Lehrstuhl für Theoretische Chemie
der Technischen Universität München

High level ab initio potential energy surfaces and vibrational spectroscopy

Kiran Sankar Maiti

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr. Notker Rösch
Prüfer der Dissertation:	
1.	Univ.-Prof. Dr. Wolfgang Domcke
2.	Univ.-Prof. Dr. Fritz Elmar Kühn

Die Dissertation wurde am 03.07.2007 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 31.07.2007 angenommen.

Acknowledgment

First of all, my heartiest thanks goes to Dr. Christoph Scheurer for offering me an interesting and important research topic; for his sincere supervision, guidance and valuable suggestions throughout this work; for his time and care in correcting the thesis and also for the personal care and affection, which I received from him.

I am very thankful to Prof. W. Domcke for his care and valuable suggestions in science and different aspect of life.

I am also thankful to Dr. Tobias Steinel for introducing me to a new field of science, his valuable discussions and suggestions.

I am grateful to all of my colleagues for their cooperation and helping hands in all directions. My heartiest thanks to Mr. Andeas Motzke, for his warm friendship and extensive help in the beginning of my life in Munich. Also a special thanks to Sabyashachi for his friendship and extensive help in correcting the thesis.

I would like to thanks Frau Mösch for her help in all semi-academic matters.

Thanks to Somnath, Lalu, Manish, Shyama, Souradip, Paramita for their friendship and help.

Above all, I am thankful to my parents, my brother and sisters and my uncle Late Panchugopal Maiti and all of my well wishers for their love and blessings.

Contents

List of Abbreviations

1D	one dimension
2D	two dimension
AO	atomic orbital
anti-o-DMB	anti-o-deutero-methyl-benzoate
aug-cc-pVXZ	augmented correlation consistent polarized valence X-tuple zeta
B3LYP	Becke's 3-parameter LYP (DFT functional)
CC	coupled cluster
cc-pVXZ	correlation consistent polarized valence X-tuple zeta
CI	configuration interaction
CCSD(T)	CC singles, doubles, and perturbative triples
DF-L-CCSD(T)	density fitting local CCSD(T)
DF-MP2	density fitting MP2
DF-SCS	density fitting spin-component scaled
DFT	density functional theory
HF	Hartree-Fock
IR	infrared
MB	methyl benzoate
MP2-R12	MP2 with linear r_{12} terms
PES	potential energy surface
syn-o-DMB	syn-o-deutero-methyl-benzoate
V-MP2	vibrational MP2
VC-MP2	vibrational corrected MP2
VSCF	vibrational self consistent field

0.1 Summary

An extended scheme for basis set extrapolation of the correlation energy is presented and analyzed for a large number of atoms and molecules. The focus in the development is on the use in generating high level potential energy surfaces (PES) for multidimensional IR spectroscopy. Methyl benzoate (MB) is studied as a model compound for the development of new IR pulse schemes with possible applicability to biomolecules. Anharmonic vibrational modes of MB are calculated on different level (MP2, SCS, CCSD(T) with varying basis sets) ab-initio PESs using the vibrational self-consistent field (VSCF) method and its correlated extensions. Dual level schemes, combining different quantum chemical methods for diagonal and coupling potentials, are systematically studied and applied successfully to reduce the computational cost.

0.2 Zusammenfassung

Ein erweitertes Schema zur Extrapolation der Korrelationsenergie zum Basissatzlimit wird vorgestellt und für eine umfangreiche Anzahl von Atomen und Molekülen analysiert. Das Ziel der Entwicklung liegt in der Anwendung zur Berechnung von exakten Potentialenergieflächen (PES) für die Simulation multidimensionaler IR-Spektroskopie. Methylbenzoat (MB) wurde als Modellsystem zur Entwicklung neuer IR Pulssequenzen, mit dem Hintergrund der Anwendbarkeit auf Biomoleküle, untersucht. Anharmonische Schwingungsmoden für MB wurden auf ab-initio PES unterschiedlicher Güte (MP2, SCS, CCSD(T) mit variierenden Basissätzen) mittels der "vibrational self-consistent field" (VSCF) Methode und Ihrer korrelierten Erweiterungen berechnet. Gemischte Ansätze, welche verschiedene quantenchemische Methoden für die Diagonal- und Kopplungspotentiale verknüpfen, wurden systematisch analysiert und erfolgreich zur Reduzierung des Rechenaufwandes verwendet.

Chapter 1

Introduction

Two of the central questions in physical chemistry and biophysics are how structures of complex molecular systems evolve, and how such dynamics are related to the important chemical and biological processes.[1] Infrared (IR) vibrational spectroscopy is among the foremost experimental tools in the exploration of molecular properties.[2] Linear IR spectroscopy has been used for a long time, however, many important properties of molecular systems cannot be determined by linear spectroscopy. On the other hand two-dimensional (2D) nuclear magnetic resonance (NMR)[3–6] can resolve the structure of complex molecules with atomic resolution, but its millisecond time scale is too slow to follow ultrafast dynamics directly,[7] which is crucial to understand many chemical and the biological processes.

The recent success in the experimental realization of coherent 2D IR spectroscopy provides a powerful new tool to study the structure and dynamics of biological systems with temporal resolution down to the sub-picosecond regime.[8–12] This technique has the ability to disentangle congested vibrational spectra of biomolecules to some extent similar to 2D NMR, but on an ultrafast timescale. The most significant feature of non-linear 2D IR spectroscopy is the possibility to directly measure coupling constants and orientations of the transition dipole moments of the IR active modes.[7] The structural and dynamical information is typically available in terms of diagonal and cross-peak shapes, locations and intensities and their respective temporal evolution.[7] The interpretation of this data in terms of a dynamical model of the biomolecule under investigation requires extensive theoretical modeling.

Until very recently, theoretical calculations of vibrational spectra were most often restricted to the harmonic approximation.[2] Such studies have been very useful, but they are also of limited significance since many biological molecules are very "floppy" and subject to strong anharmonic effects. Also, much of the interest in the molecular dynamics

is away from the equilibrium configurations, where the harmonic approximation is very poor. The main computational difficulties in the calculation of anharmonic spectra is that the different vibrational modes are not mutually separable and one has to face a large number of coupled degrees of freedom with the number of couplings growing very quickly with system size.

Several approaches are known to compute vibrational spectra for anharmonic systems. Among others, the discrete variable representation (DVR),[13–16] diffusion quantum Monte Carlo (DQMC),[17–20] and vibrational self-consistent field (VSCF)[21–23] methods proved their applicability to study anharmonic effects in systems with different sizes. The VSCF method is most successful among them to simulate the spectroscopic data for large anharmonic systems. The success of the VSCF method (and also all other methods) depends upon the accuracy of the calculated potential energy surface (PES).[24,25]

In the VSCF approach, the PES enters the calculation in the form of multidimensional integrals for the effective one-dimensional potentials which is the most difficult part of the VSCF calculations. To simplify these integrals it is possible to express the PES in terms of a hierarchical many-body expansion.[26] In typical molecular problems quadruple and higher interaction potentials have a negligible influence on the vibrational spectrum and are therefore commonly neglected in the VSCF calculations. In a recent detailed analysis it has been shown that even only few triple couplings have a significant influence.[27] For large typically molecular systems diagonal and pair potentials are usually sufficient if the PESs are calculated accurately in an appropriately chosen coordinate system. Additionally diagonal potentials may be calculated employing a higher level computational method than the pair potentials to increase the accuracy of the VSCF calculations significantly. Such a dual level computation not only improves the accuracy but also reduces the computational expenses significantly due to the large number of possible pair potential.

The PES calculated at the semi-empirical methods[28,29] seems promising but not sufficient for the high resolution spectroscopy. The quantum chemical ab initio method recently reached to a high level of accuracy for small molecules[30,31] and is expected to improve the accuracy of the PES calculation. The highly accurate ab initio calculation requires the basis functions to be as large as possible.[32] As the basis function increases, the computational expense grows much faster than the rate at which the accuracy is improved,[32] which makes the method unaffordable for the highly accurate PES calculation. On the other hand, spectroscopic accuracy can not be achieved by the smaller basis sets. In such a situation extrapolation of the energy to the basis set limit may speed up the

PES calculation if there is a general extrapolation scheme. It is well known that Hartree-Fock (HF) energy converges very quickly and essentially reaches the basis set limit with the small basis set. Therefore spectroscopic accuracy only depends upon the accurate calculation of the correlation energy, which is known to converge very slowly. Two point extrapolation of the correlation energy, proposed by Halkier et al.[31] reached to the spectroscopic accuracy but this method suggests to extrapolate from two larger basis sets, which is not feasible for large molecules. Inclusion of fifth order term from Schwartz's formula[33] provides an efficient route to extrapolate the correlation energy from small basis sets and reduces the computational expenses by two to three order of magnitude to reach the basis set limit correlation energy.[34] A further improvement on the extrapolation scheme is required for more accurate PES calculation with a reasonable effort which will lead to the VSCF, a generous method for the simulation of spectroscopic data of large anharmonic systems like proteins and peptides.

Spectroscopically the amide-I, amide-II modes and the coupling between the amide-I and β-hydrogen located in the side chain of proteins and peptides are most important because they can provide information about secondary structural motifs of proteins and peptides. The so-called amide-I band, which consists mostly of the stretch vibrational motion of the C=O group is a strong IR absorber, which is spectrally isolated from the other vibrational modes even in large proteins. The C–H or C–D stretch vibrations as backbone modes are particularly important structural probes in biological systems because they are very localized. Especially the C–D stretch frequency is an excellent structural probe since it is usually spectrally isolated even in the spectrum of large proteins. It is necessary to compute and identify these important vibrational bands efficiently.

Proteins have a too complicated structure to be studied in details. On the other hand Methyl benzoate itself is a small organic molecule and provides some of the structural similarities with proteins and peptides. Computationally it is more favorable due to its structural simplicity and can be studied in more detail. Structurally it is a stable planar molecule except for the two hydrogen atoms at the methyl group which are symmetrically out-of-plane with respect to the rest of the molecule. The C=O double bond in the carboxylic ester group behaves as a local oscillator similar to the amide-I band in proteins and provides a convenient mode to study the amide-I band structure of proteins. A potentially interesting coupling in proteins is the coupling between the amide-I and the β-hydrogen located in the sidechain. Methyl benzoate also provides a similar structure where the ortho hydrogen in the phenyl ring provides the counter-part of β-hydrogen in

protein sidechains. Due to the low potential barrier, ester group rotate about the C–C single bond in a ultrafast process and provide an efficient route to develop the new pulse scheme to study ultrafast phenomena in biomolecules. For all these above mentioned reasons Methyl benzoate is considered to be an important model system to develop the pulse sequence technique for the vibrational spectroscopy of proteins and peptides.[35]

In the following chapters we will systematically develop and improve computational methods which in the long run will allow us to study large biological systems. Chapter 2 shortly reviews the quantum chemical methods used, especially the perturbative electron correlation method (MP2) and the basis sets as these are the basis of the chapter 3.

Chapter 3 mainly discusses the convergence behavior of the correlation energy. With a short review of Schwartz's work on the He atom, it is shown that the convergence behavior of correlation energy is rather slow[33] (l^{-4}, where l is the maximum angular momentum quantum number in one-particle basis). Extrapolation of the energy towards the basis set limit is frequently used to overcome the slow convergence since it does not increase any further cost. Several extrapolation methods are discussed with their advantages and disadvantages. An extensive investigation of the proposed extrapolation scheme with a large number of atoms and molecules is presented in this chapter. It is discussed also in this chapter how efficiently the PES can be calculated by extrapolation method.

Chapter 4 mainly discusses vibrational spectroscopy of Methyl benzoate. A short review of the VSCF method and the PES expansions are presented which are the basis for the further investigations and discussions. Diagonal and off-diagonal anharmonicity calculation techniques are given here. It is also shown how the 1D and 2D PES were calculated to make the VSCF method successful. Finally it presents the detail calculations and analysis of vibrational anharmonic frequencies of Methyl benzoate and and its two isotopomers.

Chapter 2

Quantum Chemistry

2.1 Quantum Chemical Methods

2.1.1 The Electronic Problem

Molecules are nothing but many-body systems consisting of nuclei and electrons. The central problem in quantum chemistry is to solve the Schrödinger equation for this many-body system and find the wave functions. Here we will consider the solution of the non-relativistic time-independent Schrödinger equation, i.e.,

$$\mathscr{H}|\Phi\rangle = \mathscr{E}|\Phi\rangle, \tag{2.1}$$

where $\mathscr{H}$ is the Hamiltonian operator for the system of nuclei and electrons described by position vectors $\mathbf{R}_A$ and $\mathbf{r}_i$, respectively. A molecular coordinate system is shown in Fig. 2.1, where A and B denote nuclei and i and j denote electrons. Assuming the nuclei and electrons to be point masses and neglecting spin-orbit and other relativistic interactions, the complete Hamiltonian for this system, in atomic units, is given by

$$\mathscr{H} = -\frac{1}{2}\sum_{i=1}^{N} \bigtriangledown_i^2 - \sum_{A=1}^{M} \frac{1}{2m_A} \bigtriangledown_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}}, \tag{2.2}$$

where the first two terms of the right-hand-side represent the kinetic energy operator of the N-electrons and M-nuclei, respectively. While the third term describes the interaction between nuclei and electrons, the last two terms represent the inter-electronic and inter-nuclear interactions. m_A is the atomic mass of nucleus A. The Laplacian operator $\bigtriangledown^2$ involves differentiation with respect to the coordinates of the particle (i.e., electron or nucleus) and is defined as,

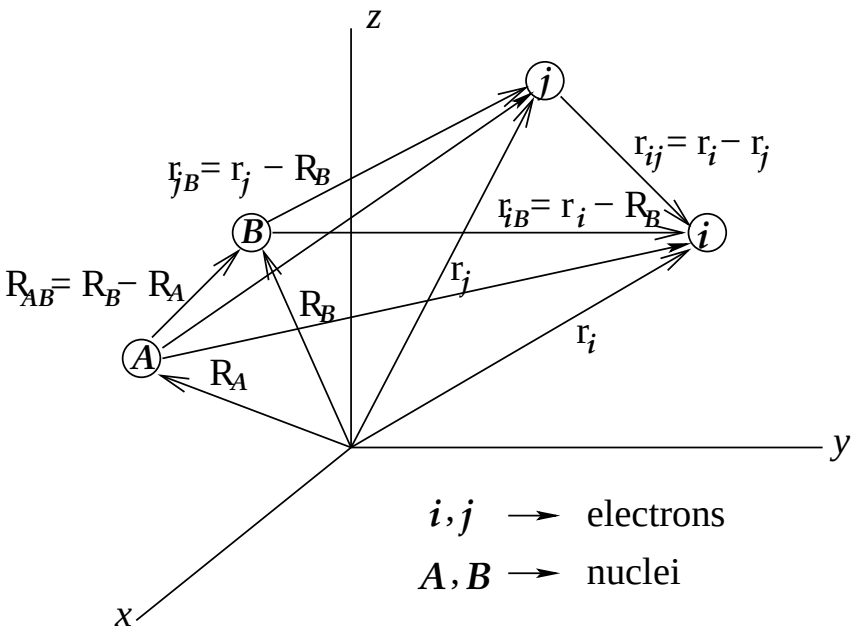

Figure 2.1: Molecular coordinate system.

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \tag{2.3}$$

in Cartesian coordinates. The restriction $j > i$ and $B > A$ avoids self interaction terms like e^2/r_{ii} as well as the overcounting of the interactions.

Born-Oppenheimer approximation

For any molecule larger than Hydrogen molecular ion (H_2^+), the molecular Hamiltonian (Eq. (2.2)) cannot be solved exactly. Born and Oppenheimer proposed a simplified approach inspired by an idea by Newton. In 1687 Newton showed that due its large mass the earth does not move towards the apple but rather the apple falls on the earth's surface from a tree. Born and his student Oppenheimer introduced a similar approach in quantum mechanics in 1927. Since nuclei are much heavier than electrons (a single proton is about 1836 times heavier than an electron), to a good approximation one can consider nuclei as fixed in space and electrons are moving under a constant electrostatic force field of the nuclei.[36] In this approximation the kinetic energy (KE) of the nuclei, i.e., the second term in the r.h.s. of Eq. (2.2) vanishes. Since nuclei are fixed, the repulsions between nuclei are constant. Any constant added to an operator only adds to the operator eigenvalues

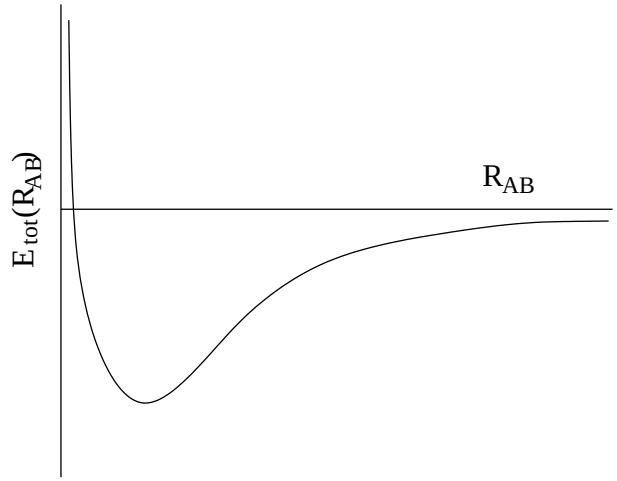

Figure 2.2: Path of nuclear motion. Here R_{AB} is the inter nuclear distance and $E_{tot}(R_{AB})$ is the molecular energy at different nuclear distances.

and has no effect on the operator eigenfunctions. Neglecting the 2nd and the 5th term in Eq. (2.2), one obtains the Hamiltonian for the purely electronic system, i.e.,

$$\mathscr{H}_{\rm el} = -\frac{1}{2}\sum_{i=1}^{N} \bigtriangledown_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}}. \tag{2.4}$$

Now the Schrödinger equation for electronic motion becomes,

$$(\mathscr{H}_{\rm el} + \mathscr{V}_{NN})\Phi_{\rm el} = \mathscr{E}_{\rm tot}\Phi_{\rm el}. \tag{2.5}$$

The energy $\mathscr{E}_{\rm tot}$ in Eq. (2.5) is the electronic energy including the internuclear repulsion ($\mathscr{V}_{NN}$). The internuclear distances $\mathbf{R}_{\rm AB}$ in Eq. (2.2) are parameters of the electronic problem. The electronic wave functions thus depend explicitly on electronic coordinates but depends parametrically on the nuclear coordinates,

$$\Phi_{\rm el} = \Phi_{\rm el}(\{r_i\};\{R_A\}). \tag{2.6}$$

Similarly, electronic energy,

$$\mathscr{E}_{\rm el} = \mathscr{E}_{\rm el}(\{R_A\}). \tag{2.7}$$

The total energy $\mathscr{E}_{\rm tot}(\{R_A\})$ provides a potential for the nuclear motion. This function represents a potential energy surface (PES) as shown schematically in Fig. 2.2. Thus the nuclei in the Born-Oppenheimer approximation move in a PES obtained by solving the electronic problem.

2.1.2 Hartree-Fock Self-Consistent Field Method

The main goal of quantum chemistry is to find the exact electronic wave function. For many-body systems, the problem becomes so complicated that the calculation of the exact wave function becomes impossible. It is only possible to find a good approximation to the exact wave function for the system which can explain most of the properties of the system accurately. The total Hamiltonian for N non interacting particles can be written as,

$$\mathscr{H} = \sum_{i=1}^{N} h(i), \tag{2.8}$$

where $h(i)$ is the operator describing the kinetic energy and potential energy of electron i, and defined as,

$$h(i) = \frac{1}{2} \bigtriangledown_i^2 - \sum_{A=1}^{N} \frac{Z_A}{r_{iA}}. \tag{2.9}$$

Considering the spin of single electron wave function (orbital) $\chi_i(q_i)$, where,

$$\chi(q_i) = \Phi(x_i, y_i, z_i)\alpha_i(\omega), \tag{2.10}$$

for the i-th electron, the time-independent Schrödinger equation separated into N one electron equations

$$h(i)\chi_j(q_i) = \varepsilon_j \chi_j(q_i). \tag{2.11}$$

Now, instead of considering non-interacting electrons, the electronic repulsion can be included in an average way and Eq. (2.9) becomes,[36]

$$h(i) = \frac{1}{2} \bigtriangledown_i^2 - \sum_{A=1}^{N} \frac{Z_A}{r_{iA}} + \upsilon_i^{\mathrm{el}}, \tag{2.12}$$

where υ_i^{el} is the average repulsive force between the i-th electron and all other electrons.

Since $\mathscr{H}$ is a sum of one-electron Hamiltonians, the corresponding wave function of the N-electron system will be the simple product of N individual wave functions. The total wave function can be written as,

$$\Phi^{\mathrm{H}}(q_1, q_2, \cdots, q_N) = \chi_i(q_1)\chi_j(q_2)\cdots\chi_k(q_N), \tag{2.13}$$

and the Schrödinger equation becomes,

$$\mathscr{H}\Phi^{\mathrm{H}} = \mathscr{E}\Phi^{\mathrm{H}}, \tag{2.14}$$

where the eigenvalue $\mathscr{E}$ is sum of the spin orbital energies of each of the spin orbitals appearing in Φ^{HF},

$$\mathscr{E} = \sum_{i=0}^{k} \varepsilon_i. \tag{2.15}$$

The best possible approximate wave function for a many electron system is obtained by making a guess for the trial functions χ_i, and solving the Schrödinger Eq. (2.11) repeatedly. This procedure is continued until the resulting wave function no longer changes. This procedure is called the Hartree self-consistence process.[37] The many electron wave function obtained by this method is still an approximate wave function, because the many-electron Schrödinger equation is not separable, so the true wave function cannot be written as the direct product of N single-electron functions.

Since the Hartree product does not take into account the indistinguishability of the electrons, Fock instead of taking the direct product of N one-electron functions, constructed an appropriate linear combination of these wave functions which will satisfy the Pauli principle. The antisymmetric wave functions can be written in determinant form, known as Slater determinant,[36]

$$\Psi(q_1, q_2, \cdots, q_N) = (N!)^{-1/2} \begin{pmatrix} \chi_i(q_1) & \chi_j(q_1) & \cdots & \chi_k(q_1) \\ \chi_i(q_2) & \chi_j(q_2) & \cdots & \chi_k(q_2) \\ \vdots & \vdots & & \vdots \\ \chi_i(q_N) & \chi_j(q_N) & \cdots & \chi_k(q_N) \end{pmatrix}, \tag{2.16}$$

where $(N!)^{-1/2}$ is the normalization factor. This Slater determinant has N electrons occupying spin orbitals $(\chi_i, \chi_j, \cdots, \chi_k)$ without specifying which electron is in which orbitals and satisfies the antisymmetry principle. In a short-hand notation this determinant is written as,

$$\Psi(q_1, q_2, \cdots, q_N) = |\chi_i \chi_j \cdots \chi_N\rangle. \tag{2.17}$$

A many electron wave function described by such a Slater determinant incorporates exchange correlation, which means that the motion of two electrons with parallel spins is correlated while being non-correlated for electrons of opposite spin.[36]

2.1.3 Concept of basis

The main objective of any of the above mentioned molecular quantum-mechanical method is to solve the time-independent Schrödinger equation, which is an integro-differential

equation. One can solve this equation analytically, but as the number of electrons increases in the system, quickly the problem goes beyond analytical solution. There is a possibility to solve the equation numerically, and a common practice is to solve atomic problem, but still there are no practical procedures presently available to solve molecular problem. Roothan in 1951 showed an elegant way to solve such a problem by using so called basis functions, which could convert the differential equation to a set of algebraic equations and efficiently solved by standard matrix techniques.[38]

Now consider a set of K known basis function $\{\phi_\mu(r)|\, \mu = 1, 2, \cdots, K\}$ and expand the unknown molecular orbitals in the linear expansion

$$\Psi_i = \sum_{\mu=1}^{K} C_{\mu i}\phi_\mu, \tag{2.18}$$

which is nothing but matrix representation of a one particle (orbital) wave function. If the set ϕ_μ is complete, the molecular orbital can be represented exactly. But in practice complete basis set is not manageable within the present technology and methods available, and therefore one has to work with a finite set of K basis functions. Then it is important to choose a basis that will provide, as far as possible, a reasonably accurate expression for the exact molecular orbitals Ψ_i.

Choice of basis functions

In the most general sense, a basis set is nothing but a mathematical representation of electron path in atomic or molecular environment. Slater developed a function to describe the atomic orbitals, the so-called Slater type orbital (STO),[39] as,

$$\phi_{1s}^{SF}(\zeta, r) = \left(\frac{\zeta^3}{\pi}\right)^{1/2} e^{-\zeta r}, \tag{2.19}$$

where ζ is the Slater orbital exponent and r is the instantaneous distance between nucleus and electron.

Another popular type of orbital is Gaussian type orbital (GTO),[36]

$$\phi_{1s}^{GTO}(\alpha, r) = \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha r^2}, \tag{2.20}$$

where α is the Gaussian orbital exponent. Although there are several different kind of basis functions, these two are more popular due to their excellent ability to express atomic and molecular orbitals and also in computational sense, they are easy to compute. The

main advantage of Gaussian type function over Slater type function is that, product of two Gaussian functions is another Gaussian function, which makes it easier to perform two electrons calculations and save enormous time of calculations. Where as calculation with Slater function such a two electron calculations are most expensive and time consuming. On the other hand, the Gaussian function could not able to describe Coulomb cusp and also decays rapidly at large distances, as compared to actual behavior of molecular orbitals. Slater type orbital, however, is a smoothly varying function at $r \rightarrow 0$ and also slowly decays at large distances, which is more physical.

Contracted Gaussian basis set

In 1950, S. F. Boys suggested to construct an approximate Slater type function using more than one Gaussian primitive.[40] This approximate Slater function defines atomic orbitals more efficiently still using Gaussian functions which simplifies the computation. Such an approximate function is known as contracted Gaussian function,

$$\phi_{\mu}^{\mathrm{CGF}}(r) = \sum_{p=1}^{N} d_{p\mu}\phi_{p}^{\mathrm{GF}}(\alpha_{p\mu}, r), \tag{2.21}$$

where N is the length of the contraction and $d_{p\mu}$ is the contraction coefficient. The p^{th} normalized primitive Gaussian ϕ_{p}^{GF} in the basis function $\phi_{\mu}^{\mathrm{CGF}}$ has a functional dependence on the Gaussian orbital exponent $\alpha_{p\mu}$. By a proper choice of the contraction length, the contraction coefficients, and the contraction exponents, the contracted Gaussian function can be made to assume Slater functional form consistent with primitive functions used. This procedure is known as STO-NG method and basis sets constructed this way are referred to as STO-NG. A very common example being the use of three Gaussian primitives: STO-3G.[41,42] Figure 2.3 shows a comparison of different contracted Gaussian functions and Slater functions. These basis sets are known as minimal basis sets, since they used one STO (or GTO) per spin-orbital. It considers one functional ($1s$) for H and He, five functional ($1s, 2s, 2p_x, 2p_y, 2p_z$) for Li $\rightarrow$ Ne, and nine functional for Na $\rightarrow$ Ar, and so on.

Extended basis sets

A minimal basis set approximates all the orbitals to be of same shape. However, in reality this is a poor approximation, since the s orbital is circular, the p is in dumbbell shape, etc. The next approach to improve the minimal basis set is to use two contracted Gaussians for each of the minimal basis functions, which differ in their orbital exponent ζ, the so-called

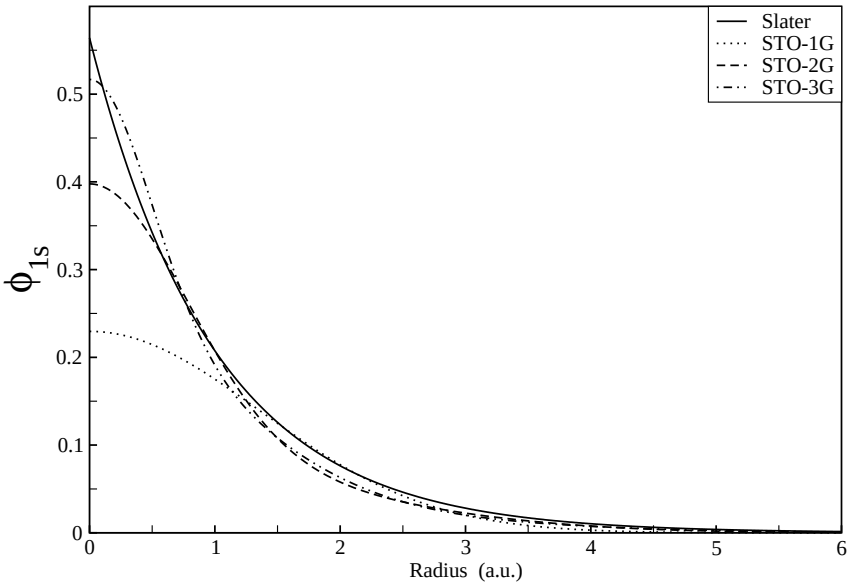

Figure 2.3: Comparison of the quality of the least-squares fit of a 1s Slater function ($\zeta = 1.0$) obtained at the STO-1G, STO-2G, and STO-3G levels. Contraction exponents and contraction coefficients are used from Ref. 36.

double-zeta (DZ) basis set.[36] The best orbital exponent of the two STOs are commonly slightly above and slightly below the optimal exponent of the minimal basis function.

$$\Phi_{2s}(r) = \Phi_{2s}^{STO}(r, \zeta_1) + d\Phi_{2s}^{STO}(r, \zeta_2) \tag{2.22}$$

The two STOs are added in some proportion. The constant d determines the contribution of each STO to the final orbital, i.e., whether the effective orbital has to be expanded or contracted. The higher order basis sets, e.g., triple-zeta, quadruple-zeta, etc., in this series are formed in the similar way.

Split valence basis

Most of the chemical properties of atoms mostly depend on the valence shell rather than the inner shells which have only some, mostly constant, contribution to the total energy. Therefore, instead of using more STOs for inner shell if one uses only a single STO, it saves enormous effort while performing calculation using the double-zeta, triple-zeta, etc., basis set which does not affect much to calculate most of the chemically interested properties like dipole moment, ionization potential, charge density, dissociation energy etc. Some popular basis sets in this series are 3-21G,[43–48] 6-21G,[43,47] 4-31G,[49–52] 6-31G etc.

For example in 4-31G basis core orbitals are constructed with four Gaussian primitives, where as three Gaussian primitives used for first STO and one for the second STO. For Li $\rightarrow$ F, the contractions are,

$$\phi_{1s}(r) = \sum_{i=1}^{4} d_{i,1s} g_{1s}(\alpha_{i,1s}, r), \tag{2.23}$$

$$\phi'_{2s}(r) = \sum_{i=1}^{3} d'_{i,1s} g_{1s}(\alpha'_{i,2sp}, r), \tag{2.24}$$

$$\phi''_{2s}(r) = g_{1s}(\alpha''_{2sp}, r), \tag{2.25}$$

$$\phi'_{2p}(r) = \sum_{i=1}^{3} d'_{i,2p} g_{2p}(\alpha'_{i,2sp}, r), \tag{2.26}$$

$$\phi''_{2p}(r) = g_{2p}(\alpha''_{2sp}, r). \tag{2.27}$$

Polarized basis sets

More than one STO basis sets may describe an isolated atomic orbital quite efficiently, but in molecular environment, orbitals will not be as they are in isolated atom. For example if one look at on an isolated hydrogen atom, it is just the spherical $1s$ orbital. Now if the hydrogen atom is placed in a uniform electric field, the electron cloud is attracted to the direction of the electric field and the orbital shape will be no longer the spherical $1s$ but a mixture of the original $1s$ orbital and a p-type orbital. Similarly when atoms are brought closer, their charge distribution shows a polarization effect due to the electric field of other atoms, which distorts the shape of the atomic orbitals. In this case the s orbitals tend to have a little of the p flavor and the p orbitals tend to have a little of the d flavor and so on. This polarization effect is described by adding higher orbitals angular-momentum in the basis sets, the so-called polarized basis sets are represented by adding an asterisk (*) after the basis set, e.g., 4-31G $\rightarrow$ 4-31G*.

Diffuse basis sets

So far all the basis sets developed mainly concerned about the inner shell electrons which are responsible for main energy of atom. This is the main area under the wave function. The area to the left of the dotted line in Fig. 2.4 indicate this region. The area to the right of the dotted line is not really a factor in calculation. But the outer shell valence electrons are mainly responsible for the molecule formation, which has less contribution in energy.

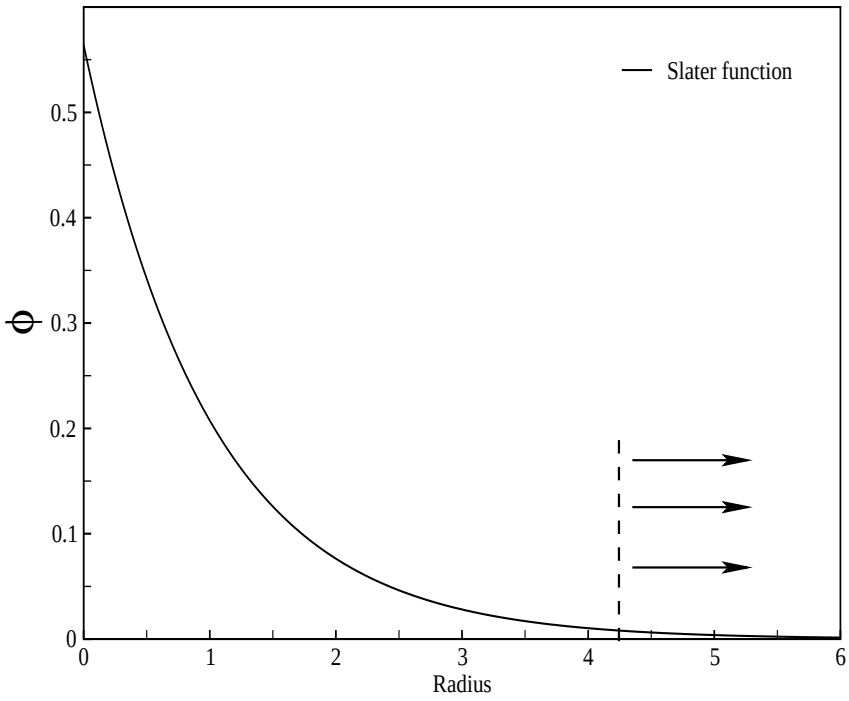

Figure 2.4: Atomic wave function. Diffuse basis set take care the wave function right to the dotted line.

