
A low-order scaling relativistic correlated electronic structure method for excited states of molecules

Maître de stage :

ANDRE SEVERO PEREIRA GOMES

andre.gomesuniv-lille.fr

Etudiant :

THOMAS LEJEUNE

thomas.lejeune@ens-paris-saclay.fr

LICENCE 3 DOUBLE DIPLÔME
VOIE FRÉDÉRIC JOLIOT-CURIE

22/05/23 - 14/07/23

Résumé

Le travail présenté dans ce rapport concerne l'implémentation de la méthode CIS(D) dans le code DIRAC. Cela débute par un travail théorique des différentes méthodes, telles que CIS et MP2, qui permettent de comprendre les équations de CIS(D). Avant d'implémenter la méthode en Fortran dans le code DIRAC, un prototype en Python a été développé et les résultats obtenus ont été vérifiés. Cette étape a permis d'établir une implémentation de référence et de se familiariser avec les outils de chimie quantique nécessaires pour mener à bien le projet. Afin d'effectuer les calculs, un cluster haute performance a été utilisé. Les résultats de la méthode CIS(D) présentés ici montrent une nette amélioration des performances de la méthode CIS.

Remerciements

Je tiens tout particulièrement à remercier André Gomes pour sa bienveillance, ses précieux conseils et enseignements ainsi que pour son encadrement durant ce stage. Je désire également remercier l'équipe PCMT et plus généralement le laboratoire PhLAM pour son chaleureux accueil, sans qui ce stage n'aurait pas été possible.

Summary :

The work presented in this report concerns the implementation of the CIS(D) method in the DIRAC code. It begins with a theoretical study of various methods, such as CIS and MP2, which help to understand the CIS(D) equations. Before implementing the method in Fortran in the DIRAC code, a Python prototype was developed, and the obtained results were verified. This step allowed establishing a reference implementation and becoming familiar with the quantum chemistry tools required to carry out the project. To perform the calculations, a high-performance cluster was used. The CIS(D) results presented here show a clear improvement in the performance of the CIS method.

Acknowledgment :

I would like to express my sincere gratitude to André Gomes for his kindness, valuable advice, and teachings, as well as for his guidance during this internship. I also want to thank the PCMT team and, more generally, the PhLAM laboratory for their warm welcome, without which this internship would not have been possible.

Table des matières

1	Présentation du PhLAM	4
2	Introduction et contexte des travaux de recherche	5
3	Aspects théorique	5
3.1	Approximations et modèles	5
3.2	Théorie de la perturbation de Møller-Plesset (MP2)	6
3.3	Configuration Interaction	6
3.3.1	Configuration Interaction Single ou Interaction de Configurations avec excitations simples (CIS)	6
3.3.2	Correction CIS-MP2	7
3.3.3	Correction CIS(D)	7
4	Détails computationnels et instrumentation	7
4.1	Propotype sur Psi4	8
4.2	Vérification des méthodes	8
4.3	Implémentation dans le code DIRAC	8
4.4	Outils informatiques	9
5	Résultats et discussions	10
5.1	Résultats CIS et CIS(D)	10
5.2	Effet de la corrélation électronique	11
5.3	Observation d'effets relativistes	12
6	Conclusion et perspectives	13
A	Equations Hartree-Fock	15
B	Équations de la théorie MP2	15
C	Équations des méthodes CIS, CIS-MP2 et CIS(D)	16

Liste des abréviations

CI Configuration Interaction

EOM-CCSD Equation Of Motion - Coupled Cluster Singles and Doubles

TDHF Time-Dependant Hartree-Fock ou Hartree-Fock dépendant du temps

CC Coupled Cluster

CIS Configuration Interaction Single ou Interaction de Configurations avec excitations simples

CIS(D) Configuration Interaction Singles with perturbatives Doubles

CISD Configuration Interaction Singles and Doubles

MP2 Théorie de la perturbation de Møller-Plesset

PhLAM laboratoire de Physique des Lasers, Atomes et Molécules

1 Présentation du PhLAM

Ce stage de recherche a été effectué dans le laboratoire de Physique des Lasers, Atomes et Molécules (PhLAM). Il s'agit d'une unité mixte de recherche (UMR 8523), sous la double tutelle du CNRS et de l'Université de Lille. Le PhLAM est né en janvier 1998 de la fusion du Laboratoire de Spectroscopie Hertzienne (LSH) et du Laboratoire de Dynamique Moléculaire et Photonique (LDMP). Au 1er janvier 2023, le laboratoire compte 173 agents, dont 71 Chercheurs et Enseignants-Chercheurs et 51 Doctorants. Les travaux du laboratoire sont orientés autour de l'interaction lumière-matière, au travers de 5 axes majeurs :

- La Physique des Atomes Refroidis par Laser (Atomes Froids),
- La DYnamique des Systèmes COMplexes (DYSCO)
- La Physico-Chimie Moléculaire Théorique (PCMT)
- La Photonique
- la Physique Moléculaire aux Interactions (PMI)

J'ai effectué mon stage au sein de l'équipe PCMT, qui se concentre sur trois axes de recherche : la dynamique de systèmes moléculaires complexes, la modélisation moléculaire des éléments lourds et les développements méthodologiques et d'outils informatiques. Mon travail portait spécifiquement sur cette dernière thématique, et plus particulièrement sur les méthodes corrélées relativistes de structure électronique.

Les travaux effectués par l'équipe PCMT sur les méthodes corrélées relativistes ont pour objectif de développer des approches de structure électronique décrivant les effets de corrélation électronique pour l'ensemble des systèmes de la classification périodique. Les méthodes employées sont basées sur la résolution de l'équation de Dirac pour les électrons. L'équipe participe activement au développement du code Dirac (<http://diracprogram.org>), dans laquelle elle implémente des méthodes comme par exemple la méthode Coupled Cluster (CC).

