Some New Applications in First-Principles Simulation of Molecular Crystals

Muhammad Murshed Siddick

Doctor of Philosophy
The University of Edinburgh
September 2005

In the name of God; the most Beneficent, the most Merciful.

To my parents...
Alhz. Nurul Islam and Shahanara Islam

Abstract

Work presented in this thesis details the development of new applications for molecular crystalline systems using first-principles simulation. In particular work has focused on the most important type of intermolecular interactions, the hydrogen bond. A new computational procedure to more accurately mimic the crystalline environment has been developed and applied to two systems: the test system ammonia and the more unusual dihydrogen bonded system BH_3NH_3. Both generated surprising results, which challenged the conventional view of bonding in the solid state.

Work has also focused on the dynamics of the hydrogen bond, resulting in the implementation of a constraint molecular dynamics (MD) algorithm for the popular simulation package, CASTEP. This code development allows molecular systems to be treated as rigid or semi-rigid bodies, thus allowing appreciable increase in the first-principles MD time step. It also allows interesting chemistry to be explored at the *ab-initio* level, which would be inaccessible by any other route. The method has been applied to the phase I structure of ammonia and a full vibrational analysis is reported.

Acknowledgements

First of all I would like to heartily express thanks and gratitude to my supervisors; Professor Graeme J. Ackland, for his patience and support in guiding me through this research and Dr. Carole A. Morrison, for her enormous help in every aspect of my PhD and also for her warm support and encouragement during difficult times both academically and personally. She has been a very good friend to both me and my family. Thank you both very much.

It has been a pleasure to be jointly supervised between Physics and Chemistry. I would like to thank Dr. Philip J. Camp and Professor David Rankin for the informal discussions that really helped me to strengthen my research. Many thanks to all the members of Physics and Chemistry, specially Alastair, Sarah, Iain and the ed@ed group, who I have found to be really kind and made Edinburgh a lovely place to work in.

Many thanks to my office mates: Alastair, Peter, Maria (in Chemistry), Derek and Arjun (in Physics), for keeping up with my hard times and extending their support in all sorts of things. I thankfully remember Khurom, for being my pal and specially for his many sweet visits to Little France, while we were having a very hard time with our daughter, Iffat Sara. My gracious thanks to all the members of the Bangladeshi Society(ife), which kept me sane while doing the hard work of PhD. The invaluable friends I made: Pradip, Baten, Tuhin, Mizan, Shahnoor, Dr. Raju and Dr. Jaigirdar, were particularly helpful to me in making my time in Edinburgh much pleasant. A special thanks to Pradip for many late-night lifts to home from KB and for free curry whenever I was not bothered to cook.

Many thanks to Khairun (again free foods), Chachi, Yati, Bima, Shamira and not to forget little Fahim (for giving me lots of his Power Ranger Dinothunder power) and Tania, who made home/Edinburgh a lovely place to live in. A heartfelt thanks to my wife, Shazna for making the big decision to move up to Edinburgh and share my ups and downs of PhD life. The arrival of our first daughter, Iffat Sara has been such a great joy!

Lastly if it was not the unconditional love and support of my parents, it would not have been possible for me to come this far. From my deep feeling, I dedicate this work to my parents.

Finally, I would like to thank God for giving me the strength to carry on with this research, and more importantly allowing me a chance to explore the beauty and balance of His creation.

Declaration

This thesis has been written entirely by me and has not been submitted in any previous application for a degree. Except where stated, all the work detailed in this thesis is carried out by me. Some original work in this thesis has been reported in the published papers listed in Appendix B.

Contents

Chapter 1

Introduction

At the beginning of 19th century, when the world of physics was thought to be very much understood, complete and saturated (according to some of the world's experts, everything in physics had been solved and there was nothing else to learn), the discovery of the theories of relativity (by Einstein, 1905 and onward) and quantum mechanics (led by Niels Bohr, 1920s and onward), set the discipline spinning in a new dimension. While relativity tried to explain things at a length scale of planets and stars, quantum mechanics looked at the fascinating world of atoms, electrons and nuclei.

Quantum theory has been proven correct over time, and has provided fundamental understanding of the laws governing atomic-scale phenomena. These are the laws obeyed by the nuclei and electrons of which matter is made. The impact of this was huge. New phenomena were discovered, fascinating experiments were demonstrated and theoretical understanding of many complicated phenomena were accomplished. Some examples are, the understanding of how the electrons orbit around the nucleus, super-fluidity, Bose-Einstein condensation, superconductivity, the two-slit experiment *etc.* to name but a few. And of course, one that inescapably changed the course of history and shaped the end of the last century - the atomic bomb.

However, amid these big discoveries, a level of frustration remained due to the failure to develop a precise working model of even very small systems. The frustration was

simply due to the fact that, if the quantum mechanical (QM) wavefunction of a system claims to contain *all* the necessary information to describe it, and the Schrödinger equation claims to predict the real time behaviour of these wavefunctions, it should therefore be possible to make a precise working model to simulate any atomic scale phenomenon from these basic principles of quantum mechanics (*aka* first-principles).

There were two main problems holding the scientific community back from such an achievement:

a) in order to simulate systems of many atoms, all the electrons ($\sim 10^{29}$) have to be treated quantum mechanically. These electrons interact with each other, and the motion of one affects the motion of all the others. There are tremendous mathematical difficulties to overcome in order to solve such a correlated electron system. Indeed such a problem is not possible to solve analytically. New theories and sensible approximations were needed to tackle this problem.

b) The wavefunction of each electron is represented by a large set of elementary functions (basis set). The numerical calculation of a large number of electrons and even larger set of basis functions is simply beyond the capability of human ability.

However, with the invention of modern electronic computer in 1950s, such numerical calculation seems approachable, which attracted scientists to develop numerical tools in order to calculate systems from first-principles. The formulation of Density functional Theory (DFT) in 1960s has been a big milestone towards achieving computational modelling from first principles, which led to the Nobel Prize in Chemistry (Walter Kohn and John A Pople, Nobel Prize, 1998)[1].

[1]Note that, from the point of theoretical development, DFT is not a QM method. However from the point of numerical implementation, it *is* essentially a QM technique.

With the very rapid improvement of computer technology in the 1970s, combined with the improvement in computational methods in 1980s, the first-principles simulation becomes accessible and a frontier of modern research. The pioneering work of Car and Parrinello in 1985 has demonstrated, for the first time, the enormous potential of a first-principles simulation in studying the real time behaviour of atomic scale phenomena with electronic level of accuracy.

Since then, in the last two decades, more and more condensed-matter scientists have become gripped by the powerful idea of first principles simulations. The idea is to read the atomic numbers of the constituent elements and use the fundamental laws of quantum mechanics to produce highly realistic simulations at the atomic scale to predict the physical and chemical properties of any given system (solid, liquid or cluster) of moderate size.

Molecular materials, where individual molecules work as the basic building unit of the solid, is one of the very interesting areas of research in the wide community of physics and chemistry. First-principles simulation can provide unparalleled insights into this field of research, which has never been possible before. There are three general areas of molecular materials to look at:

a) How the molecules are packed together to form the crystalline structure, *i.e.* the nature of the interaction that holds the molecule together in the solid form, and how this changes with pressure.

b) How the crystal packing behaves with finite temperature *i.e.* the vibrational property of the lattice, and in turn how this also changes with pressure.

c) How the molecular packing of a particular material changes from one phase to another with the change in temperature and pressure.

This thesis is largely concerned with the first two aspects, namely the crystal packing and the crystal dynamics of molecular materials, under an inter-disciplinary research project between computational physics and chemistry. New computational models and methodologies have been developed in this work for the purpose of adapting the

existing first-principles mechanism to the needs of molecular materials research. These new applications are tested on simple systems, and interesting results were obtained, which are discussed in the rest of the thesis according to the following layout:

- The theory and computational techniques of first-principles simulation are introduced in Chapter 2.
- A new computational model to study the hydrogen bonding in molecular crystals has been developed, and the nature and the strength of the hydrogen bond in ammonia phase I structure is studied in Chapter 3.
- The model developed in Chapter 3 is employed for studying a special type of hydrogen bond, the so-called dihydrogen bond in BH_3NH_3 in Chapter 4. A detailed comparison of our model with the widely used traditional method of hydrogen bond study in computational chemistry community has also been made in this Chapter.
- The basic theory and computational techniques of first-principles study of finite temperature system are presented in Chapter 5. A brief comparison between the Car-Parrinello molecular dynamics (MD) and Born-Oppenheimer MD is also presented.
- In Chapter 6, the detail report of the algorithm developed for constrained Born-Oppenheimer MD (BOMD) are discussed.
- The constrained BOMD, developed in Chapter 6, were applied and tested for solid ammonia phase I structure. The detail report of this test on the structural and dynamical properties of solid ammonia is reported in Chapter 7.
- A new methodology has been formulated in Chapter 8 in order to study the normal modes of lattice vibrations from first-principles based on the constrained MD. The detail report of Ammonia lattice frequencies and corresponding eigen-vectors from first-principles are reported in this chapter.
- The thesis is concluded with recommendations for future work in Chapter 9.

Chapter 2

First-Principles Simulation

2.1 Introduction

2.1.1 The importance of the total energy

The main task of a first-principles simulation is concerned with calculating the total energy of a system of electrons and nuclei. The ability of quantum mechanics to do this enables one to reap a tremendous benefit, as nearly all properties of matter are related to the total energies or to the differences between total energies. This comes from a fundamental law of nature that states that at zero temperature and pressure, systems like to stay in their lowest energy state.

This is best described by a simple example, namely the equilibrium lattice constant of a cubic crystal, which can be utilised to *optimise* the lattice structure. According to the law of lowest energy state, the *optimised structure* (see Figure 2.1 in the next page) is defined as the solid state structure sitting at the bottom of the minimum in the plot of energy *vs* lattice constant. Thus by minimising the energy (E) of the system to its lowest possible state we can guarantee the right lattice constant of the system. Obviously in a real case the total energy is a function of many degrees of freedom, and by minimising the energy with respect to all of these degrees of freedom, we get most of the physical and chemical properties of the system.

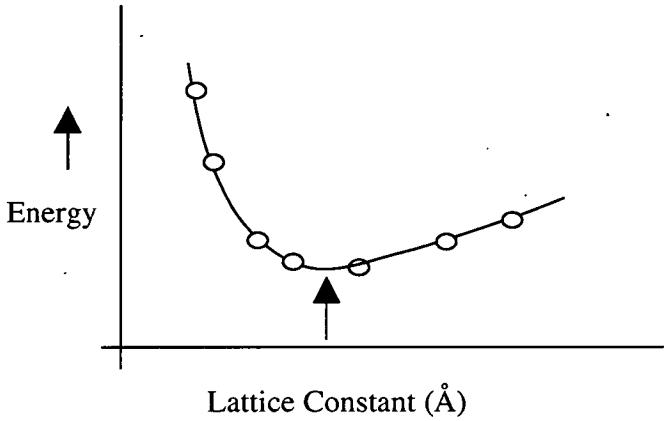


Figure 2.1: Theoretical determination of the equilibrium lattice constant. Calculations (open circles) are performed and, a smooth function is fitted through the points. The predicted lattice constant is determined by the minimum of the curve.

For properties not at 0 K, we need to minimise the free energy ($F = E - T.S$; where T is the temperature and S is the entropy), instead of the total energy. This is done in a two step process: we first optimise the system at 0 K then we run a molecular dynamics simulation to 'heat' the system to the desired temperature.

$$\text{System } (T \neq 0) = \text{System } (T=0) + \text{Molecular Dynamics } (0 \rightarrow T)$$

2.1.2 Why first-principles?

The microscopic description of the physical and chemical properties of a material is a very complex problem. There are two computational approaches available to tackle it. In general, we have a collection of atoms or molecules interacting with forces derived from some potential field. In the first approach the atoms or molecules are conveniently considered to be point particles, and the potential is some empirical potential, *e.g.* the Lennard-Jones potential. These potentials are based on some fitted parameters and although it saves tremendously on the computational cost required, the accuracy of the result obtained depends heavily on the quality of the parameters used.

In the second approach, that of first-principles simulation, the ensemble of particles, which may be isolated (molecules and clusters), extended (solids, surfaces, wires, and liquids), or a combination of both (molecules in solution), are unambiguously described by a number of nuclei and electrons interacting through Coulombic (electrostatic) potentials. Since no fitted parameters are used, first-principles calculations are more accurate, and can be applied to any system without prior knowledge of the atom-atom potential.

2.1.3 The Schrödinger equation

Formally, we can write the Hamiltonian operator of above mentioned systems in the following general form:

eqn 2.1: $$H = KE_i + KE_e + V_{ii} + V_{ee} + V_{ie} \ ,$$

where KE_i is the kinetic energy of ions or nuclei, KE_e is the kinetic energy of electron, V_{ii} is the ion-ion interaction energy, V_{ee} is the electron-electron interaction energy and V_{ie} is the ion-electron interaction energy. In proper quantum mechanical terms this can be written as:

eqn 2.2: $$H = -\sum_{I=1}^{P} \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_{i=1}^{N} \frac{\hbar^2}{2m_i}\nabla_i^2 + \frac{e^2}{2}\sum_{I=1}^{P}\sum_{J\neq I}^{P} \frac{Z_I Z_J}{|R_I - R_J|} + \frac{e^2}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{1}{|r_i - r_j|} - e^2\sum_{I=1}^{P}\sum_{i=1}^{N} \frac{Z_I}{|R_I - r_i|} \ ,$$

where $R = \{R_I\}$, $I = 1,.. ..P$; is a set of P nuclear coordinates, and $r = \{r_i\}$, $i = 1.. ..N$; is a set of N electronic coordinates. Z_I and M_I are the P nuclear charges and masses, respectively. Electrons are fermions, so that the total electronic wave function must be anti-symmetric with respect to the exchange of two electrons. Nuclei can be fermions, bosons or distinguishable particles, according to the particular problem under examination. If all these

ingredients are perfectly known and, in principle, all the properties can be derived by solving the following Schrödinger equation:

eqn 2.3: $$H\Psi(r,R,t) = E\Psi(r,R,t)\,,$$

where $\Psi(r,R)$ is the total wave function of the system. In practice, this problem is almost impossible to treat in a full quantum mechanical framework. Only in a few cases is a complete analytic solution available, and numerical solutions are also limited to a very small number of particles. There are several features that contribute to this difficulty. First, this is a multi-component many-body system, where each component (each nuclear species and the electrons) obey a particular statistics. Second, the complete wave function cannot be easily factorised because of Coulombic correlation [Fetter, 1971]. In other words, the full Schrödinger equation cannot be easily decoupled into a set of equations so that, in general, we have to deal with (3P + 3N) coupled degrees of freedom. The dynamics is an even more difficult problem, and very few and limited numerical techniques have been devised to solve it. The usual choice is to resort to some sensible approximations.

2.2 The Basic Mathematical Frame-Work

2.2.1 The adiabatic (Born-Oppenheimer) approximation

The first observation made is that the time scale associated with the motion of the nuclei is usually much slower than that of the electrons. In fact, the mass of an electron compared to that of the proton is about 1 in 2000, meaning that its velocity is much larger. The electrons can be adequately described as following instantaneously the motion of the nuclei, always staying in the same stationary state of the Hamiltonian. This stationary state will vary in time because of the Coulombic coupling of the two sets (electronic and ionic) of degrees of freedom but, if the electrons were, *e.g.* in the ground state, they will remain there forever. In other words, as the nuclei follow their dynamics, the electrons instantaneously

adjust their wave function according to the nuclear wave function. This allows the reduction of the full wave function to an expression of the type

eqn 2.4: $$\Psi(r,R,t)=\theta(R,t).\Psi_e\{r,\{R\}\},$$

where $\theta(R,t)$ is the ionic part of the wave function depending on the ionic co-ordinates and time; $\Psi_e(r,\{R\})$ is the electronic wave function depending on the electronic co-ordinate with the ionic positions, $\{R\}$, as a set of parameters. This adiabatic assumption also sets the first term of the Hamiltonian operator (eqn 2.1 and eqn 2.2), KE_i to zero. The immediate consequence of $KE_i = 0$ makes V_{ii} a constant of motion, which can then be taken out of the Hamiltonian and calculated separately. All of this can be cast in a formal mathematical framework (of the Schrödinger equation for the electronic part) by proposing a solution to eqn 2.3 of the following form:

eqn 2.5: $$\begin{cases} KE_i = 0, \\ H_e = H - V_{ii} = KE_e + V_{ee} + V_{ie}, \\ H_e\Psi_e(r,R)=E_e\Psi_e(r,\{R\}). \end{cases}$$

The total energy of the whole system (electronic plus ionic) then can be written as a simple sum

eqn 2.6: $$E = V_{ii} + E_e.$$

The total problem of electrons and ions is thus reduced down to two distinct sets. The first part is calculating the ion-ion interaction, which is treated classically. This calculation is comparatively easy, and well established methods [Allen, 1987] are available for dealing with it. The second part is the electronic Hamiltonian problem, which needs to be treated quantum mechanically. The prime interest of all first-principles *aka* ab-initio calculations is to predict this electronic part with acceptable accuracy. As it is a very

complex problem, many different methods have been developed over time to tackle it. The remainder of this chapter is devoted mainly to a discussion of these methods.

2.2.2 The common trend: towards the one electron system

With the help of the adiabatic approximation, the all-electron problem becomes much simpler, but even then the Schrödinger equation of a many electron interacting system, represented by eqn 2.5, is not exactly solvable. The only system known that can be solved exactly is a one electron with one nucleus system (the hydrogenic atoms). Unfortunately, as soon as we consider atoms with more than one electron, it becomes impossible to do the quantum mechanics exactly, and it is even worse for many atom systems. Although we know the exact form of the Hamiltonian in eqn 2.5, it is not possible to write down the exact wave function for the many electron system. The main problem behind this is the electron-electron interaction term, V_{ee}.

Although it is not possible to solve the system exactly, the basic trend of all quantum mechanical calculations is to resort to some sort of intuitive approximation so that the whole system can be described in terms of a *new* set of non-interacting electrons. This way, we simplify the many electron problem to the one electron problem (with many nuclei). This new set of electrons will be ideally different from the electrons in the real system, but they will still represent the same properties, which we are interested in, of the real system.

To illustrate this further, let us proceed, for the time being, with an ideal system with no electron-electron interactions (*i.e.* $V_{ee} = 0$). Each electron would behave as if the others were not there and we would be back close to a single electron problem (except for the presence of multiple nuclei), which can be solved exactly. Removing the cross term, V_{ee}, from the Hamiltonian also allows us to write the total electronic Hamiltonian (H_e) as a sum of one-electron Hamiltonians, as

eqn 2.7: $$H=\sum_{i=1}^{i=N} h_i \ , \quad \text{where}$$

eqn 2.8: $$h_i=-\frac{\hbar^2}{2m}\nabla_i^2-e^2\sum_{I=1}^{P}\frac{Z_I}{|R_I-r_i|} \ .$$

This simplification in the Hamiltonian immediately enforces great simplification in the total wave-function, namely that if the electrons are non-interacting, the total wave-function $\Psi_e(r_i,\{R\})$ is just the product of the wave-functions of the individual electrons, $\phi_i(r_i)$. All of this can be expressed in the following mathematical framework:

eqn 2.9: $$\Psi_e(\boldsymbol{r}_1,\boldsymbol{r}_2,.....,\boldsymbol{r}_N) = f(\phi_1,\phi_2,\phi_3.....\phi_N) \ ,$$

where ϕ_1, ϕ_2,.....ϕ_N are the one electron wave functions. In the simplest system when V_{ee} = 0, we can write

eqn 2.10: $$\Psi_e(\boldsymbol{r},\{R\}) = \phi_1(\boldsymbol{r}_1).\phi_1(\boldsymbol{r}_2).\phi_3(\boldsymbol{r}_3).....\phi_N(\boldsymbol{r}_N) \ ,$$

which reduces the all electron problem (eqn 2.5) to a one-electron Schrödinger equation of the form

eqn 2.11: $$h\phi_e(\boldsymbol{r}_e) = \varepsilon_e\phi_e(\boldsymbol{r}_e) \ ,$$

where h is defined by eqn 2.8, and with total energy given as a simple sum of one electron energy ε_e, as

eqn 2.12: $$E_e=\sum_{e=1}^{N}\varepsilon_e \ .$$

2.2.3 Solving the one-electron problem

Although the one-electron approximation allowed us to expand the total wave-function Ψ_e in terms of one-electron wave-functions ϕ_i, *we do not yet know the nature of* ϕ. The task of quantum mechanics is thus to solve eqn 2.11 in order to derive them. Once all the ϕ's are known we can construct the overall Ψ_e, and consequently derive all other system properties from it. We do this basically by using three mathematical tools: the completeness theorem, the variational principle and linear algebra [Riley, 1997].

The wave-functions, ϕ

We know the functional form of the one-electron Hamiltonian, h_i. But to proceed any further we also need to have some sort of knowledge of ϕ. The first task is thus to write down one-electron wave-functions ϕ in some reasonable form. It should be noted here the difference between the single-electron (one nucleus) problem and the one-electron (many nuclei) problem: whereas in the first case the solution is standard and known (1s, 2p, 3d,...), in the second we have to work out the specific solution to the particular ion orientation of the system.

This is where the powerful mathematical theorem of completeness steps in to help, which states that *if an infinite set of functions are linearly independent* (*i.e orthogonal to each other*) *and complete* (*i.e. the set is closed*) [Riley, 1997] *they form a complete basis set* (φ_i; $i=0,1...\infty$), *such that their linear combination can be used to construct any reasonable functions (the Dirichlet conditions for Fourier series expansion).* Thus with a *known* complete ortho-normal basis set we can expand ϕ as

eqn 2.13: $$\phi = \sum_i^{\infty} a_i \varphi_i \, .$$

The advantage of the theorem is that the basis set (also known as basis functions or eigen-functions) does not need to be unique. If a different basis set, say $\{\chi_i\}$ is also complete, it too can be used to express the one-electron wave-function ϕ in the same way. One potential problem is that because such series are infinite, it is not simply possible to calculate an infinite set of basis functions $\{\varphi_i$ or $\chi_i\}$ and their co-efficients $\{a_i\}$. However, a clever *choice* of the basis set is believed to converge to the true function ϕ rapidly so that we can truncate it at a reasonable length. One such complete basis set could be constructed from plane waves *i.e.* expanding ϕ simply in terms of the discrete Fourier series, as

eqn 2.14: $$\phi(\boldsymbol{r}) = \sum_{k=0}^{k_{cut-off}} a_k e^{i\boldsymbol{k}.\boldsymbol{r}}.$$

If we are describing an isolated molecule system, such a choice of basis set would be found to converge very slowly. An alternative and better choice for such a system is the eigen-functions that are determined for the familiar hydrogenic atomic orbitals, 1s, 2s, 2p, 3s, 3p, 3d,..... ∞, which also forms a complete basis set. In the later case we are actually building molecular orbitals (MO) from the atomic orbitals (AO) using the linear combination of atomic orbitals (LCAO) approach. Thus with a complete orthonormal set of functions available to hand, the whole problem of solving the system quantum mechanically is reduced down to solving the equations to determine the unknown co-efficients $\{a_i\}$ (real values).

The variational principle

The variational principle is probably the most important, yet the simplest principle in solving this quantum mechanical many body problems. It simply states that the energy of a system is bounded from below and the configuration corresponding to the lowest energy is the 'ground state' ε_o. Notice that this boundedness is a critical feature of quantum mechanical systems! In a classical system, one could imagine always finding a state lower in energy

than another state by simply 'shrinking the orbits' of electrons to increase the nuclear-electronic attraction while keeping the kinetic energy constant.

This is critical, because it gives us enormous freedom *from solving* the set of unknown co-efficients $\{a_i\}$ for a particular choice of basis set (i.e. LCAO, plane waves etc.) *to choosing* the unknown co-efficients $\{a_i'\}$. Suppose we have generated an infinite choice of basis set co-efficients for the *i'th* electron, $\{a_i'\}_1, \{a_i'\}_2, \ldots\ldots \{a_i''\}_\infty$. We can take each set one at a time and calculate the corresponding energy using eqn 2.8, eqn 2.11, eqn 2.12 and eqn 2.13. Since the infinite series includes all the possible combinations of real values, the set which gives the lowest energy, ε_o will be the correct set of co-efficients to describe the wave-function ϕ_i of the *i'th* electron.

In reality it is not possible to explore an infinite choice of basis set co-efficients for each of the electron, however it is implicit in the variational principle that we can start *trying* from anywhere (*i.e.* any set of choice $\{a_i'\}$) and move from one set to another in a downward energy direction in order to get to the bottom of the well. In this search each new set is apparently an *improved* guess on the previous one, as it lowers the energy. It is like a journey of a blind man who only walks down hill. The stage where he can not move further downhill (in any possible direction), he knows that he has reached home.

This can be described by the following mathematical framework. The one-electron energy (with any choice of co-efficient set) is defined as

eqn 2.15:
$$\varepsilon = \frac{\int \phi h \phi \, dr}{\int \phi^2 \, dr} \geq \varepsilon_0 .$$

If we continually change ε towards the minimum ε_o by changing the basis set $\{a_i'\}$, the minimum of the energy will be met , *i.e.* $\varepsilon = \varepsilon_o$ when

eqn 2.16: $$\frac{\partial \varepsilon}{\partial a_i} = 0 \qquad \forall i \, .$$

The secular equation

However, for our simplistic case of $V_{ee} = 0$, the exact solution could be obtained analytically from eqn 2.16 without resorting to the *trial and improve* method. Using the truncated wave-function expansion (eqn 2.13) in the definition of energy in eqn 2.15, we get

eqn 2.17: $$\varepsilon = \frac{\int (\sum_i^{N_b} a_i \varphi_i) h (\sum_j^{N_b} a_j \varphi_j) dr}{\int (\sum_i^{N_b} a_i \varphi_i)(\sum_j^{N_b} a_j \varphi_j) dr} \, ,$$

simplified to,

eqn 2.18: $$\varepsilon = \frac{\sum_{ij} a_i a_j h_{ij}}{\sum_{ij} a_i a_j s_{ij}} \, ,$$

with the matrix elements

eqn 2.19: $$h_{ij} = \int \varphi_i h \varphi_j dr \quad \text{and} \quad s_{ij} = \int \varphi_i \varphi_j dr \, .$$

N_b represents the total number of the basis functions needed to explain the one electron wavefunctions. h_{ij} and s_{ij} are known as the 'resonance integral' and the 'overlap integral', respectively. They are known values as both the operator h and the basis function φ are known. When these values are applied to eqn 2.16, we get

eqn 2.20: $$\sum_{i=1}^{N_b} a_i (h_{ki} - \varepsilon s_{ki}) = 0 \qquad \forall k \, ,$$

which is a system of N_b (one for each k) equations with N_b unknowns (running over i). [Note that we have limited our basis set to sufficient convergence, $i = 1, \ldots N_b$] This is now a problem of linear algebra. A system of N_b equations with N_b unknowns has a non-trivial solution if, and only if, the determinant formed from the co-efficients of a_i is zero, *i.e.*

eqn 2.21:
$$\begin{vmatrix} h_{11}-\varepsilon s_{11} & h_{12}-\varepsilon s_{12} & \ldots\ldots & h_{1N_b}-\varepsilon s_{1N_b} \\ h_{21}-\varepsilon s_{21} & h_{22}-\varepsilon s_{22} & \ldots\ldots & h_{2N_b}-\varepsilon s_{2N_b} \\ . & . & . & . \\ . & . & . & . \\ h_{N_b1}-\varepsilon s_{N_b1} & h_{N_b2}-\varepsilon s_{N_b2} & \ldots\ldots & h_{N_bN_b}-\varepsilon s_{N_bN_b} \end{vmatrix} = 0 .$$

Eqn 2.21 is called the *secular equation*. In general there will be N_b solutions (some could be degenerate) of ε (*i.e.* ε_α, $\alpha = 1, \ldots N_b$) for which the secular equation holds true. Each ε_α will give rise to a different set of co-efficients $\{a_{i\alpha}\}$, which can be found by solving the set of linear eqn 2.20 using ε_α. Subsequently they will define the optimal one electron wave function ϕ_α from the known basis set functions φ_i using eqn 2.13, as

eqn 2.22:
$$\phi_\alpha = \sum_i^{N_b} a_{i\alpha}\varphi_i .$$

In our simplistic case of $V_{ee} = 0$, the lowest energy molecular orbital would thus define the ground state for *the first* electron (or pair of electrons if spin is also taken into account). Note that the solution for all the other electrons will give the same set of orbitals (as the operator h is the same for all electrons). The remaining electrons will thus take the higher energy orbitals (which are excited states for the first electron). Although we have not mentioned it, it is worth noting that the variational principle also holds for the excited states as well *i.e.* the calculated energy of an excited state will be bounded from below by the true excited state energy. [MacDonald, 1933]

In summary, to solve the many electron system quantum mechanically, we require the following steps:

1. Reduce the many electron problem (Ψ_e) to an equivalent one-electron problem(ϕ).
2. Select a set of N_b basis functions $\{\varphi\}$ to expand ϕ in terms of unknown co-efficients $\{a_i\}$. Determine N_b^2 values of both h_{ij} and s_{ij}.
3. Form the secular equation and solve it to determine the N_b roots of ε. (ε_α, α=1,...N)
4. For each ε_α, solve the set of unknown co-efficients $\{a_{i\alpha}\}$. Construct one-electron wave-functions ϕ_α from $\{a_{i\alpha}\}$ and $\{\varphi_i\}$.
5. Construct the general wave-function Ψ_e from the set of one-electron wave-functions $\{\phi_\alpha\}$. The Ψ_e can then be used to calculate other system properties.

All the MOs determined this way are mutually orthogonal. Some modern methods follow alternative routes to find the correct set of co-efficients, rather than the straight-forward matrix diagonalisation of the secular equation, in order to speed up the calculation. These methods will be discussed later in Section 2.7. Apart from that, all the available methods for ab-initio calculations follow more or less the same mathematical framework to deduce the all-electron wave-function, and hence the system properties. The basic difference appears in the one-electron Hamiltonian, h_e, as the other improved methods try to calculate a more realistic system by including the electron-electron interaction term, and hence the term h_{ij} in the secular equation changes from one method to other.

2.3 Hartree Theory

The repulsion between electrons cannot be simply ignored (*i.e.* $V_{ee} \neq 0$). When the distance between two electrons is about 1 Å, their electrostatic interaction energy is roughly 14 eV (1350 kJmol^{-1}), and ignoring an energy as large as this would lead to completely wrong results [Gillan,1997]. Hartree [Hartree, 1928] suggested that although it is not possible to break down the many electron system into an equivalent one-electron system with the original electron-electron interaction present in the total Hamiltonian H, it is however possible to replace it with a static potential V_H, and still retain the simplicity of the one-electron description.

The original electron-electron potential for the whole system from the total Hamiltonian H is given in eqn 2.2 as

eqn 2.23: $$V_{ee} = \frac{e^2}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{1}{|r_i - r_j|} = \sum_{i=1}^{N} v_{ee,i} \quad where \quad v_{ee,i} = \frac{e^2}{2}\sum_{j\neq i}^{N}\frac{1}{|r_i - r_j|}.$$

Hartree proposed the following simplification:

eqn 2.24: $$V_H[\rho] = \frac{e^2}{2}\sum_{i}^{N}\int\frac{\rho(\acute{r})}{|r_i - \acute{r}|}d^3\acute{r} = \sum_{i}^{N} v_{H,i} \quad where \quad v_{H,i} = \frac{e^2}{2}\int\frac{\rho(\acute{r})}{|r_i - \acute{r}|}d^3\acute{r}.$$

The integration runs over all r', except at $r'=r_i$. Note that with this simplification the one electron 'electron-electron' potential (v_H) only depends on a static density profile rather than each individual electron (and in addition, the sum over all other electrons is replaced by an integration over space). v_H thus represents v_{ee} only in an average sense. The main point of Hartree theory is that we have replaced the correlated electron density $\rho[r, \rho(r' \neq r)]$ of the real system by a static distribution of the uncorrelated average electron density, ρ(r).

With this simplification we can still follow the same mathematical frame work described in Section 2.2, with the modification in the one electron Hamiltonian *h* as

eqn 2.25: $$h_{H,i}=-\frac{\hbar^2}{2m}\nabla_i^2 - e^2\sum_{I=1}^{P}\frac{Z_I}{|R_I - r_i|} + \frac{e^2}{2}\int\frac{\rho(\dot{r})}{|r_i-\dot{r}|}d^3\dot{r}\ ,$$

which is also written in short-hand

eqn 2.26: $$h_{H,i} = t_i + v_{ie,i} + v_{H,i}\ ,$$

where $t_i = -\frac{\hbar^2}{2m}\nabla_i^2$ and $v_{ie,i} = -e^2\sum_{I=1}^{P}\frac{Z_I}{|R_I - r_i|}$.

2.3.1 The self consistent approach

The new Hamiltonian h_H (also known as the Hartree Hamiltonian) fundamentally differs from the previous case of $V_{ee} = 0$ in that $h_H[\rho(r)]$ now depends on the density of the system (yet to be solved!). This makes the whole problem self-consistent. In order to derive the total wave-function Ψ_ρ we need to know the density of the system $\rho(r)$ first; but to know the density we must have the wave-function Ψ_ρ as

eqn 2.27: $$\rho(r)=\int \Psi^*\Psi\, d^3r\,.$$

This simply means that an analytical solution is no longer possible (as in the case of $V_{ee} = 0$) as the resonance integral ($h_{ij}[\rho(r)]$, eqn 2.19) in the secular equation *being a functional of the system density* is no longer known. However, the system could be solved iteratively using the trial and improve method as discussed in Section 2.2.3. We start with a random set of unknown co-efficients $\{a_i\}$ for a particular set of basis functions, as

eqn 2.28: $$\phi = \sum_{i}^{\infty} \acute{a}_i \varphi_i ,$$

which gives us a guess density using eqn 2.27 (and hence most likely to be wrong). Using this density we can solve the secular eqn 2.21 and get a new set of co-efficients$\{\acute{a}_i\}$. Note that the new set of co-efficients $\{\acute{a}_i\}$ is an exact solution for the wrong set-up, $\{a_i\}$. Hence, although the solved wave-function is an exact solution for *our* set-up, it is not a correct solution of the one-electron Hartree system. However being an exact solution of the starting guess state $\{a_i\}$, we expect it ($\{\acute{a}_i\}$) to generate a better density profile, which will lower the system energy. The new and better density will be used again in the same mathematical frame-work to generate an even better density profile.

Note that, this sort of down-hill motion in the energy landscape is guided by the variational principle. This means that repeating the trial and improve method iteratively for a sufficiently long time will converge the system to the correct ground-state.

Few points are important to note here with respect to the Hartree system. One: even with a simplistic form of the electron-electron interaction, the analytical solution is not available, and the system has to be solved numerically using the SCF approach. Second: the dependence of h_{H} *on* the density. The resonance integral h_{ij}, and hence the Hartree Hamiltonian are functional of the density of the system. This is because of the approximation made for the Hartree system (see eqn 2.24), and not because of the inherent nature of the true system (and the reason why Hartree theory could not be classified as a form of density functional theory!). Finally, beside missing the exact correlation, Hartree approximation $V_{ee} \simeq V_H[\rho]$, where V_H represents V_{ee} in an average sense, suffers serious blow due to the existence of strong self-interaction energy in V_H. As early as in 1934, Fermi and Amaldi [Fermi, 1934] had observed the failure of eqn 2.24 to vanish for one-electron systems due to spurious self-interaction inherent in it, and proposed the following crude correction of the form, $V_{ee} \simeq V_H(\rho) - N V_H(\rho/N)$, where N is the number of electrons in the systems, and $V_H$is given by eqn 2.24.

2.3.2 Hartree-Fock theory

Electrons have spin, so the one-electron wave-function must include spin information, *i.e.* we write the new one-electron wave-function as:

eqn 2.29: $$\chi_e = \phi_e \alpha .$$

So the total wave-function in eqn 2.10 becomes

eqn 2.30: $$\Psi_e(r,R)=\chi_1\chi_2\ldots\ldots\chi_N .$$

Such a wave function must therefore be antisymmetric (*i.e.* should change sign) upon exchange of two electronic coordinates, to comply with the Pauli exclusion principle (electrons are fermions). This simply means that the simple product wave-function in eqn 2.10 should be replaced by its Slater determinant [Slater, 1930],

eqn 2.31: $$\Psi_{SD}(r,R) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1 & \chi_2 & \ldots\ldots & \chi_N \\ \chi_1 & \chi_2 & \ldots\ldots & \chi_N \\ . & . & \ldots\ldots & . \\ . & . & \ldots\ldots & . \\ \chi_1 & \chi_2 & \ldots\ldots & \chi_N \end{vmatrix} .$$

This requirement, called 'exchange symmetry', enforces the Pauli exclusion principle into the system, which says that two electrons cannot be in the same quantum state. It also means that two electrons with the same spin cannot be found at the same place. This unsociability of electrons (as it keeps electrons away from each other) lowers the total energy and is not included in the Hartree theory. With the above correction in the wavefunction the electron-electron interaction energy term take the following form

eqn 2.32 $$V_{ee} \simeq V_H(\rho) + E_x$$

where V_H is same as in eqn 2.24 and E_x is the exchange energy due to the electron exchange symmetry. E_x consist of two part: the correct exchange energy of the electronic system and

the self-exchange energy of the electrons, which exactly cancel the Coulombic self-interaction present in the Hartree term of eqn 2.32 [Perdew, 1981]. Thus with this correction of self-interaction the mean-field Hartree-Fock theory lowers the total energy of the system towards the correct value.

Although the Hartree-Fock (HF) theory makes only a very small relative error in the total energy of atoms, it is not suitable for most solid-state and many molecular applications. The total energy is dominated by the tightly bound high-density inner-shell electrons that are well described by HF theory, but often we are more interested in the low-density valence electrons, for which correlation may be as important as exchange interaction. Although HF theory treats electron exchange exactly, it ignores the electron correlation. In addition, the long range of the Coulomb interaction produces unrealistic feature in the HF energy eigenvalues, *e.g.* vanishing density of states at the Fermi level in metals, unreliably large band gaps in insulators [Kunz, 1975], *etc.* While correlation effects can be addressed through configuration-interaction [Weiss, 1961] corrections to HF, the complexity of these corrections and their remarkable sensitivity to the choice of basis set preclude application to large systems. Density functional theory [Hohenberg, 1964; Kohn, 1965] provides an alternative to this approach. This theory which is exact in principle includes both exchange and correlation explicitly in the total energy, and describe them as a unique functional of the electronic charge density.

2.4 Density Functional Theory (DFT)

Density functional theory, established by Hohenberg and Kohn [Hohenberg, 1964] and then Kohn and Sham [Kohn, 1965], is not an improvement over Hartree-Fock theory, but a radically different line of approach. It is at the same time an *exact* theory for real systems (interacting electrons) while retaining the simplicity of the one-electron framework described in Section 2.2 for the calculation to be tractable.

It is based on two fundamental theorems that allow the *exact* equivalence of the interacting electron system to an apparently non-interacting system of electrons (Kohn-Sham electrons) in an effective external potential. The effective potential is the non-interacting potential (as seen in previous theories) plus the missing bit of the potential, which makes the system interactive. This 'extra bit' is popularly known as exchange-correlation. The theorems are

1. The total energy of a system of electrons and nuclei is a unique functional of the electron density.

2. The variational minimum of the energy, *with respect to the electron density*, is exactly equivalent to the true ground state energy.

The two theorems could be put in reverse like this: whichever density profile gives the lowest energy will therefore give the ground-state energy, and the density which gives the ground state is the *exact* electronic distribution of the system, *whether or not the wave-function associated with that density profile is a true one or has any significance.*

The beauty of DFT is that one does not need to depend on the construction of the many-body wave function Ψ_e (however, we will see later that the one electron wave-function ϕ comes into play, but this is only as an auxiliary tool to make the maths work). Instead the energy is written in terms of the electron density. This simply means that we

can stay away from the fearsome complexity of a multidimensional wavefunction (ψ) and instead we can deal with a scalar field [$\rho(r)$].

2.4.1 Theorem 1: the energy is a functional of the density

For an interacting electron system, the electronic part of the total energy (note we are still using the adiabatic separation, so the ionic energy, which will also be a functional of the density will just be an extra addition to the total energy) is

eqn 2.33: $$E_e[\rho(\boldsymbol{r})]=T_{int}[\rho(\boldsymbol{r})] + V_{ie}[\rho(\boldsymbol{r})] + V_{ee}[\rho(\boldsymbol{r})],$$

where the terms in the R.H.S. refer, respectively, to the correlated kinetic energy of electrons, the ion-electron interaction energy and the correlated electron-electron interactions. We should mention that we have no idea about the first and the third terms, because of their interacting nature.

However, theorem one of DFT simply means that whichever way we like to divide the individual energy contributions, since the total energy is a functional of the density, each component will also be a functional of the density. This offers an enormous flexibility in writing the total energy (and hence the Hamiltonian) that we can exploit immediately, *i.e.*

eqn 2.34: $$E_e[\rho(\boldsymbol{r})] = T_{ni}[\rho(\boldsymbol{r})] + V_{ie}[\rho(\boldsymbol{r})] + V_H[\rho(\boldsymbol{r})] + \Delta T_{int}[\rho(\boldsymbol{r})] + \Delta V_{ee}[\rho(\boldsymbol{r})] .$$

Here the first and third terms are the non-interacting parts of the kinetic and electron-electron interaction energy (H stands for Hartree), and the last two terms represent the missing interacting components from the kinetic and electron-electron interaction

energies, respectively. We can simplify this more by combining the last two terms into one, as

eqn 2.35: $$E_e[\rho(\boldsymbol{r})]=T_{ni}[\rho(\boldsymbol{r})] + V_{ie}[\rho(\boldsymbol{r})] + V_H[\rho(\boldsymbol{r})] + E_{xc}[\rho(\boldsymbol{r})] .$$

The advantage of writing the equation in this form is that we know exactly the first three terms, and although we do not know the last term, we still have some crucial information: it is a unique functional of the density. E_{xc} is known as the exchange-correlation functional. The power of the first theorem of DFT is that it asserts, that such a functional of density *must* exist. What it does not do, however, is tell us the exact nature of that functional or, indeed how to find it.

2.4.2 Theorem 2: the energy minimum can be reached in terms of the density

In a later section (Section 2.4.4), we will discuss what reasonable forms the exchange-correlation functional may take. For the time being let us assume E_{xc} is somehow known. The task is then to write down the remaining three terms in eqn 2.35 in explicit form of the density, and solve the problem. However, although we have claimed that first three terms are known, T_{ni} is only known in a wave-function representation (or particle representation), and not known in the density representation. Earlier attempt to find the kinetic energy in terms of the density failed [Thomas, 1926; Fermi, 1928]. The second and third terms, V_{ie} and V_H, can however be easily expressed in terms of the density as,

eqn 2.36:
$$\begin{cases} V_{ie}[\rho(\boldsymbol{r})] = -e^2 \int d^3\boldsymbol{r}_1 \rho(\boldsymbol{r}_1) \sum_{I=1}^{P} \dfrac{Z_I}{|\boldsymbol{R}_I - \boldsymbol{r}_1|} = \int d^3\boldsymbol{r}_1 \rho(\boldsymbol{r}_1) v_{ie}(\boldsymbol{r}_1) \\ \textit{where} \qquad v_{ie}(\boldsymbol{r}_1) = -e^2 \sum_{I=1}^{P} \dfrac{Z_I}{|\boldsymbol{R}_I - \boldsymbol{r}_1|} . \end{cases}$$

eqn 2.37:
$$\begin{cases} V_H[\rho(\boldsymbol{r})] = \dfrac{e^2}{2}\iint d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2 \dfrac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1-\boldsymbol{r}_2|} = \int d^3\boldsymbol{r}_2 \rho(\boldsymbol{r}_1) v_H(\boldsymbol{r}_1) \\ \textit{where} \qquad v_H(\boldsymbol{r}_1) = \dfrac{e^2}{2}\int d^3\boldsymbol{r}_2 \dfrac{\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1-\boldsymbol{r}_2|} . \end{cases}$$

For the simplicity of further mathematical manipulation it is convenient to write the first and fourth terms also in a similar fashion, which is possible by *defining* new variables $t_{ni}(\boldsymbol{r}_1)$ and $\varepsilon_{xc}(\boldsymbol{r}_1)$ as the functional derivatives [Riley, 1997] of T_{ni} and E_{xc}, respectively.

eqn 2.38:
$$\begin{cases} t_{ni}(\boldsymbol{r}_1) = \dfrac{\delta T_{ni}[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r}_1)} , \ \textit{such that} \\ T_{ni}[\rho(\boldsymbol{r})] = \int d^3\boldsymbol{r}_1 \rho(\boldsymbol{r}_1) t_{ni}(\boldsymbol{r}_1) . \end{cases}$$

eqn 2.39:
$$\begin{cases} \varepsilon_{xc}(\boldsymbol{r}_1) = \dfrac{\delta E_{xc}[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r}_1)}, \ \textit{such that} \\ E_{xc}[\rho(\boldsymbol{r})] = \int d^3\boldsymbol{r}_1 \rho(\boldsymbol{r}_1) \varepsilon_{xc}(\boldsymbol{r}_1) . \end{cases}$$

We can then re-write eqn 2.35 in the simplified form:

eqn 2.40:
$$E_e[\rho] = \int \rho(\boldsymbol{r}_1) d^3\boldsymbol{r}_1 \left[t_{ni}(\boldsymbol{r}_1) + v_{ie}(\boldsymbol{r}_1) + v_H(\boldsymbol{r}_1) + \varepsilon_{xc}(\boldsymbol{r}_1) \right] .$$

In eqn 2.35 - eqn 2.39, the term [ρ(r)] on the L.H.S. appears to reflect the meaning that energies are a *functional* of density, and that density is a *function* of position. These energies *should not be* confused as the energy at a specific position $\boldsymbol{r}$ in space. In order to avoid confusion, we have dropped the position $\boldsymbol{r}$ in denoting the total energy in the L.H.S. of eqn 2.40. However, the terms within bracket in the R.H.S. of this equation clearly represent the energy *concentrated* at a particular space position $\boldsymbol{r}_1$ due to its own (*i.e.* $\boldsymbol{r}_1$) *plus*

interactions from all other places (r_2). Integrating these energies at each point of space with the *appropriate density* gives us the correct total electron energy E_e. So eqn 2.40 can be re-written as

eqn 2.41: $$E_e[\rho] = \int \rho(r_1) d^3 r_1 \varepsilon_e(r_1) \qquad \text{where}$$

eqn 2.42: $$\varepsilon_e(r_1) = t_{ni}(r_1) + v_{ie}(r_1) + v_H(r_1) + \varepsilon_{xc}(r_1) .$$

Here $\varepsilon_e(r_1)$ is the energy concentrated at r_1. eqn 2.41 is clearly a position and space dependent representation of the total electronic energy E_e. The simplified forms of eqn 2.41 and eqn 2.42 allow us to apply the Hohenberg-Kohn variational theorem *in terms of density*, in order to define the ground state of the total energy, E_{e0}. Besides the fact that E_e is stationary with respect the variation of ρ(r), there is the additional constraint that the variation also leaves the total number of electrons, $\int dr \rho(r) = N$, unchanged at $E_e = E_{e0}$. Therefore, according to the variational principle, we have the functional derivative,

eqn 2.43: $$\left[\frac{\delta E_e[\rho] - \mu N}{\delta \rho(r)}\right]_{E_e = E_{e,0}} = 0, \quad \text{which gives,}$$

eqn 2.44: $$\left[t_{ni}(r) + v_{ie}(r) + v_H(r) + \varepsilon_{xc}(r)\right]_0 = \mu .$$

Eqn 2.44 is crucial and serves as a model equation for all further discussion. The parameter μ in eqn 2.44 is the Lagrange Multiplier associated with the particle number constraint and represents the chemical potential of the system. It has the physical meaning of the energy associated with adding or removing *a single particle* into the system, or simply the energy *per particle*. Here the subscript, '0', denotes the system is in (or restricted to) the ground state configuration. Eqn 2.44 represents a correct model for interacting ground-state electron system in a space and position based (*i.e.* density) representation. However, the severe restriction that the kinetic energy $t_{ni}(r)$ is unknown in the density representation in eqn 2.44 motivates us to seek a different interpretation of this equation.