However, when an atom is negatively charged or in an exited state, the loosely bound electrons which are responsible for the energy in the tail region of the wave function, become very important. This contribution is taken into account by using the so-called diffused basis set. Normally diffused functions are with very small orbital exponent. Diffused basis set is denoted by '+' sign, like 3-21+G, 3-21++G etc. Single '+' sign indicates that the basis sets are formed adding four highly diffused functions (s, p_x, p_y, p_z) on each first and second row elements. Double plus sign indicates that, a highly diffused s function is added for the hydrogen atom.[53–55]

Dunning basis sets

So far all basis sets developed are optimized at the HF level of computation. Dunning first pointed out that basis sets optimized at the HF level might not be ideal for correlated computations. Dunning and co-workers optimized the basis sets at singly and doubly excited configuration interaction level and named as "correlation-consistent" (cc) basis set. Over the last decade Dunning and co-workers developed a set of hierarchical sequences

of basis sets like correlation consistence polarized valence X-tuple-zeta basis set ("cc-pVXZ"), where $X = D, T, Q, 5, \cdots$. Adding diffused functions with cc-pVXZ basis set they also developed another hierarchical sequences of basis set, called as augmented cc-pVXZ (aug-cc-pVXZ).[53,54,56–59]

2.1.4 Electron Correlation

Molecular energies calculated by the Hartree-Fock (HF) method are typically in error by 0.5% to 1%. For describing chemical processes, this accuracy is far from sufficient. The main drawback of the HF approximation is that the electronic interactions are taken into account only in an average way. This is an incorrect representation, since electron-electron interaction depends on their instantaneous position. Consider for example the two electron hydrogen molecule. In HF representation when counting the interaction of one electron, the other electron's effect is symmetrically distributed over both the atoms. But in reality when one electron is situated near a hydrogen atom, it electrostatically repulse the other electron. There is only a small probability to find the second electron near the region of the first electron, which means the motion of electrons are correlated. Therefore the instantaneous electron correlation should be introduced into the wave function.

The one way to overcome the problem is the so-called configuration interaction (CI) method. The basic idea behind CI is to diagonalize the N-electron Hamiltonian in a basis of N-electron functions (Slater determinants). The exact wave function is represented as a linear combination of N-electron trial functions and the linear variational method is used. If the basis is complete, one obtains the exact energies for ground state as well as excited state of the system. But in practice, one can handle only a finite set of N-electron trial functions, consequently CI provides only upper bound to the exact energies.

Limitation of CI

Consider a system with N electrons. Now if the basis sets have $2K$ spin orbitals, then N will be occupied orbitals and $2K - N$ orbitals will be unoccupied. Now we can choose n spin orbitals from those occupied orbitals by $\binom{N}{n}$ ways. Similarly we can choose n spin orbital from $2K - N$ virtual orbitals by $\binom{2K-N}{n}$ ways. Thus total number of n-tuple excited determinants is $\binom{N}{n}\binom{2K-N}{n}$. Due to such a large determinantal form, CI calculation quickly becomes formidably large although it gives best results for a given one-electron basis set. Figure 2.5 shows how the exact nonrelativistic Born-Oppenheimer wave function is approached as the size of the one-electron and N-electron basis sets

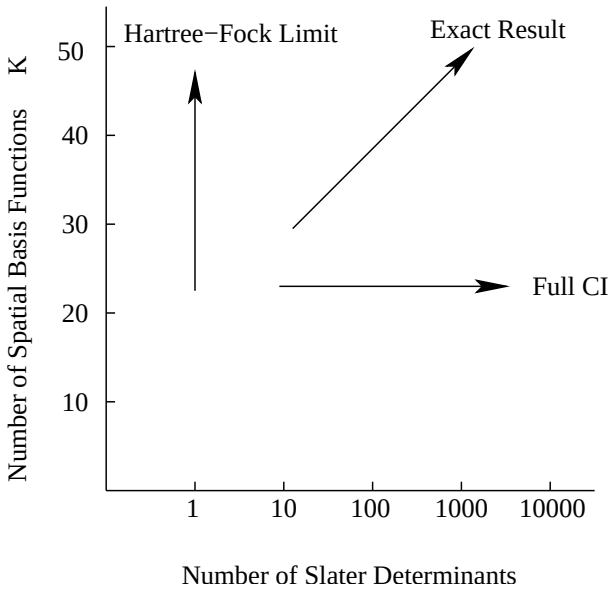

Figure 2.5: Dependence of correlations on size of one-electron and N-electron basis sets.

increase. One can obtain the best result by increasing both one-electron and N-electron basis sets sufficiently, indicated by the diagonal arrow. For a given number of one-electron basis function, N-electron basis function itself increases so rapidly that one can not reach the exact calculation. For a practical purpose to obtain a computationally viable scheme one must has to truncate the full CI matrix.

2.1.5 Møller-Plesset Perturbation Theory

The CI is a systematic procedure for going beyond the HF approximation. But due to the size-consistency problem, truncated CI is not a good choice for calculation of correlation energy. Physicist were successfully using perturbation treatment for many-body problems like crystal structure, nucleons in a nucleus, etc. and known as many body perturbation theory (MBPT). Brueckner showed that MBPT is size-consistence at each order.[60] Motivated by Brueckner's results, in 1934, Møller and Plesset proposed a perturbation treatment of atoms and molecules in which the unperturbed wave function is the HF function. This form of MBPT is called Møller-Plesset (MP) perturbation theory.[61] In this approach the total Hamiltonian is divided into two parts, the unperturbed or the zeroth order Hamiltonian (HF Hamiltonian), denoted by $\mathscr{H}$, which has known eigenfunctions (HF wave functions) and known eigenvalues, and the perturbation part describing

correlation effects.

When the perturbation is small, using power series expansion one can write the total Hamiltonian of the system as,

$$\mathscr{H}_i = \mathscr{H}_i^{(0)} + \lambda\mathscr{H}_i^{(1)} + \lambda^2\mathscr{H}_i^{(2)} + \cdots . \tag{2.28}$$

Similarly exact eigenfunctions and eigenvalues expanded as,

$$|\phi_i\rangle = |\Psi_i^{(0)}\rangle + \lambda|\Psi_i^{(1)}\rangle + \lambda^2|\Psi_i^{(2)}\rangle + \cdots , \tag{2.29}$$

$$\mathscr{E}_i = \mathscr{E}_i^{(0)} + \lambda\mathscr{E}_i^{(1)} + \lambda^2\mathscr{E}_i^{(2)} + \cdots . \tag{2.30}$$

Now substituting Eqs. (2.28), (2.29) and (2.30) in the time independent Schrödinger equation Eq. (2.1) and equating the coefficient of λ^n, one can get the different order energy correction $\mathscr{E}_i^{(n)}$.

$$\mathscr{E}_i^{(1)} = \langle\Psi_i^{(0)}|\mathscr{H}_i^{(1)}|\Psi_i^{(0)}\rangle, \tag{2.31}$$

$$\mathscr{E}_i^{(2)} = \sum_{i\neq j} \frac{|\langle\psi_i^{(0)}|\mathscr{H}_i^{(2)}|\psi_i^{(2)}\rangle|^2}{\mathscr{E}_i^{(0)} - \mathscr{E}_j^{(0)}}. \tag{2.32}$$

Zeroth and first order together yield the HF energy. When second order energy correction is added to the HF energy, the MP2 energy is obtained.[62–67] Higher order corrections give MP3, MP4,[68] MP5,[69] etc. energy.

Chapter 3

Extended scheme of the basis set extrapolation

3.1 Introduction

Ab initio quantum chemical calculations of molecular properties have reached a remarkably high level of accuracy for small molecules.[30,31] Highly accurate electronic structure calculations of molecules require the expansion of the electronic wave function to be as nearly complete as possible both in one- and n- electron space. But with the increase of the number of basis functions towards a near complete basis, computational cost grows much faster than the rate at which the accuracy is improved. In fact, it grows approximately as the fourth power in the reduction of the basis-set truncation error.[32] That means, to reduce the error of the computed energy by a factor of ten, the computational cost increases by a factor of 10^4. This tendency of a large increase in computational cost is the main obstacle to perform accurate calculations of molecular properties in most cases. The problem can be resolved either by using a wavefunction which explicitly depends on inter-electronic coordinates or by using an extrapolation scheme which should accelerate the systematic convergence to the basis set limit.[32] There have been impressive developments in the field of wave function based methods which explicitly depend on inter-electronic coordinates[70–83] during the last two decades. But due to the presence of inter-electronic terms r_{12} in the wavefunctions, these methods are also not suitable for larger molecular systems. On the other hand the extrapolation approach does not incur any further cost to estimate the basis-set limit. It may therefore be an efficient method for highly accurate calculations if a general extrapolation scheme can be devised to accurately estimate the basis set limit energy.

3.2 Theoretical background

It is well known that the Schrödinger equation for the many body problem can only be solved approximately using the variational principle which is based on an imperfect fitting of some trial function. With a sufficiently (infinitely) flexible trial function it is assumed that convergence to the exact result is achieved. Then, the question is, how rapidly does the answer converge?

Let us consider the Ritz variational principle for the energy of a bound state,[84]

$$E = \frac{(\psi H \psi)}{\psi\psi}, \tag{3.1}$$

and suppose the trial function ψ is represented by a finite linear combination of appropriate functions u_n,

$$\psi = \sum_{n=1}^{N} C_n u_n \tag{3.2}$$

Depending on the total number N of variational coefficients C_n, a sequence of approximations (E_N) according to Eq. (3.1) will be obtained which will converge to the correct energy value E. The rate of convergence depends upon how fast $|E_N - E_{N-1}|$ goes to zero as N goes to infinity.

Let us now focus on a particular example: the two electron system in the helium atom. Pekeris carried out the original Hylleraas[85] program for the calculation of the helium ground state and reached an accuracy several orders of magnitude beyond any previous attempt.[86] Schwartz analyzed the same helium atom problem and studied explicitly the convergence behavior of the second order energy in a systematic way. In the following, we will present a short summary of this analysis.

3.2.1 Correlation energy in the two electron atom

Consider the ground state of a two-electron system having nuclear charge $Z = 2$. Using perturbation theory, the zeroth order ground state wave function of the system is written as,

$$\psi_0 = \frac{4}{4\pi}\exp[-(r_1 + r_2)], \tag{3.3}$$

which is nothing else but the product of two $1s$ states of the hydrogen atom and r_1 and r_2 are the radial coordinates of the two electrons. The first order correction of the wave function ψ_1 is then given by

$$(E_0 - H_0)\psi_1 = (H_1 - E_1)\psi_0, \tag{3.4}$$

where $E_0 = -1$, $E_1 = 5/8$ and $H_1 = \frac{1}{r_{12}}$ in atomic units (see Ref. 84). The second-order energy E_2, can be calculated after ψ_1 has been obtained and is defined by

$$E_2 = \int \psi_0 (H_1 - E_1) \psi_1 dv \tag{3.5}$$

Yet, an exact solution for E_2 has never been obtained. The problem may be reduced by expanding ψ_1 in a series of Legendre polynomials

$$\psi_1 = \sum_{l=0}^{\infty} \psi_1^{(l)}(r_1, r_2) P_l(cos\theta), \tag{3.6}$$

$$E_2 = \sum_{l=0}^{\infty} E_2(l), \tag{3.7}$$

where θ is the angle between the two position vectors and l is the angular momentum quantum number. If we express Eq. (3.4) in operator form, we obtain

$$\left[\frac{1}{2}\nabla_1^2 + \frac{1}{2}\nabla_2^2 + \frac{1}{r_1} + \frac{1}{r_2} - 1\right] \psi_1^{(l)}(r_1, r_2) = \left[\frac{r_<^l}{r_>^{l+1}} - \frac{5}{8}\delta_{l,0}\right] \psi_0, \tag{3.8}$$

where the subscripts $<$ and $>$ designate the lesser and the greater of the two distances r_1 and r_2, respectively, and ∇ is the radial part of the kinetic energy operator and defined by

$$\nabla^2 \rightarrow \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - \frac{l(l+1)}{r^2}, \tag{3.9}$$

$$E_2(l) = \int \psi_0 \frac{r_<^l}{r_>^{l+1}} \psi_1^{(l)} \frac{dv_1 dv_2}{(2l+1)} \qquad (l > 0). \tag{3.10}$$

An exact solution of this two-dimensional problem is not feasible. We will seek for an approximate solution valid in the limit of large l in order to see how rapidly the series Eq. (3.7) for E_2 converges. First, we inspect the inhomogeneous term in Eq. (3.8) for $\psi_1^{(l)}(r_1, r_2)$. For large values of l the quantity $r_<^l / r_>^{l+1}$ is strongly peaked around the point $r_1 = r_2$, where the two electrons are in the same place. This is prohibited by the Pauli exclusion principle and creates a singularity in the Coulomb potential. The value of $r_<^l / r_>^{l+1}$ rapidly decays to zero as one goes away from the point $r_1 = r_2$. Thus the important contributions to the first-order corrected wave function $\psi_1^{(l)}$ are due to the immediate neighborhood of the point $r_1 = r_2$.

We introduce the Hylleraas coordinates[85]

$$s = r_1 + r_2, \qquad t = -r_1 + r_2. \tag{3.11}$$

Now $\psi_1^{(l)}$ is a function of new variables s and t and can be expressed as

$$\psi_1^{(l)} = \psi_0 f_l(s,t), \tag{3.12}$$

In this new coordinate system Eq. (3.8) becomes

$$\left[\frac{d^2}{ds^2} + \frac{d^2}{dt^2} + \frac{4s}{s^2-t^2}\frac{d}{ds} - \frac{4t}{s^2-t^2}\frac{d}{dt} - \frac{4l(l+1)(s^2+t^2)}{(s^2-t^2)^2} - 2\frac{d}{ds}\right] f_l$$

$$= 2\frac{(s-|t|)^l}{(s+|t|)^{l+1}}. \tag{3.13}$$

For simplicity, we introduce the variable $x = t/s$. Then the right hand side of Eq. (3.13) becomes

$$\frac{2}{s}\frac{(1-|x|)^l}{(1+|x|)^{l+1}}, \tag{3.14}$$

which can be expanded like

$$\left(\frac{1-x}{1+x}\right)^P = \exp\left\{-2Px\left(1 + \frac{x^2}{3} + \frac{x^4}{5} + \cdots\right)\right\}. \tag{3.15}$$

Introducing

$$y = |x|\lambda^{1/2}, \qquad \lambda = (l+\frac{1}{2})^2, \tag{3.16}$$

Eq. (3.13) becomes

$$\left[s^2\frac{d^2}{ds^2} - 2y\frac{d}{dy}s\frac{d}{ds} + y^2\frac{d^2}{dy^2} + 2y\frac{d}{dy} - 2s^2\frac{d}{ds} + 2sy\frac{d}{dy}\right.$$
$$\left. -\frac{4}{1-y^2/\lambda}\left(2y\frac{d}{dy} - s\frac{d}{ds}\right) + \lambda\frac{d^2}{dy^2} - 4\left(\lambda - \frac{1}{4}\right)\frac{(1+y^2/\lambda)}{(1-y^2/\lambda)^2}\right] f_l$$
$$= \frac{2se^{-2y}\exp\left[-(2y^3/3\lambda) - (2y^5/5\lambda^2) - \cdots\right]}{(1-y^2/\lambda)^{1/2}}. \tag{3.17}$$

In removing the absolute value sign we are left with the boundary condition

$$\left.\frac{d}{dy}f_l\right|_{y=0} = 0. \tag{3.18}$$

The variable y has the range $[\,0,\ \lambda^{1/2}]$, but because of the term e^{-2y} on the right, the effective range of f_l is only over values of y of order unity. It is then straightforward to make an expansion in inverse powers of λ:

$$f_l = \lambda^{-1}f^{(-1)} + \lambda^{-2}f^{(-2)} + \cdots \tag{3.19}$$

The equation determining $f^{(-1)}$ is to leading order in λ

$$\left[\frac{d^2}{dy^2} - 4\right] f^{(-1)} = 2se^{-2y}, \tag{3.20}$$

which is easily solved and yields,

$$f^{(-1)} = -\frac{1}{4}se^{-2y}(1+2y). \tag{3.21}$$

Similarly, we obtain the next order term

$$f^{(-2)} = -\frac{1}{4}se^{-2y}[(-\frac{1}{3}y^4 - \frac{4}{3}y^3 - y^2 + \frac{3}{4}y + \frac{3}{8}) - s(\frac{2}{3}y^3 + y^2 + y + \frac{1}{2})]. \tag{3.22}$$

The second order energy $E_2(l)$ becomes

$$E_2(l) = \frac{1}{\lambda}\int_0^\infty ds\, s^4 e^{-2s} \int_0^{\lambda^{1/2}} dy\, e^{-2y}\frac{(1-y^2/\lambda)^2}{(1-y^2/\lambda)^{1/2}} f_l \exp\left[-\frac{2y^3}{3\lambda} - \frac{2y^5}{5\lambda^2}\cdots\right]. \tag{3.23}$$

This can also be expanded as a power series in λ^{-1} and, accounting for the first two terms, Eq. (3.23) becomes

$$E_2(l) = -\frac{45}{256}\frac{1}{(l+\frac{1}{2})^4}\left[1 - \frac{19/8}{(l+\frac{1}{2})^2} + O\left(\frac{1}{l^4}\right)\right]. \tag{3.24}$$

This indicates that the second-order energy $E_2(l)$ converges very slowly (at a rate of l^{-4}) with increasing l. The second term of Eq. (3.24) indicates that one does not have to go to very large l values before this asymptotic formula is usable.

3.2.2 MP2-R12

Quantum chemical ab initio methods have two main sources of error: (1) the truncation of the one-electron basis set and (2) the use of an approximation to the "full CI" wave function in the given basis. The basis truncation problem is due to the slow convergence of the energy or properties with increasing basis set size to the basis set limit.

The slow convergence of the CI expansion is due to the inability to correctly describe the correlation cusp[87] at $r_{12} = 0$,

$$\lim_{r_{12}\to 0}\left(\frac{\partial\psi}{\partial r_{12}}\right)_{av} = \frac{1}{2}\psi(r_{12}=0), \tag{3.25}$$

which is the direct consequence of the singularity of the Coulomb repulsion.

For two-electron systems partial wave functions of the form of Eq. (3.6) converge more and more slowly with increasing l due to the increasingly singular nature of the partial wave contribution $r_<^l/r_>^{l+1}$ to r_{12}. The convergence of the partial wave expansion can be sped up considerably by including the inter-electronic terms (r_{12}) in the wave function explicitly, which can be expressed as,[88]

$$\psi(r_1, r_2) = \frac{1}{2} r_{12} \psi_0(r_1, r_2) + \chi, \tag{3.26}$$

where

$$\chi = \sum_l \chi^l(r_1, r_2) P_l(cos\theta), \tag{3.27}$$

and ψ_0 is some single-particle model reference function. Now, substituting Eq. (3.26) into the inhomogeneous differential equation Eq. (3.4), we get

$$(H_0 - E_0)\chi_1 = -(H_{12} - E_1)\psi_0. \tag{3.28}$$

The noticeable thing is that the interaction potential $g_{12} = r_{12}^{-1}$ is replaced by the residual interaction operator H_{12}, where

$$\begin{aligned} H_{12} &= [T_1 + T_2, \frac{1}{2} r_{12}] + g_{12}, \\ &= -\frac{1}{2} \frac{\mathbf{r}_{12}}{r_{12}} (\nabla_1 - \nabla_2). \end{aligned} \tag{3.29}$$

The operator H_{12} is much less singular at $r_{12} = 0$ than g_{12}.

The second order energy can be obtained from the Hylleraas functional,[85]

$$\mathscr{F}(\psi) = 2Re\langle\psi|V - E_1|\psi_0\rangle + \langle\psi|H_0 - E_0|\psi\rangle \tag{3.30}$$

If ψ_0 is the Hartree-Fock function, then $\mathscr{F}(\psi)$ decouples into a sum of pair contributions

$$\mathscr{F}(\psi) = \sum_{i<j\neq 1}^{n} f(u_{ij}). \tag{3.31}$$

Where

$$u_{ij}(1,2) = \frac{1}{2} c_{ij} \{1 - P(1)\}\{1 - P(2)\} r_{12}[ij] + \sum d_{ij}^{ab}[ab], \tag{3.32}$$

$$P(a) = \sum_k |\psi_{0k}(a)\rangle\langle\psi_{0k}(a)|, \qquad a = 1, 2 \tag{3.33}$$

where ψ_{0k} is the exact eigenfunction of the Fock operator and

$$[pq] = \frac{1}{\sqrt{2}} \{\psi_{0p}(1)\psi_{0q}(2) - \psi_{0q}(1)\psi_{0p}(2)\}. \tag{3.34}$$

The labels $i, j, k, \cdots$ refer to occupied spin orbitals, $a, b, c, \cdots$ to unoccupied and $p, q, r, \cdots$ to arbitrary spin orbitals and c_{ij} and d_{ij}^{ab} are linear variational parameters.

Substituting Eq. 3.32 into Eq. 3.30, yields

$$f(u_{ij}) = 2Re\langle[ij]|r_{12}^{-1}|u_{ij}(1,2)\rangle + \langle u_{i,j}(1,2)|F(1)+F(2)-\epsilon_i-\epsilon_j|u_{ij}(1,2)\rangle, \qquad (3.35)$$

where F is the Fock operator. Exact evaluation of Eq. (3.35) requires the evaluation of three electron integrals. Assume now ψ_{0i} and ψ_{0j} are the exact eigenfunctions of the Fock operator and that the orbital set $\{\psi_{0p}\}$ is complete in one-electron space, i.e.,

$$1 = \sum_p |\psi_{0p}(1)\rangle\langle\psi_{0p}(1)| \equiv Q(1). \qquad (3.36)$$

The result of inserting these assumptions into Eq. (3.35) is that the pair energy becomes a sum of two terms: the conventional MP2 pair energy e_{ij} and a term correcting for the incompleteness in two-electron space,

$$\sum_{k=1}^{3}\langle[ij]|r_{12}[1-Q(1)Q(2)]A_k|[ij]\rangle, \qquad (3.37)$$

where

$$A_1 = (c_{ij}-\frac{1}{2}c_{ij}^2)\frac{1}{r_{12}}, \qquad (3.38)$$

$$A_2 = -\frac{1}{4}c_{ij}^2\frac{1}{r_{ij}}(r_{ij}.(\nabla_1-\nabla_2)), \qquad (3.39)$$

$$A_3 = -\frac{1}{4}c_{ij}^2[K(1)+K(2), r_{12}]. \qquad (3.40)$$

Here K denotes the exchange operator.

In the atomic case, the MP2 partial wave expansion has increments $\sim (l+1/2)^{-4}$ for each successive completed $l-$shell in the basis set. The term A_1 looks after this deficiency, and the terms A_2 and A_3 look after the $\sim (l+1/2)^{-6}$ deficiency. Thus, if the terms A_1, A_2 and A_3 are computed, the partial wave expansion for the MP2-R12 method will have increments $\sim (l+1/2)^{-8}$. This convergence is much faster than the conventional MP2 convergence.[71,74,78,89]

Density fitting has been recently introduced to approximate all of the 4-index 2-electron integrals in the explicitly correlated MP2-R12 theory, which requires only 2- and 3-index integrals over various 2-electron operators. This method is called DF-MP2-R12[90–92] and is computationally much faster than the MP2-R12 method.

Table 3.1: Calculation time (CPU time in second) and require memory space (in MB) for different level of methods with different size of basis set for H_2O.

Basis	MP2		CCSD(T)		DF-MP2		DF-MP2-R12	
function	CPU	Mem.	CPU	Mem.	CPU	Mem.	CPU	Mem.
cc-pVDZ	0.35	2.23	0.46	6.11	2.10	3.41	3.52	2037.76
cc-pVTZ	1.01	5.14	1.54	13.01	3.66	4.21	15.88	2037.76
cc-pVQZ	7.51	40.10	13.74	104.22	10.53	6.87	84.98	2037.76
cc-pV5Z	81.18	164.23	137.37	971.75	47.18	13.89	497.85	2037.76
cc-pV6Z	675.64	863.73	987.49	6103.04	Not Available			

A further improvement over MP2-R12 has been proposed by Manby recently,[93] where the explicit inter-electronic distance r_{12} is replaced by an arbitrary function f_{12} of the inter-electronic distance. This method is known as the MP2-F12 method. It is generally more accurate and also converges faster than the MP2-R12 method.

3.2.3 Cost of different methods

We have now several highly accurate quantum chemical processes for atomic and molecular properties evaluation. Their ability and cost effectiveness bound us to choose them for different applications. The most accurate method in our hand is r_{12} dependent explicit correlation method (MP2-R12). Although this method takes care the correlation cusp and speed up the convergence behavior of basis sets but still it is restricted to apply on only very small systems because of its computational cost. Where as DF-MP2-R12 speed up the calculation further over MP2-R12 but still prohibitively costly. On the other hand highly accurate coupled cluster method is several times much faster than the DF-MP2-R12 method, see Table 3.1. We have now a huge computational power in our hand, but still performing CCSD(T) calculations with large basis sets, even if for medium size of molecule is prohibitively costly. Where as MP2 method is much much less expensive then the coupled cluster and explicit correlated methods. DF-MP2 method speed up the calculation again over regular MP2 method, see Table 3.1. Basis sets size Vs. normalized CPU time plotted in Fig. 3.1, which clearly indicates that coupled cluster and explicit correlated methods are not economical in compare to regular MP2 and DF-MP2 methods. In such a situation it is a great challenge whether extrapolation improve the results or not. Since for extrapolation one need not to pay further any cost, so it is a time to find out a general extrapolation scheme which can efficiently estimate the basis set limit results for polyatomic systems.

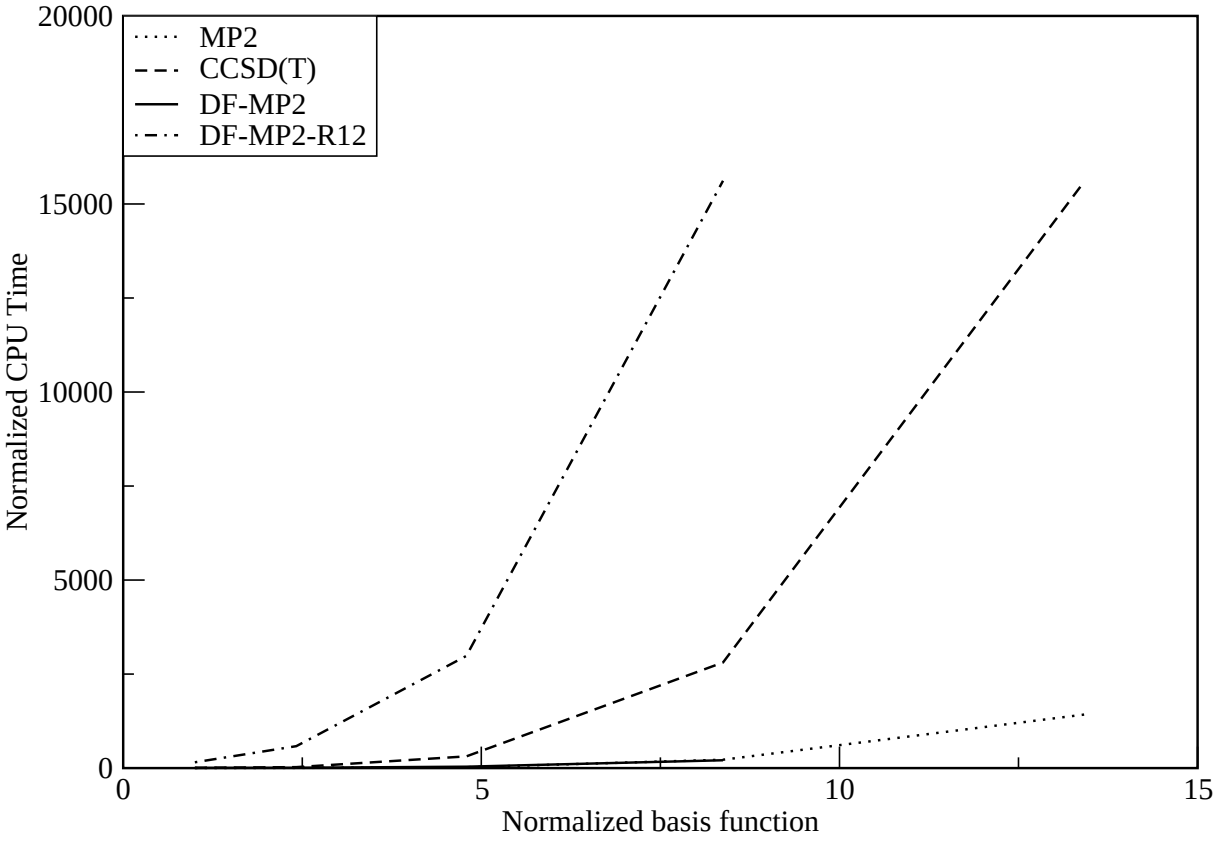

Figure 3.1: Normalized basis function vs. normalized CPU time for different level of calculations for water molecule.

3.3 Extrapolation methods

The extrapolation based methods have become very popular after the development of the family of correlation consistent basis sets cc-pVXZ and aug-cc-pVXZ (X = 2, 3, 4 $\cdots$ are the cardinal numbers of the basis sets) by Dunning and coworkers[53,54,56,57,94–97] who provided a systematic study of the convergence behavior of the energy and other molecular properties as the basis-set limit is approached. With the knowledge of basis-set convergence behavior one can predict the complete basis-set-limit results through extrapolation without actually performing computations close to the basis-set limit.

3.3.1 Exponential extrapolation

Feller used Dunning's sequential basis sets to calculate the energy of the water molecule[98] and water dimer[99] and proposed a simple three-parameter exponential extrapolation

model for the total energy

$$E_{\infty}^{\text{total}} = a + b \exp(-cX), \tag{3.41}$$

where $E_{\infty}^{\text{total}}$ is the total energy at the basis-set limit. The convergence behavior of the total energy for the H_2O molecule is shown in Fig. 3.2. It is clear from Eq. (3.41) that,

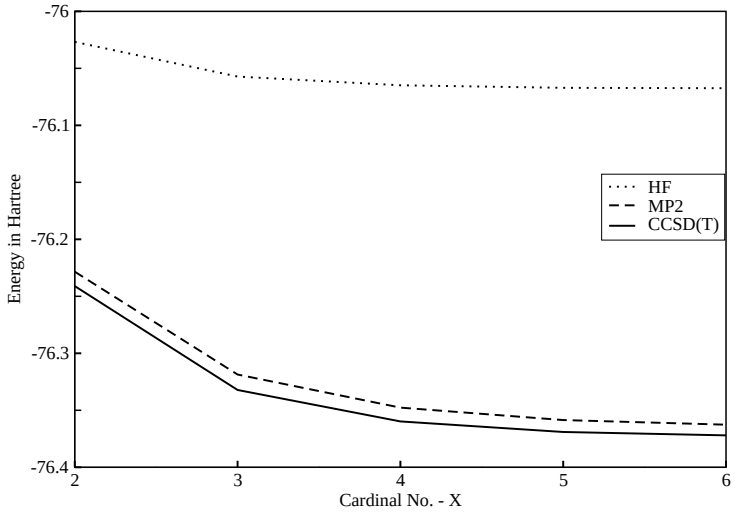

Figure 3.2: Convergence behavior of the HF, MP2 and CCSD(T) total energy with the size of the basis set for the water molecule at equilibrium geometry.

in order to calculate the basis-set-limit energy, at least three energy points are required. The exponential model was found to provide a reasonable fit for the Hartree-Fock total energy (see Refs. 98–102).

3.3.2 Power-law extrapolation of the total energy

The exponential fit proposed by Feller[99] was so successful for Hartree-Fock energies and hence for total energies, that it would appear natural to use the same extrapolation scheme for total energies from correlated methods as well. The total energy can usually be approximately described by an exponential fit since it is (for small basis sets at least) dominated by the HF energy (see Table 3.2). Yet the total energy from a correlation method does not actually follow an exponential law (see Fig. 3.4). Table 3.2 presents the total energy for different basis sets, and the energy differences between any two neighboring basis sets (ΔE) for the water molecule. For the HF energy, this difference decreases very quickly and there is essentially no gain in accuracy from the 5Z to the 6Z result. This convergence behavior of the HF energy matches reasonably well with the exponential fit proposed by

Feller. But for MP2 and CCSD(T) calculations, the convergence with basis-set size is very slow and far from being converged at the 5Z basis set (see Table 3.2). It shows some inverse power law rather than exponential behavior. A power-law convergence for correlated methods has therefore been proposed by Martin[103] where

$$E_\infty = A + B(l + 1/2)^\alpha, \qquad \alpha < 0. \tag{3.42}$$

Here l is the maximum orbital angular-momentum quantum number present in the basis set and α is an empirical parameter.

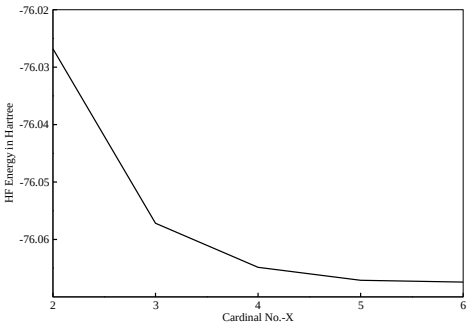

Figure 3.3: Convergence behavior of the HF energy with the size of the basis set for the water molecule at equilibrium geometry using cc-pVXZ basis sets.

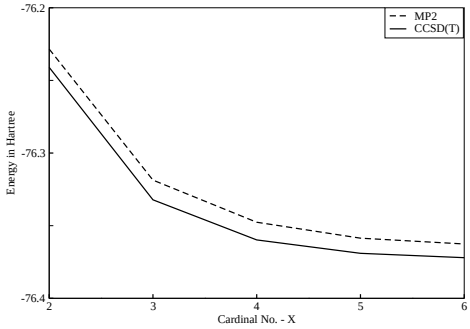

Figure 3.4: Convergence behavior of the MP2 and the CCSD(T) total energy with the size of the basis set for the water molecule at equilibrium geometry using cc-pVXZ basis sets.

Table 3.2: Convergence behavior of the HF, MP2 and CCSD(T) total energy for the H_2O at equilibrium geometry. All energies are in mE_h. ΔE is the energy difference between any two consecutive basis set energies.

