



Mémoire de fin d'études

PRESENTÉ EN VUE DE L'OBTENTION
DU DIPLOME DE : Master

Filière : Physique
Option : Physique des Matériaux

THÈME :

**Méthode de la descente du gradient avec
intégration Monte Carlo appliquée à l'atome
d'hélium.**

Préparé par : Sakhraoui Manel

Soutenu le : 28/05/2024

Devant le jury :

Président : Daoudi Salim

Rapporteur : Khalfallah Farid

Examineur : Kahoul Abdelhalim

Examineur : Koussa Walid

Pr. Université de BBA

M.C.A Université de BBA

Pr. Université de BBA

Dr. Université de BBA

Année Universitaire 2023-2024

Mémoire de Master

Méthode de la descente du gradient avec intégration Monte Carlo
appliquée à l'atome d'hélium

Département des Sciences de la Matière
Faculté des sciences et de la technologie
Université de BBA

Sakhraoui Manel.

May 28, 2024



Dédicace et remerciements

Louanges et remerciements à Dieu qui nous a aidés à faire cette humble œuvre. A cette occasion, je remercie infiniment tous ceux qui ont contribué au succès de cette thèse, laquelle je dédie à:

- *Ma source de bonheur et d'énergie, mes parents qui se sont sacrifiés pour ma réussite, qui m'ont donné du courage, du soutien, et de la confiance dans ma vie.*
- *Mes deux frères Adem et Moussa qui m'ont donné de l'esprit et de la motivation.*
- *L'âme de mes grands parents (Ali, Hafssa et Taoues), sans oublier mes chers grands-parents vivants (Yamina et Hassen), que Dieu prolonge leur âge.*
- *Mes oncles, mes tantes et toute la famille et les amies.*

Je tiens à remercier monsieur Farid Khalfallah pour son encadrement professionnel avec son suivi, ses précieux conseils, et sa disponibilité qui m'a donné la première clé vers le domaine de la recherche.

Je remercie aussi les membres du jury Pr. Daoudi Salim (Président), Pr. Kahoul Abdelhalim et Dr. Koussa Walid (Examineur) pour la lecture attentive de ce mémoire et pour les éventuelles remarques qu'ils m'adresseront lors ma soutenance.

Comme j'adresse mes sincères remerciements à tous les enseignants et administrateurs du département science de la matière, à toute l'équipe pédagogique de l'université et à tous ceux qui nous ont aidés durant nos années d'étude.

Table des Matières

Dédicace et remerciements	i
Table des Matières	ii
Liste des Figures	iv
Liste des Tableaux	v
Introduction	1
1 Méthode de Monte Carlo	3
1.1 Introduction aux méthodes Monte Carlo	4
1.2 Principes de base de la méthode Monte Carlo	4
1.2.1 Notions de probabilités et statistiques	5
1.2.2 Théorèmes de convergence	11
1.3 Simulation Monte Carlo	12
1.3.1 Génération des nombres aléatoires	13
1.4 Intégration Monte Carlo multidimensionnelle	16
1.4.1 Principe	16
1.4.2 types d'intégration Monte Carlo	17
1.5 Le problème de la dimensionalité	21
1.6 Intégration Monte Carlo : Étude comparative	22
1.6.1 Application	23
1.6.2 Présentation des résultats et Discussion	25
2 Méthode variationnelle et descente du gradient	27
2.1 Introduction	28
2.2 Méthode variationnelle en mécanique quantique	28
2.2.1 Principes	28
2.2.2 Théorème de Rayleigh-Ritz	29
2.2.3 A propos de l'optimisation	31
2.2.4 Difficultés de l'optimisation (minimisation)	31
2.3 Méthode de la descente du gradient	32
2.3.1 Le gradient	32
2.3.2 Introduction à la méthode de la descente du gradient	33

2.3.3	Algorithme de la méthode	35
3	Application à l'atome d'hélium	36
3.1	Introduction	37
3.2	Rappel théorique	37
3.2.1	Équation de Schrödinger pour l'atome d'hélium	39
3.2.2	Approximation des électrons indépendants	40
3.3	Méthode variationnelle pour l'atome d'hélium	40
3.3.1	Fonction d'onde test	40
3.3.2	Calcul de la moyenne $\langle H \rangle$	42
3.3.3	Application de la méthode de la descente du gradient	45
3.3.4	Intégration Monte Carlo	46
3.3.5	Organigramme et code de calcul :	48
3.3.6	Résultats et discussion	50
	Conclusion	56
	Bibliographie	59

Liste des Figures

1.1	Représentation graphique d'une densité de probabilité uniforme sur l'intervalle $[0,1]$. . .	9
1.2	Représentation graphique de la densité de probabilité d'une loi Gaussienne centrée réduite	10
1.3	Représentation graphique de la densité d'une distribution exponentielle avec un paramètre $\lambda = 1$	11
1.4	Méthode d'inversion pour une distribution normale (fonction gaussienne) [26].	14
1.5	Illustration de la méthode de rejet [25].	15
1.6	Histogramme des distributions uniforme (à gauche), gaussienne (au milieu) et exponentielle (à droite) respectivement pour un échantillon de 100000 événements.	15
1.7	intégration de Gauss-Legendre à 2 nœuds : Nombre de points d'intégration en fonction de la dimension n pour trois subdivision des axes : $m = 1, 2$ et 5 (en échelle log) [24]	22
1.8	Représentation graphique de la fonction $f(\vec{r}), g(\vec{r})$ à $n = 1$ dimension	24
1.9	Représentation graphique de la fonction $f(\vec{r}), g(\vec{r})$ à $n = 2$ dimensions	24
1.10	Valeur de l'intégrale IMC par les trois méthodes d'intégrations Monte Carlo avec barres d'erreurs en fonction du nombre de points N	25
1.11	L'erreur statistique relative pour les trois méthodes d'intégration Monte Carlo en fonction du nombre de points N	26
2.1	Exemple de courbes de niveaux de la fonction à 2 variables $f(x, y) = x^2 + 2y^2$ et son gradient [33].	33
2.2	Courbe-niveaux $F_X(x)$	34
2.3	Courbe du descente du gradient à 2D.	34
3.1	Représentation d'atome d'hélium	38
3.2	Organigramme de notre calcul (Calcul variationnel par descente du gradient et Monte Carlo)	49
3.3	Convergence des paramètres (α, β) et de l'énergie $E(\alpha, \beta)$ en fonction du nombre d'itérations pour $(\alpha_0, \beta_0)=(2,1.5)$ et $N = 10^4$	51
3.4	Convergence des paramètres (α, β) et de l'énergie $E(\alpha, \beta)$ en fonction du nombre d'itérations pour $(\alpha_0, \beta_0)=(2,1.5)$ et $N = 10^5$	52
3.5	Convergence des paramètres (α, β) et de l'énergie $E(\alpha, \beta)$ en fonction du nombre d'itérations pour $(\alpha_0, \beta_0)=(2,1.5)$ et $N = 10^6$	53
3.6	Convergence des paramètres α, β et représentation de l'énergie en fonction du nombre d'itérations pour $(\alpha_0, \beta_0)=(1.688, 1.687)$, $\delta = 0.1$, $N_i = 20 + 5$ et $N = 5 \cdot 10^6, 10^7$	55

Liste des Tableaux

1.1	Valeurs de l'intégrale IMC en fonction du nombre de tirages N pour les trois méthodes .	25
1.2	Valeurs de l'erreur exacte relative (%) en fonction du nombre de tirages N pour les trois méthodes MC	25
1.3	Valeurs de l'écart-type relatif (%) en fonction du nombre de tirages N pour les trois méthodes MC	25
3.1	Valeurs obtenues par le calcul fin des paramètres α, β et de l'énergie pour $(\alpha_0, \beta_0)=(1.688, 1.687)$, $\delta = 0.1$, $N_i = 20 + 5$ et $N = 5 \cdot 10^6, 10^7$	54

Introduction

Depuis plus de 60 ans d'existence, la simulation numérique a révolutionné les secteurs majeurs de l'industrie comme elle a renouvelé de nouvelles méthodes d'études et d'expérimentations de certain systèmes complexes, elle permet de prédire leurs comportements et d'optimiser leur conception, elle trouve une large gamme d'utilisation dans la plupart des domaines comme l'aéronautique, la biologie, la climatologie et la physique. Parmi les nombreux domaines de la physique où la simulation est cruciale on peut citer la physique des matériaux, la physique atomique et nucléaire et généralement toutes les branches de la physique où le calcul et la résolution des équations différentielles devient trop compliqué à cause de la complexité et le grand nombre de degré de liberté nécessitant des approximations. Les problèmes à N corps en mécanique quantique sont l'une des disciplines qui utilise le plus les méthodes d'approximations et le calcul numérique.

On distingue parmi les méthodes d'approximation en mécanique quantique, la méthode variationnelle que l'on peut coupler avec des méthodes numériques d'optimisation et d'intégration comme les méthodes du gradient et les méthodes de Monte-Carlo.

Dans cette thèse, nous allons présenter la méthode de minimisation dite méthode de la descente du gradient couplée à la méthode de Monte Carlo comme une méthode d'intégration dans un espace multidimensionnel pour l'estimation de quelques propriétés quantiques de l'état fondamental de l'atome d'hélium à l'aide de la méthode variationnelle.

Dans le premier chapitre, nous allons développer la méthode de Monte Carlo décrivant ces éléments de base ainsi que des concepts théoriques en théorie des probabilités qui sont à la base de l'intégration de Monte Carlo notamment le concept d'espérance et de variance, les nombres aléatoires et les méthodes de générations, puis on va présenter les trois types d'intégrations Monte Carlo et leurs utilisation surtout dans les intégrales multidimensionnelles et terminer par une étude comparative entre ces trois types.

Dans le deuxième chapitre, nous introduisons la théorie de la méthode variationnelle en mécanique quantique et son utilisation pour le calcul approximatif des énergies propres d'un système quantique. Nous présentons ensuite une technique numérique d'optimisation qui est la méthode de la descente du gradient et son algorithme pour la recherche du minimum d'une fonction à plusieurs variables.

Le troisième chapitre est consacré à la réalisation d'une application de ces méthodes sur

l'atome d'hélium comme un problème à trois corps. Il s'agira d'appliquer la méthode variationnelle à deux paramètres et une fonction d'essai basée sur l'approximation des électrons indépendants. La minimisation de l'énergie qui contient des intégrales 6-dimensionnelles se fera par la méthode de la descente du gradient couplée à l'intégration de Monte Carlo. A la fin du chapitre nous discutons les résultats obtenus et on les compare avec les valeurs théoriques exactes.

On termine notre mémoire par une conclusion générale ou nous résumons nos principaux résultats ainsi que quelques perspectives.

Chapitre 1

Méthode de Monte Carlo

1.1 Introduction aux méthodes Monte Carlo

Les méthodes Monte-Carlo (MC) représentent un ensemble de procédés algorithmiques utilisant le hasard pour résoudre des problèmes déterministes complexes afin de calculer une valeur numérique approchée caractérisée par un grand nombre de paramètres. Ces méthodes sont basés sur des techniques probabilistes et des simulations aléatoires.

Le nom des méthodes Monte Carlo, retourne aux jeux du hasard pratiqués au célèbre lieu de jeu le casino de Monte-Carlo à Monaco le pays européen, parce que la méthode partage les mêmes caractéristiques aléatoires qu'un jeu de roulette.

Historiquement la méthode est née en 1777 au comte du Buffon qui, avec l'estimation de calcul de la valeur numérique de π d'une façon approximative se basant sur la réalisation d'expériences répétées. Le développement des méthodes Monte-Carlo est liée à l'apparition des premiers ordinateurs et à leur utilisation dans le cadre des projets secrets du département de la défense des États Unis dans les années 40-45 en vue de la conception des premières bombes atomiques par les scientifiques de Los Alamos après la seconde guerre mondiale[1]. La technique est employée pour le calcul d'intégrales, la recherche d'extrema, la résolution de systèmes d'équations, d'équations différentielles, d'équations aux dérivées partielles et d'équations intégrales.

La première utilisation de la méthode était par le physicien mathématicien américain Nicholas Metropolis en 1947 [14], c'est également l'inventeur du nom de la méthode et d'un algorithme éponyme célèbre d'amélioration de la méthode de Monte-Carlo, avec d'autres groupes de chercheurs comme Stanislaw Ulam qui avait développé la méthode. Nicholas Metropolis a utilisé la méthode de Monte Carlo comme une méthode d'intégration pour développer ces recherches au domaine nucléaire, ou il a établi le recâblage de l'ordinateur ENIAC pour effectuer des simulations d'un noyau nucléaire en 1948[3].

Des scientifiques de plusieurs domaines trouvent que la simulation Monte Carlo compte parmi les outils efficaces et les plus utilisés dans de nombreux domaines notamment dans les mathématiques, la physique, l'astrophysique et la physique des particules, la chimie, l'informatique et l'intelligence artificielle, la gestion des projets et la tarification...ect, ou la technique a prouvé sa puissance d'évaluer de l'impact du risque dans ces domaines.

1.2 Principes de base de la méthode Monte Carlo

La méthode Monte Carlo est une méthode d'approximation qui n'a pas une définition bien précise, mais on peut dire quelle est un outil stochastique basé sur les statistiques et les probabilités et utilisant une technique d'hasardisation pour résoudre des problèmes complexes de hautes dimensions. L'application de la méthode passe par deux processus essentiels. D'abord la représentation mathématique (mise en équations) ou l'interprétation des modèles de phénomènes de la réalité physique à condition qu'ils doivent être accessi-

bles à l'analyse et au calcul, ce processus est nommé la modélisation. Puis, la simulation numérique qui nous permet de calculer sur ordinateur les solutions de ces équations ou ces modèles et donc simuler la réalité physique [4].

1.2.1 Notions de probabilités et statistiques

Les probabilités sont une branche des mathématiques qui nous permet de faire la modélisation des phénomènes aléatoires sur des systèmes ou les résultats sont pas fixes et dont on ne peut estimer avec certitude. L'apparition de cette branche renvoie au trois derniers siècles, mais les premières recherches à la théorie des probabilités ont commencés au quinzième siècle en 1494, par le mathématicien Italien "Luca pacioli" dans son article ou il a discuté les jeux du hasard, puis plusieurs chercheurs mathématiciens, physiciens ou bien des philosophes qui ont travaillé dans ce domaine comme "Girolano Cardano" qui a présenté un ensembles de règles afin d'aider à résoudre des problèmes des jeux du hasard. On trouve aussi "Galilio Galili" en 1606, Pascale en 1654 avec son théorème des points, "Ishak Newton" en 1687,"Bernoulli" 1713, et bien d'autres comme le mathématicien russe "Andrique Markov" 1922 le spécialiste dans les probabilités et qui a inventé ce qu'on appelle "les chaines de Markov". Et à partir du 19ème siècle l'application des méthodes statistiques à la science s'est accrue et une plus grande mise sous attention du concept de probabilité.

Définition : Une probabilité est une application $\mathcal{P}(\Omega)$, sur Ω l'univers des possibles ou l'ensemble des résultats possibles, dans $\mathbb{R}[0, 1]$ tel que :

- $0 \leq P(A) \leq 1$, pour tout événement $A \in \Omega$
- $P(\cup A_i) = \sum_i P(A_i)$. C-à-d que la probabilité d'un événement qui est la réunion disjointe d'événements est égale à la somme des probabilités de ces événements .
- $P(\Omega) = 1$. La probabilité de l'ensemble total des événements est égal à 1

1.2.1.1 Densité de probabilité

La densité de probabilité $\rho(x)$ est la fonction qui représente la distribution probabiliste d'une variable aléatoire continue X sur une intervalle donnée, elle est définie de $\mathbb{R} \rightarrow \mathbb{R}$ par le rapport de la probabilité dp de trouver la variable aléatoire X dans l'intervalle $[x, x + dx]$ et la longueur de dx l'intervalle.

$$\rho(x) = \frac{dp}{dx} \tag{1.1}$$

Elle est caractérisé par les deux propriétés suivantes :

1. $\rho(x)$ est toujours positive.

2. Elle est normalisée sur l'intervalle de définition (probabilités totale égale à 1)

$$\int_{-\infty}^{+\infty} \rho(t) dt = 1 \quad (1.2)$$

1.2.1.2 Fonction de répartition

La fonction de répartition est une fonction croissante sur \mathbb{R} qui prend les valeurs de 0 à 1, elle est définie comme la probabilité de trouver la variable aléatoire X dans une intervalle $]-\infty, t]$. Elle est mathématiquement représentée par la formule suivante [4] :

$$\forall x \in \mathbb{R}, F(t) = \int_{-\infty}^t \rho(x) dx \quad (1.3)$$

telle que :

1. ρ est la densité de probabilité.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$.

1.2.1.3 Variables aléatoires

Une variable aléatoire (v.a.) est une application X d'un ensemble d'événements (ou espace d'échantillons) Ω dans \mathbb{R} . Concrètement, une v.a. prend une valeur numérique résultant d'un processus ou d'une expérience dont la valeur ne peut être prédéterminée par les conditions initiales et qu'elle est due au hasard. Cet aspect aléatoire n'est pas absolu étant donné que chaque v.a. est distribuée selon une densité de probabilité donnée. L'échantillonnage des variables aléatoires consiste à tirer un ensemble fini de valeurs numériques de la variable aléatoire selon sa densité de distribution.

Les principaux paramètres statistiques liés à la variable aléatoire sont : l'espérance mathématique ou la valeur moyenne, la variance et l'écart-type ainsi que la covariance et le coefficient de corrélation. Ces paramètres permettent d'étudier et vérifier la cohérence des données d'un certain événement, voir les erreurs des résultats, la dépendance entre les variables aléatoires... etc

1.2.1.4 L'espérance mathématique

L'espérance mathématique ou bien la valeur moyenne d'une variable aléatoire est la somme des probabilités de chaque résultat possible d'une expérience multipliée par sa valeur. Elle représente la valeur moyenne pondérée de toutes les données disponibles. Elle est donnée dans le cas discret par :

$$E(X) = \langle X \rangle = \sum_i x_i P(X = x_i) \quad (1.4)$$

Et dans le cas continu, par :

$$E(X) = \int_{-\infty}^{+\infty} x\rho(x)dx \quad (1.5)$$

Une fonction $f(X)$ d'une variable aléatoire X admet une espérance mathématique notée $E[f(X)]$ à condition que l'intégrale existe et converge, donnée par la formule:

$$E[f(X)] = \langle f(X) \rangle = \int_{-\infty}^{\infty} f(x)\rho(x)dx \quad (1.6)$$

Si la densité de probabilité $\rho(x)$ est normalisée sur un intervalle fini $[a, b]$, c'est-à-dire:

$$\int_a^b \rho(x)dx = 1 \quad (1.7)$$

Alors l'espérance s'écrit:

$$E[f(X)] = \langle f(X) \rangle = \int_a^b f(x)\rho(x)dx \quad (1.8)$$

Propriétés : Soit X et Y deux variables aléatoires définies sur Ω fini et α, β deux nombres réels, on a les propriétés suivantes:

- $E(X + Y) = E(X) + E(Y)$
- $E(\alpha X + \beta) = \alpha E(X) + \beta$
- $E(XY) = E(X)E(Y)$ Si les deux variables aléatoires sont indépendantes.

1.2.1.5 La variance

La variance d'une variable aléatoire est une grandeur qui indique la dispersion de cette variable autour de sa valeur moyenne, notée $Var(X)$ ou bien σ^2 , elle est donnée par la formule de koenig suivante:

$$Var(X) = E([X - E(X)]^2) = \int_{-\infty}^{\infty} [X - E(X)]^2 \rho(x)dx \quad (1.9)$$

Telle que $\rho(x)$ est la fonction de densité de probabilité normalisée.

Propriétés : Si la variable aléatoire X admet une variance, alors:

- $Var(X) \geq 0$
- Si $Var(X) = 0$ alors X est une constante.
- $E(X) = 0$ alors la variable aléatoire X est centrée sur l'origine.
- $Var(X) = 1$ alors la variable aléatoire X est dite réduite.

- $Var(X + Y) = Var(X) + Var(Y)$
- $Var(\alpha X + \beta) = \alpha^2 Var(X)$. α, β des constantes

Formule de Huygens : On peut démontrer que la formule de Koenig peut se réduire à la formule de Huygens suivante:

$$Var(X) = E(X^2) - E^2(X) \quad (1.10)$$

En effet :

$$\begin{aligned} Var(X) &= E([X - E(X)]^2) \\ &= E(X^2 - 2XE(X) + (E(X))^2) \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2 \\ &= E(X^2) - E^2(X) \end{aligned}$$

1.2.1.6 L'écart-type

Pour avoir une bonne compréhension des dispersions, on calcul l'écart-type (nommé aussi écart quadratique moyen) qui donne une indication des écarts possibles par rapport à la valeur moyenne. Il est définie par la racine carrée de la variance et s'exprime par :

$$\sigma(X) = \sigma_x = \sqrt{Var(X)} \quad (1.11)$$

Ces paramètres énoncés dans la dernière partie, ont une grande importance dans l'évaluation des incertitudes des variables aléatoires.

