
Oscar Cosserat

Lemme du col et application en chimie quantique

Mémoire d'initiation à la recherche

Sous la direction de Monsieur **David Gontier**



Cycle Pluridisciplinaire d'Etudes Supérieures
Troisième année (L3)

Table des matières

1	Introduction	2
2	Lemme du col	3
2.1	La condition de Palais-Smale	3
2.2	Lemme de déformation	4
2.3	Lemme du col	6
3	Recherche de point col	7
3.1	Méthode de Newton	8
3.2	Algorithmes de recherches de point col	9
3.2.1	Algorithme du dimer	9
3.2.2	Algorithme élastique	10
4	Application en chimie quantique : le modèle Hartree-Fock	12
4.1	Introduction	12
4.2	Approximation de Born-Oppenheimer, principe de Pauli et précisions sur le modèle choisi	13
4.2.1	Approximation de Born-Openheimer	14
4.2.2	Principe d'exclusion de Pauli	14
4.2.3	Précision sur le modèle choisi	14
4.3	Modèle Hartree-Fock	15
4.4	Exemples	16
4.4.1	Phase stable de la molécule d'eau	17
4.4.2	Renversement de la molécule d'ammoniac	17
5	Conclusion et ouverture	20
6	Remerciements	21
7	Annexe : programmes Python et PySCF	22
7.1	Python	22
7.1.1	Calculs respectifs du gradient et de la hessienne de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ par différence finie	22
7.1.2	Algorithmes de descente : algorithme de descente de gradient à pas fixe et algorithme de Newton	22
7.1.3	Algorithme élastique	23
7.2	PySCF	24
7.2.1	Molécule d'eau H2O	24
7.2.2	Molécule d'ammoniac NH3	24

1 Introduction

L'ambition de ce mémoire est de démontrer une version du lemme du col dans \mathbb{R}^n , qui nous donne sous certaines conditions l'existence d'un point selle. La section suivante explique l'implémentation d'algorithmes tels que la méthode du Dimer et l'algorithme élastique permettant l'approximation numérique de ce point selle. Enfin, la dernière section applique ce qui précède en chimie quantique par le modèle Hartree-Fock : le lemme du col permet alors de calculer l'énergie de transition d'une molécule entre deux phases stables. On étudie ici la transition de phase de la molécule d'ammoniac NH₃.

Le principe du lemme du col dans \mathbb{R}^n est simple. Un cas typique de fonction pour lequel il s'applique est lorsque deux minima locaux x_a et x_b sont situés dans deux "vallées" distinctes, séparées par un "col". Le théorème permet alors d'affirmer l'existence d'un point critique, correspondant au point montant le plus haut d'un chemin montant le moins haut.

La preuve se fait en deux temps.

On montre d'abord un lemme de déformation de l'espace autour de x^* , qui s'applique en supposant que $f(x^*)$ n'est pas un point critique. Si on note η la déformation issue du lemme et que $f(x) < f(x^*) + \epsilon$, alors $f \circ \eta(x) < f(x^*) - \epsilon$.

La preuve du lemme du col est issue de la contradiction suivante : si γ est un chemin continu entre x_a et x_b , $\eta \circ \gamma$ est aussi un chemin continu entre x_a et x_b , ce qui contredit la minimalité de $f(x^*)$.

On expose ensuite un panel d'algorithmes permettant de trouver ce point col.

La méthode de Newton permet de trouver rapidement les minima locaux x_a et x_b de f . On évoque l'algorithme du Dimer, qui, à initialisation suffisamment proche, converge vers le point col recherché. Enfin, l'algorithme élastique délivre, pour x_a et x_b donnés, une suite minimisante de chemins entre x_a et x_b tels que le maximum du chemin converge vers $f(x^*)$.

Enfin, on donne une petite introduction au modèle de chimie quantique dit de "Hartree-Fock" pour étudier les fonctionnelles d'énergie de la molécule d'eau H₂O d'abord, de la molécule d'ammoniac NH₃ ensuite.

La molécule d'ammoniac présente l'intérêt particulier d'admettre deux phases stables, c'est-à-dire deux configurations différentes des noyaux pour lesquels la fonctionnelle admet un minimum local. On applique donc le lemme du col et les algorithmes présentés en deuxième section pour trouver le point selle entre les deux minima locaux. Ce point est digne d'intérêt en chimie parce qu'il nous donne l'énergie de transition nécessaire à la molécule pour passer d'une configuration à l'autre.

Bonne lecture !

2 Lemme du col

2.1 La condition de Palais-Smale

L'espace de départ de la fonctionnelle dont on cherche un point col est un espace de Hilbert, que l'on note \mathcal{H} .

Nous introduisons dans cette section la condition de Palais-Smale, une hypothèse importante de la preuve du lemme du col.

Définition 2.1 (Condition de Palais-Smale). Soient $f \in C^1(\mathcal{H}, \mathbb{R})$ et $c \in \mathbb{R}$. On dit que I satisfait la condition de Palais-Smale en c si $\forall (u_n) \in E^{\mathbb{N}}$,

$$\begin{cases} I(u_n) \rightarrow c \\ \nabla I(u_n) \rightarrow 0 \end{cases} \implies (u_n) \text{ converge à extraction près.}$$

Remarque 2.2. Par continuité, la limite $u \in \mathcal{H}$ de la suite extraite vérifie $f(u) = c$ et $\nabla f(u) = 0$.

Exemples 2.3. Dans \mathbb{R} , $t \mapsto t^2$ satisfait la condition de Palais-Smale en 0 tandis que $t \mapsto e^{-t}$ ne la satisfait pas, car la suite $u_n = n$ vérifie $e^{-n} \rightarrow 0$ et $u_n \rightarrow \infty$.

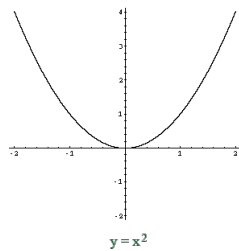


FIGURE 1 – $x \mapsto x^2$

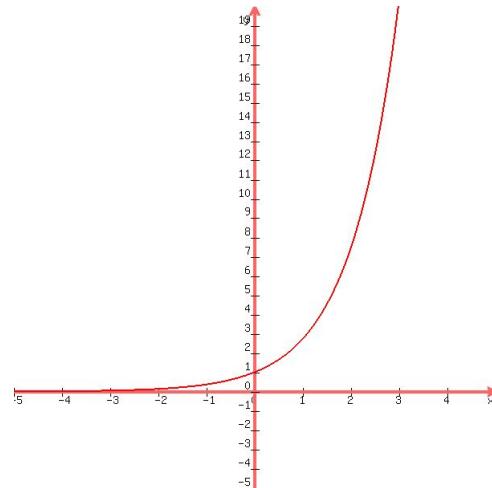


FIGURE 2 – $x \mapsto e^x$

Pour se fixer les idées, on met au clair les deux notions centrales de cette première section :

Définition 2.4 (Point critique). Dans la suite de l'article, on dira que $c \in \mathbb{R}$ est un point critique pour $f \in C^1(E, \mathbb{R})$ s'il existe $u \in \mathcal{H}$ tel que :

$$\begin{cases} f(u) = c, \\ \nabla f(u) = 0. \end{cases}$$

Définition 2.5 (Point col). *Un point col est un point critique qui n'est pas un extremum local.*

2.2 Lemme de déformation

La preuve du lemme du col repose en grande partie sur un lemme de déformation de l'espace de départ de la fonction dont on cherche un point col.

Lemme 2.6 (Lemme de déformation). *Soient \mathcal{H} un espace de Hilbert et $f \in \mathcal{C}^1(\mathcal{H}, \mathbb{R})$ satisfaisant la condition de Palais-Smale en $c \in \mathbb{R}$ et telle que $\{u \in \mathcal{H} \mid f(u) = c \text{ et } \nabla f(u) = 0\} = \emptyset$. On note*

$$\mathcal{A}_\xi := \{x \in \mathcal{H} \mid f(x) < \xi\}.$$

Alors, $\forall \bar{\varepsilon} > 0, \exists 0 < \varepsilon < \bar{\varepsilon}$ et $\eta \in \mathcal{C}(\mathcal{H}, \mathcal{H})$ tel que :

(i) pour tout $u \in \mathcal{A}_{c-\bar{\varepsilon}} \cup \mathcal{A}_{c+\bar{\varepsilon}}^c$, $\eta(u) = u$.

(ii) $\eta(\mathcal{A}_{c+\varepsilon}) \subset \mathcal{A}_{c-\varepsilon}$

Remarque 2.7. *η envoie les courbes de niveau qui sont sous $\mathcal{A}_{c+\varepsilon}$ vers $\mathcal{A}_{c+\varepsilon}$.*

On remarque aussi que la condition $\varepsilon < \bar{\varepsilon}$ est nécessaire : $\varepsilon = \bar{\varepsilon}$ est incompatible avec la continuité de η .