En juin 2023, un article de l'équipe PCMT a été sélectionné comme "Featured Article" et faisant la couverture de *Inorganic Chemistry*. Cet article rapporte et analyse pour la première fois les propriétés de luminescence d'un complexe d'uranyle(VI) coudé. Dans les dérivés d'Uranium, les effets relativistes sont très importants et ne peuvent être négligés. L'équipe a ici utilisé le code ADF (www.scm.com) qui implémente des hamiltoniens relativistes mais plutôt pour des calculs DFT.

2 Introduction et contexte des travaux de recherche

En chimie quantique, l'aspect informatique est d'une importance primordiale. Les modélisations et les calculs visent à reproduire les résultats expérimentaux pour favoriser une meilleure compréhension. Ces méthodes couvrent un pannel large d'applications, allant de l'optimisation de géométrie à la détermination de grandeurs thermodynamiques, en passant par la simulation des propriétés spectroscopiques.

C'est dans ce but qu'il est pertinent de vouloir s'intéresser à la modélisation d'états excités. En effet, l'étude de ces états permet le tracé de spectres de spectroscopie ou encore la modélisation de réactions photochimiques, particulièrement lorsque l'un des réactifs passe par un intermédiaire excité.

De plus, un enjeu important en chimie quantique est la modélisation d'éléments lourds. Des effets relativistes deviennent significatifs lorsque l'on s'intéresse à ce type d'éléments. C'est le cas par exemple du couplage spin-orbite ou de la contraction des contractions ou expansions des amplitudes des orbitales [1]. Les travaux portants sur les éléments lourds sont particulièrement important dans le domaine nucléaire, ce qui relève d'un enjeu écologique important de nos jours.

Cependant, les expériences avec des éléments radioactifs, très lourds, sont complexes. En effet, des expériences avec de tels systèmes nécessitent du matériel de protection et des procédures de sécurités très contraignantes. Puisque ce sont des systèmes avec un grand nombre d'électrons, l'application de méthodes précises mais coûteuses de chimie quantique va se révéler ardue. C'est pourquoi l'un des défis est de trouver des méthodes avec une faible complexité permettant de modéliser ces éléments lourds. On s'intéresse ici tout particulièrement à la modélisation d'états excités de ces éléments.

Afin de répondre à ces problématiques, le PhLAM développe le code DIRAC, permettant du calcul de chimie quantique basé sur des hamiltoniens relativistes. Des enjeux restants sont notamment la mise en œuvre et l'optimisation de méthodes dans ce code.

La méthode CIS(D) est une méthode perturbative à faible complexité permettant de modéliser des états excités. L'objectif du stage est alors d'implémenter cette méthode dans le code DIRAC.

3 Aspects théorique

Cette section vise à présenter le cheminement permettant d'arriver à la méthode CIS(D) qui est la méthode d'intérêt du stage.

3.1 Approximations et modèles

Durant toute l'étude théorique, nous nous plaçons dans le cadre de l'approximation de Born-Oppenheimer. Les électrons de nos systèmes sont modélisés par des spinorbitales et afin de traiter les systèmes poly-électroniques, on se place dans le cadre du modèle des déterminants de Slater.

Les méthodes présentées par la suite sont des méthodes post-Hartree-Fock. La méthode Hartree-Fock [2] est une méthode numérique approchée pour la résolution de l'équation de

Schrödinger basée sur le principe variationnel. A partir des spinorbitales, on définit l'opérateur de Fock et sa représentation matricielle, qui sera par la suite utilisées dans les méthodes post-Hartree-Fock.

3.2 Théorie de la perturbation de Møller-Plesset (MP2)

La méthode Hartree-Fock nous permet d'obtenir un premier résultat de l'énergie de l'état fondamental de notre système. Cependant, cette méthode ne prend pas en compte la corrélation électronique. Une méthode pour résoudre ce problème est d'appliquer la théorie de la perturbation de Rayleigh-Schrödinger où l'Hamiltonien non perturbé serait l'opérateur de Fock. Le potentiel perturbatif est le potentiel de corrélation, i.e la différence entre l'Hamiltonien réel du système et l'opérateur de Fock. L'expression de ce potentiel est développée dans l'équation 1, où λ est un paramètre sans dimension pouvant prendre des valeurs allant continûment de 0 (pas de perturbation) à 1 (perturbation totale). La méthode MP2 [3] consiste à développer notre fonction d'onde sur une base de déterminants doublement excités lors du calcul de la correction en énergie au second ordre. Les équations liées à cette méthode sont expliquées dans l'appendix A.

$$\hat{H} = \hat{F} + \lambda \hat{V} = \hat{F} + \lambda \left(\sum_{j>i} \frac{1}{r_{ij}} - \nu_{HF} \right) \quad (1)$$

3.3 Configuration Interaction

Les méthodes Configuration Interaction (CI) sont des méthodes post-Hartree-Fock variationnelles permettant de calcul des propriétés d'états excités [2]. La fonction d'onde est représentée par une combinaison linéaire de déterminants de Slater excités, prenant le déterminant Hartree-Fock comme référence, comme dans l'équation 2. Le détail des notations est développé dans l'appendix C.

$$|\Psi\rangle = c_0 |\Phi_0\rangle + \sum_{ia} c_i^a |\Phi_i^a\rangle + \sum_{ijab} c_{ij}^{ab} |\Phi_{ij}^{ab}\rangle + \sum_{ijkabc} c_{ijk}^{abc} |\Phi_{ijk}^{abc}\rangle + \dots \quad (2)$$

Pour former un déterminant excité, on substitue une ou plusieurs orbitales occupées par une ou plusieurs orbitales virtuelles dans le déterminant de l'état fondamental. On peut choisir le nombre de substitutions en fonction de la précision souhaitée et des capacités de calcul dont on dispose. Les méthodes CI mènent à la détermination de valeurs propres et de vecteurs propres d'un Hamiltonien.

