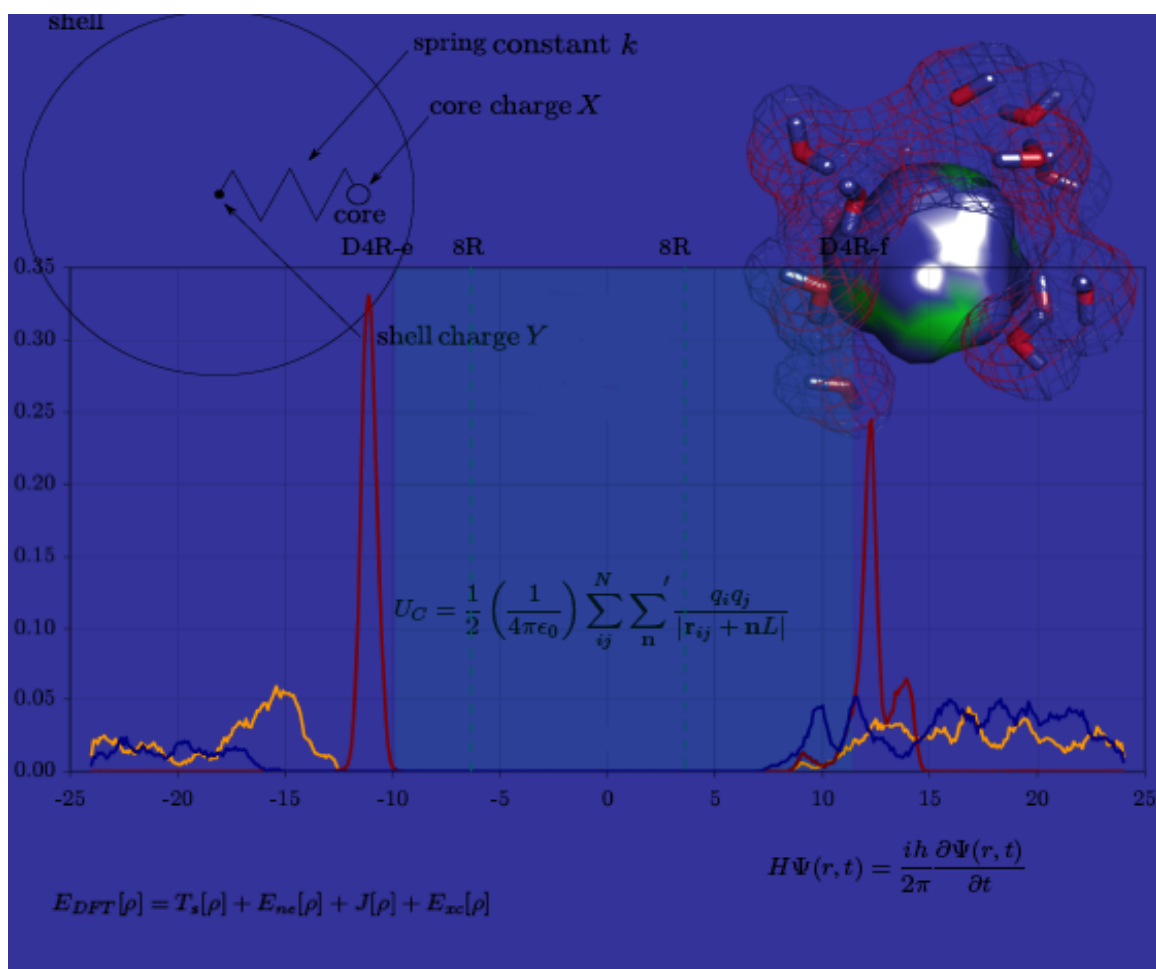# An Introduction To Computational Methods

Keith Butler, Kim Jelfs, Wojtek Gren

April 16, 2008

1   Preface

Before beginning please note that as many terms used are either unique to computational chemistry or have a unique meaning in this context a glossary of these terms, which will appear in bold face is given at the end. Also the document contains information at a number of levels; for a brief introduction to the methods used I would recommend reading sections 2 and 3 as they provide a basic conceptual background to Potential and Ab Initio methods and try to explain the differences between these two approaches. The subsections contained within section 3 give a somewhat more detailed introduction into different Ab Initio approaches. In The "Challenges In Molecular Modeling" section we have tried to outline some of the major problems which face modelers and have presented some of the solutions which have been developed, this section is by no means exhaustive but is intended to give a general feel for the kinds of work which modelers do. Finally more technical points are presented in the appendices, these sections may really only be of interest to those , who for their sins, may be forced to do some modeling. A glossary of Computational packages and codes and what they do is also provided.

I also wish to apologise in advance for any bias which is apparent towards my own field of research in Ab Initio modeling of solution phases.

Keith Butler

Contents

2   Potential Models

Potential models also referred to as classical simulation methods are able to consider a large number of atoms by treating them as point charges and therefore not accounting for the motion of the electrons, for this reason they are also commonly referred to as atomistic models. This is an application of the Born-Oppenheimer approximation, which states that nuclear motion is so slow compared to electronic motion that one need only consider the nuclear positions, as the electrons will quickly adapt to the nuclear movement. It is this approximation that makes calculations involving thousands or millions of atoms feasible. The system will be described by force fields that detail how the energy changes as a function of the atomic positions. In its simplest form a molecular mechanics method will treat a molecule as an atomic point charges with bonds represented by springs. A forcefield is divided into both the intramolecular and intermolecular potentials, and although the choice of a forcefield is arbitrary, it must parameterised to reproduce either experimental properties or first principles calculations and be as transferable as possible. The typical form of a forcefield is:

$$E_{total} = E_{electrostatic} + E_{vanderWaals} + E_{bondstretching} + E_{bondbending} + E_{torsional} \quad (1)$$

where the first two terms are intermolecular terms and the rest intramolecular. The terms for the intramolecular potentials are the two-body bond stretching term, three-body bond bending terms and the four-body torsional term. These terms equate to a consideration of how far the atoms have deviated from the ideal equilibrium value for the bond length or angle. The three-body intramolecular term for the Si-O-Si angles is particularly important in accurately reproducing the angles within zeolitic structures[2]. The form of some of the most widely used potentials is presented below.

Bond Stretching : $E_{ij} = \frac{1}{2}k_{ij}(l - l_0)^2$
Bond Bending : $E_{ijk} = \frac{1}{2}k_{ijk}(\theta - \theta_0)^2$
Torsional : $E_{l\omega} = k(l - l_0)cos\omega$

Where l, θ and ω represent distance, angle and dihedral angle respectively and a subscript 0 denotes equilibrium value.

The non-bonding intermolecular terms consist of the Coulombic electrostatic interactions and the van der Waals forces. The Coulombic electrostatic potential is particularly challenging, because it is such a long range force, with an r^{-1} dependence on distance:

$$U_C = \frac{1}{2}\left(\frac{1}{4\pi\epsilon_0}\right)\sum_{ij}^{N}\sum_{\mathbf{n}}{}'\frac{q_iq_j}{|\mathbf{r}_{ij} + \mathbf{n}L|} \quad (2)$$

where q_i and q_j are the charges of the interacting ions i and j, r_{ij} is the distance between them and ϵ_0 is the permittivity of vacuum.Since the summation of these long-range

interactions is only conditionally convergent, the Ewald summation method is commonly used. This divides the summation into one series in real space for the cell and another series in reciprocal space for the periodic images, since these series converge the overall electrostatic energy becomes computable. The Ewald method is described in more detail in the appendix.