Démonstration. Dans un premier temps, montrons qu'il existe $b > 0$ et $\hat{\varepsilon} > 0$ tel que

$$\forall u \in \mathcal{A}_{c+\hat{\varepsilon}} \setminus \mathcal{A}_{c-\hat{\varepsilon}}, \quad \|\nabla f(u)\| \geq b. \quad (2.1)$$

Par contraposée, on suppose : $\forall \hat{\varepsilon} > 0, b > 0, \exists u \in \mathcal{A}_{c+\hat{\varepsilon}} \setminus \mathcal{A}_{c-\hat{\varepsilon}}$ tels que $\|\nabla f(u)\| \leq b$. En faisant tendre $\hat{\varepsilon}$ et b vers 0, on obtient une suite (u_n) telle que :

$$\lim_n f(u_n) = c \text{ et } \lim_n \nabla f(u_n) = 0$$

Parce que f vérifie la condition de Palais-Smale en c , $\exists \phi$ extraction telle que $\lim_n u_{\phi(n)} \in \{u \in (H) \mid f(u) = c \text{ et } \nabla f(u) = 0\} = \emptyset$, ce qui est absurde par hypothèse.

Quitte à considérer f petit, on peut considérer $\hat{\varepsilon}$ assez petit : on le prend tel que

$$\hat{\varepsilon} < \min(\bar{\varepsilon}, \frac{b^2}{8}, \frac{1}{8}).$$

Soit $\varepsilon \in]0, \hat{\varepsilon}[$.

On introduit les espaces suivants :

$$A := \{u \in \mathcal{H}, f(u) < c - \hat{\varepsilon} \text{ ou } f(u) > c + \hat{\varepsilon}\}$$

et

$$B := \{u \in \mathcal{H}, c - \varepsilon < f(u) < c + \varepsilon\}.$$

On introduit aussi les fonctions suivantes (on rappelle que pour $\Omega \subset E$, $d(u, \Omega) := \inf_{\omega \in \Omega} \|u - \omega\|$) :

$$g : u \in \mathcal{H} \mapsto \frac{d(u, A)}{d(u, A) + d(u, B)},$$

$$h : x \in \mathbb{R}^+ \mapsto h(x) = \begin{cases} 1 & \text{si } x \in [0, 1] \\ \frac{1}{x} & \text{sinon.} \end{cases},$$

et

$$W : u \in \mathcal{H} \mapsto -\frac{g(u)h(\|\nabla f(u)\|)\nabla f(u)}{4}.$$

On considère l'équation différentielle suivante : Soit $\eta : [0, 1] \times \mathcal{H} \mapsto \mathcal{H}$ telle que

$$\frac{\partial}{\partial t}\eta = W(\eta) \text{ et } \eta(0, u) = u \quad \forall u \in \mathcal{H}. \quad (2.2)$$

W est localement lipschitzienne comme produit et composée de fonctions localement lipschitziennes. D'après le théorème de Picard-Lindelöf, il existe une unique solution maximale à (2.2).

Montrons que pour tout $e \in \mathcal{H}$, $\|W(e)\| \leq \frac{1}{4}$.

On a $0 \leq g(e) \leq 1$. En particulier, comme $A \cap B = \emptyset$:

$$g(e) = \begin{cases} 0 & \text{si } e \in A \\ 1 & \text{si } e \in B \end{cases}$$

Si $\|\nabla f(e)\| > 1$, $W(e) = -\frac{g(e)}{4}$ donc $\|W(e)\| \leq \frac{1}{4}$.

Sinon,

$$\|W(e)\| = \left\| \frac{g(e)\nabla I(e)}{4} \right\| \leq \left\| \frac{g(e)}{4} \right\| \leq \frac{1}{4}$$

D'après le principe de comparaison, la solution notée $\eta(.,.)$ de (2.2) est définie sur \mathbb{R} tout entier. Notre fonction $\eta(.,.)$ ainsi construite, démontrons que $\eta(1,.)$ répond aux propriétés numérotées plus haut.

1. Par (2.2) :

$$\forall u \in \mathcal{H}, \quad \eta(0, u) = u.$$

Soit u tel que $f(u) \notin [c - \bar{\varepsilon}, c + \bar{\varepsilon}]$. Alors pour tout $t \in [0, 1]$, $g(\eta(t, u)) = 0$ car $\hat{\varepsilon} < \bar{\varepsilon}$, donc $W(\eta(0, u)) = 0$. D'où, $\eta(t, u) = \eta(0, u) = u$.

2. On montre d'abord que $t \in [0, 1] \mapsto f(\eta(t, u))$ est décroissante à u fixé.

$$\begin{aligned} \frac{\partial}{\partial t}(f(\eta(t, u))) &= \nabla f(\eta(t, u)) \cdot \frac{\partial}{\partial t}\eta(t, u) \\ &= \langle \nabla f(\eta(t, u)), W(\eta(t, u)) \rangle \\ &= -\frac{h(f(\eta(t, u)))g(f(\eta(t, u)))\|\nabla f(\eta(t, u))\|^2}{4} \\ &\leq 0 \end{aligned}$$

Donc $\forall t \in [0, 1]$, $f(\eta(t, u)) \leq f(\eta(0, u)) = f(u)$.

Si $u \in \mathcal{A}_{c-\varepsilon} \subset \mathcal{A}_{c+\varepsilon}$, par ce qui précède, $\eta(t, u) \in \mathcal{A}_{c-\varepsilon}$ et la propriété est trivialement vérifiée. Soit maintenant $u \in \mathcal{A}_{c+\varepsilon} \setminus \mathcal{A}_{c-\varepsilon}$.

On a $f(u) < c + \varepsilon$.

On a aussi que pour tout t , $\eta(t, u) \notin \mathcal{A}_{c-\varepsilon}$. En effet, $\mathcal{A}_{c-\varepsilon} \subset A$ donc $g \equiv 0$ sur $\mathcal{A}_{c-\varepsilon}$. Donc si $\exists t_0$ tel que $\eta(t_0, u) \in \mathcal{A}_{c+\varepsilon}$, $\eta(t_0, u) = u$ et $c - \hat{\varepsilon} < c - \varepsilon < f(u)$, ce qui est absurde. D'où

$$\forall t, f(\eta(0, u)) - f(\eta(t, u)) \leq c + \varepsilon - (c - \hat{\varepsilon}) < 2\hat{\varepsilon} \quad (2.3)$$

Par ailleurs, il existe t_1 tel que $\eta(t_1, u) \in \mathcal{A}_{c-\varepsilon}$.

En effet, par l'absurde, supposons $\forall t$, $\eta(t, u) \notin \mathcal{A}_{c-\varepsilon}$. Alors $f(\eta(t, u)) \geq c - \varepsilon$ et par ce qui précède, $c + \varepsilon \geq f(u) \geq f(\eta(t, u))$. Donc $\eta(t, u) \in B$ et $g(\eta(t, u)) = 1$.

Par (2.2), $\forall u \in \mathcal{H}, \forall t \in [0, 1]$,

$$\|\eta(t, u) - u\| = \|\eta(0, u) + \int_0^t \frac{\partial}{\partial s} \eta(s, u) ds - \eta(0, u)\| \leq t \sup_{e \in E} \|W(e)\| \leq \frac{1}{4}.$$

Si $\|\nabla f(\eta(t, u))\| \leq 1$, $\frac{\partial}{\partial t} f(\eta(t, u)) \leq -\frac{b^2}{4}$ par (2.1).

Sinon, $\|\nabla f(\eta(t, u))\| > 1$, $\frac{\partial}{\partial t} f(\eta(t, u)) \leq -\frac{1}{4}$.

Donc $\forall t \in [0, 1]$, $\frac{\partial}{\partial t} f(\eta(t, u)) \leq \max(-\frac{1}{4}, -\frac{b^2}{4}) = -\min(\frac{1}{4}, \frac{b^2}{4})$. Cela implique, par le théorème fondamental de l'analyse,

$$f(\eta(0, u)) - f(\eta(t, u)) \geq \min(\frac{1}{4}, \frac{b^2}{4}) \quad (2.4)$$

En combinant (2.3) et (2.4), on obtient

$$2\hat{\varepsilon} > f(\eta(0, u)) - f(\eta(t, u)) \geq \min(\frac{1}{4}, \frac{b^2}{4}),$$

ce qui contredit la définition de $\hat{\varepsilon}$.

On en déduit qu'il existe t_1 tel que $\eta(t_1, u) \in \mathcal{A}_{c-\varepsilon}$. Par décroissance de $t \mapsto f(\eta(t, u))$, $\eta(1, u) \in \mathcal{A}_{c-\varepsilon}$.

On a montré que $\eta(1, \cdot)$ satisfait (i) et (ii), ce qui clôt la preuve de ce lemme. \square

Dans la suite, par abus de notation, η désignera l'application continue de \mathcal{H} dans \mathcal{H} ainsi construite et notée $\eta(1, \cdot)$ dans la preuve ci-dessus.

