CHAPTER-II

MOLECULAR MODELING

2.1. COMPUTATIONAL CHEMISTRY

Computational chemistry simulates chemical structures and reactions numerically, based on the fundamental laws of physics. Now-a-days quantum chemical calculations are performed on a wide range of molecules using advanced computer programs. The quantum chemical calculations are complement to many experimental investigations in organic, inorganic and physical chemistry as well as to atomic and molecular physics. There are two broad areas within computational chemistry: molecular mechanics and electronic structure theory. They both perform the following basic type of calculation.

- Computing the energy of a particular molecular structure (physical arrangement of atoms or nuclei and electrons).
- Performing geometry optimization, which locates the lowest energy molecular structure in close proximity to the specified starting structure. Geometry optimizations depend primarily on the gradient of the energy - the first derivative of the energy with respect to atomic positions.
- Computing the vibrational frequencies of molecules resulting from interatomic motion within the molecule. Frequencies depend on the second derivative of the energy with respect to atomic structure and frequency calculations may predict other properties, which depends on second derivatives. Frequency calculations are not possible or practical for all computational chemistry methods.
The philosophy of computational methods of vibrational spectroscopy has changed significantly when quantum mechanical programs for optimization of the geometry of a molecule and for analytical determination of its force field appeared. Harmonic force fields derived from quantum mechanics are widely used for the calculation of frequencies and the modes of normal vibrations.

Moreover, in current quantum-mechanical calculations, vibrational spectra is performed at different levels of approximation, calculated frequencies are as a rule, higher than their experimental counterparts. This outcome is due to the more or less systematic overestimation of the force constants in the Hartree-Fock method [4]. This overestimation of the force constants depend on the basis set employed [5] and to the not-so-regular discrepancies in applications of the Moller-Plesset theory [6]. These calculations required empirical corrections. To improve agreement with experiment, quantum-mechanical force fields are corrected in one way or another, e.g. using empirical corrections called scale factors, which are estimated from the experimental vibrational spectra of small molecules with reliable frequency assignments. Second-order Moller-Plesset perturbation theory (MP2) cannot fully take the correlation energy of a system into account (the more so when restricted basis sets are used). Thus, using this theory, it does not obviate the necessity of scaling force constants. Moreover, with some molecules containing heteroatoms [7], especially halogens [8,9], this approximation leads to irregular deviations of calculated frequencies from experiment. Nowadays sophisticated electron correlation calculations are increasingly available and deliver force fields of high accuracy for small polyatomics. The scaled quantum mechanical force fields [10] are of comparable accuracy with the best purely theoretical results. In addition, the scaling procedure fits the force field to observed (anharmonic) frequencies. Thus, the reproduction of observed spectra will be better with an SQM force field than with the best harmonic field.
2.1.1. Energy Minimization

The quality of the force field calculation depends on the appropriate energy expression and the accurate geometrical parameters. The potential energy is calculated by summing up the energies of various interactions as numerical value for a single conformation. The geometries and relative energies have to be optimized for energy minimization. Energy minimization is usually performed by gradient optimization i.e., the atoms are allowed to move in order to reduce the net force on them. The energy-minimized structure has small forces on each atom and therefore serves as an excellent starting point for molecular dynamics simulations. Molecular mechanics deals with the changes in the electronic energy of the molecule due to bond stretching \( (V_b) \), bond angle bending \( (V_0) \), out-of-plane bending \( (V_{00}) \), internal rotation (torsion) about bonds \( (V_\theta) \), interactions between different kinds of motions \( (V_{int}) \), van der Waals attractions and repulsion between non-bonded atoms \( (V_{dw}) \) and electrostatic interactions between atoms \( (V_{es}) \). The sum of these contributions gives the potential energy \( V \) in the molecular mechanics framework for the motion of the atoms in the molecule. It is often called the steric energy or strain energy for the motion of atoms in the molecules.

The mathematical form of this energy function (also called potential energy surface) is given below:

\[
V(X^N) = \sum_{i=1}^{k} V_i \quad \ldots (2.1)
\]

where \( V \) represents the potential energy of the molecular system, which is a function of the Cartesian coordinates of all atoms denoted as \( X^N \). The parameters of the energy functions must be known in advance for all type of energy terms comprising of the molecular systems.
The equation (2.1) can be written as,

\[ V = V_b + V_\theta + V_\phi + V_\psi + V_{es} + V_{dw} + V_{int} \quad \ldots \] (2.2)

The potential energy calculated by summing up the energies of various interactions is a numerical value for a single conformation. The geometry optimization starts with the initially assumed geometry and finds the nearest local energy minimum by minimizing the steric energy \( V \) of the equation (2.2). This equation provides as analytical form for the energy, the first and second derivatives of \( V \) can easily be evaluated analytically, which facilitates the energy minimization. Many programs have built-in searching methods that locate many low energy conformers. Force field methods are primarily geared to predict geometries and relative energies.

2.2. MOLECULAR ORBITAL THEORY

Molecular mechanics simulations use the laws of classical physics to predict the structures and properties of molecules. There are different molecular mechanics methods. Each one is characterized by its particular force field. These classical force fields are based on empirical results, averaged over a large number of molecules. Because of this extensive averaging, the results can be good for standard systems; no force field can be generally used for all molecular systems of interest. However, molecular mechanics methods cannot describe bond formation or bond breaking or predict molecular properties, which depend on molecular orbital interactions.