1.2.1.7 Quelques lois de probabilités usuelles

La loi uniforme : On dit que la variable aléatoire X suit une loi uniforme sur un intervalle $[a, b]$, et on note $X \sim U[a, b]$, si sa densité de probabilité est constante sur cet intervalle. Elle est définie comme suit [16]:

$$\rho(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a, b] \\ 0, & \text{si } x \notin [a, b] \end{cases}$$

Et dans ce cas la variable aléatoire X admet une fonction de répartition définie comme :

$$F(x) = \begin{cases} 0, & \text{si } x < a \\ \frac{x-a}{b-a}, & \text{si } a \leq x \leq b \\ 1, & \text{si } x > b \end{cases}$$

Une espérance mathématique :

$$E(X) = \frac{a+b}{2}$$

Et une variance de la forme:

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

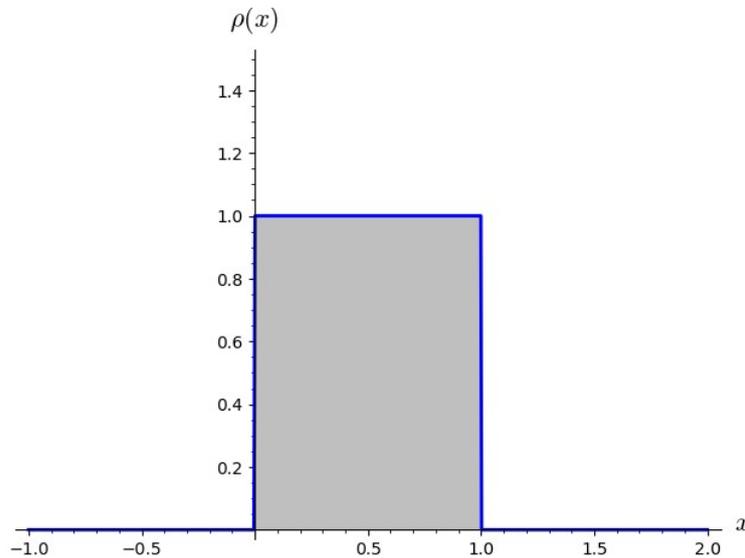


Figure 1.1: Représentation graphique d'une densité de probabilité uniforme sur l'intervalle $[0,1]$.

Loi normale, ou la loi gaussienne : Une variable est distribuée selon la loi normale ou la loi de Laplace Gauss avec des paramètres μ (espérance), et σ^2 (variance), et on note $X \mapsto \mathcal{N}(\mu, \sigma^2)$, si et seulement si X admet pour densité de probabilité une fonction $\rho_{\mu, \sigma}(x)$ pour tout $x \in \mathbb{R}$ de la forme [17]:

$$\rho_{\mu, \sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1.12)$$

Sa fonction de répartition est définie comme:

$$F_{\mu, \sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (1.13)$$

et son espérance et variance sont données par :

$$E(X) = \mu \quad \text{et} \quad \text{Var}(X) = \sigma^2 \quad (1.14)$$

Loi normale centrée réduite : On dit que la variable aléatoire suit une loi normale centrée réduite si cette variable est centrée sur $X = 0$, ce qui veut dire que sa moyenne est nulle $\mu = 0$, et si sa variance $\sigma^2 = 1$. Elle est notée $N(0, 1)$. Sa densité et sa fonction de répartition s'écrivent alors:

$$\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (1.15)$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{(-\frac{x^2}{2})} dx \quad (1.16)$$

La loi normale centrée réduite a une grande importance grâce à ces propriétés intéressantes et elle joue un rôle central dans de nombreux modèles probabilistes.

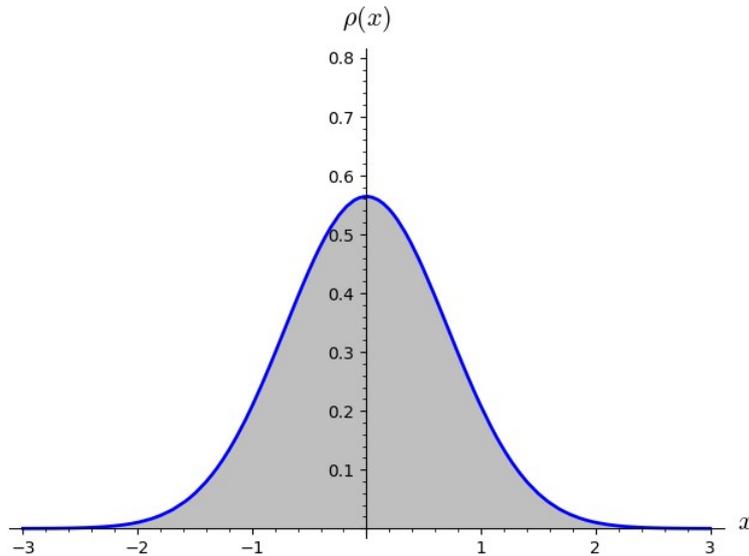


Figure 1.2: Représentation graphique de la densité de probabilité d'une loi Gaussienne centrée réduite

La loi exponentielle : Une variable aléatoire X est dite de loi exponentielle dans l'intervalle, avec un paramètre λ (réel strictement positif), si cette variable admet la densité suivante [17]:

$$\rho(x) = \begin{cases} \lambda e^{(-\lambda x)}, & \text{pour } x \geq 0 \\ 0, & \text{pour } x < 0 \end{cases}$$

Et elle admet une fonction de répartition de la forme:

$$F(x) = \begin{cases} 1 - e^{(-\lambda x)}, & \text{si } x \geq 0 \\ 0, & \text{si } x < 0 \end{cases}$$

Une espérance et une variance donnés par:

$$E(X) = \frac{1}{\lambda} \quad \text{et} \quad \text{Var}(X) = \frac{1}{\lambda^2}$$

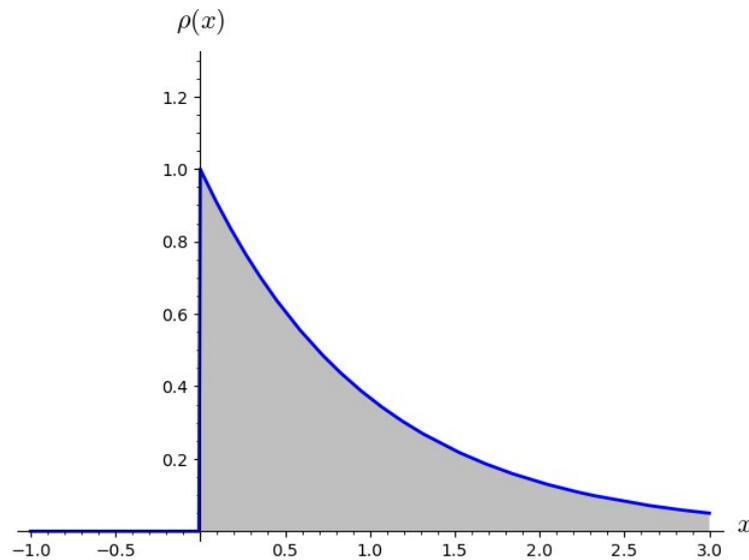


Figure 1.3: Représentation graphique de la densité d'une distribution exponentielle avec un paramètre $\lambda = 1$

Cette loi est utilisée dans de nombreuses applications tel que:

- Durée de fonctionnement d'un matériel informatique avant la première panne.
- Désintégration radioactive.
- Temps séparant l'arrivée de deux "clients" dans un phénomène d'attente[16].

1.2.2 Théorèmes de convergence

Parmi les concepts les plus importants en théorie des probabilité est celle de la convergence, cette notion est basée sur deux théorèmes essentiels qui sont la loi des grands nombres qui énonce la convergence de la moyenne empirique de variables aléatoires indépendamment distribuées et le théorème de la limite centrale limite qui indique la nature de la convergence et sa vitesse [5].

1.2.2.1 Loi des grands nombres:

Cette loi stipule que la moyenne arithmétique (ou empirique) \bar{X} d'une série de variables aléatoires (X_n) identiques et indépendantes l'une de l'autre converge vers l'espérance mathématique $E(X)$ lorsque la taille de l'échantillon tend vers l'infini :

$$E(X) \xrightarrow{N \rightarrow \infty} \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (1.17)$$

1.2.2.2 Théorème centrale limite

Soit $X_n, (n = 1, \dots, N)$ N variables aléatoires indépendantes et identiquement distribuées possédant la même espérance mathématique $\mu = E(X)$ et variance σ^2 . Soit $S_n = \sum_n X_n$ la somme de ces v.a., définissons alors la variable aléatoire Y_n suivante :

$$Y_N = \bar{X} = \frac{S_N}{N} = \frac{1}{N} \sum_n X_n$$

Le théorème centrale limite stipule que Y_N converge vers une loi normale $\mathcal{N}(\mu, \sigma_y^2)$ lorsque $N \rightarrow \infty$ [18], avec (en posant $Y = \lim_{N \rightarrow \infty} Y_N$) :

- L'espérance de Y est $E(Y) = \mu$
- L'écart-type de Y est $\sigma_y = \frac{\sigma}{\sqrt{N}}$

1.3 Simulation Monte Carlo

La simulation est un outil de calcul basé sur l'imitation d'un système réel ou d'un phénomène physique. Elle tente de donner une présentation des caractéristiques des systèmes abstraits ou physiques, puis estimer et prédire les résultats dans des conditions artificielles similaires aux conditions naturelles. Le développement de machines de calcul de plus en plus puissante d'une part et de modèles mathématiques et numériques d'autres part a permis à la simulation de faire de très grandes avancées dans un très grand nombre de domaines. [20].

La méthode Monte Carlo est une méthode stochastique basée sur l'utilisation de procédés aléatoires et de techniques probabilistes dans le but d'effectuer des calculs d'approximation numériques. Elle est basée sur l'échantillonnage de variables aléatoires distribués convenablement selon le problème étudié.

Largement utilisées en analyse numérique et calcul de probabilité, elle sert surtout à simuler des phénomènes physiques complexes avec un grand nombre de degré de libertés [12]. La simulation Monte Carlo est généralement divisée selon les problèmes étudiés en deux grands axes qui sont :

Les problèmes déterministes : représentent des problèmes qui désignent l'absence objective du hasard, dont les résultats sont entièrement prévisibles ou ces problèmes ont une solution unique et sont reproductibles. On cite quelques exemples sur ce type de problèmes: la résolution des équations non-linéaires, les équations différentielles, le calcul d'intégrales...etc

Les problèmes non-déterministes : Problèmes stochastiques dont les résultats ne sont pas entièrement prévisibles en raison de l'incertitude et l'hasardisation [21] et qui peuvent présenter différents comportements même avec des entrées identiques, conduisant à plusieurs résultats possibles[20]. On donne comme exemples: les problèmes à haute dimensionnalité en physique statistique et quantique, les problèmes à n-corps, problèmes

de diffusion et de transport en physique des rayonnements, la simulation de réactions en physique nucléaire et de détecteurs des particules ...etc.

1.3.1 Génération des nombres aléatoires

1.3.1.1 Nombres pseudo-aléatoires:

En raison de l'impossibilité mathématique des algorithmes de produire des nombres parfaitement aléatoires, ils génèrent des nombres très proches de l'aléatoire dis nombres pseudo-aléatoires, ces algorithmes sont connus sous le nom de générateurs de nombres pseudo-aléatoires (En anglais: pseudo-random number generator). Ils utilisent des méthodes déterministes qui produisent des suites de nombres vérifiant de bonnes propriétés statistiques. Les nombres pseudo-aléatoires sont caractérisés par :

1. L'indépendance: L'algorithme génère des nombres supposés indépendants les uns des autres.
2. La difficulté de prédire le comportement du groupe: Il est difficile de repérer des groupes de nombres qui suivent une certaine règle.
3. Propriétés idéales: Les nombres pseudo-aléatoires s'approchent seulement des propriétés idéales des suites aléatoires parfaites.

Un générateur de nombres pseudo-aléatoire repose sur des formules mathématiques afin de produire des séquences de nombres aléatoires. Il démarre à partir d'un état de départ et si ce dernier est connu les nombres sont certes déterministes mais très proches d'un aléatoire parfaits [2]. Les générateurs de nombres pseudo-aléatoires ont prouvés leurs efficacité à travers le temps grâce à la puissance de calcul des ordinateurs et le résultat et de très bon résultats ont été obtenu dans plusieurs domaines

1.3.1.2 Méthodes de génération

Plusieurs méthodes sont principalement utilisées par les générateurs des nombres pseudo-aléatoires employées sur ordinateurs. Ces méthodes nécessitent de faire des analyses mathématiques pour déterminer le degré d'aléatoire de ces nombres distribués selon des lois bien définies.

Méthode des congruences: Cette méthode est utilisée pour la génération de nombres aléatoires pour des lois uniformes . Elle est basée sur l'utilisation du reste de la divisions euclidienne de deux nombres. On suppose une suite (x_n) telle que $0 \leq x_n \leq m - 1$, alors la suite s'écrit comme suit:

$$x_{n+1} = ax_i + b[\text{modulo } m] \quad (1.18)$$

En augmentant alors le nombre m , on obtient plus de générateurs de nombres aléatoires ou pseudo-aléatoires d'une façon arbitraire. L'avantage de cette méthode qu'elle permette de

calculer et générer de très grands nombres. Dans certain cas, des générateurs correspondants à des nombres sous la forme de séquences binaires, donc on écrit [5]:

$$x_{n+1} = ax_i + b [\text{mod } m = 2^N] \quad (1.19)$$

Avec cette séquence et un bon choix de paramètres a, b et m , on arrive à générer des nombres de bonne qualité.

Méthode d'inversion: Parmi les méthodes les plus connues dans la génération des variables aléatoires selon une loi non uniforme est celle de l'inversion. Pour appliquer cette méthode à l'aide d'un générateur de variable aléatoire u_n uniforme $\mathcal{U}[0, 1]$. On définit la fonction inverse $F^{-1} :]0, 1[\rightarrow \mathcal{R}$ de la fonction de répartition F par:

$$F^{-1}(u) = \inf\{t : F(t) \geq u\} \quad \text{pour tout } u \in]0, 1[\quad (1.20)$$

Si la variable u suit une loi uniforme $\mathcal{U}[0, 1]$, alors la variable $X = F^{-1}(u)$ suit la loi de fonction de répartition recherchée F :

$$P(X \leq x) = P[F^{-1}(U) \leq x] = P[U \leq F(x)] = F(x) \quad (1.21)$$

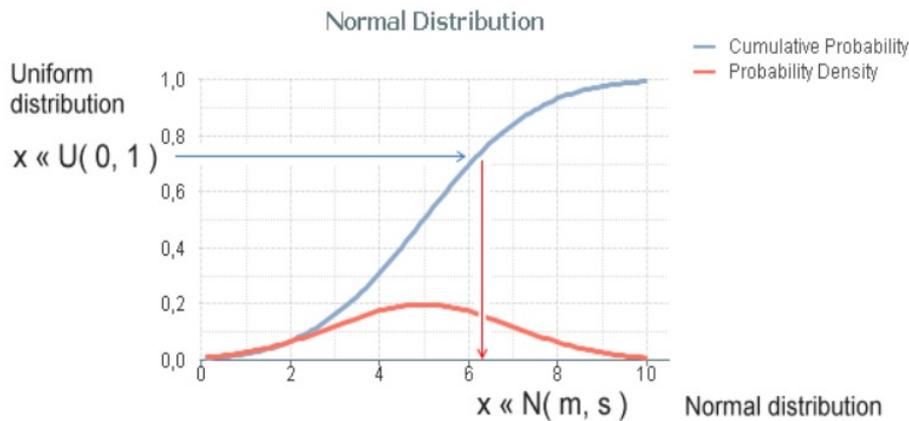


Figure 1.4: Méthode d'inversion pour une distribution normale (fonction gaussienne) [26].

Méthode d'acceptation-rejet: La simulation d'une variable aléatoire de densité $f(x)$ par la méthode de rejet nécessite la recherche d'une loi plus simple à simuler de densité $g(x)$, avec $f(x) \leq cg(x)$ [22] c'est-à-dire la nouvelle fonction $g(x)$ multipliée par une constante c doit couvrir la fonction étudiée $f(x)$, parfois $cg(x)$ est appelée l'enveloppe de $f(x)$. Et on simule après des variables X_n de densité $g(x)$ et des variables α_n uniformes sur l'intervalle $[0; 1]$, avec :

$$U_n(x) \leq \frac{f(X_n)}{cg(X_n)} \in [0, 1] \quad \Rightarrow \quad cU_n g(X_n) \leq f(X_n)$$

Ces formules conduisent à la génération des variables sous $cg(x)$ mais l'acceptation de seulement ceux qui se situent sous $f(x)$ représentés par le segment en vert sur la figure suivante qui est une illustration de la méthode de rejet.

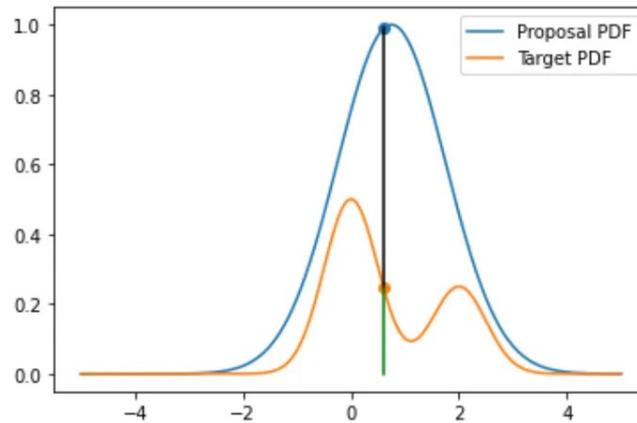


Figure 1.5: Illustration de la méthode de rejet [25].

Le point commun entre les deux méthodes d'inversion et de rejet pour les variables aléatoires continues et discrètes c'est l'utilisation d'un générateur de variables aléatoires uniforme sur l'intervalle $[0,1]$.

On montre sur les figures suivantes un exemple de simulation d'un échantillonnage selon trois distributions différentes avec la méthode de rejet.

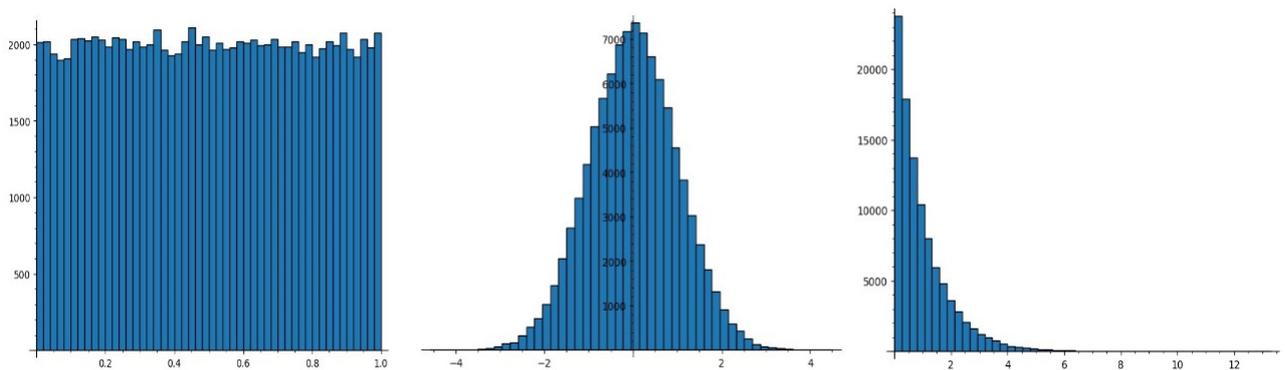


Figure 1.6: Histogramme des distributions uniforme (à gauche), gaussienne (au milieu) et exponentielle (à droite) respectivement pour un échantillon de 100000 événements.

1.4 Intégration Monte Carlo multidimensionnelle

Comme on vu, la méthode Monte Carlo repose sur la formulation des problèmes physique selon des modèles mathématique bien définie. La dimension du système étudié dépend des variables et du modèle choisi, et généralement on rencontre des problèmes de haute dimensionnalité. Les méthodes Monte Carlo travaillent avec ce type de systèmes, c'est pour ça qu'elles ce nomme: Méthodes Monte Carlo multidimensionnelles. L'une des applications les utilisées est l'intégration de fonctions multidimensionnelles.

1.4.1 Principe

Le principe de base de l'intégration Monte Carlo multidimensionnelle est l'utilisation de la propriété de l'espérance mathématique d'une fonction d'une ou plusieurs variables aléatoires. L'intégrale est transformée sous forme d'espérance de la fonction à intégrer qui est alors approximée par la valeur moyenne arithmétique selon la loi des grands nombres [1].

