptotique en $O(N_b)$. Nous renvoyons le lecteur intéressé par ces aspects aux références :

- *Domain-based parallelism and problem decomposition methods in computational science and engineering*, D.E. Keyes, Y. Saad and D.G. Truhlar (eds.), SIAM 1995.
- M. Challacombe and E. Schwegler, *Linear scaling computation of the Fock matrix*, J. Chem. Phys. 106 (1997) 5526-5536.
- J. Millam and G. Scuseria, *Linear scaling conjugate gradient density matrix search as an alternative to diagonalization for first principles electronic structure calculations*, J. Chem. Phys. 106 (1997) 5569-5577.

- S. Goedecker, *Linear scaling electronic structure methods*, Rev. Mod. Phys. 71 (1999) 1085-1123.
- G. Scuseria, *Linear scaling density functional calculations with gaussian orbitals*, J. Phys. Chem. A 103 (1999) 4782-4790.

Concernant les questions relatives à l'optimisation de géométrie des systèmes moléculaires de grande taille, nous conseillons notamment

- T. Schlick, *Molecular Modeling : An Interdisciplinary Guide*, Springer-Verlag 2002.

6.6 Exercices

Exercice 6.1 Etats propres de l'atome d'Hydrogène.

On a vu à la section 6.1 qu'une base des états propres de l'atome d'Hydrogène pouvait être recherchée sous la forme

$$\psi(r, \theta, \phi) = R(r)Y_l^m(\theta, \phi).$$

Pour l et m fixés, la fonction R est solution de l'équation aux valeurs propres radiale

$$-\frac{1}{2r} \frac{d^2}{dr^2}(r R(r)) + \frac{l(l+1)}{2r^2} R(r) - \frac{1}{r} R(r) = E R(r)$$

qu'on cherche ici à résoudre. On suppose que l est fixé.

1. Etudier le comportement asymptotique de R à l'infini.
2. On pose maintenant $R(r) = w(r)e^{-\sqrt{2E}r}$. Ecrire l'équation satisfaite par la fonction w et étudier les singularités de w en l'origine.
3. On pose $w(r) = r^l L(r)$. Ecrire l'équation vérifiée par la fonction L . Rechercher les solutions de cette équation sous la forme d'un développement en série entière. Montrer que $\sqrt{2E}$ est nécessairement un entier et que les solutions correspondantes sont des polynômes.
4. Conclure.

Exercice 6.2 Les notations sont celles de la section 6.2.1. On pose

$$A = \{D = CC^T, \quad C \in \mathcal{M}(N_b, N), \quad C^T SC = I_N\}$$

et

$$B = \{D \in \mathcal{M}_S(N_b), \quad DSD = D, \quad \text{Tr}(SD) = N\}.$$

Vérifier que $A \subset B$ puis montrer que $B \subset A$ (on pourra commencer par le cas où $S = I_{N_b}$ et chercher un argument de nature géométrique).

Exercice 6.3 Intégrales électroniques pour des OA gaussiennes.

On considère quatre orbitales atomiques gaussiennes

$$\chi_1(x) = e^{-\alpha_1|x-\bar{x}_1|^2}, \quad \dots, \quad \chi_4(x) = e^{-\alpha_4|x-\bar{x}_4|^2},$$

centrées sur quatre noyaux *a priori* différents. On cherche une expression simple des intégrales de recouvrement

$$S_{ij} = \int_{\mathbf{R}^3} \chi_i \chi_j,$$

des intégrales monoélectroniques

$$h_{ij} = \frac{1}{2} \int_{\mathbf{R}^3} \nabla \chi_i \nabla \chi_j + \int_{\mathbf{R}^3} V \chi_i \chi_j$$

et des intégrales biélectroniques

$$(ij|kl) = \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{\chi_i(x) \chi_j(x) \chi_k(x') \chi_l(x')}{|x - x'|} dx dx'.$$

1. Etablir l'égalité

$$\forall x \in \mathbf{R}^3, \quad e^{-\alpha_i|x-\bar{x}_i|^2} e^{-\alpha_j|x-\bar{x}_j|^2} = C_{ij} e^{-\gamma_{ij}|x-\bar{y}_{ij}|^2}$$

$$\text{avec } C_{ij} = e^{-\frac{\alpha_i \alpha_j}{\alpha_i + \alpha_j} |\bar{x}_i - \bar{x}_j|^2}, \quad \gamma_{ij} = \alpha_i + \alpha_j \text{ et } \bar{y}_{ij} = \frac{\alpha_i \bar{x}_i + \alpha_j \bar{x}_j}{\alpha_i + \alpha_j}.$$

2. En déduire une expression analytique de S_{ij} . On rappelle que

$$\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}.$$

3. Vérifier que

$$\int_{\mathbf{R}^3} \nabla \chi_i \nabla \chi_j = -\Delta_{\bar{x}_i} S_{ij} = -\Delta_{\bar{x}_j} S_{ij}.$$