2.2.1. Electronic Structure Methods

Electronic structure methods use the law of quantum mechanics as the basis for their computations. Quantum mechanics states that the energy and other related properties of a molecule can be obtained by solving the Schrodinger equation,
However, exact solutions to the Schrödinger equation are not practical. Electronic structure methods are characterized by their various mathematical approximations to their solutions. There are two major classes of electronic structure methods:

- **Semi-empirical methods** use a simpler Hamiltonian than the correct molecular Hamiltonian and use a parameter whose values are adjusted to fit the experimental data. That means they solve an approximate form of the Schrödinger equation that depends on having approximate parameters available for the type of chemical system in question. There is no unique method for the choice of parameter. **Ab initio**, force fields provide solutions to these problems.

- **Ab initio methods** use the correct Hamiltonian and do not use experimental data other than the values of the fundamental physical constants (i.e., c, h, mass and charges of electrons and nuclei). Moreover, it is a relatively successful approach to perform vibrational spectra.

### 2.2.2. Hartree Fock Method (HF)

D.R. Hartree introduced a procedure to calculate the approximate wave functions and energies for atoms and ions. Hartree was guided by some earlier, semi-empirical methods of the early 1920s (by E. Fues, R. B. Lindsay, and himself) based on the old quantum theory of Bohr. In the Bohr model of the atom, the energy of a state with principal quantum number \( n \) is given in atomic units as \( E = -\frac{1}{n^2} \). It was observed from atomic spectra that the energy levels of many-electron atoms are well described by applying a modified version of Bohr's formula. By introducing the quantum defect \( d \) as an empirical parameter, the energy levels of a generic atom was well approximated by the formula \( E = -\frac{1}{(n + d)^2} \). The existence of a non-zero quantum defect was attributed to electron-electron
repulsion, which clearly does not exist in the isolated hydrogen atom. This repulsion resulted in partial screening of the bare nuclear charge. These early researchers later introduced other potentials containing additional empirical parameters with the hope of better reproducing the experimental data. Hartree sought to do away with empirical parameters and solve the many-body time-independent Schrödinger equation from fundamental physical principles, i.e., ab-initio. His first proposed method of solution became known as the Hartree method.

This method is an approximate method for the determination of the ground-state wave function and ground-state energy of a quantum many-body system. This method assumes that the exact, N-body wave function of the system can be approximated by a single Slater determinant (in the case where the particles are fermions) or by a single permanent (in the case of bosons) N spin-orbitals. By invoking the variational principle, one can derive a set of N-coupled equations for the N spin orbitals. Solution of these equations yields the Hartree–Fock wave function and energy of the system, which approximate the exact ones.

Moreover, this method uses Slater type of orbitals (STO) which has the form

\[ \text{STO} = \frac{\xi^3}{\pi^{0.5}} e^{-\xi} \]  

... (2.4)

Where \( \xi \) is orbital exponent, which reflects the spatial extent of the orbital. The HF method is also called, especially in the older literature, the self-consistent field method (SCF) as the solutions to the resulting non-linear equations behave as if each particle is subjected to the mean field created by all other particles. The equations are almost universally solved by means of an iterative, fixed-point type algorithm.
It is important to remember that STO leads to very tedious calculations. Thus S.F. Boys developed an alternative type of orbital called Gaussian type orbital (GTO) for calculations, which are of the form

\[ \text{GTO} = \frac{2\chi}{\pi^{0.75}} e^{-\chi r^2} \]  \hspace{1cm} \ldots (2.5)

The difference between STO and GTO lies in the spatial coordinate \( r \). The GTO has square of \( r \) so that the product of one Gaussian gives another Gaussian. Ultimately it is found that more the combination of Gaussians, more the accuracy of the equations.

This solution scheme is not the only possible and not an essential feature of the HF method. For molecules, Hartree–Fock is the central starting point for most ab-initio quantum chemistry methods. The above discussion is only for the Restricted Hartree–Fock method, where the atom or molecule is a closed-shell system with all orbitals (atomic or molecular) are doubly occupied. Open-shell systems, where some of the electrons are not paired, can be dealt with by either restricted open-shell Hartree–Fock (ROHF) or Unrestricted Hartree–Fock (UHF) method.

2.2.3. Density Functional Theory (DFT)

The Hartree–Fock method is fundamentally based on the wave function approach. Density functional theory (DFT) is a quantum mechanical method [11-14] used to investigate the electronic structure of many body systems, in particular molecules and the condensed phases. Instead of solving for the wave function in order to calculate the energy of the molecule, density functional theory calculates the energy directly from the ground state electron density. In the framework of DFT approach, different exchange and correlation functional are routinely used. Among these, the Becke-3-Lee-Yang-Parr (B3LYP) combination [15,16] is the most used since it proved its ability in reproducing various molecular properties, including vibrational spectra. The combined use of B3LYP
functional and various standard basis sets, provide an excellent compromise between accuracy and computational efficiency of vibrational spectra for large and medium size molecules.

The main objective of density functional theory is to replace the many-body electronic wave functions with the electronic density as the basic quantity. Whereas the many-body wave function is dependent on $3N$ variables, three spatial variables for each of the $N$ electrons, the density is only a function of three variables and is a simpler quantity to deal with both conceptually and practically.

