UNIVERSIDAD DE LA LAGUNA

FACULTAD DE CIENCIAS
DEPARTAMENTO DE FÍSICA

POTENTIAL ENERGY SURFACE OF H_3^- USING ATOM-BOND PAIRWISE ADDITIVE SCHEME

Bachelor's degree thesis
2022

Supervised by::
DR. JOSÉ DIEGO BRETÓN PEÑA
DR. JAVIER HERNÁNDEZ ROJAS

Author:
ELENA GARCÍA DE LAMO

RESUMEN

La importancia del anión triatómico más simple, el H_3^-, va más allá del campo de la química cuántica y computacional. Debido a su presencia, aún no demostrada, en el medio interestelar, el campo de la astroquímica ha dedicado varios estudios a dicha molécula. Se propone en el presente trabajo una nueva superficie de energía potencial para el estado fundamental de esta molécula; consecuentemente, se confirmando que existe una configuración estable. Esta PES (nombrada así por sus siglas en inglés) se ha obtenido mediante cálculos *ab initio*; es decir, los resultados se han derivado únicamente de los valores de las constantes universales y los números atómicos.

El objetivo del trabajo es obtener una superficie de energía potencial analítica que consiga describir correctamente los datos *ab initio*. Este es un procedimiento muy usual en cuanto a la PES, con los cálculos *ab initio* se obtienen únicamente puntos configuración-energía; sin embargo realizar dicho cálculo para cada configuración posible es un esfuerzo notable. Por ello, la mejor forma de parametrizar la PES es usar un modelo de potencial analítico. De este modo, se ha obtenido además una PES ajustada a los datos ab inito.

El presente trabajo consiste en dos partes temáticas. En la primera se muestran los diferentes métodos *ab initio* utilizados para calcular la energía de la molécula, mientras que la segunda parte se presentarán las interacciones intramoleculares del anión y el modelo analítico que podrá describir dichas interacciones. Igualmente, se comprobará que los datos *ab initio* quedan adecuadamente descritos por este modelo analítico.

El primer cálculo *ab initio* utilizado, y en el que se apoyarán todos los demás, es el método variacional Hartree Fock. En él se parte de la aproximación the Born-Oppenheimer y se obtiene así la ecuación de Schrodinger electrónica. Debido a la naturaleza fermiónica de los electrones se antisimetriza la función de onda, construida por la combinación lineal de espín-orbitales, mediante determinantes de Slater. Tratándose el método HF de un método variacional, busca los coeficientes correspondientes a los espín-orbitales que minimizan la energía. El inconveniente es que la ecuación depende de sus soluciones, por tanto se ha de resolver mediante un método iterativo que obtiene la solución cuando converge.

El método Hartree Fock presenta la carencia de no tener en cuenta la correlación de electrones. Esto implica que HF calcula sobre cada electrón el efecto del resto como un promedio de la interacción electrónica, llevando esto a una sobrestimación de la energía molecular. Para corregirlo, se han desarrollado otros métodos, llamados post-Hartree Fock, que utilizan los resultados de HF para realizar un cálculo que mejore la exactitud de los cálculos. Entre ellos, los que se han llevado a cabo en el presente trabajo son Moller-Plesset, un método que utiliza teoría de perturbaciones, y Coupled Cluster, método que parte de la función de onda de Hartree Fock para construir una nueva función de onda que tenga en cuenta excitaciones de electrones. Por tanto, aunque usando procedimientos muy distintos, el fin último de ambos métodos es calcular la energía de correlación electrónica.

Para llevar a cabo los métodos *ab initio* es imprescindible aportar una base de funciones que describa los orbitales. Existen diversas familias de conjuntos de bases aunque se suelen escoger funciones gaussianas contraídas, pues facilitan los cálculos con apenas pérdida de exactitud. En este trabajo se han llevado a cabo los cálculos con la base aug-cc-pV5Z desarrollada por Dunning y sus colaboradores. Estas bases están optimizadas para describir adecuadamente la correlación electrónica y los efectos de polarización. La limitación de las bases de funciones finitas lleva a error en los resultados; sin embargo pueden en cierta medida corregirse aplicando el límite a base completa (CBS) y teniendo en cuenta el error de superposición de bases (BSSE). En el presente trabajo se da un esbozo de qué son estás correcciones y cómo llevarlas a cabo, pero no se han aplicado dichas correcciones a la PES, por lo que la nuestra se trata de una PES no corregida.

Por otro lado, la segunda parte del trabajo consiste en obtener una PES analítica parametrizada respecto a distancia entre los átomos. Para conocer qué modelo de potencial analítico puede ser propicio, se recurre a la teoría de interacciones moleculares, pues un potencial analítico adecuado debe tener en cuenta las interacciones dominantes de la molécula.

En los resultados se presentan los datos *ab initio* mediante el método HF, MP y CC, calculados mediante el software de código abierto NWChem. Dicho software precisa de un fichero de entrada del que se da un ejemplo en el Apéndice A.

Como paso previo y necesario al estudio de la molécula H_3^- y a la obtención de la PES, se ha calculado las energías y configuraciones de equilibrio de los productos de disociación, el dímero H_2 y el anión H^-. Puesto que el cálculo se realiza con los tres métodos *ab initio* se puede comprobar la exactitud de cada método y cuál de ellos es el más adecuado para los cálculos posteriores.

Para la molécula principal, el primer paso es obtener la energía y configuración de equilibrio de H_3^- por los métodos *ab initio*. Así, se conoce que la configuración de equilibrio tiene una geometría lineal asimétrica y se comprueba que los resultados son congruentes con los de la literatura. Posteriormente, se comienza a construir la PES que dado el sistema tendría tres grados de libertad; sin embargo, se han de tener en cuenta algunas consideraciones simplificatorias que llevan a modelizar el compuesto como una interacción entre H_2 y H^-, fijando un grado de libertad. Así, queda la PES con solo dos coordenadas: la distancia desde el anión H^- al centro de masas de H_2 y el ángulo que subyace entre el tercer núcleo y el eje principal del dímero. Teniendo esto en cuenta se obtiene una PES mediante puntos *ab initio* que se utilizará para ajustar la PES analítica.

El modelo analítico de potencial propuesto está formado por dos contribuciones, una contribución electrostática y otra que incluirá el resto de interacciones. El término electrostático se calcula mediante interacción coulombiana entre las cargas específicas de H_2 y H^-, mientras que para el término de interacción se usa un potencial Lennard Jones mejorado (ILJ). Específicamente, se utiliza un modelo atom-bond que permite dar cuenta de la anisotropía del potencial, debida a la molécula diatómica H_2.

Finalmente, se construye la PES analítica, ajustando los parámetros y comprobando que los datos *ab initio* están debidamente descritos. Igualmente, se compara con otras PES analíticas del H_3^- que aparecen en la literatura y los resultados no son solo coherentes sino que, en nuestra opinión, se supera la exactitud de otras PES.

AGRADECIMIENTOS

Me gustaría tramitir mi más sincero agradecimiento a mis tutores por su apoyo, dedicación y por haberme guiado a lo largo de este proceso. Por supuesto, agradezco al Departamente de Física de la ULL el uso de sus instalaciones y por proveer del NWChem.

Además, quisiera agradecer a Maren Brauner, pues de no haber sido por su buen hacer en la presentación y Trabajo de Final de Master, tal vez no me hubiera encontrado con este tema, al que le he cogido cariño.

CONTENTS

CHAPTER 1
INTRODUCTION

Resumen

Aunque el hallazgo experimental de la molécula H_3^- sucedió en las dos últimas décadas, desde el siglo pasado habían surgido diversos estudios teóricos que cuestionan la estructura y estabilidad de dicho anión. Incluso, ya entonces se barajaba la posibilidad de que la molécula tuviera interés astrofísico, pues se trata de un buen candidato para conformar nebulosas gaseosas. Hoy en día, la presencia del anión H_3^- en el medio interestelar continúa siendo una hipótesis que ha inspirado diversas investigaciones. También es evidente su importancia para la química cuántica y computacional y otros trabajos abordan la molécula desde la teoría de colisiones.

Dado el interés en la estructura del anión se pretende en el presente trabajo obtener una superficie de energía potencial, usando para ello cálculos *ab initio*. Además, se busca un modelo analítico de potencial que pueda reproducir fielmente estos datos. Para diseñar este modelo es necesario que quede respaldado por la teoría de interacciones atómicas, por tanto el modelo consta de un término de contribución electrostática y otro término de interacción representado por un potencial Lennard Jones mejorado, cuyo uso exitoso para describir interacciones similares ha sido probado en otros trabajos.

The molecular hydrogen clusters, neutral, cationic, or anionic, deserve great attention from fields such as quantum chemistry, molecular collision physics, and astrochemistry. Regardless, until relatively recently, only positive clusters, H_n^+, where $n = 3, 5, 7, ...$, were discovered experimentally [1] and successfully researched by ab initio theoretical methods[2]; in fact, the theoretical studies investigate the structure of the clusters with $(H_3^+)(H_2)_N$ form. On the other hand, anionic hydrogen clusters have been theoretically predicted to be stable since the 1990s but have only recently been detected in experiments, at least the smallest member of them H_3^-. Trihydrogen anion was finally identified from dielectric discharge plasmas by Wang and co-workers in 2003[3]. Therefore, the prior investigations were limited to theoretical methods and used quantum chemistry calculations to study and debate the stability and configuration of the H_n^- anions [4–7].

The trihydrogen anion, consisting of three protons and four electrons, is the simplest negative triatomic ion. In 1919, Bohr not only suggested its existence but also inferred its structure, as written in the reference [2]. However, the first theoretical investigation of H_3^- was reported in 1937 [8] and that study, written by Stevenson and Hirschfelder, established its linear configuration, but there was not initially a general agreement about its electronic structure. To reliably demonstrate that the H_3^- anion is stable, improving the precision of ab initio calculations was necessary, which was not feasible until the 1990s. Thus, the road to proving not only the structure of H_3^- but its very existence has been long.

For the current research, the trihydrogen anion plays a crucial role in astrophysics, specifically astrochemistry. It has been suggested that in gaseous nebulae, the conditions may be suitable for negative ion clusters to exist [6]; therefore, anionic hydrogen clusters, stable in the interstellar

medium, are suitable candidates. The hypothesis says that some diffuse interstellar absorption lines may be because of the H_3^- presence in diffuse interstellar clouds [9], although it has to be conclusively demonstrated.

Several studies have analyzed the anionic trihydrogen clusters. To sum up, they can be divided broadly into three categories: studies in astrochemistry [9] about diffuse bands, works about collisions between H_2 and H^- [5, 10] that are relevant for the chemistry of cold interstellar clouds, and studies about the structure of the anion potential energy surface [6, 11, 12].

This latter aspect of the anion provides the primary motivation for the present study. As a molecule only conformed by protons and electrons, its relative chemical simplicity makes it ideal for investigating structure and bonding characteristics. Ab initio methods have been employed for the calculations, but being H_3^- characterized by weak Van der Waals bonds, it is challenging even for these methods.

Thus, the first part of the present study is devoted to comparing different ab initio methods, MP2 and CCSD(T), and concluding which works more accurately. The adequate description of H_3^- equilibrium energy in the electronic ground state has been evaluated using a large basis set and corrected with the complete basis set (CBS) limit[1]. However, the importance of correcting the basis superposition error (BSSE) on a more extensive basis is noticeable, but our calculations lack that correction. As a result of this, our PES is uncorrected.

The principal purpose of this work is to provide an accurate enough PES of the trihydrogen anion's ground state, employing both ab initio calculations and an analytical expression. The analytical potential model has to describe the ab initio PES, so the correctness of the analytical potential model's form is also examined.

The analytical PES model proposed is a sum of electrostatic contribution and an Improved Lennard Jones [13], simulating the non-covalent contribution [14]. That process has not been proposed for H_3^- PES before, but it has been used to describe the potential energy surface for atom–molecule systems. The usefulness of this potential model in some similar molecules and other systems is remarkable [14, 15]. Therefore, this bachelor's thesis also shows a thorough test of a simple potential model, which can meet expectations in a wide range of systems.

[1] The calculation for the equilibrium energy using the CBS limit is collected in Appendix B

CHAPTER 2
THEORETICAL BACKGROUND

2.1 Ab initio methods

Resumen

En este capítulo se recoge el fundamento físico y el proceso matemático tras los métodos *ab initio* utilizados. Partimos del Hartree-Fock, el método que, aunque no muy exacto en sus resultados, servirá de precursor para los métodos Møller-Plesset y Coupled Cluster. El método HF de campo autoconsistente es un método variacional que necesita realizarse por método iterativo pues la solución depende de sí misma. La flaqueza de este método es que trata el efecto de la nube electrónica sobre cada electrón como un campo promedio, por lo que en es incapaz de calcular la correlación electrónica. Por esta razón se desarrollaron los métodos post-Hartree-Fock como MP y CC.

También se explica en este capítulo en qué consisten las bases de funciones y algunos de sus tipos, pues estas bases juegan un importante papel en la precisión de los cálculos. Además, se nombran ciertos errores que aparecen al tener que usar bases finitas y las correcciones de los resultados.

Por último, se habla sobre las interacciones intermoleculares, una introducción general y aquellas que afectan específicamente al H_3^- que se trata de un compuesto de Van der Wals. Para esta molécula se deriva cómo ha de ser su potencial de enlace débil a largo alcance, basándonos en la interacción electrostática y de inducción. Igualmente, se explica en qué consiste una PES, de cuántas coordenadas depende y cómo se puede parametrizar a través de un modelo de potencial analítico.

This section is devoted to introducing briefly the quantum ab-initio methods that are used in the calculation of intermolecular forces. To write it, we have used the following sources [16–19]. Ab initio methods aim to resolve the non-relativistic Schrödinger equation and describe the system, atom, or molecule, given the positions of all the nuclei and the number of electrons. Ab initio (or first principles) methods seek to calculate molecular properties from the beginning and use a Hamiltonian without any experimental information and only use the values of the fundamental physical constants. However, that approach "from the beginning" does not mean accurate at 100% because some approximation always has to be considered. Although the accuracy of the calculation can be systematically improved, the limit is set by the available computing resource.

An approximation used almost universally is the Born–Oppenheimer approximation, and its main consequence is the separation of electronic and nuclear motion. In addition, ab initio programs usually ignore relativistic effects of all kinds. That is generally a good approximation, especially for atoms and molecules with closed-shell electronic states. In this bachelor's thesis, the system analyzed is the H_3^- molecule in the ground state, and it is a closed-shell electronic system.