3.3.1 Configuration Interaction Single ou Interaction de Configurations avec excitations simples (CIS)

L'approche la plus simple est celle où ne sont considérées que les mono-excitations. Cette méthode a l'avantage d'être peu coûteuse et rapide à mettre en oeuvre, cependant elle peut être très imprécise. L'objectif est de calculer les éléments de matrice du type $\langle \Phi_i^a | \hat{H} | \Phi_j^b \rangle$. Les valeurs propres de la matrice obtenue sont les énergies d'excitations de notre système et les vecteurs propres donneront la contribution de chaque déterminant à l'excitation. Les vecteurs propres représentent un état excité dans une base de déterminants (ici mono-excités) donnée.

On utilise les règles de Slater-Condon [4] et on trouve le résultat de l'équation 3. Les différents termes de l'équation sont développés dans l'appendix C.

$$\langle \Phi_i^a | H - E_0 \mathbb{1} | \Phi_j^b \rangle = (\epsilon_a - \epsilon_i) \delta_{ij} \delta_{ab} + \langle a j || i b \rangle \quad (3)$$

3.3.2 Correction CIS-MP2

Afin de prendre l'effet de la corrélation électronique en compte pour l'état fondamental, la méthode MP2 est souvent employée de part sa faible complexité en $\mathcal{O}(N^5)$, avec N le nombre de fonctions de base. La correction CIS-MP2 [5] tend à corriger les énergies d'excitations provenant de la méthode CIS en assumant que la même physique régit les états excités et l'état fondamental pour la théorie MP2. Le fait que le théorème de Brillouin ne s'applique qu'à l'état fondamental rend cette correction coûteuse. En effet, on doit ici tenir compte des excitations simples et doubles par rapport à l'état CIS donc des doubles et triples excitations par rapport à l'état fondamental. La complexité de cette correction est en $\mathcal{O}(N^6)$. Les notations et équations de base de cette méthode sont rappelées dans l'appendix C

3.3.3 Correction CIS(D)

On cherche par la suite à approximer la correction CIS-MP2 pour baisser la complexité et donc le temps de calcul. La double substitution à partir de l'état CIS (donc triple substitution à partir de l'état fondamental) mets en jeu des électrons principalement inactifs dans l'excitation électronique. Ainsi, c'est ce terme que l'on va chercher à simplifier. Pour cela, en considérant que la première excitation ne va pas interagir avec les suivantes, c'est à dire que les électrons mis en jeux sont considérés inactifs dans l'excitation électronique, on va se servir des coefficients MP2 de l'état fondamental. Cette approximation va permettre de réduire drastiquement le temps de calcul. La méthode sous-jacente à cette approximation est la méthode Configuration Interaction Singles with perturbatives Doubles (CIS(D)) [5]. Les notations et équations de base de cette méthode sont rappelées dans l'appendix C

Cette correction a une complexité en $\mathcal{O}(N^5)$, ce qui la rend moins coûteuse que la correction CIS-MP2 et que la méthode Configuration Interaction Singles and Doubles (CISD). Cette méthode corrige les énergies et les coefficients obtenus par la méthode Hartree-Fock mais a une complexité en $\mathcal{O}(N^6)$ et ne respecte pas la séparabilité stricte. De plus, cette méthode respecte la propriété de séparabilité stricte contrairement à la correction CIS-MP2, qui peut être une propriété recherché si on souhaite modéliser une réaction photochimique ou encore une approche de deux réactifs, dont l'un est dans un état excité.

4 Détails computationnels et instrumentation

L'objectif est d'implémenter la méthode CIS(D) dans le code DIRAC [6], mais cela requiert une approche en plusieurs étapes. Tout d'abord, un prototype a été développé en Python en utilisant le module de chimie quantique Psi4 [7]. Ensuite, une validation des résultats a été effectuée en les comparant avec d'autres codes tels que Gaussian. Une fois cette validation réalisée, l'implémentation dans le code DIRAC en Fortran peut être envisagée.

4.1 Propotype sur Psi4

Ce prototype a plusieurs objectifs : familiarisation avec la programmation en chimie quantique, compréhension des algorithmes et fonctions utilisés dans de tels programmes afin de pouvoir implémenter des méthodes déjà établies avec des résultats de référence (RHF, UHF, MP2), et enfin, l'implémentation des méthodes d'intérêt (CIS, CIS(D)). Ce prototype permettra par la suite de confirmer les résultats du code Fortran.

Le module Psi4 est complété par un module nommé Psi4Numpy [8], permettant l'utilisation de l'infrastructure de chimie quantique de Psi4 comme le calcul d'intégrales électroniques, en combinaison avec les méthodes numériques, notamment d'algèbre linéaire, implémentées dans le module Numpy. Ce projet contient de nombreuses implémentations pouvant servir de référence (sauf CIS(D)), ainsi que des tutoriaux permettant d'appréhender plus facilement les outils mis à notre disposition, et de découvrir de nouvelles méthodes.

Parmi les outils du module numpy, la fonction `np.einsum`, qui permet de réaliser des opérations tensorielles de façon très efficace, est particulièrement utile pour la chimie quantique car c'est une discipline manipulant souvent des tenseurs. Cette fonction possède une syntaxe très similaire à celles des contractions de tenseurs utilisées dans le code DIRAC, c'est pourquoi il s'agit d'une méthode très intéressante à maîtriser. Initialement, une implémentation naïve des contractions tensorielles a été effectuée en utilisant des boucles pour bien comprendre les algorithmes. Puis une optimisation temporelle (optimisation du temps d'exécution) et spatiale (optimisation de la quantité de mémoire RAM utilisée) a été réalisée ultérieurement en utilisant le module numpy.