Although density functional theory has its conceptual roots in the Thomas-Fermi model, DFT is put on a firm theoretical footing by Hohenberg Kohn theorems [17]. The first theorem demonstrates the existence of a one-to-one mapping between the ground state electron density and the ground state wave function of a many-particle system. The second theorem proves that the ground state density minimizes the total electronic energy of the system. The most common implementation of density functional theory is through the Kohn-Sham method. Within the framework of Kohn-Sham DFT, the intractable many-body problem of interacting electrons in a static external potential is reduced to a tractable problem of non-interacting electrons moving in an effective potential. The effective potential includes the external potential and the effects of the coulomb interactions between the electrons, eg., the exchange and correlation interactions.

In many body electronic structure calculations, the nuclei of the molecules or clusters are treated as fixed, in abeyance with the Born- Oppenheimer approximation, and that generate a static potential $V$ in which electrons are moving. A stationary electronic state is described by a wave function $\psi(r_1, r_2, \ldots, r_n)$, satisfying the many electron time-independent Schrödinger equation:
\[ \hat{H}\psi = [\hat{T} + \hat{V} + \hat{U}]\psi = \left[\sum_i^N - \frac{\hbar^2}{2m} \nabla_i^2 + \sum_{i>j}^N U(\mathbf{r}_i, \mathbf{r}_j)\right]\psi = E\psi \quad \ldots (2.6) \]

where \( H \) is the Hamiltonian, \( E \) is the total energy, \( T \) is the kinetic energy, \( V \) is the potential energy from the external field due to positively charged nuclei and \( U \) is the electron-electron interaction energy. The operators \( T \) and \( U \) are called Universal operators, as they are the same for any \( N \)-electron system, while \( V \) is system dependent.

In theory, solving the Schrödinger equation yields wave functions \( \psi \) that describes the system fully. Unfortunately, the Schrödinger equation cannot be solved analytically for many body systems, so approximations are to be made. Typical solution algorithms therefore is to fix the atomic positions and then solve the Schrödinger equation for a specific value of the atomic coordinates \( r_i \). This is repeated many times to obtain the energy of the system as a function of the atomic coordinates.

Instead of solving for the wave function in order to calculate the energy of the molecule, the density functional theory calculates the energy directly from the ground state electron density \( n(\mathbf{r}) \). Here DFT provides an appealing alternative, being much more versatile as it provides a way to systematically map the many-body problem, with, onto a single-body problem with out \( \tilde{U} \). In DFT the key variable is the particle density \( n(\mathbf{r}) \) which for a normalized \( \Psi \) is given by:

\[ n(\mathbf{r}) = N \int d^3 r_2 \int d^3 r_3 \ldots \int d^3 r_N \psi^*(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N) \psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N) \ldots (2.7) \]

This relation can be reversed, i.e. for a given ground-state density \( n_0(\mathbf{r}) \) it is possible, in principle, to calculate the corresponding ground-state wavefunction \( \psi_0(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N) \). In other words, \( \psi_0 \) is a unique functional of \( n_0 \),

\[ \Psi_0 = \Psi_{n_0} \quad \ldots (2.8) \]
and consequently the ground-state expectation value of an observable $\hat{O}$ is also a functional of $n_0$

$$O[n_0] = \langle \psi[n_0] | \hat{O} | \psi[n_0] \rangle. \quad \ldots (2.9)$$

In particular, the ground-state energy is a functional of $n_0$

$$E_0 = E[n_0] = \langle \psi[n_0] | \hat{T} + \hat{V} + \hat{U} | \psi[n_0] \rangle \quad \ldots (2.10)$$

where the contribution of the external potential $\langle \psi[n_0] | \hat{V} | \psi[n_0] \rangle$ can be written explicitly

in terms of the ground-state density $n_0$

$$V[n_0] = \int V(\vec{r}) \ n_0(\vec{r}) d^3r. \quad \ldots (2.11)$$

More generally, the contribution of the external potential $\langle \psi | \hat{V} | \psi \rangle$ can be written explicitly

in terms of the density $n$,

$$V[n] = \int V(\vec{r}) \ n(\vec{r}) d^3r. \quad \ldots (2.12)$$

The functional $T[n]$ and $U[n]$ are called universal functionals, while $V[n]$ is called a non-universal functional, as it depends on the system under study. Having specified a system, i.e., having specified $\hat{V}$, one then has to minimize the functional

$$E[n] = T[n] + U[n] + \int V(\vec{r})n(\vec{r})d^3r \quad \ldots (2.13)$$

with respect to $n(\vec{r})$, assuming one has got reliable expressions for $T[n]$ and $U[n]$. A successful minimization of the energy functional will yield the ground-state density $n_0$ and thus all other ground-state observables.
The variation problems of minimizing the energy functional \( E[n] \) can be solved by applying the Lagrangian method of undetermined multipliers. First, one considers an energy functional that does not explicitly have an electron-electron interaction energy term,

\[
E_s[n] = \langle \psi_s[n] | \hat{T} + \hat{V}_s[n] \psi_s[n] \rangle
\]  

... (2.14)

where \( \hat{T} \) denotes the kinetic energy operator and \( \hat{V}_s \) is an external effective potential in which the particles are moving, so that \( n_s(\vec{r}) \equiv n(\vec{r}) \).