Basis	HF		MP2		CCSD(T)	
	Total	ΔE	Total	ΔE	Total	ΔE
cc-pVDZ	-76026.808		-76228.417		-76241.016	
cc-pVTZ	-76057.181	-30.373	-76318.630	-90.213	-76332.189	-91.172
cc-pVQZ	-76064.848	-7.667	-76347.634	-29.004	-76359.793	-27.605
cc-pV5Z	-76067.104	-2.255	-76358.598	-10.964	-76369.040	-9.246
cc-pV6Z	-76067.420	-0.316	-76362.611	-4.013	-76372.017	-2.977

Truhlar suggested another approach where the HF and correlation energy are treated separately.[104] The total energy for a correlated method is the sum of the HF and the

correlation energy, and can be written as,

$$E^{\text{tot}} = E^{\text{HF}} + E^{\text{corr}}. \tag{3.43}$$

For both contributions Truhlar attempted an inverse power law fit but with different exponents, as,

$$E_{\text{X}}^{\text{HF}} = E_{\infty}^{\text{HF}} + A^{\text{HF}} X^{-\alpha}, \tag{3.44}$$

$$E_{\text{X}}^{\text{corr}} = E_{\infty}^{\text{corr}} + A^{\text{corr}} X^{-\beta}. \tag{3.45}$$

Where E_X^{corr} is the correlation energy for a basis set of cardinal number X and E_{∞}^{corr} is the basis-set limit correlation energy. After optimizing over a small set of molecules, the following values for the parameters were obtained:[104] $\alpha = 3.4$, $\beta_{MP2} = 2.2$ and $\beta_{CCSD} = \beta_{CCSD(T)} = 2.4$.

Although this method gives a reasonable fit for double-zeta and triple-zeta basis sets, it does not follow the actual convergence behavior. Focussing on the correlation energy, it becomes clear that the asymptotic behavior of the correlated methods is due to the correlation energy as HF is already converged for smaller values of X. Table 3.3 shows the convergence behavior for MP2 and CCSD(T) correlation energy. Since the correlation

Table 3.3: Convergence behavior of the MP2 and CCSD(T) correlation energy with basis set size for water molecule at equilibrium geometry. ΔE is the energy difference between two consecutive basis set energies.

Basis	MP2 (mE_h)		CCSD(T) (mE_h)	
	Corr.	ΔE	Corr.	ΔE
cc-pVDZ	-201.609		-214.208	
cc-pVTZ	-261.449	-59.840	-275.008	-60.800
cc-pVQZ	-282.786	-21.336	-294.945	-19.937
cc-pV5Z	-291.494	-8.709	-301.936	-6.991
cc-pV6Z	-295.191	-3.697	-304.597	-2.661

part is the most expensive part of the calculation and it converges frustratingly slowly with basis set size, it is necessary to concentrate mainly on the correlation energy to find a reasonable extrapolation method.

3.3.3 Extrapolation of the correlation energy

Schwartz's work on the two-electron atom[33] (He) (sec 3.2.1) supports the empirical finding (Eq. (3.45)) that the correlation energy does not converge exponentially, but rather as an

inverse power of the highest angular-momentum quantum number present in the basis set. Investigating the second-order energy of the Z^{-1} perturbation expression for the ground state of He (Eq. (3.24)), Schwartz showed that the asymptotic convergence with respect to the orbital angular-momentum quantum number ℓ contained in the one-particle basis follows

$$E_\ell^{\rm corr} = -\frac{45}{256}(\ell+\frac{1}{2})^{-4}\left\{1-\frac{19}{8}(\ell+\frac{1}{2})^{-2}+\mathscr{O}\left[(\ell+\frac{1}{2})^{-4}\right]\right\}. \tag{3.46}$$

Carroll et al. also found empirically a similar behavior based on configuration interaction results for the helium atom,[105] which was later analyzed in more detail by Hill[106]

$$E_\ell^{\rm corr} - E_{\ell-1}^{\rm corr} = a\,(\ell+\frac{1}{2})^{-4} + b\,(\ell+\frac{1}{2})^{-5} + \mathscr{O}\left\{(\ell+\frac{1}{2})^{-6}\right\}, \tag{3.47}$$

where $a = -0.074$ and $b = -0.031$. This asymptotic expression represents the energy increment for adding a saturated shell of atomic basis functions of angular momentum quantum number ℓ. The above analysis is based on the helium atom where the basis sets at each level are completely saturated with respect to the radial part of the basis functions. For molecules, such a formula can be an approximation only, since the orbital angular momentum quantum number is not a good quantum number anymore. In fact, for the correlation-consistent polarized basis sets,[56,107] the highest angular momentum ($\ell \leq L$) in the basis is $L = X - 1$ for H and He, and $L = X$ for second row atoms (Li to Ne). Martin therefore proposed a compromise value for L of $L = X - \frac{1}{2}$.[103]

Two-point extrapolation

For a near complete-basis set, Halkier et al. proposed a very simple relation[31] based on only the first dominating term of Eqs. (3.46) and (3.47),

$$E_{\rm X}^{\rm corr} = E_\infty^{\rm corr} + AX^{-3}, \tag{3.48}$$

where A is an empirical parameter. Since there are only two unknown parameters (A and $E_\infty^{\rm corr}$), two energy points are sufficient to estimate the basis-set limit value. Halkier et. al. studied their model extensively for Ne, N_2 and H_2O and tabulated errors from extrapolations with different combinations of basis sets. Table 3.4 presents their main results, where all the errors are with respect to MP2-R12 results using a cc-pV6Z basis set. They concluded that the best extrapolations are obtained using two-point fits with the two highest cardinal numbers available. The best limit energy will thus be obtained

Table 3.4: Mean error, standard deviation, mean absolute error, and maximum absolute error of the extrapolated basis-set limits using two-point extrapolation formula for data with $X = X_{\min}, \cdots, X_{\max}$. All results are for the water molecule and errors are in mEh. [See Ref. 31.]

$X_{\min}$	$X_{\max}$	Mean *error*	Standard *deviation*	Mean abs. *error*	Max abs. *error*
2	3	13.54	5.22	13.54	22.50
2	4	8.23	4.00	8.23	15.27
2	5	5.66	3.12	5.66	11.12
2	6	4.27	2.52	4.27	8.66
3	4	0.73	2.42	2.01	5.08
3	5	0.27	1.70	1.41	3.17
3	6	0.16	1.30	1.08	2.31
4	5	-0.35	0.86	0.86	1.41
4	6	-0.17	0.69	0.63	1.26
5	6	0.06	0.52	0.42	1.07

from quintuple-zeta and sextuple-zeta basis sets. If not available, always the two highest possible basis sets $(X, X+1)$ will give the limit energy with somewhat lower accuracy but with the best result for this maximum of X.[31] It was also suggested that easily generated energies of cardinal numbers $X < 5$ should be omitted from the fit as they contain less information about the asymptotic limit than the points with $X \geq 5$ and will thus introduce more noise than information in the extrapolation. It was also strongly suggested not to use double-zeta basis set even if higher basis set energies are not possible to calculate.[31]

This approach proved to be successful for large values of X, which is to be expected from their basic assumption,[31] that "for a near complete basis set, the first term of Eqs. (3.46) and (3.47) dominates". With this in mind it becomes clear why Eq. (3.48) is successful for large X values and that the inclusion of double-zeta basis sets in the extrapolation lowers the accuracy significantly.[31]

Cost analysis of Halkier's approach.

Although the two-point extrapolation method yields accurate basis-set limits for the correlation energy,[31] it is not at all economical, since it needs very large basis sets for more accurate results. For comparison, Table 3.5 presents the required memory and compu-

tational time (CPU time) for the H_2O and CH_3NH_2 molecules, for different levels of computations (SCF, MP2, CCSD(T)) and different basis sets (cc-pVDZ to cc-pV6Z). An increase of almost one order of magnitude in memory is observed for each increment in Dunning's sequential basis sets for the water molecule. The corresponding rate of increase for the methylamine is even higher which indicates that as the molecular size, and especially the number of heavy atoms increases, so does the memory consumption (see Table 3.5). This is quite obvious, because a large number of basis functions are required for the heavy atoms. Not only the memory consumption but also the required computational

Table 3.5: CPU time and memory requirement for different levels of computation with different size of basis sets, for the water and methylamine molecules at their equilibrium geometry.

Molecule	Basis	CPU time (in Sec.)			Memory
		SCF	MP2	CCSD(T)	MB
	cc-pVDZ	0.03	0.06	0.16	8.02
	cc-pVTZ	0.23	0.10	1.51	14.91
H_2O	cc-pVQZ	2.23	0.55	14.81	106.12
	cc-pV5Z	28.30	5.98	131.31	973.65
	cc-pV6Z	60.28	17.44	410.23	6103.04
	cc-pVDZ	0.22	0.17	2.78	13.46
	cc-pVTZ	5.31	3.54	83.49	317.41
CH_3NH_2	cc-pVQZ	88.84	61.30	1206.63	4843.52
	cc-pV5Z	1080.26	618.93	10529.60	43816.96
	cc-pV6Z	Not Available			

CPU time increases steeply with the size of the basis sets. As a matter of fact, correlation energy calculations with larger basis sets soon become unaffordable (see Table 3.5). In such a situation getting a 5Z or 6Z quality basis set result for the extrapolation is prohibitively expensive. Although one may still afford this for a single-point energy calculation, for dynamics it becomes unrealistic since a large number of single-point energy calculations are required for a reasonable representation of the PES.

Fifth order correction

The asymptotic behavior of the correlation energy is dominated by the fourth order term of Schwartz's equation (Eq. (3.46)) but higher order terms can also have a significant influence on the convergence behavior. This is particularly important when extrapolating

from smaller basis sets. Following a suggestion by Varandas,[34] including the next order term from Schwartz's formula yields

$$E_{\mathrm{X}}^{\mathrm{corr}} = E_{\infty}^{\mathrm{corr}}\{1 + A_1 X^{-3}(1 + A_2 X^{-1})\}, \tag{3.49}$$

where A_1 and A_2 are two empirical parameters.

Clearly Eq. (3.49) requires at least three energy points to get the basis set limit energy. After analyzing known results for a large number of molecules, Varandas found that A_1 and A_2 are approximately exponentially correlated for a set of atoms and molecules[34] from first and second row elements. This allows us to express one of them as a function of the other parameter and again yield a two point extrapolation formula which we will call X5. This name indicates that the fifth order term has been taken into account from Schwartz's formula.

3.3.4 Gradient and Hessian calculation for the correlation method

The success of vibrational spectroscopy simulations depends upon the accuracy of the PES. To generate a smooth PES, it is necessary to calculate a large number of single energy points on a grid of well chosen coordinates. Such a large number of energy calculations are not only time consuming but also economically unfavorable. We have just shown how to reduce the computational cost by several orders of magnitude by extrapolation of the energy to the basis set limit. A further reduction of the cost can be achieved by interpolating fewer energy points on non-uniform grids. Several multidimensional interpolation methods have been available for a long time, but most of them are applicable only for regularly spaced data points. At the end of the 1960's Donald Shepard introduced an interpolation scheme based on the weighted average of the data points.[108] This method is thus capable of generating a smooth surface from irregularly spaced data points. The main requirements for the modified Shepard interpolation[108] are the gradient and the Hessians at each data point. The calculation of gradient and Hessians are the most expensive part in the computation chemistry especially with the larger basis set. Like the energy, gradient and Hessians are also converge with the basis sets size. Therefore, gradient and Hessians also can be extrapolated if there is any simple relation.

The MP2 total energy is define as,

$$E_{MP2} = E_{corr} + E_{HF}. \tag{3.50}$$

Differentiating both sides of the Eq. (3.50) with respect to the coordinate q_i we get,

$$\frac{\partial E_{MP2}}{\partial q_i} = \frac{\partial E_{corr}}{\partial q_i} + \frac{\partial E_{HF}}{\partial q_i}. \tag{3.51}$$

The MP2 gradient is the sum of the HF gradient and the gradient of the correlation part.

Differentiating both sides of Eq. (3.50) with respect to the coordinates q_i, q_j we get,

$$\frac{\partial^2 E_{MP2}}{\partial q_i \partial q_j} = \frac{\partial^2 E_{corr}}{\partial q_i \partial q_j} + \frac{\partial^2 E_{HF}}{\partial q_i \partial q_j}. \tag{3.52}$$

As for the gradient, the total Hessian for the correlated method is also the sum of the HF Hessian and the Hessian of the correlation energy.

3.3.5 Extrapolation of Gradient and Hessians

It was observed that like for the correlation energy, the correlation part of the gradient and the Hessian converged slowly. It is therefore more important to concentrate the efforts on the correlation part of the gradient and Hessian. Since the two-point extrapolation and the X5 method provide successful schemes to extrapolate the correlation energy, it is natural to apply them also to extrapolate the gradient and Hessian.

Two point extrapolation

Let us consider two energy points with basis sets of cardinal number X and Y. Then the two point extrapolation is defined as

$$E_X = E_\infty + AX^{-3}, \tag{3.53}$$

$$E_Y = E_\infty + AY^{-3}. \tag{3.54}$$

Now solving above two equations for the parameter A we get,

$$A = \frac{E_X - E_Y}{X^{-3} - Y^{-3}}.$$

Therefore Eq. (3.53) becomes,

$$E_\infty = E_X - \frac{E_X - E_Y}{X^{-3} - Y^{-3}} X^{-3}, \tag{3.55}$$

$$E_\infty = (1-p)E_X + pE_Y, \tag{3.56}$$

where

$$p = \frac{X^{-3}}{X^{-3} - Y^{-3}}.$$

Differentiating Eq. (3.56) successively we get the expressions for the basis set limit gradient and Hessian.

$$\frac{\partial E_\infty}{\partial q_i} = (1-p)\frac{\partial E_X}{\partial q_i} + p\frac{\partial E_Y}{\partial q_i} \tag{3.57}$$

$$\frac{\partial^2 E_\infty}{\partial q_i \partial q_j} = (1-p)\frac{\partial^2 E_X}{\partial q_i \partial q_j} + p\frac{\partial^2 E_Y}{\partial q_i \partial q_j} \tag{3.58}$$

Extrapolation with the X5 method

Since A_2 is a slowly varying parameter we can assume $\frac{\partial A_2}{\partial q_i} = 0$ for all q_i and the expression for the gradient and the Hessian at basis set limit becomes,

$$\frac{\partial E_\infty}{\partial q_i} = \frac{\frac{\partial E_Y}{\partial q_i}X^{-3} - \frac{\partial E_X}{\partial q_i}Y^{-3} + A_2(\frac{\partial E_Y}{\partial q_i}X^{-4} - \frac{\partial E_X}{\partial q_i}Y^{-4})}{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})}, \tag{3.59}$$

and

$$\frac{\partial^2 E_\infty}{\partial q_i \partial q_j} = \frac{\frac{\partial^2 E_Y}{\partial q_i \partial q_j}X^{-3} - \frac{\partial^2 E_X}{\partial q_i \partial q_j}Y^{-3} + A_2(\frac{\partial^2 E_Y}{\partial q_i \partial q_j}X^{-4} - \frac{\partial^2 E_X}{\partial q_i \partial q_j}Y^{-4})}{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})}. \tag{3.60}$$

(For details see Appendix A.2)

If A_1 and A_2 are both constant, then the expressions for the gradient and the Hessians simplify to

$$\frac{\partial E_\infty}{\partial q_i} = \frac{\frac{\partial E_X}{\partial q_i}}{1 + A_1 X^{-3}(1 + A_2 X^{-1})} \tag{3.61}$$

$$\frac{\partial^2 E_\infty}{\partial q_i \partial q_j} = \frac{\frac{\partial^2 E_X}{\partial q_i \partial q_j}}{1 + A_1 X^{-3}(1 + A_2 X^{-1})} \tag{3.62}$$

3.4 Choice of model systems

To verify the applicability of Eq. (3.49) we chose a few small and medium size molecules relevant to our prime objective, i.e., to obtain a parameterized model for larger molecules, mainly proteins and other bio-molecules. The basic structural unit of proteins is the amino acid. An amino acid consists of an amino group, a carboxyl group, a α-hydrogen

atom, and a distinctive R group, called sidechain, bonded to a central carbon atom which is called the α-carbon.[109] Twenty different kinds of sidechains varying in size, shape, charge, hydrogen bonding capacity and chemical reactivity occur naturally in proteins.

Since the amino acids differ from each other by the presence of different sidechains, it is necessary to study the side chain behavior. The characteristics of the sidechain mainly depend upon the characteristics of functional groups present in the sidechain. These functional groups can be represented by small model compounds. We chose some of them for our model systems along with few other atoms and molecules with first and second row elements. The model systems are H_2, He, Ne, N_2, H_2O, CH_2O, CH_4, C_2H_2, C_2H_4, C_2H_6, N_2H_2, N_2H_4, NH_3, CH_3OH, CH_3NH_2, CH_3COOH, C_6H_6, and N-methylacetamide (NMA). Some of the structural similarity of model systems and protein sidechains are shown in Fig. 3.5.

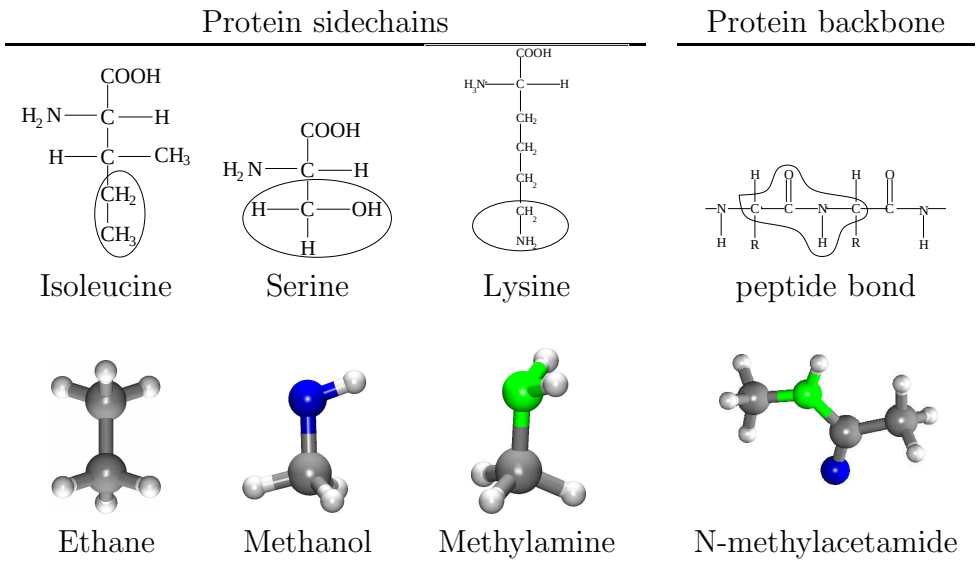

Figure 3.5: Structural similarity between functional group of protein side chain and small molecules.

3.5 Computational details

The molecular geometry of each of the chosen molecules was optimized employing density functional theory (DFT),[110] with Becke's 3-parameter Lee, Yang, Parr (B3LYP) functional[111] and aug-cc-pVTZ basis sets, using the GAUSSIAN[112] quantum chemical computational package. We have calculated total energies at the equilibrium geometry of

Table 3.6: The extrapolated basis set limits using the two point formula for the water at equilibrium geometry. Different combination of energy points were used for the extrapolation. All energies are in mE_h.

Basis	MP2	Extrapolated energy			
set	corr	DZ	TZ	QZ	5Z
DZ	-201.608				
TZ	-261.449	-286.645			
QZ	-282.785	-294.382	-298.355		
5Z	-291.494	-297.640	-299.771	-300.631	
6Z	-295.190	-297.640	-300.011	-300.414	-300.269

each molecule and also for some geometries corresponding to deformations along internal coordinates (stretch, bend, torsion) of the molecules (without partial re-optimization). All energy points have been calculated at the level of second-order Møller-Plesset perturbation theory (MP2)[61–63] and coupled-cluster theory with singles and doubles excitations and a perturbative triples correction [CCSD(T)][113] using Molpro.[114] For both levels of theory, Dunning's cc-pVXZ and aug-cc-pVXZ basis sets have been used.[53,54,56,57,94–97] Some MP2 energy points were also calculated using the TURBOMOLE[115] package. In some cases, correlation energies have been calculated with density fitting MP2 (DF-MP2) also known as resolution of identity MP2 (RI-MP2) method.[116–118]

In our study, the (QZ - 5Z) two point extrapolated results were generally used as a reference since, according to Halkier et. al., the highest possible two basis-set results yield the best possible estimate of the basis-set limit energy. Although for some small molecules and atoms in our data set 6Z basis set computations are possible this would be prohibitively expensive for most of the medium sized molecules and for molecules like benzene, CH_3COOH, CH_3NH2, CH_3OH, and N-methylacetamide (NMA) it is beyond our current computational resources. For reasons of consistency, we thus omit the 6Z basis set results for small molecules and atoms and use only up to the 5Z basis set for all systems.

A comparison of different levels of the extrapolated results for the water molecule are presented in Table 3.6. The two point extrapolation from the QZ-5Z basis is about 0.36 mE_h apart from th 5Z-6Z extrapolated result, which indicates that the two point extrapolation from the QZ-5Z generally provides a reasonable reference.

3.6 Results and discussion

The basis set extrapolation of the MP2 and CCSD(T) correlation energy has been studied for a representative set of atoms and molecules. In the following sections we discuss the results and characteristics of the X5 extrapolation method. Two possible extrapolation formula and their performance is discussed in section 3.6.1. The characteristics of the exponential A_1/A_2 relation ship for the MP2 and CCSD(T) methods is discussed in section 3.6.2. The following section shows that MP2 and RI-MP2 share identical fitting parameters if appropriate fitting basis sets are chosen. It is also discussed here how the parameterization changes for different families of basis sets. The general features of the X5 extrapolation method are described in section 3.6.4. The efficiency of the X5 method to calculate the PES with global parameters is presented in section 3.6.5. Section 3.6.6 shows how the local parameterization improves the accuracy of the PES calculation in the vicinity of a fixed point. The convergence behavior of the gradient and the Hessian are presented in section 3.6.7.

3.6.1 Choice of the extrapolation method

In analogy to Eq. (3.48) an alternative extrapolation formula for the correlation energy is conceivable which differs from Eq. (3.49) in the inclusion of E_∞^{corr}:

$$E_X^{corr} = E_\infty^{corr} + \tilde{A}_1 X^{-3}(1 + \tilde{A}_2 X^{-1}) \tag{3.63}$$

where $\tilde{A}_1$ and $\tilde{A}_2$ are again two empirical parameters. A_1 vs. A_2 and $\tilde{A}_1$ vs. $\tilde{A}_2$ for a set of molecules and different combinations of energy points are plotted in Figs. 3.6 and 3.7 respectively.

As seen from Fig. 3.6, the variation of A_1 and A_2 parameters for a set of atoms and molecules consisting of first and second row elements exhibits a noticeable correlation and one of the parameters may be expressed as a function of the other. Then, essentially, Eq. (3.49) depends only on a single empirical parameter. Therefore only two energy points are sufficient to estimate the basis set limit using Eq. (3.49) and makes it a two point extrapolation formula.

As suggested by Varandas,[34] A_2 can be express as

$$A_2 = a \exp(b\, A_1 + c). \tag{3.64}$$

On the other hand, $\tilde{A}_1$ vs. $\tilde{A}_2$ does not exhibit any correlation (see Fig. 3.7) for the same set of atoms and molecules. Thus $\tilde{A}_1$ and $\tilde{A}_2$ are independent parameters. To get

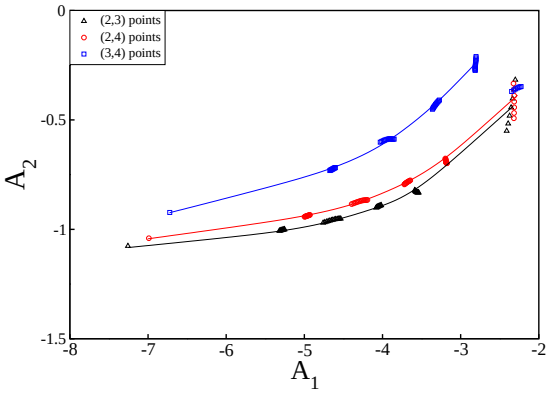

Figure 3.6: Correlation between A_1 and A_2 for different combinations of the energy points for Eq. (3.49).

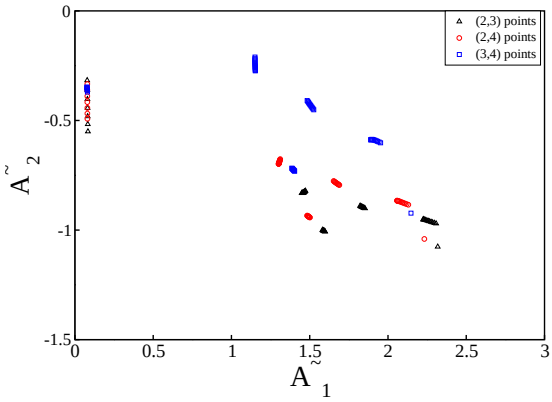

Figure 3.7: Correlation between $\tilde{A}_1$ and $\tilde{A}_2$ for different combinations of the energy points Eq. (3.63).

the basis set limit using Eq. (3.63) we need at least three energy points, which means even if we start from the DZ basis set, it is necessary to perform a QZ basis set calculation which makes the extrapolation method expensive again. For that reason we will drop Eq. (3.63) and all further investigations have been carried out only with Eq. (3.49). We will call this extrapolation model X5, which indicates that the fifth order term has been taken into account from Schwartz's formula. For further studies we strictly use the DZ-TZ basis set results to estimate the basis set limit energy using the X5 extrapolation method.

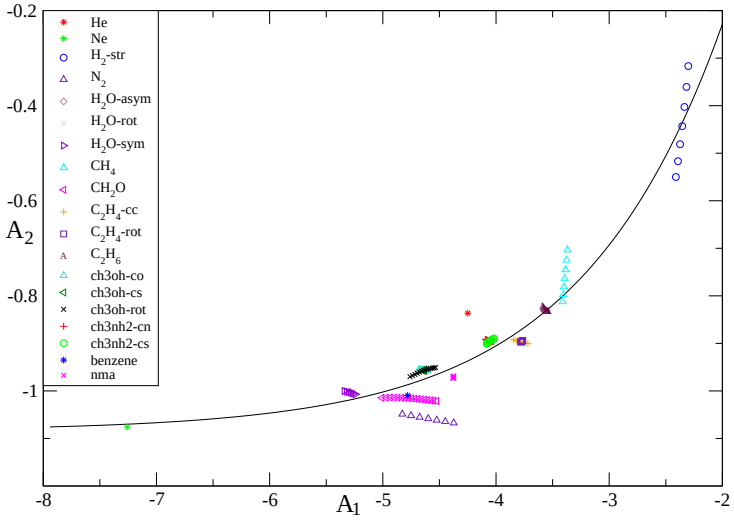

Figure 3.8: Correlation between A_1 and A_2 at the MP2 level with the cc-pVXZ basis set.

The parameters A_1 and A_2 obtained for a set of atoms and molecules are plotted in

Fig. 3.8, where energy points are calculated at MP2/cc-pVXZ level. For each molecule a set of configurations representing deformations along vibrational and rotational internal coordinates have been chosen. Least-squares fit of Eq. (3.64) over the total data set yields $a = 4.8278$, $b = 0.8397$ and $c = -1.0843$.

3.6.2 MP2 vs. CCSD(T) parameterization

A noticeable change is observed in the exponential fitting curves when the energy is calculated at different levels of theory (e.g., MP2 and CCSD(T)) employing the same family of basis sets. For a set of atoms and molecules the function (Eq. (3.64)) is plotted for the A_1 and A_2 parameters for MP2 and CCSD(T) in Fig. 3.9. The dependence on the correlation tritment is clearly visible. It is therefore necessary to parametrize them

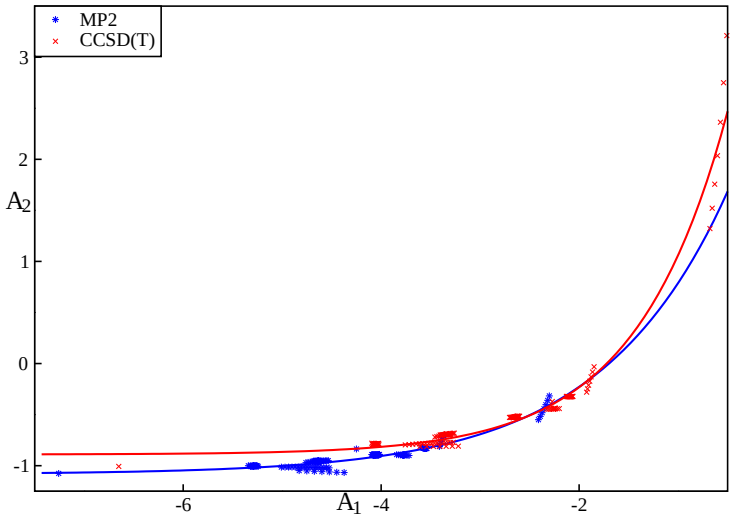

Figure 3.9: Correlation between A_1 and A_2 at the MP2 and the CCSD(T) levels with the cc-pVXZ basis set

Table 3.7: Exponential parameters for the MP2 and the CCSD(T) calculations.

Method	a	b	c
MP2	4.8278	0.8397	-1.0843
CCSD(T)	5.7873	1.0871	-0.8905

separately. The optimized parameters are presented in Table 3.7. A large difference in

parameter sets indicates that mixing of data from different computational methods would strongly affect the fitting curves. Also, parameter sets obtained for one computational method can not be used for the other. Therefore separation of the data sets and and separate optimization are the mandatory requirements for the success of the X5 method.

3.6.3 Standard MP2 vs RI-MP2 parameterization.

The effect of the DF (also called RI) approximation to the MP2 method on the extrapolation coefficients A_1 and A_2 was studied explictly. Computations using TURBOMOLE and Molpro show that the influence is negligible if appropriate fitting basis sets are chosen. The calculated A_1 and A_2 parameters for the RI-MP2 and the standard MP2 are

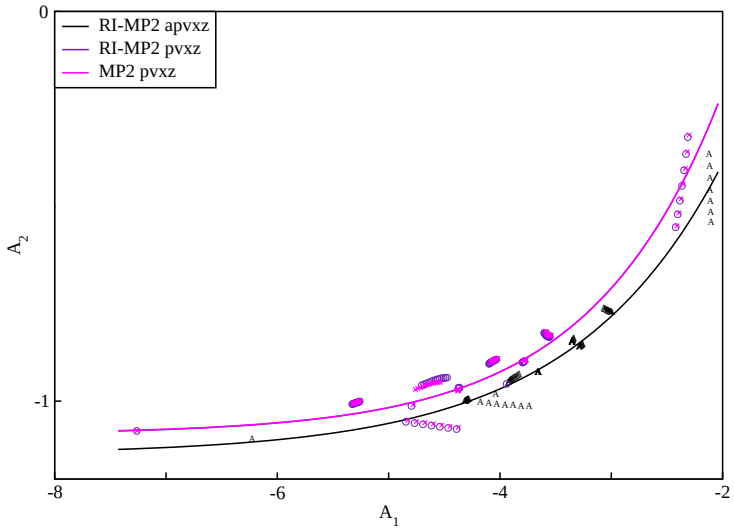

Figure 3.10: Comparison of correlation between A_1 and A_2 for the MP2 with the cc-pVXZ basis set and the RI-MP2 with the cc-pVXZ and the aug-cc-pVXZ basis set.

plotted in Fig. 3.10. The empirical parameters (A_1 and A_2) for the standard MP2 and the RI-MP2 are almost identical and their exponential fitting curves are practically indistinguishable when the energy points are calculated with the same family of basis sets. This investigation indicates that although the calculated correlation energy is slightly different for the standard and the RI-MP2 methods, mixing of data would not give any problem in extrapolation if same family of basis sets is used.

But when the calculation is performed with a particular computational method but employing a different family of basis sets, then the A_1 and A_2 parameters differ substantially. This is also shown in Fig. 3.10 where energy points calculated with the RI-MP2

Table 3.8: Exponential parameters for the standard MP2 and the RI-MP2 calculations.

	cc-pVXZ		aug-cc-pVXZ
Parameters	MP2	RI-MP2	RI-MP2
a	4.8278	4.8126	3.3145
b	0.8397	0.8503	0.7449
c	-1.0843	-1.0853	-1.1370

method employing cc-pVXZ as well as aug-cc-pVXZ basis sets are plotted. A large change in the exponential fitting curves yields different sets of exponential fitting parameters. The exponential fitting parameters for different methods and families of basis sets are presented in Table 3.8. It is obvious that exponential parameters optimize for a regular basis set can not be use for the corresponding augmented basis set and vice versa.

3.6.4 Performance of the X5 method

The extrapolated results at MP2 level for a set of atoms and molecules at their equilibrium geometries are presented in Table 3.9. The last column of Table 3.9 presents the basis set limit correlation energy obtained by the two point extrapolation from the QZ-5Z basis set results, which is taken as a reference and compare a to all other approximations. The two point extrapolated results from the DZ-TZ and the TZ-QZ are presented as well to prove the performance of the X5 extrapolation method with global parameters (exponential parameters from A_1/A_2 exponential fit). For comparison DZ, TZ, QZ and 5Z basis results are also presented here.

As seen from the Table 3.9, the DZ basis set always over-estimates the correlation energy far beyond the basis set limit value. The error reduces by almost 50% using the TZ basis set, but the computational cost increases by about one order of magnitude at the same time. The QZ basis computation, which is two orders of magnitude more expensive than for the DZ basis, is still far away from the basis set limit results and shows about 5% error. For the 5Z basis set, which is three to four order of magnitude more expensive than the DZ basis, the results improve very little over the QZ basis and still exhibit 2-3% error in the correlation energy.

The two point extrapolated correlation energy from the DZ-TZ basis sets is far from reaching the basis set limit for the correlation energy, which is to be expected and was first observed by Halkier et al.. The two point extrapolated correlation energy from DZ-TZ

basis sets is even worse than the 5Z basis result. The TZ-QZ extrapolation with the two point formula performs much better than the DZ-TZ combination and comes close to the basis set limit. The errors in basis set limit calculations are typically less than 0.5% of the correlation energy, but the computational cost increases nearly one order of magnitude over the DZ-TZ computational cost.

Table 3.9: Errors in the MP2 correlation energy (in mE_h) calculation with different size basis sets and with different extrapolation methods. All energies are compared to the two-point extrapolation from the QZ-5Z basis set results.

