A First Principles Study of Defects in Diamond

Stephen John Breuer

A theoretical thesis submitted to the University of Newcastle-upon-Tyne for the degree of Doctor of Philosophy.

August 1994

Declaration

This thesis has not previously been submitted by the candidate for a degree in this or any other university.

S. J. Breuer

August, 1994

Abstract

In this thesis a methodology for performing *ab initio* local density functional calculations using a basis of localised orbitals and norm–conserving pseudopotentials in atomic clusters is presented. The equilibrium structure of the cluster is determined by the minimisation of the total energy with respect to the atomic positions. The formalism is applied to the study of selected aspects of some of the more common and important defects in diamond by the location of these defects near the centre of large hydrogen terminated clusters.

The microscopic structure and vibrational properties of the substitutional boron defect is investigated. The equilibrium structure of the substitutional boron when passivated by a hydrogen atom is then determined and studied. An examination of the stability of the various structures proposed for the self interstitial is carried out. The structural, electronic and vibrational properties of the only stable structure, the $\langle 100 \rangle$ split interstitial, are then investigated. Aspects of the neutral and negative vacancies are studied. The migration paths for the vacancy and self interstitial are examined and the barrier energies for defect migration along these paths are found. Finally, the reorientation energy of the unique extended nitrogen–carbon bond in the neutral substitutional nitrogen defect to one of the other carbon neighbours is determined.

Acknowledgments

I would like to thank the following people.

My supervisor, Dr. Patrick Briddon, for introducing me to *ab initio* techniques and for guiding my work during the course of my PhD. Our many discussions were invaluable. The other members of the Theory of Condensed Matter Group for help with many aspects of computing, especially the more baffling, such as the menu structure of unigraph.

My parents, Diego and Jennifer Breuer, for always supporting me in all that I decide to do. My father for repeatedly telling me not to go into catering throughout my teenage years. Working in restaurants was effective aversion therapy.

All my friends in Newcastle and elsewhere for a most enjoyable three years of my life.

Contents

Chapter 1

Introduction

The work presented in this thesis is concerned with the first principles study of defects in solids on the atomic scale, and more specifically with selected structures of interest in diamond. *Ab initio* calculations using local density functional theory have become an established theoretical tool over recent years. The number of atoms that can be modelled by this approach has reached and exceeded that for which semi–empirical calculations were used for many years. A major advantage that *ab initio* calculations have over these parameterised semi–empirical approaches is that no input is required from experiment. The only information required is the atomic numbers of the atoms modelled.

Returning to the avenues of theoretical atomic simulations pursued in the past, theories such as Hartree and Hartree–Fock have long been used for studying small molecules. However, they are computationally too demanding to use for larger structures. The first approaches at attempting to model higher numbers of atoms used semi-empirical methods such as Hückel theory, CNDO (Complete Neglect of Differential Overlap) and MNDO (Moderate Neglect of Differential Overlap). All of these approaches had inherent severe approximations and required the input of parameters which forced selected results to agree with experimental work. Most currently used techniques follow on from theories published in the 1960's by Hohenberg and Kohn (1964) and Kohn and Sham (1965). Many problems had to be overcome before they could be applied to the wide range of problems that they are today. For example, in the representation of the external potential of the atomic cores in which the valence electrons move, the development of norm–conserving transferable pseudopotentials was a giant step forward. Also, the formulation of practical expressions of sufficient accuracy for the exchange–correlation energy of an inhomogeneous electron gas under the now almost universally used local density approximation was a significant contribution.

Ab initio techniques have been applied to the study of many problems on the atomic scale. Defects have been modelled in such diverse materials as semiconductors, quartz, ice and metals. Simulations of surface reconstructions have been carried out and catalysis reactions at these surfaces have subsequently been investigated. Also there have been simulations of novel molecules such as C_{60} (Faulhaber 1993a) and bis(ethylenedithio)tetrathiafulvalene (BEDT-TTF) and tetrathiafulvalene (TTF) (Faulhaber 1993b). A series of papers over recent years such as those by Liu (1989,1990) and Cohen (1993) have addressed the possible predictive power of *ab initio* calculations. These papers have dealt with a theoretical material, $\beta - C_3N_4$, which it appears should be as hard or harder than diamond. Materials such as these have been termed 'superhard' and could have obvious technological applications. The predictions resulting from computer modelling have led to attempts to produce samples of the substance. Experimental groups report varying degrees of success in trying to manufacture useful quantities and forms, but there have been some encouraging results (Niu 1993). In the same vein, some quantities which cannot be experimentally measured have been determined. This was the case with C_{60}, where the vibrational frequencies of some infra-red inactive modes were calculated using a first principles technique in order to find the electron–phonon coupling strengths of the molecule (Faulhaber 1993a), an important parameter in theories of superconductivity.

Turning to the approach that we have used in performing *ab initio* calculations, the theoretical starting point of the formalism presented in this thesis is density functional theory, which was developed by Hohenberg and Kohn (1964). The single particle equations derived by Kohn and Sham (1965) are solved self–consistently in real–space under the local density approximation. This involves the expansion of the wavefunction of a cluster of atoms using a basis of localised Gaussian orbitals. As a calculational aid, an intermediate fit to the charge density is also expanded in terms of Gaussian orbitals. The norm–conserving, highly transferable pseudopotentials generated by Bachelet, Hamann and Schluter (1982) replace the cores of the atoms in the calculation. As the approach used in this thesis simulates a defect in a solid by modelling the defect in a large cluster of atoms, it is also readily applicable to modelling molecules. As stated above, the only input required is the atomic numbers of the atoms in the structure, which makes the calculations *ab initio* or first principles. Using the theory developed we then perform state of the art electronic structure calculations on atomic systems considered to be of interest and importance in diamond. Wherever possible our theoretical results are compared with experimental results, or we deduce a value for a property which should be measurable by an appropriate experimental investigation. We present the methods that we have used to determine such observables as vibrational frequencies and diffusion barrier energies.

In this thesis we exclusively apply our *ab initio* technique to the study of defects in diamond. The possible technological applications of diamond have rapidly multiplied since the advent of techniques capable of producing thin films of the material, the most important of which is currently CVD (Chemical Vapour Deposition). The material thus produced has a wealth of possible uses and much has been published on present and possible future applications. Two of the highly desirable and useful properties of diamond are that it is exceptionally hard and has a high thermal conductivity. Research continues into forming layers of diamond on different substrates. One engineering application of this would be the coating of surfaces prone to wear by this highly abrasion resistant material. However, as diamond is a semi–conductor, the electronics industry has focussed on its suitability for use in its products. The possibility of producing electronic devices and circuits based on a material with great strength and which also dissipates the heat produced with great efficiency has obvious advantages. An obstacle to this is that although the p–type material can be manufactured using boron as an acceptor impurity, attempts to produce the n–type material have so far failed. An obvious candidate for a donor impurity would be nitrogen, which is generally the dominant impurity in natural diamond. However, the structure formed by isolated substitutional nitrogen defects removes its electrical activity.

Two other uses which have been suggested for thin diamond films are as detectors in particle physics experiments and as dosimeters for radiation in medical physics. The ionisation of the atoms by radiation would be utilised in the first application. However, the high energy radiation would also cause structural damage to the material. Due to the stronger bonds in diamond there is less damage than that in present equivalent detectors using silicon, which is one reason for the attempts to utilise diamond in this way. In the second application, the strength of the luminescence or absorption of the optical centres produced by the radiation is used as a measure of the exposure of the material (and human tissue) to radiation. This optical luminescence is accepted as being associated with vibronic transitions at the vacancy. In order to reuse the diamond sample it is necessary to anneal out these defects and so remove the optical transitions. The total process is therefore analogous to the standard solid state physics procedure of irradiating and then annealing a material before carrying out infra–red or optical absorption experiments. It is not yet established which of these radiation produced defects most readily migrate and to what trapping points. Or, of course, if another migrating species is trapped at them.

1.1 A Summary of the Thesis

In Chapter 2 we initially revue a few of the more important approaches in the development of electronic structure calculations, Hartree theory, Hartree–Fock theory and Thomas–Fermi theory. We then proceed to derive the basic equations of density functional theory and the Kohn–Sham equations, which are the basis of the theory developed in this thesis. The local density approximation is explained and the development of pseudopotentials is discussed, culminating in the prescription used to generate the norm–conserving pseudopotentials used in this work.

In Chapter 3 we present the theoretical formalism that we use in our calculations. The wavefunction is expanded using a set of localised Gaussian–type orbitals in the cluster approach. An intermediate fit to the charge density is introduced in determining the Hartree and exchange–correlation energies in order to remove the calculational difficulty of evaluating four–centre integrals. The derivation of an expression for the total energy of a cluster of atoms is presented. We explain how we obtain a self–consistent total energy for a configuration of atoms. We then discuss our approach to determining the equilibrium structure of a cluster of atoms by minimising the total energy with respect to the atomic positions. Finally, several modifications and additions to the basic formalism are presented; the incorporation of spin polarisation, applying a finite temperature and the use of symmetry.

In Chapter 4 we derive the expressions that we use to determine the frequencies of vibrational modes of defects in the harmonic approximation. The approaches used for localised vibrational modes and more delocalised resonances are presented. Expressions for the phonon density of states and the induced dipole moment of a defect are derived in terms of the perturbed lattice Green function.

In Chapter 5 we determine the microscopic structure of the important acceptor impurity, boron, located at the substitutional site in the lattice. We also investigate the stability of the boron–hydrogen defect complex formed when the substitutional boron is passivated by hydrogen and compare this case with other passivated impurities in semiconductors. The vibrational properties of the two defects are analysed.

In Chapter 6 we determine the structure which would be formed by a self–interstitial atom. This is a precursor to the calculation of a diffusion energy for the interstitial atom. We study possible Jahn–Teller distortions of the only stable structure for the interstitial, the $\langle 100 \rangle$ split interstitial, and estimate selected multiplet energies. We calculate a value for the Jahn–Teller energy of the neutral vacancy and by estimating multiplet energies of the negative vacancy we determine the ground state structure of this charge state of the defect. Finally we estimate the energy required to form a vacancy–interstitial pair.

In Chapter 7 we propose a path for the diffusion of the $\langle 100 \rangle$ split interstitial and determine the migration energy for the defect along this path. We also deduce the diffusion paths for the neutral and negative vacancies and the energy barrier at the saddle point of diffusion. These calculations are relevant to the motion of atoms during annealing, the importance of which was described in the previous section.

In Chapter 8 we determine the path taken during the reorientation of the one extended nitrogen–carbon bond in an isolated substitutional nitrogen defect to one of the other neighbouring carbon atoms. Hence, we find a value for the energy barrier to this reorientation.

In Chapter 9 we make our concluding remarks regarding aspects of carrying out first principles calculations using the approach described in this thesis.

Chapter 2

The Theory of Electronic Structure Calculations

2.1 Introduction

The problem which we would ideally like to solve is the Schrödinger equation for N interacting electrons moving under the influence of an external potential due to atomic nuclei,

$$\left[-\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i,\alpha}\frac{Z_\alpha}{|\mathbf{r_i}-\mathbf{R}_\alpha|} + \frac{1}{2}\sum_{\alpha\neq\beta}\frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha-\mathbf{R}_\beta|} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r_i}-\mathbf{r_j}|}\right]\Psi = E\Psi, \tag{2.1}$$

where Ψ is the N-electron wavefunction, $\mathbf{r_i}$ are the positions of the electrons, $\mathbf{R}_\alpha$ are the positions of the nuclei and Z_α are the charges on the nuclei. Approximations must be made in order to derive a soluble equation.

In sections 2.2 to 2.4 we outline three of the significant early approaches. Hartree and Hartree–Fock theories were formulated using the wavefunction as the fundamental variable of an electronic system. The total energy of a system is then minimised with respect to the wavefunction. Hartree–Fock theory addressed a fundamental inadequacy of Hartree theory with regards to the symmetry of the wavefunction. The importance of Thomas-Fermi theory was that it was derived using the charge density as the fundamental variable and it is the simplest version of the theory upon which this work is based, density functional theory.