An additional level of sophistication may be added by employing the Dick and Overhauser shell model to reproduce the polarisability of the oxygen ion within zeolites. In this model the the nucleus and the electron cloud are approximated as core and shell point charges connected by a harmonic spring [3]. This model improves the accuracy of the calculation, however as it splits each oxygen point charge in two it increases the complexity of the calculation and thus the time needed to complete it.
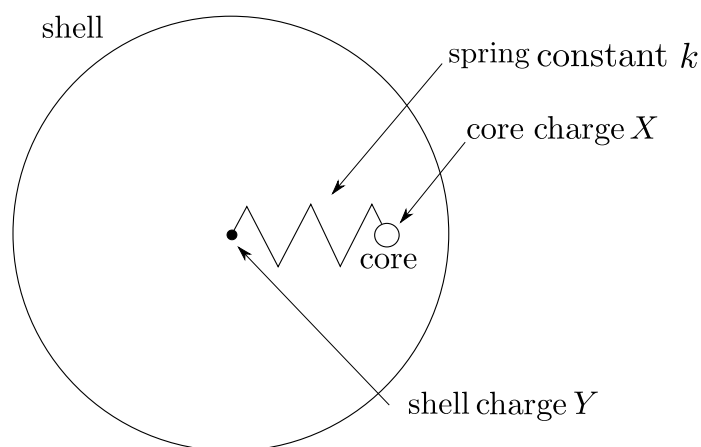

Figure 1: The shell model

3   Ab Initio Methods

Ab Initio is a Latin term meaning from the beginning. A calculation is said to be "ab initio" (or "from first principles") if it relies on basic and established laws of nature without additional assumptions or special models, this sets these methods apart from so called empirical methods which rely on experimentally determined parameters to calculate the properties of the system being studied.

For example, an ab initio calculation of the properties of liquid water might start with the properties of the constituent hydrogen and oxygen atoms and the laws of electrodynamics. From these basics, the properties of isolated individual water molecules would be derived, followed by computations of the interactions of larger and larger groups of water molecules, until the bulk properties of water had been determined. Whereas a "classical mechanics" or empirical calculation would treat the water a series of point charges connected and interacting through a series of potentials which have been parameterized to fit to experimental data.

The essential question at the heart of choosing between ab initio and empirical methods, as with most decisions in computational chemistry, is that of how much accuracy is required and how large are the computational resources available for tackling the problem. This is referred to as "scaling" and is demonstrated in figure 1. The important factors when estimating how a problem will scale are the order with which it scales O, and the prefactor with which it scales P, both with respect to the number of particles in the system. Empirical models typically scale to the order of 1 ie. the CPU time needed for N particles is proportional to N, whereas ab initio models typically scale with higher orders of $O = 2$ up to $O = 8$ for some of the highest level methods.

While empirical models have enjoyed enormous success in modelling chemical systems a number of drawbacks exist which necessitate the use of more expensive ab initio models. While empirical models are designed to behave well for the set of data (the training set) used to derive the parameters this is absolutely no guarantee that the model will behave well beyond this training set. Among the systems where empirical models can have difficulties are systems where there are many types of chemical interaction, requiring many model potentials to describe them and systems where the bonding pattern of the system changes quantitatively. An example of this is the fact that in general empirical models cannot model chemical reactions, although some recent models have claimed to overcome this barrier [4] [5].

Ab Initio methods are also often referred to as "Electronic Structure Methods" as they rely on solving the electronic state of the system, that pertaining to the electrons, to calculate various properties of the system. Electronic structure methods apply the laws of quantum mechanics[6][7] to obtain properties of chemical species. The basis of these methods is to attempt to solve the Schrödinger equation:

$$H\Psi(r,t) = \frac{ih}{2\pi}\frac{\partial \Psi(r,t)}{\partial t} \tag{3}$$
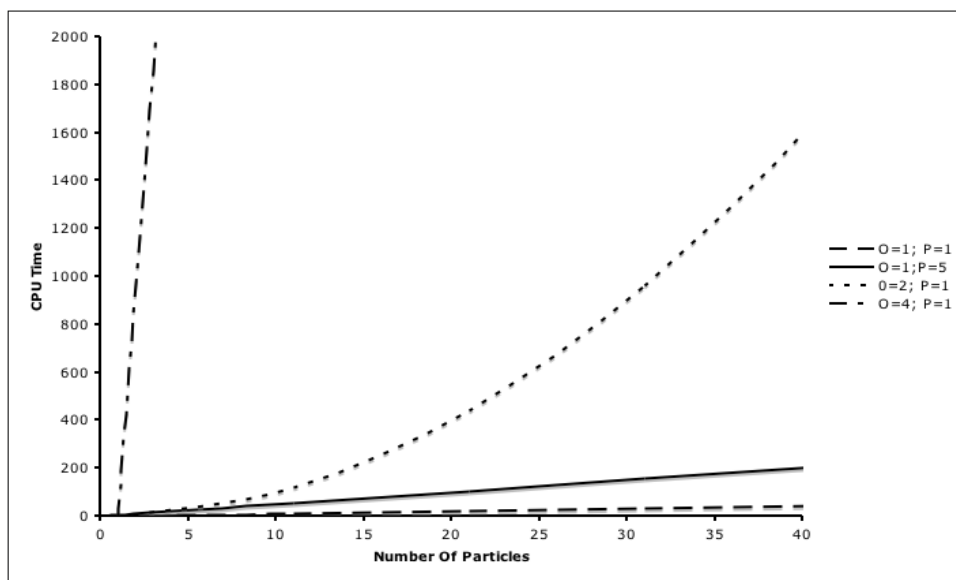
6

Figure 2: CPU time versus number of particles for a series of orders O and prefactors P; Time $= PN^O$.

where H is the **Hamiltonian operator** of the system, which is represented by the wave function Ψ, and which has a probability distribution $|\Psi^2|$, this is the probability that the system will be in a given state. If we consider that the systems potential energy, V, is time independent, we obtain the time-independent Schrödinger equation:

$$H\Psi = E\Psi \tag{4}$$

in which the energy E is the **eigenvalue** of the Hamiltonian operator. To aid calculation these eigenvalues, the systems Hamiltonian can be split up into different contributions, the Hamiltonian for this equation is:

$$H = T^e(r) + T^n(R) + V^{n-e} + V^e(r) + V^n(R) \tag{5}$$

where the superscripts n and e refer to the nucleus and electrons respectively, V is potential energy, T is kinetic energy and R and r refer to nuclear and electron position respectively. A number of approximations are necessary if we are to solve the Schrödinger equation. The first important approximation which is made is the Born-Oppenheimer (BO) approximation. The BO approximation is that given that electrons are much lighter than the nucleus and therefore move much more rapidly the motions of the nucleus and electrons can be decoupled, the electrons are assumed to be always equilibrated and thus forces on the nuclei arise only as a result of atomic positions. Thus the electronic structure part of

7

the problem is reduced to that of solving the time-independent Schrodinger equation, with the electrons assumed to be in the ground state. This removes the second term from the above equation, yielding the electronic Hamiltonian, H^e.

The next approximation necessary relates to the wave function and is called the "Linear Combination of Atomic Orbitals" (LCAO) approximation. Essentially we assume that the total wave function Ψ, can be represented by a summation of smaller functions called basis functions ϕ :

$$\Psi = \sum_i c_i \phi_i \tag{6}$$

where ϕ_i are the atomic orbitals and c_i is a factor weighting their overall contribution to the molecular orbital (MO), Ψ. The choice of the set of basis functions, called the basis set, to be used in a calculation is of crucial importance in electronic structure calculations. This relates to both the form of the basis functions and also to the number of functions to be used. The minimum amount of basis functions to be used is that which can accommodate all the electrons of the system, however for more accurate and sophisticated calculations basis functions representing unoccupied orbitals must be employed, once again which basis set to use boils down to a question of accuracy versus computational resources available.

We can now consider two of the major electronic structure methods , Hartree-Fock and Density Functional Theory.