The Hartree-Fock self-consistent field (HF-SCF) method is the simplest, although a fundamental example of an ab initio calculation as a variational method supposes that each electron is affected by an average field created by the nuclei and the rest of the electrons. The shortcoming of HF-SCF is that it does not consider the electron correlation, as we will specify later; therefore, that is not a fully satisfactory ab initio method. It serves as the precursor and stepping stone for more elaborate methods, like the Møller–Plesset perturbation theory (MP) and the Coupled Cluster (CC) methods. Those methods include the electron correlation energy using the results from the HF method. Then, they are called post-HF methods.

2.1.1 A Hartree–Fock (HF) self-consistent field (SCF) method

The goal of the HF is to resolve the time-independent Schrödinger equation. First of all, the Born-Oppenheimer approximation allows separate the motion of the nuclei from that of the electrons. The resolution starts from the electronic Schrödinger equation:

$$(\hat{H}_{el} + V_{NN})\Psi(\mathbf{r},\mathbf{R}) = E_{el}(\mathbf{R})\Psi(\mathbf{r},\mathbf{R}) \tag{2.1}$$

The term V_{NN} is the electrostatic interaction between nuclei (nucleus to nucleus), and $\hat{H}_{el}$ corresponds to the electronic energy. So the Hamiltonian can be divided into electronic and nuclei terms and, taking the second one as a constant (BO approximation), the electronic energy is obtained by solving the equation with a purely electronic Hamiltonian. The eigenvalue $E_{el}(\mathbf{R})$ is a function of the nuclei coordinates, $\{\mathbf{R}\}$, while the wavefunction $\Psi(\mathbf{r},\mathbf{R})$ depends on the electron and nuclei coordinates. The function $E_{el}(\mathbf{R})$ represents the potential energy hypersurface. It depends on the nuclear coordinates $(3N-6)$. Since the energy is a function of the positions of the nuclei, it can define certain properties of a bond, such as stable configuration of the system, given by the minima on the potential energy surface, $E_{el}(\mathbf{R})$.

The non-relativistic electronic Hamiltonian for a polyatomic molecule is expressed using atomic units (a. u.) as:

$$\hat{H}_{el} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_i\sum_\alpha \frac{Z_\alpha}{r_{i\alpha}} + \sum_i\sum_{j>i}\frac{1}{r_{ij}} \tag{2.2}$$

where $r_{i\alpha}$ is the distance from the electron i to the nucleus α, $Z_{i\alpha}$ is the number of protons of the nucleus (atomic number), and r_{ij} is the distance between two electrons. To simplify the expression, we introduce the monoelectronic operators. In this work atomic units will be used

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_\alpha \frac{Z_\alpha}{r_{i\alpha}} \tag{2.3}$$

And the bielectronic contribution

$$\hat{u}_{ij} = \frac{1}{r_{ij}} \tag{2.4}$$

The first term in Eq. (2.2) refers to the kinetic energy of the electrons. The second term is the interaction between electrons and nuclei, and the last one, corresponding to Eq. (2.4), is the interaction between electrons. Because of this term of interelectronic repulsion, the Hamiltonian is not separable. However, a scheme is proposed to carry out the method. First, an initial guess of the wavefunctions is necessary; therefore, a reasonable assumption, although not perfect because

the Hamiltonian is not separable, is that these wavefunctions are the product of different one-electron wavefunctions. However, the anti-symmetry requirements for fermions must be satisfied; therefore, wavefunctions must be a Slater determinant. That wavefunction satisfies that fermions are indistinguishable, and consequently, the Pauli principle is fulfilled.

$$\Psi(x_1,x_2,...x_N) = (\mathrm{N!})^{-\frac{1}{2}} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) & ... & \chi_N(x_1) \\ \chi_i(x_2) & \chi_j(x_2) & \dots & \chi_N(x_2) \\ \vdots & & \ddots & \\ \chi_i(x_N) & \chi_j(x_N) & \dots & \chi_N(x_N) \end{vmatrix} \tag{2.5}$$

The functions χ_i are orthonormal monoelectronic spin-orbital that describe the behaviour of electrons with spin since x_N include the spatial coordinates $\{\mathbf{r_N}\}$ of the electron and the internal spin degree of freedom $\{\sigma_N\}$. The subscript of the coordinates x is the electron label $(1,2,...,N)$. On the other hand, the subscript of the spin-orbitals χ refers to the monoelectronic quantum number.

The Hartree–Fock method looks for those orbitals $\{\chi_i(x_k)\}$ that minimize the energy $\langle\Psi|\hat{H}_{el}|\Psi\rangle$, using the wavefunction in Eq. (2.5). That implies, according to the variational principle, we have an upper bound for the energy

$$E_{HF} \leq E[\Psi] = \frac{\langle\Psi|\hat{H}_{el}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{2.6}$$

the denominator becomes $\langle\Psi|\Psi\rangle = 1$ because the wavefunction is normalized. It minimize also the variational integral

$$\delta E[\Psi] = \delta\left[\sum_{\sigma_1}\sum_{\sigma_2}...\sum_{\sigma_N}\int \Psi(x_1,x_2,...x_N)\hat{H}_{el}\Psi(x_1,x_2,...x_N)d\mathbf{r_1}d\mathbf{r_2}...d\mathbf{r_N}\right] = 0 \tag{2.7}$$

However, this method is subject to the condition that orbitals $\{\chi_i(x_k)\}$ must be orthonormal. That restriction can be treated via the technique of Lagrangian multipliers. Therefore, the value of the energy is approximated due to the variational principle. Electronic energy using HF method are expressed as

$$E_{HF} = \sum_i^N \langle\chi_i|\hat{h}_1|\chi_i\rangle + \frac{1}{2}\sum_{j,i}^N(\langle\chi_i\chi_j|\hat{u}_{12}|\chi_i\chi_j\rangle - \langle\chi_i\chi_j|\hat{u}_{12}|\chi_j\chi_i\rangle) \tag{2.8}$$

The notation $\langle\chi_i\chi_j|\hat{u}_{12}|\chi_j\chi_i\rangle$ in Eq. (2.8) implies an integral over all spatial coordinates and the summation over the spin degrees of freedom of the two electrons. The subscripts u_{12} refers to two random electrons that are involve and it does not imply only the electrons labeled like (1, 2). Equally, $\langle\chi_i|\hat{h}_1|\chi_i\rangle$ means an integral over the spatial coordinates and the summation over the spin degrees of freedom of one electron. The difference between them is that $\hat{h_1}$ is a one-electron operator while $\hat{u_{12}}$ is a two-electron operator.

The Euler-Lagrange equations for the functional in Eq. (2.7) yields the standard Schrodinger equation that, being the orbitals orthonormal, relates to the Hartree-Fock equations

$$f(x_1)|\chi_i(x_1)\rangle = \varepsilon_i|\chi_i(x_1)\rangle \tag{2.9}$$

where the energy ε_i corresponding to the spin-orbital i and the $f(x_i)$ is the Fock operator, with the expression

$$f(x_1) = \hat{h}_i(x_1) + \sum_{j=1}^{N} \left(J_j(x_1) - K_j(x_1)\right) \tag{2.10}$$

that depends on the Coulomb operator $\hat{J}$ and the exchange operator $\hat{K}$, defined as

$$J_j(x_1)\left|\chi_i(x_1)\right\rangle = \left\langle \chi_j \middle| \hat{u}_{12} \middle| \chi_j \right\rangle \left|\chi_i(x_1)\right\rangle = \left[\int \chi_j^*(x_2)\frac{1}{r_{12}}\chi_j(x_2)dx_2\right]\left|\chi_i(x_1)\right\rangle \tag{2.11}$$

$$K_j(x_1)\left|\chi_i(x_1)\right\rangle = \left\langle \chi_j \middle| \hat{u}_{12} \middle| \chi_i \right\rangle \left|\chi_i(x_1)\right\rangle = \left[\int \chi_j^*(x_2)\frac{1}{r_{12}}\chi_i(x_2)dx_2\right]\left|\chi_i(x_1)\right\rangle \tag{2.12}$$

The Coulomb operator in Eq. (2.11) is the interaction potential energy between one electron and the rest, interpreting the action of the electron cloud like a mean field effect. The exchange operator, Eq. (2.12), arises from the requirement that the wavefunction was antisymmetric with respect to electron exchange. Hartree-Fock in Eq. (2.8) energy can be rewrote as the expression

$$E_{\mathrm{HF}} = \sum_{i=1}^{N} \hat{h}_i + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{1}{2}(J_{ij} - K_{ij}) \tag{2.13}$$

$$J_{ij} = \left\langle \chi_i\chi_j \middle| \hat{u}_{12} \middle| \chi_i\chi_j \right\rangle, \quad K_{ij} = \left\langle \chi_i\chi_j \middle| \hat{u}_{12} \middle| \chi_j\chi_i \right\rangle \tag{2.14}$$

Although Eq. (2.13) depicts an interaction between electrons, besides the interaction with the core, the electrons of the HF method are not correlated because it is the action of the mean-field of the electrons over each electron.

Another shortcoming of the Hartree-Fock process is that it depends on its own solutions; therefore, to obtain the spin-orbitals energy and spin-orbitals using the Hartree-Fock equations, the spin orbitals themselves are required. Thus, it is an iterative method, and an initial guess for the spin-orbitals is needed. Using an initial guess of $\{\chi_i\}$ the matrix of the operators J_{ij} and K_{ij} can be calculated, Eq. (2.11, 2.12) respectively. With that, the Fock matrix is obtained using Eq. (2.10), and that matrix allows the resolution of Hartree-Fock equations; therefore, a new set of spin-orbitals $\{\chi_i'\}$ is obtained. The new spin-orbital set is used as the initial point for the next iteration. The process ends when the solution converges, and the difference between $\{\chi_i\}$ and $\{\chi_i'\}$ is negligible; in other words, the process continues until $\{\chi_i'\}$ is consistent. That iterative method is known as Self-Consistent Field (SCF)

A key development to calculate accurate molecular HF-SCF wavefunctions feasible was Roothaan's proposal to expand the molecular spin-orbitals as linear combinations of a set of one-electron basis functions (atomic orbitals). That is known as a linear combination of atomic orbitals or LCAO. The molecular spin-orbital χ_i expressed as a linear combination of atomic orbitals

$$\chi_i = \sum_{\mu}^{N_o} C_{\mu i}\tilde{\chi}_\mu \tag{2.15}$$

where $\tilde{\chi}_\mu$ are the atomic orbitals defined previously in a basis set, and N_o is the number of atomic orbitals in that basis set

With a larger number of basis functions, the energy value would be more accurate; then, using a large basis set with well-chosen functions, the error for the orbitals could be negligible. Although the basis set for the calculation cannot be complete, there are schemes to approximate the energy value as if it were computed in an infinite basis set, the so-called CBS limit. A more detailed explanation of basis sets is included in section 2.2. To sum up, an ab initio SCF Molecular Orbital calculation uses the approximation of taking Ψ as an antisymmetrized product of molecular spin-orbitals and uses a finite, and hence incomplete, basis set.

As aforementioned, the Hartree-Fock method neglects electron correlation, which means that Coulomb and exchange terms are considered a mean effect. Hence, each particle is subjected to the mean-field created by all other particles, making this method imprecise for atoms or molecules with several electrons. The atomic and molecular correlation energy is defined as the difference between the non-relativistic true (although unknown) energy and the Hartree–Fock (HF) non-relativistic energy calculated with a complete basis set. The following sections describe some post-Hartree-Fock methods. They use the HF-SCF method but try to include the correlation energy of the electrons.

2.1.2 Møller-Plesset second-order perturbation theory (HF-MP2)

Møller-Plesset uses the Hartree-Fock method as the initial step but attempts to estimate the correlation energy by applying perturbation theory. The unperturbed part is the HF Hamiltonian

$$H^{(0)} = H_{HF} = \sum_{t=1}^{\mathrm{N}_e} f_t \tag{2.16}$$

The perturbation term is defined as

$$H^{(1)} = H - H_{HF} = \frac{1}{2}\sum_{i\neq j}\frac{1}{r_{\mathrm{ij}}} - \sum_{ij}^{N}(J_{ij} - K_{ij}) \tag{2.17}$$

Perturbation theory has a correction to first-order, second-order, and higher-order terms. The second-order perturbation theory, depicted below, only includes until the second-order term, but higher orders can be calculated as well. It can be proved that the sum of zero-order energy, $E^{(0)}$, and first-order energy, $E^{(1)}$, is the already known Hartree-Fock energy

$$E_{HF} = E^{(0)} + E^{(1)} \tag{2.18}$$

$$E^{(2)} = \sum_{j>0}\frac{\left|\left\langle \Psi_j^{(0)}\middle|H^{(1)}\middle|\Psi_{HF}\right\rangle\right|^2}{E^{(0)} - E_j^{(0)}} \tag{2.19}$$

Therefore the energy correlation, Eq. (2.19), will be the second-order perturbation. The total energy of MP2 will be the expression:

$$E_{TOT} = E^{(0)} + E^{(1)} + E^{(2)} = E_{HF} + E^{(2)} \tag{2.20}$$

The number in the Møller-Plesset method refers to the perturbation order used. MP2 approximates to a second-order term, MP3 used until third-order term, MP4 fourth-order, and similarly for higher terms.

2.1.3 Coupled Cluster calculations (CC)

Coupled Cluster (CC) is a popular post-Hartree-Fock method that uses the Hartree-Fock wave function $|\Psi_{HF}\rangle$ as the started point to build a molecular wave function. The Coupled Cluster wave function $|\Psi_{CC}\rangle$ is written as

$$|\Psi_{CC}\rangle = e^{T}\,|\Psi_{HF}\rangle \tag{2.21}$$

the operator e^T is defined by its Taylor-series expansion and T is the so-called cluster operator, defined as $T = T_1 + T_2 + T_3 ... + T_N$. The subscript of the operators is related to the number of excitations that can produce on the HF wavefunction. T_1 implies the excitation of a single particle, substituting an occupied orbital, χ_i, with a virtual one, χ_a. The operator T_2 causes a double excitation; two occupied orbitals change into a two virtual orbitals. That actuation of these operators on the $|\Psi_{HF}\rangle$ is expressed as

$$T_1\,|\Psi_{HF}\rangle = \sum_{a=N+1}^{\infty}\sum_{i=1}^{N} t_i^a\,|\Psi_i^a\rangle\,, \quad T_2\,|\Psi_{HF}\rangle = \sum_{b=a+1}^{\infty}\sum_{a=N+1}^{\infty}\sum_{j=i+1}^{N}\sum_{i=1}^{N-1} t_{ij}^{ab}\left|\Psi_{ij}^{ab}\right\rangle \tag{2.22}$$

where Ψ is Slater determinant whose subscript implies the number of excitations and superscript corresponds to the number of virtual orbitals. This wave function has the real orbital χ_i substituted by the virtual orbital χ_a, t_i^a is a numerical coefficient that depends on the virtual orbital a and the excited electron i. Therefore, for T_2, the indices i, j are the initially occupied spin orbitals while a,b are the now occupied orbitals but virtual orbitals before. It is worth noticing the limits of the summations in Eq. (2.22). The first excitation corresponds to the subscript i, and the summation limit is $N-1$; the remaining corresponds to the excitation with subscript j. The summation over virtual orbitals begins at $N+1$, from 1 to N there are occupied orbitals, and ends at ∞. That means an infinite number of virtual orbitals.