Les tutoriaux Psi4Numpy sont disponibles gratuitement sur le site : <https://github.com/psi4/psi4numpy>. Des exemples de prototypes Psi4 effectués pendant ce stage sont disponibles sur le site : <https://github.com/ThomasLjn/phalm-13>

4.2 Vérification des méthodes

La deuxième étape était de vérifier les résultats que nous donnaient mes codes Psi4. Pour se faire, les méthodes implémentées par défaut dans Psi4 et les codes Gaussian et DIRAC ont été utilisées. C'était également une première familiarisation avec les fichiers d'entrées et de sortie de Gaussian et de DIRAC.

Afin d'avoir des résultats comparables, il fallait s'assurer de mettre les mêmes paramètres en entrée. Lors de différences entre les résultats, plusieurs points peuvent être regardés afin de trouver d'où vient l'erreur. Dans un premier temps, on peut comparer les résultats donnés par la méthode Hartree-Fock, puis la répulsion nucléaire. D'autres paramètres comme le modèle du noyau (charge ponctuelle ou distribution gaussienne de charge), le seuil des intégrales ou encore les valeurs des constantes peuvent être des facteurs provoquant des différences de valeurs.

4.3 Implémentation dans le code DIRAC

Une fois le prototype réalisé et validé, l'objectif consistait à intégrer la méthode CIS(D) dans le code DIRAC en Fortran. Étant donné que la méthode CIS était déjà disponible dans une version privée de développement de l'équipe lilloise du code DIRAC, seule la correction CIS(D) a été ajoutée. En plus de l'implémentation de la méthode, un test a été ajouté au système

de test de DIRAC. Cela permet d'avoir un résultat de référence (ici non relativiste) pour vérifier que le code fonctionne correctement.

L'implémentation de la correction CIS(D) s'est déroulée en plusieurs étapes. Après s'être familiarisé avec le langage Fortran, la première étape consistait à ajouter une syntaxe reconnue dans les fichiers d'entrée du code DIRAC. Ces fichiers d'entrée fonctionnent avec des mots-clés, et l'objectif était d'introduire un nouveau mot-clé permettant d'indiquer quels niveaux d'énergie devaient être corrigés.

Ensuite, il a été nécessaire de comprendre en détail le code CIS existant, car il servait d'exemple pour l'utilisation des bibliothèques de calcul tensoriel similaires à `np.einsum`, ici la bibliothèque `TAL_SH` (https://github.com/DmitryLyakh/TAL_SH). Une documentation détaillée de cette bibliothèque est disponible en ligne, ce qui a grandement facilité son utilisation. Néanmoins, il était important de noter que cette bibliothèque présentait quelques différences avec les syntaxes proposées en python. Par exemple, les tenseurs devaient être créés et initialisés avant de pouvoir effectuer des contractions, additions ou multiplications par un scalaire, chaque opération ayant sa propre syntaxe. Lors d'une contraction, les indices se répétant doivent forcément être sommés. Par exemple, la syntaxe `"D(a,b,c,d,i)+=L(c,i,j,a)*R(b,j,d,i)"` n'est pas valide, car le `i` devrait être sommé et donc ne plus apparaître dans le tenseur final. C'est une différence majeure par rapport à la fonction `np.einsum` qui change quelques aspects du code.

Ainsi, en Fortran, le problème du calcul des corrections peut être traité de deux manières. Soit on stocke les corrections pour tous les états considérés dans un grand tableau, soit on les calcule les unes à la suite des autres. Cette dernière option a été privilégiée car elle permet d'utiliser beaucoup moins de mémoire.

Une autre différence majeure était liée au caractère relativiste des calculs dans le code DIRAC, qui engendre des grandeurs avec des parties imaginaires non nulles car les fonctions d'ondes deviennent complexes. En python, les calculs étaient forcément non relativistes, et donc toutes les grandeurs étaient réelles. Par conséquent, lors de la mise en œuvre des calculs en Fortran, une attention particulière a été accordée à l'utilisation de la conjugaison ou de l'adjoint des matrices quand nécessaire, par exemple pour les intégrales biélectroniques antisymétrisées :

$$\langle ab || ij \rangle = \langle ij || ab \rangle^*.$$

4.4 Outils informatiques

Le laboratoire PhLAM possède un cluster informatique haute performance : Sakura. Les membres du laboratoire peuvent s'y connecter via un accès ssh et lancer des scripts linux dessus. Un système de file d'attente a été mis en place afin de répartir les différents calculs sur plusieurs processeurs et de gérer la quantité de mémoire allouée à un calcul donné. Sakura a été utilisé dans le cadre de ce stage afin de faire tourner des calculs Gaussian ou DIRAC, trop gourmands pour fonctionner sur un ordinateur personnel. De plus, l'utilisation de l'outil git a été abordé. Il s'agit d'un outil de travail collaboratif permettant de garder une trace du travail effectué, ainsi que de revenir à des versions ultérieures du code. Cet outil était particulièrement pratique car permettant de synchroniser le code source sur l'ordinateur de travail et sur Sakura.

5 Résultats et discussions

5.1 Résultats CIS et CIS(D)

Les résultats des prototypes Psi4 et du code DIRAC ont été comparés au code Gaussian utilisé comme référence. Les programmes montrent une cohérence parfaite dans l'ensemble, bien que de petites différences puissent être attribuées à des méthodes de calcul différentes, pouvant entraîner quelques approximations, notamment lors du calcul des intégrales ou de la diagonalisation de l'hamiltonien. Des essais ont été réalisés sur divers systèmes tels que le dihydrogène, l'hélium, l'eau, le formaldéhyde et le zinc. On reporte dans les tableau 1 et 2 les écarts entre les prototypes Psi4 ou le code DIRAC et le résultat fourni par Gaussian en utilisant une même base.