3.1 Hartree-Fock

Since we do not know the exact ground state solution, to solve the Schrödinger equation we have to obtain the set of constants c_i, using the **variational principle**, which states that the calculated energy will always be higher than exact solution. Hartree-Fock uses the variational principle to obtain the constants, c_i which allows us to solve the Roothan matrix equation,

$$FC = SC\epsilon \tag{7}$$

$$det|F - \epsilon_a S| = 0 \tag{8}$$

where F is the **Fock matrix**, C is the matrix of constants c_i, S is the matrix of the overlap orbitals and ϵ_a is the matrix of the energy levels of the system. The Hartree-Fock method utilizes a Self-Consistent Field (SCF) procedure to solve the variational principle equations. In an SCF the first calculation involves an educated guess for the values of c_i, this is then used for a calculation of the Fock matrix, the resulting Fock matrix is then used to generate a new better set of coefficients, this is repeated until the difference between the new and the old fock matrices is below a certian threshold value, called the convergence criterion. This procedure is represented schematically in figure 3. The choice of convergence value is another important consideration when setting up a calculation.
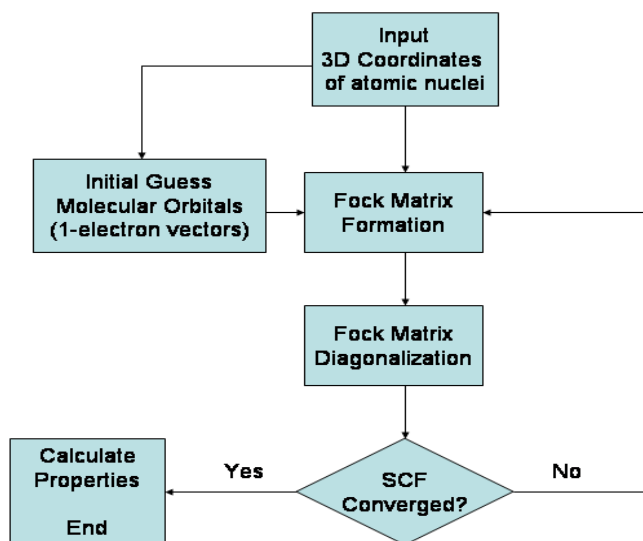
Figure 3: A Schematic Representation of The Hartree-Fock SCF Procedure.

3.2   Post Hartree-Fock Methods

The Hartree-Fock method suffers from the fact that it approximates the many electron problem as a one electron problem and implies that each electron sees the other electrons as a mean-field. Many attempts have been made to include electron correlation effects in electronic structure calculations.

The configuration Interaction (CI) method takes into account mixing of possible electronic states of the molecule in the form,

$$\Psi = b_0 \Psi_0 + \sum_s b_s \Psi_s \tag{9}$$

where $b_0\Psi_0$ is the Hartree-Fock expression and the second term on the RHS takes into account substitutions of virtual or excited states for occupied orbitals.

Another approach is the Møller-Plesset theory (MP)[8]. MP uses a perturbation on Hartree-Fock theory, to remove the error that was introduced in Hartree-Fock theory when the two electron integrals were replaced by a one electron potential. The two electron integrals are restored. Perturbation theory makes use of expanding the equation in a Taylor series with respect to the Lagrange multiplier λ. The new Hamiltonian is

$$H = H_0 + \lambda V \tag{10}$$

where H_0 is the Hartree-Fock Hamiltonian. Expanding up to the first order gives the Hartree-Fock solution. Including second order corrections leads to MP2. MP3 and MP4

are also commonly used methods. The MP series does not often converge, because in going from MP2 to MP3 and on to MP4, one does not necessarily get better approximations to the ground state wave function. This is because perturbation theory is not variational, and thus a calculated energy is not necessarily an upper bound to the true ground state energy.

Post Hartree Fock methods, however, are often simply not practical for calculations. For a system of N atoms MP2 and MP4 scale as N^5 and N^7 respectively while CI methods CISD and CISD(T) scale as N^6 and N^8 respectively. Compared to N^4 scaling for Hartree-Fock calculations [9].

3.3   Density Functional Theory (DFT)

The central focus of DFT is the electron density,ρ and not the wave function as in Hartree-Fock methods. It is called a functional theory as the energy is a function of the density, $E[\rho]$, which in turn is a function of the position, $\rho(r)$. A functional is a function of a function. The attraction of DFT is that it provides better results than Hartree-Fock, in many cases comparable with the MP methods, whilst scaling as N^4.

The basis for DFT is the *Reductio ad absurdum*[1] by Hohenberg and Kohn[10] that the ground state electronic energy is completely defined by the electron density, ρ, this proof is provided in an appendix. The foundation for the use of DFT methods in chemistry is the introduction of orbitals as suggested by Kohn and Sham (KS)[11]. Within these DFT formalisms the energy is split up as follows:

$$E_{DFT}[\rho] = T_s[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \tag{11}$$

with $T_s[\rho]$ the electron kinetic energy, $E_{ne}[\rho]$ the potential nuclear-electron interaction energy, and $J[\rho]$ the potential electron-electron interaction energy, the final term $E_{xc}[\rho]$ is the exchange-correlation energy which is calculated with an exchange-correlation functional. The choice of this last functional is where different DFT methods diverge and is crucial to the success of the method.

In the simplest cases the functional is the Local Density Approximation, this approximation however tends to break down if the electron density is not close to homogenous. The next level of approximation is the Generalized Gradient Approximation (GGA) which includes the first derivative (the gradient) of electron density when calculating $E_{xc}[\rho]$. Commonly known GGA methods include PBE[12] and B88[13]. After GGAs are the meta-GGAs which include second derivatives of electron density in calculations. Finally there are also what are known as hybrid functionals, which include a certain percentage of the exact exchange from Hartree-Fock theory. The most widely used, and very successful, hybrid functional, is B3LYP[14, 15]. The major criticism of hybrid DFT methods is that the

---

[1]Reductio ad absurdum: the premise that something impossible is actually true, and then proving it is ludicrous.

percentage of HF exchange used is defined by a parameter obtained by fitting to experimental data, so for example B3LYP which was parameterized based on heats of formation works well for energies but not necessarily well for other properties such as nuclear shielding constants.

All DFT functionals have problems reproducing physical and intermolecular interactions, like the van der Waals force. The reason for this misbehaviour is the incorrect asymptotical behaviour of the DFT interaction energy between multiple closed shell systems. DFT is a local or short-range potential, whereas intermolecular interactions are long-range. Corrections have been proposed to enhance the ability of DFT to model dispersion forces such as the DFT:MP2 hybrid approach[16]. Improvements for modelling dispersion have also been proposed by the group of Lundqvist and Langreth [17], these corrections are based on a double local density approximation. Yet another method for improved modeling of dispersion is the use of a damped dispersion function[18],

$$E_{tot} = E_{DFT} + E_{Disp} \tag{12}$$

with E_{Disp} given by a dispersion function which is dampened to remove unrealistic behavior at small distances.

4 Challenges In Molecular Modeling

The methods described thus far have been hugely successful when calculating the properties of systems in the gas phase at zero kelvin, however this clearly represents only a very small subspace of the chemical reactions and systems of interest practically. In order to be able to model "real-life" systems a number of ingenious methods have been developed, only a small fraction of which will be considered here. The two major challenges with which we shall deal here are those of molecules in solution and the search for the absolute minimum energy of chemical systems.

4.1 Molecules In Solution

When representing solvation two approaches are available, one is to model the solvent implicitly, ie to include the solvents effects directly when solving the Schrödinger equation of the system. The second is to include the solvent explicitly, ie. to include actual molecules of solvent surrounding the system of interest, of these two methods the second provides a physically more realistic scenario, however it introduces extra particles into the system, and given scaling constraints can quickly become infeasible as the number of solvent molecules rises. It is also possible to consider a hybrid of the two methods in which the molecule is surrounded by one layer of solvents which in turn includes an implicit model.

4.1.1 Implicit Solvation

The origins of an implicit model for solvation effects can be traced back to Born (1920) and Onsanger (1936). Born[19] derived the free energy of solvation for placing a charge within a spherical cavity in a solvent, Onsanger [20] extended this to a dipole within a spherical cavity. The Born model calculates solvation energy as the work done in moving a point charge from vacuum to a spherical cavity within a continuum. It is an extremely simple model however can be effective in the case of the solution of ions.

The Onsanger model is appropriate for many more species, in this model the solute dipole in side a cavity is considered. The dipole of the solvent induces a dipole in the surrounding continuum, which results in an electric field inside the cavity, this is called a reaction field. The reaction field then interacts with the solute dipole providing additional stabilization for the whole system. The reaction field model can be incorporated into quantum mechanical calculations where it is commonly known as the "Self Consistent Reaction Field" (SCRF) model. The SCRF model has been refined by Menucci and co-workers [21, 22] to use a cavity of the shape of the molecule which is built from a series of spheres centred on each atom, the solute charge is then represented as a series of charge points spread out on the resulting cavity surface.

SCRF models have been applied successfully to investigate the effects of solvation in many chemical systems and for various properties, however there are certain cases for which the continuum representation of the solvent is not sufficient, for example if hydrophobic

effects are present the continuum model will not represent them, another example is where specific solvent-solute interactions are important such as hydrogen bonding in water. In such cases it may be necessary to use explicit methods.