2.3 Lemme du col

On dit qu'un ensemble S sépare x_a et x_b si pour tout chemin γ continue allant de x_a à x_b , γ rencontre S : il existe $t_\gamma \in [0, 1]$ tel que $\gamma(t_\gamma) \in S$. On note

- $\Gamma = \{\gamma \in \mathcal{C}([0, 1], \mathbb{R}^n) : \gamma(x_a) = 0, \gamma(x_b) = 1\}$ l'ensemble des chemins continus allant de x_a à x_b ,
- $c = \inf_{\gamma \in \Gamma} \max_{0 \leq t \leq 1} f(\gamma(t))$

Théorème 2.8 (Lemme du col). *Soient $f \in C^1(\mathbb{R}^n \rightarrow \mathbb{R})$ et $x_a, x_b \in \mathbb{R}^n$.*

On suppose que f satisfait la condition de Palais-Smale au point c .

S'il existe S séparant x_a et x_b tel que $\inf\{f(x), x \in S\} > \max(f(x_a), f(x_b))$, alors c est un point critique de f .

Démonstration. Montrons que c est bien défini.

Par hypothèse, pour tout chemin $\gamma \in \Gamma$, il existe t_γ tel que $\gamma(t_\gamma) \in S$. En particulier,

$$c = \inf_{\gamma \in \Gamma} \max_{0 \leq t \leq 1} f(\gamma(t)) \geq \inf_{\gamma \in \Gamma} f(\gamma(t_\gamma)) \geq \inf_{x \in S} f(x) > \max(f(x_a), f(x_b)) > -\infty$$

On pose $\bar{\varepsilon} = \frac{1}{2} [\inf_{x \in S} f(x) - \max(f(x_a), f(x_b))] > 0$. En notant $\mathcal{A}_\xi = \{u \in \mathcal{H}, f(u) < \xi\}$, on montre que $x_a, x_b \in \mathcal{A}_{c-\bar{\varepsilon}}$:

$$f(x_a) \leq \max(f(x_a), f(x_b)) < \max(f(x_a), f(x_b)) + \bar{\varepsilon} = \inf_{x \in S} f(x) - \bar{\varepsilon} \leq c - \bar{\varepsilon}.$$

Donc $x_a \in \mathcal{A}_{c-\bar{\varepsilon}}$. La preuve est similaire pour x_b .

En vue d'appliquer notre lemme de déformation, on suppose par l'absurde que c n'est pas un point critique : $\{u \in \mathcal{H} \mid f(u) = c \text{ et } \nabla f(u) = 0\} = \emptyset$.

f satisfaisant la condition de Palais-Smale en c , on applique donc (2.6) à $\bar{\varepsilon}$:

Il existe $\varepsilon > 0$ tel que $0 < \varepsilon < \bar{\varepsilon}$ et η un homéomorphisme de \mathcal{H} tel que :

- (i) pour tout $u \in \mathcal{A}_{c-\bar{\varepsilon}} \cup \mathcal{A}_{c+\bar{\varepsilon}}^c$, $\eta(u) = u$,
- (ii) $\eta(\mathcal{A}_{c+\varepsilon}) \subset \mathcal{A}_{c-\varepsilon}$.

Par l'optimalité de c , il existe $\gamma_\varepsilon \in \Gamma$ tel que $\max_{0 \leq t \leq 1} f(\gamma_\varepsilon(t)) < c + \varepsilon$. Autrement dit :

$$\forall t, \gamma_\varepsilon(t) \in \mathcal{A}_{c+\varepsilon}$$

Par (ii), $\forall t, \eta \circ \gamma_\varepsilon(t) \in \mathcal{A}_{c-\varepsilon}$.

Par ailleurs, $\eta \circ \gamma_\varepsilon \in \Gamma$. En effet,

- $\eta(\gamma_\varepsilon(0)) = \eta(x_a) = x_a$ car $x_a \in \mathcal{A}_{c-\bar{\varepsilon}}$, de même pour x_b ,
- $\eta \circ \gamma_\varepsilon$ continue par composition.

Cela contredit la définition de c , donc c est un point critique de f . □

3 Recherche de point col

Dans la section précédente, nous avons vu des conditions suffisantes pour l'existence d'un point col. Nous présentons dans cette section quelques algorithmes permettant de l'expliciter. Par la suite, par souci d'illustration, on applique ces algorithmes à la fonction test

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} : (x_1, x_2) \mapsto f(x_1, x_2) = \frac{x_1^4}{4} + \frac{x_1^2}{2} + \frac{x_2^2}{2} + x_1 x_2$$

dont on donne les courbes de niveau en figure 3.

La fonction f est bornée inférieurement et coercive, elle satisfait donc la condition de Palais-Smale sur \mathbb{R}^2 tout entier. De plus, elle admet 2 minima locaux bien distincts. C'est donc une candidate idéale pour appliquer le lemme du col.

On trouve d'abord les 2 minima locaux par la méthode de Newton.

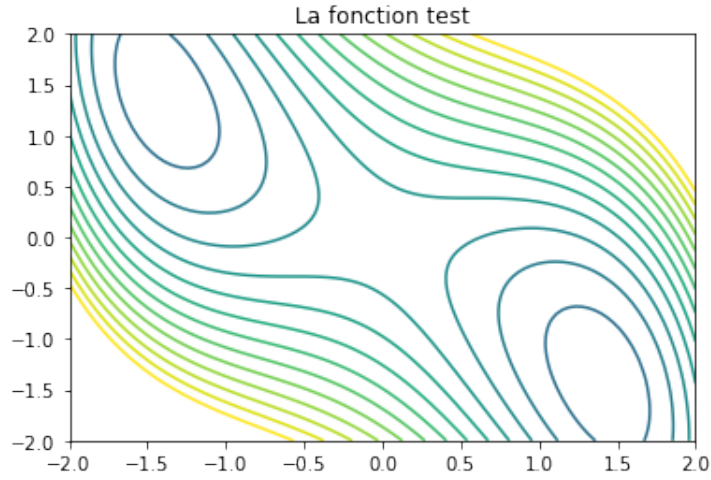


FIGURE 3 – Courbes de niveau de la fonction test.

3.1 Méthode de Newton

On rappelle que la Méthode de Newton est définie par la récurrence :

$$\begin{cases} x_0 \in \mathbb{R}^2, \\ x_{k+1} = x_k - H_f^{-1}(x_k) \nabla f(x_k), \end{cases}$$

où H_f est la dérivée seconde de f .

En l'appliquant à notre fonction test (figure 4), et selon les différentes conditions initiales, on trouve bien 2 minima locaux, vers lesquelles les itérés de la méthode de Newton convergent suivant le point d'initialisation x_0 de la suite.

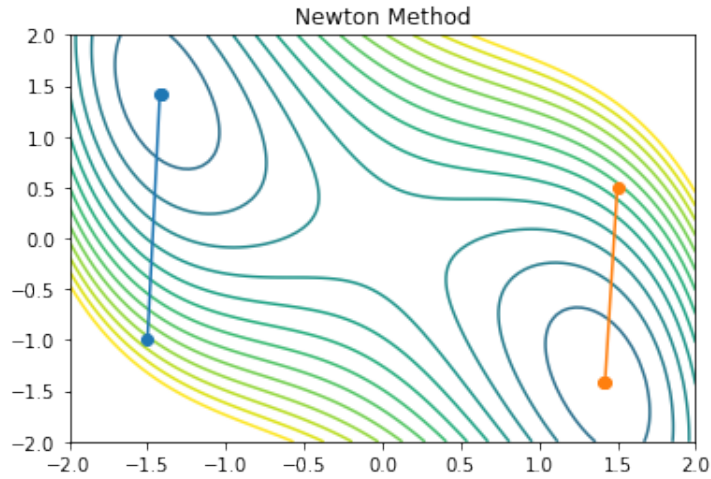


FIGURE 4 – Méthode de Newton sur la fonction test.

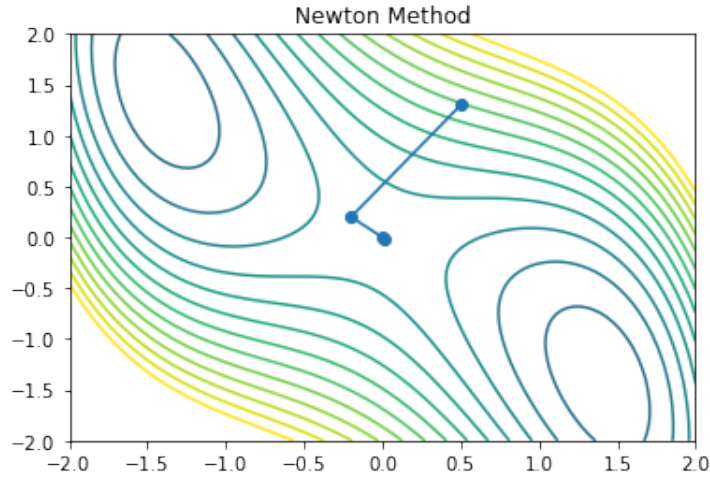


FIGURE 5 – Convergence de Newton vers un point col.