Numéro de l'état excité	Gaussian-Psi4	Gaussian-DIRAC	Psi4-DIRAC
1 (Triplet)	4.07E-08	4.35E-08	2.82E-09
4 (Singulet)	3.70E-11	2.63E-09	2.59E-09
5 (Triplet)	2.32E-08	2.59E-08	2.70E-09
8 (Triplet)	6.63E-09	1.05E-08	3.84E-09

TABLE 1 – Écarts absolus entre les énergies CIS (Hartree) calculées par les différents programmes pour la molécule d'eau, base Aug-cc-PVDZ. Les résultats donnés par les codes Psi4 et DIRAC sont ceux des codes développés pendant ce stage.

Numéro de l'état excité	Gaussian-Psi4	Gaussian-DIRAC	Psi4-DIRAC
1 (Triplet)	6.45E-06	6.10E-06	3.50E-07
4 (Singulet)	7.84E-07	5.01E-07	2.82E-07
5 (Triplet)	5.83E-06	5.48E-06	3.48E-07
8 (Triplet)	4.77E-06	4.40E-06	3.71E-07

TABLE 2 – Écarts absolus entre les énergies CIS(D) (Hartree) calculées par les différents programmes pour la molécule d'eau, base Aug-cc-PVDZ. Les résultats donnés par les codes Psi4 et DIRAC sont ceux des codes développés pendant ce stage.

On remarque des écarts de l'ordre de la précision demandée pour le SCF, soit de l'ordre de 10^{-7} . Des écarts un peu plus importants peuvent s'expliquer pour la méthode CIS(D) par une accumulation d'erreurs.

Par la suite, les résultats des méthodes CIS et CIS(D) sont comparées aux valeurs expérimentales données par le NIST (physics.nist.gov). On remarque sur l'atome de Zinc que les résultats CIS(D) sont toujours plus proches de la valeur expérimentale que les résultats CIS. Cela indique que par la prise en compte de la corrélation électronique absente dans CIS, la méthode CIS(D) est un modèle plus fidèle des processus physiques. Ces résultats sont résumés dans le tableau 3. On remarque une similarité d'erreurs entre les deux états triplets (environ 0.5eV), l'erreur est un peu plus basse pour l'état singulet.

Numéro de l'état excité	CIS	CIS(D)	NIST
1 (Triplet)	2.67559	3.51757	4.00609
4 (Singulet)	5.01666	5.48995	5.79569
5 (Triplet)	5.28640	6.06650	6.65450

TABLE 3 – Comparaison des énergies d'excitations (eV) du Zinc, base Aug-cc-pVDZ, entre les méthodes CIS, CIS(D) du code DIRAC développées pendant ce stage, hamiltonien non relativiste, et les résultats expérimentaux du NIST.

On a donc pu observer que notre méthode s'applique correctement à des hamiltoniens non relativistes, et que l'on arrive à reproduire les résultats donnés par le code Gaussian. Cependant, lorsque l'on souhaite effectuer des calculs avec le code DIRAC et un hamiltonien relativiste (X2C) [9], les résultats ne sont pas reproductibles. La source de ce comportement est encore inconnue. On remarque que les énergies d'excitations CIS et les résultats Hartree-Fock sont reproductibles : si on lance plusieurs fois le calcul avec le même fichier d'entrée, on obtient les mêmes énergies en sortie. Cependant, ce n'est pas le cas pour les énergies CIS(D) avec un hamiltonien X2C. Cela provient du fait que les vecteurs propres de l'hamiltonien peuvent changer énormément d'un calcul à l'autre, ce qui provoque de grosses variations lors du calcul de la correction. Les résultats en sortie du code DIRAC sont récapitulés dans tableau 4. Ils proviennent de 3 calculs possédant la même entrée. On remarque que les énergies ne sont plus rangées par ordre croissant, alors qu'elles le sont pour la méthode CIS. Les deux premiers états sont dégénérés entre eux (donc même énergie CIS) alors que le quatrième (troisième ligne du tableau) ne l'est pas, c'est pourquoi l'énergie CIS est différente.

Numéro de l'état excité	Calcul 1	Calcul 2	Calcul 3	Énergie CIS
1	2.81252	3.67352	2.85094	2.80950
2	2.83351	3.32588	2.88346	2.80950
4	2.82620	2.99962	2.84806	2.83099

TABLE 4 – Energie d'excitations du dihydrogène (eV) calculés par la méthode CIS(D), l'hamiltonien X2C et la base dyall.av2z

5.2 Effet de la corrélation électronique

Comme expliqué dans la section 3, la méthode Hartree-Fock ne prend pas en compte la corrélation électronique. Cependant, la correction CIS(D) tient compte de ce phénomène, mais de façon moins précise que pour d'autres méthodes comme Equation Of Motion - Coupled Cluster Singles and Doubles (EOM-CCSD) [10] qui est une autre méthode tenant compte de la corrélation électronique qui permet le calcul des énergies d'excitations. C'est pourquoi il peut être intéressant de comparer la performance de ces différentes méthodes. Pour calculer des énergies d'excitations, une méthode dérivée de Hartree-Fock existe, il s'agit de la méthode Time-Dependant Hartree-Fock ou Hartree-Fock dépendant du temps (TDHF). De part les similitudes entre les méthodes TDHF [11] et CIS, on s'attend à avoir des résultats comparables pour ces deux méthodes. On choisit de comparer les méthodes TDHF, CIS et CIS(D) à la méthode

EOM-CCSD. Le tableau 5 présente les résultats donnés par ces méthodes sur l'atome de Zinc. Sur un système avec un nombre important d'électrons comme le Zinc, la corrélation électronique n'est plus négligeable. Cela se traduit sur les résultats, où la méthode Hartree-Fock donne des résultats bien plus éloignés de l'expérience que les méthodes CIS(D) ou EOM-CCSD. On voit également que pour ce système, CIS(D) a une performance tout à fait similaire à EOM-CCSD pour les deux premiers états électroniques.