4.1.2   Explicit Solvation and Perdioc Boundary Conditions (PBCs)

If one is to model solvation explicitly by including "real" solvent molecules then an obvious question arises, how many explicit solvent molecules are necessary for a realistic representation, as the solvent molecules are obviously themselves "solvated" by other solvent molecules and these effects may be of importance. In some respect by including only a finite number of solvent molecules one is in effect only modeling a bubble of molecules within a vacuum. An elegant and relativly simple solution to this problem is to use Periodic Boundary Conditions (PBCs). This allows a relatively small number of particles to behave as if they are part of a bulk system. The molecules are placed inside a cubic box of given dimensions, this box is then surrounded by identical boxes and so on, a 2D representation of this system is given in figure   4. The coordinates of the particles in an image box can simply be calculated by adding or subtracting an integer to the coordinates of the original box, should a particle leave the box it will be replaced by its image entering from the other side, thus the contents of the box remain constant. As represented by the dashed lines in
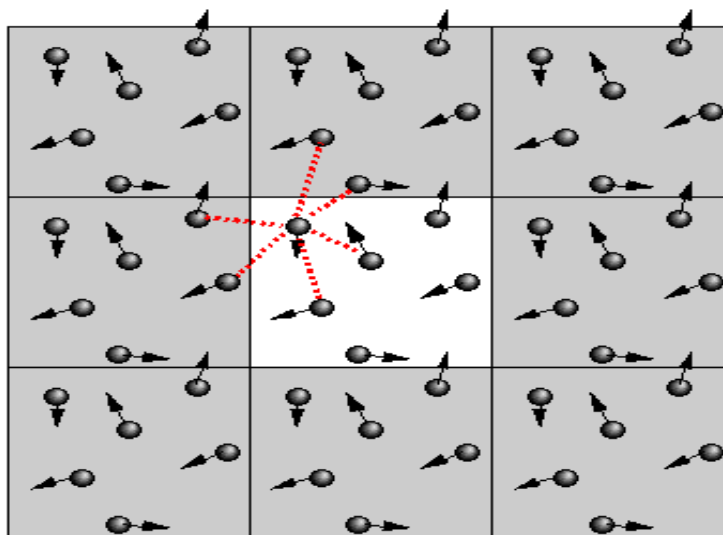

Figure 4:   A representation of a 2D system surrounded by periodic images.

figure   4 the particles in the original box now interact with particles all around them, all of this is achieved with relatively little computational cost, an certainly not of the order of just adding the extra molecules.

It should be noted that the use of PBCs is not restricted to liquid phases, indeed the technique was developed in the solid state modeling community and is particularly useful when modeling crystal structures as a periodic repetition of a basic unit cell, see sections 4.3 and   5.

4.2   Energy Minimization

The energy minimization problem is this: given a system with energy E which depends on independent variables $x_1, x_2, \ldots, x_i$, find the values of x_i for which E is a minimum. At a minimum point the first derivatives with respect to x_i of E are zero and the second derivatives are positive:

$$\frac{\partial E}{\partial x_i} = 0; \frac{\partial^2 E}{\partial x_i^2} > 0 \tag{13}$$

For standard analytical functions these values may be found by calculus methods, however given that the energy in chemical systems is dependent on the cartesian coordinates of the atoms involved the energy functions of interest are much more complex and must be found by numerical methods. Numerical methods in principle involve varying the coordinates until these conditions are met. This still seems like a simple enough proposition, however there exists another, more challenging problem, as represented in figure   5, it is this, in a potential energy landscape there generally exists more than one point where the criteria for a minimum are met. In this case we have both local and global minima, the global minimum being the minimum of the minima. The question now is how can we be sure that the minimum which we have found is a global minimum. The answer in short is we can't, not without exploring the full energy landscape. However if we can be sure that we have sampled enough of the potential energy landscape then we can be confident in our global minimum. So now we must sample the energy landscape to a suitable extent, this is not a trivial problem given the expense incurred in electronic structure calculations. In many cases it is more practical to sample the energy landscape using cheaper molecular mechanics methods followed by an electronic structure refinement of the lowest energy configurations from these calculations. However as computer resources improve and more sophisticated and cost effective electronic structure methods are developed the goal of full energy landscape exploration by Ab Initio methods is becoming more feesible. As molecular dynamics have been dealt with else where we will not look at them here but rather explore yet another restricting factor in energy landscape sampling.

This sampling problem is confounded by the energy barriers which exist in the energy landscape, it is entirely possible that during the course of a dynamics simulation a system may become stuck in an energy well which is not the global minimum, but which has an energy barrier too high to allow escape. There are many methods to overcome this issue in the next section we will look at one of the more recent and promising approaches.
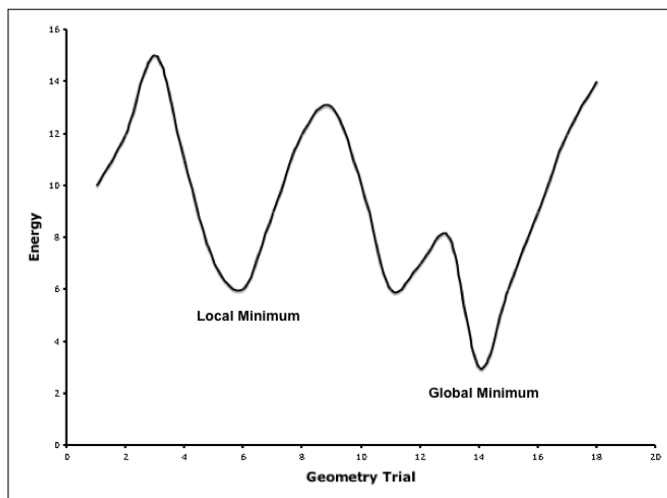
Figure 5: A 2D schematic representation of the global versus local minimum problem.

4.2.1   Molecular Dynamics (MD)

The energy minimisation technique does not take into account the kinetic energy of the species. It gives good results for the system with a temperature of zero Kelvin, even though the zero point energy is neglected. However, this approach does not include the influence of the temperature, therefore the system does not evolve in time and is likely to stay in the local minimum of energy. In the molecular dynamics method, the species have kinetic energies and therefore can mimic the situation of a real experiment in a better way. The system is not trapped in the local minimum but can evolve to reach a more stable configuration.

An MD simulation is performed in repetitive steps. However, firstly the starting configuration of the system needs to be set up. The positions of ions are initially obtained from a simple minimisation and velocities are given randomly with a distribution producing the required simulation temperature.

The next step is to calculate the force applied to each of the particles. It is obtained from the interaction energy $U(\mathbf{r}_i)$, provided by the potential model:

$$\mathbf{F}_i = -\nabla U(\mathbf{r}_i) \tag{14}$$

Having the forces, the new particle positions and velocities are obtained from the Newton's equations of motion:

$$\mathbf{a}_i(t) = \frac{\mathbf{F}_i(t)}{m_i} \tag{15}$$

15

The numerical solution of these equations leads to the scheme:

$$\mathbf{v}_i(t + \delta t) = \mathbf{v}_i(t) + \mathbf{a}_i(t)\delta t \tag{16}$$

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t \tag{17}$$

where $\mathbf{a}_i$, $\mathbf{v}_i$ and $\mathbf{r}_i$ are acceleration, velocity and position of ion i, respectively, and δt is the time step. The value of the time step δt determines the accuracy of the result, when it tends to zero the approximated solution tends to exact value. However, if the time step is too small the CPU time of the calculation is too long. In practice, the δt value is a compromise between accuracy and speed of the simulation. The size of δt is limited by the frequencies of the molecular vibrations occurring in a modelled system – the time step must be less then reciprocal of the highest frequency. Every step, after the new positions and velocities are obtained, the properties of the system are calculated. In the first several tens of time steps, the equilibration period is performed. This is necessary because the initial positions and velocities are usually far from the equilibrium ones and, therefore, in this period the velocities are scaled to obtain the required temperature.