Using $T_1^2 = T_1 \cdot T_1$ is slightly different from T_2 because the first one is the excitation of one particle twice, while the result of applying T_2 is two excitations at a time, without duplicity. On the other hand, T_N has to be the last operator since the maximum number of excitations at a time is limited to the number of electrons, N.

The equation above (2.22) would be for a complete basis set with infinite functions that describe virtual orbitals. However, that is impossible for computational calculus since the basis set has to be finite, limiting the number of virtual orbitals in the equation. Thus a minimum basis set is previously required because the electrons occupy some orbitals that the basis functions have to describe. The equation (2.22) using a finite basis yields:

$$T_1\,|\Psi_{HF}\rangle = \sum_{a=N+1}^{virt}\sum_{i=1}^{N} t_i^a\,|\Psi_i^a\rangle\,, \quad T_2\,|\Psi_{HF}\rangle = \sum_{b=a+1}^{virt}\sum_{a=N+1}^{virt}\sum_{j=i+1}^{N}\sum_{i=1}^{N-1} t_{ij}^{ab}\left|\Psi_{ij}^{ab}\right\rangle \tag{2.23}$$

Due to the orthogonality of spin orbitals $\langle\Psi_{HF}|e^T\Psi_{HF}\rangle = 1$, the energy obtained by the CC method has the expression

$$E_{CC} = \langle\Psi_{HF}|H|\Psi_{CC}\rangle = \langle\Psi_{HF}|H|e^T\Psi_{HF}\rangle \tag{2.24}$$

Evaluating the Eq. (2.24) with the wave function expression (2.21) and inserting the Taylor-series expansion of e^T, we obtain

$$\begin{aligned}\left\langle \Psi_{HF}|\hat{H}|e^{T}\Psi_{HF}\right\rangle &= \left\langle \Psi_{HF}|\hat{H}|\left(1+T+\frac{1}{2}T^{2}+\cdots\right)\Psi_{HF}\right\rangle \\ &= \left\langle \Psi_{HF}|\hat{H}|\Psi_{HF}\right\rangle + \left\langle \Psi_{HF}|\hat{H}|(T_1+T_2+...)\Psi_{HF}\right\rangle + ... \\ &= E_{\mathrm{HF}} + \left\langle \Psi_{HF}|\hat{H}|(T_1+T_2+...)\Psi_{HF}\right\rangle ...\end{aligned} \tag{2.25}$$

where E_{HF} is the Hartree–Fock SCF energy. Knowing that the Hamiltonian only has one-body and two-body operators, Eq. (2.2), and using Brillouin's theorem together with the Slater-Condon rules, see reference [18], some term goes to zero in Eq. (2.25). The expression becomes

$$E_{\mathrm{CC}} = E_{HF} + \sum_{i<j}^{\mathrm{N}}\sum_{a<b}^{\mathrm{virt}}\left(t_{ij}^{ab}+t_i^a t_j^b - t_i^b t_j^a\right)\left(\left\langle\chi_i\chi_j|\hat{u}_{12}|\chi_a\chi_b\right\rangle - \left\langle\chi_i\chi_j|\hat{u}_{12}|\chi_b\chi_a\right\rangle\right) \tag{2.26}$$

The CC calculus aims to obtain the coefficients t_i^a, t_{ij}^{ab}... in Eq. (2.26), called amplitudes, whatever the subscript i, j and the superscript a,b could be. However, the Coupled Cluster method is neither a variational one, although it uses the solution of the Hartree-Fock method that is variational, nor uses perturbations theory, as Møller-Plesset method. Furthermore, the CC method seriously improves the correlation energy, compared with the MP2 method, but requires more computing time. The improvement of the calculations using the CC method with respect to MP2 one will be confirmed in our calculations.

On the one hand, the Coupled Cluster method would need so much computational calculus that it would be barely impossible to use. However, theory shows that the most important contribution to T is made by T_2. Therefore, it is plausible that $T \approx T_2$ and the inclusion of only T_2 gives an approximate approach called the Coupled-Cluster Doubles (CCD) method. Nevertheless, other approximations can be elected. For instance, take T as $T_1 + T_2$ that will be Coupled-Cluster Single Double (CCSD) method and even include T_3 to the sum, CCSDT method. The latter gives accurate results for correlation energies and molecular properties but is very computationally demanding and only feasible for small molecules. The most widely used approximated form of CC is CCSD(T), where the triples contribution may be evaluated by perturbation theory and added to the CCSD results.

2.2 Basis sets

A finite basis set slightly changes the expressions of the calculations. For instance, in the equation (2.23) summation turns out to be finite, changing its upper limit from infinity, that would be the case that the basis set is a complete one, to the number of virtual orbitals contained in the finite basis set. The finite basis set causes the solution of the Schrödinger equation cannot be exact, but it also makes it computationally affordable. The calculation would be impossible with an infinite number of functions. For a better understanding of what should be the chosen basis set, it is necessary to provide a deeper explanation of basis sets and their importance in ab initio calculations.

The election of the basis set cannot be hazardous, and it has to be selected carefully, taking into account the characteristics of the studied molecule. The best way to make that decision is to know, based on the bibliography, what is the most suitable basis set for each type of geometry.

The basis sets usually used for diatomic molecules are basis whose functions correspond closely to atomic orbitals, and some of them are centered in one atom and the rest in the other. Each atomic orbital can be represented as a linear combination of one or more Slater-type orbitals (STOs). A similar method is used for polyatomic molecules, with the difference that the STOs have to be centered in the different nuclei. However, the presence of multiple atoms causes hardships in evaluating the needed integrals. So, to speed up the integral evaluation of molecules, new basis functions were proposed, the Gaussian-type functions (GTFs). As its name suggests, GTFs are functions with a Cartesian Gaussian form

$$g_{ijk}(r) = N_{ijk} x^i y^j z^k e^{-\alpha^2} \tag{2.27}$$

where i, j, k are nonnegative integers, α is a positive orbital exponent and N_{ijk} is a normalization constant, whose value depends of the indexes i, j, k and it is determined by an expression that can be found in [17]. The integers i, j, k rule on the form of the Gaussian described. When $i+j+k=0$ it is called an s-type Gaussian. If $i+j+k=1$ the Gaussian is a p-type. That Gaussian has three possibilities $i=1$, $j=1$ or $k=1$ and those orbitals have a privileged spatial direction. A similar rule implies that d-type Gaussian has $i+j+k=2$.

A fundamental concept of basis sets is the minimal basis set. A minimum basis set consists of one STO for each atom's inner-shell and valence-shell atomic orbital. Each atom's occupied orbitals must be described with an STO. Gaussian-type functions can describe a minimum basis set equally, but the correspondence between these functions and the atomic orbitals is not one-to-one. So, a minimum basis set of GTFs has more functions than a minimum basis set of STOs. To reduce the number of basis functions, the contracted Gaussian-type functions (CGTFs) are helpful. CGTFs are a linear combination of primitive Gaussians from Eq. (2.27), and its expression is as follows:

$$\omega_r = \sum_u d_{ur} g_u \tag{2.28}$$

where d_{ur} are contraction coefficients, constants already known from a previous optimization, and they are held fixed during the calculation.The normalized Cartesian Gaussians, g_u, are centered in the same atom and share the same indexes i, j, k but different orbital exponent α.

A minimal basis set of contracted Gaussians (2.28) consists of the same number of functions as the STO basis set; therefore, it saves computational time with little loss of accuracy. From this moment onwards, all the explanations will refer to the CGTF basis set because it is the basis set used in the present bachelor's thesis.

Unfortunately, an atomic orbital requires more than a CGTF to be appropriately described, then adding more functions for each atomic orbital would imply a substantial improvement. Double-zeta (DZ) basis set is built with the double basis functions than the minimum basis set.

Basis sets are classified by their number of basis functions; therefore, basis double-zeta (DZ), triple-zeta(TZ), and so on exist. Triple-zeta follows the same rule, but each minimum basis set orbital is replaced by three different functions. In order to reduce computational time, the split-valence (SV) basis set utilizes two CGTFs (in the case of double-zeta) for each valence atomic orbital but only one for each inner-shell atomic orbital (also known as Frozen Core).

Due to the interaction between atoms, the atomic orbitals are distorted in shape upon the formation of the molecule. In order to allow for this polarization effect, the basis set has to include more functions with the quantum number ℓ greater than the maximum ℓ number of the valence shell. Any such basis set is known as a polarized (P) basis set. Furthermore, for anions and long-range interactions such as Van der Waals forces, the result is improved by adding highly diffuse functions, a function with a very small orbital exponent, α in Eq. (2.27).

CGTF basis sets include cc-pVDZ, cc-pVTZ, cc-pVQZ, cc-pV5Z, and cc-pV6Z (developed by Dunning and co-workers [20] and collectively named cc-pVnZ with n= 2, 3, 4...). Those basis sets are used in calculation methods that include electron correlation, and the cc in their name relates to "correlation consistent". The augmented basis sets aug-cc-pVnZ are formed by adding diffuse primitive non-polarization and polarization functions to the cc-pVnZ basis sets. Those basis sets are especially suitable for anions and molecules with hydrogen bonding.

The augmented basis sets were used to carry out our calculations. The shells, number of functions and amount of orbitals of each basis set is detailed in table 2.1. The third column of this table has to be read like the orbitals needed for the number of shells and their kind. For instance, "3s2p" means that the basis set aug-cc-pVDZ has the orbitals to describe three shells s and two shells p; therefore, the shells are 5 (3s+2p), and the functions are nine because each shell p includes three orbitals. Relating the orbitals with the Contracted Gaussian that describe their, the aug-cc-pVDZ will include three s-type Gaussian ($i = j = k = 0$) and six p-type Gaussian ($i + j + k = 1$).

	Shells	Number of functions	Orbitals
aug-cc-pVDZ	5	9	3s2p
aug-cc-pVTZ	9	25	4s3p2d
aug-cc-pVQZ	14	55	5s4p3d2f
aug-cc-pV5Z	20	105	6s5p4d3f2g

Table 2.1: Functions and types of the augmented basis sets, for the H_3^- computation. The increment of the number of functions for each basis is really significant.

A basis sets repository could be found in [21]. However, the software used in this Bachelor´s Thesis for quantum chemistry calculations includes a library with the most usual basis sets.

2.2.1 Complete Basis Set (CBS) limit

Although the augmented basis sets include several basis set functions, they are finite basis sets and, however extensive, will not reach the accuracy of a complete basis set. Fortunately, an approximation to the complete basis set (CBS) limit is possible by expanding the finite basis towards an infinite set of functions [22]. This approximation can be applied using empirical extrapolation techniques, considering n of different basis sets (where $n = 2,3,4,5$) and the energy obtained using these basis sets in the calculation.

The election of the extrapolation scheme depends on the ab initio method carried out and the finite basis sets employed. The consulted reference [22] to achieve a reliable CBS limit depicts that the process, an extrapolation using the next equation that is suitable for CCSD(T) calculations and Dunning basis sets (cc-pVDZ, cc-pVTZ...).

$$E_{\text{corr}}(n) = E_{\text{corr}}(CBS) + An^{-\gamma} \tag{2.29}$$

Most of our calculations use the CCSD(T) method and the augmented Dunning basis; therefore, the extrapolation in Eq. (2.29) has been executed in that way, with few adaptations. The details of the calculations are in Appendix B. Correlation energies show an exponential dependence on the basis set cardinal number n as expressed by Eq. (2.29). That equation has an asymptotic limit, the energy at the CBS limit, $E_{\text{corr}}(CBS)$; therefore, it is pretty useful because that limit can be approximated using a fitting curve, using as the data set for fitting the energies calculated with different basis sets. The parameters obtained from the fitting would be $(E_{\text{corr}}, A, \gamma)$.

2.2.2 Basis Set Superposition Error (BSSE)

The binding energy is one measure of the strength of a chemical bond, and its definition is a subtraction between the equilibrium complex's energy and the energy of the fragments of that complex. What happens is that the ab initio energies of those fragments are higher than they should be since, in the complex, each monomer can "steal" the basis functions centered on the other monomers and lower the total energy in a way impossible for the isolated species. In other words, for the individual species, a smaller effective basis set is used, although the basis set elected for computations is the same because there are no functions centered in adjoining nuclei. The error, known as Basis Set Superposition Error, leads to over-stabilized energy of the complex.

Several ways to correct the overbinding caused by BSSE have been proposed, but the most well-known and used is Boys-Bernardi counterpoise correction [23]. The conclusion is that the effects of BSSE can be corrected by performing the calculations on the individual species with the same basis functions used for the complex, but it requires basis functions to be placed at arbitrary points in space; in fact, the points where the nuclear centers would be in the complex. Those are called "ghost functions".

When the CBS limit is made for the data, the BSSE is almost corrected. However, it appear a new type of error, Basis Set Incompleteness Error, that is due to the aplication of CBS limit. Either way, the errors using the CBS limit are not notorious.

When the ab initio data considers CBS limit and correct BSSE, it is said that those data are corrected. The ab initio PES we are calculating has uncorrected data, although the data are slightly affected by BSSE, according to the results of the references [6, 11]. Correcting the PES will be a suitable method to improve the results.

2.3 The interaction potential

The following information is mainly taken from the references [18, 24] unless another reference is given.

The molecular interactions can be classified into two major types: long-range, those interactions prevail at long distances and the energy of interaction behaves as some inverse power of r (distance between atoms or molecules depending on the system), and short-range interactions; where the energy decreases in magnitude approximately with exponential behaviour on r, so their effect is only noticeable in close enough molecules whose wavefunction's overlap.