On voit néanmoins que même EOM-CCSD n'est pas tout à fait en accord avec les résultats expérimentaux, cela se doit tout d'abord aux bases utilisées, mais aussi au fait que cette méthode ne considère pas les excitations au delà des doubles.

Numéro de l'état excité	TDHF ^a	CIS ^a	CIS(D) ^a	EOM-CCSD ^b
1	2.030	2.675	3.515	3.739
2	4.868	5.015	5.489	5.587
3	5.796	5.284	6.065	6.661

TABLE 5 – Énergies d'excitations (eV) du zinc par différentes méthodes, base Aug-cc-pVDZ. a : DIRAC, b : Gaussian, hamiltonien non relativiste.

Une autre manière de rendre compte de la corrélation électronique est de comparer les méthodes Hartree-Fock et MP2, cette dernière visant à apporter cet effet. Les résultats des méthodes Hartree-Fock et MP2 calculées à partir du logiciel Gaussian pour quelques systèmes sont référencées dans le tableau 6. On remarque que la différence entre les résultats des deux méthodes augmente avec la taille du système, ce qui est attendu puisque la corrélation électronique correspond aux interactions qu'à un électron avec ceux qui l'entourent, s'il y a beaucoup d'électrons, on s'attend à une énergie de corrélation plus importante.

Système	HF (Hartree)	MP2 (Hartree)	Énergie de corrélation (Hartree)
H2O	-76.04	-76.26	0.22
Formaldéhyde	-113.88	-114.22	0.34
Zinc	-1777.85	-1778.31	0.46

TABLE 6 – Énergies de l'état fondamental de différents systèmes, calculs réalisés avec le logiciel Gaussian et la base Aug-cc-pVDZ

5.3 Observation d'effets relativistes

Un effet de la relativité est le couplage spin-orbite. On peut voir ce couplage comme un effet provenant d'un effet d'induction, si on se place dans le référentiel d'un électron. Dans ce cas, l'électron d'intérêt est fixe, et les noyaux (et électrons, le cas échéant) se déplacent autour. On a ainsi un champ magnétique induit par le mouvement des autres particules, qui interagit avec le champ magnétique intrinsèque de l'électron (son spin). Pour des atomes lourds, ce couplage devient plus important que la corrélation électronique [12]. Alors, les règles de sélection des transitions changent. Là où pour des atomes légers, le couplage LS peut être pertinent, pour des atomes lourds le couplage spin-orbite peut être trop important et ainsi un autre couplage

comme le couplage jj (ou un couplage intermédiaire) peut être utilisé afin de rendre compte de l'expérience.

Le zinc, le cadmium et le mercure sont situés dans la même colonne du tableau périodique. Ainsi, leurs propriétés spectroscopiques sont semblables (ils possèdent des états électroniques similaires). La figure 1 représente des spectres d'absorbance pour ces trois atomes. Les énergies d'excitation ainsi que les forces d'oscillateurs ont été calculées par la méthode TDHF du code DIRAC. On utilise ici l'hamiltonien X2C afin de pouvoir observer les effets de la relativité sur nos systèmes. Pour ces trois atomes, la transition $^1P_1 \leftarrow ^1S_0$ est autorisée avec un grand moment de transition, cependant, la transition $^3P_1 \leftarrow ^1S_0$ n'est observable que pour le mercure. En effet, cette transition va à l'encontre des règles de transitions du couplage LS ($\Delta S \neq 0$), mais est autorisée en couplage jj. C'est ainsi que pour le mercure, on observe l'importance de la relativité pour le tracé de ce spectre. Formellement, la transition n'est plus interdite dès lors où le couplage jj peut s'appliquer, mais les moments de transitions restent trop faibles pour le zinc et le cadmium pour être observables.

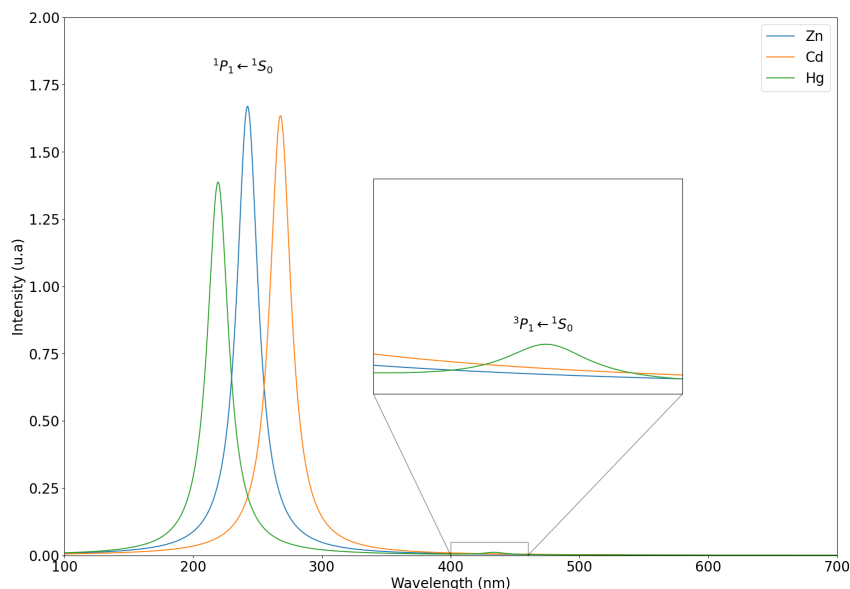


FIGURE 1 – Spectre d'absorbance calculé par la méthode TDHF du code DIRAC, base dyall.av2z et un hamiltonien relativiste (X2C)

6 Conclusion et perspectives

Les objectifs de ce stage étaient d'avoir une première expérience avec les outils de développement en chimie quantique, et des méthodes théoriques sous-jacentes et l'implémentation de la méthode CIS(D) dans le code DIRAC. Actuellement, le code CIS(D) fonctionne pour des hamiltoniens non relativistes mais ne semble pas encore déterministe pour des hamiltoniens relativistes. A terme, une fois ce problème réglé, la méthode sera rendue publique dans le code DIRAC.