Rappelons que l'espérance d'une fonction d'une variable aléatoires x distribuée selon la densité $\rho(x)$ normalisée sur un intervalle $[a, b]$ est donnée par la formule suivante:

$$E[f(x)] = \langle f(x) \rangle = \int_a^b f(x)\rho(x)dx \quad (1.22)$$

Alors par la méthode de Monte Carlo, on peut approximer cette intégrale par un estimateur de l'espérance et selon la loi des grands nombres, la formule sera écrite comme suit:

$$E[f(x)] = \int_a^b f(x)\rho(x)dx \simeq \frac{1}{N} \sum_{i=1}^N f(X_i) = \bar{Y} \quad (1.23)$$

Avec N est le nombre de tirages de la variable X . Cette quantité \bar{Y} est appelée aussi valeur moyenne arithmétique de Y .

La variance est le deuxième outil qu'on utilise après l'espérance, cette grandeur sert d'estimateur statistique de l'erreur de l'approximation, et les lois présentées au début du chapitre implique la formule :

$$Var[f(x)] = \sigma^2(f) = E[(f(x) - E[f(x)])^2] = \int_a^b (f(x) - E[f(x)])^2 \rho(x)dx = E[f^2(x)] - E^2[f(x)]$$

Et par la méthode Monte Carlo, on peut transformer cette formule et l'approximer sous N tirages aléatoires qui dépend de la fonction de distribution aussi $\rho(x)$ s'écrit:

$$Var[f(x)] = \sigma^2[Y] \simeq \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2 = \overline{(Y_i - \bar{Y})^2} \quad \text{avec : } Y_i = f(X_i) \quad (1.24)$$

Dans la partie suivante, on va détailler les cas uniforme et non uniforme de la méthode.

1.4.2 types d'intégration Monte Carlo

Il existe plusieurs types d'intégration Monte Carlo [23], on va voir dans ce qui suit essentiellement trois types.

1.4.2.1 Intégration Monte Carlo avec échantillonnage uniforme:

Commençons par une intégrale qui représente un problème déterministe à une dimension définit par :

$$I_1 = \int_a^b f(x)dx$$

Et rappelons que la loi de l'espérance mathématique s'écrit sous la forme:

$$E[f(X)] = \int_a^b f(x)\rho(x)dx$$

Avec:

$$\rho(x) = \begin{cases} 1, & x \in [a, b] \\ 0, & \text{ailleurs.} \end{cases},$$

Ou-bien d'une façon générale et après la normalisation de la densité de probabilité :

$$\rho(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}, \quad \int_{-\infty}^{\infty} \rho(x)dx = 1$$

Donc :

$$E[f(x)] = \frac{1}{b-a} \int_a^b f(x)dx \quad \Rightarrow \quad \int_a^b f(x)dx = (b-a)E[f(x)]$$

$$I_1 = \int_a^b f(x)dx \simeq \frac{b-a}{N} \sum_{i=1}^N g(X_i) = (b-a)\bar{Y}, \quad Y = f(X)$$

La variance ou l'erreur statistique de I_1 est :

$$\begin{aligned} Var(I_1) &\simeq Var\left(\frac{b-a}{N} \sum_{i=1}^N f(X_i)\right) = \left(\frac{b-a}{N}\right)^2 \sum_{i=1}^N Var(f(X_i)) \\ &= \left(\frac{b-a}{N}\right)^2 NVar(f(x)) = (b-a)^2 \frac{Var(f(x))}{N} \end{aligned}$$

D'après l'équation 1.24

$$\sigma^2(I_1) \simeq \left(\frac{b-a}{N}\right)^2 \sum_{i=1}^N (Y_i - \bar{Y})^2 \tag{1.25}$$

Et l'écart-type de la formule d'intégration est donné en fonction de l'écart-type de la fonction par :

$$\boxed{\sigma(I_1) = (b - a) \frac{\sigma(Y)}{\sqrt{N}}} \quad (1.26)$$

Généralisation à n dimensions : On considère l'intégrale à n dimensions sur un domaine D :

$$I_n = \int_D f(\vec{r}) dv, \quad \vec{r} = (x_1, \dots, x_n)$$

Si la fonction de densité de probabilité $\rho(x)$ est distribuée uniformément sur le domaine cubique $D = ([a, b], \dots, [a, b])$ de volume $V = (b - a)^n$, on peut l'écrire comme :

$$\rho(\vec{r}) = \begin{cases} 1, & \vec{r} \in ([a, b], \dots, [a, b]) \\ 0, & \vec{r} \notin ([a, b], \dots, [a, b]) \end{cases},$$

Et après normalisation:

$$\rho(\vec{r}) = \begin{cases} \frac{1}{(b-a)^n}, & \vec{r} \in ([a, b], \dots, [a, b]) \\ 0, & \vec{r} \notin ([a, b], \dots, [a, b]) \end{cases},$$

Et toujours:

$$\int_D \rho(x_1, \dots, x_n) dx_1 \dots dx_n = 1$$

L'espérance est alors :

$$E[f(x)] \simeq \frac{1}{(b-a)^n} \int_D f(\vec{r}) dv \Rightarrow \int_D f(\vec{r}) dv = (b-a)^n E[f(\vec{r})]$$

Et l'intégrale d'une fonction f à n dimensions peut être approximée par :

$$\boxed{I_n = \int_D f(\vec{r}) dv \simeq \frac{(b-a)^n}{N} \sum_{i=1}^N f(\vec{r}_i) = V\bar{Y}, \quad Y = f(\vec{r})} \quad (1.27)$$

Sa variance donc est :

$$\begin{aligned} Var(I_n) &\simeq Var\left(\frac{(b-a)^n}{N} \sum_{i=1}^N f(\vec{r}_i)\right) = \left(\frac{(b-a)^n}{N}\right)^2 \sum_{i=1}^N Var(f(\vec{r}_i)) \\ &= \left(\frac{(b-a)^n}{N}\right)^2 N Var(f(\vec{r})) = V^2 \frac{Var(f)}{N} \\ \sigma^2(I_n) &\simeq \left[\frac{(b-a)^n}{N}\right]^2 \sum_{i=1}^N (Y_i - \bar{Y})^2 \end{aligned} \quad (1.28)$$

Et son l'écart-type est représenté par:

$$\boxed{\sigma(I_n) = V \frac{\sigma(Y)}{\sqrt{N}}} \quad (1.29)$$

1.4.2.2 Intégration Monte Carlo avec échantillonnage non-uniforme

La principale propriété dans ce type d'intégrale c'est la génération des nombres aléatoires qui soient distribués selon une fonction de densité de probabilité $\rho(x)$ non uniforme. L'approximation de l'intégrale Monte Carlo à une dimension s'écrit sous la forme suivante:

$$I_1 = \int_a^b f(x)\rho(x)dx \simeq \frac{1}{N} \sum_{i=1}^N f(X_i) = \bar{Y} \quad (1.30)$$

La variance de I_1 se calcul facilement :

$$\begin{aligned} Var(I_1) &\simeq Var\left(\frac{1}{N} \sum_{i=1}^N f(x_i)\right) = \left(\frac{1}{N}\right)^2 \sum_{i=1}^N Var(f(x_i)) = \left(\frac{1}{N}\right)^2 NVar(f(x)) = \frac{Var(f)}{N} \\ Var(I_1) &\simeq \left(\frac{1}{N}\right)^2 \sum_{i=1}^N (Y_i - \bar{Y})^2 \end{aligned} \quad (1.31)$$

Et l'écart type :

$$\sigma(I_1) = \frac{\sigma(Y)}{\sqrt{N}} \quad (1.32)$$

Généralisation à n dimensions : Dans ce cas l'intégrale sur un domaine D est par :

$$I_n = \int_D \rho(\vec{r})f(\vec{r})dv = E[f(x)]$$

Son approximation par la méthode de Monte Carlo est représenté par:

$$I_n = \int_D \rho(\vec{r})f(\vec{r})dv \simeq \frac{1}{N} \sum_{i=1}^N f(\vec{r}_i) = \bar{Y}, \quad Y = f(\vec{r}) \quad (1.33)$$

La variance de I_n est:

$$\begin{aligned} Var(I_n) &\simeq Var\left(\frac{1}{N} \sum_{i=1}^N f(\vec{r}_i)\right) = \left(\frac{1}{N}\right)^2 \sum_{i=1}^N Var(f(\vec{r}_i)) = \left(\frac{1}{N}\right)^2 NVar(f(\vec{r})) = \frac{Var(f)}{N} \\ \sigma^2(I_n) &\simeq \left[\frac{1}{N}\right]^2 \sum_{i=1}^N (Y_i - \bar{Y})^2 \end{aligned} \quad (1.34)$$

Se qui implique l'écart-type :

$$\sigma(I_n) = \frac{\sigma(Y)}{\sqrt{N}} \quad (1.35)$$

1.4.2.3 Intégration Monte Carlo avec échantillonnage préférentiel (réduction de la variance)

La méthode de Monte Carlo avec échantillonnage préférentiel est une technique utilisée dans le but de réduire l'erreur statistique du calcul de l'intégrale donnée et améliorer la précision, par la réduction de la variance.[1]

A une dimension, dans ce type d'intégrales, on doit chercher une fonction approximative $g(x)$ de forme proche de la fonction étudiée $f(x)$ ($f/g \simeq Cte$), cette nouvelle fonction est alors considéré comme la FDP (la fonction de densité de probabilité) tel que :

$$g(x) \simeq f(x), \quad \rho(x) = g(x), \quad \int_{-\infty}^{\infty} \rho(x)dx = 1$$

$$\int_a^b f(x)\rho(x)dx = \int_a^b \frac{f(x)}{g(x)}g(x)dx = \int_a^b \frac{f(x)}{g(x)}\rho(x)dx$$

On calcule l'espérance de la fonction rapport $\frac{f(x)}{g(x)}$:

$$E \left[\frac{f(x)}{g(x)} \right] = \int_a^b \frac{f(x)}{g(x)}\rho(x)dx = \int_a^b f(x)dx = I_1$$

Et en utilisant l'estimateur on obtient l'approximation Monte Carlo avec échantillonnage préférentielle :

$$I_1 = \int_a^b f(x)dx \simeq \frac{1}{N} \sum_{i=1}^N \frac{f(x_i)}{g(x_i)} = \bar{\mathcal{Y}}, \quad \mathcal{Y} = \frac{f(x)}{g(x)} \quad (1.36)$$

Et de même on trouve a la variance de cette quantité I_1

$$Var(I_1) \simeq Var \left(\frac{1}{N} \sum_{i=1}^N \frac{f(x_i)}{g(x_i)} \right) = \left(\frac{1}{N} \right)^2 \sum_{i=1}^N Var \left(\frac{f(x_i)}{g(x_i)} \right) = \left(\frac{1}{N} \right)^2 N Var \left(\frac{f(x)}{g(x)} \right) = \frac{1}{N} Var(\mathcal{Y})$$

$$\sigma^2(I_1) \simeq \frac{1}{N^2} \sum_{i=1}^N (\mathcal{Y}_i - \bar{\mathcal{Y}})^2, \quad \mathcal{Y}_i = \frac{f(X_i)}{g(X_i)} \quad (1.37)$$

Son écart-type est s'écrit :

$$\sigma(I_1) = \frac{1}{\sqrt{N}} \sigma(\mathcal{Y}) \quad (1.38)$$

Généralisation à n dimensions : De la même façon, on choisit une distribution $g(\vec{r})$ proche de $f(\vec{r})$ sur l'intervalle $D = [a, b]^n$, avec $f/g \simeq Cte$

$$\rho(\vec{r}) \simeq g(\vec{r})$$

Alors on obtient presque la même formule l'approximation :

$$I_n = \int_D f(\vec{r})dv \simeq \frac{1}{N} \sum_{i=1}^N \frac{f(\vec{r}_i)}{g(\vec{r}_i)} = \frac{1}{N} \sum_{i=1}^N \mathcal{Y}_i = \bar{\mathcal{Y}}, \quad \mathcal{Y}_i = \frac{f(\vec{r}_i)}{g(\vec{r}_i)} \quad (1.39)$$

Variance de I_n aussi s'écrit:

$$\begin{aligned} \text{Var}(I_n) &\simeq \text{Var}\left(\frac{1}{N} \sum_{i=1}^N \mathcal{Y}_i\right) = \left(\frac{1}{N}\right)^2 \sum_{i=1}^N \text{Var}(\mathcal{Y}_i) = \frac{1}{N} \text{Var}(\mathcal{Y}) \\ \sigma^2(I_n) &\simeq \frac{1}{N^2} \sum_{i=1}^N (\mathcal{Y}_i - \bar{\mathcal{Y}})^2 \end{aligned} \quad (1.40)$$

Ainsi que l'écart-type:

$$\boxed{\sigma(I_n) = \frac{1}{\sqrt{N}} \sigma(\mathcal{Y})} \quad (1.41)$$

Remarque:

On remarque que pour tous les types d'intégrations Monte Carlo l'écart-type de ces intégrales sont toujours proportionnelles à l'écart-type de la fonction étudiée, et inversement proportionnelles à la racine du nombre de tirage \sqrt{N} . Donc on peut dire que la variation de l'erreur de ces intégrales dépend de l'erreur de la fonction et de la taille de l'échantillon N . Et si on veut minimiser l'erreur du calcul on doit augmenter N .

1.5 Le problème de la dimensionalité

Dans ce mémoire on utilise l'outil de l'intégration Monte Carlo qui est considéré comme l'un des moyens ou des algorithmes non-déterministes ou stochastique. Malgré ce caractère aléatoire, cela ne nous empêche pas d'arriver à des résultats précis par minimisation des propriétés probabilistes du calcul tel que la variance. Ce type de méthodes à été proposé afin d'éviter un problème qui nous empêche de faire des calculs lors de l'utilisation des autres types de méthodes dites déterministes comme les méthodes de Newton-Cotes (trapèzes, Simpson ...etc) ou les méthodes plus précises de Gauss ou toute autre méthode classique déterministe. Il s'agit du fléau de la dimensionnalité.

Lorsque la dimension n ou le nombre de degré de libertés augmente au delà de quelques dimensions, les méthodes déterministes deviennent inadaptés. En effet, ces méthodes classiques se basent sur l'estimation de la fonction à intégrer en P points d'un intervalle donné sur un axe donné. Si on cherche une meilleure précision on doit augmenter le nombre de points d'intégrations P . Le problème apparaît alors lorsque le nombre d'axe d'intégration (ou dimension n) devient trop grand, le nombre de points d'intégrations total N augmente alors d'une façon exponentielle comme c'est illustré ci-dessous:

$$N = P^n = e^{n \ln(P)}$$

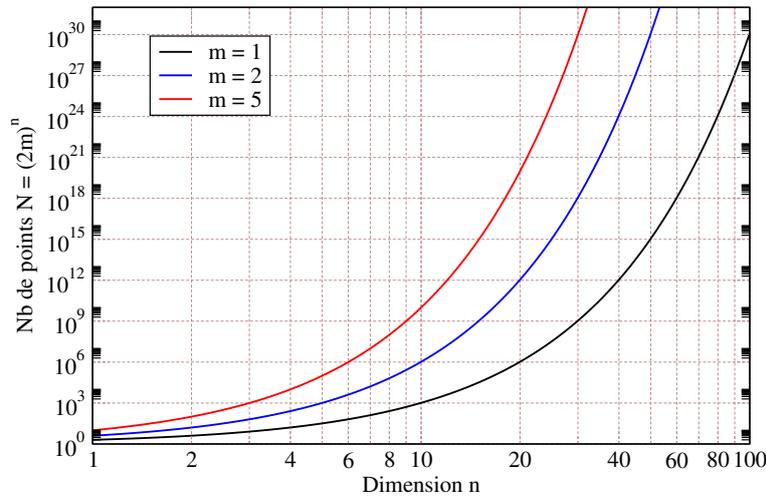


Figure 1.7: intégration de Gauss-Legendre à 2 nœuds : Nombre de points d'intégration en fonction de la dimension n pour trois subdivision des axes : $m = 1, 2$ et 5 (en échelle log) [24]

La figure 1.7 montre le problème (dit aussi malédiction) de la dimensionalité avec la méthode déterministe de Gauss-Legendre à deux nœuds. Le calcul devient très difficile voir même impossible à partir de quelque dizaines de dimensions (exemple : Gauss à 20 dimensions avec 10 points P (5 subdivisions), $n = 20 \rightarrow N = 10^{20}$), c'est pour ça qu'il est préférable d'utiliser d'autres méthodes de calcul d'intégrale de très haute dimension comme les méthodes de Monte Carlo pour une résolution rapide et plus précise.

1.6 Intégration Monte Carlo : Étude comparative

Dans cette partie nous allons présenter une brève étude comparative entre les trois méthodes d'intégrations Monte Carlo : intégration MC avec échantillonnage uniforme, non-uniforme(avec densité), et préférentiel, avec un exemple de calcul d'une intégrale multidimensionnelle de 6 dimensions de la fonction à 6 variables $\vec{r} = (x_1, x_2 \dots, x_6)$ suivante:

$$f(\vec{r}) = (1 - x_1^2)(1 - x_2^2) \dots (1 - x_6^2)e^{-(x_1^2 + \dots + x_6^2)} = \prod_{i=1}^6 (1 - x_i^2)e^{-x_i^2} \quad (1.42)$$

Ce choix particulier d'une fonction factorisable nous permet de calculer la valeur exacte de l'intégrale de cette fonction pour évaluer l'erreur de chaque méthode. Ce calcul a été réalisé à l'aide du logiciel de calcul mathématique Maple.

Maple est un logiciel de calcul formel développé dans les années 1980, édité par la société canadienne Maplesoft, écrit en langage C et java, il fait partie de la catégorie sciences et éducation, il représente des bonnes capacités et une puissance qui permettent de développer et de déployer facilement des applications mathématiques diverses tel que des générateurs de nombres aléatoires variés et robustes.

1.6.1 Application

Le but de cette application est de calculer numériquement la valeur de l'intégrale multiple suivante dans le domaine à 6 dimensions : $D = [0, 1]^6$:

$$I_6 = \int_D f(\vec{r}) dv = \int_0^1 \cdots \int_0^1 (1 - x_1^2) \cdots (1 - x_6^2) e^{-(x_1^2 + \cdots + x_6^2)} dx_1 \cdots dx_6 \quad (1.43)$$

$$I_6 = \int_D \prod_{i=1}^6 (1 - x_i^2) e^{-x_i^2} dv \quad (1.44)$$

Puis la comparer avec sa valeur exacte donnée par :

$$I_6 = \left[\int_0^1 (1 - x^2) e^{-x^2} dx \right]^6 = \left[\frac{e\sqrt{\pi} \operatorname{erf}(1) + 2}{4e} \right]^6 \simeq (0.557352)^6 = 0.299761869 \quad (1.45)$$

A l'aide du logiciel Maple, on a réalisé la simulation numérique de l'intégrale I_6 par les trois méthodes d'intégration Monte Carlo (Uniforme, Avec densité, et avec échantillon préférentiel), pour un grands nombre de points d'intégration N et pour la dimension $n = 6$.

Formules pour Monte Carlo uniforme: La FDP (fonction densite de probabilité) dans ce cas est une constante à normalisée avec la condition:

$$\int_D \rho(x) dv = 1$$

Alors l'intégrale I_6 devient d'après l'équation 1.27 :

$$I_6 = \int_D f(\vec{r}) dv \simeq \frac{(b-a)^6}{N} \sum_{i=1}^N f(\vec{r}_i) = IMC \quad (1.46)$$

Formules pour Monte Carlo non-uniforme (avec densité): On choisi la fonction de densité de probabilité suivante :

$$\rho(\vec{r}) = e^{-(x_1^2 + \cdots + x_6^2)} = \prod_{i=1}^6 e^{-x_i^2} \quad (1.47)$$

l'intégrale est alors donné selon la formule 1.33 par:

$$I_6 = \int_D F(\vec{r}) \rho(\vec{r}) dv \simeq \frac{1}{N} \sum_{i=1}^N F(\vec{r}_i) = IMC, \quad F(\vec{r}) = \prod_{i=1}^6 (1 - x_i^2) \quad (1.48)$$

Formules pour Monte Carlo avec échantillon préférentiel: On choisie une fonction $g(\vec{r})$ proche de la fonction $f(\vec{r})$, tel que cette nouvelle distribution soit la FDP $\rho(\vec{r}) = g(\vec{r})$:

Alors:

$$I_6 = \int_D f(\vec{r}) dv = \int_D \frac{f(\vec{r})}{g(\vec{r})} \rho(\vec{r}) dv$$

Avec: $g(\vec{r}) \simeq Cte f(\vec{r})$. On prend une fonction factorisable $g(\vec{r}) = \prod_i g(x_i)$. tel que $g(x)$ est approximé par son développement de Taylor de la fonction $f(x)$. On a une dimension et à l'ordre 2 :

$$e^{-x^2} \simeq (1 - x^2) \quad (1.49)$$

alors la fonction $g(x)$ est définie alors par :

$$g(x) = (1 - x^2)e^{-x^2} \simeq (1 - x^2)^2 \quad (1.50)$$

A 6 dimensions la fonction sera écrite de la forme :

$$g(\vec{r}) \simeq \prod_{i=1}^6 (1 - x_i^2)^2 \quad (1.51)$$

$$f(\vec{r}) = \prod_{i=1}^6 (1 - x_i^2) e^{-x_i^2} = \prod_{i=1}^6 \frac{(1 - x_i^2) e^{-x_i^2}}{(1 - x_i^2)^2} (1 - x_i^2)^2 = \frac{\prod_{i=1}^6 (1 - x_i^2) e^{-x_i^2}}{\prod_{i=1}^6 (1 - x_i^2)^2} \prod_{i=1}^6 (1 - x_i^2)^2 \quad (1.52)$$

$$f(\vec{r}) = \frac{f(\vec{r})}{g(\vec{r})} \rho(\vec{r}), \quad \frac{f(\vec{r})}{g(\vec{r})} = \prod_{i=1}^6 \frac{e^{-x_i^2}}{(1 - x_i^2)}, \quad \rho(\vec{r}) = \prod_{i=1}^6 (1 - x_i^2)^2 \quad (1.53)$$

Donc la formule 1.39 implique :

$$I_6 \simeq \frac{1}{N} \sum_{i=1}^N \frac{f(\vec{r}_i)}{g(\vec{r}_i)} = IMC \quad (1.54)$$

Les figures suivantes sont des représentations graphiques des fonctions $f(\vec{r})$ et $g(\vec{r})$ à une et à deux dimensions ou elles montrent qu'elles sont proches en forme l'une de l'autre.