Thus, one can solve the so-called Kohn–Sham equations of this auxiliary non-interacting system,

\[
\left[-\frac{\hbar^2}{2m} \nabla^2 + V_s(\vec{r})\right] \phi_i(\vec{r}) = \epsilon_i \phi_i(\vec{r})
\]  

... (2.15)

which yields the orbitals \( \phi_i \) that reproduce the density \( n(\vec{r}) \) of the original many-body system

\[
n(\vec{r}) \equiv n_s(\vec{r}) = \sum_i^n |\phi_i(\vec{r})|^2.
\]  

... (2.16)

The effective single-particle potential can be written in more detail as

\[
V_s(\vec{r}) = V(\vec{r}) + \int \frac{e^2 n_s(\vec{r})}{|\vec{r} - \vec{r}'|} d^3r' + V_{XC}[n_s(\vec{r})]
\]  

... (2.17)

where the second term denotes the so-called Hartree term describing the electron-electron Coulomb repulsion, while the last term \( V_{XC} \) is called the exchange-correlation potential. Here, \( V_{XC} \) includes all the many-particle interactions. Since the Hartree term and \( V_{XC} \) depend on \( n(\vec{r}) \), which depends on the \( \phi_i \), which in turn depend on \( V_s \), the problem of solving the Kohn–Sham equation has to be done in a self-consistent (i.e., iterative) way.
Usually one starts with an initial guess for $n(\vec{r})$, then calculates the corresponding $V_s$ and solves the Kohn-Sham equations for the $\phi_i$. From these one calculates a new density and starts again. This procedure is then repeated until convergence is reached.

### 2.3. APPROXIMATIONS (EXCHANGE-CORRELATION FUNCTIONALS)

The major problem with DFT is that the exact functional for exchange and correlation are not known except for the free electron gas. However, approximations exist which permit the calculation of certain physical quantities quite accurately. In physics the most widely used approximation is the local-density approximation (LDA), where the functional depends only on the density at the coordinate where the functional is evaluated:

$$
E_{XC}^{LDA}[n] = \int \varepsilon_{XC} (n) n(\vec{r}) \, d^3r.
$$

(2.18)

The local spin-density approximation (LSDA) is a straightforward generalization of the LDA to include electron spin:

$$
E_{XC}^{LDA}[n_{\uparrow}, n_{\downarrow}] = \int \varepsilon_{XC} (n_{\uparrow}, n_{\downarrow}) n(\vec{r}) \, d^3r.
$$

(2.19)

Highly accurate formulae for the exchange-correlation energy density $\varepsilon_{XC} (n_{\uparrow}, n_{\downarrow})$ have been constructed from quantum Monte Carlo simulations of free electron model.

Generalized gradient approximations (GGA) are still local but also take into account the gradient of the density at the same coordinate:

$$
E_{XC}^{LDA}[n_{\uparrow}, n_{\downarrow}] = \int \varepsilon_{XC} (n_{\uparrow}, n_{\downarrow}, \nabla n_{\uparrow}, \nabla n_{\downarrow}) n(\vec{r}) \, d^3r.
$$

(2.20)

Using the latter (GGA) very good results for molecular geometries and ground-state energies have been achieved. Potentially more accurate than the GGA functionals are the meta-GGA functionals [18-20]. These functionals include a further term in the expansion, depending on the density, the gradient of the density and the Laplacian (second derivative).
of the density. Difficulties in expressing the exchange part of the energy can be relieved by including a component of the exact exchange energy is calculated from Hartree-Fock theory. Functional of this type are known as hybrid functional.

2.3.1. Hybrid Methods

Indeed, it is common to use calculations that are a hybrid of the two methods the popular B3LYP scheme [21] is one such hybrid functional method. The mixture of exact HF exchange and approximate DFT exchange is commonly employed to increase performance. Several different mixing ratios have been advocated. Becke Half-and-Half LYP uses a 1:1 ratio of HF and DFT exchange energies. i.e.

$$E_{XE} = \frac{1}{2} E_X(HF) + \frac{1}{2} E_X(Becke88) + E_C(LYP)$$ … (2.21)

The most commonly employed hybrid method is the Becke 3-parameter scheme (B3) [22-24]. The scheme is represented as

$$E_{XE} = 0.2 E_X(HF) + 0.8 E_X(LSDA) + 0.72 E_X(Becke88) + 0.81 E_C(LYP) + 0.19 E_C(LYP)$$ … (2.22)

Becke derived the parameters by fitting to a set of thermo chemical data, the G1 molecule set. When B3 is paired with a correlation function other than LYP, the LYP coefficients are retained. Becke also developed a one parameter fit (B). Another option is to use modern valence bond methods. The MPW1PW91 method is developed by Barone and Adamo. It employs a modified PW91 exchange functional with original PW91 correlation functional and employs a HF and DFT exchange ratio of 0.25: 0.75.
2.3.2. Basis Sets

Basis sets [25] were first developed by J.C. Slater. Slater fit linear least-squares to data that could be easily calculated. The general expression for a basis function is given as:

\[
\text{Basis Function} = N \cdot e^{-\alpha r} \quad \ldots (2.23)
\]

where:

- \( N \) = normalization constant
- \( \alpha \) = orbital exponent
- \( r \) = radius in angstroms

A basis set is a set of functions used to create the molecular orbitals, which are expanded as a linear combination of similar functions with the weights or coefficients to be determined. One characteristic of a molecule that explains a great deal about the properties is their molecular orbitals. The following diagram must be considered in order to calculate the molecular orbitals.