Remarque 3.1. *La méthode de Newton vise à annuler le gradient de la fonction sur laquelle on l'applique. Ainsi, si x_0 est suffisamment proche du point col recherché, les itérés convergent vers ce point col.*

Néanmoins, la convergence des itérés de Newton vers un point col est hasardeuse, et l'on dispose de meilleurs outils.

3.2 Algorithmes de recherches de point col

3.2.1 Algorithme du dimer

Nous présentons maintenant l'algorithme du dimer [7], qui est défini par la récurrence :

$$\begin{cases} x_0 \in \mathbb{R}^2, \\ x_{k+1} = x_k - \alpha(I_n - 2v_1 \otimes v_1)\nabla f(x_k), \end{cases}$$

où :

- à la k -ième itération, v_1 est le vecteur propre associé à la plus petite valeur propre λ_1 de $H_f(x_k)$,
- $v_1 \otimes v_1$ est la matrice de rang 1 dont le coefficient en (i, j) est le produit des i -ième et j -ième coefficients du vecteur v_1 ,
- $\alpha > 0$ est un pas de descente.

L'algorithme du dimer trouve un point col en minimisant f à la manière d'une descente de gradient à pas fixe dans $(n-1)$ directions.

En effet, notons $\lambda_1 \leq \dots \leq \lambda_n$ les valeurs propres ordonnées de H_f , et v_1, \dots, v_n les vecteurs propres associés, dans une base orthonormale. Cela implique $\sum_{i=1}^n v_i \otimes v_i = I_n$.

Par théorème spectrale, on sait qu'alors,

$$H_f = \sum_{i=1}^n \lambda_i v_i \otimes v_i$$

Au voisinage d'un minimum local non dégénéré, on a $0 < \lambda_1 \leq \dots \leq \lambda_n$. Les itérés $(x_k)_k$ de la descente de gradient à pas fixe convergent vers ce minimum local, avec :

$$\begin{cases} x_0 \in \mathbb{R}^2, \\ x_{k+1} = x_k - \alpha \nabla f(x_k) = x_k - \alpha \sum_{i=1}^n v_i \otimes v_i \nabla f(x_k). \end{cases}$$

En revanche, au voisinage d'un point col, on a $\lambda_1 < 0 \leq \lambda_2 \leq \dots \leq \lambda_n$. L'idée de la méthode du dimer est de "retourner" f dans la direction pour laquelle la valeur propre associée est négative. Ainsi, les valeurs propres de la nouvelle fonction \bar{f} sont positives et on peut lui appliquer la méthode de descente de gradient à pas fixe.

Pour retourner f dans la bonne direction, on pose $\bar{\lambda}_1 = -\lambda_1 = \lambda_1 - 2\lambda_1 > 0$ de sorte que :

$$H_{\bar{f}} = \sum_{i=1}^n \lambda_i v_i \otimes v_i - 2\lambda_1 v_1 \otimes v_1,$$

et le terme général des itérés du gradient conjugué appliqué à \bar{f} est bien

$$x_{k+1} = x_k - \alpha \left(\sum_{i=1}^n v_i \otimes v_i - 2v_1 \otimes v_1 \right) \nabla f(x_k) = x_k - \alpha (I_n - 2v_1 \otimes v_1) \nabla f(x_k).$$

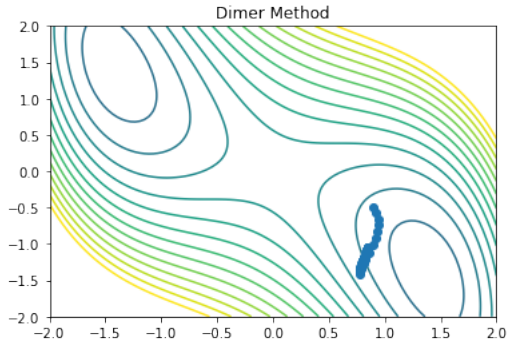


FIGURE 6 – Non-convergence de l'algorithme du Dimer.

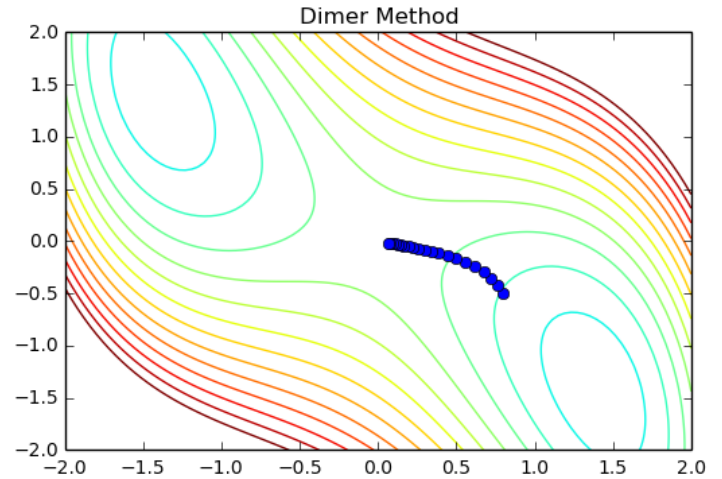


FIGURE 7 – Convergence de la méthode de Dimer vers un point col.

Remarque 3.2. Comparé à une méthode de gradient classique, la méthode du dimer est beaucoup plus sensible au choix de la condition initiale $x_0 \in \mathbb{R}^n$ (figures 6 et 7).

3.2.2 Algorithme élastique

En réponse à ce problème de contrainte sur l'initialisation de la méthode du dimer, nous présentons ici la méthode dite de la "corde élastique".

Celle-ci repose sur la minimisation par gradient à pas fixe d'une fonctionnelle de la forme

$$E : \mathbb{R}^{n \times m} \rightarrow \mathbb{R} : (z_1, \dots, z_m) \mapsto \underbrace{\rho \sum_i f(z_i)}_{(i)} + \underbrace{\sum_{i=2}^m \|z_i - z_{i-1}\|^2}_{(ii)} + \underbrace{\|x_a - z_1\|^2 + \|x_b - z_m\|^2}_{(iii)}$$

où :

- x_a, x_b sont des minima locaux de f
- z_1, \dots, z_m sont des points répartis entre x_a et x_b
- ρ est un réel strictement positif

On donne les itérés de la méthode :

$$\begin{cases} (z_i^0)_{1 \leq i \leq m} \in \mathbb{R}^2 \text{ répartis de manière uniforme entre } x_a \text{ et } x_b \\ (z_i^{k+1})_{1 \leq i \leq m} = (z_i^k)_{1 \leq i \leq m} - \alpha \nabla E(z_1, \dots, z_m) \end{cases}$$

On rappelle qu'on cherche c tel que

$$c = \inf_{\gamma \in \Gamma} \max_{0 \leq t \leq 1} f(\gamma(t)) \geq \inf_{\gamma \in \Gamma} f(\gamma(t_\gamma))$$

avec $\Gamma = \{\gamma \in \mathcal{C}([0, 1], \mathbb{R}^n) : \gamma(x_a) = 0, \gamma(x_b) = 1\}$.

La minimisation de la fonctionnelle correspond à la recherche d'un chemin linéaire par morceaux entre x_a et x_b tel que :

- (i) : le chemin ne monte pas trop haut
- (ii) : les points ne soient pas trop éloignés les uns des autres
- (iii) : les premiers et derniers points du chemin ne soient pas trop éloignés respectivement de x_a et x_b

$\rho > 0$ correspond à un poids que l'on choisit de mettre sur la contrainte (i). On dit que ρ mesure la tension de l'élastique :

- si ρ est trop petit, l'élastique est très tendu et la méthode renvoie un segment droit entre x_a et x_b , sans tenir compte des reliefs de f ,
- si ρ est trop grand, l'élastique est mou et les points $(z_i)_{1 \leq i \leq m}$ obtenus tombent de part et d'autre du point col de f .

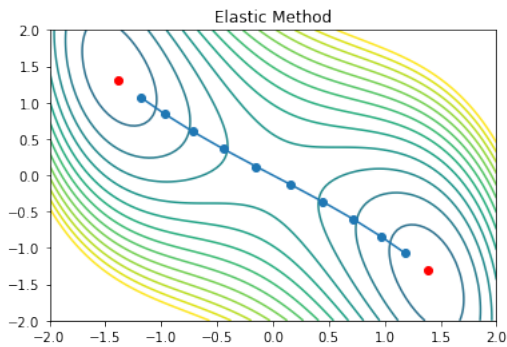


FIGURE 8 – L'élastique est très tendu

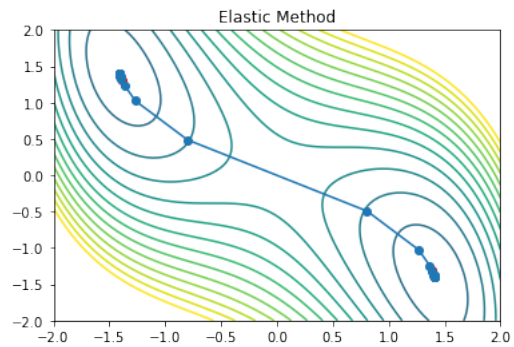


FIGURE 9 – L'élastique est mou

Enfin, pour un élastique correctement calibré, on obtient le résultat voulu.