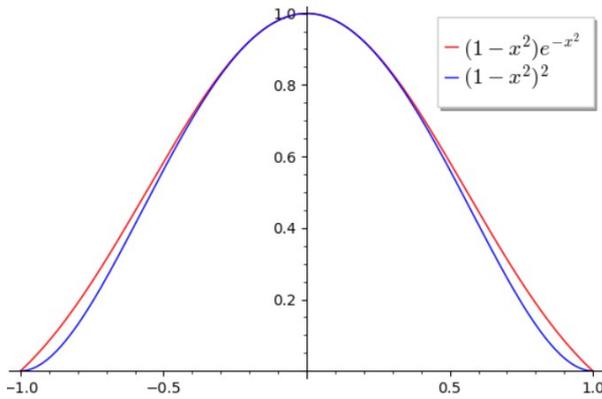


Figure 1.8: Représentation graphique de la fonction $f(\vec{r}), g(\vec{r})$ à $n = 1$ dimension

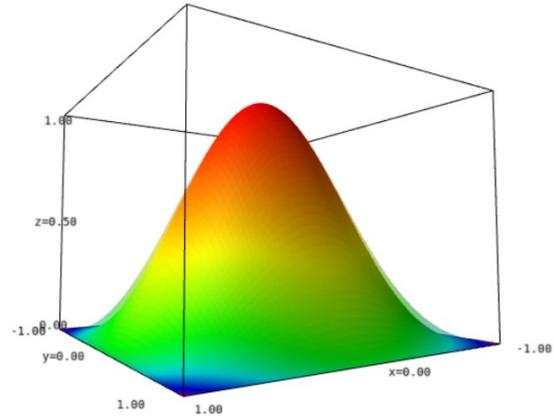


Figure 1.9: Représentation graphique de la fonction $f(\vec{r}), g(\vec{r})$ à $n = 2$ dimensions

A l'aide donc du logiciel Maple on a généré notre échantillon de N points par les trois méthodes MC selon la distribution de 6 dimensions afin d'approximer l'intégrale 1.44.

1.6.2 Présentation des résultats et Discussion

Après qu'on a introduit tous les paramètres et les formules au logiciel, et après l'affectation du calcul par Maple, les résultats sont représentés dans les tableaux et courbes suivantes:

N	10^2	10^3	10^4	10^5	10^6	10^7
IMC unif.	0.034435517	0.029657089	0.029431533	0.030035953	0.029880418	0.030004011
IMC non-unif.	0.0290222952	0.0296226087	0.0302004017	0.0301129021	0.0299455920	0.0299834153
IMC avec Ech Préf.	0.0328639013	0.0292828500	0.0299941122	0.0299244208	0.0299636132	0.02997793

Table 1.1: Valeurs de l'intégrale IMC en fonction du nombre de tirages N pour les trois méthodes

N	10^2	10^3	10^4	10^5	10^6	10^7
ErrEx unif.	14.87624244	1.064502847	1.816955078	0.1993801131	0.319481560	0.0928228555
ErrEx non-unif.	3.182164638	1.179530029	0.7479764777	0.4560794297	0.102063948	0.0241138534
ErrEx avec Ech Préf.	9.633361647	2.312958881	0.05979871136	0.1726905666	0.0419453677	0.0058449633

Table 1.2: Valeurs de l'erreur exacte relative (%) en fonction du nombre de tirages N pour les trois méthodes MC

N	10^2	10^3	10^4	10^5	10^6	10^7
Sig/I unif.	20.88543724	6.937013357	2.224223016	0.7123646264	0.2249831398	0.07125296618
Sig/I non-unif.	9.944969762	3.223654116	1.018272928	0.3199278559	0.1017279885	0.03214021965
Sig/I avec Ech Préf.	5.819418761	1.081774056	0.3838737025	0.1262780443	0.04057438769	0.01340225092

Table 1.3: Valeurs de l'écart-type relatif (%) en fonction du nombre de tirages N pour les trois méthodes MC

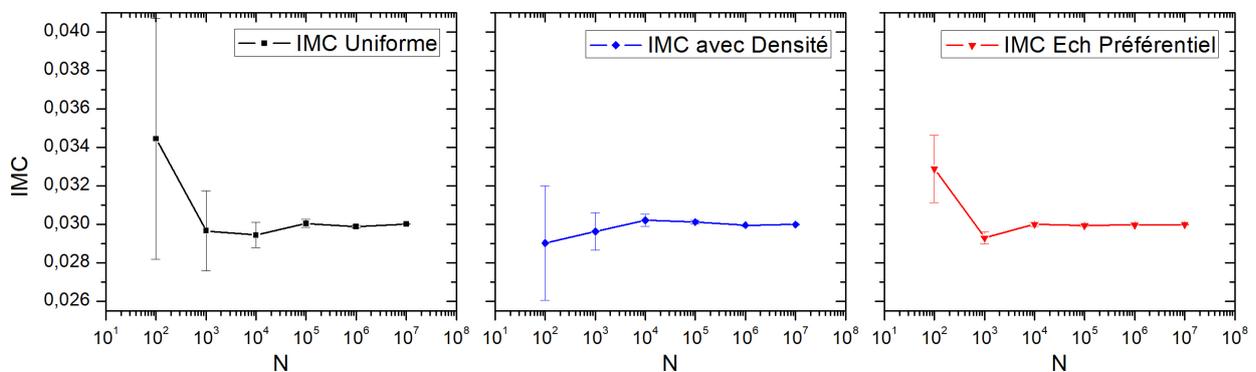


Figure 1.10: Valeur de l'intégrale IMC par les trois méthodes d'intégrations Monte Carlo avec barres d'erreurs en fonction du nombre de points N

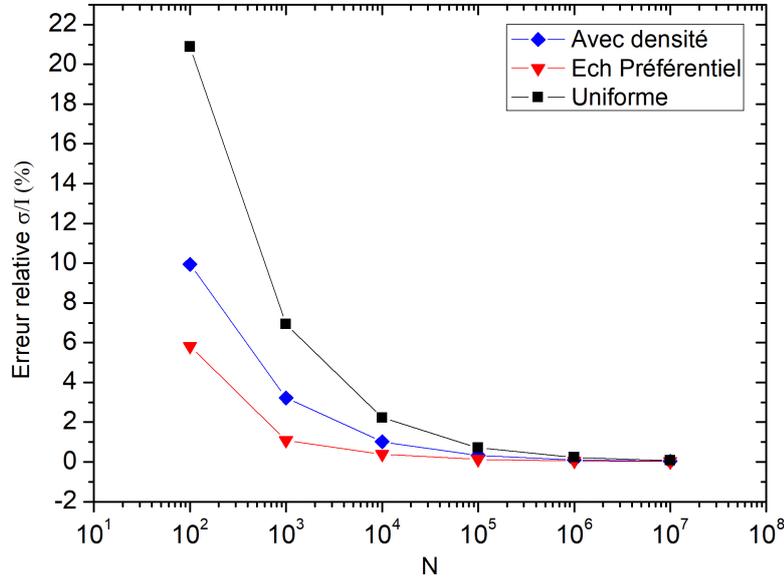


Figure 1.11: L'erreur statistique relative pour les trois méthodes d'intégration Monte Carlo en fonction du nombre de points N

Dans les trois tableaux on a présenté, pour chacune des trois méthodes d'intégration MC : la valeur de l'intégrale "IMC", l'erreur relative exacte "ErrEr" et l'erreur statistique relative σ/I . D'après ces tableaux on remarque que pour les trois méthodes, la valeur de l'intégrale Monte Carlo converge vers la valeur exacte, mais la différence se trouve dans la vitesse de convergence qui apparaît plus clairement sur les courbes avec barres d'erreurs 1.10. En effet, on remarque que la méthode qui contient des barres plus grandes est la méthode M.C avec échantillon uniforme, puis la méthode non-uniforme, et la méthode avec les plus petites barres d'erreurs est celle avec échantillonnage préférentiel.

La même chose est illustrée dans l'autre figure 1.11 qui représente l'erreur statistique relative, où l'on voit l'avantage de grandes valeurs sont associées à la technique avec échantillonnage préférentiel par rapport à celle avec densité (non-uniforme) et uniforme.

Ainsi, avec une erreur plus faible pour un même nombre de points d'intégration N , la meilleure méthode est celle avec échantillonnage préférentiel qui permet de réduire la variance suivie par la technique non-uniforme (avec densité adéquate) et en dernier celle avec échantillon uniforme. Cependant, il est parfois difficile de trouver une fonction adéquate pour un échantillonnage préférentiel notamment à haute dimension lorsqu'il n'est pas simple de deviner la forme de la fonction à intégrer, on est obligé dans ce cas de travailler avec la méthode non-uniforme (lorsqu'il est possible de trouver une bonne FDP) ou alors utiliser la méthode uniforme.

Chapitre 2

Méthode variationnelle et descente du gradient

2.1 Introduction

Dans ce chapitre nous allons présenter deux méthodes courantes et très importantes en mécanique quantique, la première est une technique fréquemment utilisée dans la résolution des problèmes qui ne possèdent pas de solutions analytiques exactes c'est la méthode variationnelle, elle nous permet de résoudre certains problèmes physiques pour lesquels l'équation de Schrodinger ne possède pas de solution exacte comme les systèmes atomiques et les problèmes à n corps.

Après la modélisation des phénomènes physiques et afin de résoudre les équations de ces modèles, on utilise la simulation sur ordinateur. Dans le cas de la méthode des variations, on a besoin de réaliser des calculs d'optimisation de l'énergie, c'est-à-dire la minimisation de la fonction qui représente l'énergie totale du système. L'optimisation est la base de la deuxième méthode qu'on va voir dans notre étude. Celle-ci, appelée méthode de la descente du gradient, connu aussi sous le nom d'algorithme de la plus forte pente ou de descente la plus rapide (en anglais "steepest descent"), est une méthode itérative qui permet de trouver un minimum local ou global de la fonction choisie.

Le but de la méthode des variations est de calculer d'une façon approximative la valeur d'énergie de l'état fondamentale (la plus basse énergie) ainsi que de certains états excités, en partant d'une fonction d'onde dite d'essai dépendante d'un ou plusieurs paramètres que l'on cherche à optimiser pour minimiser la valeur moyenne de l'énergie. Parmi les méthodes qui reposent sur le principe variationnel est la méthode Hartree-Fock.

Par la combinaison de ces deux méthodes, la méthode des variations et la méthode de la descente du gradient, on peut atteindre un objectif commun qui est la minimisation de l'énergie. C'est ce que l'on verra dans le troisième chapitre avec leur utilisation de manière conjointe au sein d'un algorithme spécifique pour le calcul de l'état fondamental de l'atome d'Hélium.

2.2 Méthode variationnelle en mécanique quantique

2.2.1 Principes

La méthode variationnelle nous permet de calculer de façon approximative les niveaux d'énergie d'un système quantique d'un système stationnaire, notamment le niveau fondamental E_0 , en optimisant l'intégrale représentant l'énergie moyenne du système et qui dépend d'un certain nombre de paramètres variationnels α par le biais de la fonction d'onde $\psi(\vec{r}, \alpha)$ [7].

Considérons un hamiltonien H indépendant du temps, de spectre discret (E_n) et non entièrement dégénéré. On a:

$$H|\varphi_n\rangle = E_n|\varphi_n\rangle \quad (2.1)$$

tel que:

$$E_0 \leq E_1 \leq E_2 \leq \dots \leq E_n \quad (2.2)$$

Tout état $|\psi\rangle$ peut se développer sur la base de ces vecteurs propres $\{|\varphi_n\rangle\}$ formant une base orthonormée associée aux énergies propres $\{E_n\}$ sous la forme suivante:

$$|\psi\rangle = \sum_n a_n |\varphi_n\rangle \quad (2.3)$$

La valeur moyenne de l'énergie du système $\langle H \rangle$ dans l'état $|\psi\rangle$ s'écrit comme suit :

$$\langle E \rangle = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\sum_n |a_n|^2 E_n}{\sum_n |a_n|^2} \quad (2.4)$$

En vertu de l'hypothèse 2.2 il est aisé de démontrer la formule suivante qui représente l'équation maîtresse de la théorie des variations :

$$\langle E \rangle = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \geq E_0 \quad (2.5)$$

En effet, on calcule $\langle \psi | \psi \rangle = \sum_n (c_n)^2$, puis la 2ème partie $\langle \psi | \psi \rangle E_0$, et parce qu'on sais que la différence entre ces deux quantités est positive, alors la formule 2.5 se réalise fortement.

L'équation maîtresse de la théorie des variations stipule que la valeur moyenne de l'hamiltonien $\langle H \rangle(\alpha)$ pour tout état $|\psi(\alpha)\rangle$ constitue une valeur approchée par excès de l'énergie du niveau fondamental. La minimisation de $\langle H \rangle(\alpha)$ donne des équations qui permettent donc de calculer les paramètres optimaux α_m et l'énergie minimale.

2.2.2 Théorème de Rayleigh-Ritz

Ce théorème permet de généraliser la méthode quelque soit le type de spectre, il donne une propriété remarquable des états propres d'un système ainsi que la valeur moyenne de l'Hamiltonien.

Le théorème stipule que la valeur moyenne de l'hamiltonien H d'un système est stationnaire au voisinage de ses états propres et prend alors comme valeur, la valeur propre de l'état en question.

D'une autre façon, l'équation 2.5 qui est une fonctionnelle de la fonction d'onde $|\psi\rangle$ passe par un minimum local, lorsque $|\psi\rangle$ est l'une des fonctions propres de H . Cette valeur extrême est alors égale à l'énergie propre correspondante à la fonction propre $|\psi\rangle$.

Démonstration: On considère l'état $|\psi\rangle$, telle que:

$$\langle H \rangle = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \quad (2.6)$$

Passé par un minimum locale. On cherche maintenant à démontrer que cet état est un état propre du système, et on considère le ket suivant dépendant du petit paramètre ε et

qui constitue une petite variation autour de $|\psi\rangle$, ce nouveau état est défini par:

$$|\psi_\varepsilon\rangle = |\psi\rangle + \varepsilon|\varphi\rangle \quad (2.7)$$

Avec $|\varphi\rangle$ représente un ket quelconque dans l'espace des états. La valeur moyenne de l'hamiltonien dans ce cas s'écrit en fonction de ce nouveau paramètre comme:

$$\langle H_\varepsilon \rangle = \frac{\langle \psi_\varepsilon | H | \psi_\varepsilon \rangle}{\langle \psi_\varepsilon | \psi_\varepsilon \rangle} \quad (2.8)$$

Cette dernière doit être extrémale pour $\varepsilon = 0$ avec les hypothèses suivantes:

$|\psi_{\varepsilon=0}\rangle = |\psi\rangle$, et $\langle H_{\varepsilon=0} \rangle = \langle H \rangle$, ce qui veut dire que:

$$\frac{d\langle H_\varepsilon \rangle}{d\varepsilon} = 0 \quad \text{pour } \varepsilon = 0 \quad (2.9)$$

On simplifie l'équation 2.8, et on l'écrit comme:

$$\langle \psi_\varepsilon | \psi_\varepsilon \rangle \langle H_\varepsilon \rangle = \langle \psi_\varepsilon | H | \psi_\varepsilon \rangle \quad (2.10)$$

En dérivant :

$$\left\langle \frac{d\psi_\varepsilon}{d\varepsilon} | \psi_\varepsilon \right\rangle \langle H_\varepsilon \rangle + \langle \psi_\varepsilon | \frac{d\psi_\varepsilon}{d\varepsilon} \rangle \langle H_\varepsilon \rangle + \langle \psi_\varepsilon | \psi_\varepsilon \rangle \frac{d\langle H_\varepsilon \rangle}{d\varepsilon} = \left\langle \frac{d\psi_\varepsilon}{d\varepsilon} | H | \psi_\varepsilon \right\rangle + \langle \psi_\varepsilon | H | \frac{d\psi_\varepsilon}{d\varepsilon} \rangle \quad (2.11)$$

Pour $\varepsilon = 0$, on a : $\frac{d\psi_\varepsilon}{d\varepsilon} = \varphi$, et l'équation 2.12 s'écrit en fonction du nouveau état $|\varphi\rangle$ comme:

$$\langle \varphi | \psi \rangle \langle H \rangle + \langle \psi | \varphi \rangle \langle H \rangle = \langle \varphi | H | \psi \rangle + \langle \psi | H | \varphi \rangle \quad (2.12)$$

Alors:

$$\langle \varphi | H - \langle H \rangle | \psi \rangle + \langle \psi | H - \langle H \rangle | \varphi \rangle = 0$$

Cette équation étant vérifiée pour tout ket $|\varphi\rangle$, elle l'est pour $|\varphi\rangle = (H - \langle H \rangle)|\psi\rangle$, ce qui implique alors que:

$$\langle \varphi | \varphi \rangle = 0 \quad \text{Donc : } |\varphi\rangle = 0$$

Alors on revient à la formule initiale du $|\varphi\rangle$:

$$|\varphi\rangle = (H - \langle H \rangle)|\psi\rangle = 0$$

Ce qui implique:

$$H|\psi\rangle = \langle H \rangle |\psi\rangle$$

En conséquence, la valeur moyenne $\langle H \rangle$ est stationnaire si et seulement si le vecteur $|\psi\rangle$ est vecteur propre de l'hamiltonien H avec la valeur propre $E = \langle H_m \rangle$ qui est la valeur stationnaire (minimale) de H.

2.2.3 A propos de l'optimisation

L'optimisation est une branche des mathématiques qui sert à trouver des solutions à des problèmes dans divers domaines tel que la physique, l'ingénierie, l'automatique, l'économie, et l'analyse numérique, en cherchant à maximiser ou à minimiser certains facteurs comme la maximisation des profits, la minimisation des coûts d'un produit ou encore la maximisation de la production en économie et la minimisation de l'énergie d'un système en mécanique ...etc. L'utilisation de ce processus d'optimisation nécessite généralement trois étapes:

1. La formulation du problème.
2. La modélisation mathématique du problème.
3. Le choix d'une méthode de résolution du problème modélisé.

On distingue deux grands types d'optimisations qui sont:

- Techniques d'optimisation sans contraintes.
- Techniques d'optimisation sous contraintes.

Pour le choix de la méthode, il existe de nombreuses techniques pour résoudre les problèmes d'optimisation selon la nature du problème et les hypothèses posées. Dans l'optimisation avec contraintes on donne la méthode du gradient projeté. Et comme exemples sur l'autre type d'optimisation sans contraintes, le type que nous intéresse, on note par exemples les méthodes suivantes:

- Le développement en série de Taylor
- Les méthodes de descente.
- Méthode de Newton et quasi-Newton
- Les Méthodes du gradient conjugué

2.2.4 Difficultés de l'optimisation (minimisation)

L'utilisation de la méthode des variations conduit à minimiser la valeur moyenne de l'énergie, en choisissant une fonction d'onde d'essai qui dépend d'un ou plusieurs paramètres variationnels α de telle sorte l'équation de minimisation est s'écrit pour chaque paramètre α :

$$\frac{\partial}{\partial \alpha} \langle H \rangle = 0$$

On obtient donc un système d'équations explicites ou implicites dont la résolution n'est pas toujours triviale, et on peut rencontrer un certain nombre de difficultés, parmi ces difficultés on note:

1. Les équations sont non-linéaires et difficile à résoudre.

2. Les équations ne sont pas explicites: On peut trouver des équations intégrales inclus par exemple dans des équations de dérivation dont la séparation n'est pas possible et bien d'autres types d'équations compliquées.

Dans ce cas on utilise des méthodes d'optimisation par itération qui permettent de s'affranchir de ces difficultés comme la méthode de la descente du gradient.