When two molecular wavefunctions overlap, the electrons, indistinguishable particles, can be interchanged between one orbital and another, and the wavefunction has to be antisymmetrized. Among other short-range phenomena, a fraction of electronic charge could be transferred between

the molecules, which is the so-called charge-transfer effect, and it is a way to indicate the wavefunctions' overlap.

However, if the molecules are far apart, the overlap between their wavefunction can be ignored, antisymmetrization is unnecessary, and only long-range interactions are significant. In practice, the overlap is never precisely zero, although the error of considering so can be small enough to be neglected.

The long-range effects include electrostatic, induction (polarization), and dispersion interaction. The effective charge distributions of the two molecules affect each other through electrostatic effects: the Coulomb interaction. On the other hand, induction effects arise from the deformation of the charges that make up a particular molecule in the electric field of all its neighbours, which is always attractive. Dispersion is an interaction that arises because the molecules' charge distributions constantly fluctuate as electrons move.

Since intermolecular forces are relatively weak, they could be most naturally accounted for by the Rayleigh-Schrödinger perturbation theory. In a system of two molecules, A and B, distanced enough to not overlap, each molecule can be described by a Hamiltonian that considers its nucleus and electrons. The unperturbed Hamiltonian consists of adding the Hamiltonians of both molecules $H_0 = H_A + H_B$. Where H_A with N_A electrons and its corresponding nucleus and H_B with N_B electrons and nucleus. The perturbation, H', consists of the electrostatic interaction between the particles of molecule A and molecule B. The expression (in atomic units) is:

$$H' = H_{AB} - H_A - H_B = -\sum_{a\in A}\sum_{b\in B}\frac{q_a q_b}{r_{ba}} \tag{2.30}$$

where q_a and q_b are the charge of those particles in a. u. and r_{ba} the distance between the particles of A an those of B. H_{AB} is the Hamiltonian of the whole system.

Using perturbation theory over the Hamiltonian H_0 with the perturbation H' gives the interaction energy between the molecules A and B. The electrostatic interaction energy comes from the first order, and the second-order energy is the induction and dispersion interaction.

That way, a computable expression of induction energy can be obtained. Stone [24] develops induction energy through multipole expansion. However, if the intermolecular distance is too short, multipole expansion may not converge. Even so, the obtained expression of induction energy[1], taking into account the polarization of A over B and the same of B over A, is

$$U_{ind} = -\frac{1}{2}\vec{\mathscr{E}}^A_\alpha(\mathbf{B})\alpha^B_{\alpha\alpha'}\vec{\mathscr{E}}^A_{\alpha'}(\mathbf{B}) - \frac{1}{2}\vec{\mathscr{E}}^B_\alpha(\mathbf{A})\alpha^A_{\alpha\alpha'}\vec{\mathscr{E}}^B_{\alpha'}(\mathbf{A}) \tag{2.31}$$

where $\vec{\mathscr{E}}^A_\alpha(\mathrm{B})$ is the electric field at B generated by the molecule A and $\alpha^B_{\alpha\alpha'}$ is the polarizability tensor of the molecule B. Similarly, $\vec{\mathscr{E}}^B_\alpha(\mathrm{A})$ is the electric field generated by B over the molecule A, and $\alpha^A_{\alpha\alpha'}$ is the polarizability tensor of the molecule A.

The H_3^- molecule can be interpreted as a dimer formed by a H_2 molecule and an H^- anion interacting weakly, as it is usually done in the literature [6, 11, 12, 16], likewise it will be confirmed by our calculation.

Having said that, the molecule is a system formed by a polarizable molecule, H_2, and an ion, H^-, and it is a perfect example for an ion-induced dipole effect. An indicative expression of its

[1]For the interaction energy, the notation U has been taken to differentiate it from the energy of the molecular species and the energy from ab initio methods. This notation is taken from [24].

induction energy can be calculated using the equation (2.31), although the dominant term is the polarization of H_2 generated by the field of the H^- ion hence the second term of the expression (2.31) will be neglected. The following explanation is based on the reference [25]. Molecular polarizability is a second-order tensor due to the configuration and symmetry of the molecule, although it is a zero-order tensor in the case of atoms. The molecular tensor of the H_2 is diagonal if the molecular xyz frame is elected and chosen the origin at the centre of H_2 dimer. The polarizability in this molecular frame, where the x-axis has been chosen in the direction of the interatomic axis, in the same fashion as Fig. 4.1, has the expression

$$\alpha^{H_2} = \begin{pmatrix} \alpha_{\parallel}^{H_2} & 0 & 0 \\ 0 & \alpha_{\perp}^{H_2} & 0 \\ 0 & 0 & \alpha_{\perp}^{H_2} \end{pmatrix} \tag{2.32}$$

where $\alpha_{zz}^{H_2} = \alpha_{yy}^{H_2} = \alpha_{\perp}^{H_2}$ and $\alpha_{xx}^{H_2} = \alpha_{\parallel}^{H_2}$.

From Eq. (2.31), using the origin has the centre of the ion H^- and calculating the polarizability value, the expression obtained for the induction energy is;

$$U_{ind}^{H_2} = -\frac{1}{2r^4}(\alpha_{\parallel}^{H_2}\cos^2\theta + \alpha_{\perp}^{H_2}\sin^2\theta) \tag{2.33}$$

where r is the distance between the midpoint of the dimer and the anion, and θ is the angle that r forms with the bond axis of H_2. Considering Eq. (2.33), the induction energy depends on r^{-4} and induction is, excluding Coulomb interaction, the dominant long-range interaction. The long-range dependence with r of the interaction will be important to construct the analytical model of the potential.

For significative distance, the interaction energy of the H_3^- is the sum of two contributions: the electrostatic interactions between the charges of the dimer H_2 and the H^- anion, and charge-induced-dipole type:

$$U = \sum_{a\in H_2}\sum_{b\in H^-}\frac{q_a q_b}{r_{ab}} + U_{ind}^{H_2} + O\left(\frac{1}{r^5}\right) \tag{2.34}$$

where the first term corresponds to the Coulomb interaction between charge of the quadrupole moment of the neutral H_2 dimer and the anion, r_{ab} is the distance between such charges. $U_{ind}^{H_2}$ is the polarization of H_2 due to the ion-charge, and the last term represents higher order interactions, such as dispersion interaction going as r^{-6} and polarization over H^- due to H_2's quadrupole. However, those terms of higher order can be neglected in comparison with the dominant interactions.

2.3.1 Construction of an analytical potential energy hypersurface

Potential energy surface (PES) is the primary focus of most experimental and theoretical studies to discuss the stability, dissociation channels, and other molecule features. A PES depicts the potential energy as a function of the relative position coordinates of the nuclei, forming a complex hypersurface of $3N-6$ dimensions, where N is the number of atoms interacting in the molecule. The "hypersurface" refers specifically to those PES that depends on more than two degrees of freedom because they are multidimensional, and their are handy because gives information about

the optimal configuration, dissociation channels, equilibrium reaction constants, and the energy barriers. On the contrary, when the PES only depends on one degree of freedom, because the rest are fixed, it is called a potential energy curve. In principle, the potential energy surface can be obtained by computing, for all the possible configurations of the system, the eigenvalue of the electronic Schrödinger equation (2.1) for each fixed nuclear geometry.

Nevertheless, doing so with ab initio computations requires striving with calculations, especially with larger systems. Therefore, the usual procedure is choosing a parametrized model potential of a certain analytic form (physically consistent) and fitting the parameters, such that calculations yield the best agreement with the ab initio data. Therefore, the number of ab initio calculations is significantly reduced.

For the case of the trihydrogen anion ($N = 3$) the PES only depends on three dimensions or degrees of freedom. The sets of coordinates chosen to describe those degrees of freedom can be diverse. One of them could be the d,the distance between the two proton of H_2 and the Jacobo coordinates, r and θ, the modulus and orientation angle of the position vector of the third nucleus regarding the centre of mass between the other two nuclei, the midpoint of d.

Depicted the complete hypersurface would be computationally demanding, so a reasonable assumption has been used to avoid the wealth of calculations. Because the H^- presence does not significantly deform the H_2 moiety, as we will show later, the distance between the two nuclei, d, is taken as fixed. Consequently, the PES dependence is reduced to a set of two coordinates (r, θ). The goal is to design a PES realistic and physically admissible; therefore, these approximations and assumptions have to be well established.

The analytical model is chosen to represent the data accurately. Analytical functions are simple enough, but they can precisely describe repulsion-exchange effects or attractive ranges with a solid physical background. Choosing a specific analytical model can be risky because the model potential taken may have a form that is not entirely correct and not flexible enough to describe the ab initio calculations perfectly. Our election of analytical potential model will be justified in section 4.2.2 and the fit of the model to the data will be tested.

CHAPTER 3
THE NWCHEM QUANTUM CHEMISTRY PACKAGE

Resumen

El NWChem es el software de química computacional que se ha utilizado para realizar los todos cálculos ab initio del trabajo. Para realizar dichos cálculos es necesario aportar un fichero de entrada en el que se aporten ciertas órdenes utilizando una sintaxis específica que se explica en este capítulo.

Además, el Apéndice A guarda una estrecha relación con este capítulo, pues muestra ejemplos de los ficheros de entrada requeridos por el programa. En estos ficheros se observan algunas de las órdenes mencionadas en el capítulo.

For the ab initio calculations, the open-source computational chemistry software package NWChem was employed. It has been developed by the Experimental Molecular Science Laboratory (EMSL) at the Pacific Northwest National Laboratory (PNNL) [26]. The NWChem software contains computational chemistry tools for quantum mechanical calculations like SCF-HF, post-HF methods, and Density Functional Theory (DFT). The program is available on the node 44 of the computer cluster Molec3 in Department of Physics at the University of La Laguna.

It is necessary to create an input file to carry on any calculation. That file has to include some commands to define the system and specify the method we will use in the calculation. Furthermore, some additional commands are necessary if the system was an open shell. For the H_3^- calculations, no special commands have been used because the molecular anion has a closed-shell electronic structure in its ground state. Some of the most important commands are summarized in the following lines [27]:

- START + *name*: initiating and assigning a name to the job and the associated auxiliary files created during calculation.
- ECHO: Useful for knowing from what calculation the output file comes because it includes the input file at the beginning.
- GEOMETRY + *input units*: allows the user to define the geometry used for a given calculation. It requires a list of the atoms and the initial cartesian coordinates of each one.
  - SYMMETRY: Point group of the molecular geometry. It is not necessary for calculations but recommended if known.
- BASIS: Choosing the basis set used for the system. There is a basis set library included in NWChem.

- CHARGE: Charge of the ion.
- TASK + method + task: Select the method and the calculation task. Possible methods are SCF, MP2, CCSD(T) and CCSDT. There are some modules of NWChem that allow making that calculation, like TCE. However, it is possible to make it directly, writing that command at the end of the input file. This method has been used for the present calculations because it requires less memory cost. Among the available tasks, *optimize* determines the equilibrium energy by varying the geometry. Task *energy* calculates the energy for whatever configuration introduced. It is used to obtain potential energy curves for fixed angular arrangements.

An example of the input file, the files used for calculating the H_3^- optimization and energy, can be found in Appendix A. The structure of the input file and some of the commands named above are depicted in the appendix.

CHAPTER 4
RESULTS AND DISCUSSION

Resumen

Los resultados comienzan con la optimización de nuestra molécula principal H_3^- y sus productos de disociación H_2 y H^-. Estos últimos son importantes para conocer la energía de disociación de la molécula, que se ha de comparar con diversas fuentes bibliográficas para asegurar la validez del cálculo. Además, también se obtiene la configuración de equilibrio del anión de trihidrógeno, lo cual permite comprobar que el sistema se puede modelizar como la interacción entre el dímero H_2 y el ión H^-.

A partir de este punto, el desarrollo se basa en el diseño de la PES. Primero se define un sistema de coordenadas con el que se parametriza la PES y que permite definir curvas de energía potencial en función de una sola coordenada. Posteriormente, se construyen las curvas *ab initio* de energía potencial que se usarán para ajustar la función analítica del potencial modelo. Estas curvas refieren a la configuración lineal de la molécula y la configuración en forma de T. Se aporta un modelo analítico de potencial que cumpla ciertas características en su forma, relacionadas con argumentos físicos. Al ajustar el potencial analítico se utilizan los datos *ab initio* y se comprueba con éxito que la PES analítica es capaz de describir satisfactoriamente los datos *ab initio* para configuraciones con distinto ángulo.

This work aims to calculate the ground electronic state of the H_3^-. An ab initio potential energy surface will be obtained. For carrying out the ab initio calculations, we used the NWChem software. All calculations have been performed using the aug-cc-pV5Z basis set, which integrates 105 functions (6s,5p,4d,3f,2g) for Hydrogen. Our results have not been corrected using CBS limit nor BSSE correction; thus, our ab initio PES is uncorrected.

Before evaluating the results of our main molecule H_3^-, we will study the possible dissociation products H_2 and H^-. Therefore, the ground state equilibrium energy of the dissociation products has been calculated. Also, the bond length has been analyzed for the H_2. This calculation has the dual intention of testing and hierarchizing the different ab initio methods and obtaining the characteristics of these species, as it is necessary to construct the PES. The energy of the stable H_2 and the H^- are well-known results, so the values can be checked with the literature. Table 4.1 reproduces the optimized equilibrium ground state energies obtained by aug-cc-pV5Z basis set and the SCF, MP2 and CCSD(T). For H_2, the bond length d is also included. As CCSD(T) requires much computational time, the configuration resulting from MP2 has been used as an initial guess for the CCSD(T) computation.

Firstly, the results from the references [16] is used to check our calculations, carried out using CCSD(T) method and the aug-cc-pVQZ basis set. That way, a comparison confirms that the two results agree. On the other hand, in the reference [28] the energy has been calculated using CBS limit; therefore, their results of the equilibrium energy are slightly better. Other references [6, 11]

	H_2	H^-	H
	Equilibrium energy		
HF	-1.133651	-0.487897	-0.499994
MP2	-1.167455	-0.517694	
CCSD(T)	-1.174292	-0.527446	
Literature [28]	-1.174252	-0.527429	-0.500000
	Bond length (d)		
MP2	1.391		
CCSD(T)	1.401		
Literature [16]	1.4		

Table 4.1: Ground state equilibrium energy of the possible dissociation products and bond length of the H_2 using different ab initio methods, HF, MP2, CCSD(T) with the same basis set. The Hydrogen is a useful calculation to evaluate the precision of the basis set. The table is in a.u.

give similar values for the energy of the species and the bond length of H_2, but they have not been included for the sake of conciseness.