One of the three major decisions is which of the basis set to use. There are two general categories of basis sets:

- Minimal basis sets
- Extended basis sets

2.3.2.1. Minimal basis sets

A basis set that describes only the most basic aspects of the orbitals. All basis set equations in the form STO-NG (where N represents the number of GTOs combined to approximate the STO) are considered to be "minimal" basis sets. The most common minimal basis set is STO-nG, where n is an integer. This n value represents the number of
Gaussian primitive functions (G) comprising a single basis function. In these basis sets, the same number of Gaussian primitives comprises core and valence orbitals. Minimal basis sets typically give rough results that are insufficient for research-quality publication, but are much cheaper than their larger counterparts. Commonly used minimal basis sets of this type are:

- STO-3G
- STO-4G
- STO-6G
- STO-3G* - Polarized version of STO-3G

### 2.3.2.2. Extended basis sets

A basis set with a much more detailed description. The "extended" basis sets, then, are the ones that consider the higher orbitals of the molecule and account for size and shape of molecular charge distributions. In these sets, there are many strategies, typically resulting in multiple sets of functions being used for part or all of the angular momentum components of a molecule to describe it. In addition, one might considered higher order angular momentum descriptions for particular molecular cases, and more diffuse functions, depending on the nature of the molecule.

There are several types of extended basis sets:

- Double-Zeta, Triple-Zeta, Quadruple-Zeta
- Split-Valence
- Polarized Sets
- Diffuse Sets
2.3.3. Double Zeta basis sets

Each atomic orbital is expressed as the sum of two Slater-type orbitals. The two equations are the same except for the value of \( \zeta \) (zeta). The zeta value accounts for how diffuse (large) the orbital is. The two STOs are then added in some proportion. The constant 'd' determines how much each STO will count towards the final orbital. Thus, the size of the atomic orbital can range anywhere between the value of either of the two STOs. For example, a 2s orbital:

\[
\Phi_{2s}(r) = \Phi_{2s}^{STO}(r, \zeta_1) + d\Phi_{2s}^{STO}(r, \zeta_2)
\]

In this case, each STO represents a different sized orbital because the zetas are different. The ‘d’ accounts for the percentage of the second STO to add in. The linear combination then gives us the atomic orbital. Since each of the two equations are the same, the symmetry remains constant.

The triple and quadruple-zeta basis sets work the same way, except use three and four Slater equations instead of two. Better accuracy implies of course more computational time required.

2.3.4. Split-Valence Basis sets

Often it takes too much effort to calculate a double-zeta for every orbital. Instead, many scientists simplify matters by calculating a double-zeta only for the valence orbital. Since the inner-shell electrons are not as vital to the calculation, they are described with a single Slater orbital. This method is called a split-valence basis. Few examples of common split-valence basis sets are 3-21G, 4-31G, and 6-31G.
In Fig 2.1, a 3-21G basis set is used to calculate for the carbon atom. This means three summing Gaussians for the inner shell orbital, two Gaussians for the first STO of the valence orbital and one Gaussian for the second STO.

This is the output file from the Gaussian Basis Set Order Form for carbon given a 3-21G basis set.

<table>
<thead>
<tr>
<th>Atom</th>
<th>$\alpha_{1s}$</th>
<th>$d_{1s}$</th>
<th>$\alpha_{2s}$</th>
<th>$d_{2s}$</th>
<th>$\alpha_{2p}$</th>
<th>$d_{2p}$</th>
<th>$\alpha_{2p}^i$</th>
<th>$d_{2p}^i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>172.256000</td>
<td>0.6175690000E-01</td>
<td>3.664980000</td>
<td>-0.395897000</td>
<td>0.236400000</td>
<td>0.195857000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25.91090000</td>
<td>0.3587640000</td>
<td></td>
<td></td>
<td>0.7705450000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.53355000</td>
<td>0.7007130000</td>
<td></td>
<td></td>
<td>0.8608190000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here is another common method of displaying data. The numbers are labelled, so it is easy to match this data with the corresponding data in the output file.

Once a basis set output file is retrieved, we can use these numbers to calculate our equations. For a carbon, we will need three equations: 1s orbital, 2s orbital, and 2p orbital.