2.4.3 Kohn-Sham electrons: mapping back to the one-electron scenario

To overcome the problem with kinetic energy Kohn and Sham [Kohn, 1965] suggested that: " the whole problem of N interacting electrons can be *exactly* mapped to an equivalent set of non-interacting fermions, such that they still fulfil eqn 2.44 in order to reproduce *exactly* the same charge density and total energy of the *real* system." These fermions were assigned the same mass, charge and total number as for the real system of electrons, and are popularly known as the Kohn-Sham (KS) electrons.

There are two key reason for the move towards the one-particle Kohn-Sham representation. Reason1: eqn 2.44 is the Euler equation for non-interacting electrons in a potential $v_{ie}(\boldsymbol{r}) + v_H(\boldsymbol{r}) + \varepsilon_{xc}(\boldsymbol{r})$. This means that it is *possible* to map these individual energy terms at a *single point* $\boldsymbol{r}$ in space, as if they are the energies due to a *single particle* at $\boldsymbol{r}$. Reason 2: since the kinetic energy $t_{ni}(\boldsymbol{r})$ is so far only solvable in the one-electron representation [Segall, 2002], it is *desirable* to treat the whole system as a non-interacting system of N particles.

The only condition in such a mapping is that, the new energies must add up to the same total energies in eqn 2.35 for the ground state, in order to maintain the correct density distribution (energy is a sole function of density). This map can be mathematically expressed as,

eqn 2.45: $$t_{ni}(\boldsymbol{r}) \rightarrow t_{ni,\alpha}, \quad \textit{such that} \quad \sum_{\alpha}^{N_\alpha} t_{ni,\alpha} = T_{ni},$$

eqn 2.46: $$v_{ie}(\boldsymbol{r}) \rightarrow v_{ie,\alpha}, \quad \textit{such that} \quad \sum_{\alpha}^{N_\alpha} v_{ie,\alpha} = V_{ie},$$

eqn 2.47: $$v_H(\boldsymbol{r}) \rightarrow v_{H,\alpha}, \quad \textit{such that} \quad \sum_{\alpha}^{N_\alpha} v_{H,\alpha} = V_H,$$

eqn 2.48: $$\varepsilon_{xc}(\boldsymbol{r}) \rightarrow \varepsilon_{xc,\alpha}, \quad such\,that \quad \sum_{\alpha}^{N_\alpha} \varepsilon_{xc,\alpha} = E_{xc},$$

eqn 2.49: $$\mu \rightarrow \varepsilon_{\alpha}, \quad such\,that \quad \sum_{\alpha}^{N_\alpha} \varepsilon_{\alpha} = E_{e,0}.$$

Here the subscript α is chosen (instead of i and j used earlier to represent electrons in Section 2.2, 2.3 and 2.4) in order to distinguish the KS fermions from the interacting electrons of the real system. Given that these KS fermions are the same in total number, mass and charge, i.e.

eqn 2.50: $$N_\alpha = N_e = N; \quad m_\alpha = m_e = m; \quad q_\alpha = q_e = e.$$

we get exactly the same form for $v_{ie,\alpha}$ and $v_{H,\alpha}$ as in eqn 2.36 and eqn 2.37. Note that so far we do not need to implement any quantum mechanical treatment in DFT, almost implying that DFT is a step back from our quest to solve the system quantum mechanically. However, the kinetic energies of electrons are only obtainable by expressing the KS particles explicitly in terms of wave-functions ($\varphi_{\alpha,0}$) at ground state, as

eqn 2.51 : $$t_{ni,\alpha} = \langle \varphi_{\alpha,0} | -\frac{\hbar^2}{2m} \nabla_\alpha^2 | \varphi_{\alpha,0} \rangle .$$

So the one electron map of our model equation (eqn 2.44) is

eqn 2.52: $$\langle \varphi_{\alpha,0} | -\frac{\hbar^2}{2m} \nabla_\alpha^2 | \varphi_{\alpha,0} \rangle + v_{ie,\alpha} + v_{H,\alpha} + \varepsilon_{xc,\alpha} = \varepsilon_{\alpha,0},$$

or more conveniently

eqn 2.53: $$\left(-\frac{\hbar^2}{2m} \nabla_\alpha^2 + v_{ie,\alpha} + v_{H,\alpha} + \varepsilon_{xc,\alpha} \right) \varphi_{\alpha,0} = \varepsilon_{\alpha,0} \varphi_{\alpha,0} .$$

eqn 2.53 is known as the Kohn-Sham equation. The main feature of this new representation is its similarity to the one-electron Schrödinger equation, immediately telling us that we can simply plug this equation in to our one-electron framework described in Section 2.2. Therefore the system of interacting particles can *now* be solved self consistently (given that ε_{xc} is known) starting with some trial functions (see Section 2.2.3), where

eqn 2.54: $$h_{ks,\alpha} = -\frac{\hbar^2}{2m}\nabla_\alpha^2 + v_{ie,\alpha} + v_{H,\alpha} + \varepsilon_{xc,\alpha} ,$$

is known as the KS Hamiltonian. Hence we see the DFT treatment now reduced to the same SCF treatment as in the original Hartree framework, with the only difference being the treatment of the resonance integral in secular eqn 2.21. The h_{ij} (in eqn 2.19) is changed by $h_{\alpha\beta}$ as

eqn 2.55: $$h_{\alpha,\beta} = \langle \varphi_\alpha | h_{ks} | \varphi_\beta \rangle .$$

Charge density and the wave-function

Note that we have cautiously avoided calling the fictitious particles (KS fermions) electrons. However they have the same mass (m), charge (e) and total number (N). They have the same total energy, and hence the same density distribution (of the electrons) of the true system, *at the ground state*. The only difference is that their wave-functions (the KS wave-function) and their Slater determinant may or may not represent the real system. However there is strong evidence that the Kohn-Sham wave-function does have physical meaning for the real system, but the interpretations need care.

With this successful mapping of Kohn-Sham (KS) electrons, the charge density can now be expressed as,

eqn 2.56: $$\rho(\boldsymbol{r}) = \sum_{\alpha=1}^{N} \varphi^*_\alpha(\boldsymbol{r})\, \varphi_\alpha(\boldsymbol{r}) .$$

Note this the definition of density in terms of KS wave-function is not absolute. It is a result of our choice of mapping to the one-electron system (a different mapping might have lead to a different definition). This emphasizes the point that in DFT the density is the central quantity, whereas the wave-function is an auxiliary quantity in order to assist the solution of the many-body problem.

2.4.4 Exchange-correlation functionals

With the above prescription we can see DFT is very much alike to the earlier Hartree-Fock(HF) theory except that the one electron Hamiltonian contains an extra term, the exchange-correlation functional ε_{xc}, which claims to contain all the missing information of the electron correlation in the HF level of theory. However, as mentioned earlier, DFT does not give any practical prescription to construct such a functional. The power of the DFT is that it sets a clear and specific direction to search and postulate such a functional, *i.e.* it is an unique functional of electron density.

Practical E_{xc} functionals are the major approximations made in DFT calculations. They are not from first-principles, rather they are postulated from physically reasonable assumptions, and their use is justified *a posteriori* by their success. However, it is surprising that simple minded approximations of this functional works, and in many cases, works incredibly well.

The most widely used approximation is the Local Density approximation (LDA), which was introduced by Kohn and Sham [Kohn, 1965], along with the KS equation. The LDA assumes that the exchange-correlation functional E_{xc} is purely a local quantity, and hence can be obtained by assuming that for each infinitesimal element of density $\rho(\boldsymbol{r})d\boldsymbol{r}$, the $\varepsilon_{xc}[\rho(\boldsymbol{r})]$ is that of a uniform electron gas of density $\rho_{gas} = \rho(\boldsymbol{r})$. So we can have

eqn 2.57
$$E_{xc} = \int d\boldsymbol{r}\,\rho(\boldsymbol{r})\varepsilon_{xc}(\rho_{gas}),$$

where $\varepsilon_{xc}[\rho_{gas}]$ is the exchange-correlation energy in a uniform gas of density ρ_{gas}. Although, LDA is wrong as the electron distribution is highly non-uniform around the atoms, it does work well for a large range of materials, particularly for those where electrons are more delocalised (*e.g.* metals). Also, our choice is somehow limited as the uniform electron gas is the only system for which E_{xc} can be calculated exactly in order to construct $\varepsilon_{xc}[\rho_{gas}]$ [Ceperely, 1980].

The LDA particularly fails for systems for which spatial variation in the electron density is significant (*e.g.* molecular solids). Functionals that include some information regarding the spatial variation of the density in the approximation have been developed to tackle this problem. This leads to a new scheme known as generalised-gradient approximation (GGA), which besides the density also includes the gradient of the density at $\boldsymbol{r}$ in the construction of $\varepsilon_{xc}[\rho_{gas}]$ [Perdew, 1991]. The GGA functionals highly improve the quality of the calculation in predicting the binding and disassociation energies, especially for hydrogen-containing systems [Perdew, 1992].

Almost all present work reported in the literature involve either LDA or GGA functionals, available in many different parameterisation frameworks. However, despite the success of LDA and GGA they are far from ideal, and finding an accurate universally acceptable functional is probably the greatest challenge for computational modelling today [Parr, 1989].

2.5 Basis set

The main focus of this PhD project is to study extended (crystal) systems. As yet there has been no mention of how to handle the problem of an infinite number of interacting electrons moving in the static field of an infinite number of ions in a crystal. Essentially, there are two difficulties to overcome: first, a wave-function has to be calculated for each of

the infinite numbers of electrons, which should ideally extend over the entire space (infinite) of the solid, and second, the basis set in which the wave-functions will be expressed will have infinite dimension.

2.5.1 Plane-wave basis set

The ions in a perfect crystal are arranged in a regular periodic way (at 0 K). Therefore the external potential (*i.e.* the ionic potential) felt by the electrons will also be periodic, the period being the same as the length of the unit cell, $\boldsymbol{a}$. The external potential on an electron can therefore be expressed as a periodic function of the form: $V(\boldsymbol{r})=V(\boldsymbol{r}+\boldsymbol{a})$. This is the requirement needed for the use of Bloch's theorem [Ashcroft, 1976], which allows the wave-function of an infinite *real space* crystal to be expressed in terms of wave-functions in a finite *reciprocal* space (k-space) crystal, *i.e.* the first Brillouin zone.

Bloch's theorem therefore reduces the *infinite* number of one-electron problem in a crystal to a one-electron problem of *finite* number, which is simply the number of electrons in the unit cell (or half that number if the electronic orbitals are assumed to be doubly occupied). Bloch's theorem states that, under a periodic external potential of period $\boldsymbol{R}$, the wave function can be written as the product of a cell periodic part and a wave-like part, *i.e.*

eqn 2.58: $$\varphi_{i,k}(\boldsymbol{r}_i)=\exp\left(ik.\boldsymbol{r}_i\right)f_i(\boldsymbol{r}_i)\ .$$

The second term is periodic, and directly reflects the periodicity of the potential *i.e.* $f_i(\boldsymbol{r}_i)=f(\boldsymbol{r}_i+\boldsymbol{R})$, where $\boldsymbol{R}$ is any lattice vector of the crystal [Ashcroft, 1976]. If we recall our discussion of completeness theorem in Section 2.2.3, the second term can be expressed in plane waves as

eqn 2.59: $$f_i(r)=\sum_G c_{i,G}\exp(iG.r),$$

where $\boldsymbol{G}$ is known as the reciprocal lattice vectors, representing the periodicity of the crystal, defined by the relation $\boldsymbol{G}.\boldsymbol{R} = 2\pi m$ for all $\boldsymbol{R}$ and m is any integer. Therefore each electronic wave function is written as a sum of plane waves

eqn 2.60: $$\varphi_{i,\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} c_{i,\boldsymbol{k}+\boldsymbol{G}} \exp\left[i(\boldsymbol{G}+\boldsymbol{k}).\boldsymbol{r}\right] .$$

2.5.2 Brillouin zone sampling

The first term in eqn 2.58 is the wave-like part, and it clearly introduces new degrees of freedom, $\boldsymbol{k}$, into the one-electron functions (eqn 2.60). As $\boldsymbol{k}$ represents all the co-ordinates in the continuous space of the first Brillouin zone of the crystal, we have infinite $\boldsymbol{k}$ points at which each $\varphi_{i,\boldsymbol{k}}$ has to be calculated. However, it is possible to sample the k-space and discretise the first Brillouin zone to a finite set of $\boldsymbol{k}$ points. This is often referred to as 'Brillouin zone sampling' and the two most common methods are those of Chadi and Cohen [Chadi, 1973] and Monkhorst and Pack [Monkhorst, 1976].

2.5.3 Cut-off energy

The electronic wave functions at each k-point are now expressed in terms of a discrete plane wave basis set. In principle the sum over G in eqn 2.60 is infinite. The coefficients for the plane waves, $c_{i,k+G}$, for each G carries a kinetic energy of $\left(\hbar^2/2m\right)|k+G|^2$, which represents how diffuse or concentrated the electrons are around the core. The plane waves with a smaller kinetic energies usually represents the valence electrons, and hence have a more important role in determining the chemistry of the system than those with a very high kinetic energy, which usually represents the core electrons. The introduction of a cut-off in the plane wave basis set in terms of their kinetic energy is thus reasonable. We therefore define a cut-off energy for the basis set as

eqn 2.61: $$E_{cut-off} = \frac{\hbar^2}{2m}|k+G_c|^2,$$

which reduces the basis set to a finite tractable size. This cut-off will obviously lead to an error in the total energy calculation of the system, but in principle it is possible to make this error arbitrarily small by increasing the size of the basis set (by setting a larger energy cut-off). In practice, the cut-off that will be used depends very much on the system under investigation.

2.5.4 Reciprocal-space representation

Another advantage of expanding the electronic wave-functions in terms of a basis set of plane waves is that the one electron equations take a particularly simple form. Substitution of eqn 2.60 in to the KS equation (eqn 2.53) gives

eqn 2.62:

$$\sum_{G'}^{G_c}\left[\frac{\hbar^2}{2m}|\boldsymbol{k}+\boldsymbol{G}|^2\delta_{GG'}+V_{ie}(\boldsymbol{G}-\boldsymbol{G}')+V_H(\boldsymbol{G}-\boldsymbol{G}')+\varepsilon_{xc}(\boldsymbol{G}-\boldsymbol{G}')\delta_{GG'}\right]c_{i,k+G'} = \varepsilon c_{i,k+G} \quad \forall G.$$

This can be written as

eqn 2.63: $$\sum_{G'}^{G_c}\left(h_{k,GG'}-\varepsilon_{i,k}\delta_{GG'}\right)c_{i,K+G'} = 0, \ \forall G, \qquad where \ \boldsymbol{G}=\boldsymbol{G}_1,....\boldsymbol{G}_c \text{ and}$$

$$h_{k,GG'} = \frac{\hbar^2}{2m}|\boldsymbol{k}+\boldsymbol{G}|^2\delta_{GG'}+V_{ie}(\boldsymbol{G}-\boldsymbol{G}')+V_H(\boldsymbol{G}-\boldsymbol{G}')+\varepsilon_{xc}(\boldsymbol{G}-\boldsymbol{G}')\delta_{GG'}.$$

The system of equations (for all $\boldsymbol{G}$) represented by eqn 2.63 forms the secular equation in k-space, which have a non-trivial solution when the determinant is zero. Hence

with the introduction of Bloch's theorem, the electronic structure problem can be solved using the self consistent matrix diagonalisation method (Section 2.2.3) in k-space. Note that the size of the determinant matrix is determined by the cut-off energy (eqn 2.61).

2.6 Ion-Electron Interactions

DFT as described above would still prove computationally too difficult for many of the systems we would like to study. Further streamlining is necessary to reduce the task to a manageable size, and one important way of achieving this is through the use of pseudopotentials. These rely on the fact that the core electrons are tightly bound to their host nuclei, and only the valence electrons are involved in chemical bonding. It is therefore possible to merge the core electronic states (wavefunctions and densities) into the bulk nuclear potential, or pseudo-potential, and only deal with the valence electrons directly.

The presence of true core electrons in any total energy calculation possess two major difficulties. The core electron's wavefunctions under the *true ionic potential* are rapidly varying with many nodes (See Figure 2.2), and therefore require a very large number of fitting functions to model them accurately. In order to remain orthogonal to the core electrons, the valence electrons are also forced to become rapidly oscillating near the atomic core, and therefore also require large numbers of fitting functions (*e.g.* plane waves). As seen in eqn 2.63 the size of the determinant to be solved is proportional to N_{pw} (where N_{pw} is the number of plane-wave coefficients), and hence the matrix diagonalisation would become increasingly expensive with the presence of these core electrons. In addition to this, the Coulombic interaction energy contribution from the core electrons would tend to swamp any total energy calculations, which tries to project on small differences in valence bond energies (*e.g.* hydrogen bond calculations). The pseudo-potential technique overcome all these problems by replacing the true core electronic states with pseudo-states or pseudo-wavefunctions (as a direct consequence of the construction of the *pseudo ionic potentials*), which are smoothly varying and usually nodeless within a certain core radius (See Figure 2.2), and much lower in energy.

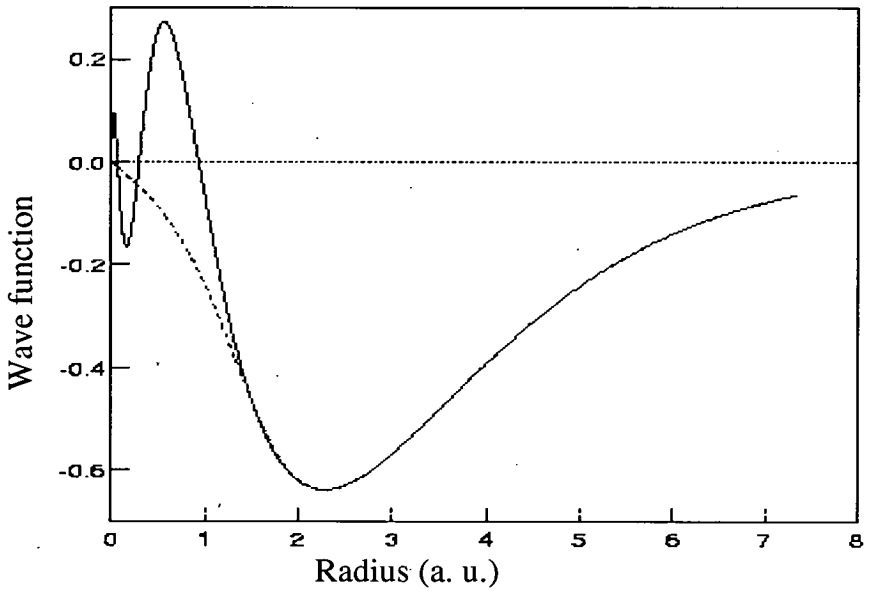


Figure 2.2 The radial 4s full- (solid) and pseudo- (dashed) radial wave-functions (in atomic units) for Ni atom. It can be seen that the rapidly varying core region has been smoothed in the pseudo-wavefunction. But beyond the cut-off radius, rc, the pseudo- and all-electron- wavefunctions are identical [Ewels, 1997].

The need for pseudo-potential become particularly important for the heavier elements when the number of core electrons vastly outweighs the valence electrons and for which relativistic effects becomes important. The removal of the core electrons allows a non-relativistic approach to be maintained.

Pseudopotentials assume that the localised core states do not take part in bonding, and so can be incorporated as a change in the charge state of the nucleus. Between the nucleus and some cut-off radius r_c, the all-electron valence wavefunctions can be replaced with a smoothed form; beyond this radius the pseudo-wavefunction must be identical to the full all-electron wavefunction. In addition there will still be exchange interactions between core and valence electrons which must be included in the potentials. Finally, if self-consistency is to be easily achieved in the calculations then the potentials must have the

correct atomic charge density outside the core (so called *'norm-conserving'* pseudo-potentials). Generation of such a pseudo-potential satisfies the equation

eqn 2.64: $$\int_0^{r_c} \varphi^*_{all}(r)\varphi_{all}(r)\,dr = \int_0^{r_c} \varphi^*_{pseudo}(r)\varphi_{pseodo}(r)\,dr \ ,$$

where ϕ_{all} is the all electron wave-function and ϕ_{pseudo} is the pseudo wave-function, guarantees the equality of the all electron and pseudo wave-functions outside the core region.

A rather more radical approach to modify the pseudopotential for use in plane-wave calculations has been suggested by Vanderbilt [Vanderbilt, 1990]. The basic aim is the same as before with the additional factor that the norm-conservation rule is relaxed(so called *'ultra-soft'* pseudo-potentials). The resulting wave function then can be expanded using a much smaller plane wave basis set, and hence further reduces the computational cost required.

2.7 Minimisation Techniques

The main drive in a computational study is not simply to find the solutions of real systems more and more accurately, but also to solve them in a faster way in order simulate bigger systems in realistic time scales. Due to this demand, many different techniques have evolved. The fact that the problem of solving the ground state configuration is a variational one *i.e.* only the minimum of the total-energy landscape matters irrespective of how we get there, makes it an open problem to look for increasingly better methods of energy minimisation. In this section we briefly introduce the different techniques followed in the relaxation of the electronic and ionic degrees of freedom to their lowest energy state.

2.7.1 Electronic energy minimisation

Although the matrix diagonalisation (Section 2.2.3) method provides a correct frame of work for the total electronic energy (E_e) minimisation, it is extremely slow and limited to systems of a few atoms only. Instead, most of the present computational packages makes use of two popular methods: the Car-Parrinello molecular dynamics (CPMD) method [Car, 1985] and the conjugate-gradients method [Stitch, 1989; Teter, 1989], which offer solution to systems of hundreds of atoms in a realistic time scale.

The Car-Parrinello molecular dynamics method

Car-Parrinello Molecular dynamics (CPMD) method of minimising the electronic energy follows an indirect scheme, where each of the coefficients a_i of the one-electron KS wave-function (see eqn 2.13) is considered as the co-ordinate of a classical 'particle'. The idea is that, if $\{a_i\}$ can be represented as a set of dynamic variables, they can be treated by the standard classical MD technique and the total electronic energy $E_e[\{a_i\}]$ can be minimised by adding an extra damping term to the equations of motion [Car, 1985]. Although there is no direct physical interpretation of such a representation, it does work and locate the energy minimum much faster than matrix diagonalisation method.

The Lagrangian of the system is written as

eqn 2.65: $$L = \sum_i \mu \langle \dot{\varphi}_i | \dot{\varphi}_i \rangle - E_e[\{\varphi_i\}, R_I] ,$$

where μ is the fictitious mass associated with the electronic wave-function and E_e is the sum of the total electronic eigenvalues. The value of μ is chosen such that it enforces the adiabatic separation between the electronic and ionic degrees of freedom. Note that this decoupling is generally very hard to obtain in metallic systems, where the band gap vanishes and consequently any small energy transfer from ionic to electronic system may cause the electronic wavefunction to go into an excited state [Payne, 1992]. Taking the above Lagrangian into the standard Lagrange equations of motion

eqn 2.66: $$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{\varphi}^*_i}\right] = \frac{\partial L}{\partial \varphi^*_i} ,$$

we get the equations of motion for our electronic system as,

eqn 2.67: $$\mu \ddot{\varphi}_i = -h_{ks}\varphi_i .$$

If we add the constraint of orthonormality between different electronic state,

eqn 2.68: $$\int \varphi^*_\alpha(\boldsymbol{r})\, \varphi_\beta(\boldsymbol{r})\, d^3\boldsymbol{r} = \delta_{\alpha\beta} ,$$

the Lagrangian the equations of motion becomes

eqn 2.69: $$\mu \ddot{\varphi}_i = -h_{ks}\varphi_i + \sum_j \lambda_{ij}\varphi_j .$$

λ denotes the Lagrange multipliers associated with the constraints. The equations of motion than can be solved with any standard integration package (*e.g.* velocity-Verlet, [Swope, 1982]). If we compare eqn 2.67 and eqn 2.69 with Newton's equation of motion ($F = ma$) it looks like the wave-function is constantly being shaped by a forcing agent h_{KS} at a constant fictitious energy surface (note from eqn 2.65 the total energy is not the same as the Kohn-Sham total energy functional). The energy minimisation could simply be incorporated by adding a damping term of the form $-\gamma \dot{\varphi}_i$ in the equations of motion or using some other technique [Payne, 1992]. A simulated annealing technique is often followed, where the system is periodically heated and then quench back to a lower temperature, in order to avoid the system getting trapped in some local minimum [Payne, 1992].

The conjugate gradients method

Both the CPMD and matrix diagonalisation methods are indirect search methods for the ground state, $E_{e,0}$. Instead of directly looking for a lower value of E_e, the downhill motion in the energy landscape is driven by the construction of a new Hamiltonian (h_{ks}) at the end of each iterative step (see eqn 2.21 and eqn 2.67). As a result, these methods risks becoming unstable, specially when the total energy (E_e) is close to the minimum (E_{e0}) [Payne, 1992].

A different and more generic approach often adopted is that of conjugate gradients [Stitch, 1989; Teter, 1989], which assess the energy landscape (E_e) in the phase space of NxN_b dimension (where N is the total number of electron and N_b is the total number of basis functions to describe each electron) directly and moves in an optimum direction to reach the minimum quickly. It was found that the energy minimum E_{e0} could be reached more quickly by assessing the value of E_e and its gradients at a set of neighbouring points, and moving along the direction of large gradients such that the new direction is always independent (conjugate) of the previous ones.

Variation of the conjugate gradients method

An earlier version of the conjugate gradients method is the steepest descents method, which, instead of moving in an independent direction, lowers the functional along each dimension (of the phase space) one at a time. Although this approach minimises the system *in a single dimension* much faster than the conjugate gradients, the strict orthogonality in search directions incorporates a bigger error in the overall move in the multi-dimensional space.

A more recent variation of the conjugate gradients method is density-mixing [Kresse, 1996] which is much faster and works particularly well with metals [Segall, 2002]. In this method the sum of the electronic eigenvalues is minimised in the fixed potential instead of the self-consistent minimisation of the total energy. The new charge density at the end of minimisation is mixed with initial density and the process is repeated until convergence is reached. The conjugate gradients approach is used to minimise the sum of eigenvalues. The disadvantage of this method is that it is also an indirect search method and is reported to be highly unstable in many cases [Segall, 2002]. Instead of using the KS total energy functional it uses a different energy functional (Harris functional; [Harris, 1985]), which does not necessarily have a variational minimum at the SCF solution.

2.7.2 Ionic energy minimisation

So far, total energy calculations of electronic system (E_e) have been described for a *fixed set* of ionic $\{R_p\}$ positions. So the minimum obtained using any of the above electronic minimisation scheme is the self consistent minimum (or electronic ground state) corresponding to that fixed set of ionic co-ordinates. In order to obtain the ground state of the whole system, obviously the ions should also be relaxed to their minimum energy configuration. This is generally done in two steps, first: calculate the forces on the ions, and second: move the ions according to the force to minimise the energy.

Force calculation

The forces on the ion are calculated using the Hellmann-Feynman theorem [Feynman, 1939], which states that: w*hen each electronic wave-function is an eigenstate of the Hamiltonian* (*i.e.* the electrons are at their electronic ground states), then the *partial* derivative of the total energy with respect to the ionic positions gives the forces on the ions. From the standard definition, the force on an ion, P, at a position R_P can be obtained from the *full* derivative of the total energy, E, given by

eqn 2.70: $$F_P = -\frac{dE}{d\boldsymbol{R}_P} .$$

As the electronic wave-functions change, the force on the ions will be altered, therefore the full derivative has to be expressed in terms of changes in the electronic wave-function too,

eqn 2.71: $$F_P = -\frac{\partial E}{\partial \boldsymbol{R}_P} - \sum_i \frac{\partial E}{\partial \phi_i}\frac{\partial \phi_i}{\partial \boldsymbol{R}_P} - \sum_i \frac{\partial E}{\partial \phi_i^*}\frac{\partial \phi_i^*}{\partial \boldsymbol{R}_P} .$$

Given that the electrons are at their electronic ground states, *i.e*

eqn 2.72: $$E = \langle \Psi_e | H | \Psi_e \rangle ,$$

taking eqn 2.72 in eqn 2.71 and using simple manipulation, we get the last two terms as,

eqn 2.73: $$\sum_i \left[\frac{\partial E}{\partial \phi_i} \frac{\partial \phi_i}{\partial \boldsymbol{R}_P} + \frac{\partial E}{\partial \phi_i^*} \frac{\partial \phi_i^*}{\partial \boldsymbol{R}_P} \right] = \sum_i \varepsilon_i \frac{\partial}{\partial R_P} \langle \phi_i | \phi_i \rangle .$$

Here ε_i is the energy of the electronic state ϕ_i. This is zero as $\langle \phi_i | \phi_i \rangle$ is a constant by normalisation. So we can write down the ionic forces as

eqn 2.74: $$f_P = -\frac{\partial E}{\partial R_P},$$

eqn 2.75: $$f_P = -\left\langle \Psi_e \left| \frac{\partial H}{\partial R_P} \right| \Psi_e \right\rangle$$

This significantly simplifies the calculation of the wave-function derivatives, and the forces on the ions can now be simply calculated from the ground-state electronic function by applying the operator $\partial H / \partial R_P$.

Ionic relaxation

New ionic positions under the influence of the Hellmann-Feynman forces are often calculated using one of the following two methods:

i) integrating the ionic equations of motions using the standard molecular dynamics (MD) methods and performing simulated annealing dynamics, exactly in the same way as in the CPMD minimisation of electrons (Section 2.7.1).

ii) direct minimisation techniques, which involves the optimisation of the Hessian matrix (partial second derivative of energy matrix, often known as force constant matrix). The most popular of this type of search method is Broydon-Fletcher-Goldfarb-Shanno or BFGS method [Broydon, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970].

A natural consequence of the Car-Parrinello (CP) formalism is that the electronic and ionic relaxation can take place simultaneously. In a simultaneous optimisation method there are two potential dangers that has to be taken care of. (i) The adiabatic separation has to be maintained between electrons and ions. This is done by assigning a very small fictious mass to the electronic degrees of freedom. (ii) The fact that the process is simultaneous (*i.e.* we are not waiting for electrons to reach their self consistency before calculating the forces on the ions), will incur some errors due to the electrons not being in the exact ground state at the time of force calculation. However, keeping the time step of the dynamics very small, and consequently keeping the error very small and bounded in time (due to error cancellation [Remler, 1990]), this problems can be avoided in CP dynamics.

Another way to avoid these problems will be to treat both systems separately using different optimisation methods, *i.e.* say conjugate-gradients for electronic minimisation and simulated annealing (MD method) for ionic minimisations. There are many different methods available for geometry optimisation including Monte Carlo based minimisation approach. Any combination of these optimisation techniques could be followed for the ionic and electronic part, so long as the two basic condition mentioned in the last paragraph are fulfilled.

2.8 Conclusion

This chapter has introduced the main features of first-principles simulation. Modeller has a range of tools at their disposal to calculate practically any chemical or physical property of a given system at any condition without bias. Impressive results are possible in frontier research areas like, bulk and surface properties of both metals and semi-conductors [Payne, 1990], minerals and oxides [Payne, 1997], molecules, molecular crystal and liquid crystals [Clark, 1997], defects [Shah, 1995], finite temperature phenomena [Ackland, 2001], phase transition, disorder and polymorphism [Clark, 1995], to name but a few.

2.9 Bibliography

1. Ackland, G. J., and S. J. Clark, *Mol. Phys.*, **99**, 855(2001).
2. Allen, M. P., and D. J. Tildesley, *Computer simulation of liquids,* (Clarendon Press, Oxford, 1987).
3. Ashcroft, N. W., and N. David Mermin, *Solid state physics*, (W.B. Saunders Company, 1976).
4. Broyden, C. G., *J. of the Inst. for Math. and Applications*, **6,** 222(1970).
5. Car, R., and M. Parrinello. *Phys. Rev. Lett.*, **55**, 2471 (1985).
6. Ceperley, D. M., and B J Alder, Phys. Rev. Lett. **45** 566(1980).
7. Chadi, D. L., and M. L. Cohen. *Phys. Rev. B*, **8**, 5747 (1973).
8. Clark, S. J., C. J. Adam, G. J. Ackland, J. A. White and J. Crain, *Liquid Crystals*, **22**, 469(1997).
9. Clark, S. J., G. J. Ackland and J. Crain, *Phys. Rev. B*, **52**, 15035(1995).
10. Ewels, C. P., *PhD Thesis*, (University of Exeter, July 1997).
11. Fermi, E., and E. Amaldi, *Accad. Ital. Rome*, **6**, 119(1934).
12. Fermi, E., *Z. Phys.*, **48**, 73(1928).
13. Fetter, A. L., J. D Walecka, *Quantum theory of many-particle systems*, (London: McGraw-Hill, 1971).
14. Feynman, R. P., *Phys. Rev.*, **56**, 340(1939).
15. Fletcher, R., *Computer Journal*, **13**, 317 (1970).
16. Fock, V., *Z. Physik,* **61**, 126 (1930).
17. Francis, G. P., and M. C. Payne, *J. Phys.: Cond. Matt.*, **2,** 4395 (1990).
18. Gillan, M. J, *Contemp. phys.*, **38**(2), 115(1997).
19. Goldfarb, D., *Mathematics of Computation*, **24**, 23 (1970).
20. Hamann, D. R., M. Schlüter, and C. Chiang, *Phys. Rev. Lett.* **43**, 1494(1979).
21. Harris, J., *Phys. Rev. B,* **31**, 1770(1985).
22. Hartree, D. R., *Proc. Cambridge. Phil. Soc.*, **24**(89), 426(1928).
23. Hohenberg, P., and W. Kohn., *Phys. Rev. B*, **136**, 864 (1964).
24. Kohn, W., and L. J. Sham. *Phys. Rev A.*, **140**, 1133 (1965).

25. Kresse, G., and J. Furthmuller, *Phys. Rev. B*, **54**, 11169(1996).

26. Kunz, A. B., *Phys. Rev. B*, **12**, 5890(1975).

27. MacDonald, J. K. L., *Phys. Rev.*, **43**, 830(1933).

28. Monkhorst, H. J., and J. D. Pack. *Phys. Rev. B*, **13**, 5188 (1976).

29. Parr, R., and W. Yang, *Density functional theory of atoms and molecules*, (Oxford University Press, 1989).

30. Payne, M. C., *20th Int. Conf. on Physics of Semiconductors*, **1**, 41,(1990).

31. Payne, M. C., M. P. Teter, D. C. Allen, T. A. Arias and J. D. Joannopoulos, *Rev. Mod. Phys.*, **64,** 1045(1992).

32. Payne, M. C., *Phase Transitions*, **61**, 41(1997).

33. Perdew, J. P., and A. Junger, *Phys. Rev. B*, **23**(10), 5048 (1981).

34. Perdew, J. P., J. A. Chevary, S. H. Vosko, K. A. Jackson, D. J. Singh and C. Fiolhas, *Phys. Rev. B*, **46**, 6671(1992).

35. Perdew, J. P., Y. Wang, *Phys. Rev. B,* **45**, 13244(1991).

36. Phillips, J. C., *Phys. Rev.*, **112,** 685(1958).

37. Remler, D. K., and P. A Madden, *Mol. Phys.*, **70**, 921(1990).

38. Riley, K. F., M. P. Hobson, S. J. Bence, *Mathematical methods for physics and engineering,* (Cambridge University Press, 1997).

39. Segall, M. D., P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark and M. C. Payne, *J. Phys.: Cond. Matt.*, **14**, 2717(2002).

40. Shah, R., A. de Vita and M. C. Payne, *J. Phys. Cond. Matt.*, **7**, 6981(1995).

41. Shanno, D. F., *Mathematics of Computation*, **24,** 647(1970).

42. Slater, J. C., *Phys Rev,* **35**, 210(1930).

43. Stitch, I., R. Car, M. Parrinello and S. Baroni, *Phys. Rev. B*, **39**, 4997(1989).

44. Swope, W.C., H.C. Andersen, P.H. Berens and K.R. Wilson, *J. Chem. Phys.*, **76**, 637 (1982).

45. Teter, M. P., M. C . .Payne and D. C. Allan, *Phys. Rev. B,* **40**, 12255(1989).

46. Thomas, L. H., *Proc. Cantab. Philos. Soc.*, **23**, 542(1926).

47. Vanderbilt, D., *Phys. Rev. B,* **41**, 7892(1990).

48. Weiss, A. W., *Phys. Rev.,* **122**(6), 1826(1961).

49. Yin, M. T., and M. L. Cohen, *Phys Rev. B,* **25**, 7403(1982).

Chapter 3

The Hydrogen Bond in Crystalline Ammonia

3.1 Introduction

The hydrogen bond is one of the most important types of intermolecular interaction, with its strength and directionality playing a crucial role in the structure, function and dynamics of nearly all molecular chemical systems [Steiner, 2002]. It has been a major area of research since its discovery nearly 100 years ago. In recent years new concepts have been established and the complexity of phenomena considered has increased dramatically, largely due to the availability of new computational techniques and increased computing performance.

Estimating the energy of such intermolecular interactions in the solid state by experimental means is a complex business, and often all the information a chemist would desire cannot be directly measured. For example, in order to identify the principal driving forces responsible for crystal packing, it would be necessary to determine the strengths of all of the different hydrogen bonds present, rather than an average value that may be obtained on the basis of thermochemical measurements. In recent years chemists have increasingly turned to computational methods, however a criticism that can be levied at traditional computational chemistry methods is that the style of computational modelling

most often adopted is not that appropriate for modelling the solid state. That is to say a dimer model system, or perhaps a small cluster, is constructed in an attempt to model a property of matter that occurs over a three-dimensional network.

Hydrogen bonding has also been described as a non-additive effect; that is the strength of a network of N interconnected hydrogen bonds is not just the sum of N isolated bonds. In these cases, it is therefore unreasonable to split up the network into individual hydrogen bonds and calculate the energy of each one.

Aside from the validity of using isolated dimer or cluster complexes to model solid state periodic properties, a further problem arises from the use of localised basis sets, which is more appropriate than the plane waves for localised system, to estimate binding energies. Any results obtained will suffer from basis set superposition error (BSSE), which will result in an overestimation of binding energies and must be corrected for. Attempts to account for this effect mostly employ the Counterpoise(CP) method [Boys, 1970], however many studies in the literature point to this correction as being rather crude and strongly dependent on the quality of basis sets used [Schwenke, 1984]. The effect of BSSE on H-bonded dimers has been studied recently in detail by Simon *et al.* [Simon, 1996] and by Muguet *et al.* [Muguet, 1995], specifically for ammonia dimers.

For all these reasons, hydrogen bond energies calculated by current computational approaches should always be taken with caution. Using a modelling approach that is capable of simulating periodic arrays of interacting molecules, such as density functional theory with non-localised (*i.e.* plane-wave) basis sets, offers hope of a better method to determine the strength of hydrogen bonds in a periodic lattice. Moreover, this style of computational modelling is satisfyingly versatile; once convergence with respect to basis set has been achieved the dimensions of the periodic cell used can sufficiently be altered to artificially mimic the gas phase or a solid surface or some other chemical environment. This method, popularly known as a super-cell calculation, has been used extensively in our work. Taking the experimental crystal structure as the starting point, the unit-cell parameters and atomic

positions are optimised alternatively until a set of convergence criteria (on energy and forces) is met. Removing all but one molecule from the periodic cell, and stretching the cell vectors (so that the remaining molecule cannot 'see' its periodic self) will allow the energy of one (effectively gas phase) molecule to be determined. As the same basis set and level of theory are employed in both cases, comparing the absolute energies obtained is valid, and yields information concerning changes in molecular conformation between the solid and gas phases and the intermolecular interaction energy. Thus, comparing the energy per molecule in the crystal environment (at 0 K, vibration-free) with that of one molecule in a large periodic cell without relaxing the atomic positions will generate a calculated value for lattice energy. Relaxing the 'gas-phase' model and correcting both phases for the effects of zero-point energy vibrations will then generate a value for the sublimation energy. In most cases for molecular crystals the sublimation energy is mostly due to hydrogen bonding †. So from knowledge of the sublimation energy and the correct crystallographic information, both of which are simultaneously obtained in a plane-wave density functional theory calculation, we can predict the strength of hydrogen bonding in molecular solids.

In this thesis we present the application of plane-wave DFT calculations to determine the lattice energies and the strengths of hydrogen bonds present in the solid state structures of ammonia and BH_3NH_3. The study of the hydrogen bonded structure of solid ammonia is discussed in the present chapter, and the study of the dihydrogen bond in BH_3NH_3 will be covered in the following Chapter.

† It is important to note that with presently available DFT functionals we can not calculate the energy due to dispersion forces (*aka* London force, Van der Waals force). This is an induced dipole - induced dipole interaction, which depends on the polarisability of the interacting molecules and is inversely proportional to the sixth power of their separation. Dispersion forces between molecules are the weakest of the intermolecular forces and is roughly about 10 to 20% of the weak hydrogen bond strength.

3.1.1 Previous works on ammonia

At ambient temperature and pressure solid ammonia phase-I (see Figure 3.1) is a cubic crystal of $P2_13$ symmetry [Boese, 1997]. Each unit cell consist of four ammonia molecules networked to each-other via tri-furcated hydrogen bonds. Each molecule both accepts and donates three hydrogen bonds, each of which deviates significantly (about 20^0) from the almost perfectly linear hydrogen bond observed in water ice [Steiner, 2002].

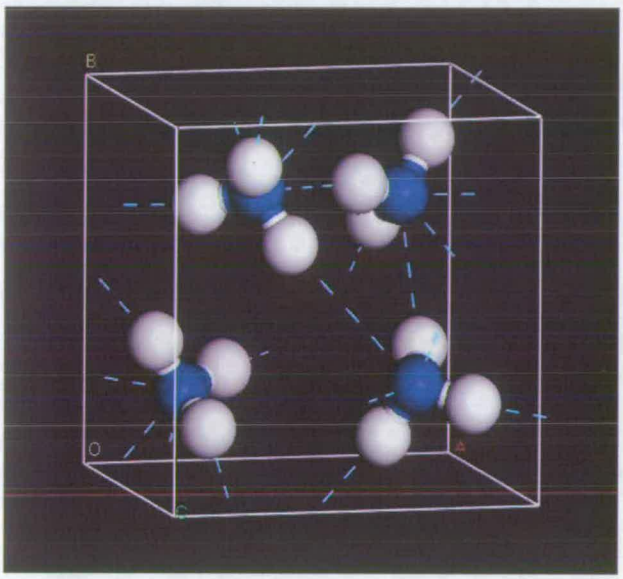


Figure 3.1 : The phase-I crystal structure of ammonia

Despite many difficulties and limitations a number of attempts have been made in determining the strength of the hydrogen bond for ammonia from experiment. Most undergraduate textbooks quote the interaction strength as 17.0 $kJmol^{-1}$, although no reference is cited as to how this value has been obtained [Shriver, 1994]. A very elegant calculation has been reported by Shipman *et al.* [Shipman, 1976], which highlights all the steps necessary to determine the lattice and sublimation energies for the phase I crystal lattice on the basis of thermodynamic measurements. Their value for the sublimation energy [29.0 ± 4.2 $kJmol^{-1}$] gives an hydrogen bond energy of 9.7 ± 1.4 $kJmol^{-1}$, assuming that all the interaction energy in the crystal lattice is due to hydrogen bonding and that each

ammonia molecule is a three-bond donor and acceptor (see Figure 3.1). Another experimental (microwave) study focuses on the gas-phase ammonia dimer [see Figure 3.2 (c)] and places an upper boundary of 11.7 $kJmol^{-1}$ on the bond dissociation energy [Fraser, 1985].