System	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	Two-point extrapolation		X5	E_∞^{corr}
					DZ-TZ	TZ-QZ		(QZ-5Z)
H_2	-7.8411	-2.5437	-1.1090	-0.5678	-0.3132	-0.0620	0.7105	-34.2245
He	-11.5523	-4.2431	-1.9027	-0.9742	-1.1656	-0.1948	0.2469	-37.3807
Ne	-133.8561	-55.0566	-25.8069	-13.2131	-21.8779	-4.4624	-6.6503	-319.3794
N_2	-113.6821	-46.3052	-21.2412	-10.8755	-17.9360	-2.9513	-4.9158	-420.0696
H_2O	-99.0221	-39.1818	-17.8455	-9.1369	-13.9859	-2.2758	-2.4221	-300.6311
CH_4	-57.4112	-20.5143	-8.7383	-4.4740	-4.9788	-0.1449	2.1513	-218.9506
CH_2O	-130.8455	-52.1052	-23.5980	-12.0822	-18.9514	-2.7955	-3.7353	-448.1022
C_2H_2	-89.2793	-34.7492	-15.3640	-7.8663	-11.7892	-1.2179	-1.2516	-343.5602
C_2H_4	-97.0844	-36.5635	-15.9259	-8.1541	-11.0811	-0.8660	0.6142	-372.4877
C_2H_6	-106.7132	-39.0910	-16.8650	-8.6349	-10.6185	-0.6459	2.4491	-409.4205
N_2H_2	-133.1312	-52.8413	-23.9402	-12.2574	-19.0351	-2.8502	-3.5195	-444.9168
N_2H_4	-149.8023	-58.5253	-26.5299	-13.5833	-20.0930	-3.1818	-2.4542	-496.6646
NH_3	-78.3940	-29.6005	-13.1363	-6.7258	-9.0559	-1.1218	0.3731	-264.6532
CH_3OH	-146.6560	-56.7313	-25.3899	-12.9996	-18.8683	-2.5191	-1.4908	-485.1429
CH_3NH_2	-126.7904	-47.7225	-20.9745	-10.7389	-14.4307	-1.4557	0.8487	-452.8129
CH_3COOH	-269.0912	-106.8199	-48.2015	-24.6792	-38.4951	-5.4259	-7.1372	-909.2622
C_6H_6	-271.3662	-104.7758	-45.9231	-23.5126	-34.6324	-2.9766	-2.4398	-1056.1373
NMA	-301.5884	-117.3309	-52.2985	-26.7768	-39.7489	-4.8424	-4.1422	-1069.4420

On the other hand, the X5 extrapolation method seems to be very efficient to estimate the basis set limit correlation energy from the DZ-TZ basis set as seen from the Table 3.9. The error for a set of atoms and molecules is typically less than 0.5% even for the large molecules like NMA, CH_3COOH, etc. It is also noticed that the extrapolated correlation energies using the X5 method are much better than the 5Z basis set results even though one pays 2 to 3 orders less in computational cost than for the 5Z basis set calculation. Especially for large molecules, where the 5Z basis set performance is prohibitively costly, one can achieve better results than with the 5Z basis set based only on the DZ-TZ basis set result. Over all samples X5 is at best as good as the TZ-QZ two-point extrapolation, but one has to pay almost one order less in computational cost than for the TZ-QZ two-point extrapolation.

The noticeable feature of the X5 method is that it is more efficient for large than for small molecules. For example, in H_2, where two point extrapolation from TZ-QZ includes only 0.06 mE_h of absolute error, the X5 includes 0.711 mE_h of the absolute error which is

$\sim$ 2% of the correlation energy. A similar situation is encountered for nitrogen, helium, neon, methane, and CH_2O in our test. For the large molecules like NMA, C_6H_6, etc. the X5 method is as good as the two point extrapolation from TZ-QZ basis set results and the error is only about 0.5% of the correlation energy. For such a large molecules the use of QZ basis sets is generally too costly, and X5 is an elegant method to achieve close to basis set limit accuracy.

Another noticeable feature of the X5 extrapolation method with the global parameter set (exponential A_1/A_2 parameter correlation for a set of atoms and molecules) is, that it gradually under-estimate the basis set limit energy when a large number of hydrogen

Table 3.10: The X5 method systematically under estimate the extrapolated correlation energy as the number of hydrogen increase in the molecule. All energies are in mE_h.

Molecule	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	Two-point extrapolation		X5	E_∞^{corr}
					2-3	3-4		(QZ-5Z)
H_2	-7.8411	-2.5437	-1.1090	-0.5678	-0.3132	-0.0620	0.7105	-34.2245
CH_4	-57.4112	-20.5143	-8.7383	-4.4740	-4.9788	-0.1449	2.1513	-218.9506
C_2H_2	-89.2793	-34.7492	-15.3640	-7.8663	-11.7892	-1.2179	-1.2516	-343.5602
C_2H_4	-97.0844	-36.5635	-15.9259	-8.1541	-11.0811	-0.8660	0.6142	-372.4877
C_2H_6	-106.7132	-39.0910	-16.8650	-8.6349	-10.6185	-0.6459	2.4491	-409.4205
N_2H_2	-133.1312	-52.8413	-23.9402	-12.2574	-19.0351	-2.8502	-3.5195	-444.9168
N_2H_4	-149.8023	-58.5253	-26.5299	-13.5833	-20.0930	-3.1818	-2.4542	-496.6646
NH_3	-78.3940	-29.6005	-13.1363	-6.7258	-9.0559	-1.1218	0.3731	-264.6532

atoms is present in the molecule. A systematic study has been carried out with small hydrocarbon and corresponding nitrogen compounds with different proportions of hydrogen atoms. Results are presented in Table 3.10. For acetylene (C_2H_2), the X5 method over-estimates the basis set limit energy by 1.252 mE_h whereas for ethylene (C_2H_4) it under-estimates by 0.614 mE_h. When two more hydrogen atoms are added to yield ethan (C_2H_6) X5 under-estimates the limit energy by 2.449 mE_h. A similar behavior is also observed for the N_2H_2, N_2H_4 and NH_3 molecules (see Table 3.10).

Although X5 is a very efficient method to extrapolate correlation energies to the basis set limit for a set of atoms and molecules at their equilibrium structure. Due to the geometry dependence of A_1 and A_2 parameters, the extrapolated result may deteriorate w.r.t the basis set limit. Especially for the highly correlated CCSD(T) method the hydrogen molecule shows a strong geometry dependency of the A_2 parameter (see Fig. 3.11). The A_2 parameters for the hydrogen molecule are in general also quite off set from all other molecules. A systematic study based on CCSD(T)/cc-pVXZ calculation has been carried and the results are presented in Table 3.11. The sixth column presents the error in extrapolation when the global parameter set is optimized including the hydrogen molecule

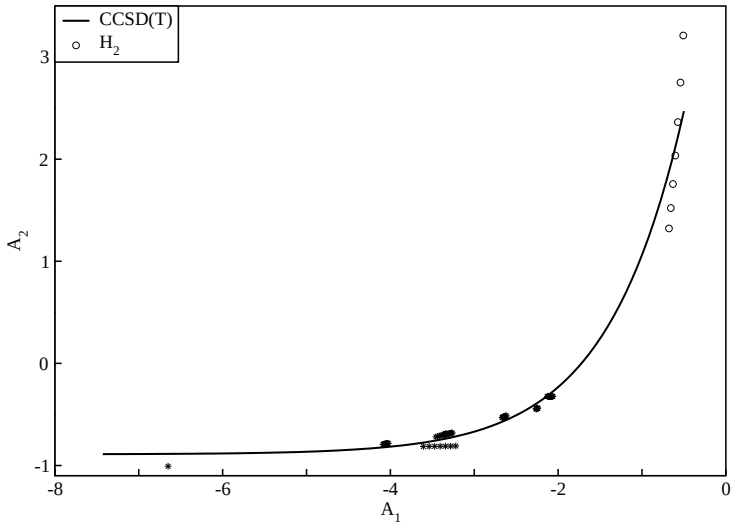

Figure 3.11: Geometry dependence of A_1 and A_1 parameters for the CCSD(T) calculation with cc-pVXZ basis set.

Table 3.11: Errors in the CCSD(T) correlation energy (in mE_h) calculation with different size of basis sets and with different extrapolation methods. All energies are compared to the two-point extrapolation from the QZ-5Z basis sets results. The terms "with H_2" and "without H_2" indicate that the parameterization has been performed including H_2 in the data set or excluding it respectively.

System	cc-pVQZ	cc-pV5Z	Two-point extrapolation		X5		E_∞^{corr}
			DZ-TZ	TZ-QZ	with H_2	without H_2	(QZ-5Z)
H_2	-0.5689	-0.2913	0.4399	0.1335	0.7252	0.8822	-40.9108
He	-1.2920	-0.6615	-0.3122	0.0345	0.0927	0.3155	-42.1886
Ne	-23.2957	-11.9274	-19.0158	-1.6752	-14.1073	-11.4071	-323.5157
N_2	-16.6637	-8.5318	-10.4284	0.0593	-6.2084	-3.8869	-430.0182
H_2O	-14.3258	-7.3348	-8.6631	0.2230	-4.9573	-2.9187	-309.2708
CH_4	-5.9629	-3.0530	-0.6922	1.3011	1.5118	2.7242	-240.9647
CH_2O	-18.5888	-9.5175	-11.1606	0.4858	-6.3014	-3.6283	-467.4811
C_2H_4	-11.2142	-5.7417	-3.5715	1.6584	0.0884	2.1017	-404.4420
C_2H_6	-11.8346	-6.0593	-2.6645	2.0289	1.4130	3.6562	-445.9864
CH_3OH	-19.8876	-10.1825	-10.3863	1.0304	-4.8614	-1.8220	-510.3177
CH_3NH_2	-15.5401	-7.9565	-5.8787	1.6893	-1.0624	1.5872	-483.4477

in the data set and the seventh column presents the error excluding hydrogen from the data set. In general, the result improves when hydrogen is excluded. This may seem a motivation to omit hydrogen from the data set. But for hydrogen rich molecules, the estimated basis set limit energy deteriorates if hydrogen is absent in the data set. For example, with hydrogen included X5 under-estimates the limit energy by 1.413 mE_h for C_2H_6, but when hydrogen is excluded, it under-estimates the basis set limit energy by

3.656 mE_h. A similar behaviors observed for the C_2H_4, CH_4 molecules. Since biomolecules are rich in hydrogen atoms, it is necessary to include the hydrogen in the data set.

3.6.5 The PES calculation

Now the obvious question arises how efficient is the X5 method to generate a PES? A systematic study has been carried out for a large number of molecules and their different vibrational and rotational motions, especially all the 30 normal modes of vibration have been studied for the NMA molecule. Although A_1 and A_2 parameters depend on the geometry of the molecule they are slowly varying parameters, and in general the estimation of basis set limit energies can be reasonably well achieved with a global parameter set for the PES calculation in the vicinity of the equilibrium structure. Table 3.12 presents the errors in the PES calculation based on the X5 extrapolation method along with the errors for the two-point extrapolation results and different basis sets for the C–O stretch vibration of methanol. Although the X5 method with a global parameter set always

Table 3.12: Error in the PES calculation along C–O vibrational band for methanol. All energies present here are in mE_h.

CO bond length in Å	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	Two-point extrapolation DZ-TZ	Two-point extrapolation TZ-QZ	X5	E_∞^{corr} (QZ-5Z)
1.265	-147.2149	-56.9331	-25.3590	-12.9838	-18.9197	-2.3184	-1.4733	-481.7091
1.285	-147.0813	-56.8810	-25.3519	-12.9802	-18.9020	-2.3441	-1.4713	-482.0564
1.305	-146.9677	-56.8377	-25.3484	-12.9784	-18.8882	-2.3697	-1.4711	-482.4293
1.325	-146.8734	-56.8024	-25.3483	-12.9783	-18.8777	-2.3953	-1.4721	-482.8269
1.345	-146.7974	-56.7750	-25.3513	-12.9799	-18.8708	-2.4205	-1.4745	-483.2479
1.365	-146.7387	-56.7546	-25.3572	-12.9829	-18.8666	-2.4455	-1.4778	-483.6912
1.385	-146.6962	-56.7409	-25.3657	-12.9872	-18.8650	-2.4703	-1.4817	-484.1557
1.405	-146.6690	-56.7333	-25.3767	-12.9929	-18.8656	-2.4948	-1.4861	-484.6400
1.425	-146.6560	-56.7313	-25.3899	-12.9996	-18.8682	-2.5191	-1.4908	-485.1429
1.445	-146.6562	-56.7345	-25.4052	-13.0074	-18.8727	-2.5432	-1.4959	-485.6634
1.465	-146.6684	-56.7425	-25.4224	-13.0163	-18.8790	-2.5672	-1.5013	-486.2001
1.485	-146.6917	-56.7549	-25.4413	-13.0260	-18.8868	-2.5909	-1.5071	-486.7519
1.505	-146.7250	-56.7715	-25.4619	-13.0365	-18.8963	-2.6144	-1.5133	-487.3178
1.525	-146.7672	-56.7918	-25.4839	-13.0478	-18.9074	-2.6377	-1.5202	-487.8966
1.545	-146.8173	-56.8156	-25.5074	-13.0598	-18.9201	-2.6608	-1.5278	-488.4873
1.565	-146.8745	-56.8427	-25.5320	-13.0724	-18.9346	-2.6837	-1.5365	-489.0890
1.585	-146.9375	-56.8727	-25.5578	-13.0856	-18.9507	-2.7063	-1.5463	-489.7007

contains an intrinsic error of ~ 1.5 mE_h, it is in general far better than the 5Z result, for which the error is about 13 mE_h. The two-point extrapolation from DZ-TZ basis sets is ~ 18.8 mE_h off for the PES calculation. The X5 method is also much more efficient than the two point extrapolation from the TZ-QZ basis sets (see Table 3.12).

3.6.6 The PES calculation with local parameters

The intrinsic error in the basis set limit energy calculation arises due to the global parameter where A_2 is an approximation only. It is then more desirable to generate a system base parameter to achieve higher accuracy. It was observed that the A_2 is a very slowly varying parameter with the geometry of the molecule, at least for near equilibrium points (see Table 3.13). This motivates an approximation where the A_2 parameter is frozen. We freeze A_2 at the equilibrium geometry (A_2^0) and use that parameter for all other geometries. Such a system based local method is named X5*, and requires, in addition to an estimate for $E_\infty^{corr}(X_0)$, only two energy points (e.g., DZ and TZ) to get the basis set limit energy for every PES point.

Table 3.13: Variation of the A_1 and A_2 parameters with the geometry change along the C=O stretch vibration of NMA. Negative sign indicates the contraction of the bond length.

C=O	A_1	A_2
0.7156	-4.349687	-0.972768
0.3577	-4.362339	-0.970373
0.0894	-4.368825	-0.967725
0.0000	-4.370373	-0.966505
-0.0428	-4.371245	-0.965885
-0.1712	-4.373243	-0.963576
-0.3423	-4.376430	-0.959431

The error in extrapolation for the PES of water is presented in Table 3.14. Table 3.15 presents the corresponding error for the NMA molecule, where the PES is calculated along the C=O stretch vibrational mode. As can be seen from Table 3.14, the calculated error is about 14 mE_h for the two-point extrapolation from the DZ-TZ basis sets, for the bending motion of water. The two point extrapolation from the TZ-QZ basis sets deviates about 2.3 mE_h from the exact result. Whereas the X5* method introduces only 0.12 mE_h maximum error along with the PES cut, which is only 0.04% of the total correlation energy.

A similar behavior is also observed for NMA. The error in the PES is about 40 mE_h when calculated with the two point extrapolation from the DZ-TZ basis sets. The 5Z basis set results are also far from the basis set limit results and one has to pay a high cost for that. The two point extrapolation from TZ-QZ is about 5 mE_h away from the limit

Table 3.14: Errors in the PES calculations for the bending mode of the water. All energies are in mE_h.

	Two-point extrapolation			X5		E_∞
Angle	DZ-TZ	DZ-QZ	TZ-QZ	A_1^0 & A_2^0	A_2^0	(QZ-5Z)
-5	13.904	6.205	2.252	-0.456	-0.103	-301.020
-4	13.925	6.217	2.259	-0.360	-0.084	-300.932
-3	13.948	6.230	2.266	-0.267	-0.064	-300.849
-2	13.971	6.242	2.273	-0.176	-0.043	-300.772
-1	13.994	6.255	2.281	-0.086	-0.022	-300.700
0	14.018	6.268	2.288	0.000	0.000	-300.633
1	14.043	6.282	2.296	0.084	0.022	-300.572
2	14.068	6.295	2.303	0.167	0.045	-300.516
3	14.094	6.309	2.311	0.248	0.069	-300.465
4	14.120	6.322	2.318	0.326	0.093	-300.420
5	14.146	6.336	2.325	0.403	0.118	-300.379

value. The X5 method with global parameters is as good as the two-point extrapolation from the TZ-QZ basis sets. The error reduces surprisingly when the X5 method is used with a fixed A_2^0 parameter, and the maximum error drops to only 0.644 mE_h.

Table 3.15: Errors in the PES calculations for the C=O stretch vibration in NMA with different approximations. The X5 extrapolation form (2,3) with $A_2 = 0$ is essentially the two-point extrapolation from the DZ-TZ basis sets. All energies are in mE_h.

Δq (C=O)	5Z	(3,4)	X5 extrapolation from (2, 3)				E_∞^{corr}
			$A_2 = 0$	X5	A_1^0 A_2^0	A_2^0	(QZ-5Z)
0.7156	-26.958	-5.101	-40.169	-4.532	1.040	-0.576	-1082.253
0.3577	-26.864	-4.999	-39.922	-4.308	0.495	-0.354	-1075.132
0.0894	-26.797	-4.907	-39.699	-4.067	0.126	-0.111	-1070.457
0.0000	-26.780	-4.880	-39.611	-3.958	0.000	0.000	-1069.055
-0.0428	-26.776	-4.873	-39.570	-3.903	-0.066	0.057	-1068.421
-0.1712	-26.767	-4.848	-39.428	-3.700	-0.276	0.267	-1066.659
-0.3423	-26.784	-4.826	-39.216	-3.339	-0.639	0.644	-1064.727

Another approximation has been applied to make the PES generation more economic. As observed, A_1 is also a slowly varying parameter with the geometry within a system at least for near equilibrium points (see Table 3.15). To a good approximation one can also freeze A_1 (A_1^0) along with A_2 for the system of interest in molecular dynamics. Then one needs only a single energy point in each configuration to get the basis set limit value.

That means one can obtain a high quality PES only performing a DZ basis set calculation at each point in the PES, which is a remarkable achievement.

Results for this approach are presented in Tables 3.14 and 3.15 as well for the bending mode of the water and the C=O stretch vibrational mode of the NMA, respectively. With this approximation the calculated basis set limits deteriorate slightly from the X5* method, but are still far better than the 5Z or the two-point extrapolation from the TZ-QZ results, which is achieved by only performing DZ basis set calculations.

3.6.7 The gradient and the Hessians

The convergence behavior of the gradient and the Hessian have been studied explicitly for the water molecule using standard MP2 and the CCSD(T) calculations (no DF) employing cc-pVXZ basis sets. Table 3.16 and 3.17 present the average error in gradient and Hessian extrapolation respectively. The average error is defined as,

$$\Delta_{av} = \sum \left(||H^{\infty} - H^{limit}|| \right) / n^2,$$

where H^{∞} is the extrapolated gradient or Hessian calculated with two-point extrapolation from the QZ-5Z basis set, H^{limit} is the approximate basis set limit value and n^2 is the number of elements in the gradient or the Hessian matrix. The two-point extrapolation results from different combinations of energy points are also included in the table.

Table 3.16: Errors in the gradient extrapolation from different combinations of basis sets with the two-point and the X5 extrapolation methods. X5* is the X5 extrapolation with A_2^0 and X5** is the X5 extrapolation with A_1^0 and A_2^0. Gradients are calculated at equilibrium, at 0.04 Å symmetric O–H stretch and at -5° H-O-H bending.

	equilibrium		δr =0.04Å		$\delta\phi = -5^\circ$	
	MP2	CCSD(T)	MP2	CCSD(T)	MP2	CCSD(T)
23	0.326	0.321	0.319	0.312	0.275	0.267
24	0.255	0.246	0.229	0.227	0.261	0.255
34	0.347	0.319	0.336	0.306	0.355	0.328
X5*	0.304	0.289	0.287	0.276	0.252	0.231
X5**	4.257	4.727	4.411	4.824	4.434	4.892

As seen from the Table 3.16, the gradient extrapolation with the X5* method from the DZ-TZ basis sets gives a similar error as the two-point extrapolation from the DZ-TZ basis sets, which means the A_2 parameter does not have any influence on the extrapolation.

Table 3.17: Errors in the Hessian extrapolation from different combinations of basis sets with the two-point and the X5 extrapolation methods. X5* is the X5 extrapolation with A_2^0 and X5** is the X5 extrapolation with A_1^0 and A_2^0. Hessians are calculated at equilibrium, at 0.04 Å symmetric O–H stretch and at $-5°$ H-O-H bending.

	equilibrium		$\delta r = 0.04$Å		$\delta\phi = -5°$	
	MP2	CCSD(T)	MP2	CCSD(T)	MP2	CCSD(T)
23	2.859	0.238	0.281	0.222	0.330	0.238
24	0.995	0.121	0.163	0.145	0.179	0.170
34	0.113	0.107	0.147	0.137	0.169	0.162
X5*	4.155	0.205	0.201	0.232	0.222	0.200
X5**	11.227	2.145	2.647	2.451	2.363	2.087

A large increase of error is observed in gradient extrapolation when A_1 and A_2 are both constant i.e., for A_1^0 and A_2^0. Since the influence of A_2 is negligible, the error is due to the fixed A_1, which means A_1 is strongly depended on the local geometry of the molecule and choosing A_1 as a constant is a poor approximation for gradient extrapolation. The two-point extrapolation with TZ-QZ also gives a similar error in the basis set limit gradient calculation for both the MP2 and CCSD(T) level, which indicates that gradient extrapolation may not be necessary.

A similar result is observed in case of the basis set limit Hessians calculation. It also indicates that the extrapolation is not necessary for the Hessian calculations.

Chapter 4

Vibrational Spectroscopy of Methyl benzoate

4.1 Introduction

Vibrational spectroscopy is among the foremost experimental tools in the exploration of molecular potential-energy surfaces (PES). Its application to biological systems has so far been severely handicapped, both by experimental difficulties and by the unavailability of adequate computational tools for quantitative interpretation. Recent success in the experimental realization of coherent multidimensional infra-red (IR) spectroscopy provides a new powerful tool to study structure and dynamics of biomolecules with a temporal resolution down to the sub-picosecond regime,[8,119,120] prepared by any of the ultrafast initiation techniques (photo-switches, pH jump, local electric field, etc.) developed recently.[121] Multidimensional IR spectroscopy has the potential to disentangle the congested vibrational spectra of biomolecules to some extent similar to multidimensional NMR[3–6] but with significantly higher temporal resolution. In nonlinear multidimensional spectra the structural and dynamical information is typically available in terms of diagonal and cross-peak shapes, locations and intensities and their respective temporal evolution. The interpretation of this data in terms of a dynamical model of the biomolecule under investigation requires extensive theoretical modeling.

The calculation of vibrational spectra of biomolecules within the harmonic approximation is very useful, but often has limited significance since many biological molecules are "floppy" and subject to strong anharmonic effects. Anharmonic effects are even much larger in weakly bound complexes of biological molecules, e.g., their hydrogen-bonded complexes with water.[2] Also, much of the interest in the PES is away from the equilib-

rium configuration, where the harmonic approximation is even less applicable. The main problem of anharmonic spectroscopic calculations is that different vibrational modes are not mutually separable like in the harmonic approximation. Therefore one has to face the task of calculating wavefunctions and energy levels for systems of many coupled degrees of freedom. Several attempts have been made to overcome this problem. Among others, the discrete variable representation (DVR),[13–16] diffusion quantum Monte Carlo (DQMC),[17–20] and vibrational self-consistent-force field (VSCF),[21–23] methods proved their applicability to study anharmonic effects in systems with different size. The VSCF method is most successful among them to effectively handle large molecular systems.

Most reactions in chemistry and biology occur in condensed phase under thermal conditions. IR absorption spectra of peptides and proteins are dominated by vibrational bands that can be described approximately in terms of oscillators localized in each repetitive unit and their mutual couplings. The most extensively studied bands are amide-A and amide-B in the region 3000–3500 cm^{-1} and amide-I and amide-II between 1500-1700 cm^{-1}, which are spectrally well separated from the remaining spectrum and exhibit a strong dependence on the structural motifs present in the investigated biomolecules.[7] The amide-I vibrational mode, which involves mainly the C=O stretch coordinate has experimentally been the most important mode due to its large transition-dipole moment and because it appears to be mostly decoupled from the remaining vibrational modes in proteins. A detail understanding of these modes are then necessary to understand the structure and dynamics of the protein and peptide. But due to the complex structure of protein a detail theoretical understanding of these modes are complicated. A simple small molecule is then necessary where these important modes can be studied in more details.

Methyl benzoate has a planar configuration, except for the two hydrogen atoms at the methyl group which are symmetrically out-of-plane with respect to the rest of the molecule (see Fig. 4.1). The methyl carboxylate group is co-planar with the phenyl ring. In Methyl benzoate, the C=O double bond in the carboxylic ester group behaves as a local oscillator similar to the amide-I band in proteins and provides a convenient mode to study the amide-I band structure of proteins. Another potentially interesting coupling in proteins is the coupling between the amide-I and the β-hydrogen located in the sidechain. Methyl benzoate also provides a similar structure where the ortho hydrogen in the phenyl ring provides the counter-part of a β-hydrogen in protein sidechains. The structural similarities of these two bands in Methyl benzoate and proteins are depicted in Fig. 4.2.

Isotopic substitution is a valuable tool in the identification of molecular structure

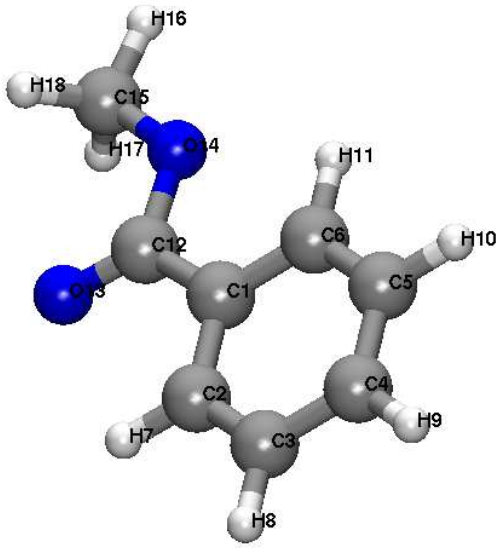

Figure 4.1: Structure of Methyl benzoate.

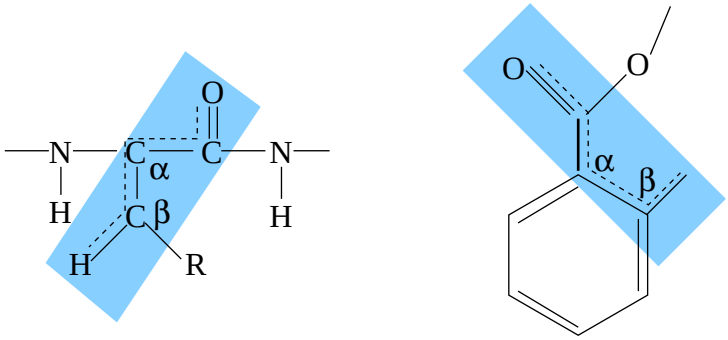

Figure 4.2: Structural similarity of protein backbone and Methyl benzoate.

and dynamics. The C–H or C–D stretch vibrations are particularly important structural probes, because they are very localized, specific and abundant. Especially the C–D stretching frequency is an excellent structural probe since it is usually spectrally isolated ($\sim$2200 cm^{-1}) even in the spectrum of large proteins. Substitution of β-hydrogen by deuterium provides a probe to obtain a better understanding of the structure and dynamics of the protein backbone.

4.2 Theoretical Background

4.2.1 Normal-mode analysis

Interest in the vibrational motion of molecules is pervasive in chemistry.[122] The traditional view of this motion is based on the Born-Oppenheimer approximation[123] which separates the motion of the electrons from that of the nuclei. The electronic motion produces an effective potential which holds atoms together and governs their vibrational motion. The complexity of this potential gives rise to the richness of much of chemistry. Thus, a theoretical picture of the vibrations of molecules is at the heart of many chemical questions.

The traditional view of vibrational motion is based on the harmonic approximation to the full nuclear potential. This very simple approximation gives rise to an extraordinarily simple and useful picture of vibrational motion, namely, that of independent vibrational modes, usually called "normal modes".[124] A normal mode in an oscillating system is the frequency at which a deformable structure will oscillate when weakly disturbed. The normal mode frequency is also known as natural or fundamental frequency of the system. A many-body system of N bodies has 3N degrees of freedom. Out of the total 3N degrees of freedom three are purely translational and three are purely rotational for any nonlinear system. These do not contribute to the oscillation of the localized oscillator of the system. Thus, there are 3N-6 remaining normal modes in an N-body nonlinear system. For a linear N-body system the number of normal modes is 3N-5.

Let us consider that such a many-body system is vibrationally excited. If we express the displacement of each body by mass-weighted Cartesian displacement coordinates (i.e., $q_1 = \sqrt{m_1}\Delta x_1$, $q_2 = \sqrt{m_1}\Delta y_1$, $q_3 = \sqrt{m_1}\Delta z_1$, $q_4 = \sqrt{m_2}\Delta x_2$, and so on), the kinetic energy of the system is defined by,[124]

$$2T = \sum_{i=1}^{3N} \dot{q}_i^2, \tag{4.1}$$

and the potential energy V to the second order in the vicinity of a minimum by

$$2V = \sum_{i,j=1}^{3N} f_{ij} q_i q_j. \tag{4.2}$$

where $f_{ij} = \left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_0$, with $f_{ij} = f_{ji}$, and $\dot{q}_i$ is the time derivative of q_i.

Since T is a function of velocity only and V is a function of the coordinates only, Newton's equations of motion can be written as,

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} + \frac{\partial V}{\partial q_j} = 0 \qquad j = 1, 2, \cdots, 3N. \tag{4.3}$$

Substituting the expression for T and V, in Eq. (4.3) one gets[124]

$$\ddot{q}_j + \sum_{i,j=1}^{3N} f_{ij} q_i = 0. \tag{4.4}$$

This is a set of $3N$ simultaneous second-order linear differential equations. One possible solution is,[124]

$$q_i = A_i \cos(\lambda^{1/2} t + \epsilon), \tag{4.5}$$

where A_i is the amplitude of oscillation, and ϵ is the phase. Now Eq. (4.4) becomes,

$$\sum_{i=1}^{3N} (f_{ij} - \delta_{ij}\lambda) A_i = 0 \qquad j = 1, 2, \cdots, 3N, \tag{4.6}$$

where δ_{ij} is the Kronecker delta function.

Only for special values of λ, Eq. (4.6) has non-vanishing solutions, for all other values of λ, the solution is trivial and corresponds to no vibration.

The special values of λ are those for which

$$\begin{vmatrix} f_{11} - \lambda & f_{12} & \cdots & f_{1,3N} \\ f_{21} & f_{22} - \lambda & \cdots & f_{2,3N} \\ \vdots & \vdots & & \vdots \\ f_{3N,1} & f_{3N,2} & \cdots & f_{3N,3N} - \lambda \end{vmatrix} = 0. \tag{4.7}$$

The elements of this determinant are the coefficients of the unknown amplitudes A_i. For a particular value of λ, say λ_k, the coefficient of the unknown variable becomes fixed, say A_{ik}, which is the amplitude of vibration for a particular normal mode of vibration. Each normal mode is a fundamental property of the system and follows the same characteristics. These are:[124]

1. Each normal mode acts like a simple harmonic oscillator.
2. A normal mode is a concerted motion of many bodies.
3. The center of mass does not move.
4. All bodies pass through their equilibrium positions at the same time.
5. Normal modes are independent; they do not interact.

4.2.2 Different approaches for normal-mode analysis

Normal-mode analysis provides the fundamental mechanical properties of many-body systems. Therefore, it is an excellent tool to study molecular structure and properties. But as the number of atoms increases in the molecule, the solution of Eq. (4.6) becomes complicated. We therefore need an efficient method for normal-mode analysis to understand the molecular structure and dynamics.

Several approaches exist to achieve this goal. Among them grid methods are successful for small polyatomic systems and clusters.[125–130] Here, a widely used technique is the discrete variable representation (DVR) method.[13–16] For small systems like H_3^+ or Ar_3 it has been established as a powerful method for vibrational analysis,[16] but it also brings out the enormous difficulties encountered owing to both computer memory as well as speed limitations for larger systems. A method that scales very well with system size is the diffusion quantum Monte Carlo (DQMC) algorithm and related techniques such as the vibrational quantum Monte Carlo approach.[17–20] The DQMC is a very effective method for obtaining the vibrational ground state of large anharmonic systems, including the extreme case of quantum clusters.[131–134] The DQMC method is, however, limited to calculation of the ground state. In specific cases it is possible to employ this method also for excited states[135] and applications to spectroscopy have been made on this basis.[136] However, a general extension of DQMC for excited vibrational states is not yet at hand.

The approach that seems at present to provide the most effective tool for spectroscopy of large polyatomic systems is the vibrational self consistent field (VSCF) approximation and its generalizations.[21–23]

4.2.3 The VSCF Approximation

The basic idea behind the VSCF approach comes from the Hartree approximation.[25] First a minimum energy configuration of the system is considered and all the normal-mode coordinates are computed for that configuration. These are then used to express the vibrational Hamiltonian operator of the system. For a system with the total angular momentum $J = 0$, and with all rotational coupling effects neglected, the vibrational Schrödinger equation can be written as,[124]

$$\left[-\frac{1}{2}\sum_{j=1}^{N}\frac{\partial^2}{\partial q_j^2} + V(q_1, \cdots, q_N)\right]\Psi_n(q_1, \cdots, q_N) = E_n\Psi_n(q_1, \cdots, q_N), \tag{4.8}$$

where N is the total number of vibrational normal modes, and V is the potential function of the system. Although normal modes are chosen to describe the vibrations of the system, the harmonic approximation is not made to describe the potential function $V(q_1, \cdots, q_N)$. The VSCF approximation is based on the Hartree product ansatz for the wave function:

$$\Psi_n(q_1, \cdots, q_N) = \prod_{j=1}^{N} \psi_j^{(n)}(q_j), \tag{4.9}$$

which, using the variational principle, leads to the single-mode VSCF equations[23,122]

$$\left[-\frac{1}{2} \sum_{j=1}^{N} \frac{\partial^2}{\partial q_j^2} + \overline{V}_j^{(n)}(q_j) \right] \psi_j^{(n)}(q_j) = \varepsilon_j^{(n)} \psi_j^{(n)}(q_j), \tag{4.10}$$

where $\overline{V}_j^{(n)}(q_j)$, the one-dimensional mean-field potential for the mode q_j, is given by:

$$\overline{V}_j^{(n)}(q_j) = \left\langle \prod_{l \neq j}^{N} \psi_l^{(n)}(q_l) | V(q_1, \cdots, q_N) | \prod_{l \neq j}^{N} \psi_l^{(n)}(q_l) \right\rangle. \tag{4.11}$$

For single mode wavefunctions, energies in Eq. (4.10), and one-dimensional effective potentials Eq. (4.11) must be solved iteratively until self-consistency is achieved. Both ground and excited SCF states of the total system can be obtained using Eq. (4.10). The VSCF approximation for the total energy is given by:

$$E_n = \sum_{j=1}^{N} \varepsilon_j^{(n)} - (N-1) \left\langle \prod_{j=1}^{N} \psi_j^{(n)}(q_j) \middle| V(q_1, \cdots, q_N) \middle| \prod_{j=1}^{N} \psi_j^{(n)}(q_j) \right\rangle. \tag{4.12}$$

The computational efficiency of the VSCF method depends upon the calculation of the single mode effective potentials, Eq. (4.11), which involves in general the evaluation of multidimensional integrals. The main computational difficulty in solving Eq. (4.11) and hence Eq. (4.10) for large systems is the need to evaluate the multidimensional integrals involving the potential function $V(q_1, \cdots, q_N)$. Below, we describe some approaches to evaluate the integrals efficiently.