This equation combines the 3 GTO orbitals that define the 1s orbital.

$$
\Phi_{1s}(r) = \sum_{i=1}^{3} d_{is} \Phi_{1s}^{GTO}(r, \alpha_{1s})
$$

$$
= 0.6176 \Phi_{1s}^{GTO}(r, 172.256) + 0.3587 \Phi_{1s}^{GTO}(r, 25.910) + 0.7007 \Phi_{1s}^{GTO}(r, 5.533) \quad \ldots (2.25)
$$

This equation combines the 2 GTO orbitals that make up the first STO of the double-zeta, plus the 1 GTO that represents the second STO for the 2s orbital.
\[
\Phi_{2s}(r) = \sum_{i=1}^{3} d_{2s_{i}} \Phi_{2s_{i}}^{GF}(r, \alpha_{2s_{i}}) + \tilde{d}_{2s} \Phi_{2s}^{GF}(r, \alpha_{2s}) \\
= -0.395 \Phi_{2s}^{GF}(r, 3.664) + 1.215 \Phi_{2s}^{GF}(r, 0.771) \\
+ 1.000 \Phi_{2s}^{GF}(r, 0.195) \quad \ldots (2.26)
\]

This equation combines the 2 GTO orbitals that make up the first STO of the double-zeta, plus 1 GTO that represents the second STO for the 2p orbital.

\[
\Phi_{2p}(r) = \sum_{i=1}^{3} d_{2p_{i}} \Phi_{2p_{i}}^{GF}(r, \alpha_{2p_{i}}) + \tilde{d}_{2p} \Phi_{2p}^{GF}(r, \alpha_{2p}) \\
= 0.236 \Phi_{2p}^{GF}(r, 3.664) + 0.860 \Phi_{2p}^{GF}(r, 0.771) \\
+ 1.000 \Phi_{2p}^{GF}(r, 0.195) \quad \ldots (2.27)
\]

Using these three equations, the linear combination of atomic orbitals (LACO) can be calculated for the carbon atom. There are many basis sets to choose from, with many combinations of functionality in optimized proportions. Many people have devoted a great deal of time to optimize basis sets for atoms using a variety of different methods and functions.

### 2.3.5. Polarized basis sets

In the previous basis sets atomic orbitals are treated as existing only as ‘s’, ‘p’, ‘d’, ‘f’ etc. Although those basis sets are good approximations, a better approximation is to acknowledge and account for the fact that sometimes orbitals share qualities of “s” and ‘p’ orbitals or ‘p’ and “d”, etc. and not necessarily have characteristics of only one or the other. As atoms are brought close together, their charge distribution causes a polarization effect (the positive charge is drawn to one side while the negative charge is drawn to the other) which distorts the shape of the atomic orbitals. In this case, “s” orbitals begin to have a little of the 'p' flavor and ‘p’ orbitals begin to have a little of the “d” flavor. One asterisk (*) at the end of a basis set denotes that polarization has been taken into account in the ‘p’
orbitals. The difference between the representation of the ‘p’ orbital for the 6-31G and the 6-31G* basis sets are given in Fig. 2.2. The polarized basis set represents the orbital as more than just ‘p’, by adding a little ‘d’.

Two asterisks (**) means that polarization has taken into account the “s” orbitals in addition to the ‘p’ orbitals. Fig. 2.3 is another illustration of the difference of the two methods.

A hydrogen atom in a molecule experiences a similar, but non-uniform, electric field arising from its non-spherical environment. By adding polarization functions to a basis set for H we directly accommodate this effect. In a similar way, d-type functions, which are not occupied in first row atoms, play the role of polarization functions for the atoms Li to F. One denotes this improvement with a star (*) or a (d) when only the heavy atoms are corrected with d-type functions, or with two stars (**) (or with a (d,p)) when the hydrogen (or helium) atom is corrected as well with p-type functions.

2.3.6. Diffuse basis sets

In chemistry, the valence electrons are the main concern, which interacts with other molecules. However, many of the basis sets that are talked about previously concentrate on the main energy located in the inner shell electrons. Sometimes, the effect of the electrons when they are far from the nucleus is important, and this cannot be taken into account by basis functions with large Gaussian exponents. To compensate these deficiencies, computational scientists use diffuse functions. These basis sets are represented by the “+” signs. One “+” means that we are accounting for the “p” orbitals, while “++” signals that we are looking at “p” and “s” orbitals.
Fig. 2.1 Example for Split-Valance basis set

Fig. 2.2 Original and modified ‘p’ orbital

Fig. 2.3 Original and modified ‘s’ orbital
2.4. NORMAL COORDINATE ANALYSIS

Normal coordinate analysis provides information concerning the nature of vibrations [26-28]. Accurate assignment is the basis for using characteristic frequencies of various functional groups. Detailed analysis of data obtained by normal coordinate analysis (NCA) has traditionally been based on the symmetry of the equilibrium structure. The optical activity and characteristics of the normal vibration can be deduced from the constructed symmetry coordinates. Recent advances in computing have a tremendous impact on analysis of vibrational spectroscopy. Vibrational modelling of polymers has traditionally involved use of a NCA based on Wilson’s GF matrices and force constants transferred from small molecules [29]. Refinements in force field are often required to better fit the polymer system. Such analyses can only involve the specific chain conformation generated with well-established bond lengths and valence and torsional angles. The first computer programs introduced in the 1960’s were followed by many variations [29,30]. The generation of structures consistent with helical parameters obtained from diffraction data usually tests the structural parameters.

A requisite for accurate NCA analysis is the availability of a reliable force field. In generally only intramolecular potential energies are needed. The Cartesian coordinate is most convenient to define all individual atoms in space. It is, however, most natural to consider the potential energy of the molecule in terms of bond, valence angles and torsional angles, commonly referred to as internal coordinates. All force constants are defined in this coordinate system. The relationship between the two sets of coordinates is easily established [31]. With few extensions, such as polyethylene and trans-1,4-polybutadienes explicit evidence of optically active intermolecular vibrations is unavailable [32]. Therefore, interchain interactions are seldom considered, even though well-defined interatomic potentials can easily be incorporated [33,34]. As the numbers of force
constants (both diagonal and off-diagonal coupling) are much larger than the number of observations, few are determined with certainty. In both intra and intermolecular interactions, the force fields are transferred from those established for small molecules. Among those refined for specific polymers, perhaps the most used are those for n-alkanes and polypeptides.