D'un point de vue personnel, ce stage fut très enrichissant, d'autant sur le point de vue théorique que technique. Les méthodes présentées dans ce rapport ont été découvertes durant le stage, ainsi que les outils de chimie quantique présentés (psi4, gaussian, DIRAC). Cela m'a aussi permis d'en apprendre plus sur monde de la recherche et sur son fonctionnement.

Références

- [1] E. J. BAERENDS et al. “Relativistic atomic orbital contractions and expansions : magnitudes and explanations”. In : *Journal of Physics B : Atomic, Molecular and Optical Physics* 23.19 (oct. 1990), p. 3225. DOI : 10.1088/0953-4075/23/19/010.
- [2] Attila SZABO et Neil S. OSTLUND. *Modern Quantum Chemistry : Introduction to Advanced Electronic Structure Theory*. Reprint Edition. Mineola, N.Y : Dover Publications, 2 juill. 1996. 480 p.
- [3] Dieter CREMER. “Møller-Plesset perturbation theory : from small molecule methods to methods for thousands of atoms : Møller-Plesset perturbation theory”. In : *Wiley Interdisciplinary Reviews : Computational Molecular Science* 1.4 (juill. 2011), p. 509-530. DOI : 10.1002/wcms.58.
- [4] C David SHERRILL. “Derivation of the Configuration Interaction Singles (CIS) Method for Various Single Determinant References and Extensions to Include Selected Double Substitutions (XCIS)”. In : (1996).
- [5] Martin HEAD-GORDON et al. “A doubles correction to electronic excited states from configuration interaction in the space of single substitutions”. In : *Chemical Physics Letters* 219.1 (mars 1994), p. 21-29. DOI : 10.1016/0009-2614(94)00070-0.
- [6] Trond SAUE et al. “The DIRAC code for relativistic molecular calculations”. In : *The Journal of Chemical Physics* 152.20 (29 mai 2020), p. 204104. DOI : 10.1063/5.0004844.
- [7] Daniel G. A. SMITH et al. “PSi4 1.4 : Open-source software for high-throughput quantum chemistry”. In : *The Journal of Chemical Physics* 152.18 (14 mai 2020), p. 184108. DOI : 10.1063/5.0006002.
- [8] Daniel G. A. SMITH et al. “Psi4NUMPY : An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development”. In : *Journal of Chemical Theory and Computation* 14.7 (10 juill. 2018), p. 3504-3511. DOI : 10.1021/acs.jctc.8b00286.
- [9] Trond SAUE. “Relativistic Hamiltonians for Chemistry : A Primer”. In : *ChemPhysChem* 12.17 (9 déc. 2011), p. 3077-3094. DOI : 10.1002/cphc.201100682.
- [10] Rodney J. BARTLETT et Monika MUSIAL. “Coupled-cluster theory in quantum chemistry”. In : *Rev. Mod. Phys.* 79 (1 fév. 2007), p. 291-352. DOI : 10.1103/RevModPhys.79.291.
- [11] Andreas DREUW et Martin HEAD-GORDON. “Single-Reference ab Initio Methods for the Calculation of Excited States of Large Molecules”. In : *Chemical Reviews* 105.11 (1^{er} nov. 2005), p. 4009-4037. DOI : 10.1021/cr0505627.
- [12] Lilian GUILLEMENEY. *Bases de chimie théorique*. URL : http://perso.ens-lyon.fr/lilian.guillemeney/wp-content/uploads/2021/10/AgCh_Atomistique_2021.pdf.

A Equations Hartree-Fock

Dans le cadre de la théorie Hartree-Fock [2], on définit :

— L'hamiltonien mono électronique :

$$\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} \quad (4)$$

— L'opérateur d'échange K et l'opérateur d'interaction coulombienne J, où χ est une fonction de la base atomique choisie :

$$K_b(1)\chi_a(1) = \left[\int dx_2 \chi_b^*(2) r_{12}^{-1} \chi_a(2) \right] \chi_b(1) \quad (5)$$

$$J_b(1)\chi_a(1) = \left[\int dx_2 \chi_b^*(2) r_{12}^{-1} \chi_b(2) \right] \chi_b(1) \quad (6)$$

— Le potentiel Hartree-Fock :

$$\nu_{HF}\chi_a(i) = \left[\sum_{b \neq a} J_b(i) - K_b(i) \right] \chi_a(i) \quad (7)$$

— L'opérateur de Fock et son équations aux valeurs propres associée :

$$\hat{F}(i)\chi_a(i) = \epsilon_a \chi_a(i) \Leftrightarrow [\hat{h}(i) + \sum_{b \neq a} J_b(i) - K_b(i)] \chi_a(i) = \epsilon_a \chi_a(i) \quad (8)$$