Power-series expansion of the potential

A very advantageous approach, in certain cases, is to expand the potential function in powers of the normal modes.[24] The expansion can be written as,

$$V(q_1, \cdots, q_N) = \sum_{m_1, \cdots, m_N} V_{m_1, \cdots, m_N} (q_1)^{m_1} \cdots (q_N)^{m_N}, \tag{4.13}$$

which requires the evaluation of one-dimensional integrals only to obtain the one-dimensional effective potentials $\overline{V}_j^{(n)}(q_j)$. Then the one-dimensional potential function is written as,[24]

$$\overline{V}_j^{(n)}(q_j) = \sum_{m_1,\cdots,m_N} V_{m_1,\cdots,m_N} \prod_{l\neq j}^{N} F_{m_l}^{(l),(n)}(q_j)^{m_j}, \tag{4.14}$$

where,

$$F_{m_l}^{(l),(n)} = \left\langle \psi_l^{(n)}(q_l) \middle| (q_l)^{m_l} \middle| \psi_l^{(n)}(q_l) \right\rangle. \tag{4.15}$$

Although the number of one-dimensional integrals increases with the system size, if the order of the expansion is modest the computational task is feasible even for thousands of modes.[25] Fourth order polynomial approximations for the PES have been successfully used.[24,137–139] Sixth order and fourth order polynomial representations for the cluster $(Ar)_{13}$ show similar results, which supports the validity of this approach.[24,25] But when higher order expansions are required, the method quickly looses its advantages.

The coefficients $V_{m_1,\cdots,m_N}$ are typically obtained by differentiation of the potential energy function at the equilibrium configuration. But higher order numerical derivatives calculated near the equilibrium configuration provide a poor approximation for anharmonic systems and the polynomial expansion method breaks down completely. The power series for the potential either diverges or requires extremely high powers for acceptable accuracy. Jung and Gerber's study on the water dimer illustrates this difficulty.[24] Figure 4.3 shows the VSCF potential for mode 12, a soft torsional mode of $(H_2O)_2$. The failure of the power-series expansion is obvious from the flat-bottom shape of the PES.

Higher order coupling

Carter et al. first suggested that the VSCF PES can be expressed in terms of a hierarchical expansion[26]

$$\begin{aligned} V(q_1,\cdots,q_N) &= \sum_{j}^{N} V_j^{(1)}(q_j) + \sum_{i<j} V_{i,j}^{(2)}(q_i,q_j) \\ &+ \sum_{i<j<k} V_{i,j,k}^{(3)}(q_i,q_j,q_k) + \cdots \\ &+ \sum_{i<j\cdots<r<s} V_{i,j,\cdots,r,s}^{(n)}(q_i,q_j,\cdots,q_r,q_s) + \cdots . \end{aligned} \tag{4.16}$$

where $V_j^{(1)}(q_j)$ is the diagonal potential, $V_{i,j}^{(2)}(q_i,q_j)$ is the pairwise potential, $V_{i,j,k}^{(3)}(q_i,q_j,q_k)$ is the triple coupling and so on.

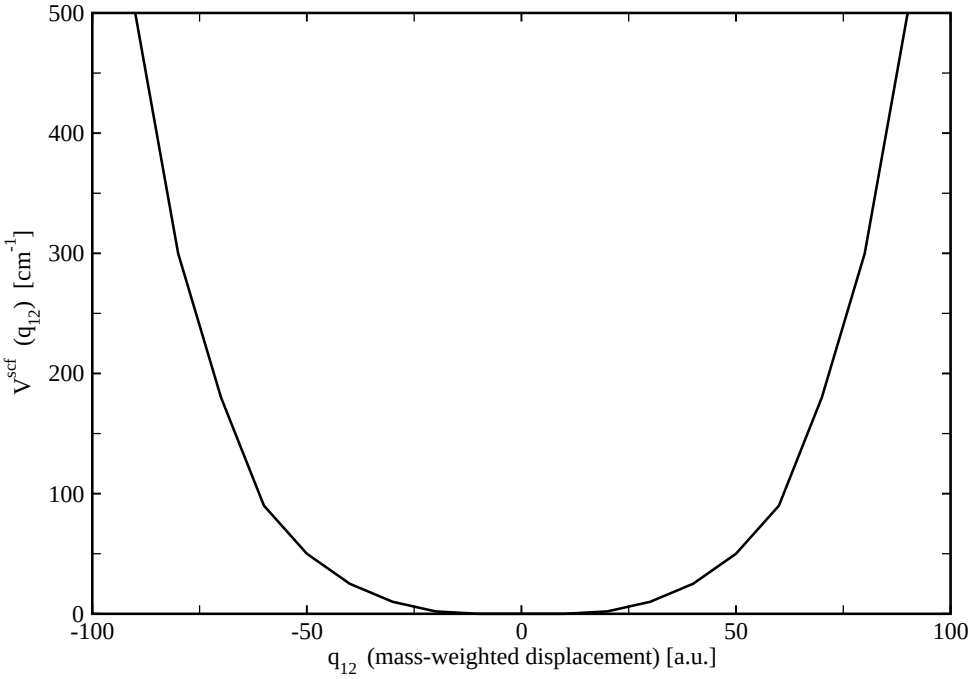

Figure 4.3: The VSCF potential for a torsional mode of $(H_2O)_2$. The normal mode here is q_{12}; Obtained from Ref. [24].

Usually, it is possible to represent the potential function $V(q_1, \cdots, q_N)$ with sufficient accuracy by assuming only pairwise interactions between normal modes.[24,140] Accounting for the pairwise interactions, the potential function can be approximated by

$$V(q_1, \cdots, q_N) = V_0 + \sum_{j}^{N} V_j^{(1)}(q_j) + \sum_{j} \sum_{i>j} V_{i,j}^{(2)}(q_i, q_j), \tag{4.17}$$

where $V_j^{(1)}(q_j)$ is the diagonal potential function defined by

$$V_j^{(1)}(q_j) = V(0, \cdots, q_j, \cdots, 0) - V_0, \tag{4.18}$$

and $V_{i,j}^{(2)}(q_i, q_j)$ are the pairwise interactions defined by

$$V_{i,j}^{(2)}(q_i, q_j) = V(0, \cdots, q_i, \cdots, q_j, \cdots 0) - V_j^{(1)}(q_j) - V_i^{(1)}(q_i) - V_0. \tag{4.19}$$

Including pairwise couplings, the effective single-mode VSCF potential $V_j^{(2)(n)}(q_j)$ reads

$$V_j^{(2)(n)}(q_j) = \sum_{i \neq j} \left\langle \psi_i^{(n)}(q_i) \middle| V_{i,j}^{(2)}(q_i, q_j) \middle| \psi_i^{(n)}(q_i) \right\rangle \tag{4.20}$$

The computation of these integrals can be efficiently done by using a grid representation for $V^{(2)}_{i,j}(q_i, q_j)$, combined with a routine technique for solving the one-dimensional Schrödinger equation for the $\psi_i(q_i)$. Grid size takes crucial role in this approach.

For typical molecular problems, quadruple ($n = 4$) interaction potentials have a negligible influence on the vibrational spectrum.[26] Most generally used models for the description of multidimensional vibrational spectroscopy in larger molecules do not even include any three mode coupling,[7] i.e., they assume $V^{(n)} = 0$ for all $n > 2$. But in a recent detailed analysis of the effect of triple contributions on high level VSCF computations for small molecules it was observed that some triple contributions may have a significant effect.[27] It was shown that the triple couplings $V^{(3)}_{ijk}$ which contribute most to the anharmonic frequencies involve modes which also exhibit large mutual pair-couplings $V^{(2)}_{ij}, V^{(2)}_{ik}$, and $V^{(2)}_{jk}$. Therefore it is justified to first analyze the pair couplings and then to include triple couplings where it is required.

4.2.4 Configuration interaction VSCF

Configuration interaction VSCF (CI-VSCF) is a very efficient and accurate method to solve the vibrational problem of polyatomic molecules proposed first by Bowman and coworkers.[141] In the CI-VSCF method,[142,143] the eigenfunctions of the vibrational Schrödinger equation Eq. (4.8) are expanded in a basis of VSCF states:

$$\psi_n(q_1, \cdots, q_N) = \sum_m C^n_m \prod_{j=1}^{N} \phi^{(m)}_j(q_j), \tag{4.21}$$

where $\phi^{(m)}_j(q_j)$ is a modal function of the VSCF Hamiltonian for mode q_j, and m is the excitation level of the state computed. It is to be noted that different VSCF states are not orthogonal. Then the VSCF energies E_n are obtained by solving the secular equation:

$$\det\left[\mathbf{H} - \mathbf{E}\mathbf{S}\right] = \mathbf{0}, \tag{4.22}$$

where $\mathbf{H}$ is the full vibrational Hamiltonian matrix and $\mathbf{S}$ is the overlap matrix. The matrix elements are defined as

$$\mathbf{H}_{m,m'} = \left\langle \prod_{j=1}^{N} \phi^{(m)}_j(q_j) \middle| H \middle| \prod_{j=1}^{N} \phi^{(m')}_j(q_j) \right\rangle \tag{4.23}$$

$$\mathbf{S}_{m,m'} = \left\langle \prod_{j=1}^{N} \phi^{(m)}_j(q_j) \middle| \prod_{j=1}^{N} \phi^{(m')}_j(q_j) \right\rangle. \tag{4.24}$$

Alternatively, it is possible to express the CI expansion by virtual VSCF states obtained from the solutions of the VSCF Hamiltonians for a fixed vibrational effective potential. The virtual VSCF states are orthogonal, and **S** becomes the unit matrix. For a sufficiently large basis, CI-VSCF becomes a rigorous method to solve the vibrational eigenvalue problem for a coupled mode system with spectroscopic accuracy. Applying the CI-VSCF method for small polyatomic molecules, an accuracy of 1 cm^{-1} is achieved.[26,137,144,145] It is also possible to incorporate rotational coupling effects (including Coriolis coupling) in CI-VSCF treatments of the vibrational state calculations.[144,145] With these generalization, this method can be applied to the full Watson Hamiltonian[146] of the vibrational problem, which is essential for small polyatomic systems, for quantitative comparisons with the experimental data, but plays a lesser role in larger systems. Although the CI-VSCF method is a very powerful tool for small systems it also has some disadvantages. First of all for some systems and states the CI-VSCF expansion may converge slowly which makes the computation costly. Secondly the computational effort in CI-VSCF is expected to scale as N^5 or worse, with the number of normal modes N, which makes the method unaffordable very soon with increasing system size.

4.2.5 Correlation Corrected VSCF

The correlation corrected VSCF (CC-VSCF) method was developed[24,138] based on the assumption that at least for low-lying vibrational excited states VSCF already yields results of very good accuracy. This suggests that correlation effects for such states are relatively small, and can be treated by perturbation theory. This approach is analogous to the very familiar Møller-Plesset method for electronic structure calculations, where perturbation theory is used to introduce correlation effects and improve the Hartree-Fock approximation. In the CC-VSCF approach, the full Hamiltonian is written as:

$$H = H^{\mathrm{SCF},(n)} + \Delta V(q_1, \cdots, q_N), \tag{4.25}$$

where,

$$H^{\mathrm{SCF},(n)} = \sum_j \overline{H}_j^{(n)}(q_j), \tag{4.26}$$

with $\overline{H}_j^{(n)}(q_j)$, the SCF Hamiltonian for the mode q_j in the state (n),

$$\overline{H}_j^{(n)}(q_j) \equiv -\frac{1}{2}\frac{\partial^2}{\partial q_j^2} + \overline{V}_j^{(n)}(q_j). \tag{4.27}$$

The potential energy change ΔV in the perturbed system in Eq. (4.25) is defined as:

$$\Delta V(q_1, \cdots, q_N) = V(q_1, \cdots, q_N) - \sum_{j=1}^{N} \overline{V}_j^{(n)}(q_j). \tag{4.28}$$

The correlation effects are all included in ΔV. Assuming correlation effects are small and the SCF part of the Hamiltonian dominates, the effect of ΔV can be calculated by perturbation theory. Using second order perturbation theory, the energy expression to the second order in ΔV is[24]

$$E_n^{\mathrm{MP2}} = E_n^{\mathrm{VSCF}} + \sum_{m \neq n} \frac{\left| \left\langle \prod_{j=1}^{N} \psi_j^{(n)}(q_j) \middle| \Delta V \middle| \prod_{j=1}^{N} \psi_j^{(n)}(q_j) \right\rangle \right|^2}{E_n^{(0)} - E_m^{(0)}}. \tag{4.29}$$

The label MP2 comes since the correlation correction term is calculated by second order Møller-Plesset perturbation theory, whereas E_n^{VSCF} is the VSCF approximation to the total energy, given by Eq. (4.12). The wave functions $\psi_j^{(m)}$ and the energies $E_m^{(0)}$ in Eq. (4.29) are calculated from the VSCF Hamiltonian $H^{\mathrm{SCF},(n)}$, corresponding to the state n.

$$E_m^{(0)} = \sum_{j=1}^{N} \varepsilon_j^{(n),m}. \tag{4.30}$$

The superscript (n) in Eq. (4.30) indicates that the levels are calculated from the Hamiltonian $\overline{H}_j^{(n)}(q_j)$, while m is the level of the state computed. $\varepsilon_j^{(n),m}$ is the m^{th} SCF energy level of the j^{th} mode, computed from the Hamiltonian $\overline{H}_j^{(n)}(q_j)$. Higher order corrections to the energy can be computed with a reasonable effort, but these corrections are probably very small[25] and sometimes diverge. The CC-VSCF method is successful even for medium and large polyatomic systems. But a serious problem arises when two or more states are nearly degenerate, i.e., $E_m^{(0)} \sim E_n^{(0)}$. This problem can be resolved by carrying out a small CI calculation in the subspace of nearly degenerate SCF states only. This is essentially equivalent to applying the degenerate perturbation theory to this subspace.[147] Then, the split levels and the corresponding states could be coupled to the other states with which they are not degenerate, by ordinary perturbation theory. The CC-VSCF method is computationally very efficient even for many hundreds of coupled modes. An interesting point concerning this is that the correlation effects are generally smaller for larger systems, at least on the average. The percentage of correlation correction is defined

as[24]

$$R = \frac{1}{N}\left[\sum_{j=1}^{N} r(j)\right] \times 100, \tag{4.31}$$

where $r(j)$ is the ratio of the MP2 correction to the j^{th} fundamental excitation and N is the number of normal modes. Figure 4.4 shows the dependence of R with cluster size for

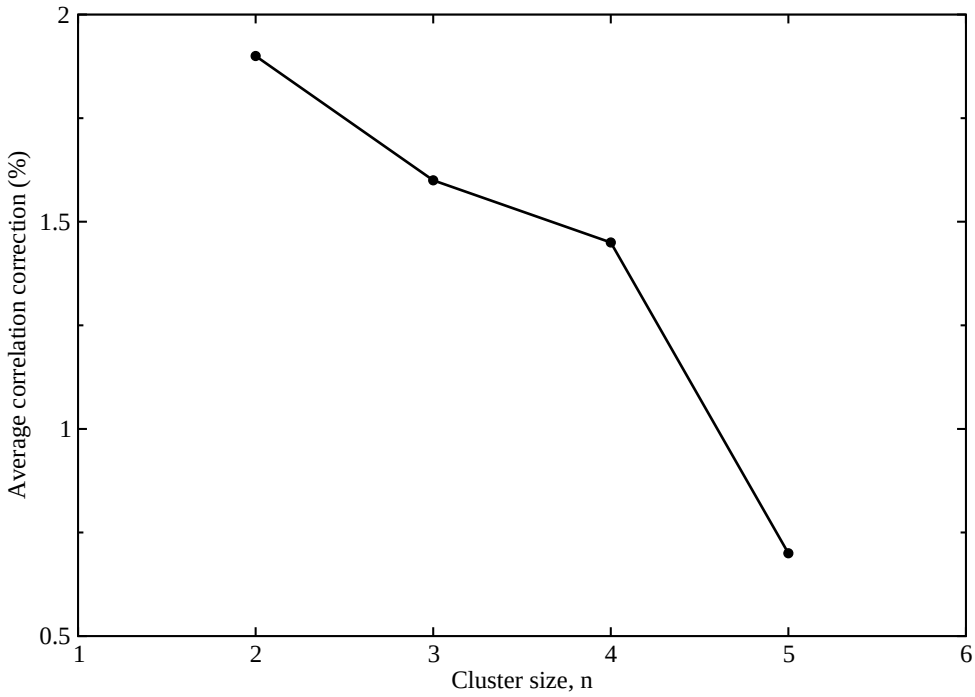

Figure 4.4: Average correlation-correction percentage versus cluster size n for $(H_2O)_n$. Obtained from Ref. 24.

water clusters $(H_2O)_n$.[24] The plot indicates that the correlation errors decrease with the cluster size. This suggests that the Møller-Plesset perturbation series should on average converge better for larger systems.

Although historically this method has been introduced as CC-VSCF it is more appropriately called vibrational MP2 (V-MP2) and, after correction for the degeneracy problem, the best way to call it is vibrational corrected MP2 (VC-MP2).

4.2.6 Anharmonicity

As discussed in section 4.2.1, the force constant k can be calculated from the curvature of the potential energy curve near equilibrium $(\frac{\partial^2 V}{\partial q^2})_0$. In practice, however, for larger

displacements from equilibrium, the potential energy does not follow a harmonic behavior, making the harmonic-oscillator approximation poor. Instead of following a parabolic path, the potential energy follows a higher order polynomial curve. This potential energy surface is called the anharmonic PES. In normal mode analysis, higher order terms than quadratic terms are neglected, which contain important information on the system. The anharmonic potential is typically less confining than the parabolic potential and energy levels can be expected to converge[148] as the vibrational quantum number increases. Not only do the the energy levels change, but also the wave functions change from those of the harmonic oscillator, therefore the selection rules, which are based on integrals over the wave functions are also modified.

Therefore it is important to know about the anharmonicity for high resolution spectroscopy. Even in low resolution work they are significant when the displacement of the nuclei reaches values far from the equilibrium. This occurs when the vibration is excited to high quantum numbers. Then a more accurate expression for the molecular vibrational energy that allows for the anharmonicity of the vibration is[149]

$$E_{\text{vib}} = (\nu + \frac{1}{2})\hbar\omega - (\nu + \frac{1}{2})^2\hbar\omega x, \tag{4.32}$$

where ωx is the anharmonicity constant. Two types of anharmonicities are often discussed, one is diagonal anharmonicity and the other is off-diagonal anharmonicity. A pictorial representation of these two is given in Fig. 4.5 where the diagonal anharmonicity is defined

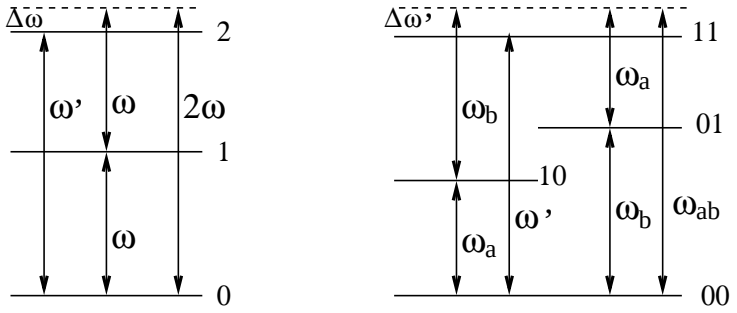

Figure 4.5: Pictorial description of anharmonicity. (a) diagonal anharmonicity, $\Delta\omega = \omega' - 2\omega$, and (b) off-diagonal anharmonicity $\Delta\omega' = \omega' - (\omega_a + \omega_b)$

as,

$$\Delta\omega = \omega' - 2\omega, \tag{4.33}$$

and the off-diagonal anharmonicity is defined as,

$$\Delta\omega' = \omega' - (\omega_a + \omega_b). \tag{4.34}$$

Here ω' is the wave number calculated from second excitation and ω is calculated from first excitation. ω_a and ω_b are wave numbers calculated from first excitation for two different normal modes. The off-diagonal anharmonicity accounts for the coupling of two different normal modes.

4.3 Computational methods

Geometry optimization of Methyl benzoate was performed using second order Møller-Plesset[61] (MP2) perturbation theory[62,63] and employing the augmented correlation-consistent polarized-valence-triple zeta (aug-cc-pVTZ) basis set. Harmonic normal-mode analysis was performed with the density fitting MP2 (DF-MP2) method using an aug-cc-pVTZ regular basis and cc-pVTZ fitting basis sets. All calculations were performed with the MOLPRO quantum chemistry program.[150]

The choice of the computational method and the basis set are not arbitrary. A standard MP2 with aug-cc-pVTZ basis set level of computation is generally a reliable method to generate an anharmonic PES. For a system like Methyl benzoate it is not suitable due to the size of the molecule and the resulting high computational cost. To find a suitable method a comparison of computations employing different basis sets has been performed and is presented in Table 4.1. The standard MP2 calculation employing an aug-cc-pVTZ basis set is given in the first row and provides the reference. From Table 4.1 it is clear that the standard MP2 method with aug-cc-pVTZ basis is beyond our scope to study all 48 normal modes for Methyl benzoate, as it takes almost one day and a very large memory space for a single-point energy calculation. If no specially optimized fitting basis sets are available for a certain regular AO basis set it is common practice for the DF-MP2 computations to use a fitting basis set one order higher than the regular basis set. The DF-MP2 with aug-cc-pVTZ regular basis and cc-pVQZ fitting basis speeds up the calculation dramatically (see Table 4.1) without sacrificing quality compared to the standard MP2/aug-cc-pVTZ results. A single-point energy calculation with this method takes only one and a half hours and requires almost 200 times less memory. Whereas the DF-MP2 method employing aug-cc-pVTZ regular and cc-pVTZ fitting basis sets gives almost identical energy values in half the time using far less disc space and is a reasonable choice for our purpose.

Table 4.1: Single point energy, calculation time and the required memory for energy calculation of Methyl benzoate at equilibrium geometry with different basis sets are presented. First row represents the normal MP2 result.

Basis		Energy in E_h			CPU time	Memory
mp2fit	jkfit	SCF	MP2	Total	in Sec.	in MB
		-457.510382	-1.760440	-459.270822	85363	6799.36
avtz	vtz	-457.510134	-1.760075	-459.270209	2822	270.39
avtz	vqz	-457.510341	-1.760277	-459.270618	5829	361.65
avqz	vqz	-457.537351	-1.860767	-459.398118	17132	3194.88
avqz	v5z	-457.537383	-1.860811	-459.398194	29014	3194.88

The anharmonic pair couplings are calculated with the DF-MP2 level of theory employing the same cc-pVDZ basis for regular and fitting basis set.

The grid was chosen using according to the harmonic frequency analysis and PM3 PES cuts. The innermost two points (-h, h) are determined from the second derivative of the PES at equilibrium corresponding to the harmonic frequencies of the normal modes. The outer two points (-a, b) are determined as those points on the 1D-PES cuts along the normal modes for which the PM3 energy reaches six times the harmonic frequency quantum w.r.t. V_0 (see Fig. 4.6). The remaining points are calculated dividing the interval with different proportion e.g., $c = \frac{-a+(-h)}{2}$, $d = \frac{b+h}{2}$, $e = \frac{c+(-h)}{2}$, $f = \frac{d+h}{2}$, $i = \frac{-a+c}{2}$, $j = \frac{d+b}{2}$, $k = \frac{e+(-h)}{2}$ and $l = \frac{f+h}{2}$. The choice of an appropriate grid size is crucial

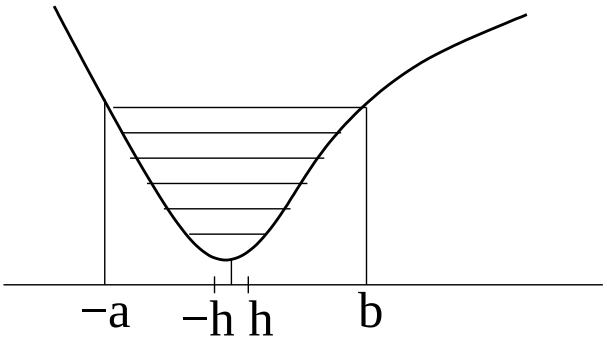

Figure 4.6: Grid limits to generate the PES. (-h, h) are determined from the second derivative of the PES at equilibrium corresponding to the harmonic frequencies of the normal modes. (-a, b) are determined as those points on the 1D-PES cuts along the normal modes for which the PM3 energy reaches six times the harmonic frequency w.r.t. V_0.

for anharmonic frequency calculations for both diagonal and pair potentials. Diagonal potentials with 6 points (see Fig. 4.7 (a)) are insufficient for a reasonable description of the PES in the calculation of anharmonic frequencies. In particular, a denser set of energy points is required near the equilibrium configuration. A comparative study has been carried out with 8 point (see Fig. 4.7 (b)) and 12 point (see Fig. 4.7 (c)) 1D grids. It indicates that the 8 point grid PES is a reasonable choice for anharmonic frequency calculations (see Table B.1). All diagonal points were first evaluated on an eight point one-dimensional grid as shown in Fig. 4.7(b), then interpolated to the equally spaced 16 point grids which were used in the collocation treatment. Pair potentials, calculated at semiempirical PM3 level, were evaluated on 8×8 point direct product grids and then interpolated to 16×16 point grids by two-dimensional cubic spline interpolation.[151] To lower the computational cost while maintaining high quality pair potentials, we performed DF-MP2/cc-pVDZ computations on irregularly spaced two-dimensional grids, as shown in Fig. 4.8. Energy points are calculated at the grid points marked with crosses. The undetermined points (marked with circles in Fig. 4.8) are filled by using non-uniform IMLS (interpolating moving list-squares)[152] interpolation, which provides a potential on 8×8 point direct product grids. The 8×8 point grid is extended to the 16×16 point grid by using 2D cubic-spline interpolation. For some of the most problematic modes the diagonal potentials are calculated with the local density fitting coupled cluster singles and doubles and perturbative triple correction DF-L-CCSD(T)[153–155] method with the cc-pVTZ basis set in a dual-level scheme. All computations were additionally performed using the density fitting spin-component scaled (DF-SCS) MP2 method[156] employing different basis sets.

The theoretical calculation of the vibrational spectra of Methyl benzoate has been performed at its minimum energy structure obtained by a standard optimization procedure using the MP2 correlation treatment and employing aug-cc-pVTZ basis sets. The Cartesian normal modes obtained from this structure are further used for the expansion of the anharmonic PES. The harmonic and anharmonic frequencies of Methyl benzoate are obtained using the non-correlated VSCF method as well as the correlation corrected VSCF method.

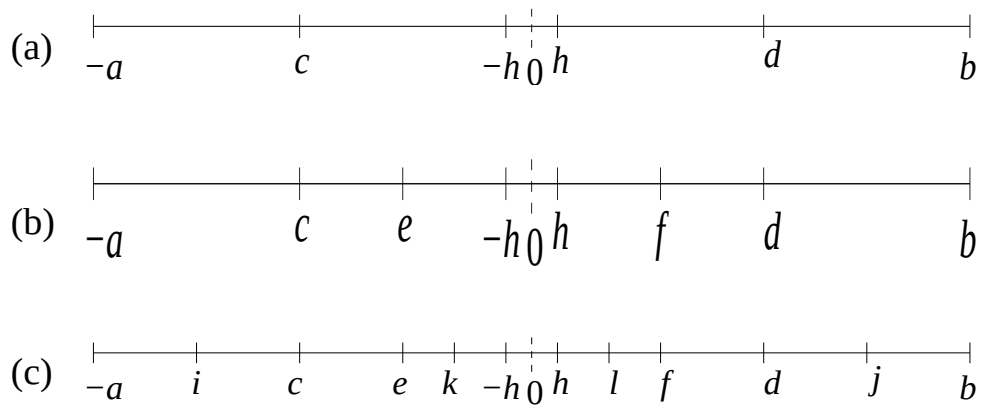

Figure 4.7: Diagonal grid with different grid size; (a) 6 points grid, (b) 8 points grid, (c) 12 points grid

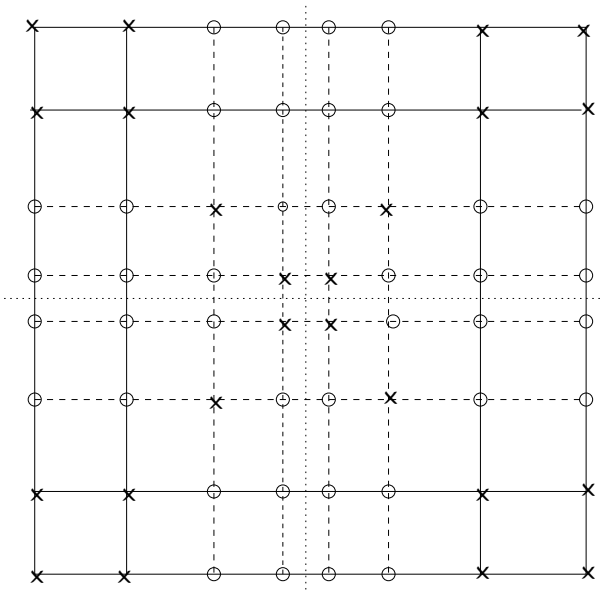

Figure 4.8: Irregular spaced 2-dimensional grid. Energy points are calculated at the grids, marked with crosses. The grids marked with circles are filled up by 2-dimensional IMLS interpolation.

4.4 Results and Discussion

4.4.1 Description of normal modes

A complete description of normal modes of Methyl benzoate are presented here with the corresponding mode number.

Mode 1 Rotation of ester group about C–C single bond.
Mode 2 Out-of-plane bending of ring along with ester group.
Mode 3 Out-of-plane C–H bending in methyl group.
Mode 4 In-plane rotation of phenyl ring as well as ester group.
Mode 5 Out-of-plane ring deformation along with out-of-plane bending of ester group.
Mode 6 In-plane deformation of ester group.
Mode 7 In-plane bending of O=C–O along with symmetric H–C–H bending in methyl group.
Mode 8 Out-of-plane ring deformation but ester group remain stationary.
Mode 9 Out-of-plane ring deformation along with out-of-plane O=C–O bending.
Mode 10 In-plane rotation of phenyl ring and ester group in opposite direction.
Mode 11 In-plane deformation of phenyl ring but ester group remain stationary.
Mode 12 In-plane deformation of phenyl ring along with in-plane bending of O=C–O.
Mode 13 Out-of-plane deformation of phenyl ring along with O=C–O out-of-plane bending.
Mode 14 Out-of-plane deformation of phenyl ring along with O=C–O out-of-plane bending.
Mode 15 O=C–O in-plane bending.
Mode 16 Out-of-plane phenyl ring deformation along with C–H out-of-bending in phenyl ring.
Mode 17 Out-of-plane phenyl ring deformation but ester group remain stationary.
Mode 18 In-plane deformation of phenyl ring along with in-plane bending of ester group.
Mode 19 In-plane ring deformation with very little coupling with ester group.
Mode 20 In plane ring deformation, ester group remain stationary.
Mode 21 Out-of-plane C–H bending in phenyl ring.
Mode 22 In-plane C–H bending in phenyl ring.
Mode 23 Out-of-plane C–H bending in phenyl ring.
Mode 24 Rocking motion of methyl group along with in-plane C–H bending in phenyl ring.
Mode 25 In-plane C–H bending in phenyl ring along with in-plane ring

deformation and C-O-C in-plane bending.

Mode 26 Asymmetric C–H bending in methyl group along with C-O-C out-of-plane bending.

Mode 27 Symmetric C–H bending in phenyl ring.

Mode 28 Out-of-plane C–H bending in phenyl ring.

Mode 29 C–O–C in-plane bending along with C–H bending in methyl group.

Mode 30 C–O stretching along with C–H bending in methyl group.

Mode 31 In-plane C–H bending in phenyl ring.

Mode 32 C–C stretching in phenyl ring.

Mode 33 C–H in-plane bending in phenyl ring as well as methyl group.

Mode 34 C–C stretching in phenyl ring along with C–H bending in methyl group.

Mode 35 Rocking motion of methyl group.

Mode 36 C–H bending in methyl group.

Mode 37 C–C stretching along with C–H in-plane bending in phenyl ring.

Mode 38 In-plane ring deformation.

Mode 39 In-plane ring deformation.

Mode 40 C=O stretching.

Mode 41 Symmetric C–H stretching in methyl group.

Mode 42 Asymmetric C–H stretching in methyl group.

Mode 43 C–H stretching in phenyl ring.

Mode 44 C–H stretching in phenyl ring.

Mode 45 Symmetric C–H stretching in methyl group.

Mode 46 C–H stretching in phenyl ring.

Mode 47 C–H stretching in phenyl ring.

Mode 48 C–H stretching in phenyl ring.

4.4.2 Vibrational spectrum of Methyl benzoate

There are not many experimental results (except for the low-resolution IR Raman spectrum of Chattopadhyay[157]) available for Methyl benzoate. Also not all modes are resolved and identified experimentally, especially for low-frequency modes and near-degenerate modes. We therefore choose our best computational result, which agrees comparatively well with the known experimental results,[157–160] as a reference and then discuss all other

levels of computation with respect to that. For this purpose, we choose the non-correlated VSCF results as a reference, where diagonal and pair potentials are calculated using the DF-SCS method employing a cc-pVDZ basis set.