2.3 Méthode de la descente du gradient

2.3.1 Le gradient

Le gradient d'une fonction $f(\vec{r})$ à plusieurs variables $\vec{r} = (x, y, z, \dots)$ est une grandeur vectorielle définie mathématiquement comme le taux de variation de la fonction en un point donné, il représente la généralisation de la notion de dérivée. Il est désigné à l'aide de l'opérateur nabla $\vec{\nabla}$ et sa direction en un point donné définit la direction de la variation (accroissement) la plus forte de f [27].

En effet, sachant que la variation d'une fonction f à plusieurs variables est donnée par :

$$df = \vec{\nabla} f \cdot d\vec{r} = |\vec{\nabla} f| |d\vec{r}| \cos \theta \quad (2.13)$$

où df est la variation infinitésimale totale de f , pour un déplacement infinitésimal $d\vec{r}$. Il est évident que df est maximale lorsque $d\vec{r}$ et le gradient $\vec{\nabla} f$ sont parallèles et dans la même direction c'est-à-dire lorsque $\theta = 0$ [28].

En coordonnées cartésiennes le gradient s'écrit :

$$\vec{\nabla} f(x, y, z) = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} + \frac{\partial f}{\partial z} \vec{k} \quad (2.14)$$

Tel que: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ représentent les dérivées partielles de f par rapport à x, y et z .

2.3.1.1 Les courbes de niveaux (équipotentiels)

Une courbe ou surface de niveau, appelée aussi équipotentielle, d'une fonction à plusieurs variables $f(x, y, z, \dots)$ est l'ensemble des points où la fonction prend une même valeur numérique constante [35]. Il s'agit de courbes dans le cas de 2 variables et de surfaces dans le cas de 3 variables. Une propriété remarquable des courbes de niveau en relation avec le vecteur gradient est le fait que ce dernier est toujours normal (perpendiculaire) à ces courbes en tout point. Cette propriété découle elle aussi de la relation de différentielle totale 2.13. En effet, sur une courbe de niveau la fonction reste constante et sa variation est donc nulle, il s'en suit que $\vec{\nabla} f \perp d\vec{r}$, or $d\vec{r}$ étant tangentiel à la courbe de niveau d'où la propriété citée.

Un exemple pratique de l'utilisation des courbes de niveau est dans le domaine de la cartographie géographique, où ces courbes représentent les points de même altitude et les

endroits de grande pente sont ceux où les lignes de niveau sont très rapprochées.

Le vecteur gradient est donc caractérisée par les propriétés suivantes:

- Si le gradient d'une fonction est nulle en un points alors ce dernier est un point stationnaire (minimum ou maximum local) de la fonction.
- Le gradient est un vecteur de même dimension que l'espace.
- Le gradient pointe vers la direction ou la fonction croît le plus rapidement.
- La norme ou l'amplitude du gradient est égale au taux de croissance dans cette direction.

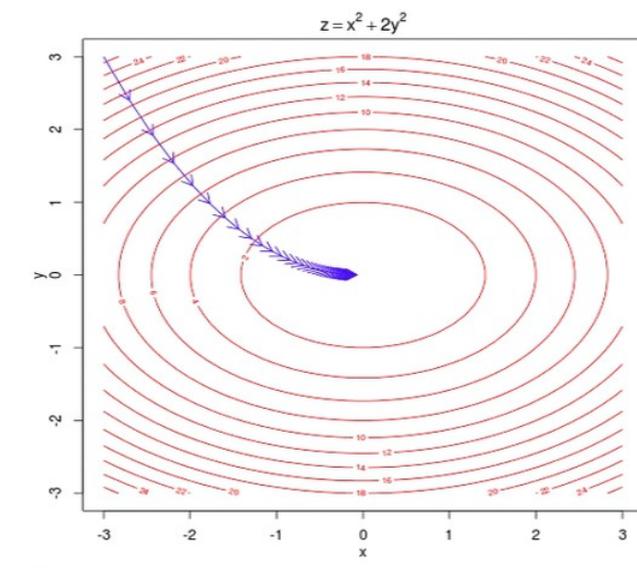


Figure 2.1: Exemple de courbes de niveaux de la fonction à 2 variables $f(x, y) = x^2 + 2y^2$ et son gradient [33].

Le gradient est une grandeur centrale qui trouve son utilisation dans plusieurs domaine notamment dans la théorie des variations, dans la résolution des équations différentiels, comme il joue un rôle important aux domaine de l'optimisation, ou il est utiliser dans la minimisation des fonctions par la méthode du descente du gradient.

2.3.2 Introduction à la méthode de la descente du gradient

L'algorithme de la descente du gradient, connu aussi sous le nom d'algorithme de la plus forte pente ou de la plus forte descente, est un processus itératif, parmi les nombreux moyens utilisés en optimisation, dont le but est d'utiliser les propriétés du gradient afin de trouver le minimum (ou maximum) d'une fonction (à condition qu'elle soit différentiable).

2.3.2.1 Principe

Comme on a vu précédemment, l'une des propriétés du vecteur gradient est qu'il pointe toujours dans la direction de croissance (ou décroissance) la plus rapide d'une fonction. Ainsi en suivant ce vecteur (ou son opposé) sur des points successifs perpendiculairement aux courbes de niveau, celui-ci nous mènera inévitablement vers un maximum (ou un minimum) de la fonction. Le principe de cet algorithme remonte au moins à Cauchy (1847) [32]. Il s'agit donc d'un procédé d'optimisation itératif qui cherche à trouver les points stationnaires suivant une direction parallèle à celle du gradient de la fonction. Les figures 2.2 et 2.3 illustre la méthode.

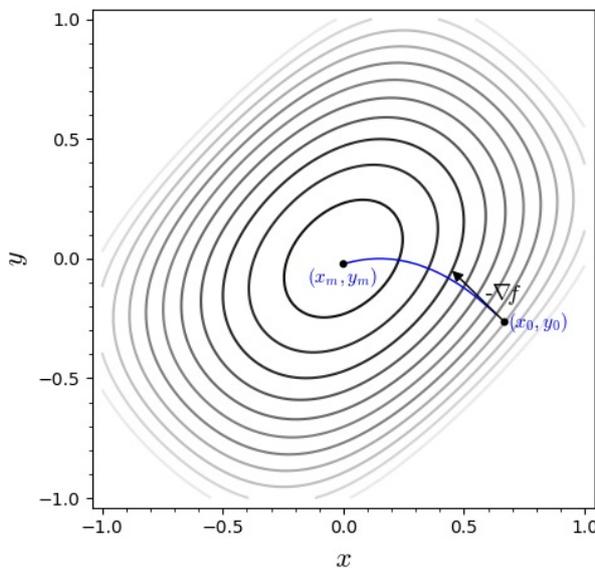
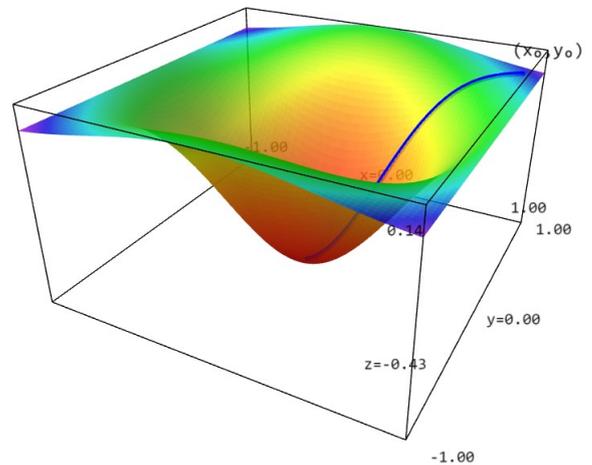
Figure 2.2: Courbe-niveaux $F_X(x)$.

Figure 2.3: Courbe du descente du gradient à 2D.

Dans le cas de recherche de minimum, la procédé consiste à chercher une suite de points $\vec{r}_k \in \mathbb{R}^n$ tel que : $f(\vec{r}_{k+1}) < f(\vec{r}_k)$ et le point \vec{r}_{k+1} étant calculé à partir du point \vec{r}_k en se déplaçant d'un petit pas δ_k dans une direction \vec{d}_k opposée au gradient. La méthode est donc donnée par la relation de récurrence suivante:

$$\vec{r}_{k+1} = \vec{r}_k + \delta_k \vec{d}_k$$

Si on choisi $\vec{d}_k = -\vec{\nabla} f|_{\vec{r}_k}$, on obtient la méthode du gradient, et la formule de récurrence sera écrite comme:

$$\vec{r}_{k+1} = \vec{r}_k - \delta_k \vec{\nabla} f|_{\vec{r}_k} \quad (2.15)$$

Le signe "-" est donc responsable de la minimisation, on sais en effet que si l'on veut maximiser une fonction, il suffira de minimiser son opposée en choisissant un pas approprié pour arriver aux plus petites valeurs de la fonction vérifiant la condition d'optimisation $\vec{\nabla} f = \vec{0}$ le plus vite que possible.

Le pas δ_k est un paramètre qui permet de moduler la correction, il ne faut pas qu'il soit trop petit au risque d'augmenter la lenteur de la convergence et pas trop grand au risque d'augmenter les oscillations.

2.3.3 Algorithme de la méthode

Pour appliquer la méthode de descente du gradient, l'algorithme passe par les étapes suivantes [31]:

1. Choix du premier points \vec{r}_0 .
2. Choix d'un pas optimal $\delta_k > 0$ le long de la direction $\vec{d}_k = -\vec{\nabla} f|_{\vec{r}_k}$ d'une façon à faire décroître f suffisamment.
3. Application de la formule 2.16:

$$\vec{r}_{k+1} = \vec{r}_k - \delta_k \vec{\nabla} f|_{\vec{r}_k} \quad (2.16)$$

4. On remplace k par $k + 1$ pour aller à l'étape suivante, jusqu'à l'arrivée à la convergence.
5. Teste d'arrêt: Si $\|\vec{\nabla} f\| \simeq 0$, implique l'arrêt de l'algorithme.

2.3.3.1 Défauts et points faibles de la méthode:

La méthode de la descente du gradient est utile à la résolution de systèmes d'équations de minimisation non explicites, mais elle rencontre quelques difficultés et points faibles comme celui de la convergence, et on note les problèmes suivants [34]:

- La recherche d'un pas δ_k optimal peut prendre un long temps.
- L'algorithme peut nécessiter de nombreuses itérations pour converger vers un minimum local, surtout si la courbe de la fonction est très différente dans des directions différentes.

Chapitre 3

Application à l'atome d'hélium

3.1 Introduction

Les méthodes de Monte Carlo ont trouvées une large gamme d'utilisation dans de nombreux domaines, ils ont fait une révolution dans la résolution des problèmes de très haute dimensionalité comme les problèmes à n cors, notamment dans des applications de calculs de certaines propriétés physiques, ou plus précisément dans le calcul des caractéristiques liées aux systèmes quantiques, statistiques et nucléaires. On remarque que ces techniques ont prouvées leur efficacités et leurs puissances à la résolution de ces problèmes surtout après le grand développement de domaines informatique et intelligence artificielle, dont les super ordinateurs avec les différents logiciels et algorithmes rendant l'utilisation des méthodes Monte Carlo et d'autres méthodes comme les méthodes d'optimisation simples, rapides et efficaces.

Les méthodes Monte Carlo sont connus pour leurs utilisation dans des systèmes à grand nombre de degré de liberté. Dans ce dernier chapitre nous nous intéressons à démontrer l'utilité de ces méthodes dans les systèmes atomiques à plusieurs électrons (problèmes à n corps), ou nous allons calculer par la méthode variationnelle une propriété principale liées à l'atome d'hélium le deuxième élément du tableau périodique et le premier exemple du problème qui n'a pas de solution précise en physique quantique, c'est l'énergie de l'état fondamental. Notre calcul variationnel utilise la méthode de la descente du gradient pour la minimisation de l'énergie et l'intégration Monte Carlo pour l'estimation des intégrales d'échanges.

Des études du même type ont été réalisés dans des mémoires précédents [24] mais en utilisant des fonctions d'ondes à un seul paramètre ce qui impacte la flexibilité du calcul variationnel et sa convergence. Pour améliorer cela, on a besoin de proposer une fonction d'onde de test approximative à deux paramètres, celle-ci sera obtenue à partir d'un modèle simple à électrons indépendants. Notons que notre calcul ne tient pas compte de la structure fine de l'atome d'hélium liée aux interactions des moments cinétiques des électrons ou aux effets relativistes.

3.2 Rappel théorique

L'hélium He est un gaz rare et inerte, il est classer comme le deuxième élément du tableau périodique et il contient deux électrons liés à un noyau de deux protons et deux neutrons avec une charge positive $Z = +2e$. Il forme environ 25 % de toute la matière dont la plupart qui se trouve dans le milieu interstellaire dans un état de plasma. L'atome d'hélium est caractérisée par une masse atomique égale à $4,002602 \pm 0,000002 u.a.$ [29], un rayon atomique de $128 pm$. L'énergie d'ionisation pour chaque électrons vaux pour le 1er: $24,587387 eV$, et pour le 2ème: $54,417760 eV$ [30]. Et la valeur de l'énergie de l'état fondamental est:

$$E_0 = -79,02 eV = -2,9040 \text{ Hartree} \quad (3.1)$$

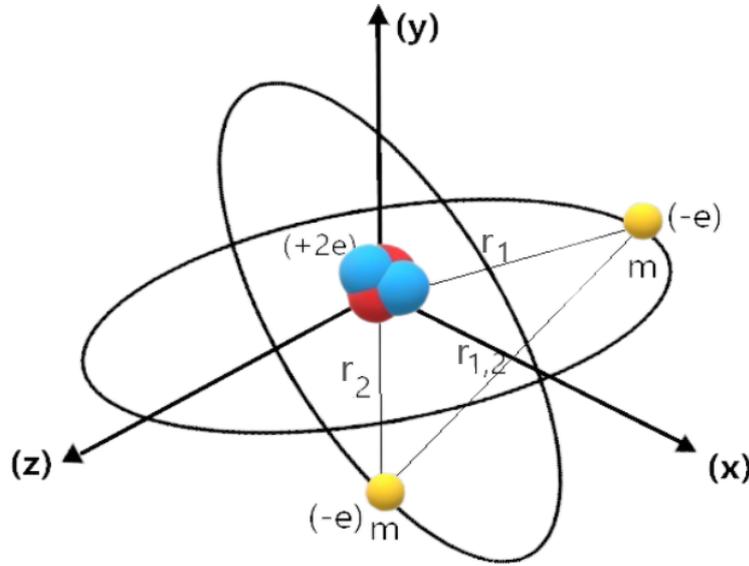


Figure 3.1: Représentation d'atome d'hélium

L'hélium est un élément qui porte une grande importance grâce à ces propriétés physiques, chimiques ou biologiques, ou il a des utilisations dans plusieurs domaines comme la sécurité, l'industrie, et la médecine.

Système d'unités atomiques (u.a): C'est un système d'unités très pratique qui facilite les calculs, il est utilisé en physique quantique, physique atomique...etc. Dans ce système d'unités, la constante de Planck réduite, la masse de l'électron, ainsi que la constante de Coulomb par le carré de la charge élémentaire sont toutes posées égales à l'unité.

$$\hbar = m_e = \frac{e^2}{4\pi\epsilon_0} = 1 \quad (3.2)$$

Avec ces définitions les unités atomiques de distance sont données en unité de a_0 rayon de Bohr et celles de l'énergie par le Hartree E_H :

$$a_0 = \frac{\hbar}{m_e c \alpha} \quad \text{et} \quad E_H = \alpha^2 m_e c^2 \quad (3.3)$$

Avec:

$$\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c} \simeq \frac{1}{137} \quad (3.4)$$

Tel que:

- $-e$: La charge le l'électron.
- m_e : La masse le l'électron.
- \hbar : La constante de Plank réduite.

- a_0 : Le rayon de Bohr, avec $a_0 = 0.529A^\circ$.
- E_H : L'énergie exprimé en Hartree, tel que $1\text{Hartree} = 27.21eV$.
- α : La constante de structure fine (Sans dimension).

3.2.1 Équation de Schrödinger pour l'atome d'hélium

Dans les systèmes atomiques à plusieurs électrons, on prend en considération les interactions coulombiennes de répulsion entre paires d'électrons et celle d'attraction entre les électrons et le noyau (on néglige les interactions dues aux moments cinétiques). En choisissant l'origine du repère sur le noyau que l'on considère immobile du fait de sa grande masse par rapport aux électrons, alors l'atome d'hélium est formalisée comme un problème de deux corps (2 électrons) dans un potentiel coulombien, et l'équation de Shrodinger s'écrit :

$$H\psi(\vec{r}_1, \vec{r}_2) = E\psi(\vec{r}_1, \vec{r}_2) \quad (3.5)$$

Avec H l'hamiltonien du système :

$$H = T_1 + T_2 + V_1(\vec{r}_1) + V_2(\vec{r}_2) + V_{ex}(\vec{r}_1, \vec{r}_2)$$

Tel que:

- $T_1 = \frac{P_1^2}{2m} = -\frac{\hbar^2}{2m}\Delta_1$ et $T_2 = \frac{P_2^2}{2m} = -\frac{\hbar^2}{2m}\Delta_2$ représentent les énergies cinétiques des électrons, avec $\vec{P} = -i\hbar\nabla$ et $\nabla^2 = \Delta$.
- $V_1(\vec{r}_1) = \frac{-2e^2}{4\pi\epsilon_0 r_1}$ et $V_2(\vec{r}_2) = \frac{-2e^2}{4\pi\epsilon_0 r_2}$ représentent l'énergie potentielle d'attractions coulombienne des électrons 1 et 2 respectivement par le noyau.
- $V_{ex} = \frac{e^2}{4\pi\epsilon_0 r_{12}} = \frac{e^2}{4\pi\epsilon_0 |\vec{r}_1 - \vec{r}_2|}$ représente le potentiel d'échange ou bien la répulsion mutuelle des deux électrons.

En remplaçant chaque terme avec sa formule :

$$-\frac{\hbar^2}{2m} [\Delta_1 + \Delta_2] \psi(\vec{r}_1, \vec{r}_2) + \left[-\frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{2e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}} \right] \psi(\vec{r}_1, \vec{r}_2) = E\psi(\vec{r}_1, \vec{r}_2) \quad (3.6)$$

Et en utilisant le système d'unités atomiques (u.a), on peut simplifier l'équation de Shrodinger qui devient :

$$\left(-\frac{1}{2}\Delta_1 - \frac{1}{2}\Delta_2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}} \right) \psi(\vec{r}_1, \vec{r}_2) = E\psi(\vec{r}_1, \vec{r}_2) \quad (3.7)$$

Rappelons l'expression du laplacien en coordonnées sphériques :

$$\Delta_i = \frac{1}{r_i^2} \frac{\partial}{\partial r_i} \left(r_i^2 \frac{\partial}{\partial r_i} \right) + \frac{1}{r_i^2 \sin\theta_i} \frac{\partial}{\partial \theta_i} \left(\sin\theta_i \frac{\partial}{\partial \theta_i} \right) + \frac{1}{r_i^2 \sin^2\theta_i} \frac{\partial^2}{\partial \varphi_i^2} \quad (3.8)$$

3.2.2 Approximation des électrons indépendants

La résolution de l'équation de Shrodinger reste très compliquée est difficile à résoudre à cause du terme d'échange $V_{ex} = \frac{e^2}{4\pi\epsilon_0 r_{12}}$ responsable de l'interaction électron-électron. Pour le choix de notre fonction test pour le calcul variationnel, on utilise l'approximation des électrons indépendants, c'est à dire qu'on néglige dans un premier temps le terme d'échange ce qui permet de séparer les variables sous la forme :

$$\left[\left(-\frac{1}{2}\Delta_1 - \frac{2}{r_1} \right) + \left(-\frac{1}{2}\Delta_2 - \frac{2}{r_2} \right) \right] \psi(\vec{r}_1, \vec{r}_2) = E\psi(\vec{r}_1, \vec{r}_2) \quad (3.9)$$

$$(H_1 + H_2) \psi_1(\vec{r}_1)\psi_2(\vec{r}_2) = (E_1 + E_2)\psi_1(\vec{r}_1)\psi_2(\vec{r}_2) \quad (3.10)$$

Sachant que l'état du système est une fonction factorisable : $\psi(\vec{r}_1, \vec{r}_2) = \psi_1(\vec{r}_1)\psi_2(\vec{r}_2)$, la résolution de chacun des systèmes séparés à 1 électrons :

$$H_1\psi_1(\vec{r}_1) = E_1\psi_1(\vec{r}_1), \quad H_2\psi_2(\vec{r}_2) = E_2\psi_2(\vec{r}_2)$$

Conduit aux fonctions d'ondes d'un atome hydrogénoïde donnée pour l'état fondamental 1s par la fonction radiale :

$$\psi(\vec{r}) = c e^{-\alpha r}$$

Une première approximation de l'énergie de l'état fondamental ($1s^2$) de l'hélium serait alors égale à

$$E = E_1 + E_2 = \frac{-2Z^2}{2} E_H = -4 \text{ Hartree}$$

par comparaison avec la valeur expérimentale de ($-2,90 \text{ Hartree}$) on constate une erreur relativement importante lorsque on néglige la répulsion entre les deux électrons. La méthode variationnelle va nous permettre d'améliorer l'approximation.