It is worth mentioning at this stage that the interaction present in the ammonia dimer (*i.e.* gas phase) is very different to that which occurs in the condensed state. In the dimer model a standard hydrogen bond occurs, where the nitrogen lone pair donates electrons to one neighbouring hydrogen atom. In the solid state the lone pair donates electrons to three neighbouring hydrogen atoms (*i.e.* a tri-furcated bond, see Figure 3.1). In any case, regardless of whether the measurements relate to the gas or solid phase, both experimental results are clearly at odds with the textbook value.

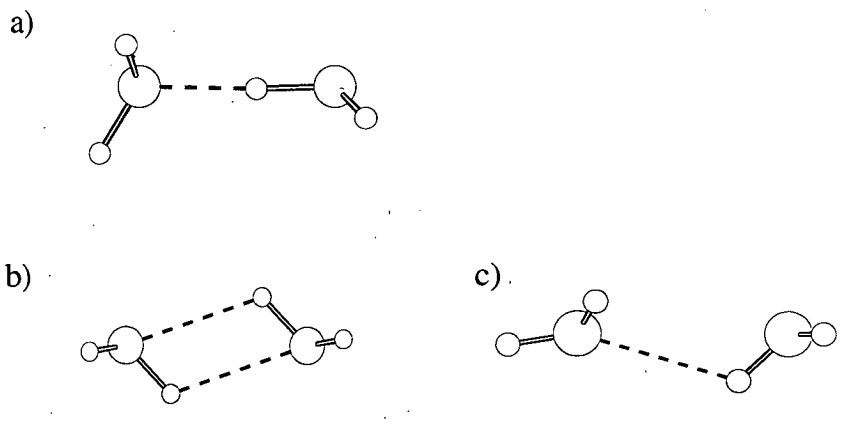


Figure 3.2: (a) classical dimer (C_s symmetry, linear H-bond); (b) cyclic dimer (C_{2h} symmetry), (c) experimental structure from microwave spectroscopy ((C_s symmetry, nonlinear H-bond)).

It is clear from the literature that the potential energy surface created when two ammonia molecules interact is very complicated. The three main dimer structures, which have been reported on the basis of *ab-inito* molecular orbital calculations and spectroscopic experiments, are shown in Figure 3.2. However, these structures appear with various conflicting reports on basis set/level of theory convergence as to which structure represent

the correct global minimum of total energy. Subsequently this leads to variations in the reported binding energy of the dimers that range from 6.6 to 13.2 $kJmol^{-1}$ [Langlet, 1995]. Most *ab-inito* calculations reported refer to a classical dimer structure with C_s symmetry and a single linear hydrogen bond [see Figure 3.2(a)] [Nelson, 1985]. Other calculations refer to a cyclic dimer with C_{2h} symmetry [two hydrogen bonds per dimer, see Figure 3.2(b)] as the lowest energy structure [Hirao, 1984]. However, none of these structures are supported by the microwave experimental study [Fraser, 1985], which postulates a different dimer structure [C_s symmetry, non-linear hydrogen bond, see Figure 3.2(c)] on the basis of measured dipole moments. Moreover, many computational studies point to the $(NH_3)_2$ system as being very sensitive to basis set superposition error, and as such only calculations employing very substantial basis sets should be considered in order to minimise the Counterpoise correction [Novoa, 1995].

Finally, several experiments [Cook, 1979; Lowder, 1970; Lambert, 1950] have reported the enthalpy of formation of the ammonia dimer from free ammonia molecules, which at 18.0 to 19.3 $kJmol^{-1}$, are inconsistent with the upper boundary value of 11.7 $kJmol^{-1}$ from the microwave study, leading Fraser *et al.* to conclude that 'the reason for this discrepancy is not well understood but these (microwave) results suggests that these (thermodynamic) measurements (of enthalpy) of weakly bound complexes are not reliable' [Fraser, 1985].

From this literature survey it is clear that there is much controversy surrounding the geometry of the ammonia dimer system. All these conflicting reports leave the determination of the intermolecular interaction energy in solid-state ammonia as an interesting challenge. It also strongly questions the acceptance of the traditional computational approach (*i.e.* using *ab-inito* molecular orbital calculations on dimer systems) to infer hydrogen bond strength in the condensed phase, particularly in the case of ammonia where the nature of bonding is so very different.

3.2 Computational Section

3.2.1 Crystal lattice calculations

Total energy density functional theory calculations have been carried out using the CASTEP package available through the Materials Studio suite of software [CASTEP, v4.2; Payne, 1992]. Periodic boundary conditions allow the valence electronic wave function to be expanded in terms of a discrete plane-wave basis set which can be continuously improved until a desired convergence level is reached; a considerable advantage over the localised basis sets approach [see Figure 3.3(a)]. For ammonia, the basis set expressed at a cut-off of 550 eV was found to be sufficient to converge total energies to better than 2.0 meV per unit cell with a finite basis set error [dE/dlog (E_{cut})] [Francis, 1990] 0.17 eV per unit cell, which is within the accepted tolerance limits. The electronic core wave function was described using the standard ultrasoft pseudopotentials available with the software package [Vanderbilt, 1990].

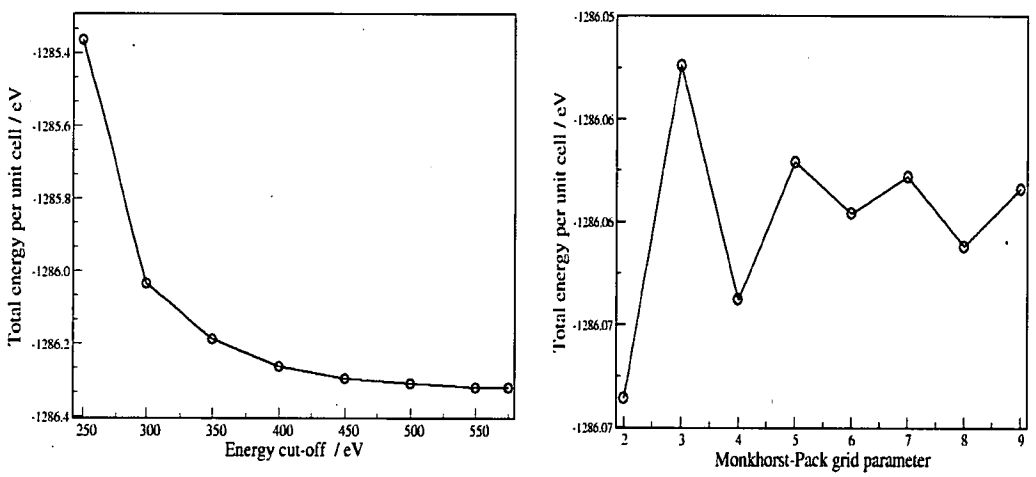


Figure 3.3: The total energy per unit cell as a function of cut-off energy and Monkhorst-Pack grid parameter. (a) The total energy decreases monotonously with the increase in cut-off energy. (b) Total energy per unit cell varies rather inconsistently with different sets of k-points due to inadequate sampling of the first Brillouin zone. With sufficiently large set of k-points the total energy becomes more consistent, and the range over which it varies diminishes.

Bloch's theorem allows the calculation to run in reciprocal space (*i.e.* k-space) which significantly improves calculation efficiency. The symmetry reduced k-point set used to sample the reciprocal space was generated using a Monkhorst-Pack grid [Monkhorst, 1976]. A 6 x 6 x 6 Monkhorst-Pack grid, giving 11 k-points in the symmetry reduced 1st Brillouin zone, was found to sufficiently converge the total energy per unit cell within 4.0 meV [see Figure 3.3 (b)]. The Fast Fourier Transformation (FFT) grid used to communicate between real and reciprocal space was set at 25% in excess of levels prescribed by the program in order to minimise wrap around errors in the transformation [Segall, 2002; CASTEP Manual, v4.2]. The GGA functional PW91 was used for the exchange and correlation functional [Perdew, 1991].

For ammonia, the initial geometry was taken from the X-ray diffraction structure by Boese *et al.* [Boese, 1997] at 160 K and ambient pressure for the phase I cubic crystal system [space group $P2_13$, a = 5.1305(8) Å, Z = 4] (Figure 3.1). Optimisation of atomic positions and unit cell parameters were performed on alternate cycles using the BFGS method until the convergence criteria were met (maximum energy change per atom = 5 x 10^{-6} eV, maximum RMS force = 0.01 eVÅ^{-1}, maximum RMS stress = 0.02 GPa and maximum RMS displacement = 5.0 x 10^{-4} Å).

3.2.2 Super cell calculations

A super cell calculation on a single, isolated ammonia molecule (effectively the gas phase) was carried out in order that the interaction energy in the solid state could be determined. Zero interaction between the nearest neighbouring cells was obtained by increasing the cell size and observing the change in total energy. An 8 x 8 x 8 Å^3 cell was found to break all intermolecular interactions, without giving rise to overly long computational time. The basis set cut-off energy used previously (550 eV) was sufficiently high to avoid the lowering of total energy due to increased number of plane waves in the basis set (as the cell volume increases) and thus the comparison of energies between the two different cell sizes is legitimate. We have estimated any error incurred per hydrogen bond

due to inconsistencies in basis set by calculating the H-bond energy difference due to different cell size. The H-bond energy changes by ~3.0×10^{-3} eV (*i.e.* ~0.3 $kJmol^{-1}$) for changing the unit cell volume form 8 x 8 x 8 $Å^3$ to 9 x 9 x 9 $Å^3$, which is negligible compared to a weak hydrogen bond strength.

3.3 Results and Discussion

Our result matches very closely with the experimental work by Hewat and Rickel [Hewat, 1979] on deutero-ammonia at 2 K and also matches very closely with the most recent work by Boese *et al.* [Boese, 1997] at 160 K . Comparison with solid deutero-ammonia at 2 K is very logical as the isotope effect is almost zero at such a low temperature.

The unit cell volume obtained is within 1.5% of the experimentally determined volume by Boese *et al.* (160 K) and within 6.5% of that determined by Hewat and Riekel (2K). The very close match with the value of Boese *et al.* should not be treated as an indication of the accuracy of the simulation as DFT calculations using GGA functionals often tends to underbind, and hence overestimate unit cell volume [Kurth, 1999; Ernzerhof, 1999; Curtiss, 1997]. The bigger cell volume obtained by Boese *et al.* is expected due to thermal expansion of the unit cell.

Table 3.1: Comparison between the experimental and *ab-initio* structure of Solid NH_3, and H-bond calculation at ambient pressure, 0K.

Parameters	Experimental results		Calculated results			
			Plane wave approach			Gaussian
	[a]Solid phase	[b]Gas phase	Solid phase	[c]Gas phase	[d]Gas phase	[e]Dimer Model
Set-up						
Theory/Method		Microwave	PW91	PW91	PW91	B3LYP
Basis set/E_cut-off [eV]		–	550	550	550	6-31++G (2d,2p)
K-point Grid	–	–	6x6x6	6x6x6	6x6x6	–
FBSC/BSSE	–	–	0.17	~0.0	~0.0	[e]Very large
Lattice [Å, °, Å^3]						
a, b, c	5.048	–	5.156	8.0	8.0	–
α, β, γ	90.0	–	90.0	90.0	90.0	–
Space/point group	$P2_13$	–	$P2_13$	P1	P1	C_s
Cell Volume	128.64	–	137.03	–	–	–
Geometry [Å, °]						
*r*N-H	1.061(5)	1.008(4)	1.029	1.029	1.022	1.017
*r*N....H	2.357(2)	–	2.329	–	–	–
*r*N.....N	3.325(2)	–	3.322	–	–	3.295
∠H–N–H	107.5(2)	107.4(2)	107.8	107.8	108.0	107.4
∠N–H...N	160.0(2)	–	161.8	–	–	–
Energy						
Total energy [eV]			–1286.34248	–321.21804	–321.22068	
Lattice energy [kJmol^{-1}]	36.4(4)		35.4	–	–	–
Sublimation energy [kJmol^{-1}]	28.9(4)		27.8	–	–	–
H-bond strength [kJmol^{-1}]	9.6(14)		9.3	–	–	11.2

[a] [Shipman, 1976; Hewat, 1979]
[b] [Weiss, 1951]
[c] Not relaxed, [d] Relaxed
[e] [Novoa, 1995]

3.3.1 Solid-phase geometry

The N-H distance from Boese *et al.* is smaller than our value by 0.2 Å and the H-bond is bigger by 0.07 Å whereas these values are in very good agreement (within 3-2%) with the Hewat and Riekel data. This might be due to the fact that the position of the H atom is poorly represented in x-ray diffraction. The N–N distance from our calculation is a very good match (within 1%) with all the experiments [Boese, 1997; Hewat, 1979; Olovsson, 1959; Reed, 1961].

Both the H–N–H and N–H...N angles are in good agreement (within 1%) with the experimental values. The H–N–H angle is smaller than the ideal value of 109.5° for sp^3 hybridisation of the nitrogen atom. This shows that bond bending of NH_3 is also present in the solid state geometry.

3.3.2 Gas-phase geometry

The molecular geometry of isolated NH_3, calculated using the super-cell approach and plane wave basis set, was also very close to the experimental gas phase values reported by Weiss *et al.* [Weiss, 1951], and no less accurate than the most advanced localised basis set calculation reported in the literature. For example, we obtained a value for the H–N–H angle of 108.0°, compared to the range 104.2°–108.3° reported by various different *ab-initio* molecular orbital calculations [Boese, 1997]. Our calculation therefore demonstrate the reliability strength of the supercell approach in plane wave DFT calculations to reproduce gas phase geometries.

3.3.3 Gas phase *vs* crystal phase

This work shows the molecular geometry of ammonia changes very little from the gas to the solid phase. The H–N–H bond angle decreases by 0.2° and N–H bond distance

increases by 0.007 Å from gas to solid. These changes are in line with experimental findings by Hewat *et al.* [Hewat, 1979] and by Olovsson *et el.* [Olovsson, 1959], as they justified that the N–H bond length should increase from gas to solid due to hydrogen bond formation. But the work of Boese *et al.* [Boese, 1997] and Reed *et al.* [Reed, 1961] shows the opposite effect. It is very hard to analyse different experimental results when the geometrical changes are very small and when the experiments are carried out by different people using different approaches. Our simulations do not suffer from this problem as both calculations (gas and solid phase) were carried out under the same conditions, and therefore the comparison of the results for the two phases is more legitimate.

3.3.4 Energy

The lattice energy has been calculated from the difference in energy of a single molecule in crystal environment (unit cell energy/Z, Z = 4) and the un-relaxed single molecule in a supercell of size 8 x 8 x 8 $Å^3$ (effectively gas phase). The lattice energy, 35.4 $kJmol^{-1}$, obtained from our calculation is a close match with the experimental lattice energy, 36.4(4) $kJmol^{-1}$, from the thermodynamic measurement by Shipman *et al.* [Shipman, 1976], and within the limit of experimental error. As it should be, a match between the experimental and computational lattice energies should be taken as a crucial test of the overall quality of the calculation.

Two additional steps are required to obtain the sublimation energy from the calculated lattice energy:

(a) relaxing the geometry of the isolated molecule to give the 'proper' gas phase geometry, and

(b) correcting for the zero-point vibration energy (ZPE). (For the detailed method of ZPE correction see Shipman, 1976 [Shipman, 1976])

The first step was easily done with a CASTEP calculation and the lattice energy is reduced down to 35.1 $kJmol^{-1}$ (See Table 3.1). However, for the ZPE correction due to the lattice vibrations, we needed to embark on the mammoth task of a full vibrational analysis of the system, which was not supported and under-development by CASTEP software package during the time of this work. Instead, we adopted the experimental value of 7.3 $kJmol^{-1}$ reported by Shipman *et al.* [Shipman, 1976] as the approximate zero-point vibrational correction. This gives us a sublimation energy of 27.8 $kJmol^{-1}$ for solid ammonia. Taking the fact that each ammonia molecule in the crystal is connected to it's neighbours by six hydrogen-bond (*i.e.* the intermolecular bond-order is 3), we estimate the average value of H-bond in the ammonia crystal to be 27.8/3 $kJmol^{-1}$ = 9.3 $kJmol^{-1}$.

It is interesting to note that, presently, three years after this work was completed, the full vibrational analysis of molecular solids is now supported in the CASTEP material studio suite [Segall, 2002]; and also in the later part of this thesis (Chapter 7 and Chapter 8), we have subsequently reported a full vibrational analysis of solid NH_3 phase I as a part of our development of the constraint first-principles molecular dynamics code.

Needless to say that the order of the ZPE correction is rather small compared to the lattice energy [Shipman, 1976], and hence it is anticipated that any error incurred, will not change our final result of H-bond strength (9.3 $kJmol^{-1}$) by a great deal. Our later work (Chapter 8, Section 8.4) has given a value of 38.52 $kJmol^{-1}$ for the ZPE of the 21 ammonia lattice modes at gamma point; and the difference between the solid phase and gas phase molecular mode ZPE is obtained as 2.081 $kJmol^{-1}$ (Chapter 7, Table 7.4). These give the ab-initio value of ZPE correction as (38.52/4 - 2.081) $kJmol^{-1}$ = 7.55 $kJmol^{-1}$. This is surprisingly close to the value, 7.3 $kJmol^{-1}$, reported by Shipman *et al.* [Shipman, 1976], which was rather crudely calculated from 12 translational mode of average frequency 129 cm^{-1} and 12 librational mode of average frequency 280 cm^{-1}, with no further consideration of ZPE correction from the molecular modes.

3.3.5 Electrostatic interaction

In the traditional view, the lone pair of nitrogen is donated towards the hydrogen atom in hydrogen-bond formation, rather than just electrostatically shared [Steiner, 2002]. However, the value of 9.3 $kJmol^{-1}$ for the hydrogen-bond in the ammonia crystal is very close to that of the standard hydrogen-bond in the gas phase ammonia dimer. This strongly suggests that the nature of the interaction in this particular hydrogen-bonded system is electrostatic *i.e.* the lone pair electron in N is equally shared by all three hydrogen atoms.

The beauty of DFT is that it gives an unique insight into the electron density of the system in order to shed more light on this problem. The error in the electron-density map from a DFT simulation is of the same order of magnitude as the error in the total energy calculation. As the electron-density is the central quantity in DFT (see Chapter 2), the electron density map so obtained is free from any sort of error incurred due to any result analysis. From the electron density map it is possible to calculate the Mulliken charges [Sanchez, 1995; Segall, 1996; Mulliken, 1955] on each of the ionic positions.

If the interaction is electrostatic, we would expect the Ewald sum [Allen, 1987] over these Mulliken charges in the unit cell to generate the same total interaction energy as that of the DFT calculation. From our calculation of the optimised ammonia structure we get Mulliken charges of 0.383 e on each hydrogen and -1.149 e on each nitrogen ionic positions. Note that, unlike the electron density, the Mulliken charges are derived quantity [Mulliken, 1955] from the Kohn-Sham single electron wave-function (written in plane wave basis), which itself is an auxiliary quantity in a DFT simulation. However, the Ewald sum over the Mulliken charges generated from our ground state optimised structure using GULP (General Utility Lattice Program, [Gale, 1997]) gives the total interaction energy as 36.1 $kJmol^{-1}$ per molecule, which is reasonably close to our DFT calculated value (lattice energy is 35.4 $kJmol^{-1}$, see Table 3.1). This means that the solid ammonia phase I is an ideal electrostatic system. Here we emphasise, that this is a very crucial outcome of this work and as far as we are aware this is the first comparative study of standard versus tri-furcated hydrogen bond strengths.

3.4 Conclusion

This is the first standard *ab-inito* calculation on a periodic system to reveal the hydrogen bond strength in a molecular solid and is in very good agreement with experimental findings [Shipman, 1976]. The result (9.3 $kJmol^{-1}$) was 'surprising' as it indicated that the trifurcated N-H..N bond in solid ammonia was not much weaker than the standard interaction observed in the gas phase dimer (11.2 $kJmol^{-1}$) [Novoa, 1995]. This raised the question that the traditional way of thinking of the nitrogen lone pair of electrons as partially shared in hydrogen bond formation for this system is incorrect, and the nature of the interaction must be mostly electrostatic. Both results from our periodic lattice model and literature isolated dimer models emphasise that the value of the hydrogen bond strength is much less than 17.0 $kJmol^{-1}$, as reported in most undergraduate chemistry textbooks [e.g. Shriver, 1994], which indicate that the texts should be carefully reviewed in light of modern computational investigation.

3.5 Bibliography

1. Allen M. P., and D. J. Tildesley, *Computer simulation of liquids,* (Clarendon Press, Oxford 1987).

2. Boese, R., M. Yu Antipin and P.R. Mallinson, *J. Phys. Chem. B*, **101**, 5794(1997).

3. Boys, S. F., and F. Bernardi, *Mol. Phys.*, **19**, 553(1970).

4. Cook, K.D., and J.W. Taylor, *Int. J. Mass. Spectrom. Ion Phys.*, **30**, 345(1979).

5. Curtiss, L. A., K. Raghavachari, P. C. Redfern and J. A. Pople, *J. Chem. Phys.*, **106**, 1063(1997).

6. Ernzerhof, M., and G. E. Scuseria, *J. Chem. Phys*, **110**(11), 5029(1999).

7. Francis, G. P., and M. C. Payne, *J. Phys.: Cond. Matt.*, **2,** 4395 (1990).

8. Fraser, G.T., D. D. Nelson, A. Charo and W. Klemperer, *J. Chem. Phys.*, **82,** 2535 (1985).

9. Gale, J. D., *GULP*, *J. Chem. Soc. Faraday Trans.*, **93**, 629(1997).

10. Hewat, A. W., and C. Rickel, *Acta Crystallogr.*, **35**, 569(1979).

11. Hirao, K., T. Fujikawa, H. Konishi and S. Yamabe, *Chem. Phys. Lett.*, **104**, 184(1984).

12. Kurth, S., J. P. Perdew , P. Blaha, *Int. J. Quant. Chem.*, **75**(4-5), 889(1999).

13. Lambert, D., and E. D. Strong, *Proc. R. Soc. London Ser. A*, **200**, 566(1950).

14. Langlet, J., J. Caillet and M. Caffarl, *J. Chem. Phys.*, **103**, 8043(1995).

15. Lowder, J. E., *J. Quant. Spectrosc. Radiat. Transfer,* **10**, 1085(1970).

16. Monkhorst, H. J., and J. D. Pack. *Phys. Rev. B*, **13**, 5188 (1976).

17. Muguet, F. F., and G. W. Robinson, *J. Chem. Phys.*, **102**(9), 3648(1995).

18. Mulliken, R. S., *J. Chem. Phys.*, *23*, 1833(1955).

19. Nelson, D. D., G. T. Fraser and W. Klemperer, *J. Chem. Phys.*, **83,** 6201 (1985).

20. Novoa, J. J., and C. Sosa, *J. Phys Chem*, **99**, 15837(1995).

21. Olovsson, I., and D. H. Templeton, *Acta Crystallogr.*, **12**, 832(1959).

22. Payne, M. C. M. P. Teter, D. C. Allan, T. A. Arias, J. D. Joannopoulos, *Rev. Mod. Phys.*, **64**, 1045(1992); CASTEP, Academic version 4.2 , (licensed under the UKCP-MSI Agreement, 1999).

23. Perdew, J. P., and Y. Wang, *Phys. Rev. B*, **45**, 13244(1991).

24. Reed, J. W., and P.M. Harris, *J. Chem. Phys.*, **35**,1730(1961).

25. Sanchez, P. D., E. Artacho and J. M. Soler, *Solid State Commun.*, **95**, 685(1995).

26. Schwenke, D. W., and D. G. Truhlar, *J. Chem. Phys.*, **82**, 2418(1984).

27. Segall, M. D., C. J. Pickard, R. Shah and M. C. Payne, *Mol. Phys.*, **89**, 571(1996).

28. Segall, M.D., P. J. D. Linden, M.J. Probert, C.J. Pickard, P. J. Hasnip, S.J. Clark and M. C. Payne, *J. Phys.: Cond. Matt.*, **14**, 2717(2002).

29. Shipman, L. L., A. W. Burgess and H. A. Scheraga, *J Phys Chem,* **80,** 52 (1976).

30. Shriver, D. F., P. Atkins, C. H. Langford, *Inorganic chemistry,* 2nd ed., (Oxford University Press, 1994).

31. Simon, S., M. Duran and J. J. Dannenberg, *J. Chem. Phys.*, **105**(24), 11024(1996).

32. Steiner, T., *Angew. Chem. Int. Ed.*, **41**, 48(2002).

33. Vanderbilt, D., *Phys. Rev. B,* **41**, 7892(1990).

34. Weiss, M. T., and M.W.P. Strandberg, *Phys Rev.*, **83**, 567(1951).

Chapter 4

The Dihydrogen Bond in Crystalline BH_3NH_3

4.1 Introduction

Classic hydrogen bonds D–H...A involve a weak hydrogen bond donor D–H, where D is a typical electrostatic atom such as nitrogen or oxygen, and a partner atom with a lone pair electrons to act as a hydrogen bond acceptor (e.g. O, N, S, halides *etc*). It is surprising that a number of transition metal – hydrogen σ bonds, M–H, are also able to act as proton acceptors to give D–H...H–M interactions with strengths comparable to standard hydrogen bonds. This was first described by Richardson and co-workers, and they suggested the term 'dihydrogen bond' to describe this type of interaction [Richardson, 1995]. Both inter-and intra-molecular versions of this bond have been cited, with typical energies in the range of 12–28 $kJmol^{-1}$ [Crabtree, 1996].

In the transition metal cases, however, we could not be sure that the acceptor was indeed the M-H σ bond because these metals also possess non-bonding valence electrons, which could in principle act as alternate H-bond acceptors [Klooster, 1999]. It was discovered that the transition metal could be replaced with boron and the resulting contact D–H...H–B was still 'unexpectedly strong' [Richardson, 1995]. This interaction has received

special attention as the absence of non-bonding valence electrons on boron negates a possible $H^{\delta+}...^{\delta-}M$ interaction that could arise if M were a 'd' block element. The H...H contact must therefore result from a D–H...σ bond type interaction. The simplest system to exhibit this interaction is solid BH_3NH_3. Here the electronegative N pulls e^- towards itself, creating a strong $N^{\delta-}$–$H^{\delta+}$ dipolar group. When this approaches a neighbouring BH_3 group it interact with the negatively charged H–B bond via N–H...H–B σ interaction [Klooster, 1999]. It is important to note that in contrast to N–H bond, both atoms of the B-H bond are negatively charged, and the bond is only weakly dipolar with the boron as the negative end [Crabtree, 1995].

Since the discovery of dihydrogen bonds in solid BH_3NH_3, many theoretical efforts have been put forward, especially using a dimer model (see Figure 4.1) as the simplest system that contains this interaction. Calculations at the PCI-80/B3LYP level were performed by Richardson, Gala, and Crabtree [Richardson, 1995] and Klooster *et. al.* [Klooster, 1999], at the levels of HF/6-31** and MP2/6-31G** by Popelier [Popelier, 1998] and by Li *et. al.* [Li, 2002] at the MP2/6-31++G** level. Popelier discussed the characteristics of the B–H...H–N interaction on the basis of electron density. Cramer and Gladfelter [Cramer, 1997] made a comparative study of the dimer conformation and binding energy with different levels of theory and discussed the potential energy surface.

All these studies helped open up a new scientific field. But the image obtained for dihydrogen bonds is yet not very clear. Do the dihydrogen bonds hold the same characteristics of classic H bonds? The question arises as most of the papers cite a 'surprisingly high' energy value for such an interaction, following on from the main computational work carried out by Crabtree and co-workers [Figure 4.1(a); Richardson, 1995; Klooster, 1999]. They have reported 25.3 $kJmol^{-1}$ for the dihydrogen bond strength from their dimer model calculations, which they then equate to the interaction present in the BH_3NH_3 crystal, *i.e.* very similar to a medium-strength hydrogen bond, which is at odds with that expected compared to standard H-bonds [Steiner, 2002; Jeffery, 1997].This leads to an interesting question that, like the ammonia system reported in the previous chapter,

inappropriate computational modelling may be resulting in anomalous results for the strengths and properties of intermolecular interaction.

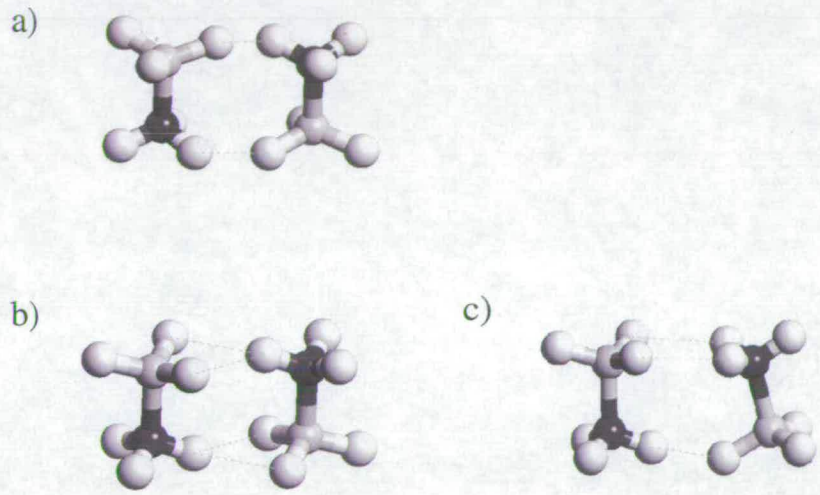


Figure 4.1 Dimer models used by (a) Crabtree and co-workers, (b) Li *et. al.* and (c) Popelier to investigate properties of the H···H contacts.

Like the ammonia case, there is also no clear agreement over which dimer conformation of $(BH_3NH_3)_2$ is the true minimum. Li *et. al.* [Figure 4.1(b); Li, 2002] and Popelier [Figure 4.1(c); Popelier, 1998] have reported two different conformations, and both claim their models reproduce *the local geometry* of the dihydrogen bond in the crystal. However the two different models produce markedly different values for the interaction energy at 13.9 $kJmol^{-1}$ and 19.0 $kJmol^{-1}$, respectively. It therefore looks like the use of the dimer model is a bad approach to rationalise the binding energy in the solid state structure of BH_3NH_3, as the potential energy surface is complex and different minima can be obtained [for further detailed discussion, see (Cramer, 1997) where an argument has been submitted whether the C_2, Figure 4.1(a), (found by Crabtree *et. al.*) or C_{2h} symmetry conformation, Figure 4.1(b), of the $(BH_3NH_3)_2$ corresponds to the global minimum]. In addition, we could

not find any report of an experimental dimer structure either, which leads to the important conclusion that all postulated models are, in all likelihood, completely artificial.

In reality, of course, the problem we wish to address looking at is one of crystallographic packing (see Figure 4.2). The dihydrogen bonds that appear in the dimer models (Figure 4.1) are quite different compared to those in the crystal (Figure 4.2). This leads to an obvious question, could Crabtree's 'surprisingly large' result just be an artefact of an inappropriate dimer model calculation? For this purpose we undertook a study of the BH_3NH_3 crystal using plane-wave DFT to simulate the exact lattice geometry, and thus determine the strength of dihydrogen bonds present in a much more relevant model system.

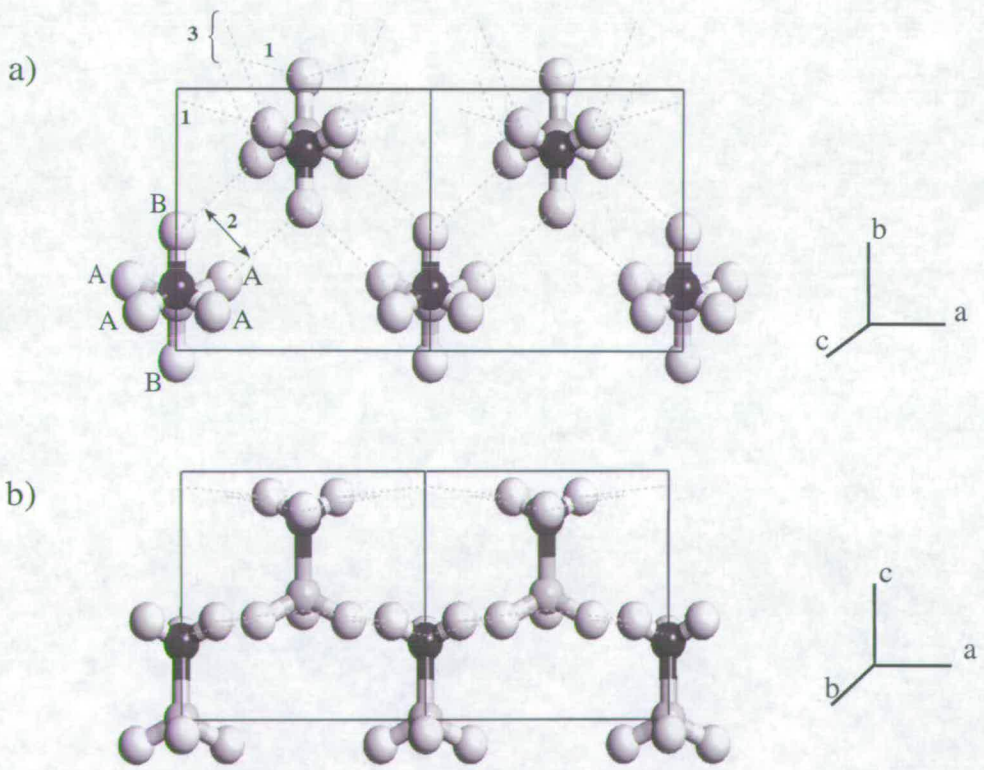


Figure 4.2 Crystal structure of BH_3NH_3 showing the dihydrogen intermolecular bonding network as viewed along (a) the c axis (numbers indicate three different types of H-bond) and (b) the b axes [Klooster, 1999]. Note N/B–H geometries along the b and a axes are not crystallographically equivalent [Klooster, 1999].

4.2 Computational Section

4.2.1 Crystal lattice calculations

Density functional theory calculations have been carried out to simulate the gas and solid phase structure of BH_3NH_3 and a dimer conformation, using the CASTEP package available through the Materials Studio suite of software [CASTEP, v4.2; Payne, 1992]. For the BH_3NH_3 crystal calculation, the basis set expressed at a cut-off of 600 eV was found to be sufficient to converge total energies to better than 3.0 meV per unit cell, with a finite basis set correction [dE/dlog (E_{cut})] [Francis, 1990] 0.22 eV per unit cell, which is within the accepted tolerance limits. The electronic core wave function was described using the standard ultrasoft pseudopotentials available with the software package [Vanderbilt, 1990]. The symmetry reduced k-point set used to sample the reciprocal space was generated using a Monkhurst-Pack grid [Monkhorst, 1976] (6 x 6 x 6, giving 27 k-points in the symmetry reduced 1st Brillouin zone). The Fast Fourier Transformation (FFT) grid used to communicate between real and reciprocal space was set at 25% in excess of levels prescribed by the program in order to minimise wrap around errors in the transformation [Segall, 2002, CASTEP Manual, v4.2]. The GGA functional PW91 was used for the exchange and correlation functional [Perdew, 1991].

The initial geometry of the BH_3NH_3 crystal was taken from the neutron diffraction structure by Klooster *et. al.* [Klooster, 1999] (An orthorhombic crystal system, space group $Pmn2_1$, a = 5.395(2), b = 4.887(2), c = 4.986(2) Å, Z = 2). Optimisation of atomic positions and unit cell parameters were performed on alternate cycles using the BFGS method until the convergence criteria were met (maximum energy change per atom = 5.0×10^{-6}, maximum RMS force = 0.03 eV $Å^{-1}$, maximum RMS stress = 0.02 GPa and maximum RMS displacement = 5.0×10^{-4} Å).

4.2.2 Super cell calculations

A super cell calculation on a single, isolated BH_3NH_3 molecule (effectively the gas phase) was carried out in order to determine the total interaction energy in the solid state. Zero interaction between the nearest neighbouring cells was ensured by increasing the cell size to 10 x 10 x 10 $Å^3$. The basis set cut-off energy (600eV) used previously for the crystal phase was sufficiently high to avoid the lowering of total energy due to increased number of plane waves in the basis set (as the cell volume increases) and thus the comparison of energies between these two different cell sizes is legitimate.

4.3 Results and Discussion

4.3.1 Crystal geometry

Our calculated geometry matches very closely with the experimental data (Table 4.1) by Klooster *et. al.* [Klooster, 1999] for the solid phase and by Thorne *et. al.* [Thorne, 1983] for the gas phase structure of BH_3NH_3.

The unit cell volume obtained is within 2.3% of the experimentally determined volume by Klooster *et. al.* (200 K). This is very good match, as DFT calculations using GGA functionals often tends to underbind, and hence overestimate the unit cell volume [Kurth, 1999; Ernzerhof, 1999; Curtiss, 1997]. The lattice vectors 'a' and 'b' obtained from calculation are in very good agreement with the experimentally determined values. However the lattice vector 'c' is about 2% bigger than that determined by experiment. This is probably due to the lack of strong intermolecular interactions along the 'c' direction that DFT could model, as all three dihydrogen bonds lie in the plane containing lattice vectors 'a' and 'b' (Figure 4.2).

Table 4.1 Comparison between the experimental and *ab-initio* structure of solid and gas BH_3NH_3 and H–bond calculation at ambient pressure, 0 K.

Values	Experimental results		Calculated result Periodic approach		Gaussian approach
	[a]Solid phase	[b]Gas phase	Solid phase	Gas phase	[c]Dimer
Set-up					
Theory/Method	Neutron Diffraction	Microwave Study	DFT-PW91	DFT-PW91	PCI-80/ B3LYP
Basis set/E_cut-off	–	–	600 eV	600 eV	Double ζ P
K-point Grid	–	–	6x6x6	4x4x4	–
FBSC/BSSE	–	–	0.22	~0.0	Not Reported
Lattice [Å, °, Å^3]					
a	5.395(2)	–	5.366	10	–
b	4.887(2)	–	4.914	10	–
c	4.986(2)	–	5.100	10	–
α, β, γ	90.0	–	90.0	90.0	–
Z	2	–	2	1	–
Volume	131.5(16)	–	134.5	1000	–
Space/point group	$Pmn2_1$	–	$Pmn2_1$	P1	C_2
Geometry [Å, °]					
*r*B–N	1.58(2)	1.6576(16)	1.580	1.624	N/A
*r*N–H	1.07(4)[d], 0.96(3)[e]	1.0140(20)	1.029, 1.034*	1.026	N/A
*r*B–H	1.15(3)[d], 1.18(3)[e]	1.2160(17)	1.218, 1.210	1.205	N/A
∠B–N–H	106(4)[d], 111(2)[e]	110.28(14)	113.0, 111.1	111.2, 110.7	N/A
∠N–B–H	114(2)[d], 112(1)[e]	104.69(11)	107.4, 109.2	105.8, 105.5	N/A
∠H–N–H	113(3)[f], 104(3)[g]	108.65(14)	107.2, 107.4	108.0	N/A
∠H–B–H	102(2)[f], 116(3)[g]	113.80(11)	110.3, 110.2	112.9	N/A
bond:1 *r*H...H	2.02(3)	–	1.909	–	1.82
∠N–H...H	156(3)	–	157.8	–	158.7
∠B–H...H	106(1)	–	100.4	–	98.8
bond:2 *r*H...H	2.21(4)	–	2.271	–	–
∠N–H...H	130(1)	–	131.8	–	–
∠B–H...H	156(3)	–	157.9	–	–
bond:3 *r*H...H	2.23(4)	–	2.174	–	–
∠N–H...H	137(2)	–	135.0	–	–
∠B–H...H	94(2)	–	87.6	–	–
Energy					
Total energy [eV]			–896.439167	–447.431416	
Total interaction energy [kJmol^{-1}]		–	76.0	–	50.6
H-bond strength [kJmol^{-1}]		–	12.7	–	25.3

[a] [Klooster, 1999] [b][Thorne, 1983] [c][Richardson, 1995]

[d] Geometry relates to the atom H_B in Fig.4.2 [e] Geometry relates to the atom H_A in Fig. 4.2

[f] Geometry relates to H_B–N/B–H_A angle in Fig.4.2 [g] Geometry relates to H_A–N/B–H_A angle in Fig. 4.2

The calculated B–N distance in the solid phase is identical as that obtained in the experiment to within the experimental error and shorter than that of the gas phase. The N–H, B–H distances and H–N–H, H–B–H angles are in reasonable agreement with the experimental values by Klooster *et. al.*

Three crystallographically distinct dihydrogen bonds (Figure 4.2) are present in the crystal lattice, with bond distances 2.02(3) Å, 2.21(4) Å and 2.23(4) Å, which are all reproduced in the calculation to within 0.1 Å. The corresponding bond angles (B/N—H...H) are also in reasonable agreement (within 1–8%) with experimental data. Although there is a reasonable agreement between the experimental and calculated geometries; as the reported experimental structure shows large estimated standard deviations [Klooster, 1999], we believe that a better structure could be obtained *experimentally*.

4.3.2 Gas-phase Geometry

The geometry obtained in the supercell calculation of gas phase BH_3NH_3 is as accurate as that obtained with advanced-level Gaussian basis sets [Thorne, 1983]. The B–N distance of BH_3NH_3 is 1.624 Å, within 0.04 Å of the experimental value, and bigger than the calculated value in the solid phase, which is presumably shortened due to σ-bond co-operativity.

4.3.3 Energy

The total interaction energy has been calculated from the difference in the energy of a single molecule in the crystal environment (unit cell energy/Z, Z = 2) and the relaxed single molecule in the supercell of size 10 x 10 x 10 $Å^3$ (effectively gas phase). The value ($kJmol^{-1}$) obtained from our calculation does not include the zero-point vibrational energy (ZPE) correction. As mentioned earlier for ammonia (Chapter 3), at the time this work was carried out we did not have access to the proper tools in the CASTEP software package to

carry out a full vibrational analysis on BH_3NH_3 crystal [CASTEP, v. 4.2; Payne, 1992]. From the literature there were no normal mode analysis of solid BH_3NH_3 available. Also, quite unfortunately, there were no experimental sublimation energy available in the literature for comparison with our calculated value.

As each BH_3NH_3 molecule in the crystal lattice is connected by twelve inter-molecular dihydrogen bonds to its neighbours (*i.e.* an intermolecular bond order of 6, see Figure 4.2) we estimate the average value for the dihydrogen bond present in the crystal structure to be 76.06/6 = 12.7 $kJmol^{-1}$. As just mentioned we neglect the ZPE correction, however we expect that this correction should be rather small compared to the total interaction energy, and any error incurred would not change the final result (12.7 $kJmol^{-1}$) by a great deal. Note that our calculated value of the dihydrogen bond strength is about half of that [25.3 $kJmol^{-1}$] reported by the dimer model calculation of Crabtree and co-workers [Richardson, 1995].

4.4 Testing The Dimer Model

The weaker dihydrogen bond found in our study may simply be due to the failure of the supercell approach in the plane-wave DFT calculation, or an underestimation of binding energies using a plane-wave basis set. But we can test if this is the case by running a calculation on a dimer model using the supercell approach within plane-wave DFT. If a similar binding energy is obtained to that found for localised Gaussian basis sets, we will know that the error lies in the choice of model (Figure 4.1), not with the calculation method. It is very important to remember that the potential energy surface in the dimer model is very different to that of the solid state and the intermolecular interactions are not solely governed by the local geometry of the dihydrogen bond but rather by the whole packing arrangement of the crystal. Our calculations will thus test whether a match between local geometry *i.e.* D–H, H...A bond distance and D–H...A bond angle of a dimer model, which has been the method pursued for so long in traditional ab-initio molecular orbital calculation, is sufficient to mimic the real hydrogen bond strength in the crystal.

4.4.1 Computational technique

Series of localised basis set (LBS) calculation were carried out using the Gaussian'98 package [Frisch, 1998] in order to reproduce the literature conformations (due to Crabtree *et.al.*, Lie *et.al.* and Poplier; Figure 4.1) of BH_3NH_3 dimer. Most successful calculation (the lowest energy minima) were obtained for the C_{2h} structure reported by Li *et.al.* The dimer-monomer calculation were performed on this (C_{2h}) dimer in order to compute the dimer binding energy *in localised basis set method.*

The Geometry optimisation of the Li structure (C_{2h} dimer) were then repeated using the plane wave (PW) supercell technique (using CASTEP package) in periodic supercells of 12 x 12 x 12 $Å^3$. The dimer-monomer (See Section 4.2.2 for detailed method) calculation were performed in order to estimate the dimer binding energy *in plane wave method.*

Both localised basis set (isolated space) and plane wave (periodic space) calculations were carried out using DFT theory (PW91 functional). In both cases the basis set taken were extensively large (Table 4.2) in order to minimise the error due to an insufficient basis set. The finite basis set correction (FBSC) is close to zero for plane wave calculation, and the localised basis set calculations were corrected for basis set superposition error, which was also very low (0.4 $kJmol^{-1}$). Very large supercells (12 x 12 x 12 $Å^3$) ensures zero-interaction between neighbouring cells in PW calculation.