The entire vibrational spectrum of Methyl benzoate can be divided into three different important frequency regions. Most interesting high frequency modes are due to the C-H stretching modes. Low frequency (up to $\sim$1000 cm^{-1}) modes are primarily due to bending motions involving the phenyl ring and the ester group. Between these two extremes there are modes with combinations of bending and stretching motions and also the spectroscopically well-separated C=O band.

The vibrational frequencies of all modes of Methyl benzoate calculated at different levels of theory are presented in Table 4.2 along with their experimentally observed values. The deviation of the vibrational frequencies calculated at different levels of theory with respect to the frequency obtained from the non-correlated VSCF/DF-SCS method are plotted in Figs. 4.9, 4.10, 4.11 and 4.12. The trend observed for the root mean squared deviation (RMSD) with respect to VSCF/DF-SCS going from harmonic (148 cm^{-1}) to the inclusion of only diagonal anharmonic potentials (151 cm^{-1}), and finally to the full potential expansion up to pair contributions at PM3 level (78 cm^{-1}) and DF-MP2/cc-pVDZ level (24 cm^{-1}) shows that a balanced description of diagonal and coupling contributions is of importance in simplified models. Surprisingly, for VC-MP2 the RMSD (45 cm^{-1}) for VSCF/DF-SCS deteriorates compared to the non-correlated VSCF treatment. Correlation correction may not be necessary for such a large system,[24] and the VC-MP2 method may suffer from a degeneracy problem due to the high density of states in Methyl benzoate. As seen from Figs. 4.10, 4.11 and 4.12, the VC-MP2 vibrational frequencies deteriorate for modes 2, 14, 16, 17, 35, 41, 45 and 46. The V-MP2 also shows degeneracy problem for these modes. Other than these modes, VC-MP2 method shows good agreement with the experimental results.

For low frequency modes dual level calculations with PM3 pair couplings are not reliable. Sometimes it overestimates and sometimes it underestimates the results. The RMSD calculated for all 40 low frequency modes is thus quite high, 73 cm^{-1}. Harmonic and diagonal frequencies (frequency calculated from the anharmonic diagonal PES) are also quite offset for few modes (modes 1, 2, 3, 13, 18, 35, 41 to 48) from the reference frequencies and give RMSDs of 35 cm^{-1} and 37 cm^{-1}, respectively. Frequencies calculated with the non-correlated VSCF method with DF-MP2/cc-pVDZ diagonals and pair coupling potentials are in good agreement with the experimental results, which yields a very

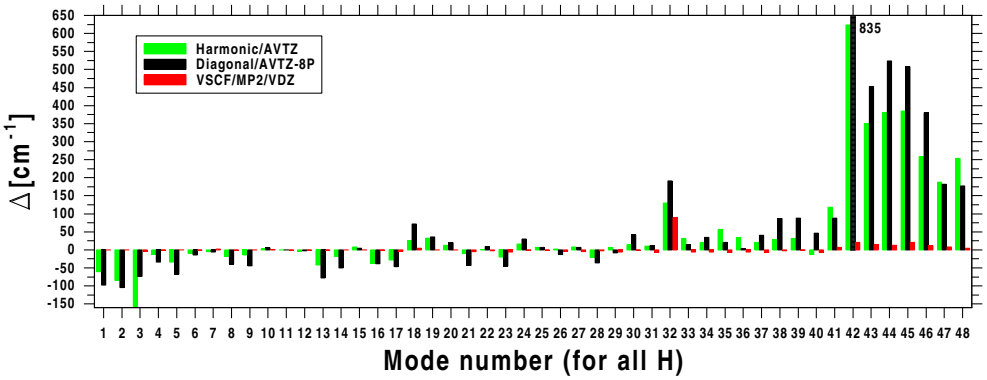

Figure 4.9: A comparison of vibrational frequencies of Methyl benzoate calculated at different levels of theory with respect to the VSCF/DF-SCS calculation. The labels indicate which level of theory was used for the calculation.

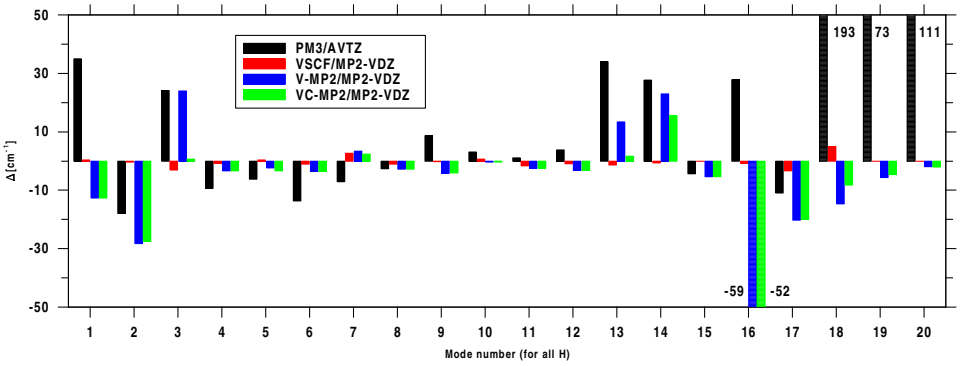

Figure 4.10: A comparison of vibrational frequencies of Methyl benzoate calculated at different levels of theory with respect to the VSCF/DF-SCS calculation for the first 20 vibrational modes. The labels indicate which level of theory was used for the calculation.

low RMSD. For the 40 low frequency modes RMSD at this level of calculation is 14 cm^{-1}. Dual level calculation using DF-MP2/aug-cc-pVTZ diagonal and DF-MP2/cc-pVDZ pair potentials, improve the results further, with a RMSD of 11 cm^{-1}. For the VC-MP2 the

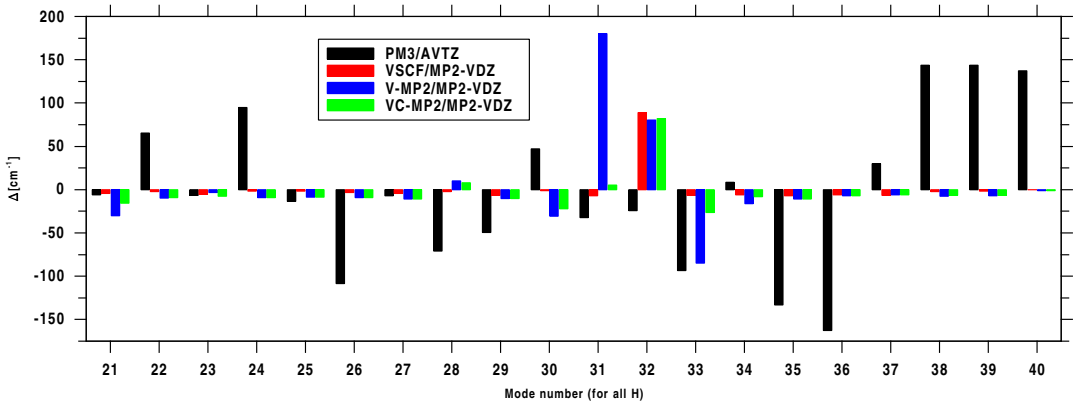

Figure 4.11: A comparison of vibrational frequencies of Methyl benzoate calculated at different levels of theory with respect to the VSCF/DF-SCS calculation for modes 21 to 40 only. The labels indicate which level of theory was used for the calculations.

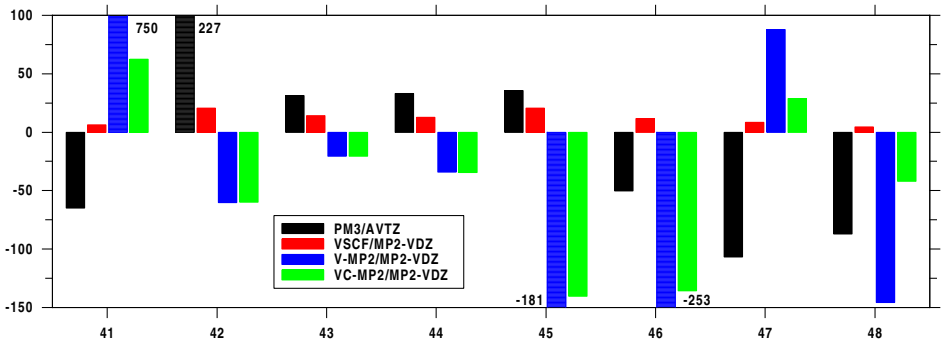

Figure 4.12: A comparison of vibrational frequencies of Methyl benzoate calculated at different levels of theory with respect to the VSCF/DF-SCS calculation for high frequency modes (41 to 48) only. The labels indicate which level of theory was used for the calculations.

RMSD slightly deteriorates to 18 cm^{-1}.

Table 4.2: Vibrational frequencies of Methyl-Benzoate with different level of theories, where diagonal grids are calculated with the DF-MP2/AVTZ level. Mode numbers are based on normal mode frequencies. The RMSD calculated with respect to the anharmonic frequency calculated with VSCF/DF-SCS method.

Mode	Harmonic	Diagonal	VSCF		VC-MP2	VSCF/	Experiment
			PM3	MP2		SCS-DZ	[Ref. 157]
1	52	82	147	103	87	112	
2	112	154	178	192	160	196	
3	185	361	366	316	352	342	
4	167	170	169	173	169	179	134
5	208	217	236	239	234	242	218
6	331	333	326	334	331	340	
7	358	358	356	366	365	363	360
8	404	411	419	416	414	421	
9	454	458	477	470	466	468	
10	481	481	481	476	475	477	
11	617	617	618	612	611	617	630
12	678	679	685	676	673	681	674
13	712	728	788	737	726	754	
14	795	801	840	808	837	813	782
15	831	830	818	818	812	823	820
16	694	760	759	781	724	731	714
17	863	894	880	856	841	891	864
18	1003	1004	1170	974	971	978	
19	1012	1012	1052	991	989	979	
20	1048	1049	1146	1029	1025	1035	1027
21	977	1000	980	978	987	986	980
22	1097	1101	1162	1082	1074	1096	1097
23	938	970	951	931	926	958	950
24	1142	1144	1220	1111	1103	1126	1128
25	1174	1189	1154	1153	1146	1167	1161
26	1185	1199	1074	1170	1161	1182	
27	1191	1198	1176	1164	1156	1183	1177
28	958	994	908	968	982	979	
29	1221	1227	1164	1200	1194	1214	1192
30	1317	1321	1350	1283	1262	1303	1278
31	1330	1334	1287	1301	1309	1320	1295
32	1472	1469	1318	1423	1415	1342	1310

continued to next page

continued from previous page							
Mode	Harmonic	Diagonal	VSCF		VC-MP2	VSCF/	Experiment
			PM3	MP2		SCS-DZ	[Ref. 157]
33	1479	1482	1354	1433	1411	1447	1435
34	1471	1475	1461	1436	1434	1452	1444
35	1509	1507	1320	1459	1454	1453	
36	1519	1517	1322	1488	1488	1485	
37	1518	1519	1529	1478	1477	1498	1474
38	1634	1636	1750	1587	1581	1606	1594
39	1634	1634	1747	1587	1580	1603	1585
40	1766	1759	1915	1724	1724	1778	1724
41	3092	3059	2909	2946	2795	2974	2998
42	3183	3255	2787	2582	2492	2560	2542
43	3205	3208	2887	2893	2849	2855	2855
44	3214	3247	2867	2872	2834	2834	2852
45	3217	3213	2868	2860	2694	2832	2845
46	3222	3233	2913	2995	2849	2963	2952
47	3226	3191	2932	3014	3035	3039	3064
48	3238	3161	2898	2933	2928	2985	3022
RMSD	148	151	78	24	45		

In the low frequency region mode 32, which corresponds to the asymmetric bending motion in the methyl group, is the most problematic mode, as seen in Fig. 4.9 and 4.11. Harmonic and diagonal frequencies for this mode are 130 and 192 cm^{-1} too high, respectively. The frequency calculated employing DF-MP2/cc-pVDZ pair potentials is also too high ($\sim$ 80 cm^{-1}) compared to experimental result.

High frequency vibrational modes

The high-frequency vibrational modes (41 to 48) are spectrally isolated from all other modes of Methyl benzoate (see Table 4.2). The deviation in frequency for these 8 modes calculated at different levels of theory with respect to the VSCF/SCS-VDZ method is plotted in Fig. 4.12. Harmonic and diagonal frequencies (green and black bars) are too high as can be seen in Fig. 4.9. The RMSD w.r.t. the VSCF/DF-SCS results for the C–H stretch vibrational modes (modes 41–48) alone is 319 cm^{-1} for the harmonic frequency, whereas it is 454 cm^{-1} for the diagonal frequencies. The large RMSDs for these eight modes from the harmonic and the diagonal frequency calculations are mostly due to mode 42, which shows an unexpected red shift to the frequency $\sim$ 2550 cm^{-1}, whereas the harmonic frequency for the mode 42 is 3183 cm^{-1} and the diagonal frequency is 3255

cm^{-1}. However, when pair couplings are included, the RMSD is reduced dramatically. Dual level calculations, with DF-MP2/aug-cc-pVTZ diagonal potentials and PM3 pair potentials (see Table 4.2) yield a RMSD of 100 cm^{-1}. In this case mode 42 is also quite high but shows a red shift from the conventional C–H stretch vibrational frequency to a frequency of 2787 cm^{-1}. Modes 43, 44, and 45 are $30-35$ cm^{-1} higher than the reference frequencies, whereas modes 46, 47, 48 are lower in frequency. Dual level non-correlated VSCF computation with DF-MP2/aug-cc-pVTZ diagonal and DF-MP2/cc-pVDZ pair potentials are surprisingly accurate for the C–H stretch vibrational frequencies. The RMSD calculated for these eight modes is 13 cm^{-1} only. The main reason for large RMSDs in the case of harmonic and diagonal anharmonic frequencies is that these neglect all coupling terms. In reality, a C–H stretch vibrational mode is strongly coupled to other C–H stretch vibrational modes. This is also the reason why the C–H stretch vibrational frequencies are obtained rather accurately from the pair coupling approximation. It is also observed that these couplings are localized, that means the coupling is only with other C–H stretch vibrations. The methyl group C–H stretch vibrations are completely isolated from the phenyl ring C–H stretch vibrations. Results for C–H stretch vibrational modes may be further improved by adding triple contributions for these modes and performing vibrational CI calculations. Our investigation with selected semi-empirical PM3 triple contributions over DF-MP2/cc-pVDZ though does not improve the results much, only 2 to 3 wave number frequency shifts are observed. Triple contributions from higher level quantum chemical methods should yield more accurate results.

A rather surprising result of the anharmonic pair-coupling calculations is that one of the C–H stretch vibrational modes is shifted to a much lower frequency than the remaining C–H stretch vibrations. No such frequency lowering can be observed in the harmonic and the anharmonic diagonal vibrational frequency calculations. This red shift of the C–H stretch vibrational mode has not been discussed in literature, so far. A close inspection of the experimental spectra,[161] however, reveals the presence of an unassigned peak at 2560 cm^{-1}, with low intensity. In a recent 1D IR experimental study of Methyl benzoate by the Steinel group,[162] a low intensity peak is observed at 2542 cm^{-1} which matches the calculated frequency.

Closer inspection of this mode reveals that this frequency originates from the asymmetric C–H bond stretching in the methyl group, where two out of plane hydrogens vibrate along the C–H bond in opposite direction. Their motion is depicted in Fig. 4.13.

Since this red shift is observed only from the pair coupling analysis and not from

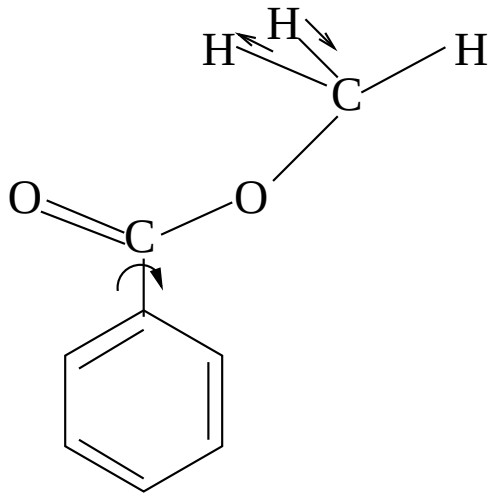

Figure 4.13: Vibrational motion of Methyl benzoate for unexpected red shift.

the harmonic and the diagonal anharmonic analysis, a noticeable coupling to some other modes must exist. It was found that the vibrational mode (mode 1) corresponding to the rotation of the ester sub-group with respect to the phenyl ring about the C–C bond is coupled to this C–H vibrational mode. The rotational motion is indicated in Fig. 4.13 by a curved arrow. This rotation pulls in one out-of-plane hydrogen to the plane of molecule and pushes the other from the molecular plane. Due to this motion force constant of this mode may change unexpectedly which may give rise to this red shift of the C–H vibrational frequency.

C=O band

The C=O stretch vibrational frequency is spectrally isolated from all other vibrational frequencies in the Methyl benzoate spectrum. Even for large proteins, the C=O vibrational band is easily resolved with high intensity. Computationally, however, the study of this vibrational mode is not straightforward. The harmonic and the anharmonic diagonal frequencies at DF-MP2/aug-cc-pVTZ level are relatively higher (1766 and 1759 cm^{-1} respectively) for this mode, compared to its experimental frequency (1724 cm^{-1}) (see Table 4.2). The harmonic frequency calculated at SCS/aug-cc-pVTZ level is found to be higher than the observed frequency as well. Energy extrapolation (HF energy extrapolated with exponential fit and correlation energy is extrapolated with the X5 method, see chapter 2) to the basis set limit yields results similar to MP2/aug-cc-pVTZ harmonic

frequency.

The anharmonic diagonal frequency at MP2/aug-cc-pVTZ with an 8 points grid (see Fig. 4.7) improves the results over the harmonic frequency and yields a C=O vibrational frequency of 1755 cm^{-1}. On the other hand, the anharmonic frequency calculated employing MP2/cc-pVDZ pair potentials is again quite high, i.e., 1771 cm^{-1}. SCS/cc-pVDZ pair potentials yield even worse results.

Dual level frequency calculations employing DF-MP2/aug-cc-pVTZ diagonal and the PM3 pair potential is even worse (see Table 4.2). The calculated C=O frequency is improved when dual level calculation are performed replacing the DF-MP2/cc-pVDZ diagonal potential for the C=O vibrational mode alone by a DF-L-CCSD(T)/cc-pVTZ diagonal potential. A perfect match with the experimentally observed frequency (1724 cm^{-1}) is obtained. For Methyl benzoate, the non correlated VSCF computation yields the C=O stretch vibrational frequency of 1726 cm^{-1}, where as both V-MP2 and VC-MP2 yield 1727 cm^{-1}. For the C=O stretch mode a high level description of the diagonal potential is thus of tantamount importance.

4.4.3 Deuterated Methyl benzoate

Vibrational frequencies are strongly affected by the change in mass due to isotope substitution. Upon isotopic substitution, the normal mode frequency changes roughly with the square root of the inverse of the reduced mass, i.e., $\nu \propto \frac{1}{\sqrt{\mu}}$. If the isotope mass difference is large, it significantly affects the vibrational frequency. For that reason the substitution of hydrogen by deuterium (atomic mass twice that of hydrogen) is an excellent method for structure determination of bio-molecules, where localized C–H vibrational frequencies are lowered (red shifted) by several hundreds of wave numbers upon deuterium substitution.

Deuteration of Methyl benzoate in the ortho position of the benzene ring yields the syn- and anti- isomers of ortho-deutero methyl benzoate (o-DMB) shown in Fig. 4.14. The ester group can rotate around the C–C single bond and yields these two main structural conformers. In the potential energy minima the ester group remains in the same plane as the benzene ring. Upon thermal excitation, the conformers interconvert by a rotation of the ester group around the C–C bond axis. The barrier height for the ester group rotation in Methyl benzoate is $\sim$ 0.25 eV, as calculated by the DF-MP2 method with aug-cc-pVTZ basis set. This is rather small and easily accessible even at room temperature and thus the conformers can easily interconvert. Since both conformers have nearly the same equilibrium structural energy, they are also equally populated. Although both conformers

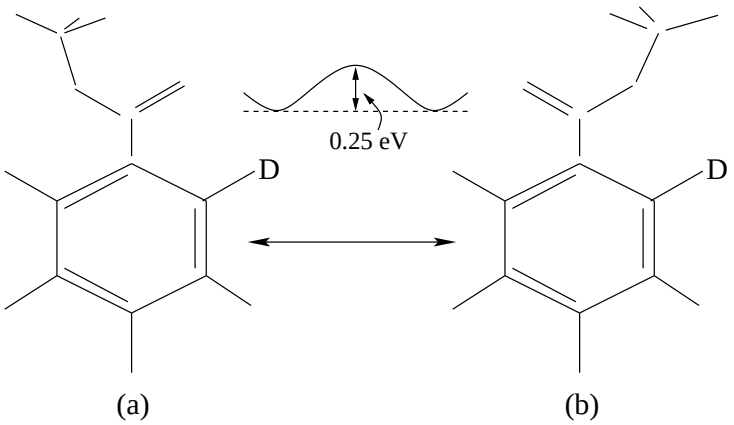

Figure 4.14: Two structural conformers of Methyl benzoate and their approximate interconversion barrier height. (a) syn-o-deutero-methyl-benzoate (b) anti-o-deutero-methyl-benzoate

have the same molecular energy, they possess different vibrational C–D frequencies due to the different couplings to the ester group. With the best possible computation (dual level VSCF) a C–D stretch frequency difference for these two conformers of about 10 cm^{-1} is found.

Identification of isotope effects by co-diagonalization

A symmetric matrix (A) can be diagonalized by an orthogonal matrix. For a set of matrices, if they have an average eigen-structure it is possible to find an orthogonal matrix which can diagonalize all the matrices as much as possible. This process is called co-diagonalization of matrices. In the co-diagonalization process, not all the off-diagonal elements become zero, but the sum of the square of all off-diagonal elements ($\sum_{i \neq j} A_{ij}^2$) is minimized.

The co-diagonalization process has an advantageous application in vibrational spectroscopy to identify the coupling of different modes and isotope effects. When an atom is substituted with its isotope, some of the eigen-vectors differ from the non-substituted molecule. In co-diagonalization these are identified by non-zero off-diagonal elements. The frequency shift for the primary isotope effect appears in the diagonal element and secondary effects appear as off-diagonal elements. Fig. 4.15 (a) and (b) depict the resulting co-diagonalized matrices of the syn- and anti-o-DMB where frequencies are on the diagonal and the residual couplings are shown in the upper left triangle on the same

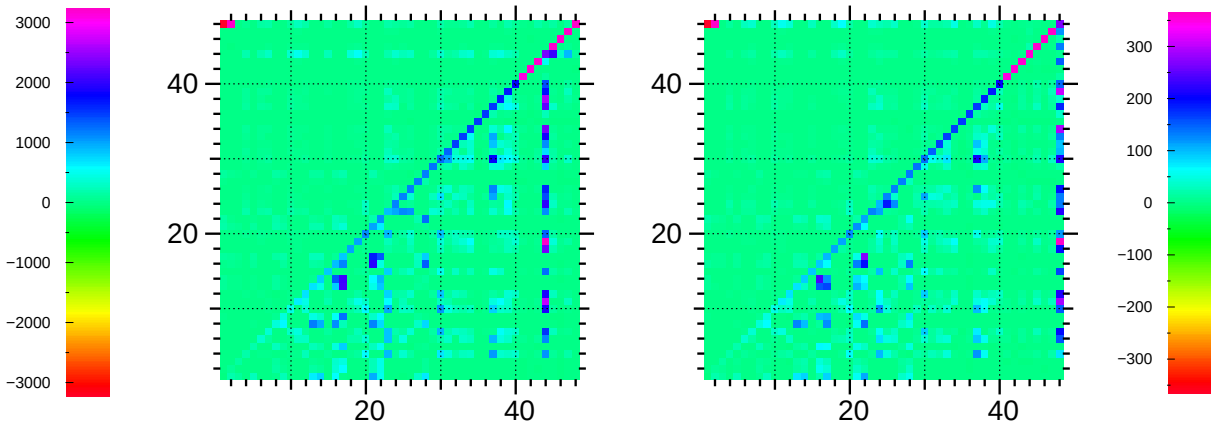

Figure 4.15: Frequency spectrum for isotopomers. (a) syn-o-deutero-methyl-benzoate and (b) anti-o-deutero-methyl-benzoate. Fundamental frequencies are plotted on the diagonal and the residual couplings are shown in upper left triangle on the same scale of diagonal. The scale is shown by left rainbow color spectrum. Magnified coupling elements are shown in lower right triangle. the corresponding scale is shown by right rainbow color spectrum.

scale as of the diagonal. But the order of magnitude is so small that they are essentially invisible. The lower right triangle shows magnification of couplings by approximately a factor of ten.

The syn- and anti-isomers show a clear primary isotope effect for the 43rd (Fig. 4.15 (a)) and 48th (Fig. 4.15 (b)) vibrational modes, respectively. The primary isotope effect is observed in to the C–D stretching vibration. The secondary isotope effects are mainly due to the coupling with the C–D vibrational modes as seen from the Fig. 4.15 (a) and (b). These appear in the low frequency region. A noticeable feature of the isotope effect is that both isotopomers show secondary isotope effects for the same normal modes (modes 8, 13, 14, 17, 18, 19, 20, 23, 27 and 31), where as primary isotope effects are observed for different normal modes (see Table 4.10).

Vibrational frequencies of anti-o-deutero-methyl-benzoate

For the deuterated species, the overall RMSD shows a similar behavior as in all-H Methyl benzoate. Calculated harmonic and anharmonic diagonal frequencies are quite offset from the reference frequencies and give high RMSD values of 137 cm^{-1} and 130 cm^{-1}, respectively. Non-correlated VSCF dual level frequency calculations based on MP2/aug-

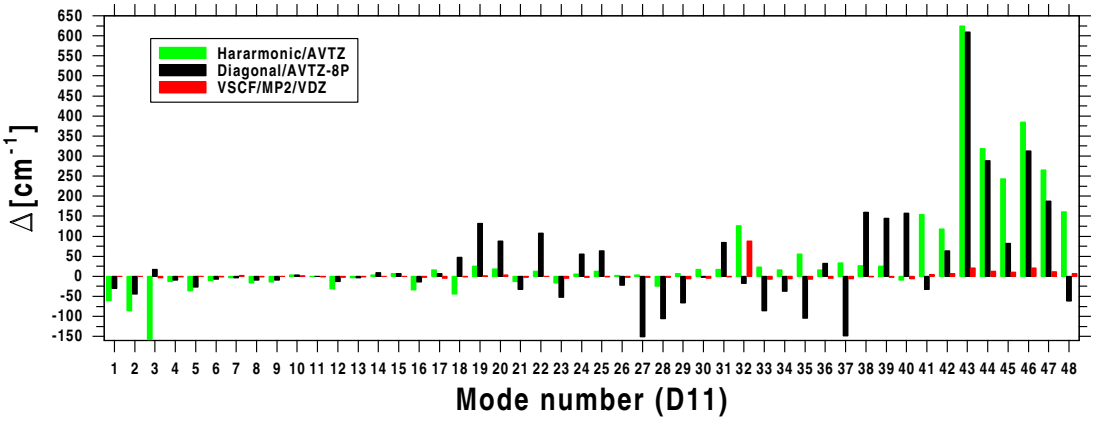

Figure 4.16: A comparison of vibrational frequencies of anti-o-deutero-methyl-benzoate calculated at different levels of theory with respect to VSCF/DF-SCS calculation.The labels indicate which level of theory was used for the calculations.

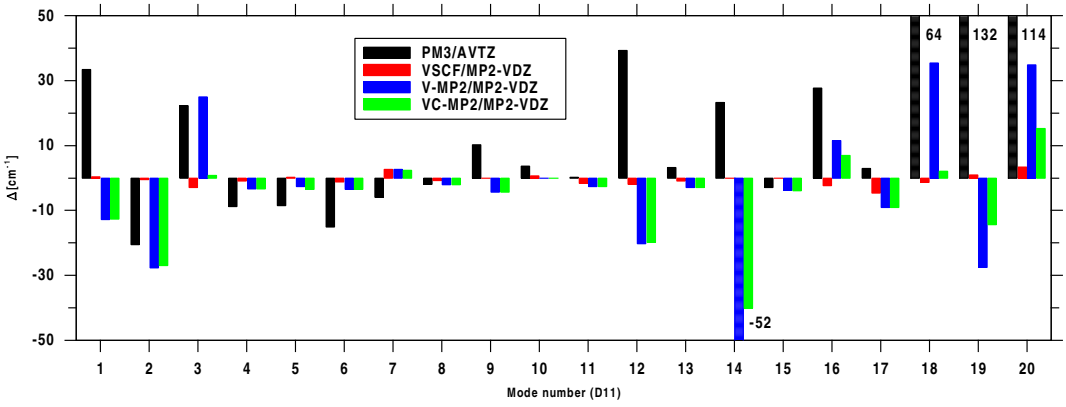

Figure 4.17: A comparison of vibrational frequencies of anti-o-deutero-methyl-benzoate calculated at different levels of theory with respect to VSCF/DF-SCS calculation for the first 20 modes only. The labels indicate which level of theory was used for the calculations.

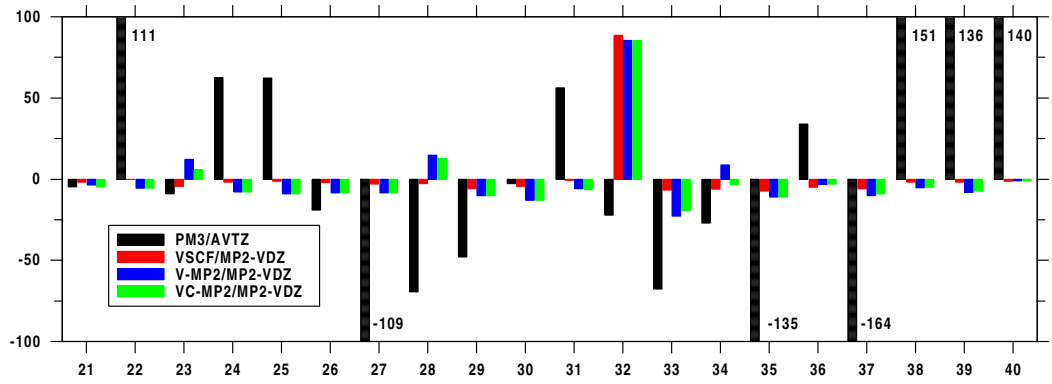

Figure 4.18: A comparison of vibrational frequencies of anti-o-deutero-methyl-benzoate calculated at different levels of theory with respect to VSCF/DF-SCS calculation for modes 21 to 40 only. The labels indicate which level of theory was used for the calculations.

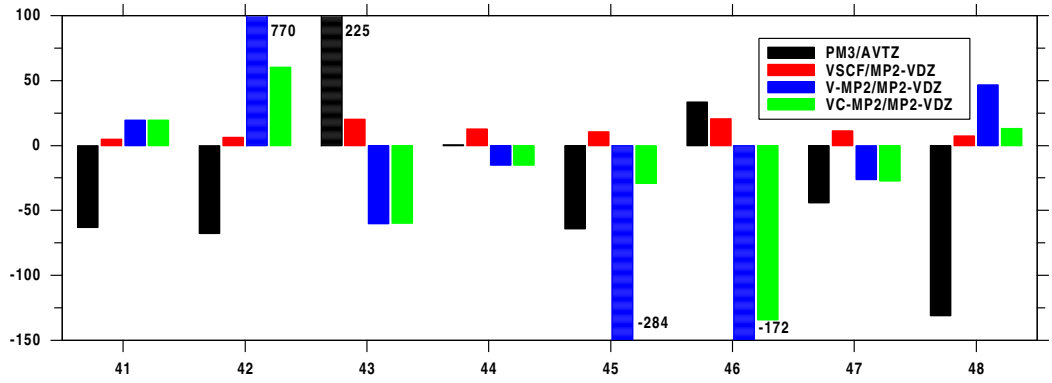

Figure 4.19: A comparison of vibrational frequencies of anti-o-deutero-methyl-benzoate calculated at different levels of theory with respect to VSCF/DF-SCS calculation for modes 41 to 48 only. The labels indicate which level of theory was used for the calculations.

cc-pVTZ diagonal and PM3 pair potentials also deviate considerably from the reference frequencies (RMSD = 76 cm^{-1}). The non-correlated VSCF frequencies employing DF-MP2/cc-pVDZ pair and DF-MP2/aug-cc-pVTZ diagonal potentials show good agreement with the reference frequencies and yield a very low RMSD of 14 cm^{-1}, for all 48 modes (see Table B.2). Figs. 4.16, 4.17, 4.18 and 4.19 depict the general features of the different computational methods for the deuterated Methyl benzoate and Table B.2 presents the vibrational frequencies for the anti-o-deutero-methyl-benzoate. Dual level frequency calculation employing DF-MP2/aug-cc-pVTZ diagonal potentials and PM3 pair potentials for all 40 low frequency modes are not reliable, but for the high frequency modes it yields quite reasonable results. The harmonic and the diagonal anharmonic frequencies are also too high in the low frequency region as well as the high frequency region. For the low frequency modes (40 modes) the non-correlated VSCF employing DF-MP2/aug-cc-pVTZ diagonal and DF-MP2/cc-pVDZ pair potentials are in good agreement with the reference frequencies, yielding a RMSD of 13 cm^{-1}. This level of computation is also efficient in the high frequency region and yields a RMSD for the 8 C–H stretch vibrational modes of 17 cm^{-1}.

C–D band

C–H stretch vibrations are very localized in the region 2850–3100 cm^{-1}. Upon isotopic substitution of a hydrogen by deuterium, the respective frequency is lowered by the several hundred wave numbers. In biological systems the C–D stretch frequency is an excellent structural probe since it is usually spectrally isolated even in the spectrum of large proteins.

In Methyl benzoate, the study of the ortho C–H stretch vibrational mode and its coupling with other modes, in particular, with the C=O stretching mode, is an important model for similar (C=O to β-hydrogen) couplings in peptides. Computationally however, the study of this vibrational mode is not straightforward. Lower level computational methods predict this vibrational frequency very poorly. Also, the experimental vibrational frequency associated with the C–D stretching mode in the deuterium substituted Methyl benzoate has not yet been reported.