Normal coordinate analysis is now a days commonly employed as an aid in the interpretation of the vibrational spectra of large molecules. In order to get meaningful results knowledge of the vibrational force field is necessary. Since the number of force constants grows quadratically with the number of atoms, one has to employ many approximations in the calculation of harmonic force field even for moderately large molecules.

Gwinn [35] developed a program for NCA analysis using mass-weighted Cartesian coordinates, which eliminates the redundancy problems arising when internal valence coordinates, are used, as in the GF-method. Gwinn’s methods have subsequently been improved [36,37]. MOLVIB [38] is based on the same fundamental idea, but differs from the above-mentioned programs in many respects. This program has been described by Sundius [39]. MOLVIB is FORTRAN program for the calculation of harmonic force fields and vibrational modes of molecules up to 50 atoms. This expresses the infrared and Raman intensities in the same units as the Gaussian program.

The normal modes can also be characterized by means of the potential energy distribution (PED), which is calculated from the L-matrix and the F-matrix in internal coordinates according to the formula:

\[ P_{ij} = (L_{ij})^2 F_{ij}/\lambda_i \quad \ldots (2.28) \]
The PED, which gives the fractional contribution of the diagonal matrix elements of $F$ to the normal modes, is used for a complete symmetry classification of the normal vibrations of the molecule.

2.5. GEOMETRY OPTIMIZATION

Generally structural changes within the molecule produce differences in its energy and other properties \[40,41\]. Its PES specifies the way the energy of a molecular system varies with small changes in its structure.

PES is the mathematical relation linking molecular structure and energy. The minimum, which corresponds to the lowest point in some limited region on the PES, is called local minimum. But, global minimum is the lowest energy point anywhere on the PES. The point, which is maximum in one direction and minimum in other direction, is called saddle point. Saddle point corresponds to the transition structure between two equilibrium structures. The minimum specifies the equilibrium structure of the molecular system with different minima corresponding to different conformations. After finding the energy and gradient in some no of cycle, finally optimization gets completed when it has converged.

2.6. GAUSSIAN – SOFTWARE OVERVIEW

Gaussian \[42\] is a computational software program initially released in 1970 by John Pople and research group at Carnegie-Mellon University. It has been continuously updated since then. The recent version is Gaussian 09. The name originates from the word Gaussian function or orbital, a choice made to improve the computing capacity of then existing software, which used a function or orbital called Slatter type. Gaussian functions are widely used in statistics where they describe the normal distributions. In this Quantum
computation, it represents the wave function of the ground state of a harmonic oscillator. The linear combinations of such Gaussian functions for a molecular orbital is called as Gaussian orbital. The Gaussian software program is used by Physicists, Chemists, Chemical engineers, Biochemists and others for research in established and emerging areas of molecular physics or chemistry. Starting from the basic laws of quantum mechanics, Gaussian predicts the energies, molecular structures, vibrational frequencies and other molecular properties derived from these basic quantities. It can be used to study molecules and reactions under a wide range of conditions, including both stable and short lived intermediate compounds. Computational techniques consist of three areas: Ab-initio methods, semi empirical methods, and molecular mechanics. Molecular mechanics utilizes classical physics to solve large systems of molecules and is considered the least accurate due to the fact that no electron behavior is factored in. Semi-empirical methods are more accurate because of utilization of quantum physics to account for some of the electron behavior, but its scope is still limited since it relies on extensive approximations and empirical parameters. Ab-initio methods are based purely on quantum physics and use no approximations from classical physics to describe the electronic structure of the molecule very accurately. The drawback of using ab-initio methods is that the computations are extremely taxing and so is limited to much smaller systems such as individual molecules. However, ab-initio methods give a lot of information on the electronic structure without having to actually synthesize the molecule experimentally. The fundamental idea behind ab-initio calculations is to solve Schrodinger’s equation with a set of mathematical functions called a “basis set”.
2.6.1. Capabilities of the software

The technical capabilities of Gaussian Software are listed below.

- The energy calculations were carried out by different methods such as Molecular mechanics calculations, Semi-empirical calculations, Self-Consistent Field calculations, Correlation energy calculations using Moller-plesset perturbation theory, Correlation energy calculation using CI, CID CISD techniques, Density functional theory.

- Analytic computation of force constants, polarizabilities, hyper polarizabilities and dipole moment derivatives analytically for RHF, DFT etc.

- Harmonic vibrational analysis and thermochemistry analysis using arbitrary isotopes, temperature and pressure.

- Analysis of normal modes in internal coordinates.

- Determination of IR and Raman Intensities for vibrational transitions.

- To determine Harmonic vibration-rotation coupling constants.

- To calculate Anharmonic vibration and vibration-rotation coupling constants.

- Mulliken population analysis, electrostatic potentials and electro-static derived charges.

- Static and frequency dependent polarizabilities and hyperpolarizabilities for HF and DFT.