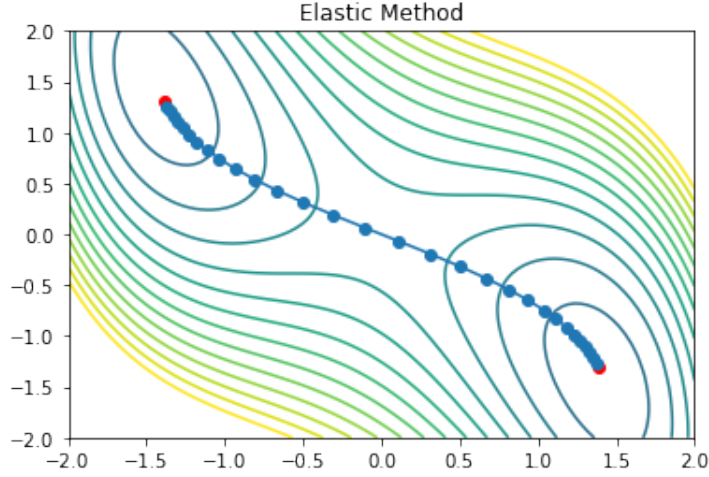


FIGURE 10 – Convergence de la méthode de l'élastique vers un chemin optimal

Le chemin obtenu $\gamma^* = (z_i)_{1 \leq i \leq m}$ approxime la limite d'une suite minimisante $(\gamma_p)_p \in \Gamma^N$ telle que

$$\max_{0 \leq t} f(\gamma_p(t)) \xrightarrow{p \rightarrow \infty} c = \inf_{\gamma \in \Gamma} \max_{0 \leq t} f(\gamma(t)).$$

L'approximation de c recherchée est donc $\max_{1 \leq i \leq m} f(z_i)$.

Nous rappelons qu'après avoir trouvé une approximation d'un point col avec cette méthode, il est encore possible d'améliorer sa position en appliquant les méthodes de Newton ou du Dimer décrites précédemment.

4 Application en chimie quantique : le modèle Hartree-Fock

Dans cette section, nous décrivons comment le modèle Hartree-Fock permet le calcul quantique de l'énergie d'une molécule. C'est le modèle qu'utilise le code avec lequel nous avons effectué nos simulations en fin de section.

4.1 Introduction

En physique classique, un système à N particules est décrit, pour un instant fixé, par les positions x_1, \dots, x_N et les moments p_1, \dots, p_N de chacune des particules. Toute chose égale par ailleurs, l'énergie du système $E : \mathbb{R}^{3N} \times \mathbb{R}^N \mapsto \mathbb{R}$ est alors :

$$E(x_1, \dots, x_N, p_1, \dots, p_N) = \sum_{i=1}^N U(p_i) + \sum_{i=1}^N V(x_i) + \sum_{i < j} W(x_i, x_j)$$

où :

- $p_i \mapsto U(p_i) = \frac{1}{2} m_i p_i^2$ est l'énergie cinétique de la particule i , avec m_i sa masse,
- $x_i \mapsto V(x_i)$ est l'énergie potentielle de la particule i ,
- $x_i, x_j \mapsto W(x_i, x_j)$ est l'énergie d'interaction entre les particules i et j .

Ici, les particules considérées sont les atomes et électrons d'une molécule. Ce modèle est déterministe et ne s'applique pas en chimie. En effet, à un instant fixé, on ne connaît ni la position d'un électron, ni son moment. D'où le modèle quantique : on considère alors les positions et moments comme des variables aléatoires.

Si $P_p((p_i)_{1 \leq i \leq N})$ est la probabilité que la particule i ait un moment p_i et $P_x((x_i)_{1 \leq i \leq N})$ la probabilité que la particule i ait une position x_i pour tout $1 \leq i \leq N$, l'énergie de la molécule est alors l'espérance d'une fonction des moments et positions des particules du système :

$$\begin{aligned} E(P_p, P_x) &= \mathbb{E} \left[\sum_{i=1}^N U(p_i) + \sum_{i=1}^N V(x_i) + \sum_{i < j} W(x_i, x_j) \right] \\ &= \int_{\mathbb{R}^{3N}} \sum_{i=1}^N U(p_i) dP_p(p_1, \dots, p_N) + \int_{\mathbb{R}^{3N}} \sum_{i=1}^N V(x_i) + \sum_{i < j} W(x_i, x_j) dP_x(x_1, \dots, x_N). \end{aligned}$$

En physique quantique, les probas P_x et P_p sont liées par la *fonction d'onde* :

Postulat 1. *Il existe une fonction $\Psi : \mathbb{R}^{3N} \mapsto \mathbb{C}$, appelée fonction d'onde du système, telle que :*

$$|\Psi(x_1, \dots, x_N)|^2 = P_x(x_1, \dots, x_N) \text{ et } |\widehat{\Psi}(p_1, \dots, p_N)|^2 = P_p(p_1, \dots, p_N)$$

où $\widehat{\Psi}$ est la transformée de Fourier normalisée de Ψ :

$$\widehat{\Psi}(p_1, \dots, p_N) = \frac{1}{(2\pi)^{\frac{N}{2}}} \int_{\mathbb{R}^{3N}} \Psi(x_1, \dots, x_N) \exp(x_1 p_1 + \dots + x_N p_N) d(x_1, \dots, x_N)$$

Le fait que P_x soit une densité de probabilité nous donne immédiatement $\|\Psi\|_2 = 1$, et par théorème de Plancherel on en déduit $\|\widehat{\Psi}\|_2 = 1$.

L'énergie du système est alors :

$$\begin{aligned} E(\Psi) &= \mathbb{E} \left[\sum_{i=1}^N U(p_i) + \sum_{i=1}^N V(x_i) + \sum_{i < j} W(x_i, x_j) \right] \\ &= \int_{\mathbb{R}^{3N}} \sum_{i=1}^N U(p_i) |\widehat{\Psi}(p_1, \dots, p_N)|^2 dp_1, \dots, p_N \\ &\quad + \int_{\mathbb{R}^{3N}} \left(\sum_{i=1}^N V(x_i) + \sum_{i < j} W(x_i, x_j) \right) |\Psi(x_1, \dots, x_N)|^2 dx_1, \dots, x_N \end{aligned}$$

4.2 Approximation de Born-Oppenheimer, principe de Pauli et précisions sur le modèle choisi

En théorie, l'énergie I d'une molécule à N particules se calcule de la sorte :

$$I = \inf \{ E(\Psi), \Psi \in L^2(\mathbb{R}^{3N}, \mathbb{C}) \text{ fonction d'onde} \}$$

Remarque 4.1. *C'est un problème de minimisation de dimension infinie complètement insoluble numériquement. Si on modélisait chaque droite réelle par 10 nœuds, ce qui est relativement peu, on obtiendrait un problème de minimisation à 10^{3N} degrés de liberté ! Par exemple, la molécule d'eau H_2O contient 10 électrons et donne au problème 10^{30} degrés de liberté.*

Voilà pourquoi en pratique, on développe des approximations simplifiant les calculs. Une première approximation est due à Born et Oppenheimer.

4.2.1 Approximation de Born-Oppenheimer

Définition 4.2 (Approximation de Born-Oppenheimer). *L'approximation de Born-Oppenheimer consiste à choisir d'observer les électrons de la molécule de façon quantique et les noyaux de la molécule de façon classique. Ainsi, notre fonctionnelle d'énergie $E(\Psi)$ prend aussi en paramètre la position $(R_j)_{1 \leq j \leq n}$ des n noyaux des atomes de la molécule :*

$$E(\Psi) = E(\Psi)((R_j)_{1 \leq j \leq n}).$$

Cette approximation se justifie par le rapport de masse entre un électron et un atome : $\frac{m_{\text{atome}}}{m_{\text{électron}}} \approx 1000$. Ainsi, du point de vue d'un électron, le noyau autour duquel il gravite est quasiment fixe. Dans la suite, on notera n le nombre de noyaux de la molécule, et R_j la position du j -ième noyau.

4.2.2 Principe d'exclusion de Pauli

On suppose les électrons indistinguables. En se rappelant que $|\Psi(x_1, \dots, x_N)|^2 = P_x(x_1, \dots, x_N)$ est la probabilité que l'électron i ait une position x_i et en notant \mathfrak{S}_N l'ensemble des permutations de $[1, N]$ et ϵ l'application signature de la permutation, on a donc :

$$\forall p \in \mathfrak{S}_N, P_x(x_1, \dots, x_N) = P_x(x_{p(1)}, \dots, x_{p(N)}).$$

Pour plus de clarté, on résume le principe d'exclusion de Pauli à l'énoncé suivant.