3.3 Méthode variationnelle pour l'atome d'hélium

Rappelons que la méthode des variations permet de calculer une approximation de l'énergie de l'état fondamental E_0 d'un système stationnaire [7] en minimisant la valeur moyenne de l'hamiltonien qui représente l'énergie du système :

$$\langle H \rangle = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \quad (3.11)$$

3.3.1 Fonction d'onde test

L'approximation des électrons indépendants nous permet de choisir une fonction test pour la méthode variationnelle donnée par : $\psi(\vec{r}_1, \vec{r}_2) = \psi_1(\vec{r}_1)\psi_2(\vec{r}_2)$. Dans notre calcul des variations nous utilisons deux paramètres variationnels (α, β) et une fonction test sous la forme :

$$\psi_{\alpha,\beta}(\vec{r}_1, \vec{r}_2) = \psi_\alpha(\vec{r}_1)\psi_\beta(\vec{r}_2) \quad (3.12)$$

Sachant que les fonctions individuelles ψ_α, ψ_β sont de la forme hydrogénoïdes ψ_{nlm} constituée de deux parties : une partie radiale R_{nl} et une partie angulaire Y_{lm} , tel que :

$$\psi_{nlm}(r, \vec{\theta}, \phi) = R_{nl}(\vec{r}) \cdot Y_{lm}(\theta, \phi)$$

On sais que pour l'état fondamental ($n = 1, l = 0, m = 0$) de l'ion d'hélium He^+ (Atome Hydrogénoïde, $Z=2$) est caractérisé par la symétrie sphérique qui veut dire qu'il n'y a pas de dépendance en θ et ϕ et que la partie angulaire de la fonction d'onde est constante :

$$Y_{lm} = Y_{00} = \frac{1}{\sqrt{4\pi}}$$

Et pour la partie radiale $R_{nl}(\vec{r}) = R_{10}(\vec{r})$, on a les fonctions exponentielles :

$$R_{10}(\vec{r}_1) = c_1 e^{-\alpha r_1} \quad R_{10}(\vec{r}_2) = c_2 e^{-\beta r_2} \quad (3.13)$$

Les fonctions individuelles ψ_α, ψ_β deviennent alors de la forme :

$$\psi_\alpha(\vec{r}_1) = C_1 e^{-\alpha r_1} \quad \psi_\beta(\vec{r}_2) = C_2 e^{-\beta r_2}$$

et l'état total du système dans l'état fondamental est :

$$\psi_{\alpha,\beta}(\vec{r}_1, \vec{r}_2) = \psi_\alpha(\vec{r}_1) \psi_\beta(\vec{r}_2) = C_1 e^{-\alpha r_1} C_2 e^{-\beta r_2} = C e^{-(\alpha r_1 + \beta r_2)} \quad (3.14)$$

La normalisation de la fonction d'onde se fait en normalisant les fonctions d'ondes individuelles par le biais de leurs parties radiales, ce qui permet de calculer les constantes C_1 et C_2

$$\int_{6D} |\psi_{\alpha,\beta}(\vec{r}_1, \vec{r}_2)|^2 dv_1 dv_2 = \int_{3D} |\psi_\alpha(\vec{r}_1)|^2 dv_1 \int_{3D} |\psi_\beta(\vec{r}_2)|^2 dv_2 = 1$$

$$\langle \psi_\alpha(\vec{r}_1) | \psi_\alpha(\vec{r}_1) \rangle = \int_{3D} |\psi_\alpha(\vec{r}_1)|^2 dv_1 = 1, \quad \langle \psi_\beta(\vec{r}_2) | \psi_\beta(\vec{r}_2) \rangle = \int_{3D} |\psi_\beta(\vec{r}_2)|^2 dv_2 = 1$$

$$\begin{aligned} \langle \psi_\alpha(\vec{r}_1) | \psi_\alpha(\vec{r}_1) \rangle &= \int_0^\infty C_1^2 e^{-2\alpha r_1} r_1^2 dr_1 \int_{4\pi} d\Omega_1 \quad \text{avec} \quad \int_{4\pi} d\Omega_1 = 4\pi \\ \Rightarrow \langle \psi_\alpha(\vec{r}_1) | \psi_\alpha(\vec{r}_1) \rangle &= 4\pi \int_0^\infty C_1^2 r_1^2 e^{-2\alpha r_1} dr_1 \end{aligned}$$

En utilisant la relation mathématique : $\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}}$, on trouve :

$$\langle \psi_\alpha(\vec{r}_1) | \psi_\alpha(\vec{r}_1) \rangle = 4\pi C_1^2 \frac{1}{4\alpha^3} = \frac{\pi C_1^2}{\alpha^3} = 1$$

Donc :

$$\Rightarrow C_1 = \sqrt{\frac{\alpha^3}{\pi}}$$

Un calcul similaire permet de déterminer la constante C_2 :

$$\Rightarrow C_2 = \sqrt{\frac{\beta^3}{\pi}}$$

Les fonctions d'ondes individuelles normalisées s'écrivent finalement :

$$\psi_\alpha(\vec{r}_1) = \sqrt{\frac{\alpha^3}{\pi}} e^{-\alpha r_1} \quad (3.15)$$

$$\psi_\beta(\vec{r}_2) = \sqrt{\frac{\beta^3}{\pi}} e^{-\beta r_2} \quad (3.16)$$

La fonction d'onde totale normalisée est alors :

$$\psi_{\alpha,\beta}(\vec{r}_1, \vec{r}_2) = \psi_\alpha(\vec{r}_1)\psi_\beta(\vec{r}_2) = \frac{\sqrt{\alpha^3\beta^3}}{\pi} e^{-(\alpha r_1 + \beta r_2)} \quad (3.17)$$

et les paramètres α, β sont alors déterminés par la méthode variationnelle en minimisant la fonction d'énergie à deux paramètres suivante :

$$E(\alpha, \beta) = \langle H \rangle(\alpha, \beta) = \langle \psi | H | \psi \rangle, \quad \psi = \psi_{\alpha,\beta} \quad (3.18)$$

Ces paramètres α, β doivent donc vérifier des équations de type :

$$\frac{\partial E}{\partial \alpha} = 0 \quad , \quad \frac{\partial E}{\partial \beta} = 0 \quad (3.19)$$

3.3.2 Calcul de la moyenne $\langle H \rangle$

Cette valeur moyenne peut s'écrire en fonction des énergies cinétiques des électrons ainsi que des énergies potentielles, en effet :

$$H = T_1 + V_1 + T_2 + V_2 + V_{ex}$$

$$\langle H \rangle = \langle T_1 \rangle + \langle V_1 \rangle + \langle T_2 \rangle + \langle V_2 \rangle + \langle V_{ex} \rangle$$

Chaque terme inclut dans l'expression de la moyenne de l'hamiltonien va être développé individuellement. D'après 3.7 :

$$\langle T_1 \rangle = \langle \psi_{\alpha,\beta} | T_1 | \psi_{\alpha,\beta} \rangle = \langle \psi_{\alpha,\beta} | \left(-\frac{1}{2} \Delta_1 \right) | \psi_{\alpha,\beta} \rangle, \quad \langle T_2 \rangle = \langle \psi_{\alpha,\beta} | T_2 | \psi_{\alpha,\beta} \rangle = \langle \psi_{\alpha,\beta} | \left(-\frac{1}{2} \Delta_2 \right) | \psi_{\alpha,\beta} \rangle$$

$$\langle V_1 \rangle = \langle \psi_{\alpha,\beta} | V_1 | \psi_{\alpha,\beta} \rangle = \langle \psi_{\alpha,\beta} | \left(-\frac{2}{r_1} \right) | \psi_{\alpha,\beta} \rangle, \quad \langle V_2 \rangle = \langle \psi_{\alpha,\beta} | V_2 | \psi_{\alpha,\beta} \rangle = \langle \psi_{\alpha,\beta} | \left(-\frac{2}{r_2} \right) | \psi_{\alpha,\beta} \rangle$$

$$\langle V_{ex} \rangle = \langle \psi_{\alpha,\beta} | V_{ex} | \psi_{\alpha,\beta} \rangle = \langle \psi_{\alpha,\beta} | \left(\frac{1}{r_{12}} \right) | \psi_{\alpha,\beta} \rangle$$

Le Laplacien est défini en coordonnées sphérique par la relation 3.8 et du fait de la symétrie sphérique de l'état 1s son expression se réduit à :

$$\Delta = \frac{\partial^2}{\partial^2 r_1} + \frac{2}{r_1} \frac{\partial}{\partial r_1}$$

Alors, pour les énergies cinétiques :

$$\begin{aligned} \langle \psi_{\alpha,\beta} | T_1 | \psi_{\alpha,\beta} \rangle &= -\frac{1}{2} \langle \psi_{\alpha,\beta} | \Delta_1 | \psi_{\alpha,\beta} \rangle \\ &= -\frac{1}{2} \int_{6D} \psi_{\alpha}^*(\vec{r}_1) \psi_{\beta}^*(\vec{r}_2) \Delta_1 \psi_{\alpha}(\vec{r}_1) \psi_{\beta}(\vec{r}_2) dv_1 dv_2. \\ &= -\frac{1}{2} \int_{3D} \psi_{\alpha}^*(\vec{r}_1) \Delta_1 \psi_{\alpha}(\vec{r}_1) dv_1 \cdot \int_{3D} \psi_{\beta}^*(\vec{r}_2) \psi_{\beta}(\vec{r}_2) dv_2 \\ &= -\frac{1}{2} \int_{3D} \frac{\alpha^3}{\pi} e^{-\alpha r_1} \left(\frac{\partial^2}{\partial^2 r_1} + \frac{2}{r_1} \frac{\partial}{\partial r_1} \right) e^{-\alpha r_1} dv_1 \\ &= -\frac{\alpha^3}{2\pi} \int_{3D} e^{-\alpha r_1} \left(\alpha^2 - \frac{2}{r_1} \alpha \right) e^{-\alpha r_1} dv_1 \\ &= -\frac{\alpha^3}{2\pi} \left(\alpha^2 \int_{3D} e^{-2\alpha r_1} dv_1 - 2\alpha \int_{3D} \frac{e^{-2\alpha r_1}}{r_1} dv_1 \right) \\ &= -\frac{\alpha^3}{2\pi} \left(\alpha^2 \int_0^{\infty} e^{-2\alpha r_1} r_1^2 dr_1 \int_0^{\pi} \int_0^{2\pi} \sin \theta d\theta d\phi - 2\alpha \int_0^{\infty} \frac{e^{-2\alpha r_1} r_1^2}{r_1} \int_0^{\pi} \int_0^{2\pi} \sin \theta d\theta d\phi \right) \end{aligned}$$

Sachant que : $\int_0^{\pi} \int_0^{2\pi} \sin \theta d\theta d\phi = \int_{4\pi} d\Omega = 4\pi$ Donc:

$$\langle \psi_{\alpha,\beta} | T_1 | \psi_{\alpha,\beta} \rangle = -\frac{\alpha^3}{2\pi} \left[4\pi \alpha^2 \cdot \frac{2}{(2\alpha)^3} - 2\alpha \cdot 4\pi \cdot \frac{1}{(2\alpha)^2} \right]$$

Et après simplification, on trouve:

$$\boxed{\langle T_1 \rangle = \frac{\alpha^2}{2}}$$

De la même façon on calcule l'énergie cinétique du deuxième électron :

$$\langle \psi_{\alpha,\beta} | T_2 | \psi_{\alpha,\beta} \rangle = -\frac{1}{2} \langle \psi_{\alpha,\beta} | \Delta_2 | \psi_{\alpha,\beta} \rangle$$

Et on trouve on trouve:

$$\boxed{\langle T_2 \rangle = \frac{\beta^2}{2}}$$

Maintenant on passe aux calcul des énergies potentielles:

$$\begin{aligned}
\langle V_1 \rangle &= \langle \psi_{\alpha,\beta} | V_1 | \psi_{\alpha,\beta} \rangle = -2 \langle \psi_{\alpha,\beta} | \left(\frac{1}{r_1} \right) | \psi_{\alpha,\beta} \rangle \\
&= -2 \int_{6D} \psi_{\alpha}^*(r_1) \psi_{\beta}^*(r_2) \frac{1}{r_1} \psi_{\alpha}(r_1) \psi_{\beta}(r_2) dv_1 dv_2 \\
&= -2 \int_{3D} \psi_{\alpha}^*(r_1) \frac{1}{r_1} \psi_{\alpha}(r_1) dv_1 \int_{3D} \psi_{\beta}^*(r_2) \psi_{\beta}(r_2) dv_2 \\
&= -\frac{2\alpha^3}{\pi} \int \frac{e^{-\alpha r_1}}{r_1} e^{-\alpha r_1} r_1^2 dr_1 \int_{4\pi} d\Omega = -\frac{2\alpha^3}{\pi} \cdot 4\pi \int_{3D} e^{-2\alpha r_1} r_1 dr_1 = -\frac{8\alpha^3 \pi}{\pi} \frac{1}{(2\alpha)^2}
\end{aligned}$$

Alors:

$$\boxed{\langle V_1 \rangle = -2\alpha}$$

De même pour $\langle V_2 \rangle$:

$$\langle V_2 \rangle = \langle \psi_{\alpha,\beta} | V_2 | \psi_{\alpha,\beta} \rangle = -2 \langle \psi_{\alpha,\beta} | \left(\frac{1}{r_2} \right) | \psi_{\alpha,\beta} \rangle$$

$$\boxed{\langle V_2 \rangle = -2\beta}$$

Pour le potentiel d'échange $\langle V_{ex} \rangle$:

$$\begin{aligned}
\langle V_{ex} \rangle &= \langle \psi_{\alpha,\beta} | V_{ex} | \psi_{\alpha,\beta} \rangle = -2 \langle \psi_{\alpha,\beta} | \left(\frac{1}{r_{12}} \right) | \psi_{\alpha,\beta} \rangle \\
&= -2 \int_{6D} \psi_{\alpha}^*(r_2) \psi_{\beta}^*(r_1) \frac{1}{r_{12}} \psi_{\alpha}(r_1) \psi_{\beta}(r_2) dv_1 dv_2 \\
\langle V_{ex} \rangle &= \frac{\alpha^3 \beta^3}{\pi^2} \int_{6D} \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 = I(\alpha, \beta)
\end{aligned}$$

Avec: $r_{12} = |\vec{r}_1 - \vec{r}_2|$ et on a noté l'intégrale suivante :

$$\boxed{I(\alpha, \beta) = \int_{6D} \frac{\alpha^3 \beta^3}{\pi^2} \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2} \quad (3.20)$$

Cette intégrale est difficile à calculer de façon exacte avec les méthodes traditionnelles, donc on effectue des calculs numériques par intégration Monte Carlo qui est adaptée aux intégrales multidimensionnelles.

Finalement la valeur moyenne de l'énergie s'écrit alors :

$$\langle H \rangle = E(\alpha, \beta) = \frac{\alpha^2}{2} + \frac{\beta^2}{2} - 2\alpha - 2\beta + I(\alpha, \beta) \quad (3.21)$$

Remarquons que la minimisation exacte par annulation des dérivées partielles 3.19 est impossible car $E(\alpha, \beta)$ contient une intégrale $I(\alpha, \beta)$ qui dépend des paramètres (α, β) . Alors on utilise une autre méthode d'optimisation par récurrence qui est la méthode de la descente du gradient.

3.3.3 Application de la méthode de la descente du gradient

On utilise la méthode de la descente du gradient comme méthode d'optimisation pour chercher le minimum d'une fonction, dans notre cas on cherche à minimiser la valeur de l'énergie de l'atome d'hélium. On applique l'algorithme de la méthode comme suit :

1. Choix dans le plan (α, β) du premier points $R_0(\alpha_0, \beta_0)$.
2. Choix d'un pas optimale $\delta > 0$ le long de la direction $\vec{d}_k = -\vec{\nabla}E(\alpha, \beta)$ de façon à faire décroître f .
3. Calcul des points $\vec{R}_k(\alpha_k, \beta_k)$ en appliquant la formule itérative 2.16 :

$$\vec{R}_{k+1} = \vec{R}_k - \delta \vec{\nabla}E(\alpha_k, \beta_k) \quad (3.22)$$

Formule itérative: L'équation 3.22 conduit au système à deux équations itératives suivantes

$$\begin{cases} \alpha_{k+1} = \alpha_k - \delta_k \frac{\partial E}{\partial \alpha} \\ \beta_{k+1} = \beta_k - \delta_k \frac{\partial E}{\partial \beta} \end{cases} \quad (3.23)$$

La formule de l'énergie 3.21 permet de trouver :

$$\begin{cases} \frac{\partial E}{\partial \alpha} = \alpha - 2 + \frac{\partial I}{\partial \alpha} \\ \frac{\partial E}{\partial \beta} = \beta - 2 + \frac{\partial I}{\partial \beta} \end{cases}$$

La formule 3.20 permet d'écrire :

$$\begin{aligned} \frac{\partial I}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \left(\frac{\alpha^3 \beta^3}{\pi^2} \int_{6D} \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 \right) \\ &= \frac{3}{\alpha} \int_{6D} \frac{\alpha^3 \beta^3}{\pi^2} \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 - \int_{6D} \frac{2\alpha^3 \beta^3}{\pi^2} r_1 \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 \end{aligned}$$

En notant :

$$I_1(\alpha, \beta) = \int_{6D} \frac{2\alpha^3\beta^3}{\pi^2} r_1 \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 \quad (3.24)$$

On obtient :

$$\frac{\partial I}{\partial \alpha} = \frac{3}{\alpha} I(\alpha, \beta) - I_1(\alpha, \beta)$$

De même pour la variable β , le même calcul donne :

$$\frac{\partial I}{\partial \beta} = \frac{3}{\beta} I(\alpha, \beta) - I_2(\alpha, \beta)$$

Avec:

$$I_2(\alpha, \beta) = \int_{6D} \frac{2\alpha^3\beta^3}{\pi^2} r_2 \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 \quad (3.25)$$

Il s'en suit :

$$\begin{cases} \frac{\partial E}{\partial \alpha} = \alpha - 2 + \frac{3}{\alpha} I(\alpha, \beta) - 2I_1(\alpha, \beta) \\ \frac{\partial E}{\partial \beta} = \beta - 2 + \frac{3}{\beta} I(\alpha, \beta) - 2I_2(\alpha, \beta) \end{cases} \quad (3.26)$$

Et finalement, en remplaçant dans le système itératif 3.23 on obtient :

$$\begin{cases} \alpha_{k+1} = \alpha_k - \delta \left(\alpha_k - 2 + \frac{3}{\alpha_k} I(\alpha_k, \beta_k) - 2I_1(\alpha_k, \beta_k) \right) \\ \beta_{k+1} = \beta_k - \delta \left(\beta_k - 2 + \frac{3}{\beta_k} I(\alpha_k, \beta_k) - 2I_2(\alpha_k, \beta_k) \right) \end{cases} \quad (3.27)$$

Les intégrales I, I_1 et I_2 devront être évalué à chaque itération k par la méthodes Monte Carlo.

3.3.4 Intégration Monte Carlo

D'après les calculs qu'on a effectué dans la partie précédente on est arrivé a la formule de l'énergie totale 3.21 et aux formules itératives 3.27 des paramètres α, β . Or ces formules contiennent les intégrales multidimensionnelles I, I_1 et I_2 . Pour les calculer nous allons utiliser la technique d'intégration de Monte Carlo. On a la choisi la variante d'intégration avec échantillonnage non uniforme, ce choix est justifié par le fait que l'intégration avec échantillonnage préférentiel (la plus précise) nécessite la recherche d'une fonction proche de la fonction a intégrer, mais cela s'avère difficile dans notre cas pour nos fonctions a 6 dimensions.