4.2.2 Monte Carlo

Monte Carlo (MC) methods unlike MD are not a deterministic technique, instead making use of random numbers to perform random moves and orientation changes to generate successive configurations. The particular advantage of MC methods is that completely unrelated configurations are generated, so it is potentially possible for the entire region of phase space to be sampled. By comparison MD methods may only sample a very small portion of the phase space close to the starting configuration unless the simulation is run for a long time or special techniques are used to overcome barriers. The simplest form of a MC simulation accepts any new configuration with a lower energy than the previous configuration, however this is an inefficient way to sample the system, since high and low energy states are generated with equal probability. Therefore sophisticated MC methods, such as the Metropolis algorithm, make use of importance sampling to ensure that the majority of the time is spent sampling the low energy configurations that are of most interest. In order to do this the configurations are weighted by the Boltzmann factor,$\beta = exp(-\Delta v/k_b T)$, which ensures higher energy configurations have a lower probability of being sampled. The Metropolis algorithm will ensure that each step only depends on the configuration of the previous step and not on any earlier configurations, hence generating what is known as a Markov chain of states. Figure 6 is a schematic which demonstrates the Metropolis Method. When a trial move is made through the calculation of a random translation in each of the x, y and z coordinates, the energy is calculated and hence the change in energy, $\Delta v = E_n - E_m$, from the previous configuration. If Δv is negative then the move is automatically accepted as it is downhill. If Δv is positive then the Boltzmann factor, β, is compared to a random number between 0 and 1. If the Boltzmann factor

16

is greater than the random number then the move is accepted, otherwise it is rejected. This process is then iterated. The greater the temperature (i.e. the greater the thermal energy) of the system, the more likely it is that an uphill move will be accepted, since the Boltzmann factor is larger.
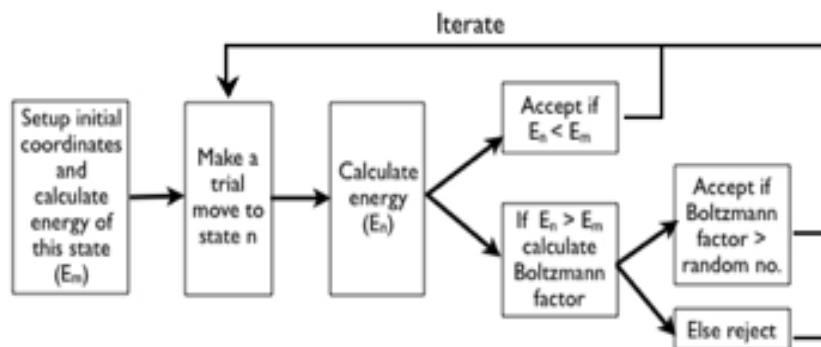

Figure 6: Schematic demonstrating how a Metropolis algorithm performs a MC simulation.

4.2.3   Metadynamics[1]

Traditional Molecular Dynamics and Monte Carlo methods have had a great impact in many fields, however due to computational costs there exist many systems for which "straight-forward" dynamics simply cannot explore the full landscape, consider the landscape presented in figure   5. If the system were to become trapped in the local minimum well, and did not have sufficient energy to overcome the barrier to get out the resultant simulation would sample only a small subspace of the system configurations and would miss the global minimum. In a traditional dynamics the configuration would propagate according to an equation something like:

$$\sigma_i^{t+1} = \sigma_i^t + \delta\sigma \frac{\phi}{|\phi|} \tag{18}$$

in which σ_i^t represents the configuration at time t, and the second term on the right hand size is related to calculating the force on the system. If this force becomes too great as the system attempts to leave the energy well it will simply slide back down the side of the well and not sample the space beyond this. The idea behind metadynamics is to build into this force a history dependent term, a kind of memory in order to avoid resampling configurational spaces. this is achieved by placing small gaussian functions at points which have already been visited, this is illustrated in figure   4.2.3 (a). The result is that as the calculation continues the wells become filled with these gaussian functions and the

energy land scape flattens out,   4.2.3 (b). The metadynamics procedure at the same time as allowing escape from local minima also provides for a backwards analysis to reveal the original energy surface, thus no information is lost and sampling power is greatly increased.
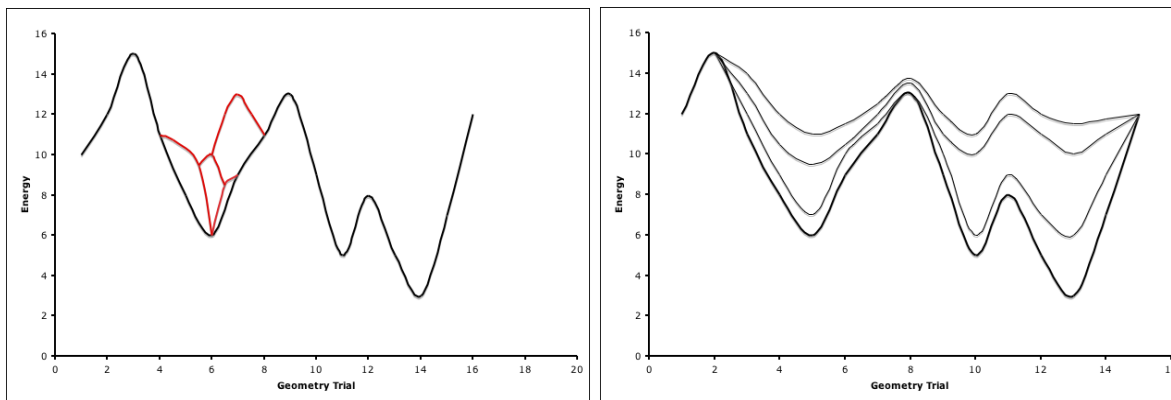

Figure 7: (a) on the left the energy profile being filled with small gaussian (red), (b) the energy profile becoming increasingly flattened as the simulation proceeds (thin lines).

4.3 Modeling Surfaces

A key part of understanding zeolite crystal growth is to look at what is happening at the surfaces. To model the processes occurring at surfaces we first need to have a reasonable model of the process. Since we are considering a crystalline (i.e. periodic) solid we will automatically require the use PBCs. There are 2 key ways we can model a surface:

1) **Periodic (3D):** Use PBCs in all three dimensions. It will consist of a slab (continuous in, say, the x and y directions) with a vacuum or some other medium between the top and bottom surface in the z direction. The size of the medium gap between the top and the bottom of the slab must be sufficiently large that the two surfaces are not chemically aware of each others presence.

2) **Aperiodic (2D):** Use PBCs only in the plane of the crystal (eg. x and y plane).

In both cases you need to ensure that your slab is the correct size. For example, imagine modelling a molecule approaching a zeolite surface, and you are using a 3D periodic cell to represent this surface. If the crystal slab is too thin, then the molecule will not only interact with the top surface, but also with the bottom surface. However we do not want to make any slab excessively thick as this will increase the computational expense of modelling the system.

When we create a surface we will be starting from a model of the materials bulk unit cell, most often the coordinates used are those characterised by an experimental technique, eg. XRD. For zeolites we rarely have as much information on the exact nature of its surfaces,
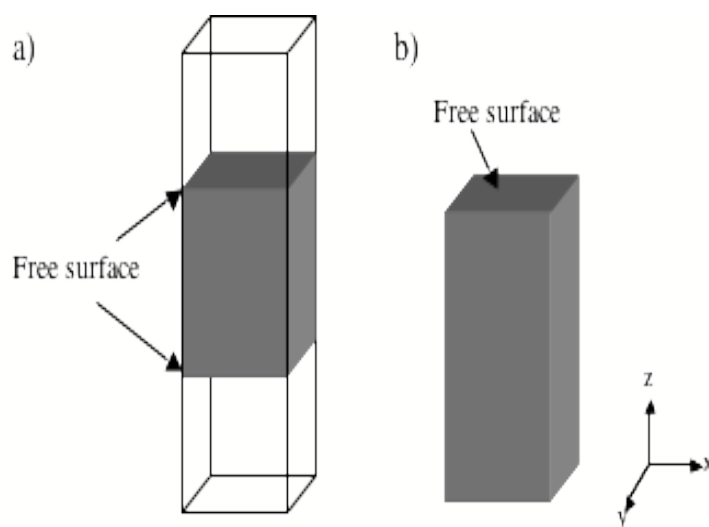
Figure 8: A representation of the two methods for surface representation; a) 3D periodic; b) 2D periodic.

although any information from HRTEM or other surface characterisation techniques is invaluable. Starting from the bulk unit cell we consider all the possible surface terminations for a particular plane. The most stable surface termination is particularly interesting since we can expect this to be the longest lived termination, hence its growth will determine the RDS of that surfaces growth and it is what we expect to be observed experimentally. Obviously, depending on what we are trying to model, it is often necessary to look at other viable surfaces as well. As a simple rule the most stable surface should be the surface that requires the least chemical bonds to be broken when it is cleaved from the bulk cell (as the more dangling bonds a surface has the less stable it would be). Another more subtle point is that a surface should be chosen so as to minimise/remove any dipole running across it (a dipole that does not exist in the true crystal, but is created merely by the making of the surface model).