Théorème 4.3 (Principe d'exclusion de Pauli). *Si $\Psi : \mathbb{R}^{3N} \rightarrow \mathbb{C}$ est une fonction d'onde représentant des électrons et x_1, \dots, x_N la position des N électrons de la molécule,*

$$\forall p \in \mathfrak{S}_N, \Psi(x_{p(1)}, \dots, x_{p(N)}) = \epsilon(p)\Psi(x_1, \dots, x_N)$$

4.2.3 Précision sur le modèle choisi

Soit Ψ une fonction d'onde. On donne ici les formules utilisées pour le calcul des différentes énergies listées plus haut, avec N le nombre d'électron de la molécule, $(R_j)_{1 \leq j \leq n}$ la position des noyaux, $(z_j)_{1 \leq j \leq n}$ leur charge.

L'énergie cinétique est

$$\frac{1}{2} \int_{\mathbb{R}^{3N}} \sum_{i=1}^N p_i^2 \|\widehat{\Psi}\|^2(p_1, \dots, p_N) dp_1, \dots, p_N.$$

Par théorème de Plancherel,

$$p_i^2 \|\widehat{\Psi}\|^2 = \|p_i \widehat{\Psi}\|^2 = \|\widehat{\nabla_{x_i} \Psi}\|^2 = \|\nabla_{x_i} \Psi\|^2.$$

Finalement, l'énergie cinétique est

$$\frac{1}{2} \int_{\mathbb{R}^{3N}} \sum_{i=1}^N \|\nabla_{x_i} \Psi\|^2(x_1, \dots, x_N) dx_1, \dots, x_N.$$

L'énergie potentielle est

$$\sum_{i=1}^N \int_{\mathbb{R}^{3N}} V(x_i) \|\Psi\|^2(x_1, \dots, x_N) dx_1, \dots, x_N,$$

avec

$$V(x) = \sum_{j=1}^n \frac{-z_j}{\|x - R_j\|}.$$

C'est l'énergie générée par l'interaction entre les électrons et les noyaux.

L'énergie d'interaction entre les électrons est

$$\frac{1}{2} \int_{\mathbb{R}^{3N}} \sum_{1 \leq i < j \leq N} \frac{\|\Psi\|^2(x_1, \dots, x_N)}{\|x_i - x_j\|} dx_1, \dots, x_N.$$

4.3 Modèle Hartree-Fock

Définition 4.4 (Fonctionnelle Hartree-Fock). *Dans le modèle Hartree-Fock, on définit la fonctionnelle d'énergie du système E^{HF} par :*

$$E^{HF} = \inf\{E(\Psi), \Psi \in \mathcal{M}\}$$

où

$$\mathcal{M} := \left\{ \Psi : \mathbb{R}^{3N} \rightarrow \mathbb{C}, \exists (\phi_k)_{1 \leq k \leq N} \in (L^2(\mathbb{R}^3))^N, \langle \phi_i, \phi_j \rangle = \delta_{ij}, \right. \\ \left. \Psi(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \det((\phi_i(x_j))_{1 \leq i, j \leq N}) \right\}.$$

Les (ϕ_i) sont appelées orbitales du système. Elles décrivent entièrement la fonction d'onde.

On remarque que le caractère alterné du déterminant implique que $\Psi \in \mathcal{M}$ vérifie le principe d'exclusion de Pauli décrit plus haut.

On vérifie que $\forall \Psi \in \mathcal{M}$, $|\Psi|^2$ est une densité en revenant à la définition du déterminant : si on note \mathfrak{S}_N l'ensemble des permutations de $\llbracket 1, N \rrbracket$ et $\epsilon : \mathfrak{S} \rightarrow \{-1, 1\}$ la signature de la permutation,

$$\Psi(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \det((\phi_i(x_j))_{1 \leq i, j \leq N}) \\ = \frac{1}{\sqrt{N!}} \sum_{\sigma \in \mathfrak{S}_N} \epsilon(\sigma) \prod_{j=1}^N \phi_{\sigma(j)}(x_j)$$

D'où :

$$\|\Psi \sqrt{N!}\|_2^2 = \left\| \sum_{\sigma \in \mathfrak{S}_N} \epsilon(\sigma) \prod_{j=1}^N \phi_{\sigma(j)}(x_j) \right\|_2^2 \\ = \left\langle \sum_{\sigma \in \mathfrak{S}_N} \epsilon(\sigma) \prod_{j=1}^N \phi_{\sigma(j)}(x_j), \sum_{\sigma \in \mathfrak{S}_N} \epsilon(\sigma) \prod_{j=1}^N \phi_{\sigma(j)}(x_j) \right\rangle \\ = \sum_{\sigma \in \mathfrak{S}_N} \sum_{\hat{\sigma} \in \mathfrak{S}_N} \epsilon(\sigma) \epsilon(\hat{\sigma}) \underbrace{\left\langle \prod_{j=1}^N \phi_{\sigma(j)}(x_j), \prod_{j=1}^N \phi_{\hat{\sigma}(j)}(x_j) \right\rangle}_{(*)}$$

Soient $\sigma, \hat{\sigma} \in \mathfrak{S}_N$.

$$\begin{aligned}
 (*) &= \int_{\mathbb{R}^{3N}} \prod_{j=1}^N \overline{\phi_{\sigma(j)}(x_j)} \prod_{j=1}^N \phi_{\hat{\sigma}(j)}(x_j) dx_1, \dots, x_N = \int_{\mathbb{R}^{3N}} \prod_{j=1}^N \overline{\phi_{\sigma(j)}(x_j)} \phi_{\sigma(j)}(x_j) dx_1, \dots, x_N \\
 &= \prod_{j=1}^N \int_{\mathbb{R}^3} \overline{\phi_{\sigma(j)}(x_j)} \phi_{\sigma(j)}(x_j) dx_j = \prod_{j=1}^N \langle \phi_{\sigma(j)}, \phi_{\sigma(j)} \rangle = \delta_{\sigma\hat{\sigma}}
 \end{aligned}$$

Finalement,

$$\|\Psi \sqrt{N!}\|_2^2 = \sum_{\sigma \in \mathfrak{S}_N} \sum_{\hat{\sigma} \in \mathfrak{S}_N} \epsilon(\sigma) \epsilon(\hat{\sigma}) \delta_{\sigma\hat{\sigma}} = \sum_{\sigma \in \mathfrak{S}_N} \sum_{\hat{\sigma} \in \mathfrak{S}_N} \delta_{\sigma\hat{\sigma}} = N!$$

Ce qui donne :

$$\|\Psi\|_2^2 = 1$$

Ainsi, on a transformé un problème de minimisation sur un gros espace, $L^2(\mathbb{R}^{3N}, \mathbb{C})$ en un problème de minimisation sur un espace plus petit, l'espace dans lequel vit le vecteur des orbitales $(\phi_i)_{1 \leq i \leq N} : L^2(\mathbb{R}^3, \mathbb{C})^N$.

Pour reprendre la remarque (4.1), si l'on modélise chaque droite réelle par 10 noeuds et que l'on étudie une molécule à N électrons, on obtient un problème de minimisation à $N \times 10^3$ degrés de libertés. En d'autres termes, l'approximation Hartree-Fock rend possible le calcul quantique de l'énergie d'une molécule parce qu'elle transforme un problème exponentiel en problème polynomial.

4.4 Exemples

Dans les exemples suivants, on cherche les positions stables de la molécule, c'est-à-dire les positions des noyaux minimisant l'énergie du système en utilisant le modèle Hartree-Fock. Le code python, issu des bibliothèques pylab (calcul scientifique) et PySCF [10] (code de chimie quantique utilisé pour le calcul de l'énergie d'une molécule), est donné en annexe.

FIGURE 11 – La molécule d'eau H2O

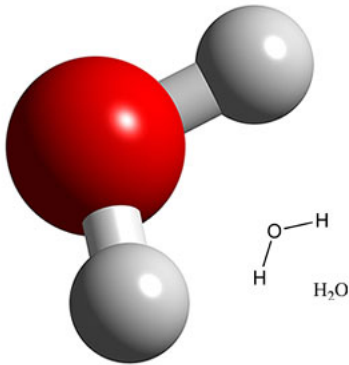
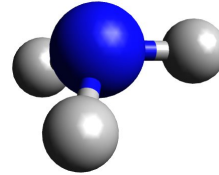


FIGURE 12 – La molécule d'ammoniac NH3



4.4.1 Phase stable de la molécule d'eau

La molécule d'eau est un système composé d'un atome d'oxygène au centre, de masse supérieure aux deux atomes d'hydrogène qui l'entourent. On s'attend donc à ce qu'il y ait un axe de symétrie passant par l'oxygène et au milieu des deux atomes d'hydrogène.