In the rest of this chapter we deal with theory directly used in the work presented in this thesis. In section 2.5 density functional theory (DFT) is formulated, which again has the charge density as fundamental variable. We then use this theory to derive self–consistent single particle equations in section 2.5.2 and the local density approximation to this theory, upon which this work is based, is explained in section 2.5.3. Finally in section 2.6 pseudopotential theory is discussed, especially in relation to its use in local density functional (LDF) electronic structure calculations.

2.2 Hartree Theory

The original derivation of Hartree Theory (Hartree 1928) started from the single particle Scrödinger equation,

$$-\frac{1}{2}\nabla^2\psi(\mathbf{r}) + U(\mathbf{r})\psi(\mathbf{r}) = \epsilon\psi(\mathbf{r}), \tag{2.2}$$

where ψ is the one electron wavefunction and $U(\mathbf{r})$ is the effective potential. Using (2.2) a suitable approximation needs to be made for the external potential exerted on each electron. Obviously it would be impossible to find a potential of such a form to accurately model the effect of a many electron system, but the basic physics is incorporated. The electrostatic attraction of the nuclei is included in the expression,

$$U_{nuc}(\mathbf{r}) = -\sum_{\alpha} \frac{Z_\alpha}{|\mathbf{r} - \mathbf{R}_\alpha|}. \tag{2.3}$$

Also, the effect of the other electrons is included in an approximate way. It is assumed that the potential due to this is a smooth field, dependent only on the averaged charge density of the other electrons, given by the expression,

$$U_{el}(\mathbf{r}) = \sum_{j \neq i} \int \frac{n_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \tag{2.4}$$

where, for each single particle electronic state, the charge density is given by,

$$n_i(\mathbf{r}) = |\psi_i(\mathbf{r})|^2, \tag{2.5}$$

where the ψ_i are the the orthonormal one electron wavefunctions. The total electronic charge density is then given by,

$$n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2, \tag{2.6}$$

where the sum is over the occupied one electron states. $U(\mathbf{r})$ is replaced in (2.2) by,

$$U(\mathbf{r}) = U_{nuc}(\mathbf{r}) + U_{el}(\mathbf{r}). \tag{2.7}$$

This leads to N single particle equations, the Hartree equations,

$$-\frac{1}{2}\nabla^2\psi_i(\mathbf{r}) + U_{nuc}(\mathbf{r})\psi_i(\mathbf{r}) + \sum_{j \neq i} \int \frac{1}{|\mathbf{r} - \mathbf{r}'|} |\psi_j(\mathbf{r}')|^2 \psi_i(\mathbf{r}) d\mathbf{r}' = \epsilon_i \psi_i(\mathbf{r}). \tag{2.8}$$

These equations for the occupied electronic levels are solved by iterative means.

The Hartree equations were later derived rigorously using the variational principle (Slater 1930) on (2.1) and minimising the energy with respect to a wavefunction of the form,

$$\Psi = \prod_i \psi_i(\mathbf{r_i}). \tag{2.9}$$

There is a fundamental flaw in Hartree theory. The total electronic wavefunction produced, being of the form given in (2.9), is always symmetric with respect to the interchange of particles. The Pauli principle states that the total wavefunction must be antisymmetric, which is impossible for a non-zero Ψ generated by this method.

2.3 Hartree–Fock Theory

In Hartree-Fock theory the total electronic wavefunction is constructed to be antisymmetric with respect to the interchange of particles. This resolves the fundamental symmetry problem of Hartree theory, but introduces considerable complexity into the implementation of the theory. The total electronic wavefunction can be represented by the determinant of an $N \times N$ matrix of single particle wavefunctions,

$$\Psi(\mathbf{r}_1 s_1, \mathbf{r}_2 s_2, \ldots, \mathbf{r}_N s_N) = \begin{vmatrix} \psi_1(\mathbf{r}_1 s_1) & \psi_1(\mathbf{r}_2 s_2) & \ldots & \psi_1(\mathbf{r}_N s_N) \\ \psi_2(\mathbf{r}_1 s_1) & \psi_2(\mathbf{r}_2 s_2) & \ldots & \psi_2(\mathbf{r}_N s_N) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \psi_N(\mathbf{r}_1 s_1) & \psi_N(\mathbf{r}_2 s_2) & \ldots & \psi_N(\mathbf{r}_N s_N) \end{vmatrix}. \tag{2.10}$$

This Slater determinant has the form of a linear combination of all possible permutations of the wavefunction obtained from (2.9) with respect to the values of i,

$$\Psi = \psi_1(\mathbf{r}_1 s_1)\psi_2(\mathbf{r}_2 s_2)\ldots\psi_N(\mathbf{r}_N s_N) - \psi_1(\mathbf{r}_2 s_2)\psi_2(\mathbf{r}_1 s_1)\ldots\psi_N(\mathbf{r}_N s_N) + \ldots \tag{2.11}$$

Applying the variational principle to (2.1) and minimising the energy with respect to a wavefunction of the form (2.10) leads to N single particle equations, the Hartree-Fock equations,

$$\begin{aligned} -\frac{1}{2}\nabla^2\psi_i(\mathbf{r}) + U_{nuc}(\mathbf{r})\psi_i(\mathbf{r}) + U_{el}(\mathbf{r})\psi_i(\mathbf{r}) & \\ - \sum_j \int \frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}')\psi_j(\mathbf{r})\delta_{s_i s_j} d\mathbf{r}' &= \epsilon_i \psi_i(\mathbf{r}). \end{aligned} \tag{2.12}$$

It should be noted that the expression used for $U_{el}(\mathbf{r})$ in (2.12) is,

$$U_{el}(\mathbf{r}) = \sum_j \int \frac{n_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}', \tag{2.13}$$

which can be compared with (2.4) as used in Hartree theory. The self–interaction terms produced by including the terms for $j = i$ are cancelled by identical terms of opposite sign arising from the fourth term on the left hand side of the equation. The potential given by (2.13) is termed the Hartree potential. The additional term which appears in (2.12) and not (2.8) relates to the exchange interaction of the electrons. There is a depletion of charge around an electron which creates an 'exchange hole' of charge with the same spin as the electron, as discussed by Slater (1974). This 'exchange' effect is due to the observance of the Pauli principle in the construction of the total electronic wavefunction. An effect that is not included in this approach is the exclusion of charge from the vicinity of an electron due to electrostatic repulsion, this effect is termed 'correlation'.

2.4 Thomas–Fermi Theory

Thomas–Fermi theory is the simplest form of density functional theory. It came about from an heuristic approach, which hindered further extensions and improvements to the theory. However, added corrections for exchange, correlation and density gradients did lead to some improvements, as discussed by Lundqvist (1974). In this approach the expressions for a uniform electron gas are used locally for a non-uniform electron gas.

A starting point for the theory is hence the determination of an expression for the charge density, n, of a uniform electron gas. The Pauli exclusion principle, together with boundary conditions, states that only 2 electrons with opposite spin can occupy a volume of h^3 in phase space. A uniform electron gas in its ground state has electrons with momenta up to p_f, the fermi momentum corresponding to the fermi energy, ϵ_f. Therefore the phase space volume occupied by an electron gas enclosed by a box of volume $\mathcal{V}$ is $\frac{4}{3}\pi p_f^3 \mathcal{V}$. The number of electrons is hence given by,

$$N = \frac{\frac{4}{3}\pi p_f^3 \mathcal{V}}{h^3} \times 2, \tag{2.14}$$

and the charge density is,

$$n = \frac{N}{\mathcal{V}} = \frac{8\pi p_f^3}{3h^3}. \tag{2.15}$$

The assumption is now made that the charge density at a point $\mathbf{r}$ in a non-uniform electron gas can be expressed as the charge density of a uniform electron gas at that point, $n(\mathbf{r})$. Therefore, (2.15) becomes,

$$n(\mathbf{r}) = \frac{1}{3\pi^2}\left(\frac{p_f(\mathbf{r})}{\hbar}\right)^3. \tag{2.16}$$

When the system is in equilibrium, the maximum total energy of an electron at a point $\mathbf{r}$ is constant. Classically it would be given by the expression,

$$\epsilon_f = \frac{p_f^2(\mathbf{r})}{2m} + V(\mathbf{r}), \tag{2.17}$$

where $\frac{p^2}{2m}$ is the classical kinetic energy of an electron with momentum p and $V(\mathbf{r})$ is the external potential. Substituting (2.16) into (2.17) in order to eliminate $p_f(\mathbf{r})$ and writing the energy in terms of the charge density,

$$\epsilon_f = \frac{\hbar^2}{2m}\left[3\pi^2 n(\mathbf{r})\right]^{\frac{2}{3}} + V(\mathbf{r}). \tag{2.18}$$

This expression can also be obtained from the minimisation of the total energy of the electronic system with respect to the charge density. An expression for the total energy of the electronic gas is constructed,

$$E_{el} = T + U_{el-el} + U_{el-nuc}, \tag{2.19}$$

where T is the kinetic energy, U_{el-el} is the energy due to the electron-electron interaction and U_{el-nuc} is the electron-nuclear interaction energy. The kinetic energy per unit volume at a point $\mathbf{r}$ can be expressed as,

$$t(\mathbf{r}) = \int_0^{p_f(\mathbf{r})} n(\mathbf{r})\frac{p^2}{2m}P(p)dp, \tag{2.20}$$

where $P(p)$ is the probability distribution or density of states of electronic momenta in the electron gas. The probability of an electron having momentum between p and $p + dp$ is the ratio of the volume of the shell at radius p in momentum space of thickness dp to the volume of the sphere enclosing all the electrons of radius p_f. i.e.,

$$P(p) = \begin{cases} \frac{4\pi p^2 dp}{\frac{4}{3}p_f^3(\mathbf{r})} & \text{for } p < p_f(\mathbf{r}) \\ 0 & \text{else} \end{cases} \tag{2.21}$$

Substituting (2.21) into (2.20),

$$t(\mathbf{r}) = \int_0^{p_f(\mathbf{r})} n(\mathbf{r}) \frac{p^2}{2m} \frac{3p^2}{p_f^3(\mathbf{r})} dp. \tag{2.22}$$

Using (2.16), the total kinetic energy is hence,

$$T = \int t(\mathbf{r}) d\mathbf{r} = \frac{3}{5} \frac{\hbar^2}{2m} (3\pi^2)^{\frac{2}{3}} \int [n(\mathbf{r})]^{\frac{5}{3}} d\mathbf{r}. \tag{2.23}$$

The electrostatic interaction terms are given by,

$$U_{el-nuc} = \int n(\mathbf{r}) V_{nuc}(\mathbf{r}) d\mathbf{r}, \tag{2.24}$$

and,

$$U_{el-el} = \frac{1}{2} e^2 \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \tag{2.25}$$

where $V_{nuc}(\mathbf{r})$ is the potential due to the nucleus. Substituting (2.23), (2.24) and (2.25) into (2.19),

$$E_{el} = \alpha \int [n(\mathbf{r})]^{\frac{5}{3}} d\mathbf{r} + \int n(\mathbf{r}) V_{nuc}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} e^2 \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \tag{2.26}$$

where,

$$\alpha = \frac{3}{5} \frac{\hbar^2}{2m} (3\pi^2)^{\frac{2}{3}}. \tag{2.27}$$

The minimisation of (2.26) with respect to $n(\mathbf{r})$ is carried out under the constraint,

$$N = \int n(\mathbf{r}) d\mathbf{r}. \tag{2.28}$$

Using a Lagrange multiplier, μ, to handle (2.28),

$$\frac{\delta}{\delta n(\mathbf{r})}(E - \mu N) = 0, \tag{2.29}$$

we obtain the expression for the Lagrange multiplier,

$$\mu = \frac{5}{3}\alpha[n(\mathbf{r})]^{\frac{2}{3}} + V_{nuc}(\mathbf{r}) + e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \tag{2.30}$$

where μ is also equivalent to the chemical potential of the system as can be seen from (2.29),

$$\mu = \frac{\partial E}{\partial N}. \tag{2.31}$$

It is noted that (2.30) can be seen to be identical to (2.18) where μ is equivalent to ϵ_f and $V(\mathbf{r})$ is taken to be equal to the sum of the nuclear and Hartree potentials. A self consistent solution of (2.30) is then found. Application of the method is discussed by Lundqvist and March (1983).