4.4.2 Results and discussion

Both the periodic dimer (plane wave) and the isolated dimer (localised basis set) calculations show the global minima (or minimum total energy) for the cyclic C_{2h} structure [Figure 4.1(b)]. Attempts to reproduce the geometry proposed by Crabtree *et. al.* were not successful even with very large Gaussian basis set. Li *et. al.* has also reported such difficulty with reproducing the Crabtree structure [Richardson, 1995]. The Table 4.2 in the next page summarises the important results obtained from the simulations.

Table: 4.2 Comparison between the *ab-initio* molecular orbital and plane-wave DFT calculations, obtained for BH_3NH_3 dimer [Figure 4.1(b)].

Values	Localised basis set approach		Plane-wave supercell approach	
	Dimer phase	Gas phase	Dimer phase	Gas phase
Set-up				
Theory Level/Method	B3-PW91	B3-PW91	DFT-PW91	DFT-PW91
Basis set / E_cut-off	6-311++G*	6-311++G*	600 eV	600 eV
K-point Grid	–	–	3x3x3	3x3x3
FBSC/BSSE	0.4 $kJmol^{-1}$	–	~0.0	~0.0
Lattice [Å, °, $Å^3$]				
a, b, c	–	–	12.0	12.0
α, β, γ	–	–	90.0	90.0
Z	–	–	2	1
Volume	–	–	1728.0	1728.0
Space/point group	C_{2h}	–	$P_{2/m}$	P1
Geometry [Å, °]				
*r*H...H	2.03	–	1.94	–
∠N–H...H	146.2	–	145.2	–
∠B–H...H	88.6	–	87.4	–
Energy				
Total energy	–166.426496 Hartree	–83.203111 Hartree	–895.502512 eV	–447.417497 eV
Dissociation Energy [$kJmol^{-1}$]	26.3		32.2	
H-bond strength [$kJmol^{-1}$]	13.2		16.2	

Geometry

The plane wave geometry optimisation reproduces the same minimum energy structure as that of the localised basis set calculations, which is a C_{2h} symmetry structure with four equivalent dihydrogen bonds (Table 4.2).The dihydrogen bond geometry obtained is very close to that of the first dihydrogen bond seen in the solid (See Table 4.1). Its length 1.94 Å, is a bit shorter (4%) than that obtained from the localised basis set calculation, and hence the bond strength in the periodic calculation is expected to be a bit stronger.

Energy

Table 4.2 shows that two different computational approaches on the same dimer conformation give very similar geometries and similar strengths of the dihydrogen bond. The dihydrogen bond strength in the periodic dimer conformation was a bit higher (3 $kJmol^{-1}$) than the isolated dimer conformation as a result of the slightly shorter interaction distance, which is a reasonable difference. Thus we can conclude that the plane-wave (periodic space) DFT calculations are capable of determining reasonable binding energies for the $(BH_3NH_3)_2$ system at the same level as the localised basis set calculation, and so should be valid for the BH_3NH_3 crystal. So we refer to the value, 12.7 $kJmol^{-1}$, calculated using plane wave DFT approach listed in Table 4.1 as the correct dihydrogen bond strength in BH_3NH_3 crystal and previously reported surprisingly high value, 25.3 $kJmol^{-1}$, of dihydrogen bond is an artefact of the dimer model.

4.5 Conclusion

This work demonstrates that a more realistic model of solid BH_3NH_3 reassigns the dihydrogen bond as a weak-moderate H-bond of strength 12.7 $kJmol^{-1}$, in line with the general prediction by the classical H-bond theory [Steiner, 2002; Jeffery, 1997]. A clear message from this detailed study is that the hydrogen bond strength in the solid is not a sole function of its local geometry (*i.e.* D–H, H...A bond distance and D–H...A bond angle). The potential map of the crystal where the true hydrogen bond lies is in general very different from that of simple dimer model. The use of dimer models to mimic the solid state can therefore at best be regarded on a trial and error basis.

4.6 Bibliography

1. Crabtree, R. H., P. E. M. Siegbahn, O. Eisenstein, A. L. Rheingold and T. F. Koetzle, *Acc. Chem. Res.*, **29**, 348(1996).
2. Cramer, C. J., and W. L. Gladfelter, *Inorg. Chem.*, **36,** 5358(1997).
3. Curtiss, L. A., K. Raghavachari, P. C. Redfern and J. A. Pople, *J. Chem. Phys.*, **106**, 1063(1997).
4. Ernzerhof , M., and G. E. Scuseria, *J. Chem. Phys*, **110** (11), 5029(1999).
5. Francis, G. P., and M. C. Payne, *J. Phys.: Cond. Matt.*, **2,** 4395 (1990).
6. Frisch, M. J., Gaussian 98, Revision A.7, (Gaussian, Inc., Pittsburgh, PA, 1998).
7. Jeffery, G. A., *An introduction to hydrogen bonding*, (Oxford University Press, Oxford, 1997).
8. Klooster, W. T., T. F. Koetzle, P. E. M. Siegbahn, T. B. Richardson and R. H. Crabtree, *J. Am Chem. Soc*, **121**, 6337(1999).
9. Kurth, S., J. P. Perdew and P. Blaha, *Int. J. Quant. Chem.*, **75**(4-5), 889(1999).
10. Li, J., F Zhao and F. Jing, *J. Chem. Phys.*, **116,** 25(2002).
11. Monkhorst, H. J., and J. D. Pack. *Phys. Rev. B*, **13**, 5188 (1976).
12. Payne, M. C., M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, *Rev. Mod. Phys.*, **64**, 1045(1992); CASTEP, Academic version 4.2 , (licensed under the UKCP-MSI Agreement, 1999).
13. Perdew, J. P., Y. Wang, *Phys. Rev. B,* **45**, 13244(1991).
14. Popelier, P. L. A., *J. Phys. Chem. A.*, **102**, 1873(1998).
15. Richardson, T. B., S. de Gala and R. H. Crabtree, *J. Am Chem. Soc*, **117**, 12875(1995).
16. Segall, M.D., P. J. D. Linden, M.J. Probert, C.J. Pickard, P. J. Hasnip, S.J. Clark and M. C. Payne, *J. Phys.: Cond. Matt.*, **14**, 2717(2002).
17. Steiner, T., *Angew. Chem. Int. Ed.*, **41**, 48(2002).
18. Thorne, L. R., R. D. Suenram and F. J. Lovas, *J. Chem.Phys.*, **78,** 167(1983).
19. Vanderbilt, D., *Phys. Rev. B,* **41**, 7892(1990).

Chapter 5

An Introduction to First-Principles Molecular Dynamics

5.1 Introduction

Molecular dynamics simulation is a computational technique to move ions, atoms or molecules under the action of realistic forces. This allows us to observe the evolution of a desired system (see Figure 5.1) at a microscopic level (*i.e.* the detailed conformation or phase points, Γ) in its phase space (6N dimensional space comprising 3N position co-ordinates and 3N momentum co-ordinates, where N is total number of atoms in the system). In real life, a system in equilibrium visits all possible phase points in a particular phase, with the probability $\rho(\Gamma)$ given by the Boltzmann distribution. The probability distribution is defined by the macroscopic parameters of the system (*i.e.* the thermodynamic state, *e.g.* NPT, NVT *etc.*). Such a system is termed 'ergodic' and the collection of the phase points (Γ) for which the probability, $\rho(\Gamma)$, is non-zero is regarded as an 'ensemble' of that thermodynamic state. Molecular dynamics (MD) can allow us to explore this phase space of the real-life system by following its *equations of motion*. Care should be taken in the choice of the MD algorithm to ensure the ergodicity [Martyna, 1992]. Averaging the desired properties over the phase trajectory (the collection of phase points generated by the equations of motion of the system) for a sufficiently long time therefore gives us valuable information about the dynamical properties of the system of interest.

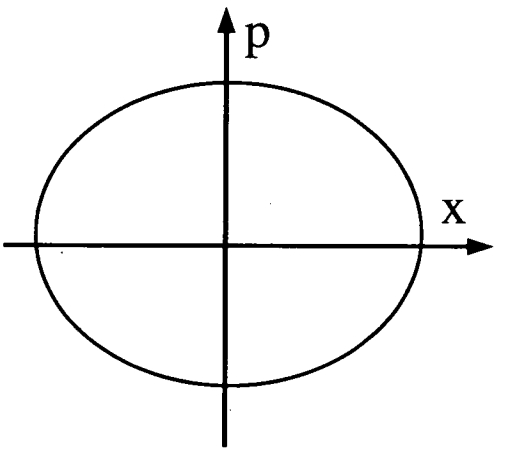


Figure 5.1: Phase space for one dimensional linear harmonic oscillator, showing a trajectory followed clockwise in time. Obviously, the phase space trajectory for a truly macroscopic system with 6N (*e.g.* 10^{23}) degrees of freedom will be very complex.

Molecular dynamics is a purely classical simulation of the atomic motion. The system's dynamics is solved using the classical equation of motion, with forces derived from the total energy of the system from all types of interaction, E.

eqn 5.1: $$\boldsymbol{F} = -\nabla E,$$

The total energy of the system could be either derived from a model potential V [Allen, 1987], or from a DFT-based quantum mechanical calculation using the Hellmann-Feynman theorem [Feynman,1939]. From Taylor's expansion we can write the position of an ion at a time $(t+h)$ linked to its position at previous time t by

eqn 5.2: $$\boldsymbol{r}(t+h) = \boldsymbol{r}(t) + h\dot{\boldsymbol{r}}(t) + \frac{h^2}{2!}\ddot{\boldsymbol{r}}(t) + \frac{h^3}{3!}\dddot{\boldsymbol{r}}(t) + \ldots\ldots + \frac{h^n}{n!}\boldsymbol{r}^n(t) + O(h^{n+1}).$$

This is an exact relation. Therefore, if we know the position of the particle at time t, and all the higher derivatives of the position $\boldsymbol{r}$, we can calculate the particle position at a later time $t+h$. In other words, if we know the system $\{\boldsymbol{r}_i(t)\}$ at time t, we can derive the system configuration $\{\boldsymbol{r}_i(t+h)\}$ at a later time $t+h$ if we know all the higher derivatives of $\{\boldsymbol{r}_i\}$. In this way systems become completely deterministic and the "true trajectory" of the system can be predicted for any desired period of time from a precisely known initial condition.

However this is not possible in a real life simulation. We can deal with the first few derivatives of position but it is impossible to know all the higher derivatives. The first derivative, $\dot{\boldsymbol{r}}$ is the velocity and can be easily derived. The second derivative, $(\ddot{\boldsymbol{r}})$ is acceleration of the ion, which can be obtained from the force exerted on that ion from Newton's equation of motion:

eqn 5.3: $$\boldsymbol{F} = m\ddot{\boldsymbol{r}},$$

where forces are derived from either an empirical potential or from quantum mechanical calculations. We have very little or no idea about other higher derivatives of $\boldsymbol{r}$ (change of force with time), *i.e.* $\dddot{\boldsymbol{r}} \ldots . \boldsymbol{r}^n$.

If, however, h is very small such that h^3 and all higher powers of h are close to zero, then we can cut down the higher terms in the Taylor expansion to give the truncated equation

eqn 5.4: $$\boldsymbol{r}(t+h) \simeq \boldsymbol{r}(t) + h\dot{\boldsymbol{r}}(t) + \frac{h^2}{2}\ddot{\boldsymbol{r}}(t).$$

Here the parameter h, known as the time step of molecular dynamics simulation, plays a central role. The parameter h has to be adjusted so that it is small enough to maintain the above prediction of eqn 5.4. This is ensured by testing that the time step is small enough to sufficiently sample the fastest mode present in the system.

5.1.1 Integrating the equations of motion

In short, a molecular dynamics calculation is all about integrating Newton's equations of motion (eqn 5.3) with the forces derived from total energy (eqn 5.1). A good integration algorithm is crucial for a good molecular dynamics simulation. All the available integration schemes are based on Taylor expansions (as in eqn 5.4) of different dynamical quantities connected together in a slightly different approach. In this work we have used the velocity-Verlet algorithm [Swope, 1982], which is not only one of the simplest, but also usually the best available MD algorithm. The Taylor expansion is used to estimate position and velocities, accurate to order h^3

eqn 5.5: $$\boldsymbol{r}(t+h) = \boldsymbol{r}(t) + h\dot{\boldsymbol{r}}(t) + \frac{h^2}{2}\ddot{\boldsymbol{r}}(t) + O(h^3),$$

eqn 5.6: $$\dot{\boldsymbol{r}}(t+h) = \dot{\boldsymbol{r}}(t) + h\ddot{\boldsymbol{r}}(t) + \frac{h^2}{2}\dddot{\boldsymbol{r}}(t) + O(h^3).$$

Using forward difference approximation, $\dddot{\boldsymbol{r}}$ can be written as

eqn 5.7: $$\dddot{\boldsymbol{r}}(t) = \frac{\ddot{\boldsymbol{r}}(t+h) - \ddot{\boldsymbol{r}}(t)}{h} + O(h^2).$$

This allows us to rewrite eqn 5.6 as

eqn 5.6: $$\dot{\boldsymbol{r}}(t+h) = \dot{\boldsymbol{r}}(t) + \frac{h}{2}\ddot{\boldsymbol{r}}(t) + \frac{h}{2}\ddot{\boldsymbol{r}}(t+h) + O(h^3).$$

If we define a new variable, $\boldsymbol{q}$, as

eqn 5.8: $$\boldsymbol{q} = \dot{\boldsymbol{r}}(t) + \frac{h}{2}\ddot{\boldsymbol{r}}(t),$$

eqn 5.5 and 5.6 can be finally written as

eqn 5.9: $$\boldsymbol{r}(t+h) = \boldsymbol{r}(t) + h.\boldsymbol{q} + O(h^3),$$

eqn 5.10: $$\dot{\boldsymbol{r}}(t+h) = \boldsymbol{q} + \frac{h}{2}\ddot{\boldsymbol{r}}(t+h) + O(h^3).$$

The newly defined variable, $\boldsymbol{q}$, is often referred to as the mid-velocity. The equations eqn 5.8 - 5.10 forms the basic structure of the velocity-Verlet algorithm.

5.1.2 The two-step algorithm

One of the main features of the velocity-Verlet algorithm is that it is a two step method, with a force calculation in between. The obvious advantage is in terms of memory storage and clarity of the coding. This algorithm requires memory storage for only three variables: position, velocity and acceleration (9N data storage). For the first step, the mid velocities $\{\mathbf{q}_i\}$ are calculated using eqn 5.8 and then the new positions $\{\boldsymbol{r}_i(t+h)\}$ are calculated using eqn 5.9. This step is often referred to as velocity-Verlet position update.

At the intermediate stage, the forces and corresponding accelerations at time $t+h$ are then computed (eqn 5.1 and eqn 5.3) from the knowledge of total energy (*i.e.* total internal energy) due to the new conformation $\{\boldsymbol{r}_i(t+h)\}$.

The second step is the velocity-Verlet velocity update. The velocities are updated at time $t+h$ from the knowledge of the new positions and forces (at $t+h$) using eqn 5.10. At this point, the kinetic energy of the system at time $t+h$ is available.

The local error with this algorithm is of order h^3 and the global error of this simulation is of order h^2 [Swope, 1982]. Since the velocities appear explicitly in this method, it is more suitable to simulate velocity-dependent phenomena. The numerical stability, convenience and simplicity makes it perhaps the most attractive method of choice [Allen, 1987].

5.1.3 Conservation criteria for numerical algorithm

While the exact solution of a Hamiltonian system follows certain conservation laws, discrete numerical scheme, in general, tends to violate it. Particular importance is the time reversible symmetry (corresponds to energy conservation) and the conservation of phase space density. For a successful MD simulation it is important to prove that these two criteria are met.

The velocity-Verlet algorithm meets both of these criteria. Recently Tuckerman *et. al.* [Tuckerman, 1992] have demonstrated how to systematically derive time-reversible, area preserving MD algorithms from Liouville formulation of dynamical mechanics. Let us consider a phase point Γ [$\boldsymbol{p}^N$(t), $\boldsymbol{r}^N$(t)] represents a classical system of N particle, where the term Γ depends on time implicitly. The time derivative of Γ is given by

eqn 5.11 $$\dot{\Gamma} = \dot{\boldsymbol{r}} \frac{\partial \Gamma}{\partial \boldsymbol{r}} + \dot{\boldsymbol{p}} \frac{\partial \Gamma}{\partial \boldsymbol{p}} \equiv iL\,\Gamma, \quad \text{where}$$

eqn 5.12 $$iL = \dot{\boldsymbol{r}} \frac{\partial}{\partial \boldsymbol{r}} + \dot{\boldsymbol{p}} \frac{\partial}{\partial \boldsymbol{p}} \equiv iL_r + iL_p.$$

iL is known as the Liouville operator, which acts as an operator to move the Hamiltonian system along its MD trajectory. By integrating the eqn 5.12 we get the operator of time evolution $exp(i\,Lt)$ of the MD system:

eqn 5.13 $$\Gamma[\boldsymbol{p}^N(t), \boldsymbol{r}^N(t)] = \exp(iLt)\Gamma[\boldsymbol{p}^N(0), \boldsymbol{r}^N(0)]$$

It is clear from eqn 5.13 that the evolution operator completely retains the time reversible symmetry, as $exp\{iL(-t)\}.exp(iLt)$ $\Gamma(0) = \Gamma(0)$, *i.e.* the exact Hamiltonian dynamics is time reversible. Now to prove that the velocity-Verlet scheme also retains this

property, we need to split the evolution operator iL into the position update operator iL_r and the velocity update operator iL_p (see eqn 5.12). Although iL_r and iL_p are noncommuting, *i.e.*

eqn 5.14 $$\exp(iL_r t + iL_p t) \neq \exp(iL_r t)\exp(iL_p t),$$

we can still use the Trotter identity [Frenkel, 2002], which gives

eqn 5.15 $$\exp(iL_r t + iL_p t) = \exp(iL_p t/2)\exp(iL_r t)\exp(iL_p t/2) + O(t^3).$$

The error of the order t^3 comes as a penalty to override the noncommuting property of iL_r and iL_p in eqn 5.14 [Tuckerman, 1992]. Now the new form of time evolution operator, which moves the system forward by a time step t, is

eqn 5.16 $$\Delta_t = \exp(iL_p t/2)\exp(iL_r t)\exp(iL_p t/2),$$

which is exactly the two step velocity-Verlet integration scheme. From eqn 5.16 it is clear that the velocity-Verlet is time reversible, as $\Delta_t \Delta_{-t} \Gamma(0) = \Gamma(0)$.

Another important property of a Hamiltonian system is that the natural time evolution of the system corresponds to a symplectic co-ordinate transformation [Frenkel, 2002], which means: *a*) each MD step produced by the evolution operator (i.e. velocity-Verlet method) corresponds to a canonical transformation of co-ordinates from $\{\boldsymbol{p}^N(0), \boldsymbol{r}^N(0)\}$ to $\{\boldsymbol{p}^N(t), \boldsymbol{r}^N(t)\}$. *b*) The size of the volume element associated with Γ in the phase space does not change by the transformation.

While most of the integration algorithm maintain the canonical transformation, *i.e.* the newly generated co-ordinates also comply with the Hamiltonian equations of motions,

many algorithm does not comply with the second condition of the incompressibility of the phase space. The symplectic condition is maintained in an MD step if the Jacobian of the co-ordinate transformation (i.e. co-ordinate update forward in time) is one [Frenkel, 2002]. For the velocity-Verlet algorithm, Jacobian of the transformation from Γ [$\mathbf{p}^N(0)$, $\mathbf{r}^N(0)$] to Γ [$\mathbf{p}^N(t)$, $\mathbf{r}^N(t)$] is simply the product of the Jacobians of the three elementary transformations (eqn 5.16). As each of these Jacobians is equal to one, the overall Jacobian is also one [Tuckerman, 1992].

5.2 The Correct MD

The above discussion explained the basic mathematical framework of a molecular dynamics integration scheme. However, apart from that, there are some conceptual issues that have to be met by an MD simulation in order to produce correct statistical predictions of the system of interest. The criterion is based on "how closely" the generated computational trajectory matches the "true trajectory" of the system. In what follows, this key issue, dominating the accuracy of a MD, will be addressed.

To have an algorithm that accurately predicts the trajectory of the system in its phase space, for both short and long times, would appear to be the most important criterion. In fact, no such algorithm exists. This is because of the approximate character of the integration algorithm (eqn 5.4) used. Although the integration error caused by the discrete algorithm is very small (but not zero) at the initial stage of the computation, it builds up as the computation proceeds, finally causing the system to deviate strongly from the true trajectory. We should therefore expect that any integration error, no matter how small, will always cause our simulated trajectory to diverge exponentially from the true trajectory compatible with the same initial condition. The deviation follows the so-called Lyapunov instability [Frenkel, 2002] ,

eqn 5.17: $$|\Delta \boldsymbol{r}_i(t)| \sim \varepsilon \exp(\lambda t) .$$

Here ε is a small difference in the initial condition $[r(0), p(0)]$ of a particle (say i) in two separate simulation runs, and $\Delta r(t)$ is their separation at time t. Eqn 5.17 illustrates the extreme sensitivity of the trajectories to small perturbations in the initial conditions.

The Lyapunov instability would be a major blow to an MD simulation, if the aim were to predict *precisely* what will happen to a system that has been prepared in a *precisely* known initial condition, but fortunately we are only interested in statistical predictions. We wish to predict the *average* behaviour of the system that was prepared in an initial state, or in other words, our aim is to sample the particular phase region of the phase space (*e.g.* the liquid phase ensemble in the phase space of water) of the system in an efficient way so that the average quantities extracted from the simulation represents the true system. So long as the error in the trajectory due to the Lyapunov instability is very small, the system remains bound in the correct ensemble of the phase space, and the statistical properties extracted from such trajectories will be correct. In this respect, an MD simulation differs fundamentally from the numerical schemes used for predicting the trajectory of a satellite through space; in the latter case, we really wish to predict the true trajectory. Hence we can conclude that for a good statistical prediction, although an MD simulation does not need to follow the true trajectory *precisely*, it is *necessary* that the trajectories obtained numerically are *close* to the true ones.

For three key reasons, an MD simulation may deviate strongly from, rather than follow closely, the true trajectory. Firstly, since the molecular dynamics is a classical approach, it can only accurately describe a real system, which obeys the laws of classical mechanics. This is usually acceptable, apart from the cases where the quantum effect [Cramer, 2003] becomes significant for light atoms and molecules in a system. This usually happens at low temperature when the de Broglie wavelength for the light particles becomes comparable to the atomic radius (*i.e.* of the order of 1 Å). Secondly, the time step of the molecular dynamics simulation has to be sufficiently small to keep the integration error within acceptable range. An infinitely small time step seems to be the ideal, but would lead to very extensive computation. Thirdly, the *derived forces* have to be very *close* to the

correct forces exerted by the *true system.* This is an essential requirement in order that our simulated system will follow the true trajectory closely, and is a very complex and central issue in any MD calculation. Forces derived from the classical model potentials suffers from the limitations due to parameterisation, whereas forces derived from first-principles also suffers from the limitations in its computational technique and due to lack of exact knowledge of the Exchange-Correlation functional. In the following sections we will discuss the limitations imposed by the modelling of potentials, which opens up the discussion for the state-of-art first-principles MD or ab-inito MD.

5.3 Classical MD

For the traditional MD simulation, more familiarly known as classical MD, forces are derived from a model potential that is written in terms of the atomic co-ordinates (their position, orientation and other properties) and a number of fitted parameters [Wallqvist, 1999]. Such a potential is based on intellectual guess and experimental evidence, and the parameters are fitted so that they can predict most of the known results. In this way once a credible model has emerged, it is used to predict the other chemical and statistical behaviours of the system, such as protein folding, cluster formations in liquid mixture, ion transport in biological system *etc.*, which may not be amenable to experimentation.

The success for such an empirical potential depends on how good it is at describing the known properties of the system. Although empirical potentials are of great value, they are limited by the fact that they can only describe a system to the accuracy of the parameterisation of the potential. In general they are good at describing the interactions of the system to which their parameters were fitted but their transferability to other environments can be quite poor. In fact in many cases it becomes very complicated to design a single model that can predict all the known behaviour of a chosen system and consequently the researcher has to resort to different models to predict different properties even for a single system. A very well known example is the molecular dynamics simulation

of liquid water with empirical potentials. Despite it being one of the best and earliest known systems, and seemingly endless efforts to build an unique potential, there is as yet no single potential that can describe all the properties of liquid water well [Gordon, 2002]. Even if a parametrised potential succeeds in describing most of the observed properties with reasonable accuracy, it cannot guarantee the correctness of its prediction for yet unknown results.

5.4 First-Principles MD

To overcome the limitations of the classical MD and to explore desired systems with greater accuracy, an approach at a more fundamental level is required. One must turn to the big task of solving the system quantum mechanically (that is solving the Schrodinger equation for the whole system) at an electronic level to extract the forces. The forces obtained this way are free from any sort of parameterisation and no *a priori* information about the system interaction is needed.

With the discovery of Density Functional Theory (DFT) by Kohn, Hohenberg and Sham [Hohenberg, 1964; Kohn, 1965], such an electronic level calculation became computationally attainable. In the mid-eighties, Car and Parrinello [Car, 1985] showed how to combine the accuracy of ab-initio density functional methods with force calculations and molecular dynamics (MD) techniques. This lead to a new computational technique, popularly known as first-principles or ab-initio Molecular Dynamics, for calculating dynamical properties of comparatively small systems with an electronic level of accuracy. In a nutshell, in ab initio MD, the nuclei are still treated as classical particles whereas the forces are calculated from first-principles using DFT.

Whereas an empirical-potential-based classical MD gives very little or no information about the electron density distribution of the system, it is the fundamental

quantity in an ab-initio MD calculation. The inter-atomic potential is determined directly by the electronic structure of the system rather than modelling the interactions in terms of a classical or semi-classical potential with fitted parameters. For these reasons, along with the continual increase of computational power, ab-initio MD became the method of choice to study a wide range of phenomena, especially where accuracy is highly desired [Alavi, 1994; Curioni, 1997; Car, 1985].

However, this accuracy is gained at the expense of high computation time. Each step requires an extensive force calculation, compared to the straight-forward force evaluation from a simple function in the case of classical MD. Before evaluating the forces on the ions, a massive minimisation calculation is required in the extremely large phase space of the basis set of the electronic wave functions. The size of this minimisation calculation overwhelms the time taken in an MD integration step. This imposes a severe restriction on the applications of ab-initio DFT to larger systems. For example, where an ab-initio MD calculation approaches the limit for size with only 10^3 atoms, an empirical calculation of a similar scale would be using ~10^7 or more particles. Currently the state of the art ab-initio MD simulation can simulate systems of a few hundreds of atoms for a pico-second time scale.

In Section 2.7.1 the methods for optimising the electronic configuration to reach the ground state configuration with respect to a fixed ionic configuration were discussed. Later in Section 2.7.2, the method for relaxing ion to their local minima was discussed. Two distinctive approaches came out of those discussions: first, the Car-Parrinello based MD method of *indirect* relaxation (see Section 2.7.1) of electrons to their ground state, and ions and cell variables to the local minima; second, a conjugate gradient scheme for *directly* minimising (see Section 2.7.1) the Kohn-Sham energy functional to the ground state and a different minimisation scheme to relax ions and cell variable to the local minima. Global minima can be obtained by repeating one of these approaches with a simulated annealing procedure.

Despite the advantages and disadvantages of both the Car-Parrinello and conjugate-gradient approach, they stand out to be the widely accepted optimisation methods. Once the system is optimised to its ground state first-principles forces on ions and cell variables are derived using the Hellmann-Feynman theorem [Feynman,1939]. There are therefore two ab-initio MD procedures based on the two different minimisation approaches. In this work we have used the conjugate-gradient based molecular dynamics but for the sake of completeness both methods are discussed in the following Section.

5.4.1 Car-Parrinello MD

The Car-Parrinello dynamics [Car, 1985] scheme can be thought of as a natural extension of the Car-Parrinello MD method of relaxing ions and electrons to the ground state[1]. Instead of quenching the heat out of the system, the system is allowed to explore its 6N dimensional phase space at a chosen temperature (see Section 2.7.2).

The Hellmann-Feynman theorem is used to calculate forces on ions and cell variables from a single total-energy functional at ground state. The theorem simplifies the calculation of the physical forces on the ions and the integrated stresses on the unit cell. However, the calculated forces are the true physical forces only when the electrons are in their ground-state configuration [Feynman, 1939]. Therefore, electrons must be relaxed to their ground-state first at each ionic time step to generate the correct dynamical trajectories. The Car-Parrinello method can be used to ensure the instantaneous ground-state electronic configuration is obtained, but it would be an extremely expensive computation to obtain dynamical simulations of ionic systems. However there is an alternative approach to performing dynamical simulations using this method. Rather than insisting that the electronic configuration be in the exact ground-state at each ionic time step, one may be able to perform dynamical simulations if the electronic configuration is only *close* to the exact ground-state. Although this implies that there are errors in the Hellmann-Feynman

1 Historically CP dynamics came first and was subsequently taken as a tool for the minimisation of the Kohn-Sham energy functional later on [Car, 1985; Payne, 1992].

forces at each time step, dynamical simulations will be successful provided that the errors in the forces remain small *and* that the effect of these error remain bounded in time. The CP-MD method can fulfil (by 'error cancellation') both these criteria [Remler, 1990; Pastore, 1991].

Limitations of CP-MD simulations

Despite the success of Car-Parrinello dynamics, it has major drawbacks. Firstly, as the approach is based on an indirect search (see Section 2.7) for the minimum of the Kohn-Sham Hamiltonian, the procedure may lead to serious instability in the search for the ground state if a long time step is used. Unfortunately the critical time step at which this instability starts to grow decreases as the system size increases [Payne, 1992]. Secondly, to have accurate Car-Parrinello dynamics it is essential to have the ionic and electronic time step decoupled. Although there are many system for which this is the case, this decoupling is generally difficult to obtain in metallic systems [Payne, 1992].

5.4.2 Conjugate gradient dynamics

An alternative approach to CP-MD is the conjugate-gradient based electronic relaxation [Arias, 1992], where the electronic structure is relaxed to its ground state using the conjugate-gradient scheme. At the ground state configuration of the electrons, the Hellmann-Feynman forces are equal to the true forces, which are then plugged into the classical integration algorithm. However, as in the Car-Parrinello dynamics scheme, it is also extremely computationally expensive to ensure that the correct electronic ground state is obtained with the conjugate-gradient scheme at each time-step. Moreover, a conjugate-gradient scheme does not benefit from the error cancellation described above when the Hellmann-Feynman requirements [Feynman,1939] are not strictly met.

A solution to this problem is achieved by moving the initial electronic configuration closer to the instantaneous ground state configuration rather than starting from a random one. Hellmann-Feynman requirements still have to be met but much less computational effort is required to converge to the exact electronic ground state, and hence a faster dynamical simulation is possible. The method has been developed by Arias *et al.* [Arias, 1992], where the new electronic configuration at the start of a new ionic step is predicted by extrapolating forward from the previous *ground-state* configuration of electrons. This method of extrapolation has been found to bring the initial electronic configuration about two orders of magnitude closer to the minimum of the Kohn-Sham total energy functional, thus reducing the total computational effort by a factor of two [Arias, 1992]. The conjugate-gradient method combined with the wave-function extrapolation technique, often known as Born-Oppenheimer molecular dynamics (BO-MD) simulation, results in a dynamical simulation comparable to the speed of a Car-Parrinello dynamical simulation (CP-MD).

5.4.3 Comparison between CP-MD and BO-MD

The Car-Parrinello and Born-Oppenheimer schemes for performing dynamical simulations are very different, and it is important to understand these differences in order to apply either technique successfully. The most important point is the difference between the time-steps used in the two methods. In this respect Born-Oppenheimer dynamics is closer to conventional dynamical simulations, in which the time step is chosen to ensure an accurate integration of the ionic equations of motion. In simulations employing empirical potentials and those using the Born-Oppenheimer scheme, the forces on the ions are, to a high precision, true derivatives of the total potential energy of the ions. In the case of empirical potentials, the only differences between the computed forces and the derivative of the total ionic energy are rounding errors due to finite machine accuracy, but in the case of Born-Oppenheimer (BO) simulations, the difference also includes a contribution due to the failure of the Hellmann-Feynman theorem [Feynman,1939] because the electronic system is not *exactly* converged to its ground state.

In the Car-Parrinello (CP) simulation, at each time step there are significantly larger errors in the Hellmann-Feynman forces, because the electronic configuration is not maintained close to its exact ground state configuration. To ensure the stability of the indirect search for the self-consistent Kohn-Sham Hamiltonian, the time step used in a CP simulation has to be much shorter than the one used in the BO approach (see Section 2.7.2). Additionally the longest time period in the electronic system must be much less than the shortest time period in the ionic system in a CP simulation, to ensure that the error in the Hellmann-Feynman forces [Feynman,1939] average to zero along the ionic trajectories [Payne, 1992].

At first sight the Car-Parrinello method and the wave-function extrapolation based conjugate-gradient relaxation (Born-Oppenheimer scheme) are rather similar, in that each essentially performs an integration of wave-functions forward in time. However, the spirit of each technique and the behaviour of the electronic wave-function co-efficients in the two cases are different. In the case of conjugate-gradient dynamics, the wave function is propagated as close as possible to the instantaneous ground state, in order to reduce the effort to fully relax them to the ground state [Arias, 1992]. In the Car-Parrinello method, a delicate adiabatic separation between the ionic and electronic degrees of freedom is preserved using two very different time scales (for details, see Section 2.7.1). The electronic co-efficients oscillate artificially about their ground state values, which leads to a cancellation of the errors in the ionic forces [Remler, 1990; Pastore, 1991].

5.5 Conclusion

Both the Car-Parrinello and the Born-Oppenheimer dynamics simulation are now popularly used by both the academic and the commercial communities, and exciting new scientific discoveries have been made through the application of these methods [Andreoni, 2000; Payne, 1989; Tarnow, 1989]. New functionalities are also being regularly added to improve the level of efficiency, as well as to tackle more theoretical issues in an attempt to cover wider area of research. In this PhD project, we have worked with the conjugate-gradient based Born-Oppenheimer dynamics in an attempt to achieve constraint dynamics from first-principles, which will be discussed in the following Chapters.

5.6 Bibliography

1. Alavi, A., J. Kohanoff and M. Parrinello, *Phys. Rev. Lett.*, **73** (19), 2599(1994).
2. Allen, M. P., and D.J. Tildesley, *Computer simulation of liquids*, (Oxford University Press, 1987).
3. Andreoni, W., and A. Curioni, *Parallel Computing*, **26**, 819(2000).
4. Arias, T. A., M. C. Payne, J. D. Joannopoulos, *Phys. Rev. B*, **45**(4), 1538(1992).
5. Car, R., and M. Parrinello, *Phys. Rev. Lett.*, **55**, 2471(1985).
6. Cramer, C.J., *Essential of Computational Chemistry*, (John Wiley & Sons Ltd, 2003).
7. Curioni, A., M. Sprik, W. Andreoni and M. Parrinello, *J. Am. Chem. Soc.*, **119**(31), 7218(1997).
8. Feynman, R.P., *Phys. Rev. B.*, **56**, 340(1939).
9. Frenkel, D., and B. Smit, *Understanding Molecular Simulation*, Second Edition (Academic Press, 2002).
10. Gordon, T. H., and G. Hura, *Chem. Rev.*, **102**, 2651(2002).
11. Hohenberg, P., and W. Kohn, *Phys. Rev. B*, **136**, 864(1964).
12. Kohn, W., and L. J. Sham, *Phys. Rev A*, **140**, 1133(1965).
13. Martyna, G.J., M.L. Klein and M. Tuckerman, *J. Chem. Phys*, **97**(4),2635(1992).
14. Pastore, G., E. Smargiassi and F. Buda, *Phys. Rev. A*, **44**, 6334(1991).
15. Payne, M.C., E. Tarnow, P.D Bristowe and J.D. Joannopoulos, *Mol. Sim.*, **4**, 79(1989).
16. Payne, M. C., M. P. Teter, D.C Allen, T.A. Arias and J.D. Joannopoulos, *Rev. Mod. Phys.*, **64**(4), 1045(1992).
17. Remler, D. K., and P. A. Madden, *Mol. Phys.*, **70,** 921(1990).
18. Swope, W.C., H.C. Andersen, P.H. Berens and K.R. Wilson, *J. Chem. Phys.*, **76**, 637 (1982).
19. Tarnow, E., P. D. Bristowe, J. D. Joannopoulos and M. C. Payne, *Proc: Atomic Scale Calculations in Materials Science*, **66**, 333(1989).
20. Tuckerman, M.E., B. J. Berne and G. J. Martyna, *J. Chem. Phys.*,**97**, 1992(1990).
21. Wallqvist, A., and R. D. Mountain, *Reviews in Computational Chemistry*, **13**, (K. B. Lipkowitz and D. B. Boyd Edition, Wiley-VCH: New York, 1999).

Chapter 6

Implementation of Constrained Dynamics in First-Principles MD

6.1 Introduction

From the previous Chapter it is clear that the high precision, parameterisation-free forces in a first-principles MD are only obtained at the expense of a mammoth computation: minimizing the electrons to their instantaneous ground state at each step of ionic dynamics. This calculation approach is thus slower than the empirical potential dynamical simulation by several orders of magnitude, consequently only small systems can be simulated in a reasonable time scale. Although ab-initio MD has very widespread use in the scientific community, for much wider applications it is necessary to explore any technical detail that may offer speed up in the calculation.

Eliminating the high frequency modes of vibrations in the dynamics of molecular systems seems to be one such technical detail. High frequency modes need very short integration time steps and very often we are simply not interested in them. It is interesting to note that the problem of integrating fast vibrational modes has also been encountered in classical MD simulations, where the common practice is to speed up the simulation by introducing constraints in the system. In polyatomic molecules, the fast internal vibrations are usually decoupled from the rotational and translational motions of the molecule and can,

therefore, be 'frozen out' by introducing a certain number of rigid bonds and angles in the skeleton of the molecule [Ryckeart, 1977]. Moreover, since quantum effects can be significant in these high frequency modes it is not clear that the standard classical MD actually offers a more realistic treatment than the simplified rigid body MD approach.

In the classical treatment of vibrations, the energy involved in different modes of vibration is a continuous function, $E_{classical} = 1/2(m\omega^2 x^2)$. The presence of any anharmonicity allows the "classical oscillators" to couple with each other, and so even an infinitesimal amount of energy can be transferred between two different modes by anharmonic effects. However, in the real case all crystal vibrations are quantized over phonon modes of vibration (quantum oscillators). A coupling between two different frequency modes occurs only when the thermal energy ($k_BT/2$) is comparable to the transition energy $\Delta E = \hbar\Delta\omega$, which is quite big compared to k_BT at room temperature for the high frequency modes. Therefore, it can be reasonably assumed that high frequency modes are fairly decoupled from the low frequency molecular modes of vibration, and freezing out those fast modes by introducing constraints may actually also improve the dynamical description of the system. Hence the constraint dynamics also brings the classical system closer to the quantum description of the system, as well as improving the time step of the dynamics.

Table 6.1: Time-step required to sample different vibrational modes.

Mode of vibration	Typical frequency range (cm^{-1})	[a]Time-step required (10^{-15}sec)
X—H bond stretching	*ca.* 3600 - 3000	*ca.* 0.6 - 0.75
Bond stretching	*ca.* 2000 - 1000	*ca.* 1.1 - 2.20
Bond bending	*ca.* 1500 - 700	*ca.* 1.5 - 3.00
Torsion and other lattice modes	below 700	*ca.* 3.15

[a]The number of complete oscillations per time step is set to 15.

Table 6.1 shows how much computer time could be gained depending on the particular choice of constraints applied to a typical molecular system. Constraining any particular mode (and all higher modes) allows us to use a much bigger time step, as we only need to integrate any modes slower than the one constrained.

6.1.1 Constraints and MD methods

Since the electronic degrees of freedom in Car-Parrinello [Car, 1985] dynamics are even faster moving than high-frequency molecular vibrations, it was believed that using any constraint in the skeleton of molecules would not allow any increase in the time step. To prevent any coupling with the ionic degrees of freedom, which could cause inaccuracy in a CP simulation, a relatively small fictitious electronic mass needs to be used. This leads to a time step of the order of 0.01 femto-second in a CP simulation, ten times smaller than required for molecular vibrations in the case where electrons are not explicitly treated in the dynamics of the system. However very recent work on liquid water by Allesch *et. al.* [Allesch, 2004] demonstrated the fact that a higher electron mass could be used if the high frequency ionic modes could be constrained. Using the rigid water approximation, they have gained about a three-fold increase in the time step compared to the flexible water simulation. The observed properties with the rigid water model simulation were found to be in better agreement with the experimental measurements [Sorenson, 2000; Soper, 2000] than the corresponding CP simulations using a flexible water model.

In contrast to CP-MD, the Born-Oppenheimer MD simulation allows a straight-forward implementation of constrained dynamics as it does not include the fast electronic degrees of freedom in the integration algorithm. It is therefore much closer to conventional classical MD simulations, in which the time step is chosen to ensure accurate integration of the ionic degrees of freedoms.

In the present Chapter we have presented a method for applying constraints in the Born-Oppenheimer type ab-initio MD in an attempt to investigate the gain in integration time-step, and thus to allow access to bigger systems for longer time scales on an ab-initio level. The implementation, coding and testing of this constrained MD approach within the CASTEP program has been a key part of this PhD project.

6.1.2 Choice of algorithms

There are a number of ways to introduce the rigid molecule approximation in an MD simulation. The classical way is to treat the constrained system in terms of generalised co-ordinates obeying constraint-free equations of motion (the so-called Lagrange-Hamiltonian formalism). This method is quite efficient for simple molecules with few degrees of freedom [Ryckeart, 1975]. However as the number of internal degrees of freedom increases it rapidly become harder to write down explicitly the appropriate equation of motion. For any molecule of moderate complexity, such an approach would be very complicated.

An alternative method for constrained dynamics is to use the methods devised by Ryckeart *et. al.* [Ryckeart, 1977], the most familiar of which is called SHAKE. SHAKE is based on the Verlet algorithm [Verlet, 1967], and has some drawbacks as it does not explicitly include the velocities in the integration of the equations of motion. The constraint technique for the velocity version of the Verlet algorithm [Swope, 1982] (which is used in this work) is an up-date of the SHAKE algorithm, called RATTLE, which also include the velocity constraints. SHAKE and RATTLE retain the simplicity of cartesian co-ordinates and avoids many of the complications of Euler angles and quaternions, while incorporating the effect of constraining the geometry of a molecule.

6.2 SHAKE and RATTLE

In the SHAKE and RATTLE algorithm, constraints appear explicitly in the Lagrangian in cartesian co-ordinates associated with the appropriate multipliers.

eqn 6.1: $$L = T - V - \lambda_1\sigma_1 - \lambda_2\sigma_2 - \lambda_3\sigma_3 \ldots\ldots ,$$

where $\lambda_1, \lambda_2, \ldots\ldots \lambda_l$ are the l Lagrange multipliers associated with the l constraints $\sigma_1, \sigma_2, \ldots\ldots \sigma_l$ in a system. These constraints may represent rigid bonds, bond-angles and dihedral-angles of molecules in the system. However, all the constraints are defined in terms of fixed inter-atomic distances, *i.e.*

eqn 6.2: $$\sigma_k(\{\boldsymbol{r}(t)\}) \equiv \{\boldsymbol{r}_j(t)-\boldsymbol{r}_i(t)\}^2-d_{ij}^2 = 0,$$

where the subscript k stand for a pair of atoms (say i and j). It is thus more convenient to express σ_k as σ_{ij} and λ_k as λ_{ij} The time derivative of the constraint equations give the velocity constraints,

eqn 6.3: $$\dot{\sigma}_{ij}(\{\boldsymbol{r}(t)\}) \equiv \{\dot{\boldsymbol{r}}_i(t) - \dot{\boldsymbol{r}}_j(t)\}.\{\boldsymbol{r}_i(t) - \boldsymbol{r}_j(t)\} = 0.$$

This means that when constraining a bond-angle or dihedral angle in a molecule, we also constrain any atomic distances that are dependent on this parameter (see Figure 6.1). This may seem an unwanted limitation, but in any real system the bond stretching is always a faster mode of vibration than a bond bending mode, and in turn the bond bending mode is always faster than a torsional mode. It would be non-physical to leave the bonds unconstrained while the angle containing these bonds is rigid and the same argument applies for fixing a torsional mode associated with the dihedral arrangement of atoms in a molecule.

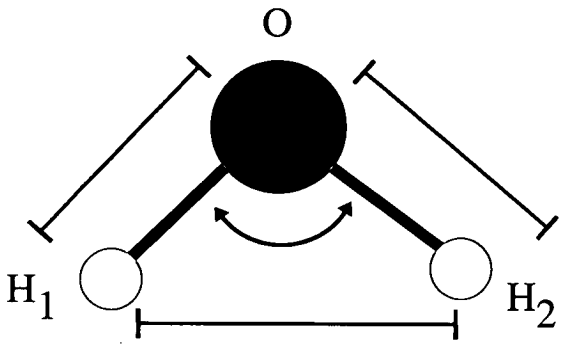


Figure 6.1: To impose a fixed HOH angle in a water molecule, we need to constrain all three distances, H_1O, H_2O and H_1H_2.