Checking with the literature, the bond length and equilibrium energy are well-reproduced with CCSD(T), but the result of the MP2 method shows a slight deviation. H_2 and H^- energy differ from the literature values at the fifth decimal using CCSD(T), but for the MP2 method, the difference begins at the second decimal. The results prove that the CCSD(T) method recuperates more energy than the MP2, as was said above. Our results of H_2 and H^- energies agree with the values accepted in literature, although there is a slight deviation at the fourth decimal figure. For the bond length of H_2, our results reflex the literature value too. Among the methods we used, CCSD(T) is the best method, and HF is the less accurate one, as expected.

For Hydrogen in table 4.1, the energy is the same whatever the method because it is a one-electron system, and electron correlation does not exist. However, the calculated energy value of H is almost equivalent to the exact one[1]. The deviation appears in the sixth decimal, almost a negligible deviation. That means that the aug-cc-pV5Z is a suitable basis set for the computations because its results are precise. Considering the results, CCSD(T) is the most accurate method for multi-electron systems; therefore, it has been used in the following calculations.

4.1 Trihydrogen anion H_3^-

The CCSD(T) method and basis set aug-cc-pV5Z predicts that the Trihydrogen anion has a linear but asymmetric geometry, which resembles all of the studies reported previously about the molecule [4, 6, 12, 16, 28]. The point group of that anion is $C_{\infty v}$ and its electronic ground state is noted $^1\Sigma$.

The table 4.2 summarizes the equilibrium configuration of H_3^- at the electronic ground state and energy value obtained from the ab initio optimizations. Excited states of H_3^- also exist, as the references [7, 16] represent, but that states are beyond the scopes of this work.

[1] Hydrogen's energy exact value is the Rayleigh constant in the proper units, a.u.

	Internuclear distance $H-H$, d	r_0^e	Equilibrium energy	Dissociation energy, D_e
Linear asymmetric configuration ($^1\Sigma$)	1.415	6.095	-1.703535	$1.832{\cdot}10^{-3}$
Literature [11]	1.421	6.069	-1.703511	$1.830{\cdot}10^{-3}$

Table 4.2: Equilibrium energy and geometry in the H_3^- electronic ground state calculated with CCSD(T) method. The dissociation energy D_e is included. Also included are the distances d, the shortest distance separating two nuclei, and r_0^e, the distance from the midpoint of d to the anion at equilibrium configuration. The results are in a.u.

An expected result in the equilibrium configuration, table 4.2, the distance d is nearly equal to the bond length of the isolated H_2 molecule. Because the H_2 moiety is not significantly deformed by the H^- presence, the distance d is fixed for the rest of the calculations. That is the reason, along with almost no charge-transfer that is neglected, for the potential modelled as an anion-dimer interaction.

The distance r_0^e in table 4.2 refers to the distance between the H^- and the center of mass of H_2 when the molecule H_3^- is at equilibrium configuration. In the Fig. 4.2 r_0^e is a significant value because it is the distance of the minimum of energy.

When H_2 and H^- are infinitely apart, that means free H_2 and free H^-, their energy is the energy of the dissociation products, $E(H_2)+E(H^-)$. The energy difference between the equilibrium energy of H_3^- and the energy of the dissociation products, is the energy needed for dissociation, D_e in table 4.2. For the reference's results [11], summarized in table 4.2, the dissociation energy coincide until five decimal figure. Considering the results in table 4.2, our calculations are in great, almost perfect, agreement with the values from the literature [11]. The reference used the CEPA method, a variation of CI ab initio calculation. However, our D_e may be not exact due to the uncorrected BSSE in our data. The BSSE produces the effect that the H_3^- energy is over-stabilized, but considering the great agreement with the literature value, it seems to be a slight deviation. Nevertheless, further comparison with other references has been made, as it is collected in table 4.3.

Other references such as [6] and [12] has similar results. The reference [6] employed the MP2 and MP4 method, while the work [12] used CCSD(T) for this calculation and corrected the results to the CBS limit. However, another reference [28] has a result that does not match ours, even when it also used CCSD(T) calculation and CBS limit correction. All the results mentioned from the literature have been summarize in table 4.3.

	Our results CCSD(T)	**CCSD(T)**	**DMC**	MP2/MP4	**DFT**
$D_e(meV)$	-48.87 -47.37	-46.27[12] -25.71[28]	-35.78[28]	-50.31[6]	- 73.26[29]
$r_0^e(\AA)$	3.27	3.32[28]		3.15[6]	2.88 [29]

Table 4.3: Our data, first column, are depicted for the comparison with the literature values. In our values of the energy D_e the first one is using a basis set aug-cc-pV5Z, while the second value is the result taking the CBS limit (details of the calculation in Appendix B). The D_e and r_0^e from the literature are organized according to the ab initio method used in each calculation. Units in meV for energy and $\AA$ for distance, the units change is due to the D_e order, much more lower than the hartree unit.

An interesting result is that our D_e using the aug-cc-pV5Z basis set is lower than the value obtained by the CBS limit in Appendix B. That seems to be a consequence of the Basis Set Superposition Error since it is an uncorrected value. Therefore, the improvement of the results using the CBS limit is confirmed. The other references that use a different ab initio methods, like [29], are useful for checking that the results obtained in this work are roughly in agreement with the values that the literature accepts. In conclusion, our D_e and r_0^e values falls in a reasonable range, considering the literature values.

4.2 Potential energy hypersurface (PES)

It has been briefly mentioned before, but, in the case of the H_3^- molecule, the distance between the third nucleus and the other two is far enough to assume no charge-transfer. Then, the total potential is considered a two-body potential between H_2 and H^-. That implies that H_3^- is held together by intermolecular attractions such as Van der Waals forces instead of chemical bonding. Van der Waals forces are weaker than covalent or other bonding; therefore, the dissociation energy of the molecule is expected to be a relatively low value when compared with the electronic energy levels. The assumption of no charge transfer between H_2 and H^- is not scrupulously valid, but the charge-transfer effect is insignificant enough to be neglected, as is done in the abovementioned literature for this molecule.

Then, to obtain a PES, it is important to establish the system's degrees of freedom, as mentioned in section 2.3.1. The three degrees of freedom will be selected (d, r, θ). The bond distances d and r are the internuclear distance of the H_2 and the distance from the center of mass of the dimer to the H^- anion centre, respectively. Futhermore, the angle θ describes the relative orientation between d and r. All the coordinates are depicted in Fig. 4.1 and the data of table 4.2 is designed to correspond with this system of coordinates, but the special notation r_0^e has been used to design the distance r at the equilibrium configuration.

The approximation of taking H_2 as a rigid molecule allows us to take the distance d as fixed. Therefore, the PES is a function of two Jacobian coordinates: r, from the centre of mass of H_2 to the anion H^- and θ angle between $\mathbf{r}$ and the interatomic axis of H_2.

The potential energy curves, that only depends on one coordinate, will be obtained by fixing angular arrangements, such as the curves for (r, 0°) or (r, 90°) .

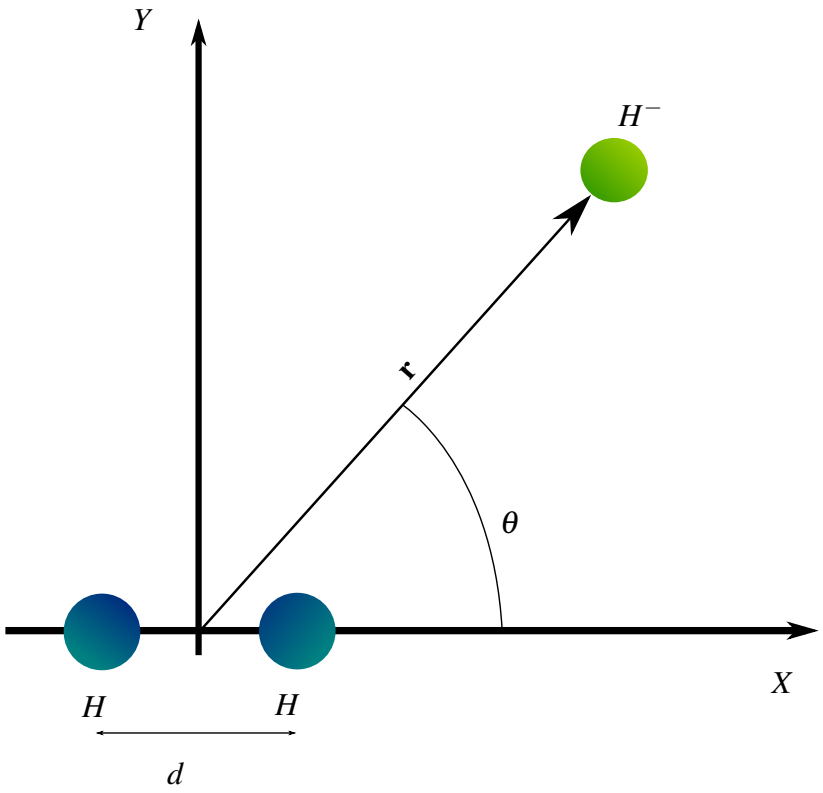


Figure 4.1: Graphic representation of the degrees of freedom of the system. The origin is placed at the center of mass of the dimer H_2 and the angle $\theta = 0°$ is the linear configuration of the molecule, while the $\theta = 90°$ is the T-shape configuration.

4.2.1 Ab initio potential energy curves

The first step we have taken in constructing an ab initio PES is to draw a potential energy curve starting from the equilibrium configuration of the molecule. Since the system is known to be an interaction of a dimer with an anion, a potential energy curve can be drawn by moving one of the nuclei, the anion, closer and further away without breaking the linear configuration of H_3^-. In other words, the potential energy curve correspond to the polar angle fixed as $\theta = 0°$. Since the curve includes the equilibrium configuration, where the energy is minimal, the shape of the curve is expected to be that of a repulsive potential at a short distance, with a minimum at the equilibrium configuration tending asymptotically at long distances to the energy of the dissociation products. That curve is depicted in Fig. 4.2 where a systematic calculation for different molecular configurations has been done[2].

The potential curve of the T-shaped distribution has also been obtained in Fig. 4.2. In that configuration, the hydrogen is distributed at the vertices of an isosceles triangle. The interest in obtaining that curve is to confirm that a triangular configuration is not stable. Since the cation H_3^+ has has a equilateral triangle equilibrium geometry geometry, it was worth studying if a local minimum could exist for that configuration of the anion. In the light of the results, neither a stable

[2]The data depicted in Fig. 4.2 appear in Fig. 4.3 too. It has been included basically for clarification about the difference between ab initio curve and parametrized PES.

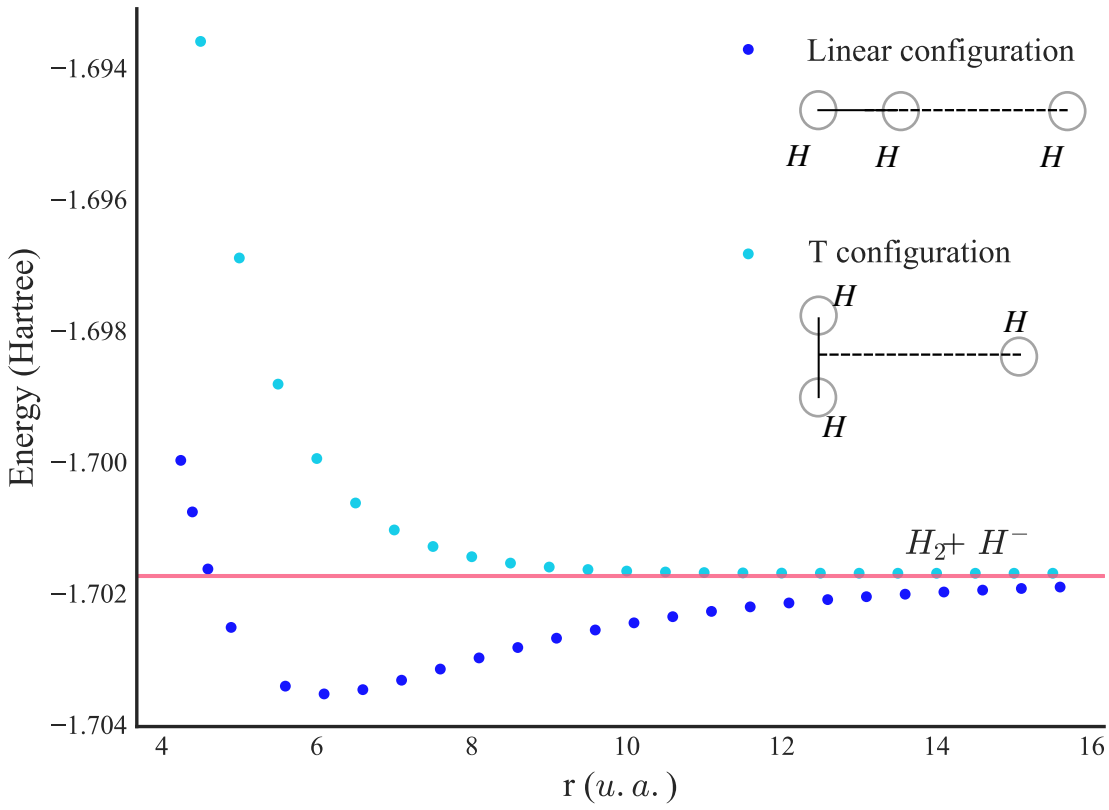


Figure 4.2: H_3^- at electronic ground state, potential energy curves as function of the distance r. The distance d remains fixed at the H_3^- equilibrium configuration. The asymptotic limit represented with a red line refers to the energy of a free H_2 and H^-, value taken from the table 4.1. In the configurations depicted at the side of the graph, the dashed line is the coordinate r while the solid line is d.

triangular configuration nor minimum energy appears; therefore, there is no dissociation energy in that direction, and the potential is purely repulsive. The anisotropy of the molecule acts in a way that not only the linear configuration is the most stable one, but also no other distribution is possible.

Both curves have the same asymptotic limit, highlighted with a red line in the Fig. 4.2, which is congruent with the dissociation channel of the molecule. However, a small distance between the points and the axis remains, and it can be explained because the distance d, table 4.2, is not exactly the same as the bond length of H_2; it differs at the second decimal figure from the equilibrium distance of the dimer H_2, table 4.1.

The ab initio energy as a function of r for the two relevant configurations is useful for the design of the analytical potential model. In the next sections, the process to obtain the adequate potential and the fitting of the parameters is described in detail.

4.2.2 Analytical potential model

First, our analytical potential model has to describe the interaction term of the PES as a function of the coordinates of the system.