2.5 Density Functional Theory

DFT was formulated by Hohenberg and Kohn (1964). As with Thomas–Fermi theory and its extensions the charge density is the fundamental variable and energies are expressed in terms of this rather than the wavefunction. This should be advantageous when minimising the total energy with respect to the fundamental variable. The charge density is only a function of three coordinates, whereas the total wavefunction depends on the wavefunction of all N electrons, each of which is a function of three coordinates. Therefore, minimisation can be carried out over three variables rather than 3N variables. The theory was then used to derive self–consistent single particle equations by Kohn and Sham (1965). Unlike Thomas–Fermi theory, this is a rigorous derivation. A complete review of the theory is presented by Lundqvist and March (1983). Atomic units are used (i.e. $e = \hbar = m_e = e^2/4\pi\epsilon_0 = 1$. The unit of energy is $2R_y$ and the unit of length is the Bohr radius, a_0.)

2.5.1 The Charge Density as Fundamental Variable

The system considered is an electron gas in a finite enclosure under the influence of an external potential, $v(\mathbf{r})$, in its non–degenerate ground state. The Hamiltonian operator for such a system is,

$$\hat{H} = \hat{T} + \hat{V} + \hat{U}, \tag{2.32}$$

where the kinetic energy term is,

$$\hat{T} = -\frac{1}{2}\sum_i \nabla_i^2, \tag{2.33}$$

the external potential term is,

$$\hat{V} = v(\mathbf{r}), \tag{2.34}$$

and the electronic interaction term is,

$$\hat{U} = \frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{2.35}$$

It is known that the total energy of an electronic system is a unique functional of the wavefunction. As DFT is formulated in terms of the charge density, the first important result that was proved was that the total energy is also a unique functional of the charge density, n($\mathbf{r}$). The proof starts from the assumption that two external potentials $v(\mathbf{r})$ and $v'(\mathbf{r})$ which have ground state wavefunctions $\Psi(\mathbf{r})$ and $\Psi'(\mathbf{r})$ give rise to the same charge density, $n(\mathbf{r})$. Providing that the potentials do not merely differ by a constant, $\Psi(\mathbf{r})$ and $\Psi'(\mathbf{r})$ must be different. The proof then continues to show that the initial assumption cannot be correct.

If E and E' are the energies corresponding to the expectation values of the Hamiltonians $\hat{H}$ and $\hat{H}'$ involving the external potentials $v(\mathbf{r})$ and $v'(\mathbf{r})$ and taken using the ground state wavefunctions $\Psi(\mathbf{r})$ and $\Psi'(\mathbf{r})$ respectively, then the total energy of the system is given by,

$$E' = \langle\Psi'|\hat{H}'|\Psi'\rangle \leq \langle\Psi|\hat{H}'|\Psi\rangle = \langle\Psi|\hat{H} + \hat{V}' - \hat{V}|\Psi\rangle, \tag{2.36}$$

as, under the assumption made, only the external potential term differs between the Hamiltonians involving $v(\mathbf{r})$ and $v'(\mathbf{r})$. Therefore, from (2.36),

$$E' \leq E + \int [v'(\mathbf{r}) - v(\mathbf{r})] n(\mathbf{r}) d\mathbf{r}. \tag{2.37}$$

Equivalently, by interchanging the primed and unprimed terms in (2.37),

$$E \leq E' + \int [v(\mathbf{r}) - v'(\mathbf{r})] n(\mathbf{r}) d\mathbf{r}. \tag{2.38}$$

Adding (2.37) and (2.38) leads to,

$$E + E' \leq E + E'. \tag{2.39}$$

Assuming a non–degenerate ground state for the system, (2.39) becomes,

$$E + E' < E + E', \tag{2.40}$$

which is a contradiction and so shows that the original assumption was incorrect. Therefore, a reciprocal functional relationship between $v(\mathbf{r})$ and $n(\mathbf{r})$ is established, where one uniquely determines the other. As $v(\mathbf{r})$ determines the Hamiltonian of which $\Psi(\mathbf{r})$ is the ground state wavefunction, it must also follow that there is a reciprocal functional relationship between $\Psi(\mathbf{r})$ and $v(\mathbf{r})$ and, therefore, also between $\Psi(\mathbf{r})$ and $n(\mathbf{r})$.

A universal functional was defined which incorporates the kinetic and coulomb interaction energies,

$$F[n] = \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle. \tag{2.41}$$

$F[n]$ is independent of the external potential, which has been shown above to be uniquely determined by $n(\mathbf{r})$. Therefore, the total energy can be defined as a functional of $n(\mathbf{r})$ for a particular $v(\mathbf{r})$,

$$E[n'] = \int v(\mathbf{r}) n'(\mathbf{r}) d\mathbf{r} + F[n']. \tag{2.42}$$

The energy functional in terms of the wavefunction is,

$$\mathcal{E}[\Psi'] = \langle\Psi'|\hat{V}|\Psi'\rangle + \langle\Psi'|\hat{T} + \hat{U}|\Psi'\rangle, \tag{2.43}$$

This attains its lowest (ground state) value for the correct value of Ψ'. If Ψ is the ground state wavefunction then, using the reciprocal functional relationship between $\Psi(\mathbf{r})$ and $n(\mathbf{r})$,

$$\mathcal{E}[\Psi] = \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} + F[n] < \mathcal{E}[\Psi'] = \int v(\mathbf{r})n'(\mathbf{r})d\mathbf{r} + F[n'], \tag{2.44}$$

so that,

$$E[n] < E[n']. \tag{2.45}$$

This then establishes that the energy also has a lower value for the ground state charge density than for any other charge density. It has been assumed in this proof that the possible $n(\mathbf{r})$ are such that,

$$\int n(\mathbf{r})d\mathbf{r} = N. \tag{2.46}$$

Another universal functional, $G[n]$, was defined. The classical coulomb interaction energy was removed from $F[n]$, giving,

$$G[n] \equiv F[n] - \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}'. \tag{2.47}$$

The energy due to the electron–electron interaction given by the expression $\frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}'$ is termed the Hartree energy. Using (2.47) in (2.42), the total energy of the electronic system can be expressed as,

$$E[n] = \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + G[n]. \tag{2.48}$$

Initial attempts at utilising the theory developed by Hohenberg and Kohn to determine charge densities centred on formulating approximations for the functional $G[n]$ which were both easy to implement and gave sufficient accuracy. This approach was later abandoned in favour of an approach based on deriving single particle equations which could be solved self consistently, as described in section 2.5.2.

2.5.2 The Kohn–Sham Equations

Following on from (2.48), Kohn and Sham (1965) formulated an exact expression for the total energy of an electronic system in terms of the charge density,

$$E_{tot}[n(\mathbf{r})] = T_s[n(\mathbf{r})] + \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{xc}[n(\mathbf{r})]. \quad (2.49)$$

This incorporates a functional which was defined by removing the kinetic energy of a non–interacting electron gas of density $n(\mathbf{r})$, $T_s[n]$, from $G[n]$ to give,

$$E_{xc}[n] \equiv G[n] - T_s[n]. \quad (2.50)$$

This is termed the exchange–correlation energy functional as it includes the energy of these electronic interactions effects for an interacting system of density $n(\mathbf{r})$.

As has been shown in section 2.5.1, $E_{tot}[n]$ attains its lowest value for the correct ground state charge density for a particular applied potential. It was also shown in this section that there is a one to one correspondence between the wavefunction, $\Psi(\mathbf{r})$, and the charge density, $n(\mathbf{r})$, of a system. The minimisation of (2.49) with respect to $\Psi(\mathbf{r})$ is, therefore, equivalent to minimising with respect to $n(\mathbf{r})$. It is assumed that the total wavefunction of an interacting electronic system can be expressed as the sum of single particle wavefunctions for the N occupied electronic levels in a non–interacting system under an effective potential, v_{eff}, so that,

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\Psi_i(\mathbf{r})|^2. \quad (2.51)$$

The grounds for making this assumption will become apparent. The minimisation of (2.49) is carried out with respect to the $\Psi_i(\mathbf{r})$, subject to the constraint,

$$\int |\Psi_i(\mathbf{r})|^2 = 1. \quad (2.52)$$

A set of Lagrange multiplier, ϵ_i, are introduced to handle (2.52), producing,

$$\frac{\delta}{\delta \Psi_i^*(\mathbf{r})}(E_{tot} - \sum_i \epsilon_i \int |\Psi_i(\mathbf{r})|^2) = 0. \tag{2.53}$$

On differentiating (2.53) becomes,

$$-\frac{1}{2}\nabla^2 \Psi_i(\mathbf{r}) + \left(v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r})\right) \Psi_i(\mathbf{r}) = \epsilon_i \Psi_i(\mathbf{r}), \tag{2.54}$$

where we have used the definition,

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}. \tag{2.55}$$

It is apparent that (2.54) is the single particle Schrödinger equation with eigenvalues $\Psi_i(\mathbf{r})$ as assumed in (2.51). Defining the effective potential as,

$$v_{eff}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}), \tag{2.56}$$

(2.54) becomes,

$$\left[-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r})\right] \Psi_i(\mathbf{r}) = \epsilon_i \Psi_i(\mathbf{r}). \tag{2.57}$$

(2.51), (2.56) and (2.57) are known as the Kohn–Sham equations. The self-consistent solution of these equations leads to the ground state charge density of the electronic system, which defines all other properties of the electronic system. Multiplying (2.57) by Ψ_i^*, integrating over all space and summing over all N occupied electronic levels,

$$\sum_{i=1}^{N} \langle \Psi_i | -\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r}) | \Psi_i \rangle = \sum_{i=1}^{N} \epsilon_i, \tag{2.58}$$

or,

$$T_s[n] + \int v_{eff}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} = \sum_{i=1}^{N} \epsilon_i. \tag{2.59}$$

Substituting (2.59) into the original expression for the total energy, (2.49), we obtain an alternative expression for $E_{tot}[n]$,

$$E_{tot} = \sum_{i=1}^{N} \epsilon_i - \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[n] - \int v_{xc}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}. \tag{2.60}$$

2.5.3 The Local Density Approximation

An approximation is needed for the exchange-correlation energy in order to turn (2.60) into a soluble problem. The approximation used in most present *ab initio* work and in this thesis is the local density approximation (LDA). The approximation is valid for slowly varying charge densities and high charge densities. It is assumed that the exchange-correlation energy contribution to the total energy, E_{tot}, can be expressed as,

$$E_{xc}[n] = \int \epsilon_{xc}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r}, \tag{2.61}$$

where $\epsilon_{xc}(n(\mathbf{r}))$ is the exchange-correlation energy per electron in a uniform electron gas of charge density $n(\mathbf{r})$. In other words, the charge density of an electron gas is assumed to be constant over a localised region of space and the exchange–correlation energy is given by that for a uniform electron gas of the same density throughout this region. As the region is made infinitesimally small, this idea can be expressed as (2.61).

The exchange-correlation potential which gives rise to $\epsilon_{xc}(n(\mathbf{r}))$ is defined as,

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}(n(\mathbf{r}))}{\delta n(\mathbf{r})} = \frac{\delta}{\delta n}\left[\epsilon_{xc}(n(\mathbf{r}))n(\mathbf{r})\right]. \tag{2.62}$$

Also, it is noted that,

$$v_{xc}(\mathbf{r}) \equiv \mu_{xc}(n(\mathbf{r})), \tag{2.63}$$

where $\mu_{xc}(n(\mathbf{r}))$ is the term in the chemical potential which arises due to the exchange-correlation interaction in a uniform electron gas. When the LDA is applied, the expression for the total energy of the electronic system under the approximation, using (2.60), becomes,

$$E_{tot} \approx \sum_i \epsilon_i - \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}|}d\mathbf{r}d\mathbf{r}' + \int n(\mathbf{r})\left[\epsilon_{xc}(n(\mathbf{r})) - \mu_{xc}(n(\mathbf{r}))\right]d\mathbf{r}. \tag{2.64}$$

The only remaining significant problem in the solution of (2.64) is to find an appropriate approximate expression for $\epsilon_{xc}(n(\mathbf{r}))$. There have been many proposed and the approximation used in this work is that formulated by Ceperley (1980). The implementation of this approximation is discussed in section 3.4.