Taking the new Lagrangian (eqn 6.1) into the fundamental form of the Lagrangian equation of motion in Cartesian co-ordinates,

eqn 6.4: $$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0,$$

where q is the position co-ordinate $\boldsymbol{r}$ and $\dot{q}$ is the velocity $\dot{r}$, yields the $3N$ Lagrangian equation of motion where the forces of constraint (G_i) appear explicitly.

eqn 6.5: $$m_i \ddot{r}_i^c = F_i + G_i = -\nabla_i V - \sum_{k=1}^{l} \lambda_k \nabla_i \sigma_k.$$

Taking the definition of constraints from eqn 6.2 into eqn 6.5 we get the new acceleration

eqn 6.6: $$\ddot{\boldsymbol{r}}_i^c = \frac{1}{m_i}\{F_i - 2\sum_j \lambda_{ij}(t)\boldsymbol{r}_{ij}(t)\}.$$

These new accelerations (eqn 6.6) contain the original forces plus the exact correction terms due to the constraints. When these constrained accelerations are used in the velocity-Verlet integration (see Section 5.1.2), they therefore generate the desired constrained dynamics of the system.

Taking these new accelerations into the velocity-Verlet integration algorithm explained in Chapter 5 (Section 5.1.2; eqn 5.8 - 5.10) we get the constrained version of the velocity-Verlet with the constrained mid-velocity and position update as

eqn 6.7: $$\boldsymbol{q}_i^c = \dot{\boldsymbol{r}}_i(t) + \frac{h}{2m_i}F_i(t) - \frac{h}{m_i}\sum_j \lambda_{ij}(t)\boldsymbol{r}_{ij}(t) ,$$

eqn 6.8: $$\boldsymbol{r}_i^c(t+h) = \boldsymbol{r}_i(t) + h\dot{\boldsymbol{r}}_i(t) + \frac{h^2}{2m_i}F_i(t) - \frac{h^2}{m_i}\sum_j \lambda_{ij}(t)\boldsymbol{r}_{ij}(t) ,$$

and the constrained velocity update as,

eqn 6.9: $$\dot{\boldsymbol{r}}_i^{\,c}(t+h) = \boldsymbol{q}_i + \frac{h}{2m_i}F_i(t+h) - \frac{h}{m_i}\sum_j \lambda_{ij}(t+h)\boldsymbol{r}_{ij}(t+h).$$

Here $\boldsymbol{r}^c$, $\boldsymbol{q}^c$, $\dot{\boldsymbol{r}}^c$ and $\boldsymbol{r}, \boldsymbol{q}, \dot{\boldsymbol{r}}$ represents the constrained and unconstrained position, mid-velocity and velocity, respectively. Comparing eqn 6.7, eqn 6.8 and eqn 6.9 with the standard velocity-Verlet integration (eqn 5.8, eqn 5.9 and eqn 5.10), the constrained values can be expressed as a correction over the un-constrained values as

eqn 6.10: $$\boldsymbol{q}_i^c = \boldsymbol{q}_i - \frac{h}{m_i}\sum_j \lambda_{ij}(t)\boldsymbol{r}_{ij}(t) ,$$

eqn 6.11: $$\boldsymbol{r}_i^c(t+h) = \boldsymbol{r}_i(t+h) - \frac{h^2}{m_i}\sum_j \lambda_{ij}(t)\boldsymbol{r}_{ij}(t) ,$$

eqn 6.12: $$\dot{\boldsymbol{r}}_i^{\,c}(t+h) = \dot{\boldsymbol{r}}_i(t+h) - \frac{h}{m_i}\sum_j \lambda_{ij}(t+h)\boldsymbol{r}_{ij}(t+h) .$$

In essence, the technique solves the equations of motion for one time step in the *absence* of the constraint forces (Chapter 5, eqn 5.8-5.10) to get the *unconstrained* positions $\boldsymbol{r}_i(t+h)$ and velocities $\dot{\boldsymbol{r}}_i(t+h)$. The unknown multipliers, $\{\lambda_{ij}(t)\}$ and $\{\lambda_{ij}(t+h)\}$, are then

calculated from these unconstrained values by taking eqn 6.11 and 6.12 in eqn 6.2 and eqn 6.3. This is followed by a step where the constrained positions, $\boldsymbol{r}_i^c(t+h)$ and the velocities, $\dot{\boldsymbol{r}}_i^c(t+h)$ are then calculated from eqn 6.10 - 6.12.

There are two ways to perform these co-ordinate corrections. The straightforward way is to solve all the unknown parameters [*e.g.* $\{\lambda_{ij}(t)\}$] to match all the constraints [*e.g.* $\boldsymbol{r}_i^c(t)$] simultaneously. This is a single step process, *i.e.* all the position and velocity co-ordinates will be 'corrected' to meet their constraints all together. This method, known as the method of undetermined multipliers, requires matrix inversion of an lxl matrix (where l is the total number of constraints), and could be very computationally expensive.

An alternative and faster approach, which is used in SHAKE and RATTLE, is to go through the constraints (eqn 6.2, eqn 6.3) one by one, cyclically, 'updating' the velocity and the atomic co-ordinate, so as to satisfy each constraint in turn. We specifically use the phrase 'updating the co-ordinates', instead of 'correcting the co-ordinates' as the SHAKE and RATTLE techniques do not correct all the co-ordinates simultaneously. It 'updates' the whole set of $6N$ co-ordinates ($3N$ positions and $3N$ velocities) towards the correct result. By iterating such co-ordinate 'updates' to self-consistence, this drastic approximation proves to converge rather rapidly to the 'correct' configuration where all constraints are met.

6.2.1 The approximation

The SHAKE and RATTLE algorithm is therefore based on treating each constraint and its linked pair of atoms one at a time. This is achieved simply by removing the sum over all the constraints in eqn 6.10, eqn 6.11 and eqn 6.12 in order that the constraints are no longer linked. This avoids the time consuming matrix inversion at each time step of ion dynamics. The deviation in position and velocity from the constrained configuration is approximated as follows

eqn 6.13: $$-\frac{h^2}{m_i}\sum_j \lambda_{ij}(t)\boldsymbol{r}_{ij}(t) \approx -\frac{h^2}{m_i}\lambda_{ij}(t)\boldsymbol{r}_{ij}(t)\,,$$

eqn 6.14: $$-\frac{h^2}{m_i}\sum_j \lambda_{ij}(t+h)\boldsymbol{r}_{ij}(t+h) \approx -\frac{h^2}{m_i}\lambda_{ij}(t+h)\boldsymbol{r}_{ij}(t+h)\,.$$

The removal of the summation sign in the approximations means that each atom is linked to one single constraint at a time, *i.e.* when the constraint forces from one constraint, say σ_{ij} on atoms i and j are calculated, the contributions from all the other constraints on atoms i and j have been ignored. Let us define, for a particular constraint, σ_{ij}

eqn 6.15: $$g_{ij} = \lambda_{ij}(t), \quad k_{ij} = \lambda_{ij}(t+h).$$

The SHAKE version of the velocity-Verlet integration algorithm for atom i due to σ_{ij} is then

eqn 6.16: $$\boldsymbol{q}_i^c = \boldsymbol{q}_i - \frac{h}{m_i}g_{ij}\boldsymbol{r}_{ij}(t)\,,$$

eqn 6.17: $$\boldsymbol{r}_i^c(t+h) = \boldsymbol{r}_i(t+h) - \frac{h^2}{m_i}g_{ij}\boldsymbol{r}_{ij}(t)\,,$$

eqn 6.18: $$\dot{\boldsymbol{r}}_i^{\,c}(t+h) = \dot{\boldsymbol{r}}_i(t+h) - \frac{h}{m_i}k_{ij}\boldsymbol{r}_{ij}(t+h)\,.$$

For the same constraint σ_{ij} the atom j is moved as

eqn 6.19: $$\boldsymbol{q}_j^c = \boldsymbol{q}_j - \frac{h}{m_j}g_{ji}\boldsymbol{r}_{ji}(t)\,,$$

eqn 6.20: $$\boldsymbol{r}_j^c(t+h) = \boldsymbol{r}_j(t+h) - \frac{h^2}{m_j}g_{ji}\boldsymbol{r}_{ji}(t)\,,$$

eqn 6.21: $$\dot{\boldsymbol{r}}_j^{\,c}(t+h) = \dot{\boldsymbol{r}}_j(t+h) - \frac{h}{m_i}k_{ji}\boldsymbol{r}_{ji}(t+h).$$

Since the constraint associated with a particular pair of atoms, say i and j, is uniquely determined [Goldstein, 1980] *i.e.* $\sigma_{ij}=\sigma_{ji}$ and $\dot{\sigma}_{ij} = \dot{\sigma}_{ji}$, the associated Lagrange multipliers are also uniquely determined, which gives $g_{ij}=g_{ji}$ and $k_{ij}=k_{ji}$ in eqn 6.19 - 6.21. We can, therefore, drop the subscript, *i.e.* $g_{ij}=g$ and $k_{ij}=k$, when referring to the same pair of atoms. With $\boldsymbol{r}_{ij} = -\boldsymbol{r}_{ji}$, we get the constrained values for the j'th atom as

eqn 6.22: $$\boldsymbol{q}_j^c = \boldsymbol{q}_j + \frac{h}{m_j} g \boldsymbol{r}_{ij}(t),$$

eqn 6.23: $$\boldsymbol{r}_j^c(t+h) = \boldsymbol{r}_j(t+h) + \frac{h^2}{m_j} g \boldsymbol{r}_{ij}(t),$$

eqn 6.24: $$\dot{\boldsymbol{r}}_j^{\,c}(t+h) = \dot{\boldsymbol{r}}_j(t+h) + \frac{h}{m_i} k \boldsymbol{r}_{ij}(t+h).$$

In a physical picture this means that the constraint forces of equal magnitudes and opposite orientations are applied to the particles i and j, in the direction of the bond between i and j.

6.2.2 The iterative technique

Position corrections

First the unconstrained equation of motion is solved to give the unconstrained positions $\{\boldsymbol{r}_i(t+h)\}$ using standard velocity-Verlet integration (eqn 5.9, Chapter 5). We then pick a constraint, say σ_{ij} of atom i and j. Let $\boldsymbol{r}_{ij}(t+h) = \boldsymbol{r}_i(t+h) - \boldsymbol{r}_j(t+h)$.

Then $\boldsymbol{r}_{ij}(t+h)$ is the initial approximation for the vector displacement of atom i and j. If $\sigma_{ij} \equiv |\boldsymbol{r}_{ij}(t+h)|^2 - d_{ij}^2$ differs from zero (see eqn 6.2) by an amount less than an acceptable tolerance, we move to a different constraint. If not, we need to generate constrained values of them, *i.e.* $\boldsymbol{r}_i^c(t+h)$ and $\boldsymbol{r}_j^c(t+h)$ using the SHAKE equations for

position correction (eqn 6.17 and eqn 6.23), such that they satisfy the constraints exactly, *i.e.*

eqn 6.25: $$\sigma_{ij} \equiv |\boldsymbol{r}_i^c(t+h) - \boldsymbol{r}_j^c(t+h)|^2 - d_{ij}^2 = 0,$$

which gives the much needed value of g (taking the positions from eqn 6.17 and eqn 6.23 into eqn 6.25)

eqn 6.26 : $$g = \frac{r_{ij}^2(t+h) - d_{ij}^2}{2h^2\{\boldsymbol{r}_{ij}(t+h).\boldsymbol{r}_{ij}(t)\}\ (1/m_i+1/m_j)},$$

where we have neglected the quantities of order g^2. Once g is obtained the constrained position is calculated using eqn 6.17 and eqn 6.23, and we move to a different constraint say σ_{jk}, and repeat the same correction. In this way the correction procedure is carried out for all constraints in the system.

However in a real system (apart from the case of a diatomic molecule) one atom can be involved in several constraints. A local solution to a particular constraint σ_{ij}, will always give rise to an additional deviation to its linked constraint, say σ_{jk}. Once all the constraints in the system have been treated one by one, those new or additional deviations will be left in the system. These correction procedure is iterated until all the constraints are satisfied to within a given tolerance. There is rigorous mathematical proof that such iterative corrections does converge to the correct results after sufficient iteration. At each stage of the iterative procedure, the positions are corrected by an amount of the proper form, and the procedure terminates when all the constraints are satisfied to within a desired accuracy.

Velocity corrections

At the next stage the constrained velocities $\dot{\boldsymbol{r}}_i^{\,c}(t+h), \forall i$, are solved by a similar iterative procedure. First the unconstrained velocities, $\dot{\boldsymbol{r}}_i(t+h), \forall i$, are

generated using the standard velocity-Verlet integration (eqn 5.10, Chapter 5). We then pick a constraint, say $\dot{\sigma}_{ij}$ for atoms i and j. The velocity constraint for the atom pair is then

eqn 6.27: $$\dot{\sigma}_{ij} \equiv \{\dot{\boldsymbol{r}}_i(t+h) - \dot{\boldsymbol{r}}_j(t+h)\}.\{\boldsymbol{r}^c_{ij}(t+h)\} = 0.$$

Note that positions $\boldsymbol{r}^c_{ij}(t+h)$ are already constrained from the procedure described in previous section. If the newly generated velocities $\dot{\boldsymbol{r}}(t+h)$ are also constrained, $\dot{\sigma}_{ij}$ should be zero (eqn 6.3). If it differs from zero by less than an acceptable tolerance, we move to a different constraint. If not, we need to generate constrained velocities $\dot{\boldsymbol{r}}_i{}^c(t+h)$ and $\dot{\boldsymbol{r}}_j{}^c(t+h)$ using the RATTLE [Andersen, 1983] equation for velocity correction (eqn 6.18 and eqn 6.24), such that they satisfy the constraints exactly, *i.e.*

eqn 6.28: $$\dot{\sigma}_{ij} \equiv \{\dot{\boldsymbol{r}}_i{}^c(t+h) - \dot{\boldsymbol{r}}_j{}^c(t+h)\}.\boldsymbol{r}^c_{ij}(t+h) = 0 .$$

Taking $\dot{\boldsymbol{r}}_i{}^c(t+h)$ and $\dot{\boldsymbol{r}}_j{}^c(t+h)$ from eqn 6.18 and eqn 6.24 into eqn 6.28 we get the value of k,

eqn 6.29: $$k = \frac{\{\dot{\boldsymbol{r}}_i(t+h) - \dot{\boldsymbol{r}}_j(t+h)\}.\boldsymbol{r}^c_{ij}(t+h)}{h\,d^2_{ij}\,(1/m_i+1/m_j)} .$$

Once the value of k is obtained the constrained velocities for atoms i and j are then calculated from eqn 6.18 and 6.24. All the velocities linked with all the constraints in the system are than corrected in the same way with this approximate Lagrange parameter.

As discussed in the case for the position constraints, once all the constraints in the system have been treated one by one, new or additional deviations will be left in the system. However if the whole procedure is iterated, system velocities will converge to the correct result. At each stage of the iterative procedure, the velocities, $\dot{\boldsymbol{r}}_i(t+h)$ are corrected by an amount of the proper form, and the procedure terminates when all the constraints $(\{\dot{\sigma}_{ij}\})$ are satisfied to within desired accuracy levels.

6.3 The Constrained Dynamics Code for the NVE Ensemble

In this section we present the algorithm for the constraint MD using the velocity-Verlet algorithm for NVE ensemble as discussed so far. The expansion of this to the NVT Nose-Hoover and NVT Langevin dynamics is fairly straight forward. However, the expansion of constrained dynamics to constant pressure (NPT) ab-initio dynamics needs to address some further theoretical questions, which have not been covered in this work.

Part 1: Initialization

1. Read the positions $\{\boldsymbol{r}_0\}$ and velocities $\{\boldsymbol{v}_0\}$ of all the atoms at time t, from an input file. Each of these variables is an array of $3N$ element, where N is the total number of atoms or ions in the system.
2. If velocities are not provided, generate $\{\boldsymbol{v}_0\}$ according to the Boltzmann distribution for a given temperature, T.
3. Calculate the forces $\{\boldsymbol{f}(\boldsymbol{r}_0)\}$ and acceleration $\{\boldsymbol{a}_0\}$ for the initial positions of the atoms using a DFT code.
4. Read the bonds to be constrained from an input file and define the bond distance, $d(i,j)=|\boldsymbol{r}_0(i)-\boldsymbol{r}_0(j)|$, for the initial position of the atoms. We would like to constrain atoms in these initial separations, $d(i,j)$, throughout the MD simulation.

Part 2: Velocity-Verlet position update

1. Do the first step of the velocity-Verlet algorithm (eqn 5.8 and eqn 5.9, Chapter 5). First calculate the mid velocity and store it to a new array $\{\boldsymbol{v}_p\}$ (this will reduce the extra effort in using another $3N$ variables),

$$\boldsymbol{v}_p(i) = \boldsymbol{v}_0(i) + \frac{h}{2}\boldsymbol{a}_0(i).$$

2. Update the position of the particle i to the next time $\boldsymbol{r}(t+h)$ using eqn 5.9 as

$$\boldsymbol{r}_p(i) = \boldsymbol{r}_0(i) + h\boldsymbol{v}_p(i) .$$

3. Repeat these updates for y and z co-ordinate of the i'th atom, and then for all atoms in the system.

Part 3: Constrain new positions (see the flow chart in Section 6.3.1)

1. Pick a single bond, k, linked to the atoms i and j and calculate the deviation caused by the velocity-Verlet position update,

$$D\_r = |\boldsymbol{r}_p(i) - \boldsymbol{r}_p(j)|^2 - d^2(i,j) .$$

2. If $D\_r > \tau$, the tolerance for the constraint to converge, apply the correction to the position and the mid-velocity for the particles i and j using the following steps. Otherwise pick a new bond and repeat the procedure. If no bond is found with $D\_r > \tau$ for all i and j (*i.e.* all constraints have been matched), then the calculation moves to the next stage of the velocity update, covered in Part 4.

3. Calculate $A = \{\boldsymbol{r}_0(i) - \boldsymbol{r}_0(j)\}.\{\boldsymbol{r}_p(i) - \boldsymbol{r}_p(j)\}$. If $A = 0$, the SHAKE correction to the position will not work for this bond. Pick a new bond and repeat the procedure. The chance that the calculation will meet $A = 0$ is very small. However, if $A = 0$ occurs repeatedly (*i.e.* the bond has rotated through a 90° angle), stop the calculation and repeat the simulation with a smaller time step.

4. Calculate $g' = D\_r/(2\mu A)$, where $\mu = 1/m(i) + 1/m(j)$. Then define

$$dx = -g'.\{r_{x0}(i) - r_{x0}(j)\},$$

$$dy = -g'.\{r_{y0}(i) - r_{y0}(j)\},$$

$$dz = -g'.\{r_{z0}(i) - r_{z0}(j)\}.$$

5. Correct the position of atoms i and j linked by bond k as

$$r_{xp}(i) = r_{xp}(i) + dx/m(i)$$
$$r_{xp}(j) = r_{xp}(j) - dx/m(j)$$

and similar corrections for y and z co-ordinates.

6. Correct the mid-velocity as

$$v_{xp}(i) = v_{xp}(i) + \frac{dx}{h.m(i)}$$
$$v_{xp}(j) = v_{xp}(j) - \frac{dx}{h.m(j)}$$

and similar corrections for y and z co-ordinates.

7. Pick a new bond and correct positions and mid-velocities to its linked atoms (step 1 to 6) to match the constraint. Do this for all bonds to be constrained.
8. Repeat steps 1 to 7 iteratively until a situation arises when no further bonds (identified in step 2) require the correction.

Part 4: Ab-initio force calculation at new positions

1. Using first principle calculation, calculate the new forces $\{f(r_p)\}$ and acceleration $\{a_p\}$ at these new positions, $\{r_p\}$. Note this is the most time consuming step in an ab-initio MD simulation.

Part 5: Velocity-Verlet velocity update

1. Update the mid-velocities to the velocities at next time step $v(t+h)$ using eqn 5.10 (see Section 5.1.2, Chapter 5) as

$$v_p(i) = v_p(i) + (h/2)a_p(i)\,, \quad \text{for all atoms in the system.}$$

Part 6: Constrain new velocities (see the flow chart in Section 6.3.2)

1. Pick a bond, k, linked to the atom i and j and calculate the deviation caused by the velocity-Verlet velocity update,

$$D\_v = \{r_p(i) - r_p(j)\}.\{v_p(i) - v_p(j)\}\,.$$

2. If $D\_v > \tau$, the tolerance for the constraint to converge, apply the correction to the new velocities, $v_p(i)$ and $v_p(j)$ using the following steps. Otherwise pick a new bond. If no bond is found with $D\_v > \tau$, both the position and velocity constraint is satisfied and the calculation can move to the next MD step.

3. Calculate $k' = \dfrac{D\_v}{\mu d^2(i,j)}$. Then define

$$dv\_x = -k'\{r_{xp}(i) - r_{xp}(j)\},$$

$$dv\_y = -k'\{r_{yp}(i) - r_{yp}(j)\},$$

$$dv\_z = -k'\{r_{zp}(i) - r_{zp}(j)\}.$$

4. Correct the velocities of atoms i and j linked to bond k as

$$\begin{aligned} v_{xp}(i) &= v_{xp}(i) + dv\_x/m(i) \\ v_{xp}(j) &= v_{xp}(j) - dv\_x/m(j) \end{aligned}$$

and similar corrections for y and z co-ordinates.

5. Repeat steps 1 to 4 for all the bonds to be constrained.

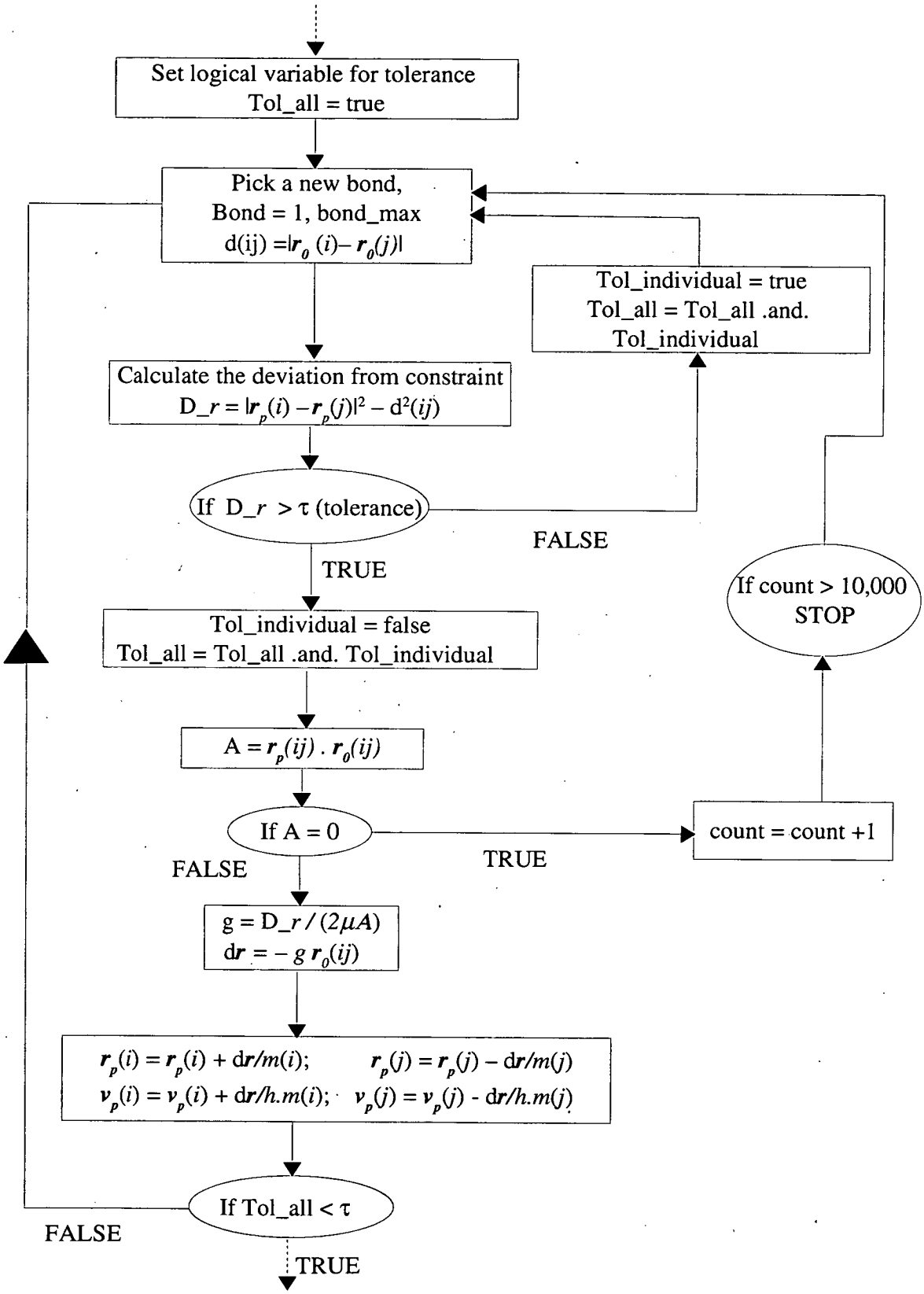
6. Repeat steps 1 to 5 iteratively until all the bonds match the constraints. (No bond is available for correction in step 3).

7. At this point store the new constrained velocities and positions to the original variables

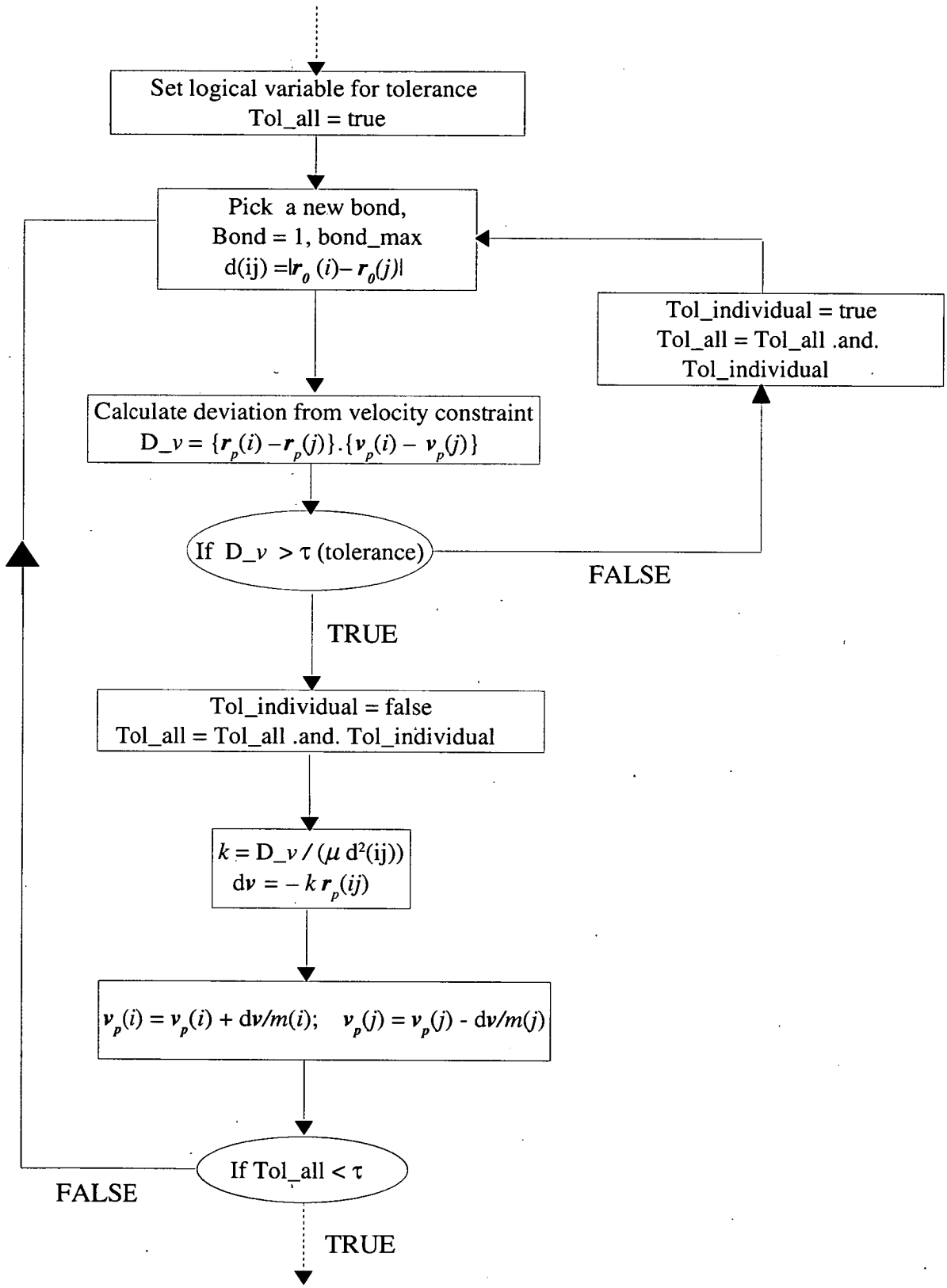
$$\boldsymbol{r}_0 = \boldsymbol{r}_p, \quad \boldsymbol{v}_0 = \boldsymbol{v}_p, \quad \boldsymbol{a}_0 = \boldsymbol{a}_p;$$

and move back to start of Part 2 for the next MD step.

6.3.1 Flow chart for part 3: constrain new positions



6.3.2 Flow chart for part 6: constrain new velocities



6.4 The Constrained Dynamics Code for the NVT Ensemble

In an NVT simulation, molecular dynamics is performed at a constant temperature. From a statistical mechanical point of view, we impose a temperature on the system by bringing it into thermal contact with a large heat bath. Under this situation the probability of finding the system in a given energy state is given by the Boltzmann distribution. This does not simply mean that the kinetic energy per particle is fixed at k_BT, rather they fluctuate around the equipartition energy according to the Boltzmann distribution. The instantaneous kinetic temperature thus also fluctuates around the imposed temperature,T.

From the numerical point of view the main difference in an NVT simulation compared to an NVE simulation is that the velocities of the particles in the system are rescaled either directly or via a stochastic collision with the heat bath (true forces are modified by random kick) to generate a canonical ensemble at constant temperature [Allen, 1987]. This means that only the second step of the velocity-Verlet algorithm (Section 6.3, Part 5) has to be modified to adapt with the velocity rescaling.

There exist two major NVT algorithms, depending on whether the velocities are rescaled directly using a thermostat or via stochastic forces. The NVT Nose-Hoover chain algorithm used in CASTEP computational package introduces two new degrees of freedom, which act as a thermostat for the rest of the system to maintain the desired temperature [Nose, 1984; Hoover, 1985; Martyna, 1992]. The implementation used 'effectively' rescales the velocities at each step of MD by a common factor, which is calculated using the Nose-Hoover chain method [Martyna, 1996]. This velocity rescaling takes place immediately *after* the velocity update in the second step (*i.e.* after Part 5, Section 6.3) of velocity-Verlet algorithm, as $\dot{r}=\dot{r}*\alpha$, where α is the Nose-Hoover re-scaling factor. Everything else remains the same. The success of the Nose-Hoover chain algorithm therefore lies in proper calculation of the scaling factor.

The NVT Nose-Hoover algorithm is fully compatible with the SHAKE and RATTLE constrain algorithm as the bond velocity constraint expressed in eqn 6.3,

$$\text{eqn 6.3} \quad \dot{\sigma}_{ij} \equiv \dot{\boldsymbol{r}}_{ij}^{\,c}(t+h).\boldsymbol{r}_{ij}^{c}(t+h) = 0,$$

is independent of the scaling. If the velocity component $\dot{\boldsymbol{r}}_{ij}^{\,c}$ is zero along the bond, the velocity component of $\alpha * \dot{\boldsymbol{r}}_{ij}^{\,c}$ is also zero along the bond. In a constrained NVT Nose-Hoover algorithm the procedure is as follows: first, the velocities are updated to the next time step without constraint (Part 5, Section 6.3), second, the constraints are then applied to the velocities (Part 6, Section 6.3). The scaling factor will be calculated using the Nose-Hoover chain for those constrained velocities. Since the calculation of the scaling factor inputs the constrained velocities rather than the unconstrained velocities, it will automatically results in the correct temperature for the reduced degrees of freedom.

In Langevin dynamics the NVT ensemble is generated using a stochastic collision of the atoms in the system with an imaginary heat bath from time to time to reflect a Boltzmann distribution of velocities [Chandrasekhar, 1943]. This is done by adding a fluctuation term to the calculated forces just *before* the 2nd step of velocity-Verlet velocity update (*i.e.* at Part 4, Section 6.3). Apart from the addition of these fictitious forces, everything else remain the same as in the NVE algorithm. A change in the forces will lead to additional velocities along the constrained bonds. When these velocities are corrected for the constraints according to SHAKE and RATTLE, velocity components along the constrained bonds are zeroed, and the overall kinetic energy will therefore be reduced for that step. However the reduced kinetic energy counts for the reduced degrees of freedom and the temperature remains the same.

The nature of the iterative minimisation approach in SHAKE and RATTLE is such that *the constraint forces* (the forces arising from the Lagrange multipliers in eqn 6.1) are always central (see eqn 6.16 - 6.18 and eqn 6.22 - 24). Such a correction is justified when the origin of deviation from the constraint is also a central field, which is the case for

6.5 Code Implementation within CASTEP

As a result of this work the position and velocity constraint algorithm have now been implemented for the commercially available first-principles calculation package, CASTEP, under the UK Car-Parrinello agreement. The two step velocity-Verlet algorithm in the CASTEP MD package is carried out in two separate subroutine (md_vv1 and md_vv2, which corresponds to Parts 2 and 5 respectively of the algorithm discussed in Section 6.3). The implementation of the constrained dynamics is done using two separate call Subroutines (constrain_position and constrain_velocity) for SHAKE position and RATTLE velocity corrections, respectively. These subroutines correspond to Parts 3 and 6 respectively, of the algorithm described Section 6.3.

At the start of MD simulation the initial velocities (generated according to the Boltzmann distribution) are constrained by the subroutine call 'constrain_velocity'. This ensures that the initial input velocities are constrained. The call for the 'constrain_position' subroutine is made immediately after the first step of the velocity-Verlet *position* update (*i.e.* after Part 2, Section 6.3) and before performing any other property calculation. Similarly, the 'constrain_velocity' subroutine is called immediately after the second step of the velocity-Verlet *velocity* update (i.e after Part 5, Section 6.3) and before undertaking any other property calculation or velocity rescaling in the case of Nose-Hoover NVT simulation.

As discussed before all the constraints (either rigid bond, angle or dihedral) are expressed in terms of rigid bonds or a combination of rigid bonds. A separate input file is supplied alongside the standard CASTEP input files (<seed>.param and <seed>.cell file) to define all the constraints in the systems in terms of rigid bonds. Below is a typical constraint input file for and MD simulation of solid Ammonia (unit cell comprising four molecules) where all the N-H bonds are fixed.

6.5.1 Input file for constraint dynamics

```
comment     : 4 ammonia molecules in a box
constrain   : true
bond_tot    : 12
pos_tol     : 0.0000001
vel_tol     : 0.0000001
max_itrn    : 10000

%BLOCK constraint
bond    N   1   :   H   1
bond    N   1   :   H   5
bond    N   1   :   H   9
bond    N   2   :   H   2
bond    N   2   :   H   7
bond    N   2   :   H   12
bond    N   3   :   H   3
bond    N   3   :   H   8
bond    N   3   :   H   10
bond    N   4   :   H   4
bond    N   4   :   H   6
bond    N   4   :   H   11
%ENDBLOCK constraint
```

The first line is a comment line for the user. The second line contains the new logical variable, constrain. If it is 'true' the constraint dynamics will be executed. The third line contain the total number of rigid bonds in the system. This is the same as the total number of lines within the constraint block. The next two lines allows the user to define tolerance levels for the convergence of position and velocity constraints in the iterative minimisation procedure. The last line of the first block gives an upper bound for the iterative minimisation technique. If either the position or the velocity constraints are not met within this number of iterations, the calculation will terminate. Each single iteration involves a complete SHAKE and RATTLE correction cycle for all the fixed bonds in the system. Each line in the block in between "%BLOCK constraint" and "%ENDBLOCK constraint":

e.g. bond N 1 : H 1

defines a rigid bond in the system that needs to be constrained. Each bond is declared with a keyword "bond" followed by the atom symbol and atom number of both atoms (participating in that bond) separated by a colon. The atom number has to be the serial number of the atom as it appears in the CASTEP <seed>.cell input file. The first nitrogen atom listed in the cell file will be N1 and so on; similarly the first hydrogen atom listed in the cell file will be H 1 and so on.

6.6 Conclusion

The application of constrained dynamics in the Car-Parrinello MD (CPMD) method has been recently demonstrated by Allesh *et. al.* [Allesh, 2004] for the rigid liquid water system. The work presented here describes the implementation of constrained MD for the Born-Oppenheimer Molecular Dynamics Method, which provides an important advantage for first-principles simulations of molecular systems (solid and liquid) where chemical reactions do not take place, and opens up the exciting possibility of investigating phenomena which would otherwise be intractable to calculation.

6.7 Bibliography

1. Allen, M. P., and D.J. Tildesley, *Computer simulation of liquids*, (Oxford University Press, 1987).
2. Allesch, M., E. Schwegler, F. Gygi and G. J. Galli, *Chem Phys*, **120** (11), 5192 (2004).
3. Andersen, H. C., *J. Comput. Phys.*, **52**, 24(1983).
4. Car, R., and M. Parrinello, *Phys. Rev. Lett.*, **55**, 2471(1985).
5. Chandarasekhar, S., *Rev. Mod. Phys.*, **15**, 1(1943).
6. Goldstein, H., *Classical Mechanics*, Second Edition, (Addison-Wesley Publishing Company, Inc., 1980).
7. Hoover, W. G., *Phys. Rev. A*, **31**, 1695(1985).
8. Martyna, G. J., M. L. Klein and M. Tuckerman, *J. Chem. Phys*, **97** (4), 2635(1992).
9. Martyna, G. J., M. Tuckerman, D. J. Tobias and M. L. Klein, *Mol. Phys.*, **87** (5), 1117 (1996).
10. Nose, S., *J. Chem. Phys.*, **81**, 511(1984).
11. Pellegrini, M., N. A. Larsen, T. O. Yeates and N. G. Jensen, *Physica A*, **261,** 224(1998).
12. Ryckeart, J. P., and A. Bellemans, *Chem. Phys. Lett.*, **30,** 123(1975).
13. Ryckeart, J. P., G. Ciccotti, and H. J. C. Berendsen, *J. Comput. Phys.* **23**, 327(1977).
14. Soper, A. K., *Chem. Phys.*, **258**, 121(2000).
15. Sorenson, J. M., G. Hura, R. M. Glaeser and T. H. Gordon, *J. Chem. Phys.*, **113**, 9149 (2000).
16. Swope, W. C., H.C. Andersen, P.H. Berens and K. R. Wilson, *J. Chem. Phys.*, **76**, 637 (1982).
17. Verlet, L., *Phys. Rev.*, **159**, 98(1967).

Chapter 7

Constrained Dynamics of Solid Ammonia (Phase I)

7.1 Introduction

In Chapter 6, we have reported an algorithm for constrained dynamics, which is now ready to be implemented into the Born-Oppenheimer first-principles calculation. This algorithm, which is an implementation of Shake and Rattle [Ryckeart, 1977; Andersen, 1983], provides an iterative minimisation technique of constraining bonds, angles and dihedrals of a molecular system during it's dynamical trajectory. The technique is particularly helpful in systems where the high frequency modes (*i.e.* bond stretching and bond bending) are usually decoupled from the low frequency modes (*e.g.* lattice modes) of vibration. Depending on the particular system of interest, a time step gain of about 4 to 10 times is possible (see Table 6.1, Chapter 6) under the constrained dynamics scheme. In the present chapter, we report the practical implementation of this algorithm to increase the time step of a molecular dynamics simulation, thus allowing access to bigger systems for a longer time scale on a quantum mechanical level.

This work reports the detailed insight obtained by this new technique for the solid-state phase I structure of ammonia at ambient pressure (~0.1 GPa) and at a temperature around 100K. The rich physical and chemical properties of ammonia have attracted a large

number of theoretical and experimental studies on both the solid (at different temperatures and pressures) and gas phases (monomer and dimer structures). The geometrical structure and dynamical behaviour of both these phases is well understood from many advanced experiments [Binbrek, 1972; Reding, 1951; Boese, 1997; Helminger, 1971]. Recent advances in first principles simulations have now attracted many physicists to cast a fresh eye on modelling ammonia at the microscopic level [Morrison, 2003; Fortes, 2003; Boese, 2003].

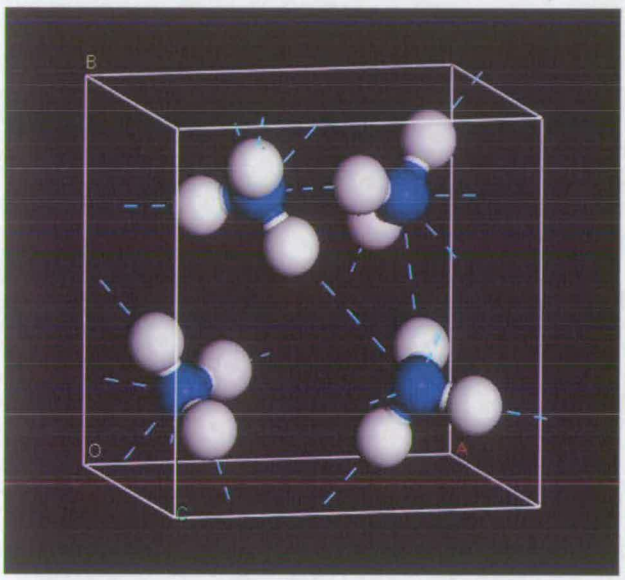


Figure 7.1 : The phase-I crystal structure of ammonia

In this work we have looked at a different aspect of the solid ammonia phase-I structure, specifically to investigate how the dynamical and structural properties of ammonia are affected upon 'freezing out' its slow moving internal degrees of freedom. The phase-I ammonia structure is a cubic crystal of $P2_13$ symmetry with lattice constant a = 5.1305(8) Å [Boese, 1997]. Each unit cell consists of four ammonia molecules networked to each other by tri-furcated hydrogen bonds (Figure 7.1), each of strength 9.3 $kJmol^{-1}$ [Morrison, 2003]. Therefore, each molecule therefore both accepts and donates three

hydrogen bonds, which deviate significantly (by ~20^0) from the almost perfectly linear hydrogen bond observed in water ice [Ebbing, 1987]. Although each lone pair orbital on the nitrogen atom is being shared between three nearest neighbours, our work (see Chapter 3 for details) showed that a tri-furcated hydrogen bond in solid ammonia is as strong as a standard hydrogen bond occurring in an ammonia dimer. We infer from this that the hydrogen bonding in solid ammonia is largely electrostatic [Morrison, 2003].

The intra-molecular vibrations, *i.e.* the N-H stretch (ca. 3300 cm^{-1}) and H-N-H angle bends (ca. 1000 cm^{-1} and 1700 cm^{-1}) [Binbrek, 1972] in solid ammonia, are much faster than the other inter-molecular modes attributed to molecular rotation and translation (known as lattice modes). It can be assumed that these higher frequency modes are fairly de-coupled from the lattice vibrational modes [Ryckeart, 1977]. Moreover in the gas phase the high frequency N-H stretch (ca. 3400 cm^{-1}) is fairly decoupled from the H-N-H angle bend modes (ca. 900 cm^{-1} and 1600 cm^{-1}) for the isolated molecule [Reding, 1951]. Thus solid ammonia is an ideal system to apply molecular constraints to, either partially (N-H bond only) or fully (rigid ammonia), and to study the effect of constraining the dynamical and structural properties of the whole system.

Since each N-H bond in ammonia is connected to its neighbours by N-H...N hydrogen bonds, constraining one may indirectly affect the other. According to covalent bond theory it seems reasonable to assume that the electron density around the N-H bond re-distributes with the change in the bond length and orientation. This then raises the interesting possibility that the associated hydrogen bonds may be affected by this re-distribution of electron density (due to the N-H bond stretching motion). If this is true, the dynamics of the whole system should be affected by the application of rigidity to the molecular skeleton.

Since such a dynamic coupling between the covalent bond and hydrogen bond interactions would be an electronic level phenomena, the empirical-potential based classical MD simulations cannot shed any light on it. On the other hand, at the time of the start of this

project, the available first-principles simulation packages (*e.g.* CASTEP [Segall, 2002], VASP [Kresse, 1996], CPMD *etc.*) did not not allow the application of molecular constraints[1]. The constrained dynamics algorithm developed in this work therefore gives us a chance to probe the chemistry of the phase I ammonia crystal. By applying constraints to the bond lengths (and consequently on their associated electronic distribution) we can now investigate the net effect of N-H bond constraints on the dynamics of the hydrogen bonds at the electronic level.

The primary aim of this chapter is to report on the effect of increasing the integration time step in BO-MD simulations of solid ammonia. To this end we have carried out a series of calculations using fixed-bond and rigid molecule approximations, and compare results with those obtained from flexible ammonia simulations (i.e without imposing any constraints). We discuss in detail the structural and dynamical properties (*i.e.* vibrational spectra) of solid ammonia and consider the effects of constraints. Although it is not the aim of this chapter, we also present the available experimental dynamical and structural properties alongside the simulation results for basic comparison. A more detailed comparison of the vibrational properties of solid ammonia will follow in the next chapter.

1 At the same time that this work was undertaken, constraint dynamics was also developed within the frame work of the CPMD program [Allesch, 2004].

7.2 Computational Method

For ammonia, the initial geometry was taken from the X-ray diffraction structure by Boese *et al.* at 160 K [Boese, 1997] (see Figure 7.1). The experimental structure was optimised at 0 K and ambient pressure using the standard plane-wave total-energy density-functional-theory (DFT) calculation package, available through the UK Car-Parrinello agreement [CASTEP, v4.2; Payne, 1992]. Details of this calculation are available in Chapter 2.