- NMR shielding tensors using the SCF and DFT methods.

- Calculation of spin-spin coupling constants at HF and DFT level.

- Vibrational circular dichroism intensities.

- To calculate nuclear quadrupole constants, rotational constants and quartic centrifugal distortion terms.

- HOMO – LUMO analysis.
2.7. GAUSSVIEW – SOFTWARE OVERVIEW

GaussView [43] is a graphical user interface designed to prepare input for submission to Gaussian and to examine graphically the output that Gaussian produces. GaussView is not integrated with the computational module of Gaussian, but rather is a front-end/back-end processor to aid in the use of Gaussian. GaussView provides three main benefits to Gaussian users [44].

First, through its advanced visualization facility, GaussView allows to rapidly sketch in even very large molecules, rotates, translates and zooms on these molecules through simple mouse operations. It can also import standard molecule file formats such as PDB files.

Secondly, GaussView makes it easy to set up many types of Gaussian calculations. It makes preparing complex input easy for both routine job types and advanced methods like ONIOM, STQN transition structure optimizations, CASSCF calculations, periodic boundary conditions (PBC) calculations, and many more.

Lastly, to run default and named calculation templates are known as schemesto speed up the job setup process.

Finally, GaussView is used to examine the results of Gaussian calculations using a variety of graphical techniques. Gaussian results that can be viewed graphically include the following:

- Optimized molecular structures.
- Molecular orbitals.
- Electron density surfaces from any computed density.
- Electrostatic potential surfaces.
- Surfaces may also be viewed as contours.
Atomic charges and dipole moments.

Animation of the normal modes corresponding to vibrational frequencies.

IR, Raman, NMR, UV and other spectra.

Molecular stereochemistry information.

Animation of geometry optimizations, IRC reaction path following and potential energy surface scans. Two variable scans can also be displayed as 3D plots.

Plots of the total energy and other data from the same job types as in the previous item.

2.8. DESCRIPTION OF THE PRESENT RESEARCH WORK

The outline of the complete work presented in the thesis involves the following algorithm:

- To cope up with the current trends in this field, the research begins with the molecule having 18 atoms and extends up to 31 atoms. Totally six molecules have been chosen for the study.

- The FT-IR spectra of the compounds are recorded in the range of 4000 to 400 cm\(^{-1}\). The spectral resolution was ± 2 cm\(^{-1}\). The FT-Raman spectra are also recorded in the same range, in the same instrument with FRA 106 Raman module equipped with Nd: YAG laser source operating at 1.064 μm line widths, with 200 mW power. The spectra are recorded with the scanning speed of 30 cm\(^{-1}\) min\(^{-1}\) with spectral width 2 cm\(^{-1}\). The frequencies of all sharp bands were accurate to ± 1 cm\(^{-1}\).

- With the aid of above discussed experimental techniques, these molecules are subjected to new trends of theoretical methods based on quantum mechanical computations such as Ab initio HF and DFT for the spectral analyses.
The quantum mechanical computations are carried out using Gaussian 03W program with appropriate basis sets. The GaussView software is used to input the necessary data.

The molecular optimization, optimized geometrical parameters such as bond length, bond angle and dihedral angles are calculated.

Potential energy surface scan and structural confirmation based on energy are also carried out.

The results obtained from the computation are sorted out, the different parameters were tabulated, and the theoretical frequencies were scaled up with appropriate scaling factors in comparison with the experimental frequencies.

Scaling of the force field is performed according to the SQMFF procedure and the Cartesian representation of the force constants were transferred to a non-redundant set of local symmetry coordinates, chosen in accordance to the recommendations of Pulay et al.

The characterization of the normal modes using potential energy distribution (PED) has been carried out with the MOLVIB-7.0 program written by Sundius. The computed normal modes are used to estimate IR and Raman intensities from dipole and polarizability derivatives, as well as vibrational averaging effects on molecular geometries and properties.

The assignments of vibrational spectra have been carried out with the help of normal co-ordinate analysis (NCA) following the scaled quantum mechanical force field (SQMFF) methodology.

The stability and charge delocalization of the molecule have been studied by natural bond orbital (NBO) analysis.
- The molecule orbital contributions are studied by using the total (TDOS), partial (PDOS), and overlap population (OPDOS) density of states.
- UV–Visible spectrum of the compound is recorded in the region 200–400 nm and the electronic properties such as HOMO and LUMO energies were determined by time-dependent TD-DFT approach. Molecular electrostatic potential interpretation are carried out.
- Fukui functions, local softness and electrophilicity indices for selected atomic sites of the title compound are determined.
- Mulliken population analysis on atomic charges is calculated.
- Thermodynamic properties (heat capacity, entropy and enthalpy) and Non-linear optical parameters (polarizabilty, anisotropy, first and second order hyper polarizabilites) are calculated for the molecules.
- In order to compare the visual view of the parameters, the graphs are plotted using appropriate software for theoretical frequencies, theoretical intensities and other structural parameters.
- The deviations between the experimental and theoretical values are noted and the deviations are analysed with the help of literatures on structurally related molecules.
REFERENCES


    Oxford University Press, 1972.
[38] MOLVIB 7.0. Calculation of harmonic forces fields and vibrational modes of 
    molecules, QCPE Program, No. 807, 2002.