2.6 Pseudopotentials

2.6.1 Introduction

To include all the atomic electrons in the implementation of LDF theory in electronic structure calculations on clusters of atoms is computationally very demanding. The core electronic wavefunctions can be rapidly oscillating and decay over a small distance compared with the valence electronic wavefunctions. The valence wavefunctions must be orthogonal to the core wavefunctions and so must also oscillate rapidly within the core region. A large basis set would therefore be needed to expand all of the atomic wavefunctions. However, generally the core states are not significantly perturbed in interatomic bonding. The concept of a pseudopotential was developed in order to replace the Hartree and exchange–correlation potentials due to the core states and the ionic coulomb potential with a potential which produces a pseudowavefunction which conforms to certain specified criteria. This pseudowavefunction from pseudopotentials generated in more recent work is identical to the valence wavefunction beyond a certain radius. Inside that radius it is no longer necessary for the pseudowavefunction to be orthogonal to core wavefunctions and indeed it need not oscillate at all. A much smaller basis set is therefore required to expand this pseudowavefunction which increases the speed of calculations. We now outline some of the more significant stages in the development of pseudopotential generation.

The first pseudopotentials to be formulated were empirical pseudopotentials in which parameters were fitted to reproduce chosen experimentally measured properties.

Phillips and Kleinman (1959) gave a prescription for producing non–empirical, non-local pseudopotentials. The valence wavefunction was fitted to a basis of orthogonalised plane waves (OPW). These were formed from plane waves and Bloch sums of the core wavefunctions taken in such a way as to make the basis functions orthogonal to the core states. The pseudowavefunction was taken to be the valence wavefunction with the core terms removed, a smooth function. The pseudopotential which has this pseudowavefunction as an eigenfunction was found from the all–electron potential and the the core–state wavefunctions. The problem with this and other similar schemes is that of the neglect of the "orthogonality hole". The consequence of this is to place too much charge in the core region, so that the normalised pseudowavefunction has an amplitude less than the true valence wavefunction outside the core region by a multiplicative constant. The effect of this in self–consistent calculations is to set up an incorrect charge distribution with respect to the core which leads to an error in the coulomb potential.

The use of *ab initio* total energy calculations was then introduced in the construction of the pseudowavefunction. Kahn (1972) and then Melius (1974) used Hartree–Fock calculations. The pseudowavefunction was taken to be a linear combination of core and valence wavefunctions. The core wavefunctions were adjusted to go to zero as r^{l+1} as $r \to 0$. The pseudopotential was fitted to a basis of Gaussian functions multiplied by powers of r which excludes the r^{-2} divergence which occurs as $r \to 0$. Similar work was carried out using LDF calculations to produce a numerical potential by inverting the Schrödinger equation (Topiol 1977).

The problem of the "orthogonality hole" was addressed for both of these approaches. The Hartree–Fock scheme (Kahn 1972, Melius 1974) was modified by Redondo (1977). The all–electron wavefunction was fitted to a basis of long and short–range functions. The long–range functions were chosen to match the normalised valence wavefunction. The short–range functions were adjusted to satisfy the criteria of nodelessness, normalisation, smoothness and r^{l+1} behaviour as $r \rightarrow 0$ for the pseudowavefunction. The pseudopotential was then fitted as in the original work. The work of Topiol (1977) was altered in similar fashion (Zunger 1979). This retained the insistence on core states being smooth and having r^{l+1} behaviour as $r \rightarrow 0$. A function of form $r^{l+1}e^{-\alpha r}$ was then added in order to make the pseudowavefunction match the valence wavefunction beyond a particular radius. Both of these approaches produced so called norm–conserving pseudopotentials. The pseudowavefunction beyond some "core" radius exactly matches the valence wavefunction and they are both normalised. Therefore within this radius there is the same charge in both cases.

Christiansen (1979) made the wavefunction equivalent to the Hartree–Fock valence function beyond a certain radius. Within that radius the wavefunction was a five term polynomial of leading term r^{l+1}, the form of which was chosen so as to match the value and the first three derivatives of the valence wavefunction at the matching radius and also to ensure normalisation of the pseudowavefunction. The matching radius was varied to ensure the pseudowavefunction was nodeless.

The r^{-2} divergence which occurs as $r \rightarrow 0$ in several of the pseudopotential constructions outlined above was merely the result of the historical insistence on the core–state wavefunctions having r^{l+1} behaviour as $r \rightarrow 0$. This singularity was removed by Kerker (1980). The pseudowavefunction was made equivalent to the valence wavefunction produced by an LDF calculation for $r > r_c$ (where r_c is the "core" radius). For $r < r_c$ polynomials and exponentials of polynomials were used for the pseudowavefunction with leading terms r^l and r^{l+2}. This ensures a non–singular pseudopotential which can be expanded much more simply in terms of basis functions.

2.6.2 The BHS Pseudopotentials

Bachelet, Hamann and Schluter (BHS) (1982) generated norm-conserving, transferable pseudopotentials for the elements hydrogen to plutonium inclusive. The pseudopotentials were derived using the relativistic form of the LDA to DFT. In their non-relativistic form these are the pseudopotentials used for the cores of the atoms modelled in this work. The Kohn-Sham equations (2.51), (2.56) and (2.57) were replaced by the their relativistic equivalent, the Dirac equations, in radial form. There are equations of this form for each j-quantum state. Originally these equations were solved for different atomic valence reference configurations. It was, however, found that the resultant pseudopotentials were generally independent of the reference configuration. These equations were solved numerically to give the valence state single particle eigenvalues and radial wavefunctions, and the self consistent LDF potential

The relativistic effects are only significant for the core electrons of heavy atoms. Consequently, for the determination of the pseudopotential for the valence electrons outside the core, the Dirac equations can be replaced by the radial Schrödinger equation,

$$\left[-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V(r)\right]\phi_j(r) = \epsilon\phi_j(r), \tag{2.65}$$

where $\phi_j(r) = rR_j(r)$ and $R_j(r)$ is the true radial wavefunction.

The following steps were carried out in the construction of the pseudopotentials.

The First Stage Pseudopotential

A first stage pseudopotential is expressed as,

$$V_j^{(1)}(r) = V(r)\left[1 - f\left(\frac{r}{r_{c_j}}\right)\right] + c_j f\left(\frac{r}{r_{c_j}}\right), \tag{2.66}$$

which removes the singularity in $V(r)$ at the atomic centre. This definition incorporates a cut-off function,

$$f\left(\frac{r}{r_{c_j}}\right) = \exp\left[\left(-\frac{r}{r_{c_j}}\right)^{\lambda}\right]. \tag{2.67}$$

After investigating the application of the construction method to various atoms the value of λ was chosen to be $\lambda = 3.5$, the reason being that it allows better analytical fits to the final pseudopotentials than higher values. The parameter c_j is adjusted to ensure that the lowest nodeless solution of (2.65) containing $V_j^{(1)}(r)$ in place of $V(r)$ has an eigenvalue equal to that of the true valence state. The resulting wavefunction, $\phi_j^{(1)}(r)$, is related to the full core valence wavefunction for $r > r_{c_j}$ by,

$$\phi_j^{(1)}(r) = \gamma_j \phi_j(r), \tag{2.68}$$

where γ_j is a normalisation constant. r_{c_j} delimits the core of the pseudoatom and was termed the "quality" parameter as it is an indication of the accuracy of the pseudopotential. For small values of r_{c_j}, the pseudopotentials constructed are strongly dependent on j with the maximum possible accuracy in matching the true valence potential. For large values of r_{c_j}, the pseudopotentials are smooth, quite independent of j and with larger deviations of the pseudowavefunction from the true wavefunction outside the core region. The choice of r_{c_j} is restricted to be between the outermost node and stationary point, r_{max}, of the wavefunction. BHS expressed the r_{c_j} for each atom as,

$$r_{c_j} = \frac{r_{max}}{cc}, \tag{2.69}$$

where the values of cc are tabulated.

Normalisation

The wavefunction, $\phi_j^{(1)}$ which arises from the first stage pseudopotential is then normalised in order to make it identical to the full core valence wavefunction for $r > r_{c_j}$. The new wavefunction is given by,

$$\phi_j^{(2)}(r) = \gamma_j \left[\phi_j^{(1)}(r) + \delta_j r^{(l+1)} f\left(\frac{r}{r_{c_j}}\right)\right]. \tag{2.70}$$

The value of δ_j is given by the smaller solution of,

$$(\gamma_j)^2 \int \left[\phi_j^{(1)}(r) + \delta_j r^{(l+1)} f\left(\frac{r}{r_{c_j}}\right)\right]^2 dr = 1. \tag{2.71}$$

This step ensures the norm-conserving nature of the pseudopotential. A consequence of this is that the charge within the core radius of the pseudoatom is equal to that in the real atom.

The Second Stage Pseudopotential

The pseudopotential corresponding to the normalised pseudowavefunction, $\phi_j^{(2)}(r)$ is now found. The radial Schrödinger equation is inverted to give the final screened pseudopotential, $V_j^{(2)}(r)$. An analytical method was given,

$$V_j^{(2)}(r) = V_j^{(1)}(r) + \frac{\delta_j r^{l+1} f(x)}{2\phi_j^{(2)}(r)} \left[\frac{\lambda^2 x^{2\lambda} - [2\lambda l + \lambda(\lambda+1)]x^{\lambda}}{r^2} + 2\epsilon_j - 2V_j^{(1)}(r)\right]. \tag{2.72}$$

The final unscreened pseudopotential, $V_j^{ion}(r)$, is found by removing the Hartree and exchange-correlation contributions from $V_j^{(2)}(r)$,

$$V_j^{ion}(r) = V_j^{(2)}(r) - \int \frac{n(r')}{|r-r'|} dr' - \frac{\delta E_{xc}[n]}{\delta n(r)}, \tag{2.73}$$

where $n(r)$ is the valence pseudo charge density.

Fitting the BHS Pseudopotentials to Analytic Functions

BHS defined an average pseudopotential for each l-quantum state, replacing the pseudopotential for each j-quantum state with,

$$V_l^{ion}(r) = \frac{1}{2l+1}\left[l V_{l-\frac{1}{2}}^{ion}(r) + (l+1) V_{l+\frac{1}{2}}^{ion}(r)\right], \tag{2.74}$$

where the weighting factors arise from the degeneracies of the different j states. $V_l^{ion}(r)$ is consequentially a non-relativistic potential. They also defined a spin-orbit pseudopotential which incorporates relativistic effects. For the atoms dealt with in this work relativistic effects are negligible so this term will not be considered. They finally split this function into long and short range parts and defined,

$$V_l^{ion}(r) = V^{core}(r) + \Delta V_l^{ion}(r). \tag{2.75}$$

The long range coulombic term is fitted using,

$$V^{core}(r) = -\frac{Z_v}{r}\left[\sum_{i=1}^{2} c_i^{core}\mathrm{erf}\left[(\alpha_i^{core})^{\frac{1}{2}} r\right]\right], \tag{2.76}$$

where Z_v is the valence charge and $c_1^{core} + c_2^{core} = 1$. This term is l-independent. The short range l-dependent term is fitted using,

$$\Delta V_l^{ion}(r) = \sum_{i=1}^{3}\left(A_i + r^2 A_{i+3}\right) e^{\alpha_i r^2}. \tag{2.77}$$

BHS tabulated the parameters c_i^{core},α_i^{core}, A_i and α_i for the atoms H to Pu. The fitting of $V_l^{ion}(r)$ to the analytic functions chosen, i.e. erf, Gaussian and $r^2\times$ Gaussian, allows the direct evaluation of the matrix elements of the pseudopotential between basis functions in methods involving the expansion of the wavefunction in terms of plane waves or Gaussians.