The MD simulations were performed using the new CASTEP package [new CASTEP, v2.2; Segall, 2002] on a 2x2x2 super-cell (of the optimised unit cell) with P1 symmetry, containing 32 molecules (128 atoms) in a cubic box of lateral dimension 10.311 Å. Running simulations on a super-cell is necessary in order to visualise possible lattice vibrations of the crystal system, which have a periodic repeat length greater than the crystallographic 1x1x1 cell. This will also help in spotting/indicating the LO-TO active modes [Ashcroft, 1976]. An MD simulation run from an unit cell would not show any LO-TO splitting due to the fact that long-range polarization field are not compatible with the periodic boundary condition of a single unit cell. Usually large/elongated supercells are used to recover the long range polarisation effect (*i.e.* LO-TO splitting) in an ab-initio calculation [Sarnthein, 1997]. However, the long range polarisation fields are artificially suppressed in such a supercell calculation (due to the missing very long range component of the electrostatic interaction). In a 2x2x2 supercell of ammonia, although the main contribution of the long range field will be missing, we might expect to see small build up of polarisation over the supercell, with the LO branch more likely to be wrongly positioned as the size of the LO-TO splitting will be artificially reduced.

The calculations were performed with the GGA-PBE functional [Perdew, 1996] to describe the exchange-correlation potential and the ultra-soft pseudo-potentials [Vanderbilt, 1990] to describe the wave-functions of the individual atoms. The accuracy in the description of the electronic system can be continuously improved by increasing the basis

set cut-off energy until a desired convergence level is reached. For the MD simulations reported here, a plane-wave basis set cut-off energy of 300 eV was found to be sufficient as it converges the forces on the atoms to less than 0.1 eV/Å per atom, well within the level of acceptance in a typical MD simulation. Further improvements in the cut-off energy to 350 eV resulted in only very small changes in the forces (*ca.* 0.01 eV/Å for hydrogen and 0.02 eV/Å for nitrogen) The finite basis set correction [dE/dlog (E_{cut})] [Francis, 1990] for the 300 eV cut-off was 0.125 eV per atom, which is again within accepted tolerance limits.

With respect to k-point sampling of the first Brillouin zone, four k-points, generated using a 2x2x2 Monkhorst-Pack grid [Monkhorst, 1976], were found to show convergence [with forces on an average atom at its relaxed position changing by just 0.0001 eV/Å due to further improving the Monkhurst-Pack grid to 3x3x3 (14 k points)]. The Fast Fourier Transform Grid was used to communicate between real and reciprocal space. CASTEP grid parameter was set to 1.50 as prescribed by the package to minimise the wrap around error in the transformation [Segall, 2002; CASTEP Manual, v4.2].

At the start of the MD simulation the initial electronic configuration was generated from a random choice of weighting co-efficients for the plane waves used to describe the electronic wave-function. The coefficients were then optimised using the density mixing conjugate gradient optimisation method [Kresse, 1996] to generate the lowest energy configuration. At each successive time step of the ionic dynamics, instead of a completely random choice, the new initial electronic configuration was generated by extrapolating forward from the previous *ground-state* electronic configuration of the last ionic step [Arias, 1992].

The forces on the ions were calculated from the ground-state electronic configuration using the Hellman-Feynman theorem [Feynman, 1939]. These forces were then used by the velocity-verlet algorithm to generate the MD trajectory for the next ionic configuration. Three different MD trajectories were generated starting from the same initial configuration of the 2x2x2 ammonia super-cell to investigate the effects of constraining the

dynamics of the system. The time step for each simulation was chosen such that each complete oscillation of the fastest available mode of vibration is sampled at least 20 times during the simulation. The fastest mode in a flexible ammonia simulation is the asymmetric N-H stretch of 3370 cm^{-1}; for the bond fixed simulation it is the asymmetric H-N-H angle bend mode of 1679 cm^{-1} [Binbrek, 1972]; and finally for the rigid ammonia simulation the fastest mode is expected to fall below 800 cm^{-1}. Thus, time steps of 0.5 fs seconds for the first, 1fs for the second and 2.0 fs for the third simulation were *anticipated* to be appropriate to integrate the dynamics of these systems correctly. All three different simulations were run to generate NVE ensembles from the same initial temperature, 200 K. The type and description of all the constraints to be applied were read from a separate constraint input file (see Section 6.5.1, Chapter 6). The constraints were then applied using the Shake and Rattle algorithm (see Section 6.3 and 6.4, Chapter 6) developed for CASTEP in the course of this work.

7.3 Results and Discussion

The systems tended to equilibrate (*i.e.* the kinetic energy and potential energy become constant) very fast (at around 100 K) upon initiating the MD simulation from the optimised ammonia structure. We equilibrated the systems for 1 ps, before collecting the MD data for ca. 7 ps for the fixed-bond and rigid ammonia simulations, and for 5 ps for the flexible ammonia system. Although we initialised the velocities with the same starting temperature (200 K) and from the same initial geometry, the temperatures of the three different simulations (and hence the pressures) were slightly different because of the different constraints applied to the three different systems. After equilibration the temperature of the fixed-bond simulation stabilised about 100 K, whilst both the flexible and rigid ammonia simulations stabilised about 107 K. Since the initial box size for these simulations corresponded to the optimised structure at 0 K, the fixed volume simulations at higher temperature obviously generates some pressure (of about 0.1 GPa) on both constrained and unconstrained systems (see Figure 7.2).

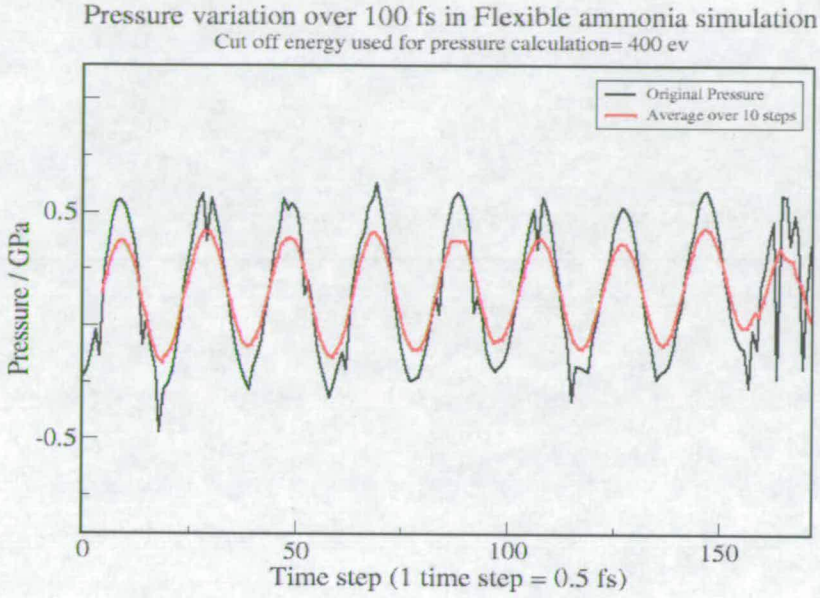


Figure 7.2: The pressure of the flexible ammonia system at 110 K

Figure 7.2 shows the pressure variation of the 2x2x2 flexible ammonia system along the simulation trajectory. The solid black line shows the pressure at each time step. For clarity, we have also drawn the average pressure (solid red line), which is the average of instantaneous pressure over every ten time steps along the simulation length. From previous theoretical and experimental work [Eckert, 1984; Gauthier, 1986; Yurtseven, 2004] it has been confirmed that the ammonia phase I structure is stable at temperatures up to 217 K and pressure up to 1.4 GPa (14 kbar). Therefore despite a pressure of 0.1 GPa, our simulation systems stay well within the stability region expected for phase-I ammonia.

7.3.1 Conservation of energy and temperature - justifying the time steps

The time steps for the three simulations were confirmed by careful observation of the thermodynamic quantities, energies and temperature, of the system. We analysed the conservation of each of these quantities by plotting their averages over every 100 fs of the simulation length.

The 0.5 fs time step was confirmed as adequate for the flexible ammonia simulation as the average total energy stayed almost constant, with only a very slight tendency to increase. The average temperature over the whole simulation run is 107.91 K with a standard deviation of 5.87 K [see Figure 7.3(a)]. For the constrained N-H bond simulation, the chosen time step of 1.0 fs was found to produce very good stability in the energy and temperature curves [see Figure 7.3(b)]. The average total energy stayed fixed, and average temperature over the whole simulation run is 100.12 K with a standard deviation of 6.35 K. Finally, the energy and temperature curves for the fully constrained simulation [see Figure 7.3(c)] show that a 2.0 fs time step, representing a four-fold increase compared to the first simulation, is possible. Energy conservation is still maintained with temperature at 106.91 $\pm$ 9.08 K. In all three cases, the temperature fluctuation stays well within the prescribed theoretical value, *i.e.* of the order of $1/\sqrt{N}$, which is about 11 K in this case (N is the total number of particles in the system). Note that the slightly larger fluctuation in the fixed bond and the rigid ammonia simulation temperature is attributed to the reduction in the total number of degrees of freedom available to the system.

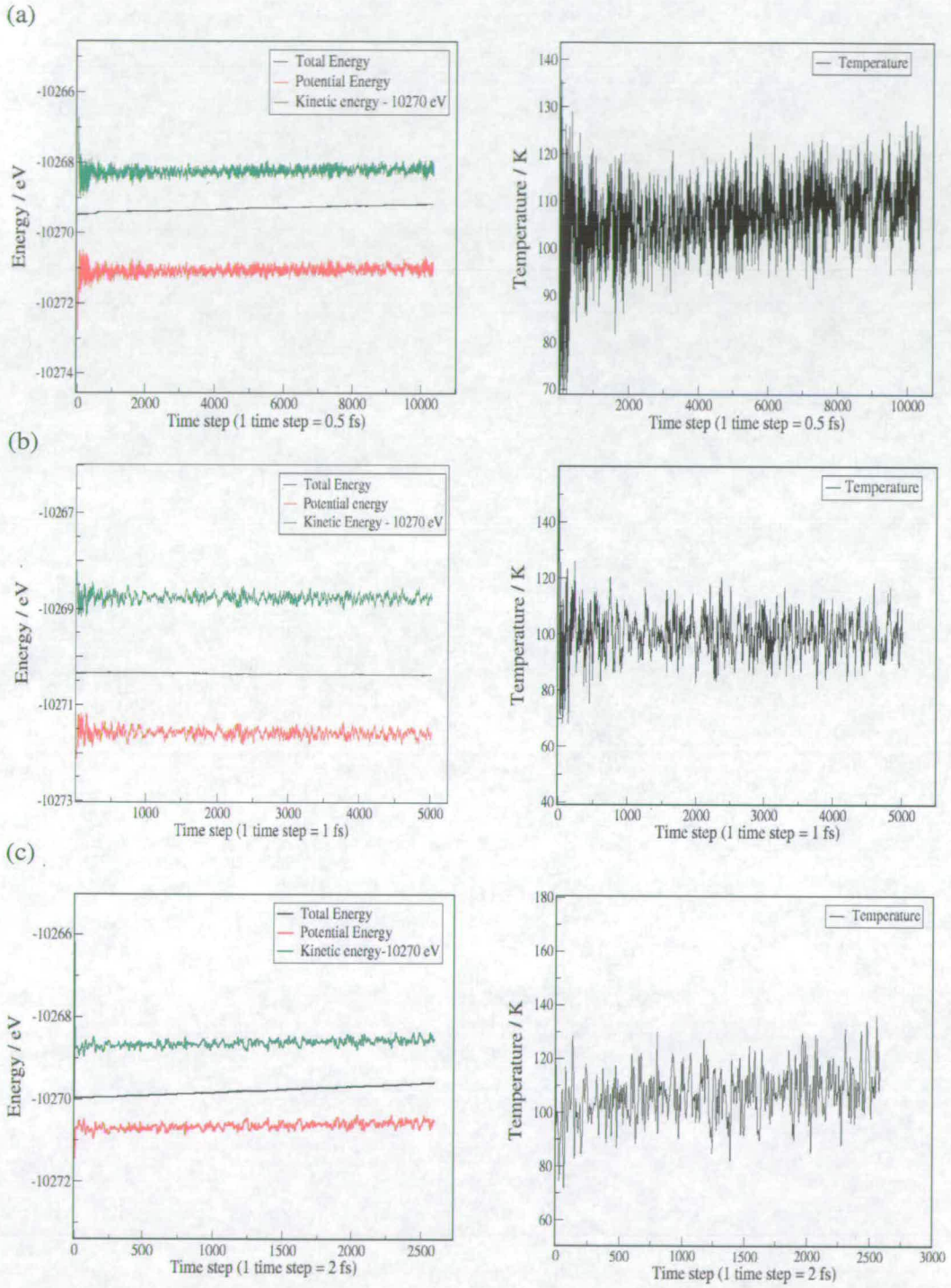


Figure 7.3: The instantaneous energies and temperatures for the (a) flexible, (b) fixed-bond and (c) rigid ammonia simulations over 5 ps.

The maximum deviations in the total energy curves were found to be ~0.22 eV, ~0.04 eV and ~0.25 eV for the above flexible, bond fixed and rigid ammonia simulations respectively; over 5 ps simulation length after equilibration. The use of a time step greater than 2.0 fs for the rigid ammonia simulation was found to create an instability in the energy and the temperature curves (see Figure 7.4). For a time step of 3 fs, the *average* total energy and temperature (over every 100 fs) shows a clear upward trend, indicating that the maximum gain in time step possible for the ammonia system due to the introduction of molecular constraints is a four-fold increase (*i.e.* a time step no greater than 2.0 fs).

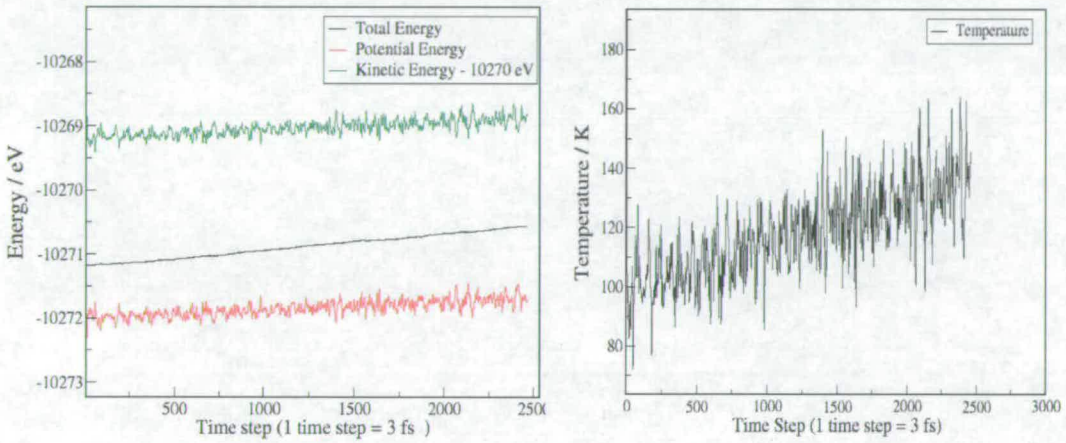


Figure 7.4: The instantaneous energies and temperature for the rigid ammonia simulation over 5 ps, with a time step increment of 3 fs.

It is important to note that no thermostat was used in these calculations. The energy drifts seen here would be trivially corrected by the use of thermostat, which is the normal procedure in ab-initio MD.

7.3.2 Structural properties of the ammonia crystal

The structure of a molecular system in a dynamical simulation is best described by a set of distribution functions for the positions and orientations of atoms in (the same or different) molecules. The simplest and most commonly used distribution functions are the atom-atom radial distribution functions, g(r), and the angular distribution functions, g(θ)

[Allen, 1987]. In this work, we have calculated the nitrogen-nitrogen, hydrogen-nitrogen, hydrogen-hydrogen radial distribution functions (RDF), and the angular distribution functions for the nitrogen atoms and the hydrogen bonds, from the atomic trajectories generated by the ab-initio simulations. The results for the flexible, fixed-bond and rigid molecule simulations are discussed in turn in the following sections.

The N-N radial distribution and the N-N-N angular distribution functions

The nitrogen-nitrogen radial distribution and its angular distribution in the first and second co-ordination shells, form a complete representation of the crystal packing arrangement in solid ammonia phase-I at temperature ~110 K and ~1kbar pressure (see Figure 7.5 and Figure 7.6).

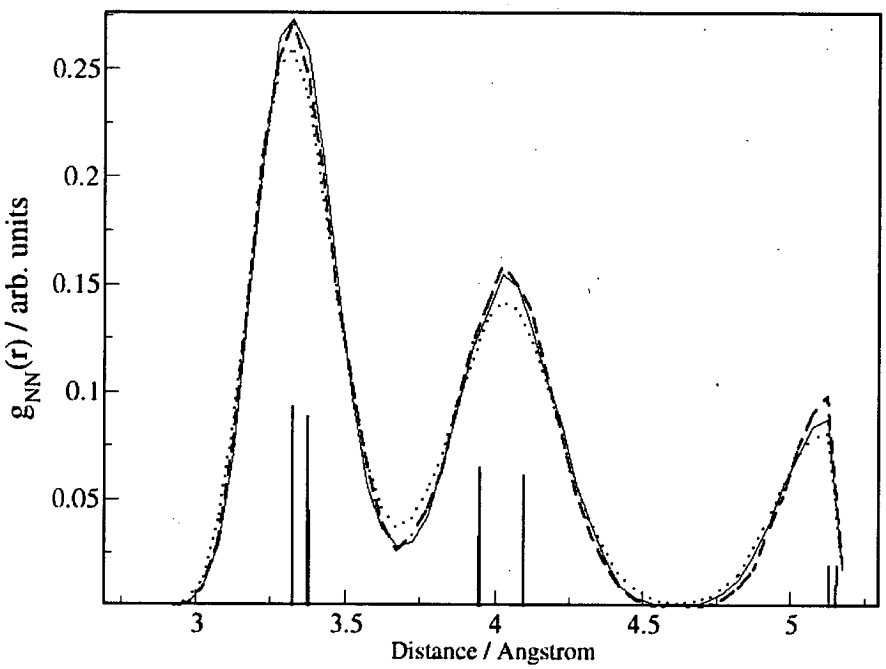


Figure 7.5: Nitrogen-nitrogen radial distribution functions from MD at ~100 K. The solid, dashed and dotted lines show the flexible, fixed-bond and rigid molecule results, respectively. The red bars represent the experimental values and the blue bars represent the ab-initio optimised structure values.

The peak positions of the nitrogen-nitrogen radial distribution functions from the three different simulations were found to be within 0.02 % of each other, indicating the fact that the geometrical effects of imposing molecular constraints on the system are minimal. Figure 7.5 shows the first peak at 3.325(1) Å, the second peak at 4.025(1) Å and the third peak at 5.125(1) Å. The N-N peak positions from the MD simulations are in very good agreement with the experimental values at 160 K (maximum shift is found as 1.9%, for the second peak) [Boese, 1997]. The MD peak positions also show very small shifts from the ab-initio optimised structure values at 0 K. The first peak is very close to the first N-N distance in the optimised structure at 0 K, the second and third peaks, however, are shifted to lower distances (by 1.75% and by 0.6%, respectively).

The angular distribution of ammonia molecules in the lattice were investigated for both the first nearest neighbour distance [$g_{NNN}(\Theta_1)$, Figure 7.6(a)] and second nearest neighbour distance [$g_{NNN}(\Theta_2)$, Figure 7.6(b)] with respect to the nitrogen atom, which is taken as the reference centre for each ammonia molecule. Both distributions show five distinct peaks, suggesting that the nearest neighbour nitrogen atoms are distributed around five distinct angles: 60.0(0.7)°, 74.2(0.5)°, 101.0(1.0)°, 117.0(1.5)° and 133.2(0.7)° in the first co-ordination shell, and 48.5(0.5)°, 60.2(0.3)°, 79.0(0.5)°, 108.1(1.2)° and 117.5(1.5)° in the second co-ordination shell (Figure 7.6). The nitrogen orientation along the 60° orientation are found exceptionally strong in both first and second co-ordination shells. All the other orientation angles (in both first and second co-ordination shells) show a slight shift from their corresponding experimental [Boese, 1997] and ab-initio optimised structure values [Morrison, 2003]. The second peak at first co-ordination shell shows the maximum shift from the experimental (by 3.25%) and ab-initio optimised structure (by 3.00%) values.

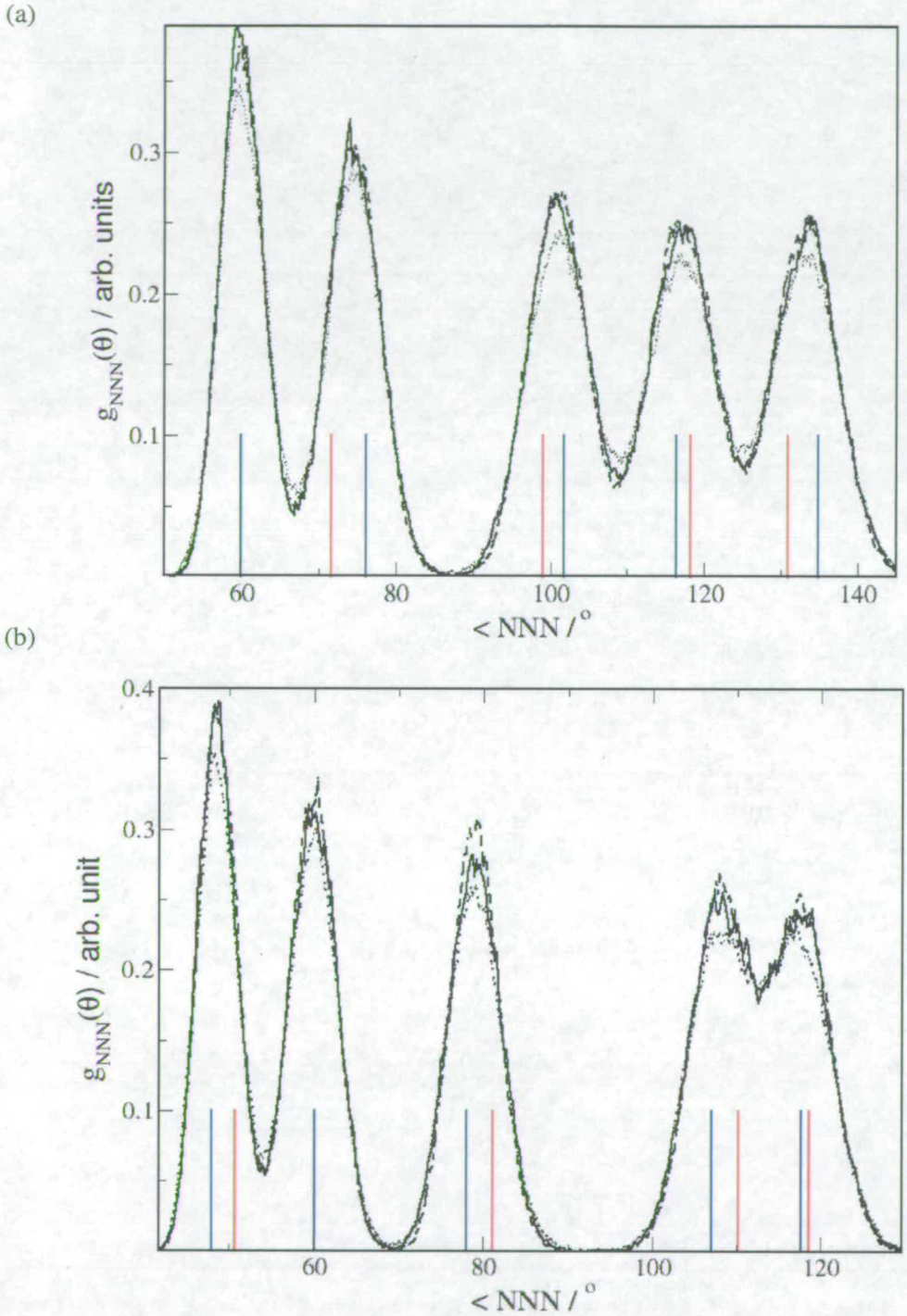


Figure 7.6: The angular distribution functions of nitrogen atoms at (a) first nearest neighbour and (b) second nearest neighbour. The solid, dashed and dotted lines show the flexible, fixed-bond and the rigid molecule results, respectively. The red bars represent the experimental values and the blue bars represent the ab-initio optimised structure values.

There is, of course, an obvious difference between the three data sets: the experimental structure refers to 160 K [Boese,1997], whilst the ab-initio optimised structure is at 0 K, and the ab-initio MD structure distribution has an averaged ensemble temperature of ~110K. However, the results from all three cases are found to be very close to each other for the case of nitrogen-nitrogen radial and angular distribution. This reflects the fact that the crystal packing in solid ammonia phase I structure is quite stable, even with lattice vibrations and an increase in temperature and pressure, particularly in relation to the first co-ordination shell and along the 60^0 orientation. We can also postulate that the minor discrepancies in the RDF's between the experimental structure (at 160 K) and the MD data (at ~110 K) can be reasonably attributed to the calculated pressure rise in the MD simulation of ~1kbar (Figure 7.2).

Beside these agreements with the experimental and the ab-initio optimised structure, the agreement within the three simulations with different level of constraints is very close. Figure 7.5 and Figure 7.6(a,b) clearly show that there is almost no variation in crystal packing between the constrained and unconstrained structures. This reflects our assumption that constraining the fast N-H vibration and internal geometry of the ammonia molecule does not affect the crystal packing of the system, nor does it appear to disrupt the nature of the hydrogen bonding network.

The N-H radial distribution function

The first peak at 1.025 Å in the nitrogen-hydrogen radial distribution curve (see Figure 7.7) relates to the internal geometry of the ammonia molecule. The peak is single valued in the case of the fixed-bond and rigid molecule simulations, thus assuring that the N-H distance is properly constrained in these two calculations. Although this distance is of course unconstrained in the flexible molecule MD simulation, the first peak is very sharp, around the value of the constrained distance (see Figure 7.7); an indication that constraining the fast N-H vibration will have little effect on the overall system dynamics. The second peak at 2.35 Å overlaps almost exactly in all three simulations and corresponds to the

hydrogen bond distance in the static solid structure. The high degree of coincidence for the three simulations recorded for the second peak is again testament to the fact that the constraint does not have any observed effect on the hydrogen bond distance, or on the overall structure of the ammonia solid at finite temperature.

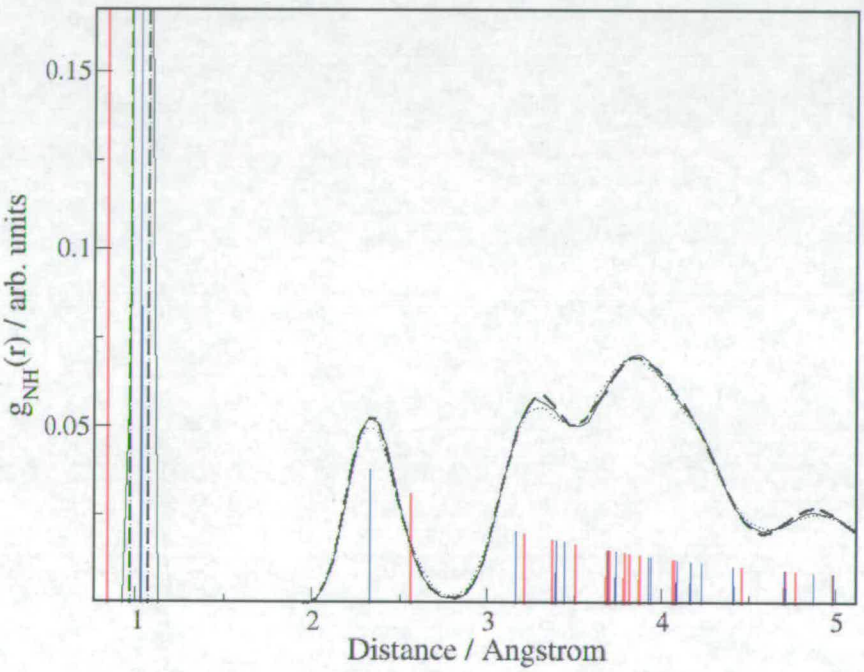


Figure 7.7 : The nitrogen-hydrogen radial distribution functions. The solid, dashed and dotted curve show the flexible, fixed-bond and the rigid molecule results, respectively. The red bars represent the experimental values and the blue bars represent the ab-initio optimised structure values.

The H-H radial distribution function

The first peak at 1.675 Å in the hydrogen-hydrogen radial distribution curve (see Figure 7.8) represents the non-bonded hydrogen-hydrogen distance within the ammonia molecule. A single valued peak in the case of the rigid molecule simulation thus reflects that the H-N-H intra-molecular angle is properly constrained. The second peak distance corresponds to the non-bonded distance between hydrogen atoms of two neighbouring hydrogen-bonded ammonia molecules. The fact that the second (and subsequent peaks) in

the H-H radial distribution curves from the flexible, fixed-bond and rigid molecule simulations almost superimpose on each other, reflects once again that constraining the fast moving intra-molecular bonds and angles in the ammonia molecules does not affect the overall crystal structure.

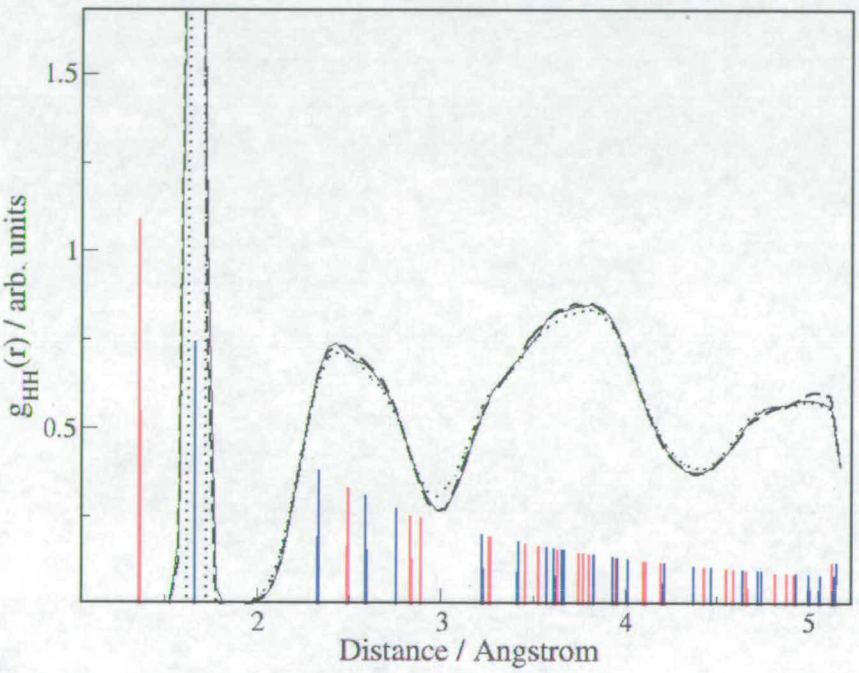


Figure 7.8: The hydrogen-hydrogen radial distribution functions. The solid, dashed and dotted curve show the flexible, fixed-bond and the rigid molecule results, respectively. The red bars represent the experimental values and the blue bars represent the ab-initio optimised structure values.

Angular distribution function of the N-H..N hydrogen bond

The hydrogen-bond angle (*i.e.* N-H---N angular distribution, Figure 7.9), with peak value at 160.0(1.1)°, are distributed similarly in the constrained and unconstrained cases. This, as well as the second peak of the N-H radial distribution function, is again consistent with our hypothesis that the hydrogen-bonded network in the ammonia crystal is not effected by the introduction of rigidity in the molecular structure. This is a very important observation for a molecular crystal from a quantum mechanical level of calculation.

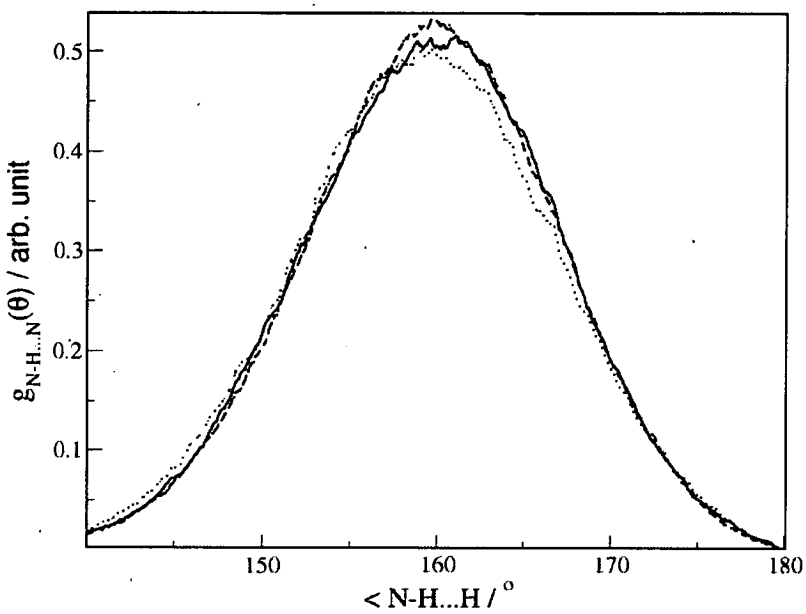


Figure 7.9 : The angular distribution of the hydrogen-bond angle. The solid, dashed and dotted curves show the flexible, fixed-bond and the rigid molecule results, respectively.

Since the ab-initio MD simulation is based on an electronic structure calculation, the implication of any rigidity in the molecular skeleton *could possibly* affect the electronic configuration and hence the system's dynamics (which is dominated by the dynamics of the hydrogen bonds) directly. If the N-H intra-molecular bonds and N---H inter-molecular hydrogen bonds are linked in such a way, the ab initio dynamical simulation should have picked it up. However a very close match is observed between the N---H radial distribution curves (second peak in Figure 7.7) and the N-H---N angular distribution curves (Figure 7.9) from the constrained and un-constrained cases, demonstrating that no such links exist. This also confirms our previous view that the hydrogen bonding in ammonia is largely electrostatic (see Section 3.4, Chapter 3).

We can summarise the findings of our variable constraint MD simulations in the following Table. The static structure from X-ray and neutron diffraction experiment as well as the calculated geometry at 0 K is presented for basic comparison.

Table 7.1: The Atom-Atom radial distributions compared with optimised ab-initio and experimental structures

Atom Group (Å, °)	Peak positions at ~110 K (Flexible, Fixed-bond and Rigid simulations)	[a]Optimised Structure at 0 K	[b]X-ray structure at 160 K	[c]Neutron Diffraction Structure at 77 K
N-N	3.325(1)	3.322	3.375	3.342(8)
	4.025(1)	4.097	3.949	-[d]
	5.125(1)	5.155	5.131	-
N-H	1.025	1.029	0.838	1.039(9)
	2.350(25)	2.330	2.564	2.393(9)
H-H	1.675	1.663	1.364	-
	2.450(25)	2.328	2.490	-
N-H---N	160.0(1.1)	161.8	164.1	160.7(6)

[a] [Table 3.1, Chapter 3], [b] [Boese, 1997], [c] [Hewat, 1979] , [d] Not reported

The first peaks of the N-H and H-H distributions simply represent the molecular geometry of the ammonia molecule; whilst their second peaks, the N-H---N angular distributions and the N-N distribution summarise the inter-molecular interactions in the crystal lattice (Table 7.1). The geometry information from the three different simulations (constrained and unconstrained) have been listed in a single column (2nd column, Table 7.1), as their peak positions shows almost no change from each other. The peak positions for the molecular geometry (*i.e.* the first peaks of N-H and H-H distributions) from the three simulations were found to match exactly; and those corresponding to the inter-molecular distribution were found to match within 1%.

Table 7.1 show that the peak positions from the dynamical simulations are in close agreement with the ab-initio optimised structure at 0 K (see 3rd column, Table 7.1). This reflects the fact that the crystal geometry of the dynamic system (constrained and

unconstrained) at ~110 K does not change much from the static structure at 0 K due to the temperature rise, and the 0 K structure serves well as a starting equilibrium structure for the MD simulation.

However, the peak positions, except for the case of N-N distribution, from the simulations differ significantly from the X-ray diffraction structure at 160 K (see 4th column, Table 7.1) [Boese, 1997]. This is, most likely, due to the difficulty in locating reliable hydrogen positions in an X-ray diffraction experiment, where distances relate to the electron density, rather then the nuclear positions. Thus the hydrogen atom, with its sole electron, is always located experimentally with a degree of uncertainty, and usually close to the atom to which it is bonded.

The N-H and H-H first peaks are out by *ca.* 22% from the corresponding X-ray structure value (see Table 7.1). Due to the electro-negativity effect, the valence electron of the hydrogen is shifted towards the highly electro-negative nitrogen atom. As X-rays counts the distance from the electron-density maxima points rather than the nuclear positions, X-ray experiment typically sees a bond-length less than 1 Å whereas alternative methods based on nuclear positions (*e.g.* neutron diffraction or ab-initio calculations) find a bond-length (*e.g.* N-H bond-length in Table 7.1) greater than 1 Å. Thus we see, the peak positions are in better agreement with the neutron powder-diffraction study (see Table 7.1) by Hewat *et. el.* at 77 K [Hewat, 1979].

A less pronounced effect in the hydrogen-hydrogen (and even less pronounced in nitrogen-hydrogen) radial distribution curves (Figures 7.8 and 7.7) is that the second and subsequent peak heights and depths are slightly decreased in the rigid molecule simulation compared to that obtained in the flexible molecule case, showing that the rigid molecule approximation causes an overall decrease in the order of the crystal; a phenomenon also observed by Allesh *et al.* in the constrained simulation of liquid water [Allesh, 2004]. However this effect is not obvious in our fixed-bond (partially constrained) simulation and so can be attributed to the bond angle constraint.

7.3.3 Dynamical properties of the ammonia crystal

The effect of the constraints on the dynamical behaviour of the system can also be observed from the normal mode analysis of the crystal dynamics. At finite temperature the dynamics of each atom in a crystal is not random, but rather are highly correlated to one another. The motion of each atom in a crystal can be described fully in terms of a set of 3N eigenvectors of the solid system. Although these eigenvectors are quite hard to investigate, each is associated with an experimentally observable eigen-mode or normal-mode frequency. The dynamical properties of any molecular crystal can therefore be represented by its normal mode frequency distribution (also known as the Phonon Density of States), or the vibrational spectra of the crystal, which will be a characteristic for that system only.

Group theory analysis of ammonia normal modes

The structure and symmetry of the gas [Helminger, 1971] and solid [Hewat, 1979; Boese, 1997] phases for ammonia are well known. A full group theory analysis for the phase I structure, detailing the total number of normal modes, their respective symmetry type and full breakdown into translational, rotational and molecular modes for the primitive unit cell is also available from the literature [Reding, 1951].

The gas phase dynamics of a single ammonia molecule shows four (2A + 2E) distinct normal modes of vibrations, labelled $\nu_3, \nu_1, \nu_4, \nu_2$. These are the asymmetric (E) and symmetric stretching (A) modes of the N-H bonds at 3450 cm^{-1} and 3336 cm^{-1}, respectively; and the asymmetric (E) and symmetric (A) bending modes of the H-N-H bond-angle at 1627.5 cm^{-1} and 968.1 cm^{-1}, respectively. These four modes also account for the main features observed in the solid state vibrational spectra, with additional features arising from the effect of the crystalline field on the molecules and from various couplings of the motions of the molecules.

Solid ammonia Phase I crystallises in the primitive cubic group $P2_13$ (T^4) and has four molecules of C_{3v} symmetry in the unit cell on special positions of C_3 symmetry. Each molecule is oriented with its principal axis along one of the {111} Miller index directions, *i.e.* the body diagonal of the cube. In the solid ammonia cubic primitive cell there are therefore 48 normal modes (three times the number of atoms, 16, in the super-cell), out of which 24 modes, known as molecular modes, are split into the four fundamental bands [2A + 2E] corresponding to those mentioned above, coupled with lattice motions [F] (as shown in Table 7.2).

Table 7.2: Group theoretical prescription of the molecular modes of vibration in solid ammonia.

Normal Mode Frequency	Symmetry Coordinates	Symmetry			Total number of modes
		Isolated molecule, C_{3v}	Lattice Translation, T	Crystal	
ν_3	Asymmetric stretch	E		E + 2F	2 + 6
ν_1	Symmetric stretch	A_1	F	A + F	1 + 3
ν_4	Asymmetric bending	E		E + 2F	2 + 6
ν_2	Symmetric bending	A_2		A + F	1 + 3

The remaining 24 modes arise due to inter-molecular interaction in the lattice, and hence known as lattice modes. Three out of the 24 lattice modes are acoustic modes. At gamma point these correspond to rigid translation of the crystal and have zero frequency [Ashcroft, 1976]. The remaining 21 zone centre normal modes are classified as: nine lattice translational normal modes [A + E + 2F] and 12 lattice librational modes [A + E + 3F]. Taking degeneracy into account means that only nine peaks (*i.e.* four translational and five librational) will be visible in the vibrational spectrum of the ammonia crystal, as summarised in Table 7.3.

Table 7.3: Group theoretical prescription of the lattice modes of vibration in solid ammonia [Anderson, 1965].

Type of free molecule motion	Symmetry			Total number of modes
	Molecular,C_{3v}	Site, C_3	Crystal	
Translational	A_1	A }	A + E + 2F	1 + 2 + 6
	E	E }		
Rotation	A_2	A	A + F	1 + 3
	E	E	E + 2F	2 + 6

Normal mode calculations

The spectral distribution of a crystal can be obtained from three main sources:

a) spectroscopic investigation using Raman, Infra-red or neutron scatterings;

b) the dynamical matrix or force constant calculation, using empirical (molecular mechanics force field models), semi-empirical (AM1, PM3) or first-principles (Finite Displacement Method, Linear Response method) methods;

c) and finally from the auto-correlation functions from classical and quantum mechanical molecular dynamics simulations.

Out of all these methods, the ab-initio MD calculation provides a unique opportunity for an authentic investigation of the effect of "constraints in a molecule" on the lattice dynamics. The experimental techniques usually offer very little flexibility in their set-up, compared to computational methods. The dynamical matrix calculation imposes the harmonic approximation on the lattice dynamics, which can introduce significant error for hydrogen-bonded systems. Although the MD technique is free from these limitations, it generally requires calculation on a larger system for a very long time to achieve reasonable statistical sampling. DFT-based quantum mechanical simulation is obviously free from parameterisation of potentials as in classical MD, and with the continuous advance in techniques and computer power it is now possible to generate sufficient sampling in reasonably short CPU time. However it is still a hard task to simulate larger systems for a longer time with ab-initio MD, and so every effort to improve the speed of calculation is highly desired.

The basic quantity for calculating the phonon distribution of a crystal in an MD simulation is the simplified form of the velocity time auto-correlation function [Allen, 1987]

eqn 7.1: $$C_{vv}(t) = \langle V(t)V(0)\rangle = \frac{1}{t_{max}} \sum_{t_0=1}^{t_{max}} V(t_0).V(t_0 + t),$$

where V is the ionic velocity. Once the auto-correlation function [$C_{vv}(t)$] of the system has been calculated for a sufficiently large duration of time, the Fourier transform of the data [$C_{vv}(t)$] produces the occupied phonon density of states (Phonon DOS x Boltzmann factor) *i.e.* the desired vibrational spectra of the system. However, the following points need particular consideration.

A direct consequence of using a larger cell size (the 2x2x2 super-cell) is that our generated velocity auto-correlation function [$C_{vv}(t)$] produces the normal mode distribution at all available k points in the first Brillouin zone, which in this case is the obvious zone centre k point (k = 1,1,1; frequently known as gamma point) plus the k points at zone boundaries [k={1/2,0,0}, {1/2,1/2,0}, (1/2,1/2,1/2)]. Information from all the k-points can be used to draw the phonon dispersion relation of solid ammonia from the *ab-initio* MD route. Figure 7.10 shows a typical phonon-dispersion relation along the high symmetry directions in diamond [Ashcroft, 1976]. The lower branches in Figure 7.10 that crosses the origin are

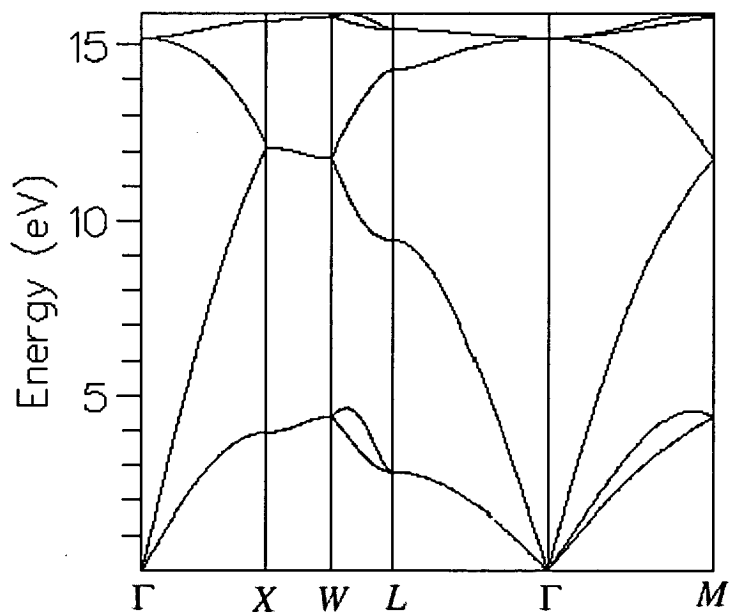


Figure 7.10: Schematic phonon dispersion curve along the high-symmetry directions in diamond.

called acoustic phonon branch, as their frequencies become small at the long wavelengths (small k) and correspond to sound waves in the lattice. All the upper branches, known as optical phonon branch, always have some minimum frequency even at very small k, and in ionic crystals they can interact with electro-magnetic waves. In practice, frequencies at each of the k-point along the high symmetry directions, *i.e.* at a very large number of k-points are needed to draw a dispersion curve (see Figure 7.10). In present case, this means that a very large supercell calculation and consequently a very intensive calculation is required to extract the dispersion relation using *ab-initio* MD method. Alternative first principles techniques, *i.e.* force constant methods are thus generally preferred to study the phonon dispersions of solids.