On remarque que l'intégrale I (qui n'est autre que la valeur moyenne du potentiel d'échange) et les intégrales I_1 et I_2 peuvent être écrits sous la forme :

$$I(\alpha, \beta) = \int_{6D} \frac{\alpha^3 \beta^3}{\pi^2} \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 = \int_{6D} \rho_{\alpha, \beta}(\vec{r}_1, \vec{r}_2) f(\vec{r}_1, \vec{r}_2) dv_1 dv_2 \quad (3.28)$$

$$I_1(\alpha, \beta) = \int_{6D} \frac{2\alpha^3 \beta^3}{\pi^2} r_1 \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 = \int_{6D} \rho_{\alpha, \beta}(\vec{r}_1, \vec{r}_2) f_1(\vec{r}_1, \vec{r}_2) dv_1 dv_2 \quad (3.29)$$

$$I_2(\alpha, \beta) = \int_{6D} \frac{2\alpha^3 \beta^3}{\pi^2} r_2 \frac{e^{-2(\alpha r_1 + \beta r_2)}}{r_{12}} dv_1 dv_2 = \int_{6D} \rho_{\alpha, \beta}(\vec{r}_1, \vec{r}_2) f_2(\vec{r}_1, \vec{r}_2) dv_1 dv_2 \quad (3.30)$$

Avec

$$f(\vec{r}_1, \vec{r}_2) = \frac{1}{|\vec{r}_1 - \vec{r}_2|} \quad (3.31)$$

$$f_1(\vec{r}_1, \vec{r}_2) = \frac{2r_1}{|\vec{r}_1 - \vec{r}_2|} \quad (3.32)$$

$$f_2(\vec{r}_1, \vec{r}_2) = \frac{2r_2}{|\vec{r}_1 - \vec{r}_2|} \quad (3.33)$$

et

$$\rho_{\alpha, \beta}(\vec{r}_1, \vec{r}_2) = |\psi_{\alpha, \beta}(\vec{r}_1, \vec{r}_2)|^2 = |\psi_{\alpha}(\vec{r}_1)|^2 |\psi_{\beta}(\vec{r}_2)|^2 = \frac{\alpha^3 \beta^3}{\pi^2} e^{-2(\alpha r_1 + \beta r_2)} \quad (3.34)$$

n'est autre que la densité de probabilité de présence des électrons de l'atome d'hélium. Celle-ci qui est déjà normalisée va être utilisée comme densité non-uniforme dans l'intégration Monte Carlo. Les formules de Monte Carlo avec densité non-uniforme donne alors les approximations suivantes des trois intégrales :

$$I \simeq \frac{1}{N} \sum_{i=1}^N f(\vec{r}_{1i}, \vec{r}_{2i}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{|\vec{r}_{1i} - \vec{r}_{2i}|} \quad (3.35)$$

$$I_1 \simeq \frac{1}{N} \sum_{i=1}^N f_1(\vec{r}_{1i}, \vec{r}_{2i}) = \frac{1}{N} \sum_{i=1}^N \frac{2r_{1i}}{|\vec{r}_{1i} - \vec{r}_{2i}|} \quad (3.36)$$

$$I_2 \simeq \frac{1}{N} \sum_{i=1}^N f_2(\vec{r}_{1i}, \vec{r}_{2i}) = \frac{1}{N} \sum_{i=1}^N \frac{2r_{2i}}{|\vec{r}_{1i} - \vec{r}_{2i}|} \quad (3.37)$$

Les N points d'intégration $(\vec{r}_{1i}, \vec{r}_{2i})$ doivent être générés selon la densité normalisée $\rho_{\alpha, \beta}(\vec{r}_1, \vec{r}_2)$ dans tout l'espace \mathbb{R}^6 . Cette densité peut être factorisée selon les variables indépendantes de chaque électron en deux densités radiales :

$$\rho_{\alpha, \beta}(\vec{r}_1, \vec{r}_2) = \rho_1(\vec{r}_1) \rho_2(\vec{r}_2) = \frac{\alpha^3}{\pi} e^{-2\alpha r_1} \frac{\beta^3}{\pi} e^{-2\beta r_2} = |\psi(\vec{r}_1, \vec{r}_2)|^2$$

Ce qui revient à générer les N points dans les 2 sous espaces 3-dimensionnels selon les distributions ρ_1 et ρ_2 . Pour cela écrivons en coordonnées sphériques la probabilité de présence de l'électron 1 dP_1 dans le volume élémentaire $dv_1 = r_1^2 dr_1 d\Omega_1$, ($d\Omega_1 = \sin \theta_1 d\theta_1 d\varphi_1$ étant la partie angulaire) :

$$dP_1 = \rho_1 dv_1 = 4\alpha^3 r_1^2 e^{-2\alpha r_1} dr_1 \times \frac{1}{4\pi} d\Omega_1 = \rho_{1r} dr_1 \times \rho_{\Omega_1} d\Omega_1 \quad (3.38)$$

Donc la coordonnée radiale du point \vec{r}_1 est générée selon la distribution radiale

$$\rho_{1r} = 4\alpha^3 r_1^2 e^{-2\alpha r_1}$$

avec $r_1 \in [0, \infty[$ et la coordonnée angulaire $\Omega_1 = (\theta_1, \varphi_1)$ est générée de façon isotrope dans l'espace. On a évidemment les relations de normalisation des densités radiales et angulaires:

$$\int_0^\infty \rho_{1r} dr_1 = 1, \quad \int_{4\pi} \rho_{\Omega_1} d\Omega_1 = 1$$

On a des relations similaires pour l'électron 2.

$$\rho_{2r} = 4\beta^3 r_2^2 e^{-2\beta r_2}$$

3.3.4.1 Valeurs théoriques exactes par la méthode variationnelle

Notons qu'il existe un procédé exacte [36] mais fastidieux pour calculer les intégrales I, I_1 et I_2 . Avec ce calcul exacte, le calcul variationnel abouti aux valeurs des paramètres α, β et de l'énergie suivantes [37] :

$$\alpha = \beta = \frac{27}{16} = 1.6875 \quad (3.39)$$

$$E = - \left(Z - \frac{5}{16} \right)^2 = -2.84765 \text{ Hartree} \quad (3.40)$$

Ces valeurs permettrons de comparer nos calculs qui utilisent l'intégration Monte Carlo.

3.3.5 Organigramme et code de calcul :

Le calcul de minimisation de l'énergie a été réalisé à l'aide d'un code Maple. Le but est donc de calculer numériquement de proche en proche les paramètres (α_m, β_m) qui minimisent l'énergie $E(\alpha, \beta)$ (formule 3.21) ainsi que la valeur de cette énergie minimale $E(\alpha_m, \beta_m)$ à l'aide de la méthode de la descente du gradient. Ceci est réalisé à l'aide d'une boucle itérative sur les (α, β) (formule 3.27). A chaque itération, les intégrales I, I_1 et I_2 sont calculés par intégration Monte Carlo avec échantillonnage non-uniforme sur l'espace 6-dimensionnel des deux électrons avec une densité égale à la densité de probabilité de présence des électrons $\rho(\vec{r}_1, \vec{r}_2) = |\psi_{\alpha, \beta}|^2$.

Le fonctionnement du code est représenté dans l'organigramme suivant :

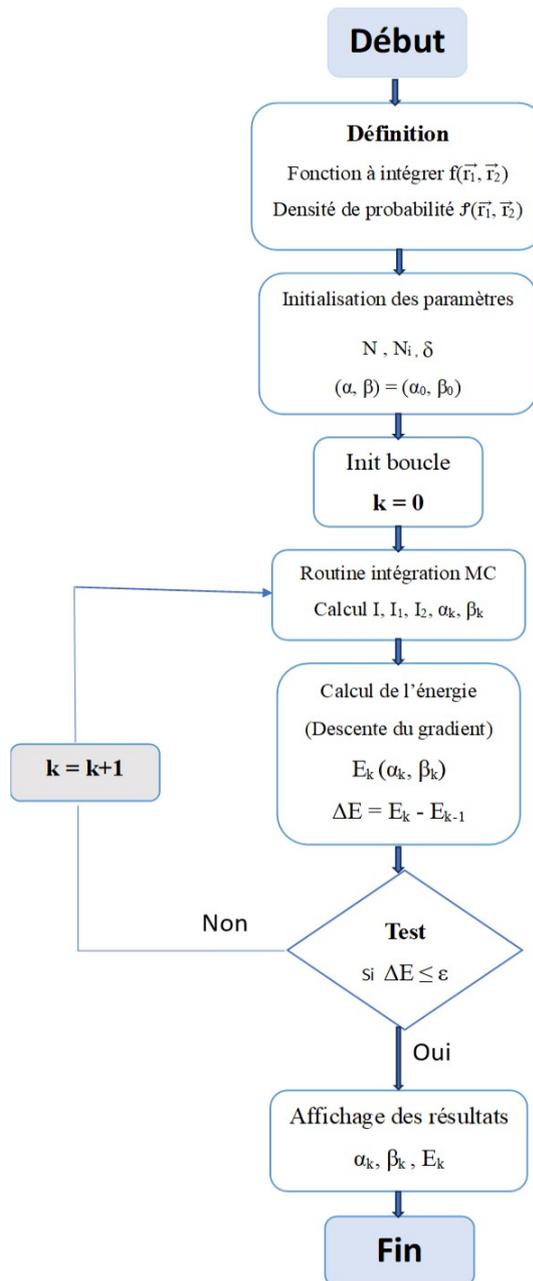


Figure 3.2: Organigramme de notre calcul (Calcul variationnel par descente du gradient et Monte Carlo)

On commence par la définition des fonctions à intégrer par Monte Carlo ainsi que la fonction densité de probabilité, on passe ensuite à la déclaration des valeurs initiales des paramètres de calcul (α_0, β_0) , du pas de la descente du gradient δ , du nombre de

points d'intégrations N et du nombre d'itérations N_i . Après l'initialisation de la boucle, à chaque étape k le programme commence par calculer les intégrales I, I_1 et I_2 à l'aide de la routine Maple "MCInteg" puis les paramètres (α_k, β_k) et l'énergie $E(\alpha_k, \beta_k)$. La boucle est terminée à l'aide d'une condition de convergence.

Notons que les valeurs des paramètres variationnels initiaux (α_0, β_0) ont été choisi à partir du modèle des électrons indépendants pour lequel le terme $\langle V_{ex} \rangle = I(\alpha, \beta)$ est nul, la minimisation de 3.21 par dérivation devient alors possible et donne $(\alpha, \beta) = (2, 2)$. En effet :

$$I(\alpha, \beta) \Rightarrow \langle H \rangle = E(\alpha, \beta) = \frac{\alpha^2}{2} + \frac{\beta^2}{2} - 2\alpha - 2\beta \quad (3.41)$$

La minimisation directe donne alors

$$\frac{\partial E}{\partial \alpha} = 0, \quad \frac{\partial E}{\partial \beta} = 0 \Rightarrow \alpha - 2 = 0, \quad \beta - 2 = 0 \Rightarrow \alpha = \beta = 2 \quad (3.42)$$

La même valeur pour α et β est prévisible vu la symétrie du problème. Pour notre calcul et pour plus de flexibilité nous avons choisi $(\alpha_0, \beta_0) = (2, 1.5)$

3.3.6 Résultats et discussion

Les résultats de calculs sont présentés dans les figures ci-dessous. Sur le coté gauche on montre l'énergie en fonction du nombre d'itérations et sur le coté droite la convergence des paramètres (α, β) avec un zoom sur la zone de convergence. Nous avons effectué le calcul pour 3 différentes valeurs du nombre de points d'intégration MC : $N = 10^4, 10^5, 10^6$ et pour chacun de ces cas nous avons utilisé différentes combinaisons du pas δ et du nombre d'itérations pour la descente du gradient.

Vu la symétrie de la fonction d'onde $\psi_{\alpha, \beta}$ 3.17 par rapport aux paramètres (α, β) et des équations 3.27 on s'attend que les paramètres optimaux (α_m, β_m) soient égaux $\alpha_m = \beta_m$, la droite d'équation $\alpha = \beta$ tracé sur les figures permet de voir la zone de convergence à l'intersection avec cette droite.

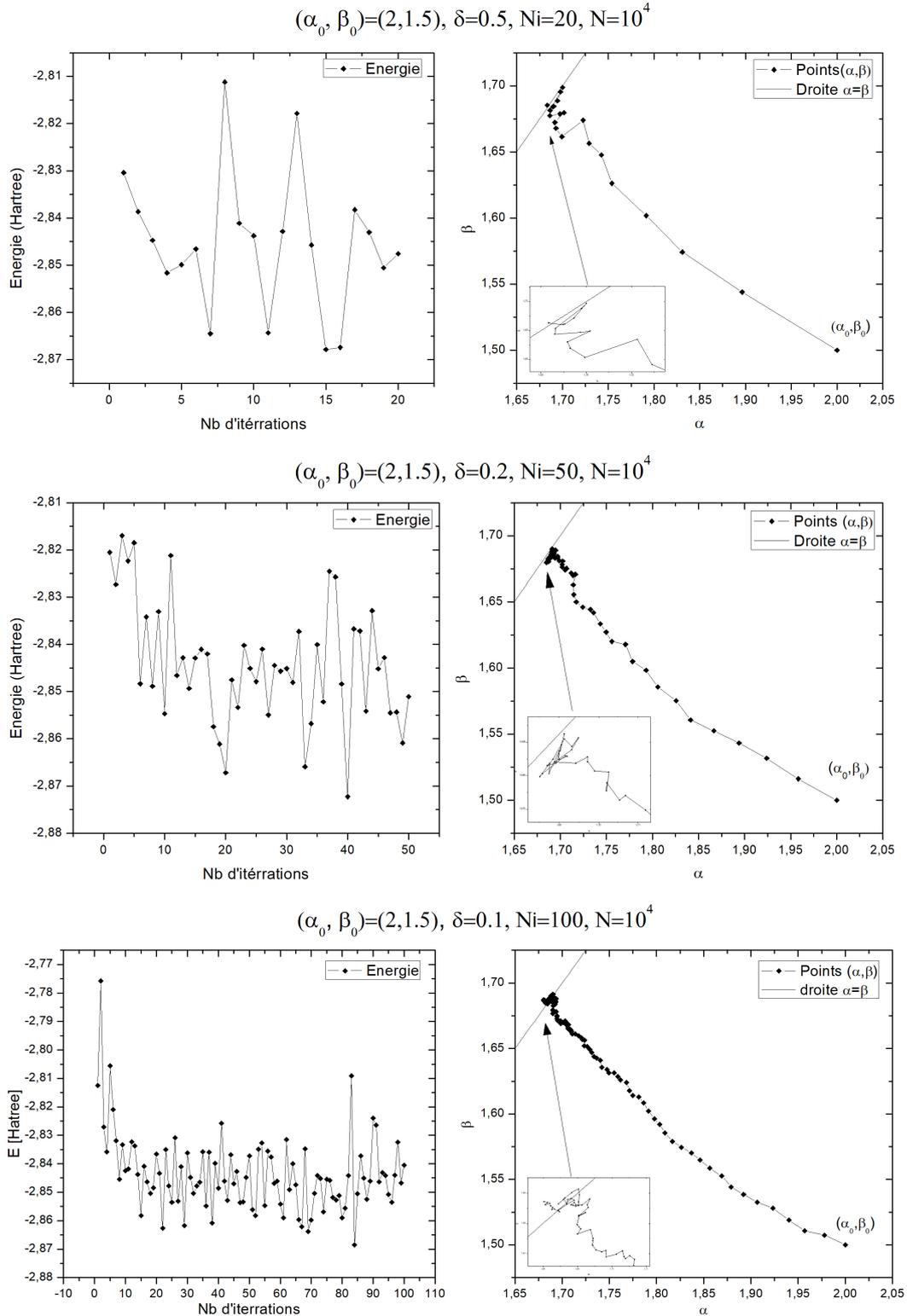


Figure 3.3: Convergence des paramètres (α, β) et de l'énergie $E(\alpha, \beta)$ en fonction du nombre d'itérations pour $(\alpha_0, \beta_0) = (2, 1.5)$ et $N = 10^4$

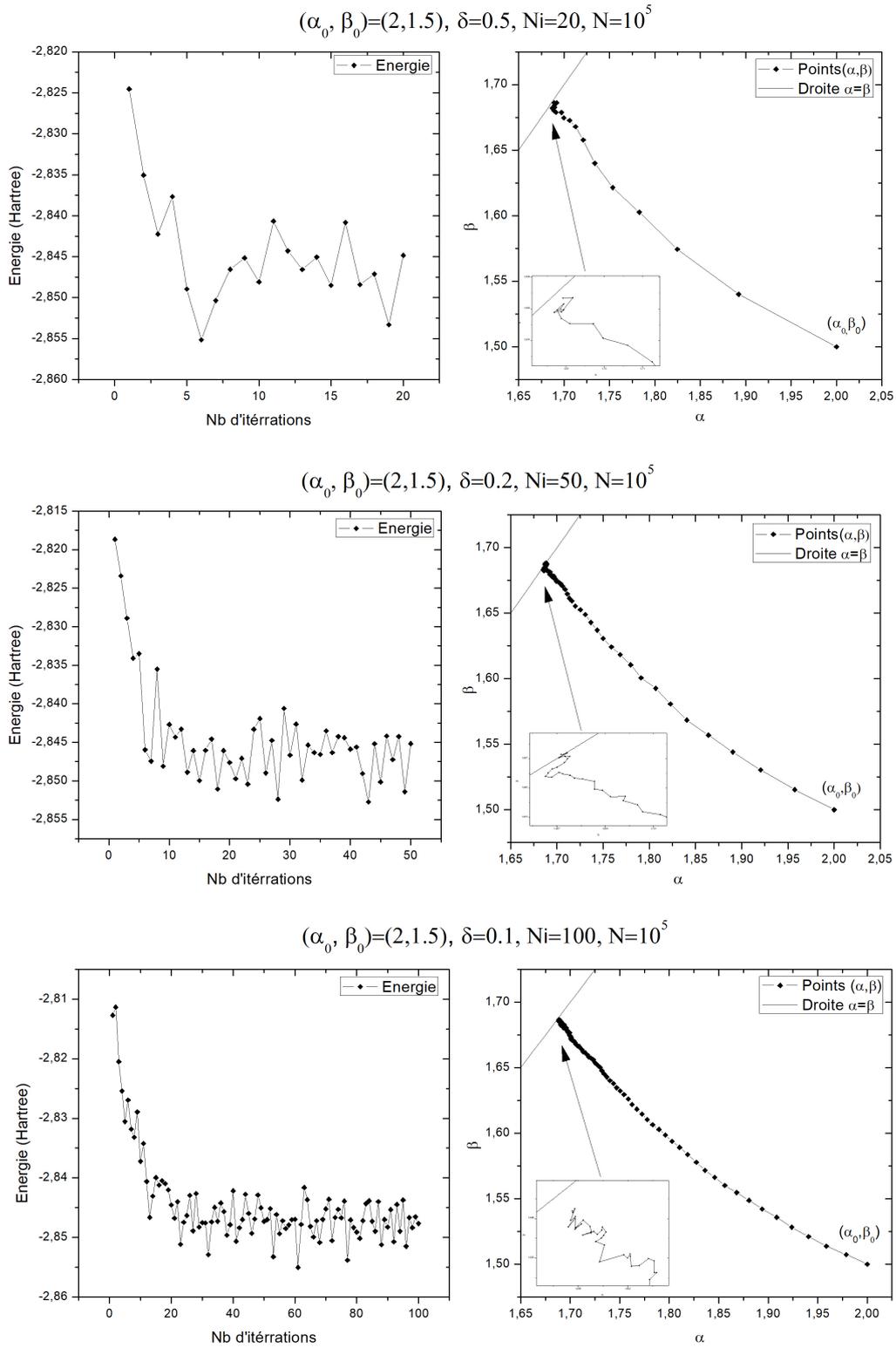


Figure 3.4: Convergence des paramètres (α, β) et de l'énergie $E(\alpha, \beta)$ en fonction du nombre d'itérations pour $(\alpha_0, \beta_0) = (2, 1.5)$ et $N = 10^5$

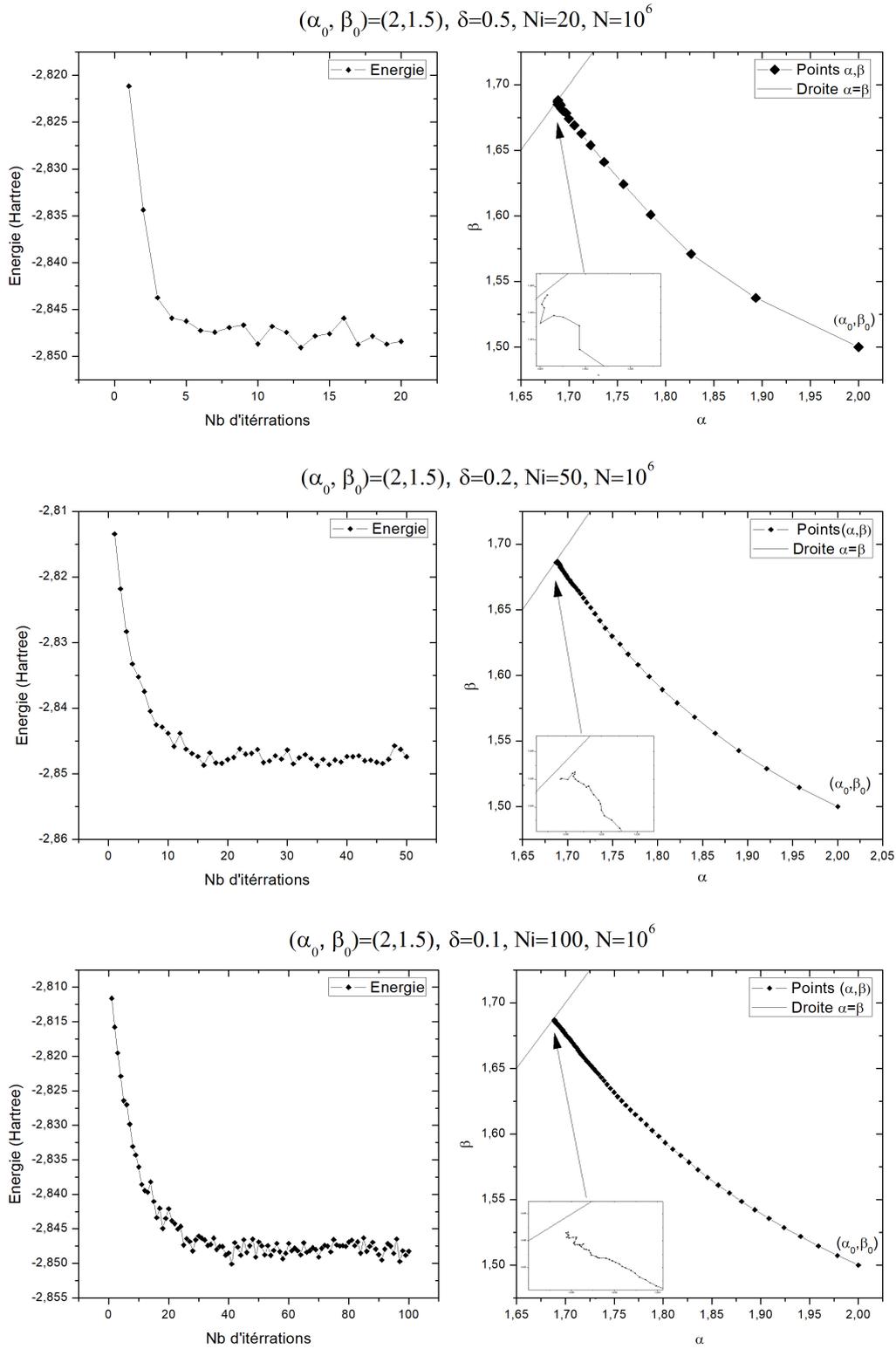


Figure 3.5: Convergence des paramètres (α, β) et de l'énergie $E(\alpha, \beta)$ en fonction du nombre d'itérations pour $(\alpha_0, \beta_0)=(2,1.5)$ et $N = 10^6$