Dès lors, les paramètres que l'on fait varier ici sont la distance, en Ångström, des hydrogènes à l'oxygène et l'angle, en radian, entre les hydrogènes.

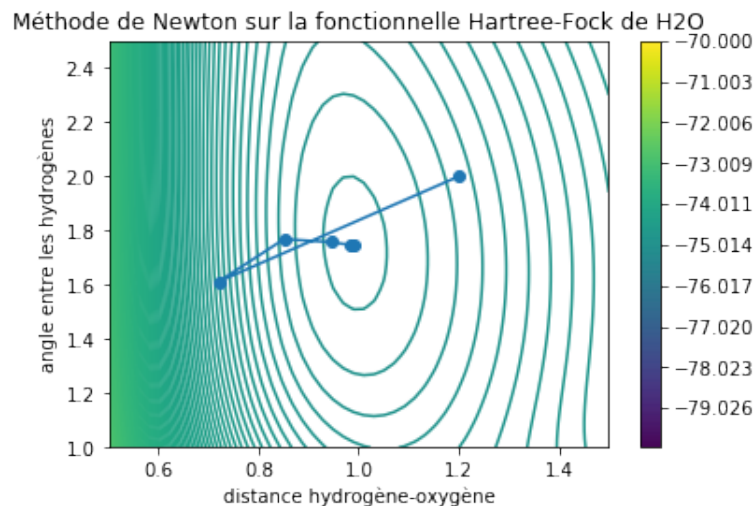


FIGURE 13 – Recherche de phases stables de H2O par méthode de Newton.

Si on note d la distance OH, θ l'angle entre les hydrogènes et I l'énergie du système, la simulation illustrée par la figure 13 donne :

- $d = 0.9893$ Ångströms
- $\theta = 1.7457$ radians

Wikipedia donne les valeurs :

- $d = 0.9584$ Ångströms
- $\theta = 1,8151$ radians

La précision de la simulation est donc de l'ordre du dixième de radian et du centième d'Ångströms.

4.4.2 Renversement de la molécule d'ammoniac

La molécule d'ammoniac est composée de trois hydrogènes autour de l'azote central. Là encore, on suppose que les hydrogènes sont tous sur le même plan, et à égale distance les uns des autres.

Soient donc $G = (a, 0, 0)$, $a \in \mathbb{R}$, le barycentre du triangle équilatérale décrit par les trois hydrogènes et $h > 0$ la longueur des côtés de ce triangle.

On trouve 2 phases stables, illustrées par la figure 14 :

- $a = \pm 0.5421$ Ångströms
- $h = 0.9403$ Ångströms

Pour connaître la quantité d'énergie dont a besoin la molécule pour passer d'une phase à une autre et se "retourner", on applique le lemme du col dans \mathbb{R}^n .

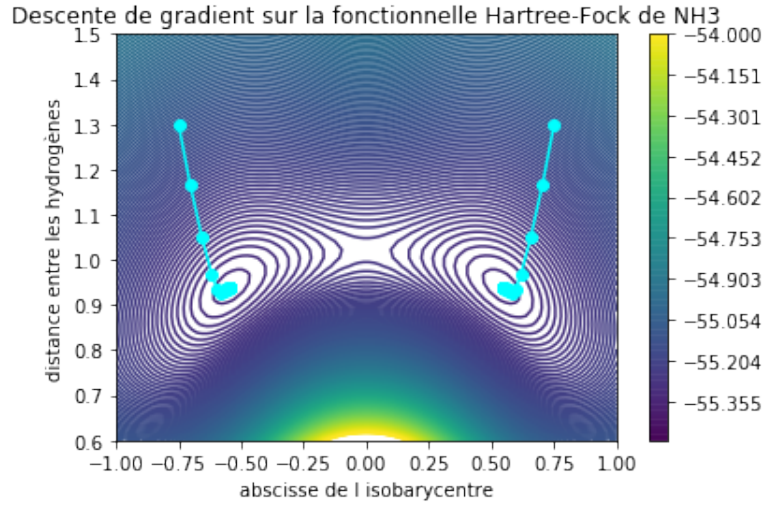


FIGURE 14 – Recherche de phases stables de NH3 par descente de gradient à pas fixe.

Soient $x_0 = (-a^*, h^*)$ et $x_1 = (a^*, h^*)$, $V \ni x_0, x_1$ un voisinage de 0 dans \mathbb{R}^2 , $I : V \rightarrow \mathbb{R} : (a, h) \mapsto I(a, h)$ la fonctionnelle d'énergie de NH3.

— I est bornée inférieurement et strictement croissante quand $x \rightarrow \partial V$, donc I vérifie la condition de Palais-Smale.

— Comme de plus x_0 et x_1 sont les deux seuls minima de I , l'hypothèse de séparation est aisément vérifiée.

Par lemme du col, on a donc l'existence d'un point col, noté $c = I(x^*)$, tel que :

— $\nabla I(x^*) = 0$

— $I(x^*) > \max(I(x_0), I(x_1))$

L'énergie nécessaire au retournement de NH3 est alors la différence entre l'énergie du système en phase stable et au point de transition :

$$\Delta E = I(x^*) - I(x_0) = I(x^*) - I(x_1)$$

Pour trouver ces valeurs, on applique l'algorithme "élastique", qui trouve le plus court chemin entre x_0 et x_1 . Le résultat est donné en figure 15.

L'énergie est en Hartrees (Ha).

On trouve $I(x^*) = -55.2722$ Ha, $I(x_a) = I(x_b) = -55.3118$ Ha, d'où

$$\Delta E = -55.2722 - (-55.3118) = 0.0396 \text{ Ha.}$$

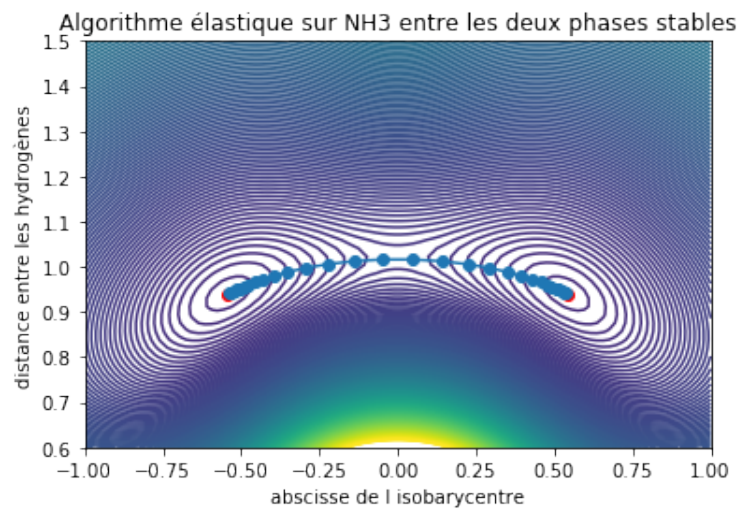


FIGURE 15 – Recherche du point de transition de NH3 par algorithme élastique.

5 Conclusion et ouverture

Voilà comment, à l'aide de méthodes itératives, on applique le lemme du col au calcul de l'énergie nécessaire à la transition de phase d'une molécule.

Par ailleurs, chaque section peut être prolongée indépendamment.

Ainsi, le lemme du col dans \mathbb{R}^n peut se généraliser à des espaces de Hilbert ou de Banach et offre une porte d'entrée vers l'étude des problèmes variationnels, c'est-à-dire l'étude de l'existence et/ou de l'unicité d'équations dont l'espace de départ est de dimension infinie. Le calcul variationnel offre un large champ d'applications.[2][6][5][4][3][2]

L'algorithme élastique que l'on utilise en sections 2 et 3 est une version "douce" : la fonctionnelle que l'on minimise est dérivable et la résolution du problème est équivalent à l'annulation du gradient de cette fonctionnelle. C'est la raison pour laquelle il nous suffit d'appliquer une méthode itérative de descente.

Cependant les contraintes du problème de recherche de point col ne sont pas différentiables, et Messieurs Moré et Munson [6] livrent une version "dure" de l'algorithme élastique, plus fidèle au problème posé.

Enfin, le modèle Hartree-Fock mériterait que l'on s'y intéresse de plus près. Il reste à éluder à l'aide d'analyse fonctionnelle le calcul des orbitales décrivant la fonction d'onde.[8]

FIGURE 16 – Douglas Hartree (1897-1958),
mathématicien et physicien anglais



FIGURE 17 – Vladimir Fock (1898-1974),
physicien soviétique



6 Remerciements

Je remercie chaudement David Gontier pour son optimisme constant et sa capacité à me sauver en très peu de temps de mes errances mathématiques chroniques. A quelques jours du rendu du mémoire, je m'apprêtais à rendre 20 pages sur le lemme du point col vierges de toute définition d'un point col...

C'est grâce à lui que j'ai pu tremper un orteil dans les eaux profondes de la chimie quantique, et ce domaine vient s'ajouter aux autres domaines d'applications des mathématiques que je connaissais auparavant pour épaissir ma conviction selon laquelle les mathématiques sont une porte sur le monde.