Nevertheless, the vibrational frequencies at zone centre k point (gamma point, k=1,1,1) are also very crucial in testing the properties of a particular substance, and can be used in order to make a direct comparison with group theoretical predictions and other reported experimental works. In this work, in order to extract the normal modes at gamma point, we have thus projected the velocity data onto the gamma point before calculating the normal modes for each of the three simulations of solid ammonia.

As our simulation runs are finite, the truncation of $C_{vv}(t)$ after a finite time (t_{max}) poses further difficulty in the Fourier transform. Spurious features are obtained in the vibrational spectrum, which can obscure the true feature of the spectrum. In particular, the truncation causes rapidly varying side lobes around a peak, and loss of resolution [Allen, 1987]. In order to avoid such problem, we have used 'windowing functions', which are weighting functions applied to $C_{vv}(t)$ to reduce the order of discontinuity at the truncation point t_{max}. A detail comparison of different windowing functions is presented by Harris [Harris, 1978]. In this work we have used Blackman window, which means that each value of $C_{vv}(t)$ was multiplied by the windowing function before carrying out the Fourier transformation.

eqn 7.2: $$C_{vv}(t) \rightarrow C_{vv}(t)W(t) \quad \text{where}$$

eqn 7.3: $$W(t) = 0.42 - 0.5\cos\left(\frac{\pi t}{t_{max}}\right) + 0.08\cos\left(\frac{2\pi t}{t_{max}}\right).$$

The final result is shown in Figure 7.11. Note that peak *heights* and *widths* are an arbitrary function of the sampling time. The important feature of this figure is the peak positions.

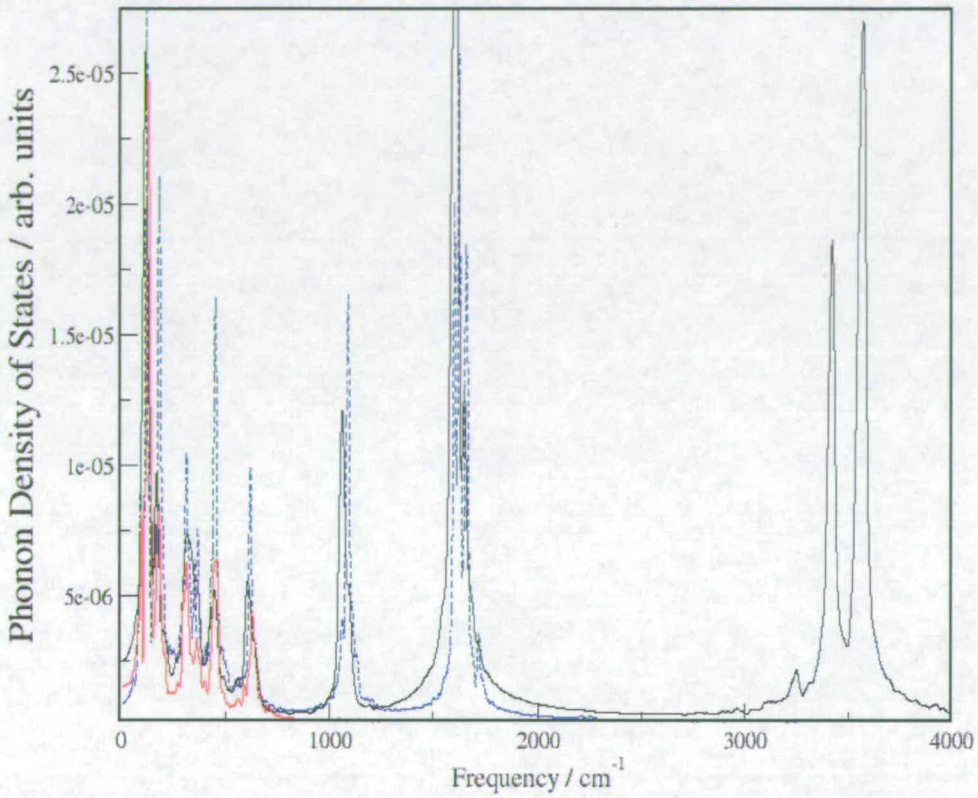


Figure 7.11: The normal mode frequencies of solid ammonia I, at gamma point calculated from ab initio MD. The black, blue and red lines show the flexible, fixed-bond and the rigid ammonia calculation, respectively.

Molecular modes of the ammonia crystal at gamma point

The frequencies above 800 cm^{-1} correspond to the molecular modes of motion, whereas those below 800 cm^{-1} are the lattice vibrational modes. The rigid molecule simulation obviously shows no peaks in the molecular mode region as all the internal motions in each molecule have been frozen using rigid bonds and angles. Similarly the bond stretching modes are absent in the fixed-bond simulation. The flexible ammonia simulation clearly shows all four molecular modes (Figure 7.11).

The solid state and gas phase experimental values, along with our calculated values of the molecular modes are presented in Table 7.4 for direct comparison. Our calculated normal modes are reasonably close to the experimental findings: the stretching modes are within 6.5% and the bending modes are within 2.0% of the experimental values quoted by Binbrek and Anderson [Binbrek, 1972].

Table 7.4: Molecular modes of vibration for solid and gas phase ammonia

Assigned Motion	Calculated Frequencies		[a]Raman spectra at 80 K	[b]Gas Phase IR spectra	Gas Phase[c] Calculation
	Flexible	Fixed Bond			
Asymmetric stretch	3555(10)	—	3369.5(6.5)	3450	3584.7
Symmetric stretch	3415(15)	—	3204(36)	3336	3463.4
Asymmetric bending	1645(5)	1655(5)	1679(8)	1627.5	1751.1
Symmetric bending	1095(8)	1084(10)	1072(14)	968.1	1124.7

[a] [Binberk, 1972], [b] [Reding, 1951], [c] this work, b3lyp/6-311G*

Although the frequency shifts are comparatively small between the gas and the crystal phase spectra, the experimental study shows that the stretching mode frequencies decrease while the bending mode frequencies increase due to the effect of crystallisation (see 4th and 5th columns of Table 7.4).

In order to study the effect of the hydrogen bond network on the internal modes, a hybrid DFT calculation using the b3lyp functional and 6-311G* basis set was carried out using GAUSSIAN'98 [Frisch, 1998]. The result is shown in Table 7.4. In line with the experimental observation, our calculations also show that the ammonia stretching frequencies decreases upon crystal formation. However the b3lyp/6-311G* calculation for the isolated ammonia molecule shows higher values for the bending modes. This is not unexpected, as the isolated molecule calculations are, in general, reported to produce higher values for the internal modes due to inadequate modelling of the electron correlation effect [Rauhut, 1995]. This is fairly well documented for isolated molecule calculations at the ab-

initio SCF level (for both ab-initio and hybrid DFT methods) and has resulted in the use of scaling factors to reduce the calculated values of molecular modes [Baker, 1998].

On the other hand, the pure DFT functionals used in the solid state calculation includes a more complete description of the total electron correlation, although in a somewhat empirical fashion, than the pure ab-initio or hybrid DFT methods [Baker, 1998]. We should also consider that the ab-initio SCF level calculation is a force constant calculation; it misses out the anharmonic effect, which usually lowers the normal mode frequencies [Rauhut, 1995].

The most important aspect of this calculation is that the bending mode frequencies from the fixed bond calculation match those from the flexible ammonia simulations within their calculation error (see Table 7.4). This clearly demonstrates that there is very little or no coupling between the N-H stretching and the H-N-H bending motions from an ab-initio perspective. Eliminating the N-H stretch in the fixed-bond simulation does not shift the H-N-H bending mode. Both the flexible and the fixed-bond simulations show the bending mode at the same value within the limit of calculation uncertainty.

Lattice vibrations of the ammonia crystal at gamma point

Sampling the velocity data for a sufficiently large duration of time is particularly crucial for the correct prediction of the lattice modes. We found at least 7 ps of MD data (after equilibration) is needed in order to to produce a good spectrum of lattice modes. Because of the very short time step (0.5 fs) required for the flexible molecule simulation, we only managed to collate 5 ps of data during the course of this work, which in turn incorporated a larger uncertainty in the peak positions. This is why upon comparison much broader peaks were observed for the lattice modes of flexible ammonia, particularly in the translational region. Figure 7.12, shows the extended view of the lattice modes obtained from the three different simulations.

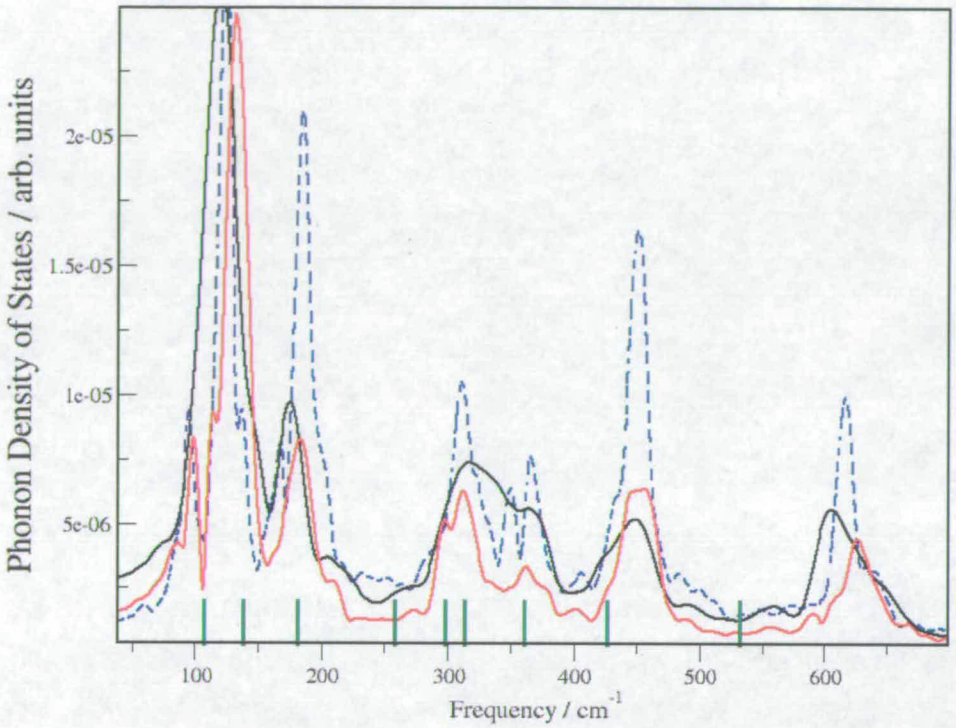


Figure 7.12: Expanded view of the low frequency region of Figure 7.11. The vibrational lattice mode of solid ammonia I at gamma point. The black, blue and red curve show the flexible, fixed-bond and the rigid ammonia calculation, respectively. The green bars represent the experimental results from Binbrek and Anderson.

Our results are compared with the most illustrated and cited work on the ammonia vibrational spectrum by Binbrek and Anderson [Binbrek, 1972] obtained using both IR and Raman spectroscopy (see Table 7.5). Both works show good agreement with the group theory predictions. We saw the nine total distinct modes expected in the lattice region, well separated in terms of four translational and five librational modes. There is also a good agreement between most of our calculated frequencies and the experimental result. The calculated lattice modes as well as their experimental counter-parts are listed in Table 7.5 for basic comparison.

Table 7.5: The lattice modes of phase I ammonia crystal

	[a] Assigned Symmetry	[b] First principles MD values (cm^{-1})			[a] Raman frequencies at 80 K [cm^{-1}]	[a] Infra-red frequencies at 107 K [cm^{-1}]
		Flexible (5 ps) [cm^{-1}]	Fixed-bond (7 ps) [cm^{-1}]	Rigid (7 ps) [cm^{-1}]		
Translational	A	100(5)	97(3)	99(3)	107	—
	E	120(2)	122(1)	127(4)	138	—
	F	145(5)	136(5)	140(3)	—	138
	F	182(9)	179(6)	179(5)	—	181
Librational	F (?)	—			260	258
	A	290 – 340	290(6)	296(4)	298	—
	E	290 – 340	310(10)	310(3)	(313 at 18 K)	—
	F	350(2), 372(6)	349(3), 365(5)	340 – 370	358	361
	F	445(5)	435 – 460	445(5), 460(5)	426	—
	F	612(8)	616(4)	622(8)	—	532

[a][Binbrek, 1972], [b][This work].

However some significant issues arise from our calculations. A very crucial disagreement is that all of our three calculations completely fail to assign a lattice mode at 260 cm^{-1}, which was reported in both experiments but as a very weak signal [Binbrek, 1972]. Another disagreement follows from the Binbrek and Anderson's observation of a doublet at 360 cm^{-1} (F), which they ignored to be a possible LO-TO splitting of 360 cm^{-1} mode. Rather they have suggested that their observed mode at 426 cm^{-1} is the LO branch of the 360 cm^{-1} (F) mode. In the Fourier analysis of ab initio supercell simulation, although we expect to see the LO-TO splitting, we do not expect to see any doublet due to combination or overtone band. However doublets are observed for both 352 cm^{-1} and 432 cm^{-1}. This, as well as the fact that we could not observe one of the nine lattice frequencies at 260 cm^{-1} predicted by the experiments, points us towards considering that both 352 cm^{-1} and 432 cm^{-1} could be separate fundamental of F modes, which accounts for the full 21 lattice modes at gamma point. Finally, the last lattice mode is found at a much higher value, 617 cm^{-1}, than the experimental prediction at 530 cm^{-1}. A more detailed comparison with the experimental works will follow in the next chapter.

However, the main feature of these results is that the lattice mode frequencies calculated from flexible, partially constrained and fully constrained ammonia systems are just the same (see Figure 7.12 and Table 7.5). We have also calculated the lattice vibrations at the Brillouin zone boundaries. They all agree with our central argument that the lattice dynamics of ammonia are unaffected by the intra-molecular constraints.

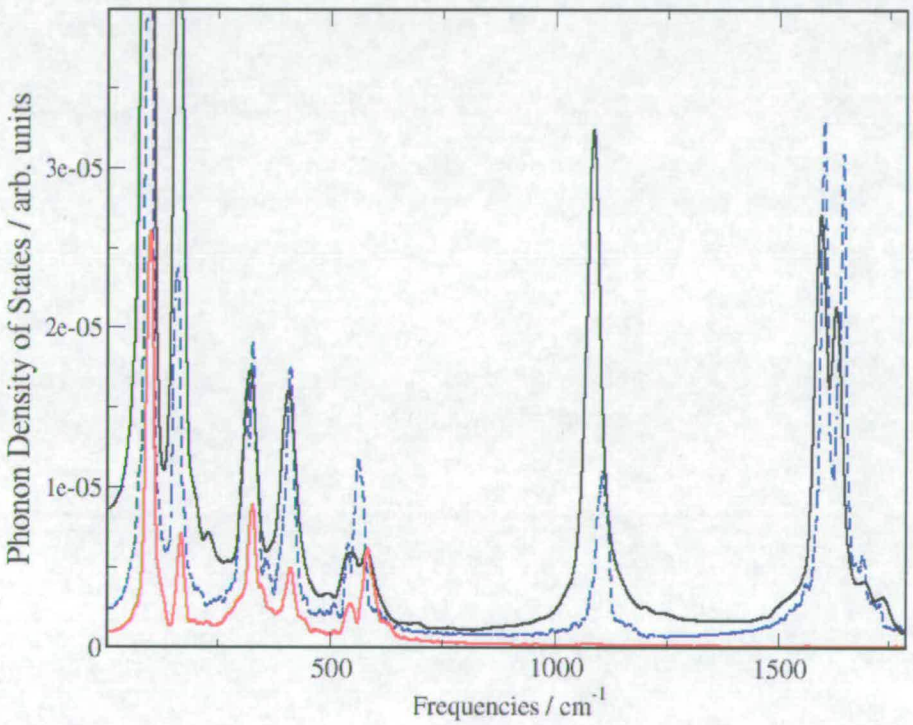


Figure 7.13: The vibrational lattice mode of solid ammonia I at k = (1/2,1/2,1/2). The solid, dashed and dotted curve show the flexible, fixed-bond and the rigid ammonia calculation, respectively. (The details of how this is done are discussed in the next chapter).

Indeed, we see a better agreement between constrained and un-constrained dynamics at the k = (1/2, 1/2, 1/2) zone boundary. For clarity of comparison lattice frequencies are shown only up to 1800 cm^{-1}, at k = (1/2, 1/2, 1/2) in Figure 7.13.

7.4 Conclusions

The main conclusion of this chapter is that the structure and the dynamics of the ammonia phase I crystal is unaffected due to the application of molecular constraints. As the crystal lattice of the ammonia molecule is mainly dominated by hydrogen bonding, it implies that there is no significant link between the internal flexibility of the molecules and their associated hydrogen bond network. This removes the possibility that the electron density around the N-H bond could be coupled to its associated hydrogen bond during the stretching motion. This finding is in line with our previous result that the nature of the hydrogen bonding is electrostatic for this system. Thus our implementation of the constrained MD algorithm has given us unique insight to the chemistry of solid ammonia, which would have been impossible to obtain by any other route.

This work thus clearly demonstrates the strength of constrained dynamics in ab-initio MD simulation. Where the CPU time scale of an ab-initio simulation is weeks or months, a four-to-five fold gain in the total computational time using constrained dynamics simulation offers a much better chance to explore bigger systems using first-principles simulations.

7.5 Bibliography

1. Allen, M. P., and D. J. Tildesley, *Computer simulation of liquids*, (Oxford University Press, 1987).
2. Allesch, M., E. Schwegler, F. Gygi and G. Galli, *J. Chem, Phys*, **120**(12), 5192(2004) .
3. Andersen, H. C., *J. Comput. Phys.*, **52**, 24(1983).
4. Anderson, A., and S. H. Walmsley, *Mol. Phys.*, **9**, 1(1965).
5. Arias, T. A., M. C. Payne, J. D. Joannopoulos, *Phys. Rev. B*, **45**(4), 1538(1992).
6. Ashcroft, N. W., and N. D. Mermin, *Solid state physics*, (W.B. Saunders Company, 1976).
7. Baker, J., A. A. Jarzecki and Peter Pulay, *J. Phys. Chem. A*, **102**, 1412(1998).
8. Binbrek, O. S., and A. Anderson, *Chem. Phys. Lett.*, **15**(3), 421(1972).
9. Boese, A. D., *J. Chem. Phys.*, **119**(12), 5965(2003).
10. Boese, M., Yu Antipin and P.R. Mallinson, *J. Phys. Chem. B*, **101**, 5794(1997).
11. de Leeuw, S. W., and M. F. Thorpe, *Phys. Rev. Lett.*, **55**(26), 2879(1985).
12. Ebbing, D. D., *General Chemistry, Second Edition,*(Houghton Mifflin Company, 1987).
13. Eckert, J., R. Mills and S. Satija, *J. Chem. Phys.*, **81**, 6034(1984).
14. Feynman, R.P., *Phys. Rev. B.*, **56**, 340(1939).
15. Fortes, A. D., J. P. Brodholt, I. G. Wood and L. Vocadlo, *J. Chem. Phys.*, **118**(13), 5987 (2003).
16. Francis, G P, and M. C. Payne, *J. Phys.: Cond. Matt.*, **2**, 4395 (1990).
17. Frisch M. J., *Gaussian 98, Revision A.7*, (Gaussian, Inc., Pittsburgh, PA, 1998).
18. Gauthier, M., Ph. Pruzan, J. C. Chervin and J. M. Besson, *Phys. Rev. B.*, **37**, 2102 (1988).
19. Gauthier, M., Ph. Pruzan, J.M. Besson, G.Havel and G.Syfosse, *Physica 139 & 140 B*, **218**(1986).
20. Goyal, P. S., B. A. Dasannacharya, C. L. Thaper and P. K. Iyengar, *Phys. Stat. Sol. B*, **50**, 701(1972).

21. Harris, F.J., *Proc. IEEE*, 66, 51(1978).

22. Helminger, P., F.C. De Lucia and W. Gordy, *J. Mol. Spectroscopy*, **39**, 94(1971).

23. Hewat, A. W., and C Rickel, *acta Cryst. A*, **35**, 569(1979).

24. Kresse, G., and J. Furthmuller, *Phys. Rev. B*, **54**, 11169(1996).

25. Monkhorst, H. J., and J. D. Pack, *Phys. Rev. B*, **13**, 5188 (1976).

26. Morrison, C. A., and M. M. Siddick, *Chem. Eur .J.*, **9**(3), 628(2003).

27. Nye, C. L., and F. D. Medina, *J. Chem. Phys.*, **87**(12), 6890(1987).

28. Nye, C. L., and F. D. Medina, *Phys. Rev. B*, **32**(4), 2510(1985).

29. Payne, M. C., M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, *Rev. Mod. Phys.*, **64**, 1045(1992); *CASTEP, Academic version 4.2,* (licensed under the UKCP-MSI Agreement, 1999).

30. Perdew, J. P., K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, **77**, 3865 (1996).

31. Powell, B. M., G. Dolling, G. S. Pawley and J. W. Leech, *Can. J. Phys.*, **58**, 1703 (1980).

32. Rauhut, G., and Peter Pulay, *J. Phys. Chem*, **99**, 3093(1995).

33. Reding, F. P., and D. F. Hornig, *J. Chem. Phys.*, **19**(5), 594(1951).

34. Reding, F. P., and D. F. Hornig, *J. Chem. Phys.*, **22**(11), 594(1954).

35. Righini, R., N. Neto, S. Califano and S. H. Walmsley, *Chem. Phys.*, **33**, 345(1978).

36. Ryckeart, J. P., G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, **23**, 327(1977).

37. Sarnthein, J., A. Pasquarello and Roberto Car, *Science*, **275**, 1926(1997).

38. Segall, M. D., P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark and M. C. Payne, *J. Phys.: Cond. Matt.*, **14**, 2717(2002).

39. Steiner, T., *Angew. Chem. Int. Ed.*, **41**, 48(2002)

40. Vanderbilt, D., *Phys. Rev. B*, **41**, 7892(1990).

41. VASP, Kresse G., J. Furthmueller, *Comp. Mater. Sci.*, **6**, 15(1996).

42. Yurtseven, H., *Chinese J. Phys.*, **42**(2), 209(2004).

Chapter 8

Phonon Modes of Ammonia,

construction of a new methodology from first-principles MD

8.1 Introduction

In the last chapter we reported on the full vibrational analysis for solid ammonia phase I in order to explain *the effect of constraints* on the dynamical properties of the system. This valuable information was obtained from the Fourier transformed auto-correlation function of the ionic velocities. However, new approaches had to be developed and some careful post-data analysis had to be carried out in order to successfully report on the values of the vibrational frequencies (Figures 7.11, 7.12; Tables 7.4, 7.5; Chapter 7). A fuller report of the methodologies developed are therefore presented in this chapter. Although our principal aim is to gain a clear insight into the ammonia lattice vibrations at the gamma point (irrespective of constraints), this new methodology can be pursued for other molecular materials.

A detailed review of the previous works is presented alongside our findings to assess the strength of the constrained ab-initio MD simulation in correctly reproducing the ammonia normal modes. Once the vibrational frequencies are known, it is also possible to extract the eigenvectors for each of these modes from the positional or velocity data. The

extracted eigenvectors reveal invaluable information on the exact nature of the normal mode vibrations that have not previously been extracted from ab initio molecular dynamics in any system and thus acts to solve some of the discrepancies with previous works raised by our calculation.

8.2 Eigenvalue Extraction

The eigenvalues of lattice vibration (*i.e.* the lattice frequencies) are routinely calculated from the velocity auto-correlation, c_{vv}. For a single atom it is defined as

eqn 8.1:
$$c_{vv}(t) = \frac{\langle \boldsymbol{V}(t+t_0) \,.\, \boldsymbol{V}(t_0)\rangle_{time} - \langle \boldsymbol{V}\rangle^2_{time}}{\langle \boldsymbol{V}^2\rangle_{time} - \langle \boldsymbol{V}\rangle^2_{time}} ,$$

where the averages are defined over a sufficiently large finite time, as

eqn 8.2:
$$\langle \boldsymbol{V}\rangle_{time} = \frac{1}{t_{0max}} \sum_{t_0=1}^{t_{0max}} \boldsymbol{V}(t_0) ,$$

eqn 8.3:
$$\langle \boldsymbol{V}(t+t_0).\boldsymbol{V}(t_0)\rangle_{time} = \frac{1}{t_{0max}} \sum_{t_0=1}^{t_{0max}} \boldsymbol{V}(t+t_0).\boldsymbol{V}(t_0) .$$

When $t_{0\,max}$ is very large, for the ergodic system, the time average for a single atom approaches it's true ensemble average. $c_{vv}(t)$ is also known as the normalised auto-correlation function, and ranges from zero to one. The temporal Fourier transform of the velocity auto-correlation produces the desired spectrum,

eqn 8.4:
$$\hat{c}(w) = 2\Re \int_0^\infty dt\, c_v(t)\, e^{-iwt} ,$$

which is proportional (according to the Boltzmann factor) to the occupied density of normal modes, and is often referred to as the occupied phonon density of states (Phonon DOS) in solids [Allen, 1987]. We have taken the limit from zero to infinity as the classical auto-correlation $c_{vv}(t)$ is an even function of time [Allen, 1987].

The time independent terms in the velocity auto-correlation $c_{vv}(t)$ (eqn 8.1) will be constants of integration in the final spectra produced by eqn 8.4. Since we are only interested in the spread and positions of the peaks of the vibrational spectrum, it is easier to work with the non-normalised functions extracted directly from molecular dynamics, and projected onto a particular velocity component of atom i as

eqn 8.5: $$C_{vv}(t) = \langle \boldsymbol{V}_i(t+t_0) \,.\, \boldsymbol{V}_i(t_0) \rangle_{time} \,.$$

The final spectra will thus be

eqn 8.6: $$\hat{C}_i(w) = 2\Re \int_0^{\infty} dt\, e^{-iwt} \left(\sum_{t_0=1}^{t_{0max}} \boldsymbol{V}_i(t+t_0).\boldsymbol{V}_i(t_0) \right).$$

This will produce the phonon spectrum projected onto a particular atom, say i, at all available wave-vectors permitted by the size of the supercell used in the simulation. The ensemble average can be obtained by summing $\hat{C}_i(\omega)$ for all atoms in the system.

8.2.1 Problem with the standard technique

The spectrum obtained using the above standard method, from a supercell bigger than the unit cell, is a mix of normal mode frequencies from all the k-points compatible with the supercell size. In other words, a supercell calculation produces phonons whose wavelengths ($2\pi/k$) 'fit' the supercell; thus in our simulation of the 2x2x2 ammonia supercell we get phonon modes at the k-points (in units of $2\pi/a$, where a is the lattice constant): (0,0,0), (1/2, 0, 0), (0, 1/2, 0), (0, 0, 1/2), (1/2, 1/2, 0), (1/2, 0, 1/2), (0, 1/2, 1/2), (1/2, 1/2, 1/2).

In order to extract a meaningful set of frequencies, we need to have well-defined peaks disentangled from each other. However, with the contribution from as many as 8 *k*-points the overall picture is very crude. The resultant spectrum is a superposition of eight different sets of normal modes, a total of 3x8x4x4 = 192 modes for the 2x2x2 ammonia supercell. This makes it difficult to assign individual peak values to particular phonon modes.

Note that, if all the phonon eigenvectors were known, a definitive solution would be to extract the frequencies by projecting the atomic velocities onto each individual eigenvectors. However, while the full set of eigenvectors might be known from the symmetry for simple systems (with few atoms in the unit cell), for our NH_3 supercell this is much too complex.

In this work, a theory has been developed in order to project the phonon spectrum onto a particular *k*-point and thereby obtain a much clearer picture of the spectrum. This gives us the added advantage of comparing our result with the experimental spectra at the gamma point, e.g. IR and Raman spectra. Such a comparison with the already available results gives a much deeper insight into the dynamical behaviour of solid ammonia and provides an obvious benchmark against which to test the accuracy of our simulation.

8.2.2 Theory of *k*-point separation

In order to extract the spectrum at a particular *k*-point, we need to project the basic dynamical quantity (*i.e.* the ionic velocities) on to the desired *k*-point. We seek a transformation of the form

eqn 8.7: $$\boldsymbol{V}_i \Rightarrow \boldsymbol{V}_{i,k} \,.$$

To achieve this, it is convenient to label individual atoms (say *i*) in terms of its lattice vector $\boldsymbol{R}$ and basis vector $\boldsymbol{b}$, mathematically

eqn 8.8: $$\boldsymbol{V}_{b,R} \equiv \boldsymbol{V}_i \,.$$

These velocities in a crystal environment are not independent. According to the classical theory of harmonic crystal the atoms, under the influence of a phonon of wave vector $\boldsymbol{k}$ and frequency w, move according to the following pattern,

eqn 8.9: $$\left[\boldsymbol{V}_{b,R}(t)\right]_{k,w} = \boldsymbol{\varepsilon}_{b,k,\omega}\, e^{i(\boldsymbol{k}.\boldsymbol{R}-\omega t)} \,,$$

where $\varepsilon_{b,k,w}$ is the velocity eigenvector of the basis atoms $\boldsymbol{b}$ corresponding to a particular phonon ($\boldsymbol{k}$, w). For a mono-atomic basis this is just a 3D vector, but for a complicated basis like four NH_3 molecules it has 3x4x4 = 48 components and is unknown, although Group Theory may give some indications. The *observed* velocity of the atoms under all the phonon modes would be

eqn 8.10: $$\boldsymbol{V}_{b,R}(t) = \sum_{k} \boldsymbol{V}_{b,k}(t)\, e^{i\boldsymbol{k}.\boldsymbol{R}} \,,$$

where

eqn 8.11: $$\boldsymbol{V}_{b,k}(t) = \sum_{w} \boldsymbol{\varepsilon}_{b,k,\omega}\, e^{-i\omega t} \,.$$

Here $\boldsymbol{V}_{b,k}(t)$ is the velocity component at a particular wave-vector $\boldsymbol{k}$ due to all normal modes. The velocity component $\boldsymbol{V}_{b,k}(t)$ does not depend on the unit cell $\boldsymbol{R}$ in question: the pattern of velocites in each cell is identical, differing only in phase ($e^{i\boldsymbol{k}.\boldsymbol{R}}$). Thus despite the lack of information about $\varepsilon_{b,k,w}$, we can still isolate all the modes at a particular wave vector $\boldsymbol{k}$ by using the identity for the crystal lattice [Ashcroft, 1976],

eqn 8.12: $$\sum_{\{\boldsymbol{R}\}} e^{i\boldsymbol{k}.\boldsymbol{R}} = N\,\delta_{k,0} \,,$$

where N is the total number of unit cells and $\boldsymbol{R}$ is the lattice vector. By multiplying both sides of eqn 8.10 with a phase factor ($e^{-i\boldsymbol{k}'.\boldsymbol{R}}$) of the desired wave vector $\boldsymbol{k}'$ and summing up over all the lattice vectors ($\{\boldsymbol{R}\}$) we get

eqn 8.13: $$\sum_{\{R\}} V_{b,R}(t)\, e^{-ik'.R} = \sum_{\{R\}} \left(\sum_{k} V_{b,k}(t)\, e^{ik.R} \right) e^{-ik'.R} .$$

Since $V_{b,k}(t)$ is the same for all the unit cells, we can re-write this as

eqn 8.14: $$\sum_{\{R\}} V_{b,R}(t)\, e^{-ik'.R} = \sum_{k} V_{b,k}(t) \left(\sum_{\{R\}} e^{i(k-k').R} \right),$$

and using the identity of eqn 8.12 in eqn 8.14, we get the desired transformation (see eqn 8.7) of the velocities at a particular *k*-point.

eqn 8.15: $$\frac{1}{N} \sum_{\{R\}} V_{b,R}(t)\, e^{-ik'.R} = V_{b,k'}(t),$$

where $\boldsymbol{R}$ runs over all the unit cells of the super-cell used. Instead of using the phase independent velocities to calculate the velocity-correlation, we can now use the velocities projected onto a specific $\boldsymbol{k}$ vector (eqn 8.15), which will produce a spectra for the specific $\boldsymbol{k}$- point only. So the vibrational spectrum at a specific *k*-point would be

eqn 8.16: $$\hat{C}_k(w) = 2\Re \int_0^{\infty} dt\, e^{-iwt} \sum_{t_0=1}^{t_{0max}} V_{b,k}(t+t_0).V_{b,k}(t_0) \equiv 2\Re \int_0^{\infty} dt\, e^{-iwt} I_k(t)$$

where $I_k(t) = \sum_{t_0=1}^{t_{0max}} V_{b,k}(t+t_0).V_{b,k}(t_0)$, and $V_{b,k}(t)$ is given by eqn 8.15. In order to obtain a much stronger spectrum we can add up all the spectra from each of the basis atoms in a unit cell according to their mass ratio, *i.e.*

eqn 8.17: $$I_k(w) = \sum_{b} m_b \hat{C}_k(w),$$

where m_b is the mass of atom at basis *b*. Note that this ensures equal contributions to $I_k(t)$ from each particle at t=0. Eqn 8.17 will give a much clearer spectrum at any desired *k*-point

and comparison with experimental work is now possible. Figure 8.1 shows the vibrational spectrum from the raw velocity data (using eqn 8.6) and the spectrum projected at the Brillouin zone centre (using eqn 8.16).

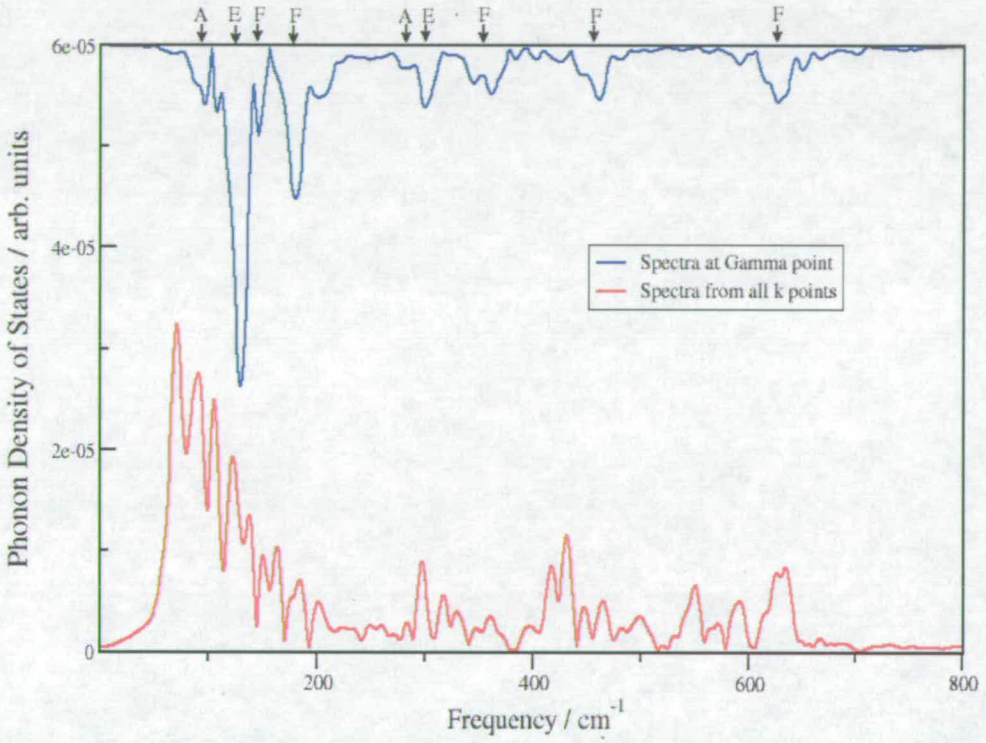


Figure 8.1: The lattice mode frequencies as obtained from fully constrained ammonia simulation (2x2x2 supercell). Note that the gamma point spectrum is shown inverted for the clarity of comparison and symmetry is assigned to it according to Binbrek and Anderson [Binbrek, 1972].

From the 'all *k*-point' spectrum in Figure 8.1, it is not possible to distinguish the nine independent lattice modes (classified as 2A + 2E + 5F, excluding the three acoustic modes. See Section 7.3.3) predicted by the Group Theory analysis of solid ammonia phase I structure. However, after the spectrum is projected onto the gamma point (Figure 8.1) the presence of these nine modes is clearly visible. Further analysis was carried out in the vicinity of each peak in the projected spectrum in order to clarify the position of each peak from the surrounding noise, which is described in the following section.

8.2.3 Statistical analysis for frequencies

The infinite limit in eqn 8.16 means that there is always a shortage of data in any MD simulation as the theoretical delta function peaks are broadened. Although with the advantage of computational power this is no longer a problem for classical MD, the shortage of enough CPU time is still acute for any first principles MD simulations. This results in a poorer sampling of the system and error in the overall prediction of any dynamical property.

In the present case this means that even the projected spectrum (Figure 8.1) is not easy to interpret. A straightforward interpretation would lead to very large uncertainty in the result. To overcome such a problem and to gain a greater confidence in the predicted results it is important to understand the source and nature of the error involved in the vibrational analysis at each stage of the statistical interpretation.

The first set of errors comes from eqn 8.5, the calculation of the auto-correlation $C(t)$

eqn 8.18: $$\langle \boldsymbol{V}(t+t_0).\boldsymbol{V}(t_0)\rangle_{time} = \frac{1}{t_{0\,max}} \sum_{t_0=1}^{t_{0max}} \boldsymbol{V}(t+t_0).\boldsymbol{V}(t_0), \quad \text{where}$$

eqn 8.19: $$t_{0\,max} + t = t_{max}\,.$$

Given that we have a fixed range of data, say up to t_{max} (*e.g.* 5 ps for rigid ammonia simulation) after the equilibration, for every choice of t we can average $C(t)$ from a maximum of $t_{0\,max} = (t_{max} - t)$ terms [in fact the total number of effective terms is even less: $(t_{max} - t)/\tau$, where τ is the correlation length of the dataset. The value of τ depends very much on the system of interest]. Thus while the short time correlation may be determined with greater statistical precision, the long time correlations become less and less reliable as t increases [inversely proportional to the square root of N, where $N = (t_{max} - t)/\tau$]; with the extreme case of $C(t = t_{max})$ we have only one term in the summation of eqn 8.18.

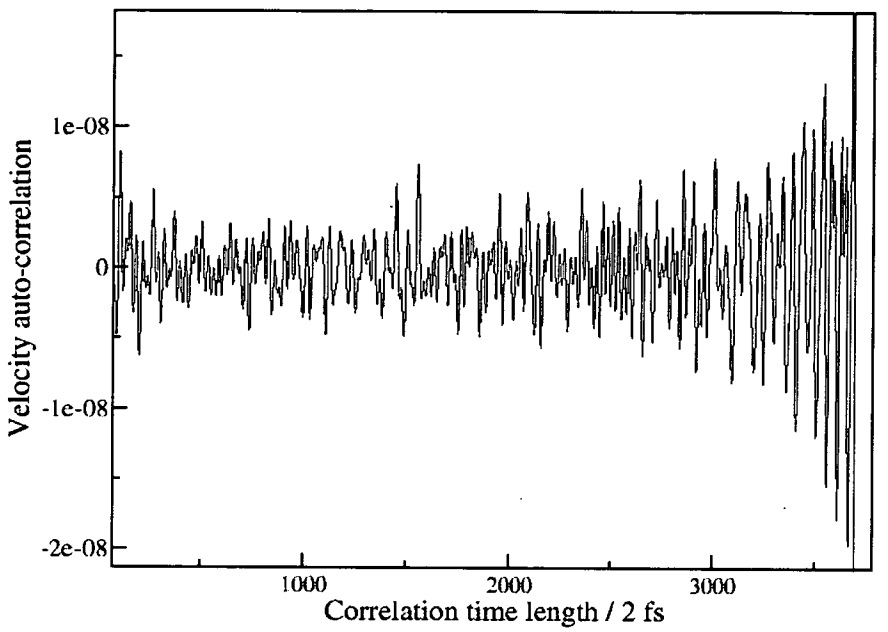


Figure 8.2: The velocity auto-correlation of a H atom from the rigid ammonia simulation. The maximum available time step for calculating the auto-correlation is 3700 (7.4 ps). The accuracy of the data falls sharply as the correlation time length, t, approaches t_{max} = 7.4 ps.

The problem could be avoided by truncating $C(t)$ at a maximum tolerable length of time, say t_{tol} (in Figure 8.2). However, truncating an already limited data set (due to limited CPU time) poses a new error in next stage of the statistical interpretation *i.e.* in evaluating eqn 8.16, which now becomes

eqn 8.20:
$$\hat{C}_k(w) = 2\Re \int_0^{t_{tol}} dt\, C'_{vv}(t)\, e^{-iwt},$$

where $C'_{vv}(t)$ is the velocity auto-correlation data projected at a specific k-point and corrected with the Blackman windowing function (see Section 7.3.3) for abrupt truncation at the t_{tol}. Choosing a very low tolerance (t_{tol}), although offering a better supply of quality data (from eqn 8.18), will incorporate a large uncertainty in the Fourier transformation (from eqn 8.20). Ideally we would want t_{tol} to be infinity. This leads to an optimal choice of the time length, say t_{opt} in Figure 8.2, which will be long enough to pick up the desired frequency upon Fourier transformation, while the noise contained in such a length of data is tolerable.

There is, however, no straightforward way to define such an optimal choice of time length for all the modes in the spectrum. Rather, it appears that the optimal length is different for different modes. The different vibrational modes are different in their strengths of amplitudes and time periods. Vibrations with large amplitudes are more tolerant to the noise and can cope with a much longer data set for $C'_{vv}(t)$ (eqn 8.18), whereas fast vibrations have very small time-periods and a shorter data set is sufficient to nicely pick up the mode in its Fourier transform (eqn 8.20). Consequently the small amplitude and the low frequency vibrations are the most difficult to interpret.

A lack of clear knowledge of the vibrational spectrum of ammonia means that we had to start with a poor guess of t_{opt} for the whole series. Once the peak positions were roughly defined, they were subsequently refined by individually targeting each of the vibrational modes and looking at their response with changes in t_{opt}.

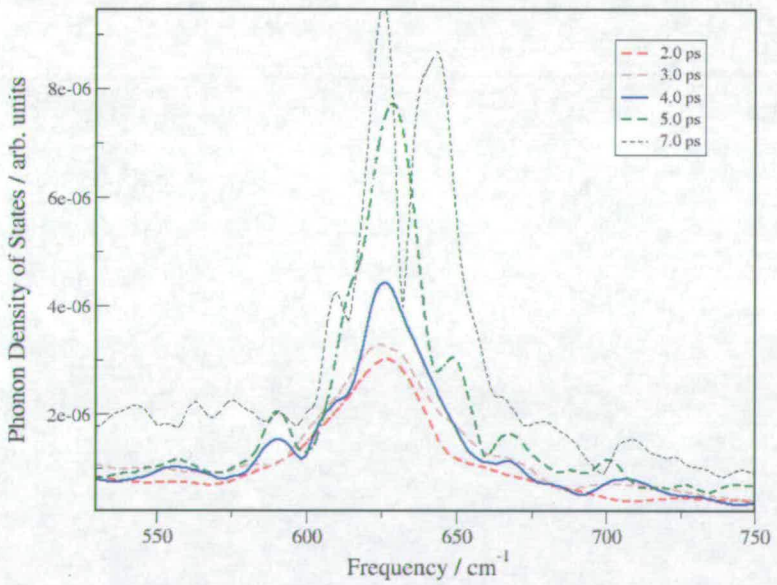


Figure 8.3: The normal mode at around 620 cm^{-1} for rigid ammonia simulation with different choice of correlation time length, t_{opt}, from Figure 8.2. The maximum available time length is 7.4 ps as in Figure 8.2. The vibrational spectra from all the basis atoms were added together to get a stronger response.

Figure 8.2 shows 5 different spectra for 5 different values of t_{opt} where we aim to refine the peak position for the normal mode at around 620 cm^{-1}. The best result is obtained for the blue peak at t_{opt} = 4 ps (Figure 8.3). For t_{opt} less than 4 ps we see a flattening of the peak due to a poor FT; using a longer data set (say t_{opt} = 5 ps, see Figure 8.3) results in symmetric ripples appearing on both sides of the main peak due to poor quality of data, $C'_v(t)$. The symmetric appearance of ripples reflects the fact that the errors in $C'_v(t)$ are of random nature, and are becoming significant. With even longer t_{opt} (say t_{opt} = 7 ps, see Figure 8.3) the ripples starts to grow, and might even take on the appearance of major peaks. Once the peak position is successfully identified, the maximum variation of the peak position (arising due to the change in the length of the data set) is reported within brackets, for example in this case the value is 622(8) cm^{-1}. Therefore a small number within the brackets will mean the peak is defined with much greater confidence, as the peak positions varies only slightly with change in t_{opt}. However, in cases where the peak positions varied widely with t_{opt}, along with any ambiguity in the total number of peaks, the whole range is reported, e.g. 340 – 370 cm^{-1} in our fully constrained (rigid) simulation (Table 7.5, Chapter 7).