En déduire une expression analytique de l'intégrale d'énergie cinétique

$$K_{ij} = \frac{1}{2} \int_{\mathbf{R}^3} \nabla \chi_i \nabla \chi_j.$$

4. Ramener l'intégrale $(ij|kl)$ à une intégrale à deux centres de la forme

$$(ij|kl) = C \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} e^{-\gamma|x-\bar{y}|^2} \frac{1}{|x-x'|} e^{-\gamma'|x'-\bar{y}'|^2} dx dx'.$$

5. Montrer que pour tout $x \in \mathbf{R}^3$

$$e^{-\lambda|x|^2} = \frac{1}{(2\pi)^3} \left(\frac{\pi}{\lambda}\right)^{3/2} \int_{\mathbf{R}^3} e^{-k^2/4\lambda} e^{ik \cdot x} dk$$

et

$$\frac{1}{|x|} = \frac{1}{(2\pi)^3} \int_{\mathbf{R}^3} \frac{4\pi}{|k|^2} e^{ik \cdot x} dk.$$

6. En déduire que

$$(ij|kl) = \frac{2\pi^2 C_{ij} C_{kl}}{(\gamma_{ij} \gamma_{kl})^{3/2} |\bar{y}_{ij} - \bar{y}_{kl}|} \int_0^{+\infty} \frac{\sin u}{u} e^{-u^2/4w} du$$

$$\text{avec } w = \frac{\gamma_{ij} \gamma_{kl}}{\gamma_{ij} + \gamma_{kl}} |\bar{y}_{ij} - \bar{y}_{kl}|^2.$$

7. Montrer que

$$\int_0^{+\infty} \frac{\sin u}{u} e^{-u^2/4w} du = \sqrt{\pi w} \int_0^1 e^{-ws^2} ds.$$

8. Exprimer $(ij|kl)$ en fonction de

$$F(w) = \int_0^1 e^{-ws^2} ds.$$

9. En déduire l'expression en fonction de F des intégrales monoélectroniques h_{ij} .

Exercice 6.4 Soit H une matrice symétrique de taille $N_b \times N_b$ et $\epsilon_1 \leq \epsilon_2 \leq \dots \leq \epsilon_{N_b}$ ses valeurs propres. Soit $D \in \mathcal{M}_S(N_b)$ solution de

$$\begin{cases} HC = SCE, & \text{avec } E = \text{Diag}(\epsilon_1, \dots, \epsilon_N) \\ C^T SC \\ D = CC^T \end{cases} \quad (6.37)$$

où $C \in \mathcal{M}(N_b, N)$ et S est une matrice symétrique définie positive. On cherche à montrer que

$$D = \text{arginf} \{ \text{Tr}(HD'), \quad D' \in \mathcal{P}_N \} \quad (6.38)$$

avec

$$\mathcal{P}_N = \{ D' \in \mathcal{M}_S(N_b), \quad D' S D' = D', \quad \text{Tr}(S D') = N \}.$$

On note $\gamma = \epsilon_{N+1} - \epsilon_N$.

1. On suppose dans un premier temps que $H = \text{Diag}(\epsilon_1, \epsilon_2, \dots, \epsilon_N)$, que S est la matrice identité, et que $\gamma > 0$.

a) Donner l'expression de l'unique matrice D solution de (6.37).

b) Montrer que pour toute matrice $D' = [D'_{ij}] \in \mathcal{P}_N$, $0 \leq D'_{ii} \leq 1$ pour tout $1 \leq i \leq N_b$.

c) Montrer que la matrice D solution de (6.37) est telle que pour toute matrice $D' = [D'_{ij}] \in \mathcal{P}_N$,

$$\|D' - D\|^2 = 2 \sum_{i=N+1}^{N_b} D'_{ii},$$

où $\|\cdot\|$ désigne la norme matricielle de Fröbenius définie par $\|A\| = \text{Tr}(AA^T)^{1/2}$.

d) Montrer que pour toute matrice $D' = [D'_{ij}] \in \mathcal{P}_N$,

$$\mathrm{Tr}(HD') \geq \sum_{i=1}^N \epsilon_i + \sum_{i=1}^N (\epsilon_N - \epsilon_i) (1 - D'_{ii}) + \gamma \sum_{i=N+1}^{N_b} D'_{ii}.$$

2. En déduire que si $S = I_{N_b}$, on a pour toute matrice D solution de (6.37) l'inégalité

$$\forall D' \in \mathcal{P}_N, \quad \mathrm{Tr}(HD') \geq \mathrm{Tr}(HD) + \frac{\gamma}{2} \|D' - D\|^2, \quad (6.39)$$

d'où découle immédiatement (6.38)

3. Etendre l'argument au cas où S est une matrice symétrique définie positive quelconque.

Méthodes mathématiques en chimie quantique. Une introduction

Cancès, E.; Le Bris, C.; Maday, Y.

2006, XVI, 412 p. 34 ill., Softcover

ISBN: 978-3-540-30996-3