Merci aussi à Guillaume Vigerat pour avoir supervisé notre année de CPES avec un calme rassurant en toute circonstance.

Merci à mon père pour ses questions (et même quelques réponses) sur mon mémoire tout au long de son écriture, et à ma mère pour ses encouragements.

Enfin, merci à PSL de nous permettre d'effectuer ce CPES 3 après 2 années à Dauphine, et en particulier à l'Université Paris-Dauphine et aux enseignants que j'ai eu la chance de côtoyer : Madame Hénon, Messieurs Doss, Lamboley, Féjoz, Legendre, Frouvelle, Lehec... En ce moment, le rythme avec lequel j'élargis mon horizon mathématique s'accélère et vous y êtes forcément pour quelque chose.

7 Annexe : programmes Python et PySCF

7.1 Python

7.1.1 Calculs respectifs du gradient et de la hessienne de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ par différence finie

```
def diff_finie(f,t,eps_diff=1e-8):
    x = t[0]
    y = t[1]
    t_x = array([x + eps_diff,y])
    t_y = array([x,y + eps_diff])
    return array([(f(t_x) - f(t))/eps_diff,(f(t_y) - f(t))/eps_diff])

def diff_finie2(f,t,eps_d2=1e-5):
    x = t[0]
    y = t[1]
    dxx = (f([x+2*eps_d2,y]) - 2*f([x+eps_d2,y]) + f([x,y]))/eps_d2**2
    dxy = (f([x+eps_d2,y+eps_d2]) - f([x+eps_d2,y]) - f([x,y+eps_d2]) + f([x,y]))/eps_d2**2
    dyy = (f([x,y+2*eps_d2]) - 2*f([x,y+eps_d2]) + f([x,y]))/eps_d2**2
    return array([[dxx,dxy],[dxy,dyy]])
```

7.1.2 Algorithmes de descente : algorithme de descente de gradient à pas fixe et algorithme de Newton

```
def GradientMethod_diff_finie(f,z0,step,tol,niterMax=100):
    zn = z0
    path = hstack([zn,f(zn)])
    dfz = diff_finie(f,zn)
    niter = 0
    #iteration
    while(norm(dfz)>tol and niter<niterMax):
        niter += 1
        zn = zn - step*dfz
        dfz = diff_finie(f,zn)
        path = vstack([path,hstack([zn,f(zn)])])
    if niter < niterMax:
        print("Gradient Method converged in ",niter,"iterations\n")
    else:
        print("Gradient Method did not converge")
    return array(path)

def NewtonMethod_diff_finie(f,z0,tol,niterMax=100):
    zn = z0
    path = hstack([zn,f(zn)])
    dfz = df(zn)
    ddfz = ddf(zn)
    niter = 0
    #iteration
    while(norm(dfz)>tol and niter<niterMax):
```

```

        niter += 1
        zn = zn - dot(inv(diff_finie2(f,zn,1e-1)),diff_finie(f,zn,1e-1))
        dfz = diff_finie(f,zn)
        ddfz = diff_finie2(f,zn)
        path = vstack([path,hstack([zn,f(zn)])])
    if niter < niterMax:
        print("Newton Method converged in ",niter,"iterations\n")
    else:
        print("Newton Method did not converge")
    return array(path)

```

7.1.3 Algorithme élastique

```

    #init of the functional
    z0 = array([1, 1.5])
    z1 = array([-1,1.5])
    step,tol = 0.2, 0.001
    array_GM_a = GradientMethod_diff_finie(ammoniac,z0,step,tol,niterMax=100) #minimum proche de z0
    array_GM_b = GradientMethod_diff_finie(ammoniac,z1,step,tol,niterMax=100)
    #minimum proche de z1
    a,b = array(array_GM_a[-1][0:2]), array(array_GM_b[-1][0:2])

    n = 30
    #tab is a discrete uniformed range of n points between a and b, a and b are excluded
    tab = zeros((2,n))
    for i in range(1,n+1) :
        zi = a*(n+1-i)/(n+1) + b*(i)/(n+1)
        tab[:,i-1] = zi

    alpha = 100./n #

def Elastic(f, tab, alpha, a, b):
    #The fonctional has the form : sum(f(zi)) + alpha*(sum(norm(zi-z(i+1))^2) + norm(z1-a)^2 + no
    n = size(tab,1)
    energy = 0
    for i in range(n):
        energy += alpha * f(tab[:,i]);
        if (i != n-1):
            energy += norm(tab[:,i+1] - tab[:,i])**2
    energy += norm(a - tab[:,0])**2
    energy += norm(b - tab[:,n-1])**2
    return energy;

def dElastic(f, tab, alpha, a, b):
    n = size(tab,1)
    dE = zeros((2,n))
    #end points first
    dE[:,0] = diff_finie(f,tab[:,0]) + 2*alpha*(2*tab[:,0]-tab[:,1]-a)
    dE[:,n-1] += diff_finie(f,tab[:,n-1]) + 2*alpha*(2*tab[:,n-1]-tab[:,n-2]-b)

```



```

    for i in range(1,n-1):
        dE[:,i] = diff_finie(f,tab[:,i]) + 2*alpha*(2*tab[:,i]-tab[:,i-1]-tab[:,i+1])
    return dE

step = 0.01
tol = 0.1

def Adapted_GradientMethod(Eelastic,dEelastic, f, tab, alpha, a, b,step,tol,niterMax=500):
    tabn = tab
    dftab = dEelastic(f, tab, alpha, a,b)
    niter = 0
    #iteration
    while(norm(dftab)>tol and niter<niterMax):
        E = Eelastic(f, tabn, alpha, a, b)
        if niter%5 == 0:
            print('iteration n = '+str(niter)+'\t energy = '+str(E))
            niter += 1
            tabn = tabn - step*dftab
            dftab = dEelastic(f, tabn, alpha, a, b)
    if niter < niterMax:
        print("Gradient Method converged in ",niter,"iterations\n")
    else:
        print("Gradient Method did not converge")
    return tabn

path_GM = Adapted_GradientMethod(Eelastic,dEelastic,ammoniac, tab, alpha, a, b, step, tol)

```

7.2 PySCF

7.2.1 Molécule d'eau H2O

```

def energieH2O(x):
    #retourne l'énergie de H2O où la distance H-O est 1, et l'angle H-O-H est theta
    l, theta = x[0],x[1]
    m = gto.Mole()
    m.atom = [['O', [0,0,0]],['H',[1,0,0]], ['H',[1*cos(theta),1*sin(theta),0]]]
    m.basis = 'sto-3g'
    m.build()
    e = scf.RHF(m)
    conv, etot, mo_e, mo, mo_occ = scf.hf.kernel(scf.hf.SCF(m))
    return etot

```

7.2.2 Molécule d'ammoniac NH3

```

def ammoniac(x):
    #a isobarycentre des hydrogènes (distance entre N et le plan des H)
    #h distance entre le centre des H et un H.
    a,h = x[0],x[1]
    am = gto.Mole()

```

```

am.atom = [['N', zeros(3)],
           ['H', a, h/2, -h*sqrt(3)/2],
           ['H', a, -h/2, -h*sqrt(3)/2],
           ['H', a, 0, h]]
am.build()
e = scf.RHF(am)
etot = e.kernel()
print("a = ",a," , h = ",h," . ")
return etot

```

Références

- [1] Ivar Ekeland, *On the Variational Principle* UER Mathématiques de la Décision, Université Paris IX, 1974.
- [2] Antonio Ambrosetti et Paul H. Rabinowitz, *Dual Variational Methods in Critical Point Theory and Applications* Scuola Normale Superiore, Pisa, University of Wisconsin, Madison, 1973.
- [3] Harris Enniss, *A Refined Saddle-Point Theorem and Applications*. Harvey Mudd College, 2012.
- [4] Youssef Jabri, *The Mountain Pass Theorem : Variants, Generalizations and Some Applications* Cambridge University, 2003.
- [5] Fengying Li, Bingyu Li, Shiqing Zhang, *Generalized Mountain Pass Lemma Related with a Closed Subset and Locally Lipschitz Functionals* School of Economic and Mathematics, Southwestern University of Finance and Economics, 2014.
- [6] Jorge J. Moré et Todd S. Munson, *Computing Mountain Passes*. Mathematics and Computer Science Division ARGONNE NATIONAL LABORATORY, 2002.
- [7] Antoine Levitt, Christoph Ortner *Convergence and Cycling in Walker-type Saddle Search Algorithms* 2016.
- [8] Eric Cancès, Claude Le Bris, Yvon Maday, *Méthodes mathématiques en chimie quantique. Une introduction* Springer-Verlag Berlin Heidelberg 2006.
- [9] Jean-Louis Basdevant, Jean Dalibard, *Mécanique quantique* Ecole Polytechnique 2002.
- [10] *PySCF — Python-based simulations of chemistry framework* <http://sunqm.github.io/pyscf/>