Chapter 3

The Solution of the Kohn–Sham Equations and the Optimisation of Structure

3.1 Introduction

In Chapter 2 we derived a set of self–consistent single particle equations, the Kohn–Sham equations, from which the charge density of an electronic system can be calculated. In this chapter we present the formalism that we have used in the solution of these equations for our chosen electronic system, the valence electrons of a cluster of atoms. The core electronic states and nuclei of the atoms are replaced by pseudopotentials generated using the prescription given in section 2.6.2. We also describe how we determine the equilibrium structure of a cluster of atoms by minimising the total energy of the cluster with respect to the atomic positions.

The two most widely used schemes for the *ab initio* modelling of configurations of atoms at present are those using supercells and atomic clusters. We will initially discuss their relative merits and shortfalls and then proceed to develop the methodology for our chosen approach, that of expanding the wavefunction and charge density of a cluster of atoms using a basis of localised Gaussian orbitals.

3.1.1 Plane Waves and Supercells

The wavefunction has been expanded in terms of plane waves, localised orbitals and a mixture of both under the supercell formalism. As plane waves are predominantly used, we will discuss this approach which is a typical Fourier–type expansion.

The expansion of the wavefunction in terms of plane waves in a unit cell is tailor–made for perfect or bulk materials. The study of bulk materials is, however, of little interest today. Much of the research using *ab initio* methods involves the investigation of defects. In order to adapt this approach for this purpose the defects are sited at the centre of an extended unit cell known as a supercell. This produces an infinite lattice composed of defect centred supercells. As periodic boundary conditions are used, the termination at a boundary with free space need not be considered, as it must be in the cluster approach. The concentration of defects in this approach is far higher than that which would be found in reality. A typical supercell currently used in the modelling of defects in diamond would be of the order of 60 atoms. If a point defect was sited in a supercell of this size it would give a 2 % defect concentration (considering only one atom as comprising the defect). The effect of defect–defect interactions between the supercells causes the broadening of defect energy levels into bands. This effect is more pronounced the smaller the supercell used. The convergence of results must be checked with respect to supercell size. One of the largest calculations performed using this approach was that using a 700 effective atom supercell (400 silicon atoms and a vacuum layer) with a cut–off energy of 6 Ry to study the (7×7) surface reconstruction of Si(111) (Brommer 1992). A parallel machine with an effective rate of 600 Mflops took several hundred hours to model the system and had a memory requirement of 60 Mwords. The statistics associated with the calculation show that calculations involving this many atoms are not suitable for extensive investigations due to the massive computational requirement and the capabilities of computers generally available.

The only parameter associated with the plane wave expansion with which convergence of results must be checked is the cut–off frequency of the plane waves. This becomes two parameters if the higher frequency waves are dealt with using Löwdin perturbation theory, as they often are. When attempting to model materials with a high charge density, such as diamond, the plane wave cut–off energy has to be much higher than in materials such as silicon. In order to make calculations feasible, a sacrifice must generally be made in one or more areas of the formalism. The supercell size can be reduced leading to the inherent increase in defect–defect interactions. Alternatively, the plane wave cut–off energy can be reduced, although great care must be taken if this is carried out as strictly this should be raised until convergence of results is assured. Finally softer pseudopotentials can be used. This leads to a less transferable pseudopotential and produces a less accurate pseudo–wavefunction. As the work presented in this thesis involves the investigation of defects in diamond, this problem is of great significance.

With the plane wave expansion the same basis is applied throughout the supercell. This means that if there is any space in the supercell not occupied by atoms then this has the same basis as those regions which are occupied by atoms. This is especially significant when considering surfaces, when half the supercell is empty, and molecules such as C_{60}, when the majority of the supercell would be empty. The presence of atoms which are difficult to accurately model, such as transition metals or oxygen, again drastically increases the basis applied over the whole of the supercell.

3.1.2 Localised Orbitals and Atomic Clusters

The expansion of the wavefunction in terms of localised orbitals in a cluster of atoms is tailor–made for molecules. Therefore it is ideal for modelling discrete structures such as C_{60} (Faulhaber 1993a) and bis(ethylenedithio)tetrathiafulvalene (BEDT-TTF) and tetrathiafulvalene (TTF) (Faulhaber 1993b). The orbitals are centred on atoms, and also on interatomic bond–centres when considered necessary, in our approach.

In order to use a finite cluster to model defects in a solid, two problems associated with the limited extent of the cluster must be addressed. Firstly, there can be defect–surface interactions which cause variations in determined properties with changing size of cluster. It must, therefore, be ensured that the properties of a defect studied converge with cluster size. Secondly, the effect of the boundary with free space must be considered. The surface of a cluster simulating a solid has dangling bonds which would be attached to the next shell of atoms out in the solid. These bonds can give rise to levels in the band gap which can interact with any defect levels present. The approach that we have used to saturate these bonds is to attach hydrogen atoms to them.

An advantage of localised orbitals is that the basis on any atom or bond–centre is variable. A larger basis can be put on atoms involved in or surrounding defects than on those atoms close to the surface of the cluster. Also, if there are atoms which require a large basis to accurately expand the wavefunction and charge density, this increased basis is only applied to those atoms, and no basis need be applied to free space. This allows larger clusters to be used than equivalent supercells or simply faster calculations in cases where there is a significant restriction caused by the uniformity of the basis inherent in the plane wave approach. We have performed calculations using a 304 atom cluster (Chapter 6) which, using a 23 Mflop machine, took 20–40 hours to produce a self–consistent charge density and had a memory requirement of 30

Mbytes.

3.2 The Kohn–Sham Equations and a Basis of Localised Gaussian Orbitals

The Kohn–Sham equations (2.51), (2.56) and (2.57) must be solved self–consistently in order to determine the charge density for an electronic system. The total electronic wavefunction is assumed to be the sum of single particle wavefunctions for each of the λ occupied valence electronic states, Ψ_λ. The wavefunction for the λ'th electronic level is expanded in terms of a set of localised basis functions,

$$\Psi_\lambda(\mathbf{r}) = \sum_i c_i^\lambda \phi_i(\mathbf{r}). \tag{3.1}$$

As the basis set chosen is that of Gaussian orbitals, the functions have the form,

$$\phi_i(\mathbf{r}) = F e^{-\alpha_i(\mathbf{r}-\mathbf{R})^2}, \tag{3.2}$$

where F is a function which imparts the appropriate character of an atomic electronic orbital to $\phi_i(\mathbf{r})$. For s-type orbitals, $F = 1$, resulting in a spherically symmetric $\phi_i(\mathbf{r})$. For p-type orbitals, $F = (x - R_{ix})$ or $F = (y - R_{iy})$ or $F = (z - R_{iz})$, depending on which axis the orbital is aligned along. The charge density of the electronic system can be written as,

$$n(\mathbf{r}) = \sum_\lambda f_\lambda |\Psi_\lambda(\mathbf{r})|^2, \tag{3.3}$$

where f_λ is the occupancy of the λ'th electronic level. Defining,

$$b_{ij} = \sum_\lambda f_\lambda c_i^\lambda c_j^\lambda, \tag{3.4}$$

(3.3) becomes,

$$n(\mathbf{r}) = \sum_{ij} b_{ij} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}). \tag{3.5}$$

The matrix elements of the Hamiltonian for the Kohn–Sham equations, between wavefunction basis functions, can be written as,

$$H_{ij} = T_{ij} + V_{ij}^{ext} + V_{ij}^{H} + \mu_{ij}^{xc}, \tag{3.6}$$

where,

$$T_{ij} = \int \phi_i(\mathbf{r}) \left[-\frac{1}{2}\nabla^2\right] \phi_j(\mathbf{r}) d\mathbf{r} \tag{3.7}$$

$$V_{ij}^{ext} = \int \phi_i(\mathbf{r}) v^{ext}(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r} \tag{3.8}$$

$$V_{ij}^{H} = \int \phi_i(\mathbf{r}) \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \phi_j(\mathbf{r}) d\mathbf{r} d\mathbf{r}' \tag{3.9}$$

$$V_{ij}^{xc} = \int \phi_i(\mathbf{r}) \mu_{xc}(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r}. \tag{3.10}$$

The external potential, v^{ext}, is that due to the pseudopotentials representing the atomic cores. The evaluation of the terms T_{ij} and V_{ij}^{ext} can be carried out straightforwardly. The same is not true of V_{ij}^{H} and V_{ij}^{xc} and sections 3.3 and 3.4 discuss the approximations made in order to carry out these calculations.

3.3 The Hartree energy

The Hartree energy is given by,

$$U^{H} = \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{3.11}$$

From (3.5) it can be seen that in order to evaluate this expression a four centre integral of the wavefunction bases must be carried out. This would be a serious limiting factor for the computational implementation of this process. In fact, presently the calculation of this integral for clusters of a useful size is intractable. In order to reduce this problem to the evaluation of at most a three centre integral a technique has been developed which uses an intermediate fit to the charge density (Sambe 1975). The intermediate fit is expressed as,

$$\tilde{n}(\mathbf{r}) = \sum_k c_k g_k(\mathbf{r}), \tag{3.12}$$

where g_k is a Gaussian type function, the exact form of which is discussed below. The difference between the Hartree energy due to the true and intermediate fit charge densities is,

$$\Delta U^H = \frac{1}{2}\int \frac{[n(\mathbf{r}) - \tilde{n}(\mathbf{r})]\,[n(\mathbf{r}') - \tilde{n}(\mathbf{r}')]}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{3.13}$$

It can be seen that as $\tilde{n}(\mathbf{r}) \to n(\mathbf{r})$, $\Delta U^H \to 0$. We can write,

$$U^H = \tilde{U}^H + \Delta U^H, \tag{3.14}$$

where $\tilde{U}^H$ is the Hartree energy due to the intermediate fit charge density. Substituting (3.11) and (3.13) into (3.14),

$$\tilde{U^H} = \int \frac{n(\mathbf{r})\tilde{n}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' - \frac{1}{2}\int \frac{\tilde{n}(\mathbf{r})\tilde{n}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{3.15}$$

If we define,

$$t_{ijk} = \int \frac{\phi_i(\mathbf{r})\phi_j(\mathbf{r})g_k(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \tag{3.16}$$

and,

$$g_{kl} = \int \frac{g_k(\mathbf{r})g_l(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \tag{3.17}$$

then (3.15) becomes,

$$\tilde{U^H} = \sum_{ijk} c_k b_{ij} t_{ijk} - \frac{1}{2}\sum_{kl} c_k c_l g_{kl}. \tag{3.18}$$

The values chosen for the coefficients, c_k, are those which minimise (3.13). Substituting (3.11) and (3.18) into (3.14) and differentiating,

$$\frac{\partial \Delta U^H}{\partial c_k} = \frac{\partial}{\partial c_k}\left[\frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' - \sum_{ijk} c_k b_{ij} t_{ijk} + \frac{1}{2}\sum_{kl} c_k c_l g_{kl}\right] = 0. \tag{3.19}$$

This leads to the following set of linear equations, the solution of which yields the required values of c_k,

$$\sum_{ij} t_{ijk} b_{ij} = \sum_{l} g_{kl} c_l. \tag{3.20}$$

The procedure followed is that the components of t_{ijk} are evaluated using (3.16) and contracted with b_{ij} to form a vector q_k. The values of the coefficients c_k are then determined from,

$$c_k = \left(g^{-1}\right)_{kl} q_k. \tag{3.21}$$

Substituting (3.20) into (3.18) leads to the final expression for the approximation to the Hartree energy,

$$\tilde{U}^H = \frac{1}{2}\sum_k c_k c_l g_{kl}. \tag{3.22}$$

There are two types of fitting functions used for the intermediate fit to the charge density for the determination of the Hartree energy, so that,

$$\tilde{n}(\mathbf{r}) = Z\left(\frac{\alpha}{\pi}\right) e^{-\alpha(\mathbf{r}-\mathbf{R})^2} + \sum_k c_k \left[1 - \frac{2}{3}\alpha_k(\mathbf{r}-\mathbf{R})^2\right] e^{-\alpha_k(\mathbf{r}-\mathbf{R})^2} \tag{3.23}$$

The first term is a standard Gaussian. The integration of this function leads to a constant charge. This term only needs to be evaluated once for each self–consistent energy as it has a constant coefficient. It is then added to the pseudopotential. The set of k modified Gaussians give rise to a potential which is short ranged. The integration of each of these terms produces a charge of zero. These terms must be evaluated in each self-consistency step. The choice of values for the exponents and coefficients, α, α_k and c_k, of the fitting functions is discussed in section 3.13.