Such an analysis was found to be very reassuring as we were limited by a reasonable but comparatively small amount of data. Following this careful analysis we have identified the following lattice frequencies: 99(3), 127(4), 140(3), 179(5), 296(4), 310(3), 340–370, 445(5), 460(5) and 622(8) cm^{-1} from the fully constrained (rigid ammonia) simulation of 2x2x2 supercell (Table 7.5, Chapter 7). Further valuable insight about the nature of these modes is gained upon careful review and comparison of all the previous experimental work on gamma point vibrational spectra.

8.2.4 Temperature and pressure effects

As we expect our results to differ slightly from the experimental values due to the fact that our simulations were run under slightly different conditions of temperature (~100 K) and pressure (0.1 GPa), it is important to have some understanding of the effect that temperature and pressure will have on the lattice mode values. The effect of pressure on lattice mode frequencies is measured in terms of the Gruneisen parameter, γ, given by,

eqn 8.21: $$\gamma = -\frac{\partial \ln \omega}{\partial \ln V},$$

where ω is the frequency and V is the molar volume [Nye, 1985]. These parameters are critically dependent on the form of the intermolecular potential of the particular system. The pressure effect on the lattice modes of ammonia has been investigated in detail by Nye and Medina [Nye, 1985]; their investigation on three particular lattice modes is summarised in Table 8.1.

Table 8.1: The effect of pressure on ammonia lattice modes: 98.5 cm^{-1}, 128.4cm^{-1} and 276.4 cm^{-1} at 195 K

[a] Molar volume (cm^3/mol)	[b] Pressure Change (kbar)	[b] Translational Frequencies (cm^{-1})		[b] Librational frequencies (cm^{-1})
20.62	Ref. Pressure	98.5	128.4	276.4
20.35	0.76(1)	102.1	132.5	279.4
20.29	0.93(1)	102.7	133.2	280.0
[a] Gruneisen parameters		2.7(1)	2.7(1)	2.4(1)

[a] [Nye, 1985]

[b] Pressure change has been calculated from molar volume using the reported bulk modulus, 5.8(6) GPa at 195 K, of solid ammonia phase I [Fortes, 2003]

The Table shows that for phase I solid ammonia the lattice mode frequencies increase by a few cm^{-1} due to a pressure rise of ~1 kbar. From the Raman frequency experiment by Yurtsven, 2004 [Yurtsven, 2004] the 268 cm^{-1} lattice mode of phase II solid ammonia (at 224.1 K and 0.365 GPa) also shows a frequency shift of similar magnitude.

Similarly, the 268 cm^{-1} librational mode increases by 10 cm^{-1} due to an increase in pressure of about 3 kbar (0.3 GPa). The pressure of our calculated sample increased by about 0.1 GPa (i.e. 1 kbar) due to fixed volume simulation. According to these results we might expect an upward frequency shift in the translational modes of the order of 5 cm^{-1}. It is also pointed out by Nye and Medina [Nye, 1985] that the frequency shift for librational modes are much smaller than for the translational modes.

The effect of temperature on the frequency shift of lattice modes was also well-studied [Nye, 1987]. This showed that the frequency falls by a few cm^{-1} for a significant rise in the temperature [e.g. the 141.0(4) cm^{-1} translational modes at 18 K falls to 130.9(16) cm^{-1} at 213.5 K]. As our simulation temperature is higher than the experimental temperature of Binbrek and Anderson [Binbrek, 1972] by only a few degrees (10 – 20 K), any effect, which we expect to be small, should counter-balance the frequency rise in our simulation due to slightly higher pressure.

8.2.5 Review of previous work

It is now timely to compile a brief summary of relevant literature work. The most cited experimental work for ammonia phonon modes are the IR and Raman experiments by Binbrek and Anderson [Binbrek, 1972]. Many other works are based on this reference, either to clarify some unsolved findings by them, or to gain more insight into the vibrational properties of crystalline ammonia.

Righini *et. al.* [Righini, 1978] made a successful attempt to recalculate these values using an empirical model calculation (often referred to as an atom-atom, multipole-multipole potential model). This calculation [Righini, 1978] tries to best-fit their model in order to closely reproduce the experimental work and by doing so they gained valuable insight of the system. Powell *et. al.* [Powell, 1980] used coherent inelastic neutron scattering (C-INS) in conjunction with the atom-atom empirical model to explore other k-points as well as the gamma point spectrum in order to systematically study the presence and absence of certain modes over different k-points. This technique gives a comprehensive answer for the symmetry assignment of normal modes. Both the empirical model calculation [Righini, 1978] and the C-INS spectra [Powell, 1980] were found to be in good agreement with the earlier work of Binbrek and Anderson [Binbrek, 1972], except that they suggested some changes in the original symmetry assignment made by the earlier work. Some other earlier attempts to resolve the spectra include: near-IR and Raman experiments by Reding and Hornig [Reding, 1951; Reding, 1954], far-IR work by Anderson and Walmsley [Anderson, 1965] and an inelastic neutron scattering (INS) study by Goyal *et.al.* [Goyal, 1972].

Although there are slight disagreements about frequencies and symmetry assignments between these mentioned works, they have established a clear picture of the lattice dynamics of ammonia. The work presented in this thesis is the first-known attempt to investigate the ammonia lattice modes from first-principles. We found much in accord with the established picture but a number of interesting disagreements were unearthed, which are reported in the following Sections.

The first few modes in the range ~100 cm^{-1} to ~160 cm^{-1}

Table 8.2 below lists all the peaks in the range ~100 cm^{-1} to ~160 cm^{-1} from previous works and our MD simulation of rigid ammonia.

Table 8.2: Ammonia lattice mode frequencies in the range ~100 cm^{-1} to ~160 cm^{-1}.

Method	Ammonia	Deutero-Ammonia	Ratio
[a]**Reding and Hornig**			
Near-IR (83 K)	–	–	–
Raman (191 K)	99 s, 129 m	91.5 s, 121 m,	1.08, 1.06,
		142-161vw	–
[b]**Binbrek and Anderson;** [c]**Anderson and Walmsley**			
Far-IR (77 K)	138 w	132 w	1.04
IR (107 K)	138(10)	131(10)	1.07
Raman (80 K)	107(5), 138(6)	100(4), 129.5(5)	1.07, 1.07
[d]**Atom-atom model**	122, 130,	113, 122,	1.08, 1.07
	143	131	1.09
[e]**C-INS** (95 K)	106.5, 141.5 **	98.6(0.7), 131(2.3)	1.08**
Ab initio MD	99(3) s, 127(4) vs,	–	–
(~110K)	140(3) m		

** Approximated values were obtained from corresponding ND_3 values using the standard ratio of pure translational modes.

vs = very strong, s = strong, m = medium, w = weak, vw = very weak.

[a][Reding, 1951; Reding, 1954], [b][Binbrek, 1972], [c][Anderson, 1965], [d][Righini, 1978], [e][Powell, 1980].

The fourth column of the Table 8.2 lists the ratio of the frequencies between ammonia and deutero-ammonia lattice modes. The combined result from ammonia and deutero-ammonia gives significant evidence to predict any mixing between the translational and librational modes. The ratios of vibrational frequencies between solid ammonia and deutero-ammonia should be 1.08 (square root of the ratio of molecular masses) for pure

translational modes, and ~1.4 (square root of the ratio of molecular moment of inertia) for pure librational modes. If any significant mixing between translational and vibrational motion appears in a certain ammonia mode the ratio will differ significantly from these standard values.

From the Table 8.2 it is immediately clear that all the vibrational modes appear in the range, ~100 cm^{-1} to ~160 cm^{-1}, are purely translational. Binbrek and Anderson [Binbrek, 1972] assigned them to the following symmetry movement as

A (107 cm^{-1}) , E (138 cm^{-1}) and F (138 cm^{-1}) symmetry.

The empirical model calculation [Righini, 1978] found similar values but their model predicted a different order of symmetry that is,

E (122 cm^{-1}), F (130 cm^{-1}) and A (143cm^{-1}) symmetry.

This confusion in the symmetry assignment was then lifted by the coherent INS work [Powell, 1980], which shows (using selection rules for intensities at multiple *k*-points) that the original assignment by Binbrek and Anderson [Binbrek, 1972] was correct. For deutero-ammonia the C-INS work claimed

A (98.6 cm^{-1}), E (98.6 cm^{-1}) and F (131 cm^{-1}) symmetry.

However, unlike Binbrek and Anderson's prediction, which claims the second peak lies very close to the third peak, the coherent INS [Powell, 1980] work claims that the second peak actually lies very close to the first peak. They found the lower peak splits at low temperature and reveals the presence of two different modes [Powell, 1980] as,

A (99.33 cm^{-1}), E (101.33 cm^{-1}) and F (133.6 cm^{-1}) symmetry at 20 K.

From our calculation we find all three peaks are well distinguished from each other (~99 cm^{-1}, ~127 cm^{-1}, ~140 cm^{-1}) and matches rather closely with the Raman work by Reding and Horning.

Lattice modes within the range ~160 cm^{-1} to ~280 cm^{-1}

Table 8.3 below shows all the peaks in the range ~100 cm^{-1} to ~160 cm^{-1}.

Table 8.3: Ammonia lattice mode frequencies in the range ~160 cm^{-1} to ~280 cm^{-1}.

Method	Ammonia	Deutero-Ammonia	Ratio
Reding and Hornig			
Near-IR (83 K)	250*	200*	1.25*
Raman (191 K)	–	–	–
Binbrek and Anderson; Anderson and Walmsley			
Far-IR (77 K)	180 w, 268 w	167 w, 200 w	1.08, 1.34
IR (107 K)	181(14), 258(12)	171(14),193(14)	1.08
Raman (80 K)	260(5)	192(9)	1.34
Atom-atom model	181, 273	168, 196	1.08, 1.4
C-INS (95 K)	179.9, 201.5/261.24**	166.6(1.7), 186.6(6.7)	1.08, 1.08/1.4**
Ab initio MD (~110K)	179 vs	–	–

*Not directly observed. Observed in combination band.
**Approximated values were obtained from corresponding ND_3 values using the standard ratio of pure translational/librational modes.
vs = very strong, w = weak, vw = very weak.

From Table 8.3, a common agreement of the peak at around 180 cm^{-1} is very clear and there is a indication for the second peak at around 260 cm^{-1}. All previous works agreed on F symmetry for these two modes. From their ratio with the corresponding deutero-ammonia mode, we can further suggest that the mode at around 180 cm^{-1} is the last translational mode of the series and the second peak at around 260 cm^{-1} is the first librational peak. The large gap between ~180 cm^{-1} and ~260 cm^{-1} [Goyal, 1972] is also consistent with the non-mixing between the translation and librational modes of solid ammonia .

Uncertainty in the 260 cm^{-1} mode

It is quite surprising that although the ab-initio calculation matched the other experimental peaks quite successfully, it clearly shows no possibility of any peak occuring around 260 cm^{-1}.

However, the confirmation of the second peak around 260 cm^{-1} in the range is also not evident from experiments. It was first suggested by Reding and Hornig from their near-IR (83 K) experiment [Reding, 1951]. The presence of this peak (at 250 cm^{-1}) in their work was *derived* from the combination bands, rather than detected directly from the scattering.

A very weak band at 268 cm^{-1} was first observed by Anderson and Walmsley [Anderson, 1965] in the far-IR spectra (77 K) but they could not detect the peak in their deutero-ammonia spectra. In their words: "Assuming the cubic structure for IR-work also, it is *felt* that 268 cm^{-1} represents a separate fundamental."

The presence of a mode at ~260 cm^{-1} was then *confirmed* using both Raman and IR data by Binbrek and Anderson [Binbrek, 1972] for both ammonia and deutero-ammonia samples. Close observation, however, shows that the mode registering in their work [Binbrek, 1972] is of very, very weak intensity, and could easily be considered to be within the error limit of the experiment. The 260 cm^{-1} mode was later reproduced by the atom-atom, multipole-multipole potential model calculation [Righini, 1978] and the coherent inelastic scattering [Powell, 1980] at 186.6 cm^{-1} for deutero-ammonia (186.6 times 1.4 gives 260.4 for the ammonia).

However, the presence of the 260 cm^{-1} peak in ammonia cannot be verified from the appearance of the 186.6 cm^{-1} peak in deutero-ammonia in the C-INS experiment [Powell, 1980] due to the following two reasons:

a) the C-INS spectra shows only one peak (instead of two) in this range of interest with two insignificant shoulders, one of which is 166 cm^{-1} (claimed as the translational F mode) whereas the other shoulder was ignored.

b) While the 166 cm^{-1} shoulder can be attributed to error in the experiment, the peak 186.6 cm^{-1} can be matched to either a translational F mode (a choice of ratio, 1.08 predicts the mode at 201.5 cm^{-1}) or a librational F mode (a different choice of ratio, 1.4 predicts the mode at 260.4 cm^{-1}) of ammonia.

Two peaks at about 300 cm^{-1}

Table 8.4 below compares all the modes observed near 300 cm^{-1}.

Table 8.4: Ammonia lattice mode frequencies in the range ~280 cm^{-1} to ~320 cm^{-1}.

Method	Ammonia	Deutero-Ammonia	Ratio
Reding and Hornig			
Near-IR (83 K)	–	–	–
Raman (191 K)	284 vs	213 vs	1.33
Binbrek and Anderson; Anderson and Walmsley			
Far-IR (77 K)	–	–	–
IR (107 K)	–	–	–
Raman (80 K)	298(8), 310(9)	227.5(5.5), 242(-)	1.31, 1.3
Atom-atom model	299, 306	219, 223	1.37, 1.37
C-INS (95 K)	295.1, 315.9 **	227(1.7), 243(3.3)	1.3 **
Ab initio MD (~110K)	290 m, 310 m	–	–

** Approximated values were obtained from their corresponding ND_3 values using the ratio observed in Raman(80 K) experiment.
vs = very strong, m = medium.

The existence of two closely-spaced peaks around 300 cm^{-1} is agreed by all the experiments and our ab-initio MD result. The initial symmetry assignment by Binbrek and Anderson [Binbrek, 1972] was

A (298 cm^{-1}) and E (310 cm^{-1}) symmetry.

This assignment, however, did not match the atom-atom, multipole-multipole potential model calculation as their calculation predicted the correct distribution only if the symmetry assignments were interchanged [Righini, 1978], that is

E (299 cm^{-1}) and A (306 cm^{-1}) symmetry.

This observation was supported by a more rigorous Coherent INS experiment on deutero-ammonia [Powell, 1980]. As they claimed: "Our observations also suggest that the A- and E-mode identification of Binbrek and Anderson, at 231 [corresponds to 298 cm^{-1} in NH_3] and 246 cm^{-1} [corresponds to 310 cm^{-1} in NH_3] respectively, should be interchanged. Their identification scheme would lead, ..., to a large predicted neutron intensity for the mode at 231 cm^{-1} [corresponds to 298 cm^{-1} in NH_3] at reciprocal lattice point (220), whereas no intensity is observed, and would also predict zero intensity for the mode at 246 cm^{-1} [corresponds to 310 cm^{-1} in NH_3], whereas intense scattering is observed. Interchanging the two modes resolves this difficulty."

F modes in the range 320 cm^{-1} to 500 cm^{-1}

All the works generally agrees to the presence of two distinct peaks in this range, one around 360 cm^{-1} and the other one around 430 cm^{-1} (See table 8.5).

The first peak is assigned to an F-mode. However, the assignment to the second peak around 430 cm^{-1} is not very clear from previous works. Binbrek and Anderson [Binbrek, 1972] first suggested that this could be attributed to either a combination mode or an LO branch of the 360 cm^{-1} mode. From the very strong presence of the 360 cm^{-1} mode in the IR spectrum, and the absence of the 430 cm^{-1} mode in the IR spectrum, Binbrek and Anderson concluded that such a large LO-TO splitting (~70 cm^{-1}) could be possible.

Table 8.5: Ammonia lattice mode frequencies in the range ~320 cm^{-1} to ~500 cm^{-1}.

Method	Ammonia	Deutero-Ammonia	Ratio
Reding and Hornig			
Near-IR (83 K)	362 s	–	–
Raman (191 K)	325-375 w, 430-462 vw	280-312 w, 330-380 w	1.6-1.2
Binbrek and Anderson; Anderson and Walmsley			
Far-IR (77 K)	362 s, –	297 s, –	1.22, –
IR (84 K)	361(26), –	271(30), –	1.33, –
Raman (80 K)	358(21), 426(27)	267(16), 315(20)	1.34, 1.35
Atom-atom model	360, 423	262, 306	1.38, 1.38
C-INS (95 K)	373.8, 457.8**	267(3.3), 327(3.3)	1.4**
Ab initio MD (~110K)	340 – 370* s; 440 – 465 s splits in 445(5) and 460(5)	–	–

* Shows possibility of more than one peak in the range.
** Approximated values of the ammonia modes obtained using the standard ratio of pure librational modes in NH_3 and ND_3 solid.
s = strong, w = weak, vw = very weak.

The empirical model calculation by Righini *et. al.* [Righini, 1978] later claimed that the suggested LO-TO splitting of the 360 cm^{-1} mode is in good agreement with their model. However, the dipole moment of the system was changed by $0.06x10^{-18}$ e.s.u from the gas phase value in their calculation to improve the initial splitting from 43 to 63 cm^{-1} [Righini, 1978]. Although, the change in the dipole moment was made in an attempt to increase the observed LO-TO splitting, no physical interpretation of this change in dipole moment were presented. The more comprehensive work of C-INS [Powell, 1980] spectroscopy also claims these two modes [at 371.8 cm^{-1} and the 457.8 cm^{-1} (Table 8.5)] as LO-TO branches of an F-mode. It is important to note that LO-TO splitting of F modes in C-INS work is also predicted using the empirical model proposed by Righini *et. al.* [Righini, 1978; Powell, 1980].

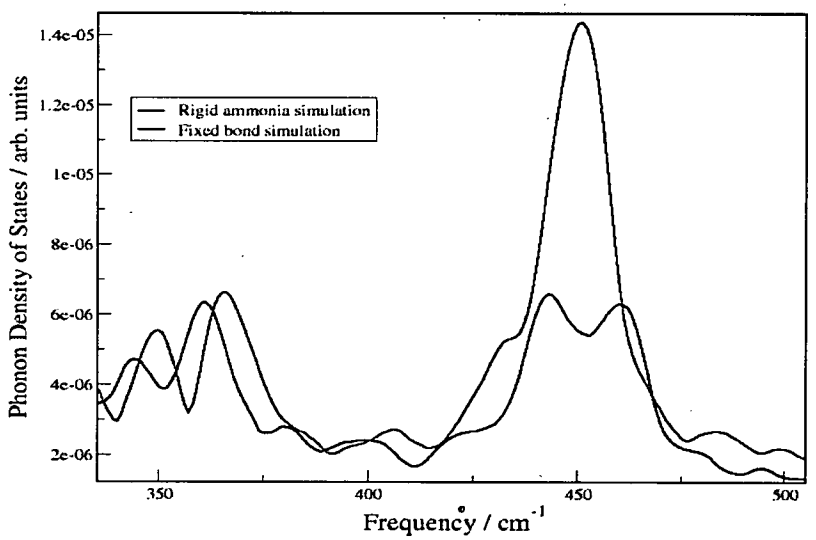


Figure 8.4: The lattice modes at around 350 and 450 cm^{-1} from fixed-bond and rigid ammonia simulation for a dataset $C(t)$ of length 5ps. (Note that, both simulations, however, exhibit the possibility of doublets at both frequency ranges.)

Although there is no clear evidence from previous works that these two frequencies are the LO-TO counterpart of a single F-mode (or two different fundamental modes), the MD simulation of a 2x2x2 supercell is expected to pick up a significant part of the long

range interactions and thus should be able to predict any such splitting. From our simulation (see Figure 8.4) we see clear presence of at least two broad peaks in the range 340-370 cm^{-1} and 440-465 cm^{-1}, which, in line with the previous works, could be considered as LO-TO branch of the suggested F mode.

An alternative assignment (of 340-370 cm^{-1} and 440-465 cm^{-1})

A closer investigation of our data, however, reveals that each of these broad peaks (340-370 cm^{-1} and 440-465 cm^{-1}) are further divided into doublets (narrow splitting of broad peaks). The doublets appear rather consistently in our constrained and unconstrained simulations (see Table 8.5 and also Table 7.5, Chapter 7), which therefore rules out the possibility of them appearing simply due to poor data quality [*i.e.* ripples appearing as a major peak in the form of doublets (see Figure 8.5)]. The previous assignments [Binbrek, 1972; Righini, 1978; Powell, 1980] of 340-370 cm^{-1} range to an LO branch, and 440-465 cm^{-1} range to a TO branch of a single F-mode, thus leaves an open question to these distinct splittings (of ~15 cm^{-1}) of each peaks.

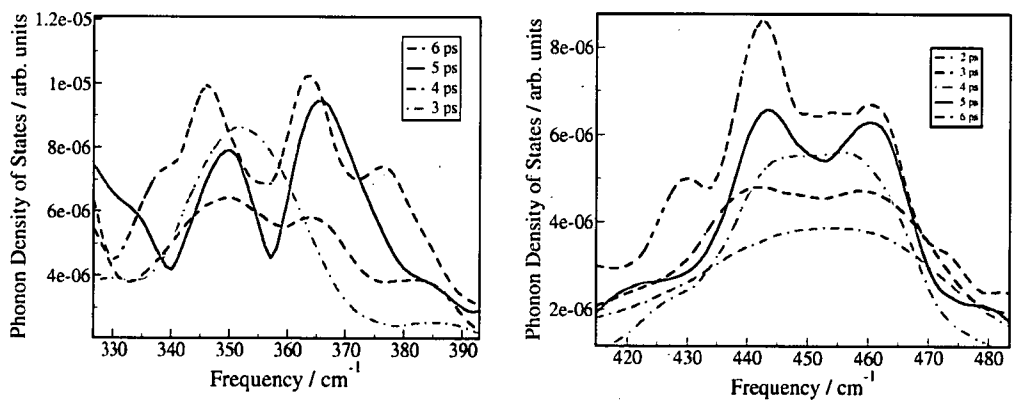


Figure 8.5: The presence of doublets becomes evident with increase in the length of dataset C(t), at ~350 cm^{-1} for fixed bond simulation and at ~450 cm^{-1} for rigid ammonia simulation.

A doublet was also observed for the lower range (at around 360 cm^{-1}) by Binbrek and Anderson [Binbrek, 1972], but they did not attribute this to LO-TO splitting as they have observed the splitting in both IR and Raman spectra. (IR is insensitive to such an effect). Very broad peaks were also observed by Reding and Hornig in their Raman (191 K) spectra [Reding, 1954] in both of these two ranges for both solid NH_3 and ND_3 (Table 8.5), also suggesting the possibility of more than one peak in each of these ranges (340-370 cm^{-1} and 440-465 cm^{-1}) of spectrum.

A different conclusion might be possible, where each range (340-370 cm^{-1} and 440-465 cm^{-1}) represents two separate F modes, based on the following reasons:

a) it explains the doublets observed at each of these ranges. Combination or overtone bands will not be detected by Fourier analysis of the MD simulation, therefore the only possibility of these doublets are to be LO-TO splitting of two individual F-modes.

b) The large gap [70 cm^{-1} in Raman experiment (80 K) and 85 cm^{-1} in C-INS spectra (95 K)] between these two modes seems quite out of place compared to the LO-TO splitting of all the other F modes of ammonia, which are very small or show no splitting effect [Binbrek, 1972; Powell, 1980].

c) Finally, considering them as two separate modes counts for the full nine independent lattice modes from our ab-initio MD simulation (as our ab-initio MD simulation completely misses the presence of a peak around 260 cm^{-1}).

Lattice mode above 500 cm^{-1}

Only one lattice vibrational mode was observed above 450 cm^{-1}.

Table 8.6: Ammonia lattice mode frequencies above 500 cm^{-1}.

Method	Ammonia	Deutero-Ammonia	Ratio
Reding and Hornig			
Near-IR (83 K)	527 vs	406 vs	1.3
Raman (191 K)	–	–	–
Binbrek and Anderson; Anderson and Walmsley			
Far-IR (77 K)	–	–	–
IR (84 K)	532(18)	405(14)	1.31
Raman (80 K)	–	–	–
Atom-atom model	524	383	1.37
C-INS (20 K)	574**	410(3.3)	1.4**
Ab initio MD (~110K)	622(8) vs	–	–

** Approximated values were obtained from the corresponding ND_3 values using the standard ratio of pure librational modes.
vs = very strong.

The normal mode of F symmetry at around 530 cm^{-1} was only observed by two experimental works (see Table 8.6) of IR spectra [Reding, 1951; Binbrek, 1972]. They, however, strongly suspected it as an overtone of 260 cm^{-1} or a combination band. The atom-atom empirical model [Righini, 1978] and the C-INS [Powell, 1980] work later confirmed it as a fundamental F mode. However, the C-INS spectra gives a much higher value (574 cm^{-1}). The presence of a peak within 500 to 600 cm^{-1} is clearly absent in our calculation; rather we see a very sharp peak at 622(8) cm^{-1}.

In summary, from the careful review of previous works we propose the following assignment of the nine modes for phase I solid ammonia at gamma point: A-99(3), E-127 (4), F-140(3), F-179(5), E-296(4), A-310(3), F-340–370, F(TO)-445(5), F(LO)-460 (5) and F-622(8) cm^{-1}. (Note that the symmetry assignment are not from first-principles, they are simply a best match to the reports available form earlier works).

8.3 Eigenvector Extraction

Given that the gamma point frequencies are now known, it is now possible to proceed further and extract their corresponding eigenvectors. The theory assembled in this work to achieve this aim is reported in the remainder of this Section.

8.3.1 The theory

According to the classical theory of harmonic crystal the motion of an atom in a crystal follows the particular pattern,

eqn 8.22: $$\boldsymbol{V}_{b,\boldsymbol{R}}(t) = \sum_{\boldsymbol{k}} \sum_{\omega} \boldsymbol{\varepsilon}_{b,\boldsymbol{k},\omega}\, \mathrm{e}^{i(\boldsymbol{k}.\boldsymbol{R}-\omega t)},$$

where $\boldsymbol{V}_{b,\boldsymbol{R}}$ is the velocity of an atom, say i, of basis $\boldsymbol{b}$ at a lattice site $\boldsymbol{R}$. $\varepsilon_{b,\boldsymbol{k},\omega}$ is the velocity eigenvector of the normal mode $(\boldsymbol{k}, \omega)$, and is same for all atoms with same basis, $\boldsymbol{b}$.

The first step to extract the eigenvector of a normal mode at a specific $(\boldsymbol{k}, \omega)$, is to project the velocity data at the desired $\boldsymbol{k}$. In this case we are particularly interested in gamma modes at $\boldsymbol{k}$ = (0,0,0), which could be obtained using the method described in Section 8.2.3 (eqn 8.15), as

eqn 8.23: $$\sum_{\{\boldsymbol{R}\}} \boldsymbol{V}_{b,\boldsymbol{R}}(t) = V_{b,\Gamma}(t)\,.$$

So the projected velocity at gamma point would be

eqn 8.24: $$\boldsymbol{V}_{b,\Gamma}(t) = \sum_{\omega} \boldsymbol{\varepsilon}_{b,\Gamma,\omega}\, \mathrm{e}^{-i\omega t}.$$

Each atom would have its initial phase (at time t = 0) for each individual mode, (Γ, ω). Including the phase the velocity at the gamma point due to all modes would be

eqn 8.25:
$$V_{b,\Gamma}(t) = \sum_{\omega} \varepsilon_{b,\Gamma,\omega}\, e^{i(-\omega t+\phi_{\Gamma,\omega})}.$$

Multiplying the projected velocity at gamma point ($V_{b,\Gamma}$) with the phase factor of the known frequency of the crystal $e^{i\omega_0 t}$, and taking a normalised sum over all the available data range, gives us

eqn 8.26:
$$\frac{1}{N_{tot}} \sum_{t=0}^{t_{tot}} V_{b,\Gamma}(t) e^{i\omega_0 t} = \frac{1}{N_{tot}} \sum_{t=0}^{t_{tot}} \left(\sum_{\omega} \varepsilon_{b,\Gamma,\omega}\, e^{i(\omega_0-\omega)t + i\phi_{\Gamma,\omega}} \right),$$

where $N_{tot} = t_{tot}/\Delta t$ and Δt is the time step of the simulation. Eqn 8.26 could be re-written as

eqn 8.27:
$$\frac{1}{N_{tot}} \sum_{t=0}^{t_{tot}} V_{b,\Gamma}(t) e^{i\omega_0 t} = \sum_{\omega} \varepsilon_{b,\Gamma,\omega}\, e^{i\phi_{\Gamma,\omega}} \left(\frac{1}{N_{tot}} \sum_{t=0}^{t_{tot}} e^{i(\omega_0-\omega)t} \right).$$

For very short time step ($\Delta t \rightarrow 0$, which is true for the MD to produce correct trajectory) and very long interval ($t_{tot} \rightarrow \infty$), the term in brackets will represent the Kronecker delta function, as

eqn 8.28:
$$\lim_{N_{tot}\rightarrow\infty} \frac{1}{N_{tot}} \sum_{t=0}^{t_{tot}} e^{i(\omega_0-\omega)t} = \delta_{\omega_0,\omega}.$$

Taking this identity in eqn 8.27,

eqn 8.29:
$$\lim_{N_{tot}\rightarrow\infty} \frac{1}{N_{tot}} \sum_{t=0}^{t_{tot}} V_{b,\Gamma}(t) e^{i\omega_0 t} = \varepsilon_{b,\Gamma,\omega_0}\, e^{i\phi_{\Gamma,\omega_0}}.$$

Eqn 8.29 is our working equation to extract the eigenvectors at desired frequency ω_0. We have worked with the velocity data as it is easier to handle in this particular case. However, once the velocity eigenvector is extracted the position eigenvector can easily be derived using the relation

eqn 8.30: $$\boldsymbol{\Sigma}_{b,k,\omega} = \frac{\boldsymbol{\varepsilon}_{b,k,\omega}}{\omega(\boldsymbol{k})},$$

where $\boldsymbol{\Sigma}_{b,k,\omega}$ is a position eigenvector and $\omega(\boldsymbol{k})$ is the frequency at the particular value of $\boldsymbol{k}$.

8.3.2 Limitations

Although the above mechanism should allow us to extract the eigenvectors at desired frequencies, limitations in both theory and simulation makes it rather complex. The first difficulty comes from the limited range of the data set. Eqn 8.29 is exact only at an infinite limit. Consequently the error in the sum of eqn 8.29 becomes tolerable only for an extensively large (ideally infinite) data set, but such a long dataset is hard to achieve from first-principles simulation.

The fact that the eigenvectors are extracted only at a particular frequency and *k*-point, is often inadequate. If a normal mode is degenerate (*e.g.* F mode) the above method would extract a linear combination of all the degenerate modes at this frequency, which would then be rather hard to assign with symmetry.

Given these shortcomings, the method derived for eigenvector extraction gives a unique way to gain much a closer look into the crystal vibration. Although it may not reveal all the detailed information with exact certainty, nevertheless, it would undoubtedly extract the main features of the desired eigenmodes.

8.3.3 Viewing the eigenvectors

Eqn 8.29 and eqn 8.30 gives the eigenvectors entangled with the initial phase of the the mode, at the start time $t_{start} = 0$. Removing the phase is, however, not straightforward. Taking the real and imaginary part of eqn 8.29 we have,

eqn 8.31: $$\boldsymbol{A} \equiv \Re\left(\lim_{N_{tot}\to\infty} \frac{1}{N_{tot}}\sum_{t=0}^{t_{tot}} \boldsymbol{V}_{b,\Gamma}(t)\mathrm{e}^{i\omega_0 t}\right) = \boldsymbol{\varepsilon}_{b,\Gamma,\omega_0}\cos\phi_{\Gamma,\omega_0},$$

eqn 8.32: $$\boldsymbol{B} \equiv \Im\left(\lim_{N_{tot}\to\infty} \frac{1}{N_{tot}}\sum_{t=0}^{t_{tot}} \boldsymbol{V}_{b,\Gamma}(t)\mathrm{e}^{i\omega_0 t}\right) = -\boldsymbol{\varepsilon}_{b,\Gamma,\omega_0}\sin\phi_{\Gamma,\omega_0}.$$

From the ratio of $\boldsymbol{A}$ and $\boldsymbol{B}$, the phase of the eigenvectors could be extracted as,

eqn 8.33: $$\phi_{\Gamma,\omega_0} = \arctan\left(\frac{-\boldsymbol{B}}{\boldsymbol{A}}\right).$$

However, the true phase information gets lost here in eqn 8.33 as the *arctan* function is bound between 0°→180°. This make it impossible to remove the initial phase, $\phi_{\Gamma,\omega}$ from the eigenvectors analytically.

Alternatively we can construct the complete time evolution of each mode from eqn 8.31 by repeating the process of calculations over many times, using slightly different start times t_{start} at each stage. The $3N$ set of vectors (N is number of atoms in the unit cell), $\boldsymbol{\varepsilon}_{b,\Gamma,\omega_0}\cos\phi_{\Gamma,\omega_0}$, in eqn 8.31 represents the true displacement of the system at the time $t = t_{start}$ due to a *single* mode under consideration (designated by the frequency, ω). Thus changing the t_{start} successively over the length of it's complete time period ($2\pi/\omega$) will generate successive displacement of the system under that particular mode over it's whole period. The results is, however, not easily presentable in this thesis format and hence supplied (with comments and explanations) in electronic form [as animated gif (Graphics Interchange Format) file and also as xyz data file] for the purpose of assessment. The material (in a compact disc) is available in a pocket inside the back cover of the thesis. The result is also available for viewing at the following URL(universal resource locator): http://www.ph.ed.ac.uk/~gja/Ammonia_Modes.html, provided by the School of Physics, The University of Edinburgh official website.

8.3.4 Results from solid ammonia simulation

The complete motion of the system (over a whole time period) were extracted for each of the molecular modes at frequencies, ω = 3555(asymmetric stretch), 3415 (symmetric stretch), 1645(asymmetric bending) and 1095(symmetric bending) cm^{-1} from the fully flexible ammonia simulation. The original motion had to be amplified for clearer observation, and are available as animated gif files within the supplementary materials. The result clearly shows the six molecular modes (including the degeneracy) with each molecule differing in phase from the others. This is due to the fact that the four molecules in the unit cell give rise to an additional four fold degeneracy resulting in a total of 6x4 = 24 molecular modes per unit cell of ammonia.

Our extracted motions show very good match with the results available in the standard texts for these molecular modes [Ebsworth, 1991]. The error was found to be almost negligible. For example, the ammonia stretching modes show less than 0.01° change in the H–N–H angle and the bending modes show less than 0.001 Å change in the N–H bond length. The strong match between the extracted modes with the established picture of gas phase ammonia modes serves as a test of accuracy of the method developed and provides the essential confidence to interpret the results further for the lattice vibrations.

The complete set of eigenvectors (available as animated gif files within the supplementary materials) were also extracted at each of the calculated lattice frequencies in the earlier sections (Section 8.2) from the fully constrained (rigid) simulation. It is now timely to put forward a detailed report of the symmetry assignments obtained using these extracted eigenvectors. An in depth analysis of the symmetry of the phase I ammonia structure is needed in order to establish the desired match between the observed motion and symmetry, which at this stage is, however, beyond the scope of this PhD project. From our limited analysis we would, however, like to address the following interesting conclusions:

1) For the non-degenerate translational modes, we expect the atoms to approach the equilibrium position once at every half cycle. For degenerate translational modes, however, due to the initial phase difference of constituent modes, atoms will come to equilibrium once at a complete cycle.

Based on these criteria, 127 cm^{-1} clearly demonstrates itself as a non-degenerate translational mode, which is surprisingly in contrast with all the previous predictions [Binbrek, 1972; Righini, 1978; Powell, 1980]. On the other hand, the nitrogen atoms at mode 99 cm^{-1} mode shows slight elliptical motions, which is only possible if two or more non-degenerate translational modes are mixed. We therefore strongly suggest reconsidering these two modes as

E (99 cm^{-1}) and A (127 cm^{-1}) symmetry.

2) From our observation of extracted lattice modes (particularly the motions of Nitrogen atoms), it is clear that the 99(3), 127(4), 140(3) and 179(5) cm^{-1} modes are translational modes, and the higher ones are librational modes, which confirms all the previous predictions. It is strongly felt that a further symmetry analysis will also reveal the level of mixing between the translation and librational modes.

3) From our obtained spectrum in Section 8.2 (Figure 8.1) we do not obtain peaks at 260 and 532 cm^{-1} as predicted by experiments [Binbrek, 1972]. In search for a definitive answer to this anomaly we extracted the atomic motions corresponding to these values (irrespective to the existence of any peak). The obtained motion was found much (3 to 10 times) weaker compared to the motions extracted at distinguished peaks, which validates our findings that 260 and 532 cm^{-1} do not represent any real phonon mode.

4) However, when the extracted motions at 260 and 532 cm^{-1} were amplified many fold, their motions were found to bear some distinct patterns, rather than showing random oscillation. This is, however, not surprising, as a closer look shows that they actually bear the patterns of the neighbouring modes, reflecting the fact that these values (260 and 532 cm^{-1}) indeed lie on the tails of other peaks. The 260 cm^{-1} motions show a mixture of weak translational and strong librational motion, which indicate that the librational tail of 296 cm^{-1} is much stronger than the translational tail of 179 cm^{-1} at this value.

   This effect is also evident even far away from the presence of strong peaks. The extracted motion at 700 cm^{-1}, where we expect to see no systematic motion, although very weak bears a resemblance to the 622 cm^{-1} mode.

5) The mixing of different modes due to the extended tails of the phonon peaks (Figure 8.1) has made it hard to distinguish the non-degenerate modes from the degenerate ones in the librational region (based on the criteria that the rotational amplitudes of different molecules under degenerate mode would be different due to the initial phase difference of the constituent modes). Particular effort has been made to distinguish between the LO-TO modes but no clear conclusion could be drawn.

Although the level of error and/or mixing between different modes (due to each of the peaks having extended tails) could be significant, from the above findings it is strongly felt the residual effect from other modes could be removed with due assistance from the symmetry analysis.

8.4 Conclusion

The methodology formulated in this chapter allows us to study the lattice dynamics of molecular materials in a comprehensive way from first principles†. The methodology is applied successfully to cubic ammonia phase I structure and reveals all the nine eigen-modes of ammonia at gamma points, which are E-99(3), A-127(4), F-140(3), F-179(5), E-296(4), A-310(3), F-340–370, F(TO)-445(5), F(LO)-460(5) and F-622(8) cm^{-1} as prescribed by the Group Theory analysis. (Note that, the assignment of the first two modes has been interchanged from what has been assigned previously in Section 8.2.4 from the review of previous works).

Phonon eigenvectors are not experimentally accessible, so theoretical work can complement experimental investigation, for example by following the changing nature of modes under pressure [Hsueh, 1995; Hsueh, 1996]. Although in this chapter, we have restricted ourselves from the symmetry analysis of those extracted eigenvectors, which will vary much depend on the system under study, nevertheless this work establishes a general method of complete phonon analysis from first-principles MD. The use of constraint dynamics has made this method well comparable to the other first principles techniques (*e.g.* the linear response theory and the force constant calculation from first principles), and even better in some respects (*e.g.* in treating anharmonic modes).

†Theories developed in Sections 8.2.2 and 8.3.1 (eqn 8.16 and eqn 8.29) are original works of this thesis. During the course this thesis was produced, we were not aware of any similar work reported on k-point projection and eigen vector extraction for periodic solid systems. However, on completion of this thesis, we have now come across some analogous works by Boon and Yip [Boon, 1980] for liquid systems.

8.5 Bibliography

1. Allen, M. P., and D. J. Tildesley, *Computer simulation of liquids*, (Oxford University Press, 1987).

2. Anderson, A., and S. H. Walmsley, *Mol. Phys.*, **9**, 1(1965).

3. Ashcroft, N. W., and N. D. Mermin, *Solid state physics*, (W.B. Saunders Company, 1976).

4. Binbrek, O. S., and A. Anderson, *Chem. Phys. Lett.*, **15**(3), 421(1972).

5. J.P Boon and S. Yip, *Molecular Hydrodynamics*, (Dover, New York, 1980).

6. Ebsworth, E. A. V., D. W. H. Rankin and S. Cradock, *Structural methods in inorganic chemistry*, (Blackwell Publications, 1991).

7. Fortes, A. D., J. P. Brodholt, I. G. Wood and L. Vocadlo, *J. Chem. Phys.*, **118**(13), 5987 (2003).

8. Goyal, P. S., B. A. Dasannacharya, C. L. Thaper and P. K. Iyengar, *Phys. Stat. Sol. B*, **50**, 701(1972).

9. Hsueh, H. C., H. Vass, S. J. Clark and J. Crain, *Europhys. Lett.*, **31**,151(1995).

10. Hsueh, H. C., M. C. Warren, H. Vass, G. J. Ackland, S. J. Clark and J.Crain, *Phys. Rev. B*, **53**,14806(1996).

11. Nye, C. L., and F. D. Medina, *J. Chem. Phys.*, **87**(12), 6890(1987).

12. Nye, C. L., and F. D. Medina, *Phys. Rev. B*, **32**(4), 2510(1985).

13. Powell, B. M., G. Dolling, G. S. Pawley and J. W. Leech, *Can. J. Phys.*, **58**, 1703 (1980).

14. Reding, F. P., and D. F. Hornig, *J. Chem. Phys.*, **19**(5), 594(1951).

15. Reding, F. P., and D. F. Hornig, *J. Chem. Phys.*, **22**(11), 594(1954).

16. Righini, R., N. Neto, S. Califano and S. H. Walmsley, *Chem. Phys.*, **33**, 345(1978).

17. Yurtseven, H., *Chinese J. Phys.*, **42**(2), 209(2004).

Chapter 9

Future Work

> 'And if all the trees on earth were pens, and the sea [were] ink, with seven [more] seas yet added to it, the words of God would not be exhausted: for, verily, God is almighty, wise.' [*The Qur'an*, **31**, 27]

During the course of this PhD work, a powerful tool of constrained molecular dynamics (MD) has been developed that will allow the study of bigger systems for longer time scales than was not previously accessible from first-principles. This has been implemented for one of the most popular first-principles packages, CASTEP, which is utilised globally by both the academic and commercial communities.

It is therefore anticipated that many exciting and innovative research projects will benefit directly from this work. One example, which is of immediate interest to us is the cluster formation of methanol in water at various methanol-water ratios. There has been a strong indication from empirical model simulations that the nature of cluster formation at different ratios is entropy driven, *i.e.* the inter-molecular interactions play the main role [Dixit, 2002; Finney, 2003]. Realistic modelling of such systems will require simulations of the order of 100 ps with a few hundreds of atoms [van Erp, 2001], and can therefore only be achievable from first-principles using the constrained dynamics method.

Another interesting area ripe for future work follows from our successful development of a methodology for extracting phonon modes from first-principles MD calculations. This will allow direct observation of how the lattice vibrational properties change with variable conditions of temperature and pressure. The real-time change in an anharmonic vibration (called mode softening) which can be reliably observed only by the MD route, has a key role to play in predicting possible phase transitions [Warren, 1998]. From the observation of a change in eigen-vectors, this will also allow us to study directly from first-principles the mechanism for such transitions (*e.g.* the motion of particular atoms triggering the transition) [Clark, 1997]. This tool therefore offers much promise in offering unique insight into the study of systems under extreme conditions (*i.e.* at high pressure and temperature) [Karki, 1997].

9.1 Bibliography

1. Clark, S. J., and G. J. Ackland, *Phys. Rev. B*, **56**, 47(1997).

2. Dixit, S., A. K. Soper, J. L. Finney and J. Crain, *Europhys. Lett.*, **59**(3), 377(2002).

3. Dixit, S., J. Crain, W. C. K. Poon, J. L. Finney and A. K. Soper, *Nature*, **416**, 829 (2002).

4. Finney, J. L., D. T. Bowron, R. M. Daniel, P. A. Timmins and M. A. Roberts, *Biophys. Chem.*, **105**, 391(2003).

5. Karki, B. B., M. C. Warren, L. Stixrude, G. J. Ackland and J. Crain, *Phys. Rev. B*, **55**, 3465(1997).

6. van Erp, T. S., and E. J. Meijer, *Chem. Phys. Lett.*, **333**, 290(2001).

7. Warren, M. C., G. J. Ackland, S. J. Clark and B. B. Karki, *Miner. Magazine*, **62**, 585 (1998).

Appendix A

Conference and Poster Presentations

1. 10th International Conference on the Applications of Density Functional Theory in Chemistry&Physics, Sept, 2003, Brussels, Belgium (includes poster presentation).

2. CCP5 Conference, "Modelling of Materials: Atomistic and Ab-Initio Approaches", Mansfield College, Oxford, April, 2003 (includes poster presentation).

3. CCP5 Summer school, "Methods in Molecular Simulation" King's College, London, July, 2002 (includes poster presentation).

4. CASTEP Developer Workshop, "Nuts and Bolts of First-Principles simulation" Durham, December, 2001.

Appendix B

List of Publications

1. M.M. Siddick, G. J. Ackland & C. A. Morrison, Constrained MD of Solid Ammonia from first principles, *Phys. Rev. B*, Manuscript to be submitted.

2. C. A. Morrison, M. M. Siddick, P. J. Camp & C. C. Wilson, Towards understanding mobile proton behavior from first principles: the short hydrogen bond in crystalline urea-phosphoric acid, *J. Am. Chem. Soc.*, 127(11), 4042 (2005).

3. C. A. Morrison & M. M. Siddick, Dihydrogen bonds in solid BH_3NH_3, *Angew.Chem.*, 43, 4780(2004).

4. C. A. Morrison & M. M. Siddick, Determining the Strengths of Hydrogen Bonds in Solid-State Ammonia and Urea: Insight from Periodic DFT Calculations, *Chem. Eur .J.*, 9(3), 628(2003).