5   Case Study: ZEBEDDE

The computer program ZEBEDDE (Zeolites By Evolutionary *De Novo* Design) was developed by Lewis and Willock in 1996 with the aim of predicting templates for a given zeolitic structure[1,2,3]. The principle being that the better the fit of a molecule within the framework the better that molecule will actually template to a particular structure. The program was designed to build a template from scratch (de novo); the procedure used is summarised as follows; the program places a small fragment (seed) within the void of a zeolite (termed the host) and then carries out a series of actions upon this guest. These actions are:

[1] BUILD add on a new fragment (from a library of possible fragments).

[2] ROTATE rotate the last bond added.

[3] SHAKE move the guest a random distance (between set limits) in a random direction.

[4] ROCK rotate the guest a random amount.

[5] RING FORMATION take two end atoms and join them to form a ring.

[6] MINIMISE GAS PHASE Perform an energy minimisation (as if the zeolite host was not visible to the guest molecule).

[7] MINIMISE IN HOST Perform an energy minimisation (with the zeolite host visible to the guest molecule).

[8] TWIST this is a recent addition to ZEBEDDE that rotates a randomly selected bond.

When an action is carried out it can be testing to find out whether it is an improvement by comparing the current and previous energy, to see whether the action has resulted in a reduction in energy (i.e. it is an energy minimisation tool, see Section   4.2). More recently a Monte Carlo function was added into ZEBEDDE. This can be chosen instead of the energy minimisation so that the simulation can overcome energetic barriers (consider Figure   5).

ZEBEDDE is an atomistic technique (see Section   2), this is necessary because it needs to carry out typically 1000+ actions and often each simulation would need to be repeated many times to ensure you had sampled all possible answers. It would be too time consuming using electronic structure methods and this level of accuracy is not deemed necessary for the task of testing the fit. ZEBEDDE requires PBC (Section   4.1.2) to

be able to model the zeolite (since it is a periodic solid). To calculate the electrostatic (Coulombic) interactions it makes use of the Ewald summation method (Section   9), necessary due to the PBCs.

There are many of simulation parameters that can be used to affect how the simulation runs, for example: the fragments available for addition, the bias towards the addition at certain sites (ie. to influence the amount of branched vs linear molecules that would be grown), the minimisation tool used (for Actions 6 7), the bias towards certain actions over others, how often an action is accepted without testing etc etc.

ZEBEDDE is not limited to being used for the growth of templates. It may also just be used to take a given template and then carry out actions (only 3,4,6,7,8) so as to find the most energetically favourable position for the template within a particular void or on a surface (using a 3D periodic cell, see Section   4.3). Similarly ZEBEDDE does not have to be carried out on a zeolite and template, it can be used for any guest molecule within/on a host.

6   Case Study: Kinetic Monte-Carlo Method

6.1   The Model

The kinetic monte-carlo method randomly chooses a particular growth site based on certain probabilities, and the number of that particular site which are available, and adds a growth unit to this site. Using these probabilities the program can then be applied to calculate the rates, activation energies and Free energies of crystal growth. As we have substantial information realting to Zeolite A we have chosen this as our prototype model.

Thus far one type of growth unit has been implemented in the program in order to model the following:

- Zeolite A crystal growth program in 3D.

- Zeolite A crystal growth and dissolution program in 3D

- Zeolite A crystal growth and dissolution program on (100) face in 2D (monolayer spreading growth)

- Zeolite A crystal growth and dissolution program on (110) face in 2D (monolayer spreading growth)

The growth unit is modeled as 1.2nm cubes, which corresponds to the experimentally determined terrace height in zeolite A, comprising a sodalite cage and double four ring (D4R).

The simulation performs one growth unit addition or dissolution per iteration, and then updates the site type of the 24 nearest neighbors of the site which has been affected. The
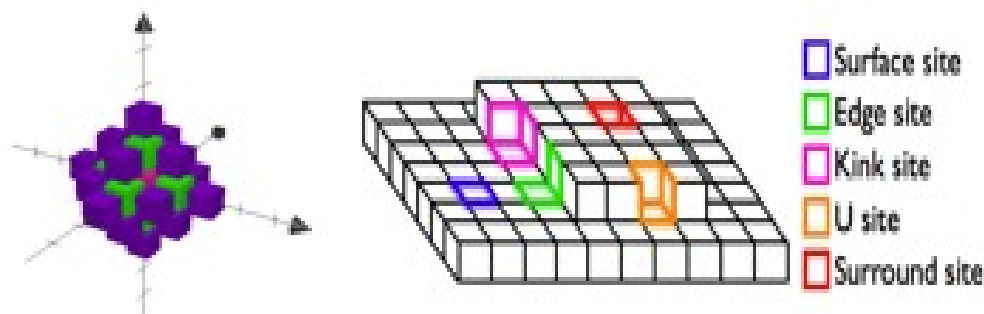
Figure 9: (a)Each cube shows a growth unit. Growth/dissolution site (pink) is surrounded by 6 nearest neighbours (green) and 18 second nearest neighbours (purple). (b) The Kossel model.

nearest neighbor sites are categorized as having either first or second sphere coordination to the affected site see Figure   9 (a) The chosen growth/dissolution site has six nearest neighbours, which have 5 categories corresponding to the Kossel model (Fig   ??) . In order to generate bevelled edges (ie, 110 faces) and/or corners (ie, 111 faces) of a cube, the second nearest neighbours are also considered. This results in 223 possible "site types" or categories, however for simplicity we consider only the six key site types.

6.2   Development

The challenge for the program is to control both morphology and topology using a certain set of probabilities. In 3D, although crystal morphology could be controlled,topology needed a certain shape to be more realistic.

6.2.1   Zeolite A crystal growth and dissolution program in 3D

The problem was it was hard to treat 192 site types (excluding bulk) to control the morphology and topology at the same time. Also, the small limited computer memory environment caused a further difficulty to compare simulation images with real crystal images. Hence, 2D simulation program on (100) face was created to cope with above problems.

6.2.2   Zeolite A crystal growth and dissolution program on (100) face in 2D

The simulation was focused on mono-layer surface spreading (ie, the top of the surface terrace is supposed to be simulated). NB: multi-layer shows 3D because the second nearest neighbours are considered. As a result, we could say the curved corners were generated under the dissolution; however, to decide the certain probability set, we need more data

(eg, more real images and their size of radius at the corner (Fig,3), and also the ratio between growth rate and dissolution rate).

6.2.3   Zeolite A crystal growth and dissolution program on (110) face in 2D

Zeolite A crystal growth and dissolution program on (110) face in 2D As we have observed the (110) faces by means of AFM and SEM, the 2D simulation program for (110) face was also created to decide the (110) surface topology by a certain probability set.

Probability sets generated by the 2D programs will be applied for probability sets on the 3D program, considering the ratio of site types to make the crystal morphology as well as surface topology.