3.4 The Exchange-Correlation Energy

As with the Hartree energy evaluation, an intermediate fit to the charge density is employed in the approach used to determine the exchange–correlation energy. This intermediate fit is expanded in terms of a different set of Gaussian functions to those in (3.23),

$$\tilde{n}_{xc}(\mathbf{r}) = \sum_k c_k^{xc} h_k(\mathbf{r}), \tag{3.24}$$

where the form of $h_k(\mathbf{r})$ is discussed below.

Using the local density approximation for the exchange–correlation energy, a further approximation is made by replacing $n(\mathbf{r})$ by $\tilde{n}_{xc}(\mathbf{r})$,

$$E \approx \tilde{E}_{xc} = \int \tilde{n}_{xc}(\mathbf{r})\epsilon_{xc}(\tilde{n}_{xc}(\mathbf{r}))d\mathbf{r} \tag{3.25}$$

$$= \sum_k c_k^{xc} \int h_k(\mathbf{r})\epsilon_{xc}(\tilde{n}(\mathbf{r}))d\mathbf{r} \tag{3.26}$$

The mean value of $\tilde{n}_{xc}(\mathbf{r})$ is given by the first moment of the charge density,

$$m_{1k} = \frac{\int h_k(\mathbf{r})\tilde{n}_{xc}(\mathbf{r})d\mathbf{r}}{\int h_k(\mathbf{r})d\mathbf{r}} \tag{3.27}$$

$$= \sum_l \frac{c_l^{xc} H_{kl}}{I_k}, \tag{3.28}$$

where,

$$I_k = \int h_k(\mathbf{r})d\mathbf{r}, \tag{3.29}$$

and,

$$H_{kl} = \int h_k(\mathbf{r})h_l(\mathbf{r})d\mathbf{r} \tag{3.30}$$

Using Rollé's theorem to expand $\epsilon_{xc}(\tilde{n})$,

$$\epsilon_{xc}(\tilde{n}) = \epsilon_{xc}(m_{1k}) + (\tilde{n} - m_{1k})\epsilon_{xc}{}'(m_{1k}) + \frac{1}{2}(\tilde{n} - m_{1k})^2\epsilon_{xc}{}''(\bar{n}), \tag{3.31}$$

where $\bar{n} \neq m_{1k}$, but is otherwise unknown. We then obtain an expression for the kth integral of (3.26),

$$\int h_k(\mathbf{r})\epsilon_{xc}(\tilde{n})d\mathbf{r} = \int h_k(\mathbf{r}) \left[\epsilon_{xc}(m_{1k}) + \frac{1}{2}\int(\tilde{n} - m_{1k})^2\epsilon_{xc}{}''(\bar{n})\right] d\mathbf{r}, \tag{3.32}$$

as the second term on the right hand side of (3.31) becomes,

$$\int h_k(\mathbf{r})(\tilde{n} - m_{1k})\epsilon_{xc}{}'(m_{1k})d\mathbf{r} = 0. \tag{3.33}$$

The function ϵ_{xc} is convex, so that $\epsilon_{xc}'' < 0$. Therefore, a function f_k (where $f_k < 0$) can be found, so that we can express (3.32) exactly as,

$$\int h_k(\mathbf{r})\epsilon_{xc}(\tilde{n})d\mathbf{r} = \epsilon_{xc}(m_{1k})\exp(f_k)I_k. \tag{3.34}$$

In order to determine f_k, firstly we note that the single particle exchange–correlation energy can be fitted to,

$$\epsilon_{xc}(n) = An(\mathbf{r})^{s_0}, \tag{3.35}$$

where the Ceperley–Alder fit has the values $A = 1$ and $s_0 = 0.30917$ (Ceperley 1980). Replacing n by $\tilde{n}_{xc}$ and substituting (3.35) into (3.34),

$$\int h_k(\mathbf{r})\tilde{n}_{xc}(\mathbf{r})^{s_0} d\mathbf{r} = {m_{1k}}^{s_0} \exp(f_k) I_k \tag{3.36}$$

If we replace the constant, fitted value, of s_0 by a variable, s, from (3.36) we obtain the function,

$$f_k(s) = \ln\left[\frac{\int h_k(\mathbf{r})\tilde{n}_{xc}(\mathbf{r})^s d\mathbf{r}}{{m_{1k}}^s I_k}\right]. \tag{3.37}$$

We require the value of $f_k(s_0)$, but the numerical evaluation using a grid of points would be too time consuming. Instead we follow an interpolation approach using easily evaluated values of $f_k(s)$. The value of $f_k(s)$ is determined at three points,

$$f_k(0) = f_k(1) = 0, \tag{3.38}$$

and,

$$f_k(2) = \ln\left(\frac{m_{2k}}{{m_{1k}}^2}\right). \tag{3.39}$$

This involves the second moment of the intermediate fit to the charge density, which is given by,

$$m_{2k} = \frac{\int h_k(\mathbf{r})\tilde{n}_{xc}(\mathbf{r})^2 d\mathbf{r}}{\int h_k(\mathbf{r}) d\mathbf{r}} \tag{3.40}$$

$$= \sum_{lm} \frac{c_l^{xc} c_m^{xc} U_{klm}}{I_k}, \tag{3.41}$$

where,

$$U_{klm} = \int h_k(\mathbf{r}) h_l(\mathbf{r}) h_m(\mathbf{r}) d\mathbf{r}. \tag{3.42}$$

Using the Lagrange formula for polynomial interpolation to fit a quadratic to the values of $f_k(s)$ given by (3.38) and (3.39),

$$f_k(s) = \frac{1}{2}s(s-1)f_k(2). \tag{3.43}$$

The minimum of (3.43) is at $s = 0.5$ and, typically, $f_k \sim -0.1$. Therefore the evaluation of $f_k(s_0)$ should be in a region in which $f_k(s)$ is slowly varying and so not affected significantly by the fitting of $f_k(s)$ to only three points. The resulting approximation used for the exchange–correlation energy is,

$$\tilde{E}_{xc} = \sum_k c_k^{xc}\epsilon_{xc}(m_{1k})\exp(f_k(s_0))I_k \tag{3.44}$$

The values of c_k^{xc} for the exchange–correlation intermediate fit are chosen to be those which minimise the difference between the intermediate and true charge densities. Using a least–squares minimisation,

$$\frac{\partial}{\partial c_k^{xc}}\int\left(\tilde{n}_{xc}(\mathbf{r}) - n(\mathbf{r})\right)^2 d\mathbf{r} = 0, \tag{3.45}$$

therefore,

$$\int 2\tilde{n}_{xc}(\mathbf{r})\frac{\partial\tilde{n}_{xc}(\mathbf{r})}{\partial c_k^{xc}}d\mathbf{r} - 2\int n(\mathbf{r})\frac{\partial\tilde{n}_{xc}(\mathbf{r})}{\partial c_k^{xc}}d\mathbf{r} = 0. \tag{3.46}$$

Differentiating (3.24),

$$\frac{\partial\tilde{n}_{xc}(\mathbf{r})}{\partial c_k^{xc}} = h_k(\mathbf{r}). \tag{3.47}$$

Substituting (3.5), (3.24), (3.30) and (3.47) into (3.46), we obtain,

$$\sum_k c_k^{xc}H_{kl} = \sum_{ij} b_{ij}\int\phi_i(\mathbf{r})\phi_j(\mathbf{r})h_k(\mathbf{r}')d\mathbf{r}d\mathbf{r}'. \tag{3.48}$$

The form of the functions $h_k(\mathbf{r})$ are chosen to be standard Gaussians with the same exponents as for the modified Gaussians used in the intermediate fit to the charge density for the Hartree energy, (3.23),

$$\tilde{n}(\mathbf{r}) = \sum_k c_k^{xc}e^{-\alpha_k(\mathbf{r}-\mathbf{R})^2}. \tag{3.49}$$

The time dominant step in the determination of the c_k^{xc} values from (3.48) is the evaluation of the three centre integral on the right hand side of the equation. Referring to part of the equivalent integral for the Hartree energy in (3.16), for the modified Gaussian fitting functions,

$$\int \frac{\left[1 - \frac{2}{3}\alpha(\mathbf{r}' - \mathbf{R})^2\right] e^{-\alpha_k(\mathbf{r}'-\mathbf{R})^2}}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' = \frac{2\pi}{3\alpha_k} e^{-\alpha_k(\mathbf{r}-\mathbf{R})^2}. \tag{3.50}$$

Therefore the three centre integral in (3.48) has already effectively been evaluated. The vector q_k in (3.21) is multiplied by $\dfrac{3\alpha_k}{2\pi}$ to form a vector p_k. The values of the coefficients c_k^{xc} are then determined from,

$$c_k^{xc} = \left(H^{-1}\right)_{kl} p_k. \tag{3.51}$$

This selection of functions for $h_k(\mathbf{r})$, therefore, significantly reduces computer time. The choice of values for the coefficients of the fitting functions, c_k^{xc}, is discussed in section 3.13.

3.5 The Matrix Elements of the Hartree and Exchange–Correlation Potentials

The potential, V, corresponding to a particular energy contribution, $U[n]$, is given by the functional derivative,

$$V = \frac{\delta U[n]}{\delta n(\mathbf{r})}. \tag{3.52}$$

The matrix elements, V_{ij}, of a potential, V, are given by,

$$V_{ij} = \int \phi_i(\mathbf{r}) V \phi_j(\mathbf{r}) d\mathbf{r}. \tag{3.53}$$

Substituting (3.52) into (3.53),

$$V_{ij} = \int \phi_i(\mathbf{r})\phi_j(\mathbf{r}) \left[\frac{\delta U[n]}{\delta n(\mathbf{r})}\right] d\mathbf{r}. \tag{3.54}$$

Differentiating (3.5),

$$\frac{\partial n(\mathbf{r})}{\partial b_{ij}} = \phi_i(\mathbf{r})\phi_j(\mathbf{r}). \tag{3.55}$$

Substituting (3.55) into (3.54),

$$V_{ij} = \int \frac{\partial n(\mathbf{r})}{\partial b_{ij}} \frac{\delta U[n]}{\delta n(\mathbf{r})} d\mathbf{r} \tag{3.56}$$

$$= \frac{\partial U}{\partial b_{ij}}. \tag{3.57}$$

Using (3.57) the matrix elements of the Hartree potential are given by,

$$\tilde{V}^H_{ij} = \frac{\partial \tilde{U}^H}{\partial b_{ij}}. \tag{3.58}$$

Substituting (3.22) into (3.58),

$$\tilde{V}^H_{ij} = \frac{\partial c_k}{\partial b_{ij}} c_l g_{kl}. \tag{3.59}$$

Differentiating (3.21),

$$\frac{\partial c_k}{\partial b_{ij}} = \left(g^{-1}\right)_{kl} t_{ijl}, \tag{3.60}$$

hence,

$$\tilde{V}^H_{ij} = c_k t_{ijk}. \tag{3.61}$$

Again using (3.57), the matrix elements of the exchange–correlation potential are given by,

$$\mu^{xc}_{ij} = \frac{\partial \tilde{E}_{xc}}{\partial b_{ij}}. \tag{3.62}$$

Substituting (3.44) into (3.62),

$$\mu^{xc}_{ij} = \sum_k \left(\frac{\partial c^{xc}_k}{\partial b_{ij}}\right) \epsilon_{xc}(m_{1k}) \exp(f_k) I_k + \sum_k c^{xc}_k {\epsilon_{xc}}'(m_{1k}) \left(\frac{\partial m_{1k}}{\partial b_{ij}}\right) \exp(f_k) I_k$$
$$+ \sum_k c^{xc}_k \epsilon_{xc}(m_{1k}) \left(\frac{\partial f_k}{\partial b_{ij}}\right) \exp(f_k) I_k. \tag{3.63}$$