For the analytical model that should describe the ab initio PES some references [13–15] have been consulted. Those are about the analytical model in different systems, some of them similar to H_3^- and their usage of it. In addition, other sources about H_3^- analytical and ab initio PES have corroborated the physical assumptions and approximations carried out [6].

The Hamiltonian for the interaction of this system is formed by an electrostatic term and another non-covalent term:

$$V(r,\theta) = V_{elect}[H_2 - H^-] + V_{Nc}[H_2 - H^-] \tag{4.1}$$

It is also known that, at a long distance, the form of the energy has to be similar to Eq. (2.34).

The potential will have an electrostatic part, the first term in Eq. (4.1), calculated as a Coulomb interaction between the charges that make up H_2 and H^-. H_2 would be modelled by effective charges: two positive charges +Q in the position of the nuclei and a negative one, -2Q, in the center of mass of H_2. Q is the effective charge used for modelled H_2 and has the value $Q = 0.48226$ in a.u., and k is the Coulomb constant in the proper units. The values of these charges have been chosen to reproduce the quadrupole moment of the neutral H_2 molecule, as described in [15]. On its part, H^- would be a negative charge. The Coulomb interaction has the analytical expression:

$$V_{elect}(r,\theta) = -Q\cdot k\left(\frac{1}{\sqrt{r^2+d^2-2dr\cos\theta}} + \frac{1}{\sqrt{r^2+d^2+2dr\cos\theta}} - \frac{2}{r}\right) \tag{4.2}$$

The Coulomb interaction also depends on the distance d, but since it is a fixed value it is not taken a variable of the function $V_{elect}(r,\theta)$. The Eq. (4.2), when the approximation $d \ll r$ is used, recovers the charge-quadrupole interaction.

The non-covalent term, second term in Eq. (4.1), contains the induction interaction and the interaction due to Van der Waals forces. This interaction has a short-range repulsion-exchange interaction, and at long-range attraction, where the dominant term is due to the polarization of H_2. The atom-bond pairwise additive model has been used to represent this term analytically. This model has been built with an improved Lennard Jones model that also considers anisotropy.

A simply Lennard Jones function had not been adequated because it would have been excessively repulsive at short-range; the repulsion part of LJ goes with r^{-12}, in other words, $n = 12$. The LJ model also describes the attractive part with a r^{-6} dependence, $m = 6$. The improved Lennard Jones allows other values of n and m and the atom-bond model. The following expression taking from Pirani´s atom bond model ([14]) has a n depending on θ and r:

$$V_{Nc}(r,\theta) = \varepsilon(\theta)\left[\frac{m}{n(r,\theta)-m}\left(\frac{r_e(\theta)}{r}\right)^{n(r,\theta)} - \frac{n(r,\theta)}{n(r,\theta)-m}\left(\frac{r_e(\theta)}{r}\right)^{m}\right] \tag{4.3a}$$

$$n(r,\theta) = \beta + 4\left(\frac{r}{r_e(\theta)}\right)^2 \tag{4.3b}$$

The angular dependence in the equations depicts as:

$$\begin{aligned} \varepsilon(\theta) &= \varepsilon^{\perp}\sin^2(\theta) + \varepsilon^{\parallel}\cos^2(\theta) \\ r_e(\theta) &= r_e^{\perp}\sin^2(\theta) + r_e^{\parallel}\cos^2(\theta) \end{aligned} \tag{4.3c}$$

For the Eq. (4.3a), the value of the exponent $m = 4$ is used, which corresponds to a charge-induced dipole interaction. The reason for this is the form of the long-range interaction, Eq. (2.34), justified in the previous chapter.

This analytical model is computationally simple and has been used in molecules with a similar structure to H_3^-, [15]. However, for this anion, it had yet to be implemented, although the results are expected to be satisfactory because the atom bond pairwise additive approach adequately describes the anisotropy of the molecule.

The parameters ε and r_e correspond to the depth of the energy well and its position, respectively. As the molecule is anisotropic, it is necessary to know these parameters in the direction parallel ($r_e^{\parallel}$, $\varepsilon^{\parallel}$) and perpendicular to the intermolecular axis of H_2 ($r_e^{\perp}, \varepsilon^{\perp}$). The dependence of ε and r_e on the angle allows an accurate calculation of the energy values for any angles. Likewise, the parameter β in Eq. (4.3b) is related to the hardness of the potential barrier.

The Eq. (4.1) in combination with Eq. (4.2) and Eq. (4.3), that each one represents one term of interaction, yield to:

$$V(r,\theta) = V_{elect}(r,\theta) + V_{Nc}(r,\theta) \tag{4.4}$$

where $V(r,\theta)$ wil be the interaction potential energy.

In Eq. (4.4), specifically the atom bond potental term, five parameters have to be obtained ($r_e^{\parallel}$, $\varepsilon^{\parallel}$, $r_e^{\perp}$, $\varepsilon^{\perp}$, β). To achieve the value of the parameters a parallel approximation is useful. That means that, firstly, we make $\theta = 0^\circ$ in Eq. (4.3c), the the expression is simplified like

$$\begin{aligned} \varepsilon(0^\circ) &= \varepsilon^{\parallel} \\ r_e(0^\circ) &= r_e^{\parallel} \end{aligned} \tag{4.5}$$

that potential energy curve for $\theta = 0^\circ$ can be described using the Eq. (4.4). To the term of Eq. (4.3) the simplification of Eq. (4.5) is used; therefore, ILJ only depends of three parameters ($r_e^{\parallel}$, $\varepsilon^{\parallel}$, β).

At that point, the ab initio curve for the linear configuration, depicted in Fig. 4.2, is useful. Once the Coulomb interaction of Eq. (2.11) and the dissociation products' energy are removed, that ab initio data correspond to the non-covalent interaction, represented by the atom bond model, Eq. (4.3). These data are used to fit the ILJ model and obtain the triad of parameters via optimisation.

The later scheme is done with the potential energy curve $\theta = 90^\circ$ too. This time, the ab initio curve is the T configuration of the Fig. 4.2. Now the remaining parameters in Eq. (4.3c) are

$$\begin{aligned} \varepsilon(90^\circ) &= \varepsilon^{\perp} \\ r_e(90^\circ) &= r_e^{\perp} \end{aligned} \tag{4.6}$$

The parameters obtained with that curve are ($r_e^{\perp}$, $\varepsilon^{\perp}$), β has already been calculated so its value is fixed. The ab initio data corresponding to the non-covalent interaction is fitted to obtain the rest of the parameters.

A Python programming was designed for solving the curve fitting by least-square, using the SciPy library. SciPy is an open-source library used for solving mathematical, scientific and engineering problems. Specifically, package optimize was used for the fitting.

4.2.3 Obtaining model parameters: Fitting ab initio uncorrected potential curves

The Potential Energy Curves used for the parallel approach and the parameters obtained are depicted in Fig. 4.3. The behavior of both curves has an important difference, the linear configuration has a minimum, but the T-shape configuration has not. The potential curve behavior of the T-shape configuration is due to the electrostatic interaction, strongly repulsive in this configuration, that domains over the rest of the effects.

The ab initio data help find an optimal set of parameters for the expression. The non-linear least-squares procedure is used to fit the ILJ potential model as described above.

As mentioned above, the potential energy curves have been calculated with ab initio methods CCSD(T) using the base aug-cc-pV5Z. The energy values are not corrected with the CBS limit, nor is BSSE eliminated. Thus the ab initio PES is uncorrected.

Firstly, the parameters obtained with optimization are β, $r_e^{\parallel}$ and $\varepsilon^{\parallel}$ from Eqs. (4.5, 4.3a). For the atom-bond model, the fitting to the ab initio data of triangular configuration has to be performed similarly. Nevertheless, the parameter β depends on the type of molecules, so it has to be the same without matter the direction. Therefore, that is the reason β is not optimized again. Instead, the optimal set of parameters is $r_e^{\perp}$ and $\varepsilon^{\perp}$. The results of these calculations are summarized in table 4.4.

In the following the units will be $\AA$ for distance and meV for energy insted of the a.u. used until now. The equivalence between the units is 1 Hartree= 27211.38 meV and 1a.u.= 0.529177$\AA$. The change is due to the next results are energies of the order of D_e, energies really small compared with Hartree unit. This change of units is taking into account the usual unit system used in the literature.

	ε	r_e	β
Parallel approach ($\parallel$)	4.156925	10.948366	3.109370
Perpendicular approach ($\perp$)	4.579968	6.247498	

Table 4.4: Values of the parameters ε, r_e and β obtained by the curve fitting, using least-squares procedure. The fitting is been used for the linear and the T-shape configuration, although β is a optimized parameter just the first time. The β value is fixed for the fitting of the T-shape data. The units of the table are meV for energy, ε, and $\AA$ for distance, r_e. β is an adimensional parameter.

The β value in the literature [14, 15] is between $2 < \beta < 10$ depending on the element of the atoms. The β obtained in the fitting, depicts in table 4.4, is congruent and has the same order as the literature example.

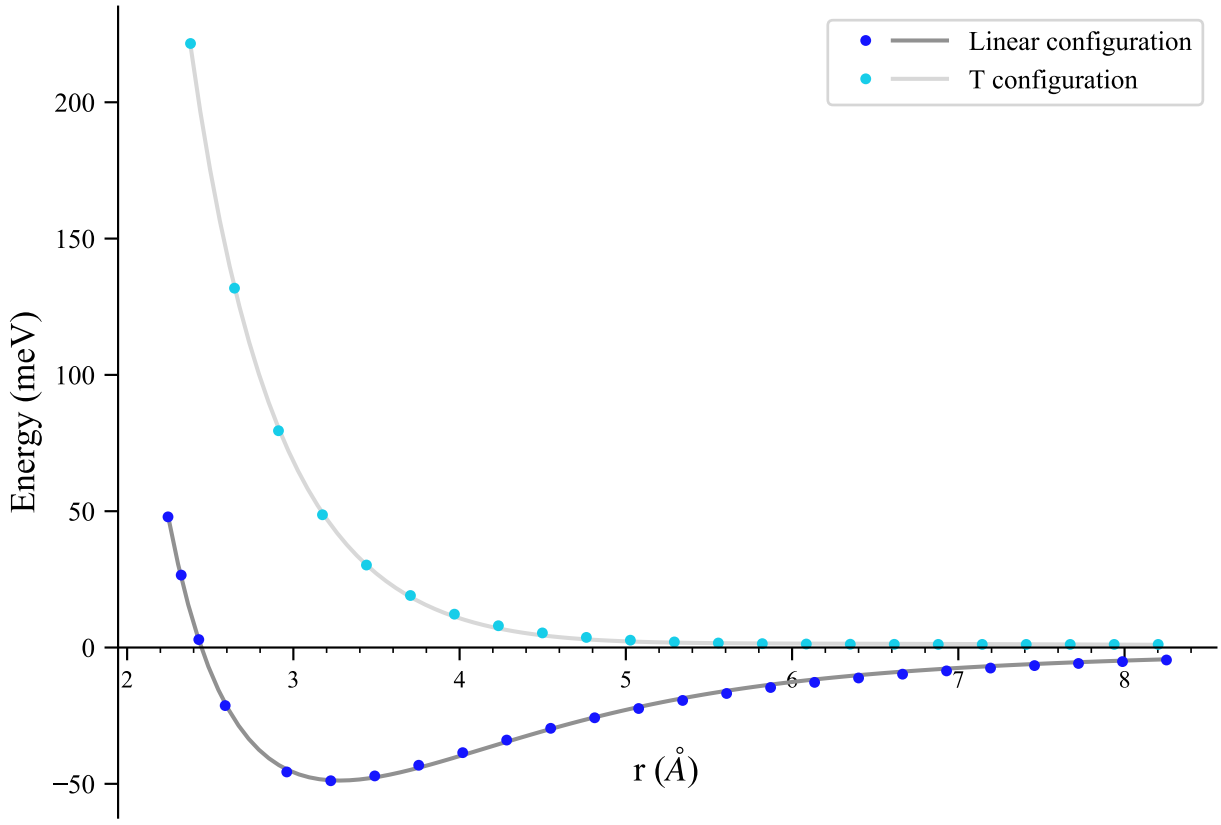


Figure 4.3: H_3^- ground state potential energy curves as function of the distance r depicted in Fig. 4.1. The distance d remains fixed. The zero level corresponds to the energy of a free H_2 and H^-, from the table 4.1. The ab initio data of each potential curve are depicted together with the associated fitted function.

The points obtained from ab initio calculations and the fitted potential is depicted in Fig. 4.3. The fitting is excellent, and, in both directions, the points' behavior is adequately described.

Eq. (4.4) gives the analytical function for the PES, built with the optimal parameters. Given the angular dependence of ε and r_e in the potential model, Eq. (4.3c), the parametrized PES has to accurately describe the ab initio calculation for different values of θ.

4.2.4 Accuracy of atom-bond potential for angular configurations

The Eq. (4.4) give an analytical PES can be constructed depending on r and the angle θ. The ab initio calculation for selected angles has been done. These results, alongside the analytical curve for these angles, are depicted in Fig. 4.4.