B Équations de la théorie MP2

Dans tout ce paragraphe, i et j correspondent à des orbitales occupées, a et b à des orbitales virtuelles. $|\Phi_{ij}^{ab}\rangle$ le déterminant de Slater où les orbitales i et j du déterminant Hartree-Fock sont substituées par les orbitales a et b. $|\Phi_0\rangle$ correspond au déterminant Hartree-Fock de l'état fondamental. V correspond au potentiel défini dans l'équation 1 Dans le cadre de la théorie MP2 [3, 5], on définit :

— Les intégrales biélectroniques antisymétrisées, où les fonctions ψ sont des spinorbitales (ou spinneurs dans le cadre relativiste) :

$$\langle pq || rs \rangle = \int \int \psi_p^*(1) \psi_q^*(2) r_{12}^{-1} [\psi_r(1) \psi_s(2) - \psi_s(1) \psi_r(2)] d\tau_1 d\tau_2 \quad (9)$$

— L'opérateur de doubles excitations T_2 et les amplitudes de double substitution de l'état fondamental au premier ordre de la théorie de la perturbation (premier ordre en λ) a_{ij}^{ab}

$$T_2 \Phi_0 = \frac{1}{4} t_{ij}^{ab} \Phi_{ij}^{ab} \quad (10)$$

$$t_{ij}^{ab} = - \frac{\langle \Phi_{ij}^{ab} | V | \Phi_0 \rangle}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \quad (11)$$

$$t_{ij}^{ab} = - \frac{\langle ab || ij \rangle}{\Delta_{ij}^{ab}} \quad (12)$$

— L'énergie MP2

$$E^{MP2} = \langle \Phi_0 | V | T_2 \Phi_0 \rangle = \frac{1}{4} \sum_{ijab} t_{ij}^{ab} \langle ij || ab \rangle \quad (13)$$

Dans l'équation de l'énergie MP2, on remarque qu'on peut traduire cette équation sous forme de produit tensoriel, c'est pourquoi les librairies tensorielles sont si importantes en chimie quantique.

C Équations des méthodes CIS, CIS-MP2 et CIS(D)

Dans tout ce paragraphe, i, j, k correspondent à des orbitales occupées, a, b et c à des orbitales virtuelles. $|\Phi_{ijk}^{abc}\rangle$ le déterminant de Slater où les orbitales i, j et k du déterminant Hartree-Fock sont substituées par les orbitales a, b et c . $|\Phi_0\rangle$ correspond au déterminant Hartree-Fock de l'état fondamental. V correspond au potentiel défini dans l'équation 1

La méthode CIS se ramène au calcul des termes d'un hamiltonien. Pour le calcul des éléments de l'Hamiltonien, différents cas doivent être séparés :

— Cas $i=j, a=b$:

$$\langle \Phi_i^a | H | \Phi_i^a \rangle = E_0 - F_{ii} + F_{aa} - \langle ia || ia \rangle$$

— Cas $i \neq j, a=b$:

$$\langle \Phi_{ij}^a | H | \Phi_{ij}^a \rangle = -F_{ij} - \langle ia || ja \rangle$$

— Cas $i=j, a \neq b$:

$$\langle \Phi_i^a | H | \Phi_i^b \rangle = F_{ab} - \langle ai || ji \rangle$$

— Cas $i \neq j, a \neq b$:

$$\langle \Phi_i^a | H | \Phi_j^b \rangle = \langle aj || ib \rangle$$

On utilise les règles de Slater-Condon [4] et on trouve le résultat de l'équation 3 C.

$$\langle \Phi_i^a | H - E_0 \mathbb{1} | \Phi_j^b \rangle = (\epsilon_a - \epsilon_i) \delta_{ij} \delta_{ab} + \langle aj || ib \rangle \quad (14)$$

où

- E_0 est l'énergie de l'état fondamental (Hartree-Fock)
- $\mathbb{1}$ la matrice identité
- H l'hamiltonien CIS
- ϵ_p l'énergie Hartree-Fock de la spinorbitale p
- $\langle aj || ib \rangle$ les intégrales biélectroniques antisymétrisées

Dans le cadre de la théorie de la méthode CIS-MP2 [5], on définit :

- Les opérateurs de simple (U_1) doubles (U_2) et triples (U_3) substitutions au premier ordre de la théorie de la perturbation pour les états excités tels que : $\langle \Phi_i^a | H_{CIS} | U_1 \Phi_0 \rangle = \omega_i^a$
- L'énergie CIS-MP2 :

$$E = \langle \Phi_{CIS} | V | U_2 \Phi_0 \rangle + \langle \Phi_{CIS} | V | U_3 \Phi_0 \rangle$$

Dans le cadre de la théorie CIS(D), on définit :

— Les coefficients CIS b_i^a pour une correction donnée

—

$$u_{ij}^{ab} = \sum_c [\langle ab||cj \rangle b_i^c - \langle ab||ci \rangle b_j^c] + \sum_k [\langle ka||ij \rangle b_k^b - \langle kb||ij \rangle b_k^a]$$

— L'énergie CIS(D) :

$$E_{CIS(D)} = \langle \Phi_{CIS} | V | U_2 \Phi_0 \rangle + \langle \Phi_{CIS} | V | T_2 U_1 \Phi_0 \rangle$$

— L'énergie d'excitation CIS(D) : $\omega = E_{CIS(D)} - E_{MP2}$

$$\omega_{CIS(D)} = E_{CIS(D)} - E_{MP2} = -\frac{1}{4} \sum_{ijab} \frac{(u_{ij}^{ab})^2}{\Delta_{ij}^{ab} - \omega_{CIS}} + \sum_{ia} b_i^a \nu_i^a \quad (15)$$

— Le tableau ν_i^a :

$$\nu_i^a = \frac{1}{2} \sum_{jkb} \langle jk||bc \rangle (b_i^b t_{jk}^{ca} + b_j^a t_{ik}^{cb} + 2b_j^b t_{ik}^{ac})$$