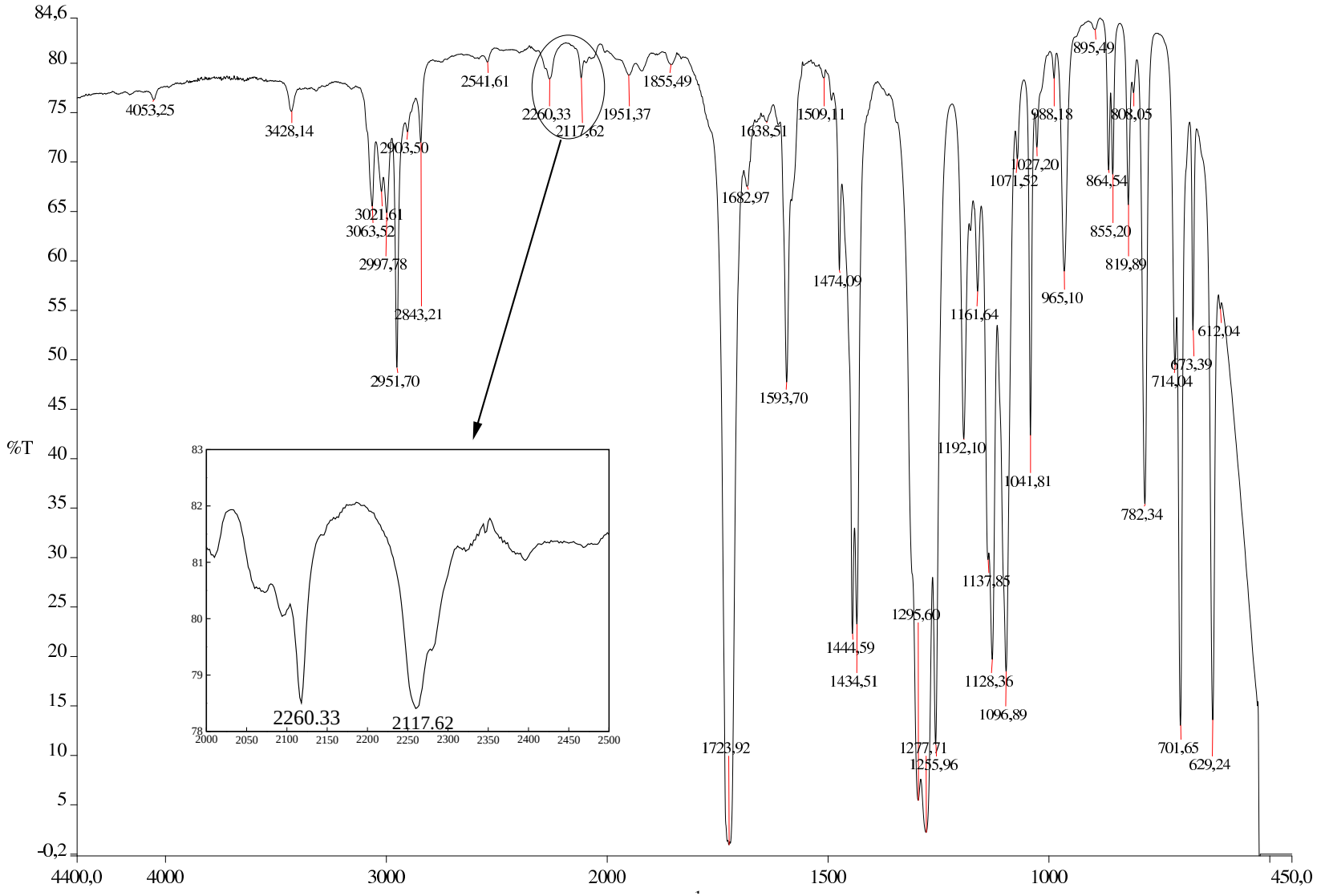

Figure 4.20: 1D IR experimental spectrum of deuterated Methyl benzoate. [165]

The harmonic C–D vibrational frequencies at DF-MP2/aug-cc-pVTZ level are rather large for the two possible conformers, at 2384 and 2392 cm^{-1} for the syn- and anti-o-DMB, respectively. The harmonic analysis at SCS/aug-cc-pVTZ level also predicts similar results. When the 1D PES is generated by extrapolating the HF energy by the exponential extrapolation and the correlation energy by the X5 extrapolation, the harmonic frequency calculation yields slightly lower C–D frequencies, of 2381 cm^{-1} and 2389 cm^{-1} for the syn- and anti-o-DMB, respectively. The anharmonic diagonal frequencies at DF-MP2/aug-cc-pVTZ level are lower them by about 70 cm^{-1} from the harmonic frequency.

It is noteworthy that the C–D stretching frequency is significantly improved by using the pair potentials calculated at DF-MP2/cc-pVDZ level. The non-correlated dual level VSCF method yields C–D vibrational frequencies of 2171 cm^{-1} and 2161 cm^{-1}, for the syn- and anti-isomer, respectively. The VC-MP2 at DF-MP2/cc-pVDZ level yields 2194 and 2185 cm^{-1} for these same isomers.

An interesting result is observed when the dual level calculations are performed with systematically improved diagonal PES for the C–D stretch vibrational mode. With the improved basis set (for all atoms same basis sets are used) for the diagonal PES, the C–D anharmonic vibrational frequency exhibits a red shift. The frequency change is rather large for the DZ to TZ basis change, and it converges quickly for larger basis sets. For example, when the diagonal potential for C–D (in anti-o-DMB) vibrational mode is calculated with augmented basis sets, a 32 cm^{-1} frequency shift is found going from AVDZ to AVTZ. On the other hand, a basis set improvement from AVTZ to AVQZ shows just 1 cm^{-1} frequency shift (see Table 4.3). When the 1D PES is generated by extrapolating the total energy (the HF energy by exponential extrapolation and correlation energy by the X5 extrapolation), the C–D frequency is lowered by 4 cm^{-1}. The C–D shows also a red shift when highly correlated methods are used. A non-correlated VSCF calculation where the diagonal potential for the C–D vibrational mode is calculated at the DF-L-CCSD(T)/cc-pVTZ level, yields C–D frequencies of 2161 cm^{-1} and 2171 cm^{-1} for the syn- and anti-o-DMB respectively. When the DF-L-CCSD(T) calculation is performed using larger basis sets (cc-pVQZ) only for the deuterium and the directly connected carbon (for all other atoms cc-pVTZ basis are used), the frequency is lowered further by 5 cm^{-1} with respect to the DF-L-CCSD(T)/cc-pVTZ result (see Table 4.3).

A detailed analysis shows that only a few modes are coupled with the C–D stretch vibrational mode. Table 4.4 shows the coupling effect of different modes on the C–D stretch vibrational mode. It is observed that modes 17 and 19 which correspond to the

Table 4.3: The C–D vibrational frequency for anti-o-deutero-methyl-benzoate. DF-MP2/cc-pVDZ diagonal and the pair potentials are used for all modes. The diagonal potential for the C–D stretch vibrational mode alone replaced by the C–D stretch PES calculated at different level of theory and employing different size of basis sets. Selected PM3 triple points are (i=17, j, k=41, where j = 12, 13, 14, 15, 16, 18, 20, 23) are also used for triple level of computation.

isotope	Method	Basis	Diagonal	SCF	V-MP2	VC-MP2
	MP2	DZ	2367	2233	2243	2247
syn	MP2	TZ	2336	2187	2210	2210
o-DMB	MP2	ATZ	2198	2180	2202	2201
	CCSD(T)	TZ	2308	2161	2185	2184
	MP2	DZ	2374	2242	2257	2257
	MP2	TZ	2343	2196	2220	2219
	MP2	QZ	2338	2190	2214	2214
	MP2	ADZ	2359	2221	2239	2239
anti	MP2	ATZ	2235	2189	2212	2212
o-DMB	MP2	AQZ	2335	2188	2212	2211
	Ext	AVXZ	2331	2183	2208	2207
	CCSD(T)	TZ	2315	2171	2194	2194
	CCSD(T)	T/QZ	2310	2166	2190	2189
	PM3	selected triple		2172	2197	2197

C–D bending modes, are strongly coupled with the C–D stretching mode. Also some other modes like 12, 13, 14, 15, 16, 18, 21, 24, have some weak coupling with the C–D stretch vibrational mode (see Table 4.4). All these modes are involved with some kind of C–D bending motions (see Table 4.10). All other modes have negligible influence on the C–D stretching mode. It is therefore to be expected that higher level calculations for these selected coupling potentials will further improve the result for the C–D stretch vibrational frequency.

One might expect that higher order terms in the many-body expansion of the PES have a significant influence on the vibrational frequency of the anharmonic system. A recent analysis[27] shows though that only those triple couplings contribute most to the anharmonic frequencies which involve modes that also exhibit large mutual pair-couplings. The PM3 triple coupling potentials were used for such an analysis. Selected PM3 level triple couplings were added to the DF-MP2/cc-pVDZ pair coupling. The lowers the C–D stretching frequency negligibly (see Table 4.5) indicating a minor influence of triple couplings for this particular mode.

Table 4.4: Strongly coupled modes with C–D vibrational mode for anti-o-deutero-methyl-benzoate. Although the frequency differ a little amount, but the qualitative result does not depend upon the isotopomers.

Coupled modes	SCF	V-MP2	VC-MP2
all decoupled	2316.40	2316.44	2316.44
17	2287.35	2295.99	2295.85
19	2294.76	2294.72	2294.72
17,19	2265.62	2274.69	2274.52
12,17,19	2243.13	2258.98	2258.71
12,14,17,19	2232.27	2249.85	2249.54
12,14,15,17,19	2226.44	2244.31	2243.99
12,14,15,17,19,21	2219.96	2238.80	2238.46
12,14,15,17,19,21,24	2215.64	2234.14	2233.79
12,13,14,15,17,19,21,24	2211.43	2230.34	2229.98
12,13,14,15,16,17,19,21,24	2204.51	2224.82	2224.43
12,13,14,15,16,17,18,19,21,24	2199.23	2220.17	2219.76
8,12,13,14,15,16,17,18,19,21,24	2192.27	2214.07	2213.63
7,8,12,13,14,15,16,17,18,19,21,24	2187.93	2209.99	2209.54
7,8,12,13,14,15,16,17,18,19,21,24,30	2183.82	2205.96	2205.49
7,8,9,12,13,14,15,16,17,18,19,21,24,30	2180.06	2202.65	2202.16

Table 4.5: Strongly coupled PM3 triple contributions for C–D vibrational mode of anti-o-deutero-methyl-benzoate. Although the frequency differ a little amount, but the qualitative result does not depend upon the isotopomers.

	Coupled modes	SCF	V-MP2	VC-MP2
	all decoupled	2003.2	2038.9	2038.3
17,j,40	19	2167.98	2192.07	2191.50
17,j,40	12,19	2168.40	2192.41	2191.84
17,j,40	12,14,19	2168.01	2192.01	2191.44
17,j,40	12,14,15,19	2172.64	2196.65	2196.09
17,j,40	12,14,15,19,21	2172.16	2196.17	2195.60
17,j,40	12,14,15,19,21,24	2173.87	2198.86	2198.29
17,j,40	12,13,14,15,19,21,24	2173.06	2198.59	2198.02
17,j,40	12,13,14,15,16,19,21,24	2172.46	2198.02	2197.44
17,j,40	12,13,14,15,16,18,19,21,24	2171.63	2197.20	2196.62

There are two strong peaks observed in the linear IR absorption spectra (see Fig. 4.20) which are equally separated from the calculated C–D stretch frequency. It seems that the

C–D stretch frequency mode is coupled with some low vibrational modes and thus it is shifted equally in both the directions and appears as two peaks. Vibrational CI calculation may be necessary to describe this feature.

C=O band

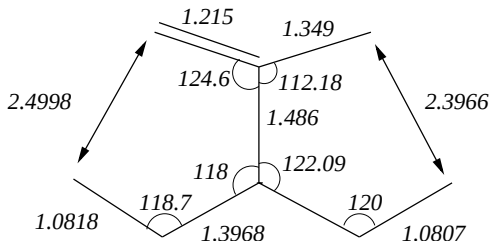

Figure 4.21: Possibility to form five member rings between ortho hydrogen of the phenyl ring and the ester group. Bond lengths are given in Å and bond angles in degree.

Table 4.6: Anharmonic vibrational frequency for the C=O stretch mode calculated at dual level VSCF method with DF-L-CCSD(T)/cc-pVTZ diagonal and DF-MP2/cc-pVDZ pair potentials.

isotope	SCF	V-MP2	VC-MP2	Experimental
syn-o-DMB	1724.5	1724.4	1724.5	
anti-o-DMB	1724.3	1724.1	1724.1	1724
MB	1725.6	1727.2	1727.2	

Due to the structural arrangement, there is a possibility to form a strained five membered ring between the ester and the phenyl ring, which may induce an additional coupling between the C=O and the C–D band. The possible structure is shown in Fig. 4.21. Both the oxygen atoms in the ester group may take part in five membered rings along with the two ortho hydrogens of the phenyl ring. The calculated possible non-bounded distances are 2.499 Å and 2.397 Å for the syn- and anti- positions, respectively, which are much larger than the required bond length to form a true five membered ring. Therefore, isotopic substitution at the ortho position in the phenyl ring does not have any influence on the C=O vibrational frequency. The C=O vibrational frequency remains largely unchanged for both syn- and anti-o-DMB. Calculated anharmonic frequencies in dual level calculations yield the same frequency of 1724 cm^{-1} for both syn- and anti- conformers, which is in agreements with the experimental results.

4.4.4 Anharmonicity observed in the VSCF calculations

Table 4.7: Diagonal anharmonicity for vibrational frequencies of Methyl benzoate and its two isotopomers are calculated at the non-correlated VSCF method.

Mode	Isotopomer			Mode	Isotopomer		
No.	MB	syn-	anti-	No.	MB	syn-	anti-
1	17.67	18.98	17.57	25	13.86	13.19	12.17
2	17.45	16.79	16.67	26	9.72	5.72	1.34
3	83.80	78.25	79.38	27	10.37	10.25	9.97
4	4.16	4.14	3.83	28	22.96	22.88	19.70
5	9.05	8.95	8.80	29	5.65	6.16	5.88
6	3.78	3.77	3.73	30	-3.19	-8.34	-6.33
7	-5.71	-2.53	-2.24	31	4.96	7.17	4.45
8	3.66	4.42	3.56	32	1.51	10.88	9.98
9	5.20	5.81	5.30	33	5.02	4.20	6.72
10	0.29	0.73	0.79	34	3.81	6.24	5.20
11	1.94	1.92	1.92	35	1.17	0.93	1.22
12	-0.47	-0.92	-1.31	36	-10.23	-10.05	-10.13
13	21.11	21.71	19.16	37	2.92	3.39	2.37
14	15.04	14.94	15.46	38	-2.05	-1.32	-1.85
15	1.32	-0.12	0.30	39	0.04	0.93	0.26
16	7.82	7.00	5.51	40	-6.85	-7.09	-6.60
17	25.71	4.03	3.69	41	-120.76	-28.66	-29.07
18	1.58	9.65	9.94	42	-278.75	-123.61	-123.67
19	11.41	12.56	15.04	43	-155.03	-214.25	-216.96
20	3.32	-5.11	-5.94	44	-198.70	-170.12	-124.85
21	26.03	29.62	24.43	45	-88.52	-114.61	-68.48
22	5.73	1.67	3.41	46	-102.29	-90.28	-87.35
23	21.06	20.36	21.55	47	-88.28	-87.37	-85.05
24	-2.10	0.10	4.42	48	-67.71	-70.31	-94.47

The diagonal anharmonicities in frequency calculations are presented in Table 4.7 where mode numbers are based on the harmonic normal mode analysis. The diagonal anharmonicities are very high for the high frequency C–H stretch vibrational modes, up to a few hundred wavenumbers. Other than the C–H stretch vibrational modes, diagonal anharmonicities are rather small (< 10 cm^{-1}). The calculated anharmonicities excied the value of 10 cm^{-1} only for some low frequency modes. These vibrational modes are mostly involved with C–H bending motions. The assignment of these modes is presented in Table 4.8.

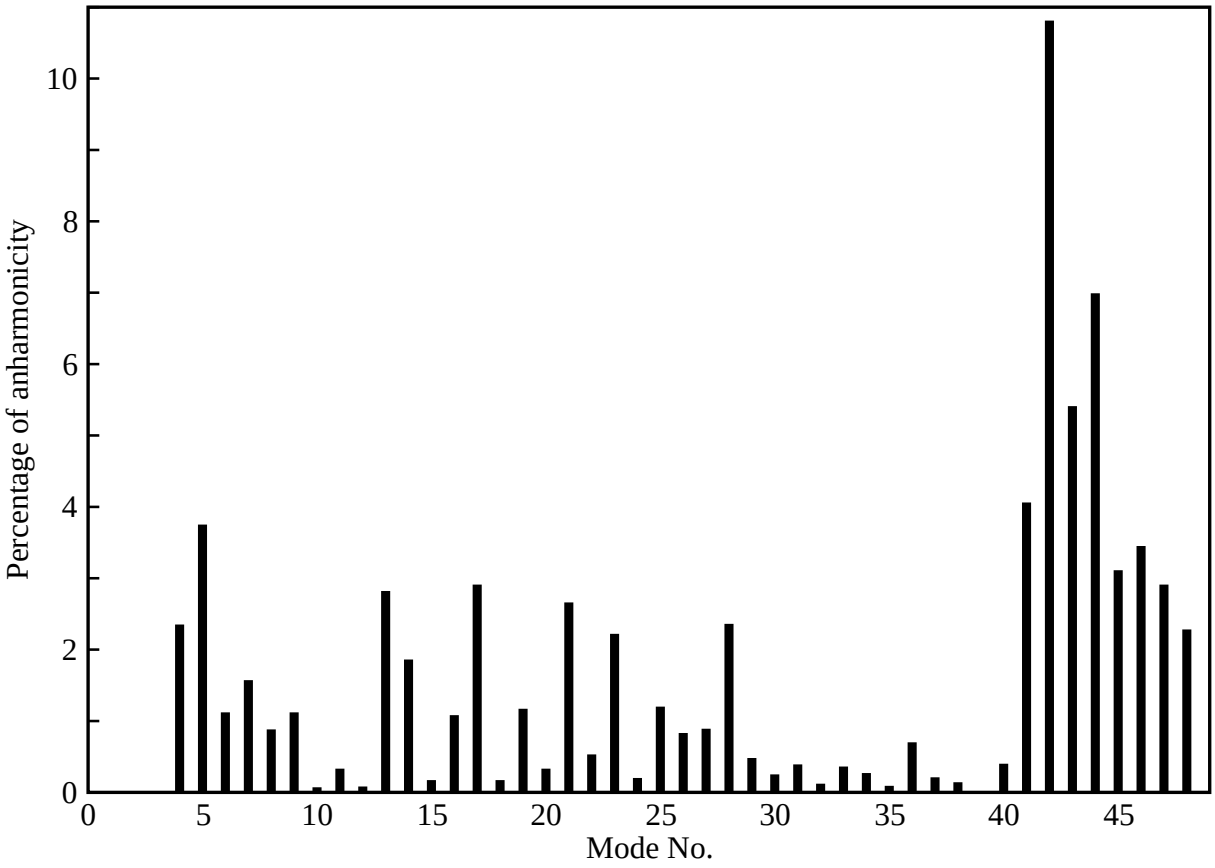

Figure 4.22: Percentage of anharmonicity vs. the mode number for the fundamental transitions of Methyl benzoate

The relative anharmonicities with the mode numbers are plotted in Fig. 4.22. For the first three modes the calculated vibrational frequencies are not reliable and for that reason anharmonicities are also unreliable and omitted from the plot. As seen in the Fig. 4.22, the relative anharmonicities are very high for C–H stretch vibrational modes. Especially mode 42, which is unexpectedly red shifted, shows maximum percentage of anharmonicity $> 10\%$. For most of the low frequency modes anharmonicities are rather small.

Off-diagonal anharmonicity

Off-diagonal anharmonicities calculated (see Section 4.2.6) for the few spectroscopically most important modes are presented in Table. 4.9. Calculated off-diagonal anharmonicities (see Fig. 4.5) for the C=O vs. C–D coupling modes for both the isotopomers are less than a wave number. Such a negligibly small off-diagonal anharmonicity also indicates

Table 4.8: The modes for which diagonal anharmonicities are observed larger than 10 cm^{-1}.

Mode	Isotopomer			Assignment
	MB	syn-o-DMB	anti-o-DMB	
13	754	708	704	phenyl opd O=C-O opb
14	813	829	834	phenyl opd O=C-O opb
17	891			phenyl opr
18	978	759	771	phenyl ipd, ester ipd
19	979	973	973	phenyl ipd
21	986	984	984	phenyl opd, P-CH opb
23	958	956	980	phenyl opd, P-CH opb
25	1167	1171	1173	pCH ipb
26	1182	1180	1181	CH_3 sopb, C-O-C opb
27	1183	1149	1141	pCH sipb
28	979	969	984	pCH opb
36	1485	1485	1485	CH_3 sb

Symbols: opr = out-of-plane rotation, ipb = in-plane bending, opb = out-of-plane bending, ipd = in-plane deformation, opd = out-of-plane deformation, sb = symmetric bending, sipb = symmetric in plane bending, pCH = phenyl CH bond.

Table 4.9: Off diagonal anharmonicities for the C=O and C–D coupling modes calculated with non-correlated VSCF method with DF-MP2/AVTZ diagonal and DF-MP2/VDZ pair potentials.

Coupling	level	ω_a	ω_b	ω'	$\Delta\omega$
CO CD_7	SCF	1771.03	2233.49	4003.846	0.67
	V-MP2	1768.81	2242.75	4012.076	-0.52
	VC-MP2	1768.81	2246.59	4015.142	0.36
CO CD_{11}	SCF	1770.62	2242.33	4012.771	0.18
	V-MP2	1768.75	2257.22	4025.613	0.36
	VC-MP2	1768.75	2257.22	4025.654	0.32

that there is negligible coupling between the C=O and the C–D vibrational modes.

4.4.5 Assignment of vibrational frequencies

Table 4.10 represents a complete assignment of the vibrational frequencies of Methyl benzoate along with the deuterated species. The normal mode along with the mode number 17 and 20 in deuterated species are completely different from the non-deuterated

species. Frequencies for these two modes are kept blank for deuterated species in the Table 4.10.

Table 4.10: Assignment of the vibrational frequencies of Methyl-Benzoate and its two deutorated species.

Mode	Isotopomer			Assignment
	MB	syn-o-DMB	anti-o-DMB	
1	112	111	111	Phenyl-ester opr
2	196	197	198	Phenyl opb, ester opb
3	342	344	343	CH_3 asb
4	179	177	177	Phenyl-ester ipr
5	242	243	242	Phenyl-ester opb
6	340	339	340	Phenyl-Ester ipr
7	363	361	361	O=C-O ipb, phenyl ipb
8	421	402	403	phenyl opb
9	468	458	461	phenyl opb, O=C-O-C opb
10	477	474	471	phenyl ipr, O=C-O-C ipr
11	617	612	612	phenyl ipd
12	681	678	680	phenyl ipd O=C-O ipb
13	754	708	704	phenyl opd O=C-O opb
14	813	829	834	phenyl opd O=C-O opb
15	823	822	821	O=C-O ipb, phenyl, CH_3 ipb
16	731	734	725	phenyl opd
17	891			phenyl opr
18	978	759	771	phenyl ipd, ester ipd
19	979	973	973	phenyl ipd
20	1035			phenyl ipd, P-CH sipb
21	986	984	984	phenyl opd, P-CH opb
22	1096	1047	1048	phenyl ipd, P-CH ipb
23	958	956	980	phenyl opd, P-CH opb
24	1126	1112	1121	OCH_3 r, O=C-O ipb

continued to next page

continued from previous page

Mode	Isotopomer			Assignment
	MB	syn-o-DMB	anti-o-DMB	
25	1167	1171	1173	pCH ipb
26	1182	1180	1181	CH_3 sopb, C-O-C opb
27	1183	1149	1141	pCH sipb
28	979	969	984	pCH opb
29	1214	1212	1213	OCH_3 r, O=C s, C-O s
30	1303	1298	1307	OCO ipb pCH ipb
31	1320	1268	1259	pCH ipb
32	1447	1444	1447	CH_3 asb
33	1342	1348	1345	pCC s
34	1452	1445	1455	pCC s
35	1453	1454	1453	CH_3 wagging
36	1485	1485	1485	CH_3 sb
37	1498	1480	1483	pCC s, pCH ipb
38	1606	1602	1600	phenyl ipd-sy
39	1603	1598	1604	phenyl ipd-sy
40	1725	1724	1724	C=O s
41	2974	2974	2973	mCH ss
42	2560	2561	2559	mCH as
43	2855	2809	2889	pCH as
44	2834	2171	2971	pCH as
45	2832	2858	2832	mCH ss
46	2963	2831	2957	pCH as
47	3039	3017	3065	pCH as
48	2985	2989	2161	pCH as

Symbols: R = rotation, r = rocking, ipr = in-plane rotation, opr = out-of-plane rotation, ipb = in-plane bending, opb = out-of-plane bending, ipd = in-plane deformation, opb = out-of-plane deformation, as = asymmetric stretch, ss = symmetric stretch, asb = asymmetric bending, sb = symmetric bending, sipb = symmetric in plane bending, mCH = methyl CH bond, pCH = phenyl CH bond, sy = symmetric.

Chapter 5

Conclusions

The purpose of the present work is to improve the efficiency of the computational methods to explore the molecular properties more efficiently and accurately and then use the improved computational method to study a model system. The exploration of dynamical properties of molecules depends upon the accurate calculation of the PES along different vibrational normal modes. In the first part of the present work we developed a computational technique to generate the PES more accurately and more efficiently with a reasonable expense.

We have performed a systematic study of the convergence behavior of the MP2 and CCSD(T) correlation energies with basis set size. The proposed X5 method proved to be an efficient method to estimate the basis set limit correlation energy from the DZ-TZ basis sets, which is much better in accuracy than the 5Z basis set result. It reduces the computational cost by two to three orders of magnitude to achieve this accuracy, which is a remarkable achievement for the high quality ab initio calculation. The X5 method also proved its efficiency to calculate basis set limit correlation energy for larger molecules than small molecules and atoms.

The proposed extrapolation method (X5 method) needs two model parameters (A_1 and A_2) to estimate the basis set limit correlation energy which show an exponential correlation. The global parameter based X5 method is specific to the computational method as well as family of basis set. The A_1 vs. A_2 fitting curves for a set of atoms and molecules consisting the first and second row elements differ substantially between MP2 and CCSD(T). This indicates that mixing of data from different levels of computation affects the performance of the X5 method negatively and that parameters optimized for one level can not be used for the other. It is also necessary to optimize the exponential fitting parameters separately for different families of basis sets as the correlation of A_1

and A_2 differs substantially.

The X5 method gradually under-estimates the basis set limit energy as the number of hydrogen atoms increases in the molecule. This has been investigated systematically for carbon and nitrogen based molecules with different proportions of hydrogen atoms. Although the A_2 parameter for the hydrogen molecule is quite off-set from the other atoms and molecules and its absence in the optimization of correlation parameters deteriorates the performance of the X5 method for hydrogen dominated molecules.

The PES generation has been carried out with the X5 method for several molecules for their different rotational and vibrational motions, and in particular, for all 30 normal modes of the NMA molecule. An excellent performance of the X5 method has been observed with the global parameter set, but it always includes an intrinsic error in the PES. It has been observed that the A_1 and A_2 parameters are slowly varying parameters w.r.t. internal coordinate of the molecule and to a good approximation they are constant within the system for small deviations in the molecular geometry. The constant A_1 and A_2 parameters calculated at the equilibrium geometry (A_1^0 and A_2^0) may be used in the extrapolation of correlation energies with the X5 method for all other configurations in the PES. This approach yields an improved PES just performing DZ basis set calculations. The local parameter based X5 extrapolation removes the intrinsic error of the global parameter method and nearly reaches to the spectroscopic accuracy (< 0.5 mE$_h$).

The convergence behavior of the correlation part of the gradient and Hessian have been analyzed with the MP2 and CCSD(T) methods for the water molecule. A rather small change has been observed in the correlation part of the gradient and Hessian with the basis set size. It appears that basis set extrapolation is not necessary but a higher correlation method may provide more accurate gradient and Hessian. This accurate gradient and Hessian can be used to generate a smooth PES from non-uniform energy points by modified Shepard interpolation.

Such highly accurate PESs have been used in the harmonic and anharmonic vibrational analysis of Methyl benzoate and its sys- and anti-ortho-deutero isotopomers.

It has been observed that harmonic frequencies are a very poor approximation to assign the vibrational frequencies of Methyl benzoate and its two isotopomers. Anharmonic diagonal frequencies show some improvements but were still not sufficient to reach a reasonable assignment. The fundamental transition frequencies calculated with the VSCF method from a pair potential energy surface expansion seem promising, especially when the 2D PES are calculated at a sufficiently high ab initio level. The success of the VSCF

frequency calculations depends upon the accuracy of the PES, in particular near the equilibrium where denser grid points are required. Dual level computations in which the diagonal anharmonic potential along a single vibrational mode is calculated using higher level ab initio methods than for coupling potentials provide an efficient route to the computation of the PES expansion in the VSCF framework. Such a dual level VSCF calculation with DF-L-CCSD(T)/cc-pVTZ diagonal and DF-MP2/cc-pVDZ pair coupling potentials provided a nearly perfect agreement of the C=O vibrational frequency with respect to the experimental result.

Carrying out a systematic study, we have shown that not all pair couplings are necessary to describe a particular vibrational band. Using only 14 coupling potentials for the C–D stretch vibrational mode the computed frequency is in good agreement with the result based on the full set of 1128 couplings. Such a reduced coupling potential may provide a novel method to study even larger systems with reasonable computational cost. Non-uniform IMLS interpolation has been successfully used to reduce the computational cost for potential energy surface generation even further.

An unexpected red shift has been observed for a C–H stretch vibrational mode when the VSCF calculation has been performed with pair coupling potentials. The inability to find this mode in harmonic and diagonal anharmonic calculations indicates that this is a concerted anharmonic effect and pair and higher order couplings are in fact necessary to understand this feature. Our investigations for the C–H(D) stretch vibrational modes with selected triple couplings at PM3 level do not improve the results much over the ab initio pair coupling calculations. This could indicate that either PM3 triple coupling potentials are in sufficient or that higher order coupling effects are negligible in this model.

The correct assignment of the C–D stretch frequency still poses a problem to both the theoretician and experimentalist. The calculated C–D stretch frequency just sits between the two strong peaks observed in the linear IR absorption spectra at the expected C–D frequency region. It seems that the C–D stretch frequency mode is coupled with some low vibrational frequency modes and thus it is shifted equally in both directions and appears as two peaks. Vibrational CI calculation may necessary to describe this feature. Still now, the best possible calculation identifies the syn- and anti- ortho deutero Methyl benzoate by a frequency difference of 10 cm^{-1} for the C–D stretch vibrational mode.

The diagonal and the off-diagonal anharmonicities have been calculated by non-correlated VSCF method. Other than for a few low frequency modes, which are involved in the C–H bending modes, diagonal anharmonicities are very small. The negligible off-diagonal

anharmonicity for C=O and C–D coupling modes indicates that these modes are mostly decoupled.

Outlook

The proposed extrapolation method (X5 method) accounts up to the fifth oder term from the Schwartz formula. Inclusion of higher order terms may improve the efficiency of the extrapolation. Although A_2 is a slowly varying parameter for a small perturbed system, further improvements on the PES calculations are made feasible by interpolation of the A_2 parameter if there is any simple relation between the A_2 and the geometry of the molecule. Similarly better gradient and Hessian calculations may also be possible by basis set extrapolation using interpolated A_2 parameter.

The description of strongly anharmonic modes and especially modes exhibiting large amplitude motions still poses a problem in the VSCF framework. These are mainly C–H stretching modes of the phenyl ring and methyl group, C–C stretching modes in the phenyl ring, and some other low frequency modes. Diagonal potentials at a higher level of theory and inclusion of selected triple coupling potentials may solve the problem. It may be possible to describe the C–D band much better by performing the CI calculation for the C–D stretch vibration mode.

Appendix A

A.1 Calculation of A_1 and A_2

After obtaining E_∞^{cor} from Eq. (3.48), only two unknown parameters remain in Eq. (3.49): A_1 and A_2. We can calculate them with two consecutive basis sets. Let the cardinal number of two consecutive basis sets be X and Y, then Eq. (3.49) becomes

$$\frac{E_X^{cor} - E_\infty^{cor}}{E_\infty^{cor}} = A_1 X^{-3} + A_1 A_2 X^{-4} \tag{A.1.1}$$

$$\frac{E_Y^{cor} - E_\infty^{cor}}{E_\infty^{cor}} = A_1 Y^{-3} + A_1 A_2 Y^{-4} \tag{A.1.2}$$

Multiplying eq.(A.1.1) by Y^{-4} and eq.(A.1.2) by X^{-4} we get,

$$\frac{E_X^{cor} - E_\infty^{cor}}{E_\infty^{cor}} Y^{-4} = A_1 X^{-3} Y^{-4} + A_1 A_2 X^{-4} Y^{-4}$$

$$\frac{E_Y^{cor} - E_\infty^{cor}}{E_\infty^{cor}} X^{-4} = A_1 X^{-4} Y^{-3} + A_1 A_2 X^{-4} Y^{-4}$$

therefore,

$$A_1 = \frac{E_{Xd} Y^{-4} - E_{Yd} X^{-4}}{X^{-3} Y^{-4} - X^{-4} Y^{-3}} \tag{A.1.3}$$

where,

$$E_{Xd} = \frac{E_X^{cor} - E_\infty^{cor}}{E_\infty^{cor}}$$

$$E_{Yd} = \frac{E_Y^{cor} - E_\infty^{cor}}{E_\infty^{cor}}$$

$$A_2 = \frac{E_{Xd} - A_1 X^{-3}}{A_1 X^{-4}} \tag{A.1.4}$$

A.2 The gradient and Hessian extrapolation with the X5 method

For the two basis sets of cardinal number X and Y, the X5 method can be written as,

$$E_X = E_\infty\{1 + A_1X^{-3}(1 + A_2X^{-1})\}, \tag{A.2.5}$$

$$E_Y = E_\infty\{1 + A_1Y^{-3}(1 + A_2Y^{-1})\}. \tag{A.2.6}$$

Now eliminating A_1 from above two equations we get,

$$\frac{E_X - E_\infty}{E_Y - E_\infty} = \frac{X^{-3}(1 + A_2X^{-1})}{Y^{-3}(1 + A_2Y^{-1})}.$$

$$\therefore \quad E_XY^{-3}(1+A_2Y^{-1})-E_\infty Y^{-3}(1+A_2Y^{-1}) = E_YX^{-3}(1+A_2X^{-1})-E_\infty X^{-3}(1+A_2X^{-1})$$

$$\therefore \quad E_\infty\{X^{-3}(1+A_2X^{-1})-Y^{-3}(1+A_2Y^{-1})\} = E_YX^{-3}(1+A_2X^{-1})-E_XY^{-3}(1+A_2Y^{-1})$$

$$\therefore \quad E_\infty = \frac{E_YX^{-3}(1 + A_2X^{-1}) - E_XY^{-3}(1 + A_2Y^{-1})}{X^{-3}(1 + A_2X^{-1}) - Y^{-3}(1 + A_2Y^{-1})}$$

$$\therefore \quad E_\infty = \frac{(E_YX^{-3} - E_XY^{-3}) + A_2(E_YX^{-4} - E_XY^{-4})}{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})} \tag{A.2.7}$$

Differentiating both sides of Eq.-(A.2.7) with respect to the coordinate q_i we get,

$$\frac{\partial E_\infty}{\partial q_i} = \frac{\frac{\partial E_Y}{\partial q_i}X^{-3} - \frac{\partial E_X}{\partial q_i}Y^{-3} + \frac{\partial A_2}{\partial q_i}(E_YX^{-3} - E_XY^{-3}) + A_2(\frac{\partial E_Y}{\partial q_i}X^{-4} - \frac{\partial E_X}{\partial q_i}Y^{-4})}{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})}$$

$$- \frac{\{(E_YX^{-3} - E_XY^{-3}) + A_2(E_YX^{-4} - E_XY^{-4})\}\frac{\partial A_2}{\partial q_i}(X^{-4} - Y^{-4})}{\{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})\}^2}, \tag{A.2.8}$$

since A_2 is geometry dependent.