7 Appendix A: Proof of the Hohenberg-Kohn Theorem

The Hohenberg-Kohn theorem demonstrates that there is a functional relationship between the electron density, ρ, and all observable properties of the interacting system of particles, meaning that all properties of the system may be determined if the electron density is known. We start by assuming that there exist two different potemtials v and v' which yield the same electron density ρ, these two external potentails generate different operators $\hat{V}$ and $\hat{V}'$ which in turn yield two different Hamiltonians $\hat{H}$ and $\hat{H}'$. These systems also will have different wave functions Φ and Φ' with energies E and E' the energies are given by :

$$
\begin{aligned}
E &= \langle\Phi|\hat{H}|\Phi\rangle \\
E' &= \langle\Phi'|\hat{H}'|\Phi'\rangle
\end{aligned}
\tag{19}
$$

Now by applying the variational principle we can say that the exact Hamiltonian for a system always yields lower energy than an approximate one thus:

$$
\begin{aligned}
E' &= \langle\Phi|\hat{H}|\Phi\rangle < \langle\Phi|\hat{H}'|\Phi\rangle \\
E &= \langle\Phi'|\hat{H}'|\Phi'\rangle < \langle\Phi'|\hat{H}|\Phi'\rangle
\end{aligned}
\tag{20}
$$

We can also say that

$$
\begin{aligned}
\langle\Phi|\hat{H}'|\Phi\rangle &= \langle\Phi|\hat{H} + \hat{V}' - \hat{V}|\Phi\rangle \\
\langle\Phi'|\hat{H}|\Phi'\rangle &= \langle\Phi'|\hat{H}' + \hat{V}' - \hat{V}|\Phi'\rangle
\end{aligned}
\tag{21}
$$

assuming that

$$
\langle\Phi'|\rho|\Phi'\rangle = \rho(r) = \langle\Phi'|\rho'|\Phi'\rangle
\tag{22}
$$

which is our original assumption, we obtain that

$$
\begin{aligned}
E' &< \langle\Phi|\hat{H}|\Phi\rangle + \int dr[v'(r) - v(r)]\rho(r) = E + \int dr[v'(r) - v(r)]\rho(r) \\
E &< \langle\Phi'|\hat{H}'|\Phi'\rangle + \int dr[v(r) - v'(r)]\rho(r) = E' + \int dr[v(r) - v'(r)]\rho(r)
\end{aligned}
\tag{23}
$$

Adding these two inequalities yields:

$$
E' + E < E + E'
\tag{24}
$$

This result is inconsistent and thus proves that the original assumption is *false*.

8   Appendix B: Basis Sets

It would be negligent when introducing Ab Initio methods to overlook an explanation of basis sets, which after the choice of method are the most important factor in setting up a calculation, and a continual source of controversy and debate.

As stated earlier basis sets are sets of basis functions which are combined to give a wave function. There exist two major types of basis sets, plane wave basis sets and atom centred basis sets. We will concentrate primarily on the latter as they are more commonly used in liquid and gas phase calculations. Atom centred basis sets are comprised of atomic functions representing the electron population at a given distance from the nucleus and can be thought of as being atomic orbitals. There are a number of different types of atom centred basis functions available. One obvious choice are Slater type orbitals (STOs) which are of the form:

$$R(r) = Nr^{n-1}e^{-\zeta r} \tag{25}$$

where n is the principal quantum number, N is a normalization constant, r is the distance from the nucleus and ζ is a constant related to the atomic charge of the nucleus. Unfortunately STOs are often impractical for quantum chemical calculations as some of the integrals are difficult or even impossible.

By far the most popular type of basis sets used in quantum chemistry are those based on Gaussian functions, Gaussian type orbitals (GTOs), these have the form:

$$R(r) = x^a y^b z^c e^{-\alpha r^2} \tag{26}$$

α determines the spread of the function, r is the distance from the nucleus, x, y, z are cartesian variables and a, b, c determine the order of the function. If $a + b + c = 0$ then the function is zeroth order, and one such function exists, if $a + b + c = 1$ the function is first order and three such functions exist, zeroth order functions are equivalent to s orbitals, first order functions to p_x, p_y and p_z orbitals and so on up the orders.

A minimal basis as stated earlier is that which contains just enough functions to accommodate all of the electrons present., however as stated earlier in practice minimal basis sets are often not good enough for the task at hand. The basis set can be expanded in a number of ways. One way is to split each of the functions in the basis set, if the functions are split once this is known as a double zeta basis set, basis sets up to quadruple zeta are commonly used. An alternative approach which is highly popular is to split only the basis sets used for valance electrons, the rational being that the chemical properties of interest are effected by the valance rather than the core electrons, this approach is known as the split valance approach and is so popular that it has its own notation, exemplified by the label 3-21G, this means that core electrons are teated with 3 functions, whilst valance electrons are treated with functions split into two contracted Gaussians and one diffuse Gaussian. Popular examples of this type of basis set are 3-21G, 4-31G and 6-31G.

Simply increasing the number of basis functions may not necessarily improve the model, and there are other issues which must be considered when issues such as orbital mixing

are involved, for example in an isolated hydrogen atom the electron cloud is spherical, however as the atom approaches another hydrogen atom this becomes distorted and the cloud takes on some p orbital character or is said to be sp hybridized. in order to account for this we introduce what are called polarization functions. Polarization functions have higher angular quantum numbers, thus the polarization function for a hydrogen atom corresponds to a p orbital function, and to a d orbital function for first and second row atoms. The use of polarization functions is denoted by an asterix. Hence the 6-31G basis with added polarization is the 6-31G* basis set.

A final complication which we will consider is that of species with lone pairs of electrons. These lone pairs reside far away from the nucleus and are poorly represented by Gaussian functions, as they decay quickly when moving away from the nucleus. The solution to this is to add highly diffuse functions. These are denoted by a +, so the 6-31G* basis set with added diffuse functions for heavy atoms is denoted 6-31+G*.

When modeling periodic systems basis sets of the lane wave form are usually chosen, the general form of such basis sets is:

$$\psi(r) = \sum_G a_G exp(i(k+G)\dot{r}) \tag{27}$$

In this case the basis function is continuous and not centred on the atom as with gaussian type basis sets.

There exist many other types of basis sets and methods for improving them, however for the purposes of this introductory document it is considered sufficient to discuss only the most commonly encountered basis sets.

9   Appendix C: Ewald Summation

The sum in equation 2 converges slowly because the electrostatic potential, due to point charges, decays as $1/r$. The way to overcome that was presented by Ewald [?]. In this method the density of point charges, represented by a sum of δ functions, is modified by an additional sum of diffuse charge distributions around each ion. The diffuse distribution is selected such that it is opposite in sign to the point charge and thus the total charge is cancelled out. As a result, the electrostatic potential due to a single ion is a fraction of original potential that is not screened by diffuse charge. This remaining fraction rapidly converges to zero at long distances, which allows the use of a direct summation for the screened charges. The screening charge distribution around ion i usually has the form of a Gaussian

$$\rho_g(r) = -q_i \left(\frac{\alpha}{\pi}\right)^{\frac{3}{2}} \exp(-\alpha r^2) \tag{28}$$

where r is position relative to the centre of the distribution and parameter α sets the width of the distribution. The interaction energy due to the screened charge distribution can be

calculated from following equation

$$U_1 = \frac{1}{2} \left(\frac{1}{4\pi\epsilon_0} \right) \sum_{ij}^{N} \sum_{\mathbf{n}}' \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}L|} \mathrm{erfc}(\sqrt{\alpha}|\mathbf{r}_{ij} + \mathbf{n}L|) \tag{29}$$

where erfc is the complementary error function

$$\mathrm{erfc}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\mathrm{d}t \tag{30}$$

obtained by solving Poisson's equation for Gaussian charge distribution. For large arguments this function tends to zero.
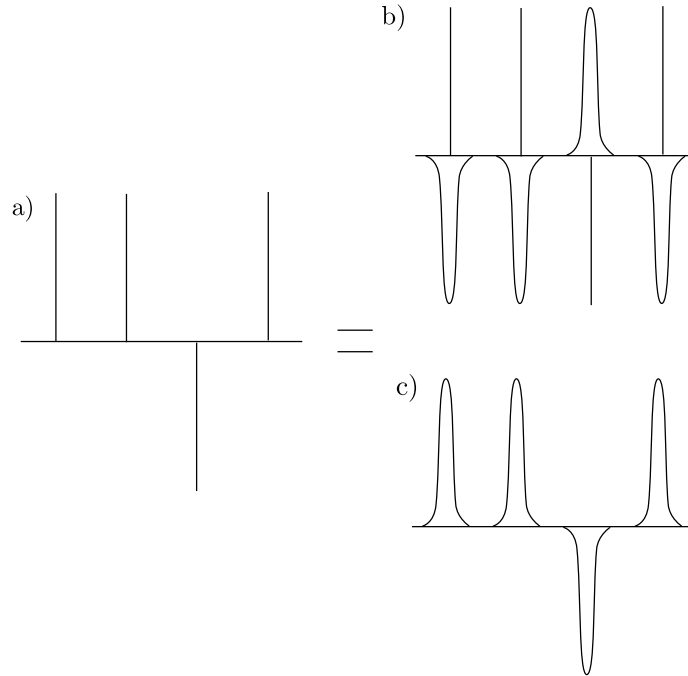