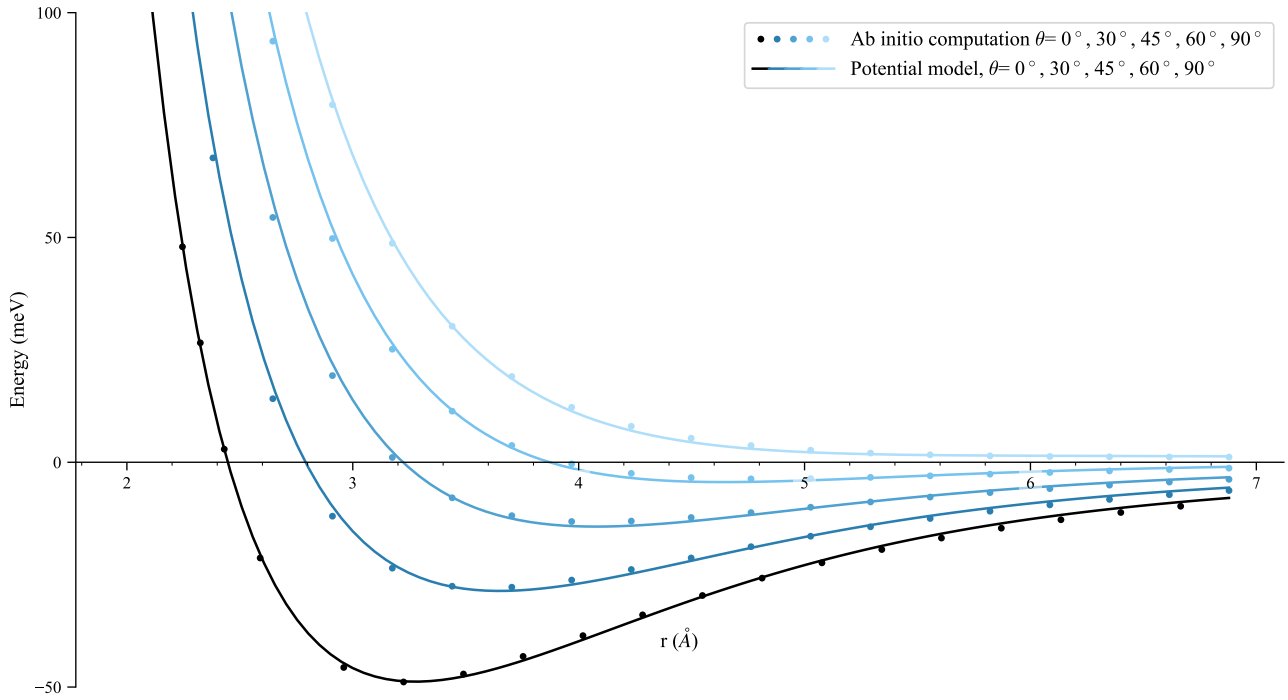


Figure 4.4: H_3^- ground state potential energy as a function of the distance r for fixed angular arrangements, following the diagram 4.1. The five potential curves (points) represent the ab initio results for each θ ($\theta = 0°, 30°, 45°, 60°, 90°$) together with the analytical function for each angle (solid line). The fitted parameters in table 4.5 are used to calculate the behaviour of the analytic function. The distance H-H is fixed in the H_3^- equilibrium configuration

	$\mathbf{r_0}(\AA)$	V_e (meV)	**Analytic function value** (meV)
$\theta = 0°$	3.225	-48.871	-48.714
$\theta = 30°$	3.704	-27.792	-28.573
$\theta = 45°$	3.969	-13.182	-14.193
$\theta = 60°$	4.763	-3.701	-4.362

Table 4.5: Energy minimum of each potential curve depicted in Fig. (4.4). r_0 corresponds to the distance of the minimum energy in ab initio data for each curve, unit $\AA$. ε is the depth of the potential well considering ab initio data, in meV. The last column is about the analytical function in meV, Eq. (4.4).

The fitting is almost flawless except for the slight deviation at a short distance in the repulsion section. Before the energy minimum, the fitted potential does not match the data exactly, in Fig. 4.4; therefore, the obtained analytical PES would be unsatisfactory for studies about short-distance effects, such as the scattering process [5, 6]. It was to be expected because the short-range interaction has an exponential expression, and the ILJ, although describing very well the general behaviour, only depends on the inverse of the distance. Putting aside that limitation, the behaviour

	Analytic function value (meV)	**Electrostatic contribution** (meV)	**Non-covalent term** (meV)
$\theta = 0°$	-48.714	-58.854	10.140
$\theta = 30°$	-28.573	-23.960	-4.614
$\theta = 45°$	-14.193	-7.674	-6.519
$\theta = 60°$	-4.362	2.285	-6.647

Table 4.6: The table shows the analytical PES in the minimum of each curve. Therefore, the results are about the analytical function in the meV unit. The first one is the value of the analytical potential model, Eq. (4.4). The second and third ones are the the electrostatic contribution, Eq. (2.11), and the atom-bond pairwise contribution, Eq. (4.3).

is globally satisfactory, and the process seems valuable considering the few parameters fitted. The behaviour of the anisotropy is well described with the atom-bond pairwise additive approach.

Compared with PES in the literature, reference [12] has a similar goal. However, their PES has corrected to the CBS limit, and the analytical model lacks any explicit angular-dependent term, although it has more terms in the potential model than ours. In our view, the accuracy of our analytical PES is better than the analytical potential model for this molecule found in reference [12], especially in the configurations with higher θ, the more repulsive configuration.

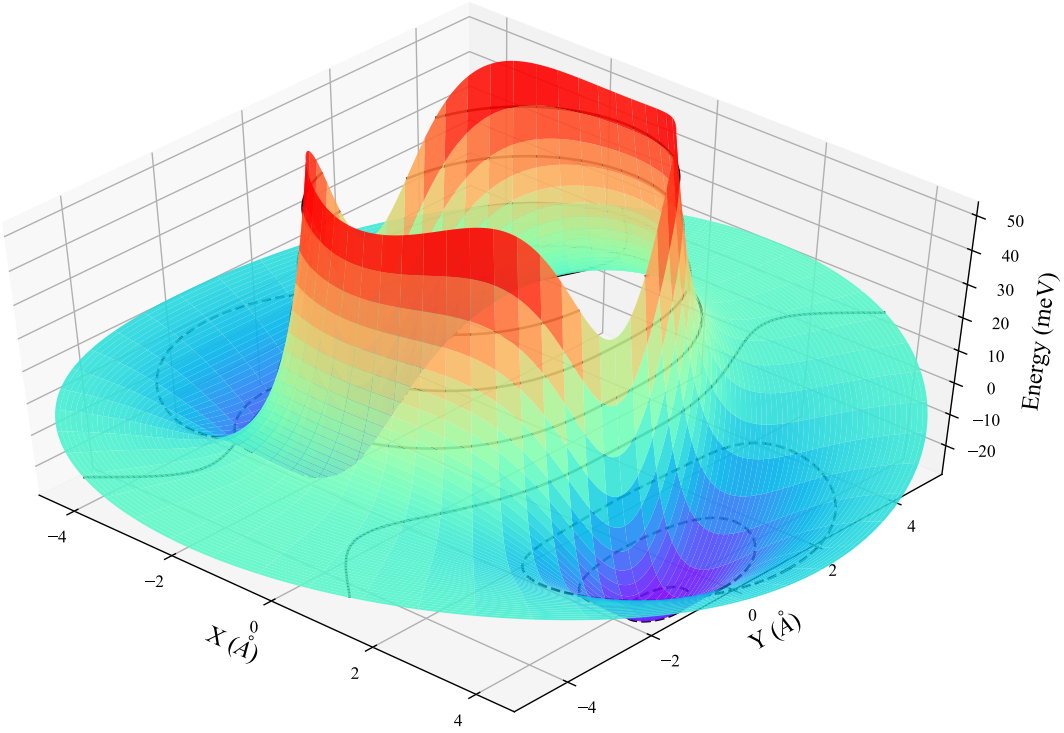
Considering the results in table 4.5, it is not unexpected how the ab initio potential well, V_e is less deep and farther away with bigger angles. The table C.1 in Appendix C serves to illustrate this quantitatively. The potential curves are gradually changing to a purely repulsive interaction, the behaviour at $\theta = 90°$, and, considering the Fig. 4.4, the limit angle θ is barely over $60°$. Considering the contibutions of the analytical PES in table 4.6 it makes obvious that the Coulomb interaction domains over the non-covalent term.

The deviation between the ab initio data and the analytical function, whose values can be compared in table 4.5, is really slight. The values differ by about $1meV$ at most. Therefore, the proposed function form of the PES is quantitatively a good method to describe the interaction.

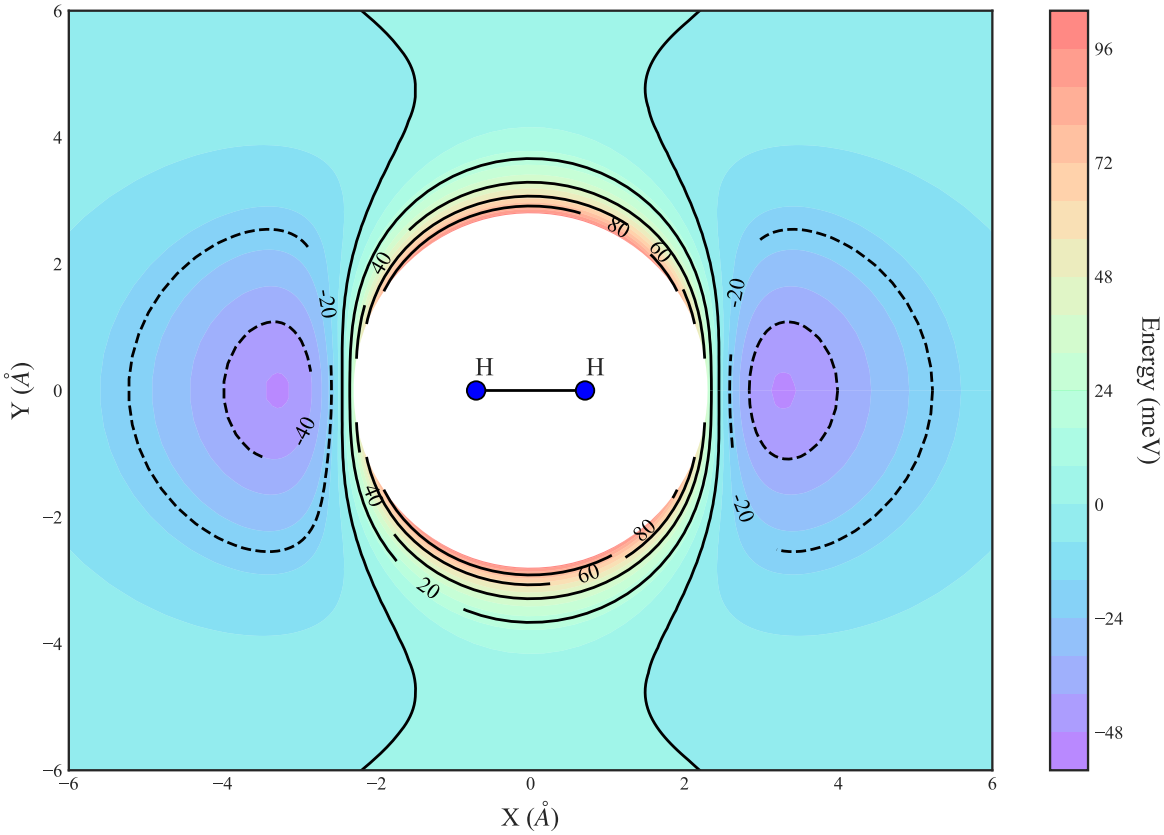
4.2.5 Analytical potential energy surface

The analytical potential model reproduces the molecule's behaviour very well; therefore, it is possible to depict the potential energy surface, PES. Being the origin of the centre of mass of H_2, the PES for the H_3^- molecule is the Fig. (4.5a). The figure depicts a favourable direction for the H_3^- as the direction of linear configuration with two wells. Both wells are due to the same linear geometry, but the molecule's symmetry makes two equivalents configurations. On the contrary, the triangular configuration, or T-shaped, appears with a purely repulsive behaviour. Consequently, the well smoothly disappears when bigger angles θ are analyzed.

Figure 4.5: Representation of the analytical PES parameterized



(a) Representation of the H_3^- ground electronic state analytical PES as a function of the plane XY. Value of the energy using the analytical potential model, Eq. (4.4).



(b) Equipotential lines of the analytical PES for H_3^-. The orientation of the H_2 molecule is shown in the center of the figure.

Also, the shape of the equipotential lines can be obtained using a contour representation, Fig. 4.5b. That way, the wells' shape, and depth can be roughly observed. In addition, it is notorious how the wells are gradually flattened. It is easy to notice that at T-shape and even lower angles the potential is purely repulsive.

CHAPTER 5
CONCLUSIONS

Resumen

En este capítulo se repasan los resultados más importantes del trabajo, dando una visión global de lo que se ha logrado en el desarrollo del mismo.

Aunque los resultados son en general impecables y se muestran de acuerdo con la mayoría de las fuentes bibliográficas, aún aparece el impedimento de considerar el costo computacional de los métodos ab initio, en vista a una mejora de resultados.

También se esbozan algunas ideas sobre cómo se podría mejorar la exactitud de la PES y se proponen nuevos sistemas sobre los que proseguir el estudio.

In the light of the results in table 4.1, the different methods, HF, MP2 and CCSD(T), have an obvious hierarchy. The improvement in the quality is especially significant with larger systems. For instance, for the H_3^- equilibrium energy, table 4.2, the importance of electron correlation is clear and the post-HF methods obtained a better result than HF. CCSD(T) is proved to be the best method for these calculations, more energy was recuperated in table 4.1, and for that reason, it was selected over the MP2 method.

Even more, using the aug-cc-pV5Z basis set in combination with the CCSD(T) method the value of H_2 equilibrium energy is accurate up to the fifth decimal figure, table 4.1. This result is well enough, but there is always room for improvement. The largest basis set aug-cc-pV6Z could be used. The aug-cc-pV6Z basis set was suggested as a candidate for this work calculations, but it was finally discarded due to the computational cost. However, it is undeniable that this basis set would improve the calculations. As commented before, a way to make the calculations more precise is to correct them by using the CBS limit and eliminating the BSSE.

The ground state energy of the trihydrogen anion H_3^- follows the literature values. The molecule's geometry differs a little from the literature result [28], but considering several references [6, 12, 29] our configuration falls in an acceptable range of accurateness. The calculations of the potential curve for different configurations, with different values of θ, have made it possible to obtain a potential energy surface representing the ground electronic state of the H_3^-. A natural suggestion is trying to do so with excited states of H_3^-. Furthermore, a possible next step would be searching the equilibrium configuration of other molecules H_n^- like in the references [28, 29].

The analytical potential model fitting goes beyond expectations. Almost all the ab initio points fall nicely on the analytical curve except for some points at a short distance. The potential model reproduces the data even better than some examples in the literature, an impressive result taking into account the simplicity of the atom-bond pairwise approach. Maybe some variations in the potential could describe the data better, but using the potential model in more complex systems would be even more interesting.

The computation cost is a point to ponder while calculating the H_3^- molecule. Even if the anion is a relatively small molecule, the computation time for the CCSD(T) optimization was impressive, about one day more or less. The problem arises when using large basis sets like aug-cc-pV5Z or aug-cc-pV6Z, although it also depends on the method. The computation time needed for each method is proportional to N^4 for HF, N^5 for MP2, and N^8 for CCSD(T), which is N the number of functions in the basis set. Therefore, is not unexpected, attending to table 2.1, that the difference in computational cost is significant between one basis set and the next. A way to achieve accurateness is to extrapolate to the CBS limit and correct the BSSE, as in the Appendix B. However, it has been done only for the equilibrium configuration. An arguably way to improve the result would be to extend the CBS limit and correct BSSE to all the potential energy curves and the potential energy surface.