Differentiating both side of Eq.-(A.2.8) again with respect to q_j we get,

$$\begin{aligned}
\frac{\partial^2 E_\infty}{\partial q_i \partial q_j} &= \frac{\frac{\partial^2 E_Y}{\partial q_i \partial q_j} X^{-3} - \frac{\partial^2 E_X}{\partial q_i \partial q_j} Y^{-3} + \frac{\partial^2 A_2}{\partial q_i \partial q_j}(E_Y X^{-4} - E_X Y^{-4}) + \frac{\partial A_2}{\partial q_i}(\frac{\partial E_Y}{\partial q_j} X^{-4} - \frac{\partial E_X}{\partial q_j} Y^{-4})}{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})} \\
&+ \frac{\frac{\partial A_2}{\partial q_j}(\frac{\partial E_Y}{\partial q_i} X^{-4} - \frac{\partial E_X}{\partial q_i} Y^{-4}) + A_2(\frac{\partial^2 E_Y}{\partial q_i \partial q_j} X^{-4} - \frac{\partial^2 E_X}{\partial q_i \partial q_j} Y^{-4})}{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})} \\
&- \frac{\{\frac{\partial E_Y}{\partial q_i} X^{-3} - \frac{\partial E_X}{\partial q_i} Y^{-3} + \frac{\partial A_2}{\partial q_i}(E_Y X^{-3} - E_X Y^{-3}) + A_2(\frac{\partial E_Y}{\partial q_i} X^{-4} - \frac{\partial E_X}{\partial q_i} Y^{-4})\}\frac{\partial A_2}{\partial q_j}(X^{-4} - Y^{-4})}{\{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})\}^2} \\
&- \frac{\{\frac{\partial E_Y}{\partial q_j} X^{-3} - \frac{\partial E_X}{\partial q_j} Y^{-3} + \frac{\partial A_2}{\partial q_j}(E_Y X^{-3} - E_X Y^{-3}) + A_2(\frac{\partial E_Y}{\partial q_j} X^{-4} - \frac{\partial E_X}{\partial q_j} Y^{-4})\}\frac{\partial A_2}{\partial q_i}(X^{-4} - Y^{-4})}{\{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})\}^2} \\
&- \frac{\{(E_Y X^{-3} - E_X Y^{-3}) + A_2(E_Y X^{-4} - E_X Y^{-4})\}\frac{\partial^2 A_2}{\partial q_i \partial q_j}(X^{-4} - Y^{-4})}{\{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})\}^2} \\
&+ \frac{\{(E_Y X^{-3} - E_X Y^{-3}) + A_2(E_Y X^{-4} - E_X Y^{-4})\}\frac{\partial A_2}{\partial q_i}(X^{-4} - Y^{-4})\frac{\partial A_2}{\partial q_j}(X^{-4} - Y^{-4})}{\{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})\}^3}
\end{aligned} \tag{A.2.9}$$

Since A_2 is a slowly varying parameter with respect to the geometry, for a small perturbation the first and second derivatives of A_2 essentially become zero, and Eqs.-(A.2.8) and (A.2.9) become,

$$\frac{\partial E_\infty}{\partial q_i} = \frac{\frac{\partial E_Y}{\partial q_i}X^{-3} - \frac{\partial E_X}{\partial q_i}Y^{-3} + A_2(\frac{\partial E_Y}{\partial q_i}X^{-4} - \frac{\partial E_X}{\partial q_i}Y^{-4})}{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})}, \tag{A.2.10}$$

and

$$\frac{\partial^2 E_\infty}{\partial q_i \partial q_j} = \frac{\frac{\partial^2 E_Y}{\partial q_i \partial q_j}X^{-3} - \frac{\partial^2 E_X}{\partial q_i \partial q_j}Y^{-3} + A_2(\frac{\partial^2 E_Y}{\partial q_i \partial q_j}X^{-4} - \frac{\partial^2 E_X}{\partial q_i \partial q_j}Y^{-4})}{(X^{-3} - Y^{-3}) + A_2(X^{-4} - Y^{-4})}. \tag{A.2.11}$$

Eqs.-(A.2.10) and (A.2.11) are the expression for the gradient and the Hessians at basis set limit.

Appendix B

B.1 Convergence of diagonal frequency with grid size.

An appropriately grid size is crucial for the success of anharmonic frequency calculation, spatially a dense energy points are required near the equilibrium configuration. A systematic study have been carried out with 6 points, 8 points and 12 points diagonal grid (see Fig. 4.7). The improvement of frequency calculations are presented in Table B.1. A significant frequency change is found going from 6 points grid to 8 points grid. On the other hand when calculations are performed with 12 points grid a very small change in frequency are observed. It indicates that 8 points grid which contain more points near equilibrium (see Fig. 4.7) is sufficient for the reasonable description of the 1D PES in anharmonic vibrational calculation.

Table B.1: Diagonal frequencies of Methyl benzoate with different grid size and PES calculated at MP2/AVTZ level. For only 6 and 12 points grid the frequency differences from 8 points grid are presented here.

Mode	Methyl benzoate			syn-o-DMB			anti-o-DMB		
	8p	8p-6p	8p-12p	8p	8p-6p	8p-12p	8p	8p-6p	8p-12p
1	82	2.66	0.45	82	2.52	0.37	81	2.52	0.36
2	154	1.10	1.65	153	0.92	1.45	154	0.92	1.46
3	361	1.29	3.66	360	0.87	3.24	360	0.87	3.24
4	170	1.39	0.21	168	1.21	0.03	168	1.21	0.03
5	217	2.52	0.39	216	2.35	0.19	215	2.33	0.19
6	334	0.55	0.65	331	0.22	0.27	332	0.21	0.32
7	359	-0.15	0.49	356	-0.49	0.14	357	-0.50	0.14

continued to next page

continued from previous page

Mode	Methyl benzoate			syn-o-DMB			anti-o-DMB		
	8p	8p-6p	8p-12p	8p	8p-6p	8p-12p	8p	8p-6p	8p-12p
8	411	1.87	0.44	394	1.69	-0.00	394	1.59	-0.01
9	459	1.39	0.51	452	1.05	0.05	452	1.05	0.04
10	482	0.55	0.48	478	0.07	-0.01	475	0.08	-0.00
11	618	0.86	0.65	612	0.24	0.03	612	0.24	0.03
12	679	0.87	0.69	695	4.50	0.22	691	4.67	0.25
13	728	4.67	0.82	676	0.21	0.01	676	0.20	0.01
14	802	3.35	0.76	743	0.85	-0.04	734	0.62	-0.04
15	831	1.55	0.67	829	0.56	-0.12	827	0.58	-0.12
16	761	0.29	0.23	816	4.11	0.04	820	3.67	0.08
17	824	-62.16	1.34	858	-13.11	0.16	864	-16.08	0.09
18	1142	138.32	1.63	816	40.48	0.02	818	36.99	0.03
19	1057	44.68	1.08	1119	123.33	0.04	1105	107.31	0.04
20	1145	96.76	1.32	1053	49.42	0.41	1071	69.54	0.45
21	934	-59.24	1.01	874	-47.36	0.13	878	-54.24	0.26
22	1159	59.49	1.16	1159	96.59	0.17	1155	93.09	0.20
23	904	-57.57	0.95	927	-65.23	0.20	928	-54.88	0.08
24	1220	76.64	1.54	1181	56.12	0.14	1177	47.44	0.03
25	1150	-33.77	1.18	1194	31.38	0.11	1205	49.51	0.25
26	1032	-161.36	1.39	1153	-34.30	0.02	1151	-34.33	0.03
27	1178	-17.60	1.17	1029	-161.76	0.30	1031	-162.21	0.31
28	866	-118.46	0.57	884	-101.96	-0.18	878	-108.94	-0.25
29	1148	-76.90	1.92	1147	-77.77	0.79	1147	-78.13	0.76
30	1381	60.48	1.62	1246	-47.76	0.04	1257	-23.29	0.06
31	1287	-46.24	1.33	1419	97.36	0.28	1392	65.16	0.16
32	1322	-145.89	1.82	1327	-140.94	0.39	1327	-140.28	0.49
33	1333	-148.01	1.42	1426	-40.42	0.04	1362	-110.10	0.04
34	1458	-16.45	1.50	1340	-138.72	0.06	1418	-54.59	0.03
35	1351	-158.43	1.42	1349	-159.78	0.07	1349	-159.78	0.07
36	1338	-180.60	1.51	1518	15.54	0.03	1516	15.23	0.03
37	1529	10.10	1.56	1337	-181.61	0.18	1337	-181.95	0.18
38	1759	123.44	1.82	1752	119.97	0.04	1760	130.94	0.06
39	1755	120.81	1.89	1752	125.40	0.10	1748	118.62	0.17
40	1938	175.94	2.42	1937	175.00	0.51	1935	173.28	0.51
41	3040	-36.72	2.70	2196	-148.92	1.40	2205	-148.02	1.42
42	3171	-60.49	4.49	3037	-39.59	-0.28	3037	-39.49	-0.28

continued to next page

continued from previous page

Mode	Methyl benzoate			syn-o-DMB			anti-o-DMB		
	8p	8p-6p	8p-12p	8p	8p-6p	8p-12p	8p	8p-6p	8p-12p
43	3122	-77.38	5.18	3168	-63.71	1.27	3168	-63.71	1.27
44	3150	-85.34	3.43	3109	-85.87	2.18	3177	-54.66	0.64
45	3149	-55.45	5.86	3149	-70.08	1.16	3054	-134.99	0.70
46	3144	-84.94	3.33	3141	-62.75	2.71	3145	-59.56	2.73
47	3038	-164.40	2.88	3037	-161.14	-0.05	3145	-83.93	0.25
48	2944	-227.40	5.38	2942	-229.04	2.55	3003	-196.92	-0.27

B.2 Vibrational frequencies for anti-deutero-oDMB

Table B.2: Vibrational frequencies of anti-o-deutero-methyl-benzoate at DF-MP2/AVTZ level. The RMSD calculated w.r.t. the non-correlated VSCF frequencies at SCS/cc-pVDZ level.

Mode	Harmonic	Diagonal	VSCF		CC-VSCF	Exact	Observed
			PM3	MP2		SCS	[Ref. 157, 165]
1	51	81	145	112	99	111	
2	112	154	177	197	171	198	
3	185	360	365	339	343	343	
4	165	168	169	177	174	177	
5	206	215	233	242	238	242	
6	330	332	325	339	336	340	
7	357	357	355	364	364	361	
8	387	394	401	402	401	403	
9	447	452	471	461	457	461	
10	475	475	475	472	471	471	
11	612	612	612	610	609	612	629
12	672	691	743	702	684	704	686
13	676	676	683	679	677	680	713
14	728	734	748	725	685	725	855
15	828	827	818	821	817	821	820
16	801	820	862	832	841	834	
17	873	864	860	852	848	857	864
18	727	818	835	770	773	771	782
19	998	1105	1106	974	959	973	965
20	1002	1071	1098	987	999	984	980
21	898	878	906	908	906	910	
22	1061	1155	1159	1048	1042	1048	1047

continued to next page

continued from previous page

Mode	Harmonic	Diagonal	VSCF		CC-VSCF	Exact	Observed
			PM3	MP2		SCS	[Ref. 157, 165]
23	963	928	971	975	985	980	
24	1127	1177	1183	1119	1113	1121	1097
25	1154	1205	1204	1140	1132	1141	1161
26	1176	1151	1154	1171	1165	1173	1128
27	1185	1031	1072	1179	1173	1181	1177
28	960	878	915	982	997	984	
29	1220	1147	1165	1207	1202	1213	1192
30	1277	1257	1257	1255	1246	1259	1295
31	1324	1392	1364	1306	1301	1307	1256
32	1471	1327	1323	1434	1431	1345	1310
33	1470	1362	1380	1441	1428	1447	1435
34	1471	1418	1428	1449	1452	1455	
35	1509	1349	1318	1446	1442	1453	1445
36	1500	1516	1517	1478	1481	1483	
37	1519	1337	1321	1479	1476	1485	1474
38	1627	1760	1751	1598	1595	1600	1585
39	1629	1748	1740	1602	1596	1604	1594
40	1767	1935	1917	1771	1769	1777	1724
41	3092	3037	2906	2980	3034	2973	2940
42	3183	3168	2784	2579	2499	2559	2542
43	3207	3177	2889	2902	2874	2889	2855
44	3214	3054	2907	2981	2942	2971	2952
45	3217	3145	2866	2853	2698	2832	2845
46	3222	3145	2913	2968	2929	2957	2998
47	3226	3003	2934	3072	3078	3065	3064
48	2392	2205	2174	2242	2257	2166	2118
RMSD	137	130	76	14	28		

Bibliography

[1] P. Tian, D. Keusters, Y. Suzaki, W. S. Warren, Science 300 (2003) 1553–1555.

[2] R. B. Gerber, B. Brauer, S. K. Gregurick, C. G. M., Phys. Chem. Comm. 21 (2002) 142–150.

[3] C. Scheurer, A. Piryatinski, S. Mukamel, J. Am. Chem. Soc. 123 (2001) 3114–3124.

[4] C. Scheurer, S. Mukamel, J. Chem. Phys. 115 (2001) 4989–5004.

[5] C. Scheurer, S. Mukamel, J. Chem. Phys. 116 (2002) 6803–6816.

[6] C. Scheurer, S. Mukamel, Bull. Chem. Soc. Jpn. 75 (2002) 989–999.

[7] M. Bounouar, C. Scheurer, Chem. Phys. 323 (2006) 87–101.

[8] S. Woutersen, P. Hamm, J. Phys.: Condens Matter 14 (2002) R1035.

[9] O. Golonzka, M. Khalil, N. Demirdöven, A. Tokmakoff, Phys. Rev. Lett. 86 (10) (2001) 2154–2157.

[10] J. B. Asbury, T. Steinel, C. Stromberg, S. A. Corcelli, C. P. Lawrence, J. L. Skinner, M. D. Fayer, J. Phys. Chem. A. 108 (7) (2003) 1107–1119.

[11] T. Steinel, J. B. Asbury, S. A. Corcelli, C. P. Lawrence, J. L. Skinner, M. D. Fayer, Chem. Phys. Lett. 386 (2004) 295–300.

[12] J. B. Asbury, T. Steinel, C. Stromberg, K. J. Gaffney, I. R. Piletic, A. Goun, M. D. Fayer, Phys. Rev. Lett. 91 (23) (2003) 237402.

[13] Z. Bačíc, R. B. Gerber, M. A. Ratner, J. Phys. Chem. 90 (16) (1986) 3606–3612.

[14] J. C. Light, I. P. Hamilton, J. V. Lill, J. Chem. Phys. 82 (3) (1985) 1400–1409.

[15] J. R. Henderson, J. Tennyson, B. T. Sutcliffe, J. Chem. Phys. 98 (9) (1993) 7191–7203.

[16] N. J. Wright, J. M. Hutson, J. Chem. Phys. 110 (2) (1999) 902–911.

[17] J. B. Anderson, J. Chem. Phys. 63 (4) (1975) 1499–1503.

[18] V. Buch, J. Chem. Phys. 97 (1) (1992) 726–729.

[19] R. N. Barnett, K. B. Whaley, J. Chem. Phys. 99 (12) (1993) 9730–9744.

[20] Z. Bačíc, M. Kennedy-Mandziuk, J. W. Moskowitz, K. E. Schmidt, J. Chem. Phys. 97 (9) (1992) 6472–6480.

[21] J. M. Bowman, J. Chem. Phys. 68 (2) (1978) 608–610.

[22] R. B. Gerber, M. A. Ratner, Chem. Phys. Lett. 68 (1979) 195–198.

[23] R. B. Gerber, M. A. Ratner, Adv. Chem. Phys. 70 (1988) 97–132.

[24] J. O. Jung, R. B. Gerber, J. Chem. Phys. 105 (1996) 10332–10347.

[25] R. B. Gerber, J. O. Jung, The vibrational self-consistent field approach and extensions: Method and applications to spectroscopy of large molecules and clusters, in: P. Jensen, P. R. Bunker (Eds.), Computational Molecular Spectroscopy, John Wiley & Sons, 2000, pp. 365–390.

[26] S. Carter, J. M. Bowman, L. B. Harding, Spectrochim. Acta A 53 (8) (1997) 1179–1188.

[27] G. Rauhut, J. Chem. Phys. 121 (2004) 9313.

[28] J. J. P. Stewart, J. Comput. Chem. 10 (2) (1989) 209–220.

[29] J. J. P. Stewart, J. Comput. Chem. 10 (2) (1989) 221–264.

[30] O. L. Polyansky, A. G. Császár, S. Shirin, N. F. Zobov, P. Barletta, J. Tennyson, D. W. Schwenke, P. J. Knowles, Science 299 (2003) 539–542.

[31] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, A. K. Wilson, Chem. Phys. Lett. 286 (1998) 243–252.

[32] W. Klopper, K. Bak, P. Jørgensen, J. Olsen, T. Helgaker, J. Phys. B 32 (1999) R103–R130.

[33] C. Schwartz, Phys. Rev. 126 (3) (1962) 1015–1019.

[34] A. Varandas, J. Chem. Phys. 113 (20) (2000) 8880–8887.

[35] C. Scheurer, T. Steinel, Chem. Phys. Chem. 8 (2007) 503–505.

[36] A. Szabo, N. S. Ostlund, Modern Quantum Chemistry, Dover Publications, INC., 1996.

[37] S. M. Blinder, Am. J. Phys. 33 (6) (1965) 431–443.

[38] C. C. J. Roothaan, Rev. Mod. Phys. 23 (2) (1951) 69–89.

[39] J. C. Slater, Phys. Rev. 36 (1) (1930) 57–64.

[40] B. F. S., Proc. Roy. Soc. [London] A 200 (1950) 542–554.

[41] W. J. Hehre, R. F. Stewart, J. A. Pople, J. Chem. Phys. 51 (6) (1969) 2657–2664.

[42] J. B. Collins, P. v. R. Schleyer, B. J. S., J. A. Pople, J. Chem. Phys 64 (12) (1976) 5142–5151.

[43] J. S. Binkley, J. A. Pople, W. J. Hehre, J. Am. Chem. Soc. 102 (3) (1980) 939–947.

[44] M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, W. J. Hehre, J. Am. Chem. Soc. 104 (10) (1982) 2797–2803.

[45] W. J. Pietro, W. J. Francl, M. M.and Hehre, D. J. DeFrees, J. A. Pople, J. S. Binkley, J. Am. Chem. Soc. 104 (19) (1982) 5039–5048.

[46] K. D. Dobbs, W. J. Hehre, J. Comp. Chem. 7 (3) (1986) 359–378.

[47] K. D. Dobbs, W. J. Hehre, J. Comp. Chem. 8 (6) (1987) 861–879.

[48] K. D. Dobbs, W. J. Hehre, J. Comp. Chem. 8 (6) (1987) 880–893.

[49] R. Ditchfield, W. J. Hehre, J. A. Pople, J. Chem. Phys. 54 (2) (1971) 724–728.

[50] W. J. Hehre, R. Ditchfield, J. A. Pople, J. Chem. Phys. 56 (5) (1972) 2257–2261.

[51] M. S. Gordon, Chem. Phys. Lett. 76 (1) (1980) 163–168.

[52] P. Hariharan, J. Pople, Mol. Phys. 27 (1) (1974) 209–214.

[53] D. E. Woon, T. H. Dunning Jr., J. Chem. Phys. 98 (2) (1993) 1358–1371.

[54] R. A. Kendall, T. H. Dunning Jr., R. J. Harrison, J. Chem. Phys. 96 (9) (1992) 6796–6806.

[55] M. J. Frisch, J. A. Pople, J. S. Binkley, J. Chem. Phys. 80 (7) (1984) 3265–3269.

[56] T. Dunning Jr., J. Chem. Phys. 90 (2) (1989) 1007–1023.

[57] K. A. Peterson, D. E. Woon, T. H. Dunning Jr., J. Chem. Phys. 100 (10) (1994) 7410–7415.

[58] A. K. Wilson, T. v. Mourik, T. H. Dunning Jr., J. Mol. Struc. (THEOCHEM) 388 (1997) 339–349.

[59] E. R. Davidson, Chem. Phys. Lett. 260 (1996) 514–518.

[60] K. A. Brueckner, Phys. Rev. 97 (5) (1955) 1353–1366.

[61] C. Møller, M. S. Plesset, Phys. Rev. 46 (1934) 618–622.

[62] M. S. Gordon, J. A. Pople, M. J. Frisch, Chem. Phys. Lett. 153 (6) (1988) 503–506.

[63] M. J. Frisch, M. S. Gordon, J. A. Pople, Chem. Phys. Lett. 166 (3) (1990) 275–280.

[64] M. J. Frisch, M. S. Gordon, J. A. Pople, Chem. Phys. Lett. 166 (3) (1990) 281–289.

[65] M. Head-Gordon, T. Head-Gordon, Chem. Phys. Lett. 220 (1–2) (1994) 122–128.

[66] S. Saebø, J. Almlöf, Chem. Phys. Lett. 154 (1) (1989) 83–89.

[67] N. C. Handy, H. F. Schaefer III, J. Chem. Phys. 81 (11) (1984) 5031–5033.

[68] R. Krishnan, J. A. Pople, Int. J. Quant. Chem. 14 (1) (1978) 91–100.

[69] R. Krishnan, J. A. Pople, E. S. Replogle, M. Head-Gordon, J. Phys. Chem. 94 (14) (1990) 5579–5586.

[70] W. Cencek, J. Rychlewski, J. Chem. Phys. 98 (2) (1993) 1252–1261.

[71] W. Klopper, J. Chem. Phys. 102 (15) (1995) 6168–6179.

[72] W. Klopper, W. Kutzelnigg, Chem. Phys. Lett. 134 (1) (1987) 17–22.

[73] W. Klopper, W. Kutzelnigg, J. Phys. Chem. 94 (14) (1990) 5625–5630.

[74] W. Klopper, W. Kutzelnigg, J. Chem. Phys. 94 (3) (1991) 2020–2030.

[75] W. Klopper, R. Rohse, W. Kutzelnigg, Chem. Phys. Lett. 178 (5,6) (1991) 455–460.

[76] W. Klopper, Chem. Phys. Lett. 186 (6) (1991) 583–585.

[77] J. Noga, W. Kutzelnigg, W. Klopper, Chem. Phys. Lett. 199 (5) (1992) 497–504.

[78] W. Klopper, J. Almlöf, J. Chem. Phys. 99 (7) (1993) 5167–5177.

[79] J. Noga, W. Kutzelnigg, J. Chem. Phys. 101 (9) (1994) 7738–7762.

[80] W. Cencek, J. Rychlewski, J. Chem. Phys. 102 (6) (1995) 2533–2538.

[81] J. Komasa, W. Cencek, J. Rychlewski, Phys. Rev. A 52 (6) (1995) 4500–4507.

[82] J. Komasa, J. Rychlewski, Chem. Phys. Lett. 249 (1996) 253–256.

[83] V. Termath, W. Klopper, W. Kutzelnigg, J. Chem. Phys. 94 (3) (1991) 2002–2019.

[84] C. Schwartz, Estimating convergence rates of variational calculation, in: B. Alder (Ed.), Methods in Computational Physics, Vol. 2, Academic Press, New York and London, 1963, pp. 241–266.

[85] A. Hylleraas, Z. Physik 54 (1929) 347–366.

[86] C. L. Pekeris, Phys. Rev. 112 (5) (1958) 1649–1658.

[87] W. Kutzelnigg, Theoret. Chim. Acta 68 (1985) 445–469.

[88] W. Kutzelnigg, W. Klopper, J. Chem. Phys. 94 (3) (1991) 1985–2001.

[89] M. J. Bearpark, N. C. Handy, R. D. Amos, P. E. Maslen, Theor. Chim. Acta 79 (1991) 361–372.

[90] F. R. Manby, J. Chem. Phys. 119 (9) (2003) 4607–4613.

[91] H. J. Werner, F. R. Manby, J. Chem. Phys. 124 (2006) 054114.

[92] F. R. Manby, H. J. Werner, T. B. Adler, J. Chem. Phys. 124 (2006) 094103.

[93] A. J. May, F. R. Manby, J. Chem. Phys. 121 (10) (2004) 4479–4485.

[94] D. E. Woon, T. H. Dunning Jr., J. Chem. Phys. 100 (4) (1994) 2975–2988.

[95] D. E. Woon, T. H. Dunning Jr., J. Chem. Phys. 103 (11) (1995) 4572–4585.

[96] A. Wilson, T. v. Mourik, T. H. Dunning Jr., J. Mol. Struct. (THEOCHEM) 388 (1996) 339–349.

[97] A. Wilson, D. E. Woon, K. A. Peterson, T. H. Dunning Jr., J. Chem. Phys. 110 (16) (1999) 7667–7676.

[98] D. Feller, J. Chem. Phys. 98 (9) (1993) 7059–7071.

[99] D. Feller, J. Chem. Phys. 96 (8) (1992) 6104–6114.

[100] S. S. Xantheas, T. H. Dunning Jr., J. Chem. Phys. 97 (1) (1993) 18–19.

[101] S. S. Xantheas, T. H. Dunning Jr., J. Chem. Phys. 97 (25) (1993) 6616–6627.

[102] K. A. Peterson, T. H. Dunning Jr., J. Chem. Phys. 102 (5) (1995) 2032–2041.

[103] J. M. L. Martin, Chem. Phys. Lett. 256 (1996) 669–678.

[104] D. Truhlar, Chem. Phys. Lett. 294 (1998) 45–48.

[105] D. Carroll, H. Silverstone, R. Metzger, J. Chem. Phys. 71 (10) (1979) 4142–4163.

[106] R. N. Hill, J. Chem. Phys. 83 (3) (1985) 1173–1196.

[107] T. Helgaker, W. Klopper, H. Koch, J. Noga, J. Chem. Phys. 106 (23) (1997) 9639–9646.

[108] D. Shepard, Proc. ACM Natinal Conference (1968) 517–524.

[109] L. Stryer, Biochemistry, W. H. Freeman and Company, San Francisco, 1981.

[110] W. Kohn, Rev. Mod. Phys. 71 (5) (1998) 1253–1266.

[111] A. D. Becke, J. Chem. Phys. 98 (7) (1993) 5648–5652.

[112] M. J. Frisch, et al., Gaussian 98, Gaussian, Inc., Pittsburg, PA, 2001.

[113] J. A. Pople, M. H. Gordon, K. Raghavachari, J. Chem. Phys. 87 (10) (1987) 5968–5975.

[114] H.-J. Werner, P. J. Knowles, et al., Molpro 2002.6, a package of *ab initio* programs, see http://www.molpro.net (2003).

[115] R. Ahlrichs, A. Bär, H. P. Baron, R. Bauernschmitt, S. Böcker, N. Crawford, P. Deglmann, M. Ehrig, K. Eichkorn, S. Elliott, F. Furche, F. Haase, M. Häser, H. Horn, C. Hättig, C. Huber, U. Huniar, M. Kattannek, A. Köhn, C. Kölmel, M. Kollwitz, K. May, P. Nava, C. Ochsenfeld, H. Öhm, H. Patzelt, D. Rappoport, O. Rubner, A. Schäfer, U. Schneider, M. Sierka, O. Treutler, B. Unterreiner, M. v. Arnim, F. Weigend, P. Weis, H. Weiss, Turbomole, V5.8 (November 2005).

[116] J. L. Whitten, J. Chem. Phys. 58 (10) (1973) 4496–4501.

[117] B. I. Dunlap, J. W. D. Connolly, J. R. Sabin, J. Chem. Phys. 71 (8) (1979) 3396–3402.

[118] O. Vahtras, J. Almlöf, M. W. Feyereisen, Chem. Phys. Lett. 213 (5,6) (1993) 514–518.

[119] I. V. Rubtsov, J. Wang, R. M. Hochstrasser, Proc. Natl. Acad. Sci. 100 (2003) 5601–5606.

[120] S. Woutersen, R. Pfister, P. Hemm, Y. Mu, D. S. Kosov, G. Stock, J. Chem. Phys. 117 (14) (2002) 6833–6840.

[121] M. Volk, Eur. J. Org. Chem. 14 (2001) 2605–2621.

[122] J. M. Bowman, Acc. Chem. Res. 19 (1986) 202–208.

[123] M. Born, R. Oppenheimer, Annalen der Physik 84 (1927) 457–484.

[124] E. B. Wilson Jr., J. C. Decius, P. C. Cross, Molecular Vibrations, Dover Publications, New York, 1980.

[125] J. M. Hutson, Annu. Rev. Phys. Chem. 41 (1990) 123–154.

[126] Z. Bačíc, J. C. Light, Annu. Rev. Phys. Chem. 40 (1989) 469–498.

[127] W. Yang, A. C. Peet, J. Chem. Phys. 92 (1) (1990) 522–526.

[128] N. Halberstadt, S. Serna, O. Roncero, K. C. Janda, J. Chem. Phys 97 (1) (1992) 341–354.

[129] J. Zúñiga, A. Bastida, A. Requena, N. Halberstadt, J. A. Beswick, J. Chem. Phys. 98 (2) (1993) 1007–1017.

[130] R. C. Cohen, R. J. Saykally, Annu. Rev. Phys. Chem. 42 (1991) 369–392.

[131] M. A. McMahon, R. N. Barnett, K. B. Whaley, J. Chem. Phys. 99 (11) (1993) 8816–8829.

[132] E. Cheng, K. B. Whaley, J. Chem. Phys. 104 (9) (1996) 3155–3175.

[133] D. Blume, M. Lewerenz, F. Huisken, M. Kaloudis, J. Chem. Phys. 105 (19) (1996) 8666–8683.

[134] A. Vegiri, M. H. Alexander, S. Gregurick, A. B. McCoy, R. B. Gerber, J. Chem. Phys. 101 (3) (1994) 2577–2591.

[135] J. B. Anderson, J. Chem. Phys. 65 (10) (1976) 4121–4127.

[136] P. Sandler, V. Buch, J. Sadlej, J. Chem. Phys. 105 (23) (1996) 10387–10397.

[137] H. Romanowski, J. M. Bowman, L. B. Harding, J. Chem. Phys. 82 (9) (1985) 4155–4165.

[138] L. S. Norris, M. A. Ratner, A. E. Roitberg, R. B. Gerber, J. Chem. Phys. 105 (24) (1996) 11261–11267.

[139] A. Roitberg, R. B. Gerber, R. Elber, M. A. Ratner, Science 268 (1995) 1319–1322.

[140] G. M. Chaban, J. O. Jung, R. B. Gerber, J. Chem. Phys. 111 (5) (1999) 1823–1829.

[141] J. M. Bowman, K. Christoffel, F. Tobin, J. Phys. Chem. 83 (8) (1979) 905–912.

[142] M. A. Ratner, V. Buch, R. B. Gerber, Chem. Phys. 53 (3) (1980) 345–356.

[143] T. C. Thompson, D. G. Truhlar, Chem. Phys. Lett. 75 (1) (1980) 87–90.

[144] S. Carter, S. J. Culik, J. M. Bowman, J. Chem. Phys. 107 (24) (1997) 10458–10469.

[145] S. Carter, J. M. Bowman, J. Chem. Phys. 108 (11) (1998) 4397–4404.

[146] J. K. G. Watson, Mol. Phys. 15 (1968) 479–490.

[147] R. McWeeny, Methods of Molecular Quantum Mechanics, 2nd Ed., Academic Press, San Diego,, 1992.

[148] P. W. Atkins, Molecular Quantum Mechanics (2nd Ed.), Oxford University Press, New York, 1983.

[149] I. N. Levine, Quantum Chemistry (4th Ed.), Prentice-Hall International, Inc., 1991.

[150] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, Molpro, version 2006.1, a package of ab initio programs, see http://www.molpro.net (2006).

[151] C. d. Boor, A Practical Guide to Splines, Berlin, Springer (1st Ed.), 1978.

[152] Lancaster, Salkauskas, Curve and Surface fitting, Acad. Press, 1987.

[153] M. Schütz, H. J. Werner, Chem. Phys. Lett. 318 (2000) 370–378.

[154] M. Schütz, Phys. Chem. Chem. Phys 4 (2002) 3941–3947.

[155] M. Schütz, F. R. Manby, Phys. Chem. Chem. Phys. 5 (2003) 3349–3358.

[156] S. Grimme, J. Chem. Phys. 118 (20) (2003) 9095–9102.

[157] S. Chattopadhyay, Indian J. Physics 42 (1968) 335–343.

[158] J. H. S. Green, D. J. Harrison, Spectrochim. Acta 33A (1977) 583–587.

[159] E. B. Wilson, Phys. Rev. 45 (1934) 706–714.

[160] D. H. Whiffen, J. Chem. Soc. (1956) 1350.

[161] S. D. for Organic Compounds.

[162] T. Steinel, Private communication.

[163] J.F. Cardoso, Perturbation of joint diagonalizers., Tech. rep., Télécom Paris (1994).

[164] J.-F. Cardoso, A. Souloumiac, SIAM J. Mat. Anal. Appl. 17 (1) (1995) 161–164.

[165] P. Lemman, Private communication.