3.3.6.1 Calcul fin

Dans le but d'améliorer encore la précision et obtenir une valeur plus fine de l'énergie, nous avons lancé un dernier calcul avec les paramètres $(\alpha, \beta) = (1.688, 1.687)$ obtenu avec les calculs précédants et avec un pas $\delta = 0.1$. Vu la lourdeur des calculs avec le matériel informatique à notre disposition, les intégrales des 20 premières itérations ont été calculé avec $N = 5 \cdot 10^6$ points et les 5 dernières avec $N = 10^7$ points pour améliorer encore plus la convergence. Les valeurs obtenue sont présentées dans le tableau suivant :

N_i	α	β	$E(\alpha, \beta)$
1	1.688	1.687	-2.848106886
2	1.687952759	1.687018566	-2.848086931
3	1.687915037	1.687006409	-2.847563898
4	1.687878207	1.687099225	-2.847497932
5	1.687811743	1.687115319	-2.847705482
6	1.687817436	1.687161782	-2.847251928
7	1.687821854	1.687186169	-2.847639068
8	1.687779435	1.687209543	-2.847971909
9	1.687763548	1.687263389	-2.847283825
10	1.68773757	1.687342955	-2.847593257
11	1.687724336	1.68736178	-2.848038333
12	1.687694898	1.687373992	-2.847509079
13	1.687691976	1.687392078	-2.847570061
14	1.687661976	1.687418207	-2.847769043
15	1.687632545	1.687425460	-2.847473481
16	1.687628255	1.687442521	-2.847839087
17	1.687622235	1.687482322	-2.847506946
18	1.687611437	1.687498542	-2.847733101
19	1.687598141	1.687512580	-2.847852920
20	1.687595701	1.687530705	-2.847578421
21	1.687591351	1.687539231	-2.847680123
22	1.687589544	1.687547512	-2.847595992
23	1.687580215	1.687555789	-2.847662051
24	1.687578640	1.687563232	-2.847642512
25	1.687565640	1.687569932	-2.847648525

Table 3.1: Valeurs obtenues par le calcul fin des paramètres α, β et de l'énergie pour $(\alpha_0, \beta_0) = (1.688, 1.687)$, $\delta = 0.1$, $N_i = 20 + 5$ et $N = 5 \cdot 10^6, 10^7$.

Ces résultats sont tracés dans la figure ci-dessous

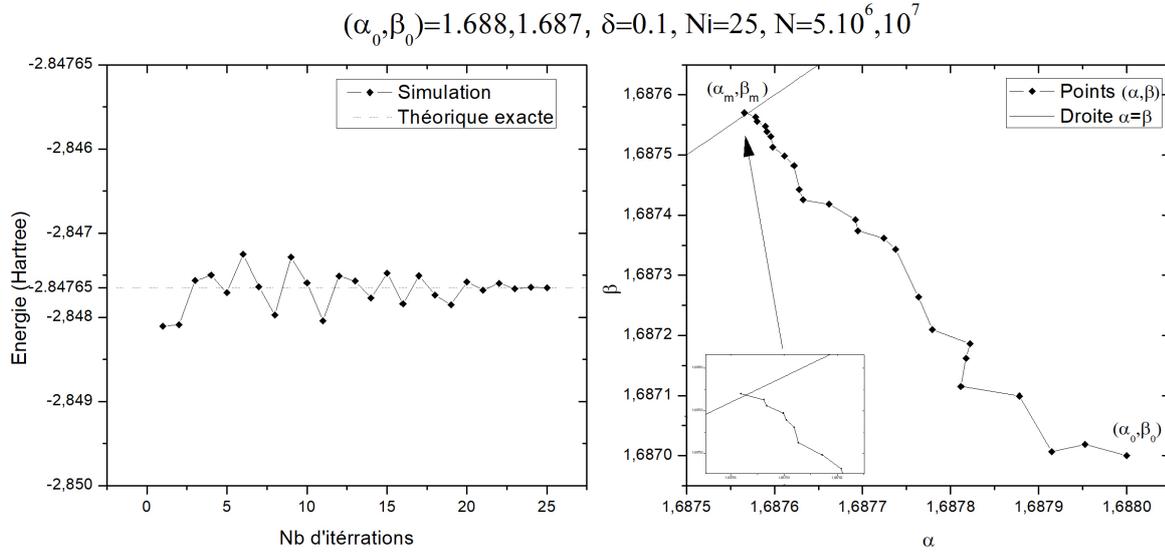


Figure 3.6: Convergence des paramètres α, β et représentation de l'énergie en fonction du nombre d'itérations pour $(\alpha_0, \beta_0) = (1.688, 1.687)$, $\delta = 0.1$, $Ni = 20 + 5$ et $N = 5.10^6, 10^7$

3.3.6.2 Discussion

D'après les figures 3.3, 3.4 et 3.5, on peut faire les observations suivantes:

- Des fluctuations dans la convergence du calcul qui sont plus visibles sur la courbe de l'énergie, ces fluctuations sont due aux caractère stochastique de l'intégration Monte Carlo. Il est donc prévisible qu'elle diminue avec l'augmentation des points d'intégrations N comme on le voit sur les figures. D'après 3.3, 3.4 et 3.5, par comparaison des trois groupes de figures les meilleurs résultats sont ceux de $N = 10^6$. En effet la précision de l'intégration Monte Carlo étant inversement proportionnelle à \sqrt{N} 1.35, plus N augmente plus on améliore la précision .
- Pour chaque groupe de figures, on remarque aussi des fluctuations qui diminuent lorsqu'on passe d'un grand pas $\delta = 0.5$ vers un petit pas $\delta = 0.1$, un grand pas provoquant un "over-shooting" dans la minimisation couplé au caractère stochastique de l'intégration Monte Carlo conduit à beaucoup de fluctuations dans la convergence.
- Comme prévu, les paramètres (α, β) convergent vers la zone d'intégration $\alpha = \beta$.
- La meilleur convergence a été obtenue avec $N = 10^6$ et $\delta = 0.1$. Dans ce cas on obtient les valeurs : $(\alpha, \beta) = (1.688, 1.687)$, $E = -2.8481068 Hartree$. Ainsi, on peut dire que l'obtention des meilleurs résultats nécessite une augmentation des points d'intégrations N et une diminution de la valeur du pas δ .

Le calcul fin nous a permis d'améliorer la convergence, en changeant les valeurs des paramètres initiaux (α_0, β_0) avec les dernières resultats obtenus $(1.688, 1.687)$ et un pas $\delta = 0.1$

D'après la figure 3.5 on observe que la convergence des paramètres (α, β) est très bien selon les dernières valeurs obtenues (α_m, β_m) et l'intersection avec la droite $(\alpha = \beta)$. Concernant les valeurs de l'énergie aussi on peut dire que les fluctuations ont diminué au minimum jusqu'à ce que les résultats finaux de la simulation soient presque identiques à la valeur théorique exacte comme la figure le montre ainsi que le tableau 3.1, avec une précision de $\simeq 10^{-4}$.

Le choix des fonctions d'ondes d'essai à deux paramètres (α et β) a amélioré la précision du calcul variationnel, avec une valeur finale de l'énergie de l'état fondamentale égale à $E = -2.847648525$ *Hartree* en parallèle la valeur exacte avec ce type de fonction test est $E_{ex} = -2,84765$ *Hartree*, c'est-à-dire avec une erreur relative :

$$\frac{\Delta E}{E} = \left| \frac{-2.847648525 + 2,84765}{2,84765} \right| = 5.18 \times 10^{-5}$$

La méthode de la descente du gradient couplé à la méthode Monte Carlo nous a donc permis de retrouver la valeur exacte de l'énergie avec une très bonne approximation. Le choix de l'utilisation d'une distribution non-uniforme revient aux difficultés de définir une fonction proche de notre fonction initiale avec la technique d'échantillonnage préférentiel.

Pour finir, il faut mentionner que la valeur exacte de l'énergie donnée par la méthode variationnelle est la meilleure valeur obtenue en utilisant ce type de fonction test avec des intégrales exactes. Ce qui ne correspond pas exactement à la valeur expérimentale, étant donné que nous avons négligé plusieurs termes dans l'hamiltonien (interaction spin orbit, spin-spin etc...). Le but de notre calcul était de tester l'efficacité de l'optimisation par descente du gradient ainsi que l'intégration par Monte Carlo, ce but a été atteint avec succès. Pour approcher encore mieux la valeur expérimentale, il va falloir inclure les autres interactions et choisir une fonction test plus proche de la réalité.

Conclusion

Les algorithmes non déterministes comme les méthodes de Monte Carlo représentent un aspect fascinant et crucial en plusieurs domaines notamment à la physique et à l'étude des propriétés intrinsèque des éléments offrant des solutions aux problèmes complexes. Ces algorithmes reposent sur un échantillonnage aléatoire répété et sont caractérisés par leurs stochasticité.

Ce travail avait deux principaux objectifs, le premier c'est de présenter deux importantes méthodes de simulation numériques l'une pour l'optimisation qui est la méthode de la descente du gradient et l'autre pour l'intégration qui est la méthode de Monte Carlo. Le deuxième objectif était d'appliquer ces méthodes dans le calcul quantique variationnel à deux paramètres d'une propriété reliée à l'atome d'hélium, il s'agit de l'énergie de l'état fondamental.

Dans le premier chapitre, nous avons présenté la méthode Monte Carlo ainsi que quelques concepts mathématiques dont nous avons besoin pour la compréhension de la méthode, puis la deuxième partie du même chapitre a été consacré à l'intégration Monte Carlo multidimensionnelle et ses trois variantes avec échantillonnage uniforme, non-uniforme et préférentiel. En fin de chapitre nous avons présenté une étude comparative avec un exemple de calcul d'intégrale à 6 dimensions d'une fonction à 6 variables. Le calcul a été réalisé à l'aide du logiciel Maple, dont le résultat a montré la supériorité de l'intégration Monte Carlo avec échantillonnage préférentiel qui a permis de réduire la variance par rapport aux deux autres méthodes.

Le deuxième chapitre a été consacré dans sa première partie à la présentation de la méthode variationnelle en mécanique quantique et ces principes de base comme le théorème de Ritz. Dans la deuxième partie nous avons abordé la méthode d'optimisation connue sous le nom de la descente du gradient avec son algorithme de calcul en montrant des figures illustratrices expliquant la technique.

Dans le troisième chapitre qui représente le cœur de notre mémoire, nous avons appliqué la méthode variationnelle à deux paramètres couplée à la méthode de la descente du gradient et à l'intégration MC pour calculé la valeur de l'énergie de l'état fondamental de l'atome d'hélium. Le but principal était de tester l'efficacité de la méthode de minimisation par la descente du gradient et celle de l'intégration par Monte Carlo. Ce calcul nécessite l'évaluation de la valeur moyenne de l'hamiltonien contenant plusieurs termes.

Ceux des énergies cinétiques et potentielles individuelles ont été calculé de façon exacte, et celui de l'énergie potentielle d'échange, difficile à intégrer, a été calculer numériquement par Monte Carlo. La dépendance de l'intégrale d'échange des paramètres variationnels empêche la minimisation directe, celle-ci a été effectué par la méthode itérative de la descente du gradient.

Les calculs nous ont permis de retrouver une convergence vers les valeurs exactes des paramètres variationnelles (α, β) ainsi que l'énergie avec plusieurs essais et après une augmentation de nombres de points d'intégrations, dont les résultats obtenus sont très satisfaisants par rapport à des calculs à un paramètres dans des mémoires précédents, en témoigne la faible valeur d'erreur relative, qui prouve le succès du processus de simulation.

Par conclusion, les deux méthodes de simulation, descente du gradient et intégration Monte Carlo ont prouvés leurs efficacités au calcul de minimisation et à l'évaluation des intégrales multidimensionnelles malgré qu'elles nécessitent un matériel informatique performant pour générer des échantillons de points d'intégration de haute dimension et de grande taille. Pour améliorer encore le calcul, on peut citer comme perspectives l'utilisation d'une fonction test plus adaptée avec un plus grand nombres de paramètres et en incluant les interactions que l'on a négligé comme les interactions liés au moment cinétique orbital et de spin, mais cela ne peut se faire qu'avec un outil de calcul informatique puissant.

Bibliographie

- [1] Eric Veach, Monte-Carlo Intégration, Stanford University, 1997.
- [2] Michel Roger. Méthodes de Monte-Carlo, octobre 2008.
- [3] N. Metropolis et S. Ulam, The Monte Carlo Methode, *Journal of the American Statistical Association*, vol.44, n°247, 1949, p.335-341.
- [4] J. Escoffier, Probabilités et statistiques pour le CAPES externe et l'Agrégation interne, 2010.
- [5] B. Lapeyre - E. Pardoux - R. Sentis, Méthodes de Monte-Carlo pour les équations de transport et de diffusion, Verlag Berlin Heidelberg, 1998.
- [6] D. Fredon, M. Maumy-Bertrand, F. Bertrand. *Mathématiques Statistique et probabilités*, 2009.
- [7] C. Cohen-Tannoudji, B. Diu, F. Laloe, *Mécanique quantique (Tom II)*, CNRS, 2018.
- [8] Edwin K.P. Chong, Stanislaw H. Zak - *An Introduction to optimisation*, fourth Edition, 2013.
- [9] G. Allaire - *Analyse numerique et optimisation-Une introduction à la modelisation mathematique et à la simulation numerique* French.
- [10] J. Hladik, M. Chrysos, P. Hladik, L. Ugo Ancarani - *Mécanique quantique, Atomes et noyaux, applications technologiques - cours et exercices corrigés*-Dunod, 2009.
- [11] N. Zettili - *Quantum Mechanics- Concepts and Applications*-Wiley, 2009.
- [12] J. Senegas, école des mines de Paris, thèse de Doctorat, *Methodes de Monte Carlo en vision stioscopique, Application à l'étude de modèles numériques de terrain*, 2002.
- [13] S.Kant Mishra, B. Ram - *Introduction to Unconstrained Optimization with R*-Springer, 2019.
- [14] M. Hayoun, *La methode de Monte Carlo Metropolis*, Paris, 2002
- [15] R. Fakhreddine, *Méthodes de Monte Carlo stratifiées pour l'intégration et la simulation numériques*, thèse de Doctorat, univ de Grenoble, 2005 - 2006.

- [16] A. Perrut, Cours de probabilités et statistiques, Université Claude Bernard Lyon 1, 2010.
- [17] G. Baudrand. Mathématiques résumés du cours ECE, 2008.
- [18] P. Dusart, Cours de Probabilités, 2013.
- [19] Dictionnaire unifié des termes de biologie, Tunis:Bureau de coordination de l'arabisation, 1993, p. 239.
- [20] F. Legrand. Magazine scientifique (Informatique Appliquée aux Sciences Physiques), <https://www.f-legrand.fr>.
- [21] D. Andre, Article: Qu'est-ce qu'un Algorithme Non Déterministe?décembre 28, 2023.<https://www.allaboutai.com>
- [22] A. Joulin, Méthodes de Monte Carlo, INSA Toulouse, 2019 - 2020.
- [23] Jinhong Du, Monte Carlo Imtegration, <https://cran.r-project.org>
- [24] B. Boulahlib, Calcul de quelques propriétés quantiques de l'atome d'hélium par intégration Monte Carlo, Thèse de Master, univ. BBA 2021/2022.
- [25] D. Unzueta, Méthode d'acceptation-rejet, <https://towardsdatascience.com>
- [26] M. Belkasmi, Méthode Monte Carlo quantique par diffusion : Application à l'oscillateur harmonique, Thèse de Master, Univ. BBA. 2020/2021.
- [27] Le Gradient, article sur le site Wikipédia: <https://fr.wikipedia.org>, Gradient, sur [larousse.fr](https://www.larousse.fr)
- [28] Gradient From Wikipedia, the free encyclopedia, Gradient, sur [larousse.fr](https://www.larousse.fr)
- [29] David R. Lide, CRC Handbook of Chemistry and Physics, CRC Press Inc, 2009.
- [30] W. M. Haynes, Ionization Energies of Atoms and Atomic Ions, in CRC Handbook of Chemistry and Physics, 91st Edition 2011, ed., CRC Press/Taylor.
- [31] B. Ali, Descente suffisante et conergence globale de la methode du gradient conjugue, 2014.
- [32] A. Cauchy, Méthode générale pour la résolution des systèmes d'équations simultanées, Comptes Rendus de l'Académie des Sciences de Paris.
- [33] Descente de gradient pour les réseaux de neurones, <https://lucidar.me/fr>
- [34] Algorithme du gradient, Article sur wikipedia l'encyclopedie, <https://fr.wikipedia.org>.
- [35] Élie Lévy, Dictionnaire de physique, PUF, 1988, page 304.

- [36] Jim Branson, Quantum Physics, (UCSD Physics 130) 2003.
- [37] J. V. Lambers, Explorations in numerical analysis, WSPC 2018.

Résumé :

Ce travail est consacré au calcul variationnel de l'état fondamental de l'atome d'Hélium à l'aide de la méthode d'optimisation de la descente du gradient en plus de la méthode d'intégration multidimensionnelle de Monte Carlo. Dans les deux premiers chapitres, après une présentation détaillée des deux méthodes et comparaison des différentes variantes de Monte Carlo, Dans le troisième chapitre nous appliquons, avec succès, la méthode variationnelle à deux paramètres avec une fonction d'onde test issue du modèle des électrons indépendants. Le calcul effectué à l'aide d'un programme Maple montre des résultats satisfaisants et une convergence rapide des paramètres variationnels et de l'énergie vers leurs valeurs exactes à condition d'utiliser un petit pas δ et un grand nombre de points d'intégration afin de minimiser les fluctuations stochastiques.

Mots clés : Atome He, Descente Gradient, Intégration MC, Méthode variationnelle.

Abstract :

This work is devoted to the variational calculation of the ground state of the Helium atom using the gradient descent optimization method in addition to the multidimensional Monte Carlo integration technique. In the first two chapters, after a detailed presentation of the two methods and comparison of the different variants of Monte Carlo, in the third chapter we successfully apply the two-parameter variational method with a test wave function from the independent electrons model. The calculation carried out using a Maple program shows satisfactory results and a rapid convergence of the variational parameters and the energy towards their exact values provided that a small step δ and a large number of integration points are used in order to minimize stochastic fluctuations.

Keywords : He atome, Gradient descent, MC intégration, Variational method

المخلص:

هذا العمل مخصص للحساب التفاضلي للحالة الأساسية لذرة الهليوم بالاستعانة بطريقة التحسين النزول بالتدرج إضافة إلى طريقة التكامل متعددة الأبعاد مونت كارلو. في الفصلين الأولين قمنا بعرض مفصل للطريقتين ومقارنة مختلف المتغيرات مونت كارلو. في الفصل الثالث طبقنا بنجاح طريقة التفاضل ثنائية المتغيرات مع دالة اختبار موجية مستخرجة من نموذج الإلكترونات المستقلة. نفذت الحسابات باستخدام برنامج Maple الذي أظهر نتائج مرضية وتقارب سريع للمتغيرات التفاضلية والطاقة إلى قيمتها الدقيقة بشرط استخدام خطوة صغيرة وعدد كبير من نقاط التكامل من أجل التقليل من التقلبات العشوائية.

كلمات مفتاحية: ذرة الهليوم، النزول بالتدرج ، تكامل مونت كارلو، لحساب التفاضلي