Figure 10: Scheme of the Ewald method; a) point charges; b) screened point charges; c) Gaussian charge distribution

The electrostatic potential due to screened charges does not describe the full interaction between point charges, therefore, a correction must be introduced. The method for correcting the interaction is shown in Figure 10. Point charges are surrounded by screening charges and are then compensated by a charge distribution smoothly varying in space. As the compensating charge distribution is a periodic function it can be represented by a

rapidly converging Fourier series. The density of the compensating charge is formulated by a sum of Gaussians:

$$\rho_1(r) = \sum_{j=1}^{N} \sideset{}{'}\sum_{\mathbf{n}} q_j \left(\frac{\alpha}{\pi}\right)^{\frac{3}{2}} \exp\left[-\alpha|\mathbf{r} - (\mathbf{r}_j + \mathbf{n}L)|^2\right] \tag{31}$$

which in reciprocal space becomes:

$$\rho_1(\mathbf{k}) = \frac{1}{V} \sum_{j=1}^{N} q_j \exp(-i\mathbf{k} \cdot \mathbf{r}_j) \exp\left(\frac{-k^2}{4\alpha}\right) \tag{32}$$

Applying Poisson's equation yields the electrostatic potential:

$$\Phi(k) = \frac{4\pi}{k^2} \frac{1}{V} \sum_{j=1}^{N} q_j \exp(-i\mathbf{k} \cdot \mathbf{r}_j) \exp\left(\frac{-k^2}{4\alpha}\right) \tag{33}$$

where $\mathbf{k} \neq 0$. The reciprocal space contribution becomes:

$$\Phi(r) = \frac{1}{V} \sum_{k \neq 0} \sum_{j=1}^{N} \frac{4\pi q_j}{k^2} \exp[-i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}_j)] \exp\left(\frac{-k^2}{4\alpha}\right) \tag{34}$$

and the contribution to the interaction energy is equal to:

$$U_2 = \frac{V}{2} \sum_{k \neq 0} \frac{4\pi}{k^2} |\rho(\mathbf{k})|^2 \exp\left(\frac{-k^2}{4\alpha}\right) \tag{35}$$

where $\rho(\mathbf{k})$ is defined as:

$$\rho(\mathbf{k}) = \frac{1}{V} \sum_{i=1}^{N} q_i \exp(-i\mathbf{k} \cdot \mathbf{r}_i) \tag{36}$$

As equation 35 includes the interaction between the charge cloud around an ion and the ion charge itself, another correction is introduced. The form of this self-interaction correction is evolved to be

$$U_s = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \sum_{i=1}^{N} q_i^2 \tag{37}$$

and does not depend on ions positions.

Considering all contributions, the Coulomb interaction is given

$$U_C = U_1 + U_2 - U_s \tag{38}$$

A detailed derivation of above equations can be found in Frenkel and Smit [23].

28

10   Glossary

De Novo Design The design of compounds by incremental construction of a model within a model of the crystal site.

Eigenvalue:A scalar value λ that permits nonzero solutions y in equations of the form
$Ly = \lambda y$;

where L is an operator and where y can represent a vector or a function that is subject to certain boundary conditions (eg that y is zero at a certain point).

Fock Matrix: Is defined by the Fock Operator, which is equivalent to a Hamiltonian for a single electron in a poly-electron system.

Hamiltonian: An operator representing the energy of the electrons and nuclei in a molecule. The Hamiltonian operator acts on the wave function yielding the energy.

Operator : a function, that operates on (or modifies) another function.

Potential Energy Surface: A plot of the potential energy of a system in relation to the parameters which define the potential energy. It can be visualized as a landscape in which the height changes with position, thus the energy would correspond tho height and the variables would correspond to location on the north south east west grid. For this reason it is also commonly referred to as a potential energy landscape. In reality in a chemical system the surface will be of more than three dimensions; given that each atom in the system has 3 degrees of freedom the surface will be in 3N dimensions, where N is the number of atoms

Variational Principle: The expectation value of the Hamiltonian for a trial wavefunction must be greater than or equal to the actual ground state energy. Or in other words:

$$E_{ground} \leqslant \langle \phi | H | \phi \rangle \tag{39}$$

By definition, the ground state has the lowest energy, and therefore any trial wavefunction will have an energy greater than or equal to the ground state energy.

Wave Function: A mathematical function for a state of a system, from a space that consists of the possible states of the system.

11   Glossary of Programs and Codes

Gaussian: A commercially available package for quantum and semi-empirical calculations.

GAMESS: Essentially the same as Gaussian but open-source, so somewhat less polished.

GULP: General Utility Lattice Program. Program to perform atomistic calculations for molecules and solids. It provides a variety of simulation approaches, mostly using

analytical techniques, and can be applied to a variety of systems eg molecule, surface, periodic solid.

METADISE: Atomistic code for modeling surface structure and stability.

SIESTA: Spanish Initiative for Electronic Simulations with Thousands of Atoms. Program to perform electronic structure (Density Functional Theory) calculations and Molecular Dynamics simulations of molecules and solids.

Quickstep (cp2k): Open-source code using a mixture of gaussian and plane wave basis sets for efficient molecular ab initio dynamics simulations of large systems.

ZEBEDDE: ZEolite By Evolutionary /De Novo/ DEsign. Program to design a template for a specific zeolite structure. In principle has wider use for locating/building any 'guest' within/on a 'host'.

References

[1] Laio, A. and Parrinello, M., *Proceedings of the National Academy of Sciences*, 2002, **99**(20), 12562–12566.

[2] Henson, N. J.; Cheetham, A. K. and Gale, J. D., *Chemistry of Materials*, 1994, **6**(10), 1647–1650.

[3] Dick, A. and Overhauser, B., *Physical Reviews*, 1958, **73**, 90–103.

[4] Nielson, K.; vanDuin, A.; Oxgaard, J.; Deng, W.-Q. and Goddard, W., *Journal of Physical Chemistry A*, 2005, **109**(3), 493–499.

[5] Cole, D. J.; Payne, M. C.; Csányi, G.; Spearing, S. M. and Ciacchi, L. C., *The Journal of Chemical Physics*, 2007, **127**(20), 204704.

[6] Szabo, A. and Ostlund, N., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory 2nd ed.*, Dover Publications, 1996.

[7] Atkins, P. and Friedman, R., *Molecular Quantum Mechanics 4th ed.*, Dover Publications, 1996.

[8] Møller, C. and Plesset, M. S., Oct , (1934), **46**(7), 618–622.

[9] Leach, A., *Molecular Modelling Principles and Applications, 2nd Ed.*, Prentice Hall: Harlow, 2001.

[10] Hohenberg, P. and Kohn, W., *Phys. Rev.*, 1964, **136**.

[11] Kohn, W. and Sham, L., *Phys. Rev.*, 1965, **140**.

[12] Perdew, J.; Burke, K. and Ernzerhof, M., *Phys. Rev. Lett.*, 1996, **77**, 3865.

[13] Becke, A., *Phys. Rev. A*, 1988, **98**, 3098.

[14] Becke, A., *J. Chem. Phys.*, 1995, **98**, 5648.

[15] C. Lee, W. Yang, G. P., *Phys. Rev. B*, 1988, **37**, 785.

[16] Tuma, C. and Sauer., J., *Phys. Chem. Chem. Phys.*, 2006, **8**, 3995.

[17] Hult, E.; Rydberg, H.; Lundqvist, B. I. and Langreth, D. C., *Phys. Rev. B*, 1999, **59**, 4708.

[18] Zimmerli, U.; Parrinello, M. and Koumoutsakos, P., *The Journal of Chemical Physics*, 2004, **120**, 2693.

[19] Born, M., *Z. Physics*, 1920, **1**, 45.

[20] Onsanger, L., *J. Am. Chem. Soc.*, 1936, **58**, 1486.

[21] Miertu˘ s, S.; Scrocco, E. and Tomasi, J., *Chem. Phys.*, 1981, **55**, 117.

[22] Cancés, E.; Mennucci, B. and Tomassi, J., *J. Chem. Phys.*, 1997, **107**, 3031.

[23] Frenkel, D. and Smit, B., *Understanding Molecular Simulations*, Academic Press, San Diego, 1996.