Differentiating (3.28),

$$\frac{\partial m_{1k}}{\partial b_{ij}} = \sum_l \left(\frac{\partial c^{xc}_l}{\partial b_{ij}}\right) \frac{H_{kl}}{I_k}, \tag{3.64}$$

and differentiating (3.43),

$$\frac{\partial f_k(s)}{\partial b_{ij}} = \frac{1}{2}s_0(s_0-1)\left[\left(\frac{\partial m_{2k}}{\partial b_{ij}}\right)\frac{1}{m_{2k}} - 2\left(\frac{\partial m_{1k}}{\partial b_{ij}}\right)\frac{1}{m_{1k}}\right], \tag{3.65}$$

where, by differentiating (3.41),

$$\frac{\partial m_{2k}}{\partial b_{ij}} = 2\sum_{lm}\left(\frac{\partial c_l^{xc}}{\partial b_{ij}}\right)c_m^{xc}\frac{U_{klm}}{I_k}. \tag{3.66}$$

Differentiating (3.51),

$$\frac{\partial c_k^{xc}}{\partial b_{ij}} = \left(H^{-1}\right)_{kl}\frac{3\alpha_k}{2\pi}t_{ijl}. \tag{3.67}$$

Combining these terms leads to the final expression,

$$\mu_{ij}^{xc} = \sum_{kl}\frac{3\alpha_k}{2\pi}\left(H^{-1}\right)_{kl}t_{ijl}\left[\exp(f_k)\epsilon_{xc}(m_{1k})I_k + \sum_m c_m^{xc}\exp(f_m)F_{km}\right] \tag{3.68}$$

where,

$$F_{kl} = \left[\epsilon_{xc}{}'(m_{1k}) - \left(\frac{s_0(s_0-1)}{m_{1l}}\right)\epsilon_{xc}(m_{1l})\right]H_{kl} + \frac{s_0(s_0-1)}{m_{2l}}\epsilon_{xc}(m_{1l})\sum_m c_m^{xc}U_{klm}. \tag{3.69}$$

3.6 The Total Energy

The expression for the total energy of a cluster of atoms, using (3.6) and the results obtained in sections 3.3 and 3.4, is,

$$E_{tot} = \sum_{ij}\left(T_{ij} + V_{ij}^{ext}\right)b_{ij} + \frac{1}{2}c_k c_l g_{kl} + \sum_k c_k^{xc}\epsilon_{xc}(m_{1k})\exp(f_k)I_k + E_{ion-ion}, \tag{3.70}$$

where $E_{ion-ion}$ is the energy of the coulomb interaction between the ions, known as the Madelung energy. We now wish to minimise (3.70) with respect to the wavefunction coefficients, c_i^λ. The constraint imposed on the minimisation is,

$$\sum_{ij} c_i^\lambda c_j^\lambda S_{ij} = 1, \tag{3.71}$$

where S_{ij} is the overlap matrix,

$$S_{ij} = \int \phi_i(\mathbf{r})\phi_j(\mathbf{r})d\mathbf{r}. \tag{3.72}$$

(3.71) ensures the normalisation of the wavefunctions of the one–electron levels. Using a Lagrangian multiplier, ϵ_λ, to handle (3.71),

$$\frac{\partial}{\partial c_i^\lambda}\left[E_{tot} - \epsilon_\lambda \sum_{ij} c_i^\lambda c_j^\mu S_{ij}\right] = 0. \tag{3.73}$$

Therefore, substituting (3.6) into (3.73),

$$\sum_j \left(T_{ij} + V_{ij}^{ext} + V_{ij}^{H} + \mu_{ij}^{xc}\right) c_j^\lambda = \epsilon_\lambda \sum_j S_{ij} c_j^\lambda. \tag{3.74}$$

This is a generalised eigenvalue problem which can be written in matrix notation as,

$$(\mathbf{H} - E\mathbf{S})\mathbf{c} = \mathbf{0}. \tag{3.75}$$

In order to solve (3.75), we could multiply by $\mathbf{S}^{-1}$, but this leads to a nonsymmetric Hamiltonian, which is undesirable. Instead, we proceed via the Choleski decomposition of $\mathbf{S}$ into an upper triangular matrix with its transpose,

$$\mathbf{S} = \mathbf{U}^T\mathbf{U}. \tag{3.76}$$

Substituting (3.76) into (3.75) and multiplying on the left by $(\mathbf{U}^T)^{-1}$,

$$(\mathbf{H}' - E\mathbf{I})\mathbf{d} = \mathbf{0}, \tag{3.77}$$

where,

$$\mathbf{H}' = (\mathbf{U}^T)^{-1}\mathbf{H}\mathbf{U}^{-1}, \tag{3.78}$$

and,

$$\mathbf{U}\mathbf{c} = \mathbf{d}. \tag{3.79}$$

From the standard eigenvalue problem, (3.77), it is a straightforward matter to determine **d** and hence **c**. The values found for the new wavefunction coefficients are then substituted in (3.4) to find b_{ij}^{out}. The coefficients of the fit to the charge density for the Hartree energy are then found using (3.20) and those for the fit to the exchange–correlation energy using (3.51). The new values for the charge density coefficients obtained (we shall refer to them as d_k and d_k^{xc} to avoid confusion with the original c_k and c_k^{xc}) will only be the same as those entered into the equations if self–consistency has been reached. Multiplying (3.74) by c_i^λ and summing over i and λ, then, in terms of the output wavefunction coefficients, we obtain,

$$\sum_{ij} \left(T_{ij} + V_{ij}^{ext} + V_{ij}^{H} + \mu_{ij}^{xc} \right) b_{ij}^{out} = \sum_{\lambda} \epsilon_\lambda. \tag{3.80}$$

Subtracting (3.80) from (3.70), also expressed in terms of the output wavefunction and charge density coefficients, we obtain an expression for the total energy in terms of the output coefficients,

$$\begin{aligned} E_{tot} = \sum_{\lambda} \epsilon_\lambda - \sum_{ij} \left(V_{ij} + \mu_{ij} \right) b_{ij}^{out} + \frac{1}{2} \sum_{kl} d_k d_l g_{kl} \\ + \sum_{k} d_k^{xc} \epsilon_{xc}(m_{1k}) \exp(f_k) I_k + E_{ion-ion}. \end{aligned} \tag{3.81}$$

3.7 The Determination of a Self Consistent Total Energy

The stage that we have reached in the solution of the Kohn–Sham equations is such that we are now able to calculate the total energy of a cluster of atoms. This is not, however, a self–consistent total energy. An approximation to the charge density,

$$\tilde{n}^{in}(\mathbf{r}) = \sum_k c_k g_k(\mathbf{r}) \qquad \text{and} \qquad \tilde{n}_{xc}^{in}(\mathbf{r}) = \sum_k c_k^{xc} g_k(\mathbf{r}), \tag{3.82}$$

is input into the equations (3.61),(3.68) and (3.69) to construct the Hamiltonian. This is diagonalised and, following the approach described in section 3.6, we obtain the total energy of the system along with a new approximation to the charge density,

$$\tilde{n}^{out}(\mathbf{r}) = \sum_k d_k g_k(\mathbf{r}) \qquad \text{and} \qquad \tilde{n}^{out}_{xc}(\mathbf{r}) = \sum_k c^{xc}_k g_k(\mathbf{r}). \tag{3.83}$$

In order to calculate a self consistent total energy we require that,

$$\Big| E_{tot}[\tilde{n}^{in}(\mathbf{r})] - E_{tot}[\tilde{n}^{out}(\mathbf{r})] \Big| \leq \Delta E, \tag{3.84}$$

where ΔE is the accuracy to which the total energy is to be determined, the criterion adopted for convergence.

What is needed is a method for bringing the total energy to convergence in as few iteration steps as possible, but which is robust enough to succeed in almost all cases. Beginning with the simplest, possible iteration schemes are appraised. Only the coefficients for the intermediate fit to the charge density in calculating the Hartree energy will be considered in relation to the convergence of the total energy. The reason for this will be given below.

(1) The values of d_k obtained from the previous iteration could be entered as the values of c_k for the next. This is unsuitable as it causes erratic fluctuations in the total energy.

(2) The input charge density coefficients for the next iteration, c'_k, could be determined in terms of the input and output charge density coefficients of the previous iteration,

$$c'_k = c_k + w(d_k - c_k), \tag{3.85}$$

where w is a weighting factor which regulates the possible change in the c'_k's. The question then would be what to use for w.

(i) w could be fixed as a constant. A balance must be struck if this is chosen as decreasing the value of w leads to slower convergence and increasing it leads to greater instability in the convergence.

(ii) w could be a variable, so that the total energy is a function of w, $E^{tot}(w)$. This can then be minimised with respect to w using a suitable technique. One such method is that formulated by Brent (1973) which is very reliable.

The actual method used in our calculational procedure is based on this last approach. It is, in general, quicker but less robust than the last scheme described. Brent's method itself can be used if the procedure described below fails. This was never found to be necessary in the work described in this thesis.

We define the charge density coefficients, c_k, of the nth iteration as the components of an N–space vector, $\mathbf{c}_n$, where N is the number of coefficients. We also define an operator, $\hat{L}$, such that,

$$\hat{L}(\mathbf{c}_n) = \mathbf{d}_n. \tag{3.86}$$

Therefore the perfect self–consistency criterion can be expressed as,

$$\hat{L}(\mathbf{c}_n) = \mathbf{c}_n, \tag{3.87}$$

which would be the case where the input charge density produced an identical output charge density. (3.85) can be modified to give,

$$\mathbf{c}_{n+1} = \mathbf{c}_n + w(\mathbf{d}_n - \mathbf{c}_n). \tag{3.88}$$

Operating with $\hat{L}$,

$$\hat{L}(\mathbf{c}_{n+1}) = \hat{L}\left(\mathbf{c}_n + w(\mathbf{d}_n - \mathbf{c}_n)\right). \tag{3.89}$$

Considering $(\mathbf{d}_n - \mathbf{c}_n)$ to be small in relation to c_n, an approximation is made that the right hand side of (3.89) can be linearised,

$$\begin{aligned} \hat{L}(\mathbf{c}_{n+1}) &= \hat{L}(\mathbf{c}_n) + w\hat{L}'(\mathbf{d}_n - \mathbf{c}_n) \\ &= \mathbf{d}_n + w\hat{L}'(\mathbf{d}_n - \mathbf{c}_n), \end{aligned} \tag{3.90}$$

where $\hat{L}'$ is another operator. We define a vector, which is connected to the error in self–consistency,

$$\begin{aligned}\mathbf{x} &= \hat{L}(\mathbf{c}_{n+1}) - \mathbf{c}_{n+1} \\ &= \mathbf{d}_n - \mathbf{c}_n + w(\hat{L}' - \hat{1})(\mathbf{d}_n - \mathbf{c}_n).\end{aligned} \tag{3.91}$$

Finally $\mathbf{x}^T\mathbf{G}\mathbf{x}$ is minimised with respect to w in order to give the value of w to be used, where the elements of $\mathbf{G}$ are given by (3.17). The quantity $\mathbf{x}^T\mathbf{G}\mathbf{x}$ is the electrostatic energy associated with the vector $\mathbf{x}$.