APPENDIX A
SAMPLE OF NWCHEM INPUT FILE

```
memory heap 200 mb stack 1000 mb global 2800 mb
start H3_anion
title "H3-"
echo
charge -1
geometry units au
H     0.00000000     0.00000000    -4.18600000
H     0.00000000     0.00000000     1.38000000
H     0.00000000     0.00000000     2.38000000
end
basis
  H library aug-cc-pV5Z
end
task CCSD(T) optimize
```

```
memory heap 200 mb stack 1000 mb global 2800 mb
start H3_anion
title "H3-"
echo
charge -1
geometry units au
H     0.00000000     0.00000000    -4.56281973
H     0.00000000     0.00000000     1.32378342
H     0.00000000     0.00000000     2.73903630
end
basis
  H library aug-cc-pV5Z
end
task CCSD(T) energy
```

APPENDIX B
CBS LIMIT FOR H_3^- EQUILIBRIUM CONFIGURATION

The equilibrium energy of H_3^- has been calculated for different basis set cardinal numbers n (n=2, 3, 4, 5) using the CCSD(T) method. The aug-cc-pV6Z basis set has not been employed because it required too much computational cost due to its large extension. The energy value for the 6Z basis can probably be obtained by approximation, but this is beyond the scope of this thesis.

The curve fit has been calculated once again using Python programming language and the non-linear least-square process. The optimized parameters set in Eq. (2.29) is A, γ and $E_{tot,CBS}(H)$. Table B.1 depicts the results of the parameters optimization. The $E_{tot,CBS}(H)$ is an asymptotic limit and the representation of that fitting curve is shown in Fig. B.

	$\mathbf{E_{tot,CBS}(H)}$ (Hartree)	**A(H)**	γ	$\mathbf{D_{e,CBS}}$ (meV)
H^-	-0.527660	0.028624	2.979576	
H_2	-1.174477	0.178897	4.223246	
H_3^-	-1.703878	0.193666	3.881485	47.37

Table B.1: Energy of H^-, H_2 and H_3^- evaluated at the complete basis set limit. The optimized parameters of the Eq. (2.29) are summarized for each molecule because their usefulness for the CBS limit calculation. $\mathbf{E_{tot,CBS}(H)}$ is the energy at CBS limit, in hartree., while A and γ are parameters of the equation, without physical meaning. Dissociation energy in meV has also been included.

Considering the results in Fig. B, the aug-cc-pVnZ basis set family, except in the H^- case, fall nicely on the curve. The deviation of the H^- from the asymptotic limit was expected but also implied that the energy value of the 6Z basis set may be significant to correct the equilibrium energy at the CBS limit.

The dissociation energy of H_3^- evaluated in the CBS limit appears in the literature [12] as $D_{e,CBS}(H_3^-) = -46.27meV$. That is congruent with our result in table B.1. However, another reference [28] upholds that at the CBS limit the dissociation energy is $D_{e,CBS}(H_3^-) = -25.71meV$. The value falls well below our results in table B.1. Seems that the source has underestimated the equilibrium energy at the CBS limit for some reason.

Since the equilibrium energy at the CBS limit is more accurate than that energy using the 5Z basis set, it is expected that the potential curve being again evaluated at the CBS limit will generate a better PES for H_3^- molecule.

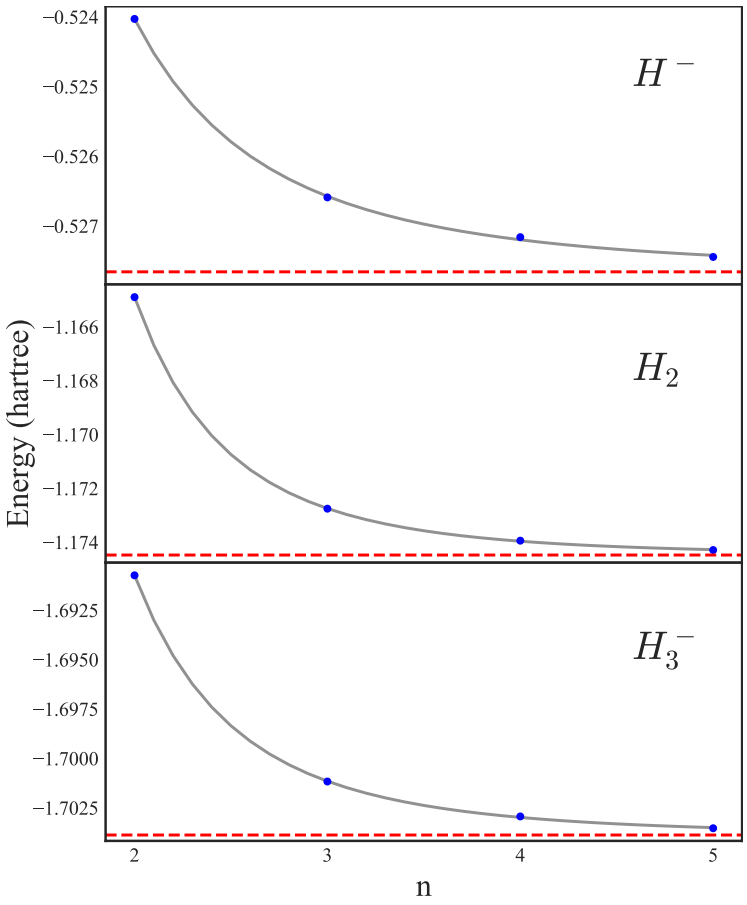


Figure B.1: Fitted curve of Eq. (2.29) as a function of the cardinal number n. The curve follows an asymptotic behaviour whose limit is the $E_{tot,CBS}(H)$ in table (B.1). That limit is depicted in each graph with a red dash line.

APPENDIX C

TABLE: AB INITIO CALCULATIONS

r ($\theta = 0°$)	E($\theta = 0°$)	r	E ($\theta = 10°$)	E ($\theta = 20°$)	E ($\theta = 30°$)	E ($\theta = 45°$)	E ($\theta = 60°$)	E ($\theta = 90°$)
2.246	47.923	2.117	107.569	133.042	171.494	242.015	310.323	376.537
2.325	26.561	2.381	19.850	38.903	67.695	120.548	171.801	221.557
2.431	2.911	2.646	-22.286	-7.796	14.133	54.452	93.657	131.805
2.590	-21.298	2.910	-40.171	-28.963	-11.992	19.278	49.770	79.521
2.960	-45.644	3.175	-45.692	-36.895	-23.556	1.070	25.144	48.692
3.225	-48.871	3.308	-45.946	-38.107	-26.220	-4.247	17.256	38.310
3.490	-47.099	3.440	-45.213	-38.205	-27.573	-7.904	11.366	30.255
3.754	-43.185	3.704	-42.034	-36.377	-27.792	-11.881	3.739	19.080
4.019	-38.568	3.969	-37.842	-33.224	-26.204	-13.182	-0.373	12.229
4.283	-33.946	4.233	-33.470	-29.655	-23.857	-13.082	-2.470	7.989
4.548	-29.639	4.498	-29.310	-26.131	-21.289	-12.291	-3.412	5.351
4.812	-25.773	4.763	-25.533	-22.858	-18.784	-11.196	-3.701	3.709
5.077	-22.371	5.027	-22.184	-19.914	-16.453	-10.003	-3.627	2.683
5.342	-19.412	5.292	-19.260	-17.829	-14.356	-8.833	-3.365	2.046
5.606	-16.859	5.556	-16.727	-15.054	-12.503	-7.741	-3.021	1.654
5.871	-14.665	5.821	-14.548	-13.096	-10.886	-6.750	-2.649	1.418
6.135	-12.785	6.086	-12.676	-11.414	-9.479	-5.869	-2.279	1.283
6.400	-11.171	6.350	-11.071	-9.960	-8.265	-5.088	-1.930	1.209
6.665	-9.781	6.615	-9.685	-8.705	-7.207	-4.402	-1.610	1.170
6.929	-8.576	6.879	-8.487	-7.617	-6.289	-3.795	-1.320	1.151

Table C.1: Calculated values of energy for different distance, r, and angles, θ. Not only the points depicted in Fig. 4.4, but also extra points and angles.The 5Z basis set is used too and it is in units of $\AA$ and meV. The reference zero used is the same as table 4.5. For all the angles except $\theta = 0°$, the same distance has been used in order to simplify the table. The cell with the minimum energy is colored to highlight it.

BIBLIOGRAPHY

[1] T. Oka. "Chemistry, astronomy and physics of H_3^+". In: *Phil. Trans. R. Soc. A.* 370 (2012), pp. 4991–5000. URL: https://doi.org/10.1098/rsta.2012.0243.

[2] Helge Kragh. "The childhood of H_3 and H_3^+". In: *Astronomy & Geophysics* 51 (2010), pp. 6.25–6.27. URL: https://doi.org/10.1111/j.1468-4004.2010.51625.x.

[3] W. Wang et al. "Observations of H_3^- and D_3^- from dielectric barrier discharge plasmas". In: *Chemical Physics Letters* 377 (2003), pp. 512–518. URL: https://www.sciencedirect.com/science/article/pii/S0009261403012107.

[4] H.H. Michels and J.A. Montgomery Jr. "The electronic structure and stability of the H_3^- anion". In: *Chemical Physics Letters* 139.6 (1987). URL: https://www.basissetexchange.org/..

[5] J. Stirck and W. Meyer. "Ab initio potential energy surface for the collisional system H^- + H_2 and properties of its van der Waals complex". In: *Chemical Physics* 176 (1993), pp. 63–95.

[6] Grzegorz Rick Chalasiñski et al. "Ab initio studies of the structure and energetics of the hydride (hydrogen) complex". In: *The Journal of Chemical Physics* 91.24 (1987), pp. 6151–6158. URL: https://doi.org/10.1021/j100308a019.

[7] J. C. Rayez et al. "Theoretical study of the H_3^- cluster". In: *The Journal of Chemical Physics* 75 (1981), pp. 5393–5397. URL: https://doi.org/10.1063/1.441939.

[8] David Stevenson and Joseph Hirschfelder. "The Structure of H_3, H_3^+, and of H_3^-. IV". In: *The Journal of Chemical Physics* 5.933 (1937). URL: https://doi.org/10.1063/1.1749966.

[9] L. Huang et al. "Anionic hydrogen clusters as a source of diffuse interstellar bands (DIBs)". In: *Chemical Physics Letters* 377 (2019), pp. 512–518. URL: https://arxiv.org/abs/1912.11605.

[10] Edward A. Mason and Joseph T. Vanderslice. "Interactions of H^- Ions and H Atoms with Ne, A, and H_2". In: *The Journal of Chemical Physics* 28.1070 (1958). URL: https://doi.org/10.1063/1.1744345.

[11] M. Ayouz et al. "Potential energy and dipole moment surfaces of H_3^- molecule". In: *The Journal of Chemical Physics* 132.194309 (2010). URL: https://doi.org/10.1063/1.3424847.

[12] F. Calvo and E. Yurtsever. "The quantum structure of anionic hydrogen clusters". In: *The Journal of Chemical Physics* 148.102305 (2018). URL: https://doi.org/10.1063/1.4990612.

[13] F. Pirani et al. "Beyond the Lennard-Jones model: a simple and accurate potential function probed by high resolution scattering data useful for molecular dynamics simulations". In: *Physical Chemistry Chemical Physics* 10 (2008), pp. 5489–5503. URL: https://doi.org/10.1039/B808524B.

[14] F. Pirani et al. "Atom–bond pairwise additive representation for intermolecular potential energy surfaces". In: *Chemical Physics Letters* 394 (2004), pp. 37–44. URL: https://doi.org/10.1016/j.cplett.2004.06.100.

[15] Josu Ortiz de Zárate et al. "Snowball formation for Cs^+ solvation in molecular hydrogen and deuterium". In: *Physical Chemistry Chemical Physics* 21 (2019), pp. 15662–15668. URL: https://doi.org/10.1039/C9CP02017A.

[16] Maren Brauner. "The Electronic Structure of H_3^+ , H_3 and H_3^- analyzed with NWChem Quantum Chemistry Package and Hückel's Model". Master Thesis. Universidad de La Laguna, 2021.

[17] Ira Levine. *Quantum Chemistry*. 7th ed. Advanced Chemistry. Pearson, 2004.

[18] C. J. Cramer. *Essentials of Computational Chemistry*. 2nd ed. England: John Wiley & Sons, 2004.

[19] F. Jensen. *Introduction to Computational Chemistry*. 2nd ed. England: John Wiley & Sons, 2007.

[20] T. H. Dunning et al. "Gaussian basis sets for use in correlated molecular calculations. I.The atoms boron through neon and hydrogen". In: *The Journal of Chemical Physics* 90 (1989), pp. 1007–1023. URL: https://doi.org/10.1063/1.456153.

[21] K. L. Schuchardt et al. "Basis Set Exchange: A Community Database for Computational Sciences". In: *Journal of Chemical Information and Modeling* 47 (2007), pp. 1045–1052. URL: https://www.basissetexchange.org/..

[22] Group of Prof. Zipse. "Complete Basis Set (CBS) Extrapolation Schemes". Internal report.

[23] C. David Sherrill. "Distinguishing Basis Set Superposition Error (BSSE) from Basis Set Incompleteness Error (BSIE)". Internal report. School of Chemistry and Biochemistry, Georgia Institute of Technology, 2017.

[24] Anthony J. Stone. *The Theory of Intermolecular Forces*. 2nd ed. United Kingdom: Oxford University Press, 2013.

[25] J. Bretón and J. Hernandez Rojas. *On the role of the molecular polarizability anisotropy in the 3B polarization terms in the trimer cation-*$(H_2)_2$. Private communication. 2022.

[26] E. Aprà et al. "NWChem: Past, present, and future". In: *The Journal of Chemical Physics* 152.184102 (2020). DOI: 10.1063/5.0004997.

[27] *NWChem User Documentation*. 2022. URL: https://nwchemgit.github.io/Home.html..

[28] A. Mohammadi et al. “Coupled Cluster and Quantum Monte-Carlo study of anionic hydrogen clusters $H_n^-(3 < n(odd) < 11)$”. In: *Chemical Physics Letters* 744 (2020). URL: https://www.sciencedirect.com/science/article/pii/S0009261420301317..

[29] Lulu Huang et al. “Ion Induced Dipole Clusters H_n^- $(3 < n-odd < 13)$: Density Functional. Theory Calculations of Structure and Energy”. In: *The Journal of Chemical Physics* 115 (2011), pp. 12445–12450.