This approach can be extended to include information from previous iterations. (3.88) can be rewritten as,

$$\mathbf{c}_{n+1} = \mathbf{c}_n + \sum_{i=0}^{2} w_i(\mathbf{d}_{n-i} - \mathbf{c}_{n-i}). \tag{3.92}$$

Therefore, following the same procedure as before,

$$\mathbf{x} = \mathbf{d}_n - \mathbf{c}_n + \sum_{i=0}^{2} w_i(\hat{L}' - \hat{1})(\mathbf{d}_{n-i} - \mathbf{c}_{n-i}). \tag{3.93}$$

This time the minimisation of $\mathbf{x}^T\mathbf{G}\mathbf{x}$ with respect to the three w_i leads to three simultaneous equations,

$$\begin{aligned}\sum_{i=0}^{2} w_i \left[(\hat{L}' - \hat{1})(\mathbf{d}_{n-i} - \mathbf{c}_{n-i})\right]^T \mathbf{G} \left[(\hat{L}' - \hat{1})(\mathbf{d}_{n-i} - \mathbf{c}_{n-i})\right] = \\ -(\mathbf{d}_n - \mathbf{c}_n)^T \left[(\hat{L}' - \hat{1})(\mathbf{d}_{n-i} - \mathbf{c}_{n-i})\right].\end{aligned} \tag{3.94}$$

The solution of which gives the weighting factors, w_i. The only unknown factor in (3.94) is the operator $\hat{L}'$. Using (3.88) and (3.90), we can write,

$$\mathbf{c}_{n+1} = \mathbf{c}_n + w_c(\mathbf{d}_n - \mathbf{c}_n), \tag{3.95}$$

and,

$$\mathbf{d}_{n+1} = \mathbf{d}_n + w_c\hat{L}'(\mathbf{d}_n - \mathbf{c}_n), \tag{3.96}$$

where w_c is a constant which is set, at present, to 0.1. This enables the determination of $\hat{L}'(\mathbf{d}_n - \mathbf{c}_n)$. In the calculations that we have performed the value of ΔE has usually been set to 10^{-5} a.u.

In the determination of the values of the w_i only the charge density coefficients for the Hartree energy are used. The values of w_i calculated are then also used for the determination of the new charge density coefficients for the exchange–correlation intermediate fit. This is because the change in the contribution to the total energy from the Hartree energy is typically several orders of magnitude larger than that from the exchange–correlation energy. Therefore any major instabilities in convergence of the energy will be due to the Hartree energy. However, the Hartree energy cannot converge without the exchange–correlation energy also converging.

There are two situations which have been encountered in which this method fails to reach self–consistency.

(a) The first arises when the highest energy occupied level and the lowest energy unoccupied level have closely spaced energies and there is an even number of electrons. The lowest energy level of the two would be doubly occupied at the end of an iteration. The energy of this level can be lifted above the next highest level during the next iteration, if the energy difference is small enough. This is due to the electrostatic repulsion of the electrons in a localised state. At the end of this iteration the new lower energy level is now doubly occupied. This produces a discontinuous change in the total energy, with self consistency never being reached. In order to overcome this problem, a variable occupancy can be allowed as described in section 3.12.

(b) The second is due to an overcomplete basis in the Gaussian functions used in the intermediate fit to the charge density. If two Gaussian functions which are close to each other spatially have exponential decay constants which are too similar, then the matrix g_{kl} becomes almost singular. The determination of the inverse of this matrix then becomes unstable. To avoid this problem, the exponential coefficients for atoms are chosen to be amply different. This situation can still occur if an atom moves so that it is on top of a bond–centred fitting function.

3.8 Structural Optimisation

3.8.1 Minimisation Techniques

Using the method described up to this point we are able to find the self–consistent charge density and total energy for a cluster of atoms given their coordinates and atomic types. What we now wish to find is the minimum energy of the cluster with respect to the atomic positions, possibly with some constraints placed upon the atomic motion as will be discussed in section 3.10 and Chapters 7 and 8. In the minimisation of the total energy we employ two items of information, the total energy and the forces on the atoms (the derivative of the total energy with respect to atomic position). An explanation of the calculation of the forces is given in section 3.8.2.

A discussion of numerical methods for the minimisation of functions is given by Press (1992). The methods centre on a choice of a vector for a "search direction" along which the function is minimised (a line minimisation), this step being repeated until the value of the function does not change by more than a specified tolerance in a step. In this way a many parameter minimisation is simplified to a series of single parameter line minimisations. The algorithm for the line minimisation is discussed in section 3.9. The most obvious way in which to incorporate forces in the procedure is termed "the method of steepest descent". In this case each line minimisation is along the direction of the force. The problem associated with this approach is that the next search direction must be perpendicular to the one just searched along. This method is least efficient in a long, narrow "valley" in a function, when many line minimisations are generally required as the search directions zig–zag down the valley, each one orthogonal to the last.

We use a conjugate gradient method in order to generate the next search direction along which a line minimisation is to be carried out. The equation used to generate the next search direction, $\mathbf{D}^k_{i\alpha}$, is,

$$\mathbf{D}^k_{i\alpha} = \mathbf{F}^k_{i\alpha} + \frac{\left(|\mathbf{F}^k_{i\alpha}|^2 - \mathbf{F}^k_{i\alpha}.\mathbf{F}^{k-1}_{i\alpha}\right)\mathbf{D}^{k-1}_{k\alpha}}{|\mathbf{F}^{k-1}_{i\alpha}|^2}, \tag{3.97}$$

where $\mathbf{F}^k_{i\alpha}$ is the force vector in the k'th iteration of the minimisation. The vectors $\mathbf{F}^k_{i\alpha}$ and $\mathbf{D}^k_{i\alpha}$ have $3N$ components, where N is the number of atoms, being the force or search direction along each of the Cartesian axes, α, for each atom, i. This equation is derived and discussed by Press (1992). In this approach information from all previous chosen search directions is effectively utilised in order to ensure that each search direction is conjugate to all previous search directions and each force is orthogonal to all previous forces. This results from simply ensuring that each $\mathbf{F}^k_{i\alpha}$ is orthogonal to $\mathbf{F}^{k-1}_{i\alpha}$ and each $\mathbf{D}^k_{i\alpha}$ is conjugate to $\mathbf{D}^{k-1}_{i\alpha}$ and is proven by Polak (1971). This an advantage over the method of steepest descent in which the new search direction depends only on the previous search direction and so utilises increasingly less of the available information.

3.8.2 The Calculation of Forces

The forces on the atoms are calculated by the analytical determination of the derivative of the total energy with respect to the positions of the atoms, $\frac{\partial E_{tot}}{\partial \mathbf{R}}$. The self consistent total energy of a cluster in terms of the output charge density coefficients is given, with reference to (3.70), by,

$$E_{tot} = \left(T_{ij} + V^{ext}_{ij}\right) b_{ij} + \frac{1}{2} d_k d_l g_{kl} + E_{xc}(d^{xc}_k) + E_{ion-ion}. \tag{3.98}$$

Differentiating (3.98),

$$\frac{\partial E_{tot}}{\partial \mathbf{R}} = \left(\frac{\partial T_{ij}}{\partial \mathbf{R}} + \frac{\partial V^{ext}_{ij}}{\partial \mathbf{R}}\right) b_{ij} + \left(T_{ij} + V^{ext}_{ij}\right)\frac{\partial b_{ij}}{\partial \mathbf{R}} \\ + d_k g_{kl}\frac{\partial d_l}{\partial \mathbf{R}} + \frac{1}{2} d_k d_l \frac{\partial g_{kl}}{\partial \mathbf{R}} + \frac{\partial E_{xc}(d^{xc}_k)}{\partial \mathbf{R}} + \frac{\partial E_{ion-ion}}{\partial \mathbf{R}}. \tag{3.99}$$

Dealing with the term involving E_{xc}, differentiating (3.44),

$$\frac{\partial E_{xc}}{\partial \mathbf{R}} = \frac{\partial E_{xc}}{\partial d_k^{xc}}\frac{\partial d_k^{xc}}{\partial \mathbf{R}} + d_k^{xc}\epsilon_{xc}'(m_{1k})\exp\left[\frac{s_0(s_0-1)}{2}\ln\left(\frac{m_{2k}}{m_{1k}^2}\right)\right]\frac{\partial m_{1k}}{\partial \mathbf{R}}I_k$$
$$+\frac{s_0(s_0-1)}{2}d_k^{xc}\epsilon_{xc}(m_{1k})\exp\left[\frac{s_0(s_0-1)}{2}\ln\left(\frac{m_{2k}}{m_{1k}^2}\right)\right]$$
$$\times\left(\frac{1}{m_{2k}}\frac{\partial m_{2k}}{\partial \mathbf{R}} - \frac{2}{m_{1k}}\frac{\partial m_{1k}}{\partial \mathbf{R}}\right)I_k, \tag{3.100}$$

where, differentiating (3.27),

$$\frac{\partial m_{1k}}{\partial \mathbf{R}} = \frac{d_l^{xc}}{I_k}\frac{\partial H_{kl}}{\partial \mathbf{R}}, \tag{3.101}$$

and, differentiating (3.40),

$$\frac{\partial m_{2k}}{\partial \mathbf{R}} = \frac{d_l^{xc}d_m^{xc}}{I_k}\frac{\partial U_{klm}}{\partial \mathbf{R}}. \tag{3.102}$$

The first term in (3.100) involves $\frac{\partial E_{xc}}{\partial d_k^{xc}}$,which is obtained by differentiating (3.44),

$$\frac{\partial E_{xc}}{\partial d_k^{xc}} = \epsilon_{xc}(m_{1k})\exp(f_k)I_k + \sum_l d_l^{xc}\epsilon_{xc}'(m1l)\left(\frac{\partial m_{1l}}{\partial d_k^{xc}}\right)\exp(f_l)I_k$$
$$+\sum_l d_l^{xc}\epsilon_{xc}(m_{1l})\left(\frac{\partial f_{1l}}{\partial d_k^{xc}}\right)\exp(f_l)I_l, \tag{3.103}$$

where, differentiating (3.27),

$$\frac{\partial m_{1l}}{\partial d_k^{xc}} = \frac{H_{kl}}{I_k}, \tag{3.104}$$

and, differentiating (3.43),

$$\frac{\partial f_l}{\partial d_k^{xc}} = \frac{s_0(s_0-1)}{2}\left[\frac{2\sum_m d_m^{xc}U_{klm}}{\sum_{mn} d_m^{xc}d_n^{xc}U_{lmn}} - \frac{2}{m_{1l}}\frac{\partial m_{1l}}{\partial d_k^{xc}}\right]. \tag{3.105}$$

(The summations are explicitly included as l is not summed over in the last term.) Returning to (3.99), we require an expression for $\frac{\partial d_k}{\partial \mathbf{R}}$. Differentiating (3.20),

$$\frac{\partial d_k}{\partial \mathbf{R}} = (g^{-1})_{kl}\left(\frac{\partial t_{ijl}}{\partial \mathbf{R}}b_{ij} + t_{ijl}\frac{\partial b_{ij}}{\partial \mathbf{R}} - \frac{\partial g_{lm}}{\partial \mathbf{R}}d_m\right). \tag{3.106}$$

Similarly, in (3.100) we require an expression for $\dfrac{\partial d_k^{xc}}{\partial \mathbf{R}}$, so differentiating (3.51),

$$\frac{\partial d_k^{xc}}{\partial \mathbf{R}} = (H^{-1})_{kl}\left[\frac{3\alpha_k}{2\pi}\left(\frac{\partial t_{ijl}}{\partial \mathbf{R}}b_{ij} + t_{ijl}\frac{\partial b_{ij}}{\partial \mathbf{R}}\right) - \frac{\partial g_{lm}}{\partial \mathbf{R}}d_m\right]. \tag{3.107}$$

If we gather together the terms including $\dfrac{\partial b_{ij}}{\partial \mathbf{R}}$, we obtain,

$$\left(T_{ij} + V_{ij}^{ext} + d_k t_{ijk} + \frac{3\alpha_k}{2\pi}(H^{-1})t_{ijl}\frac{\partial E_{xc}}{\partial d_k^{xc}}\right)\frac{\partial b_{ij}}{\partial \mathbf{R}} = H_{ij}\frac{\partial b_{ij}}{\partial \mathbf{R}}. \tag{3.108}$$

Considering the right hand side of (3.108),

$$H_{ij}\frac{\partial b_{ij}}{\partial \mathbf{R}} = 2\sum_\lambda H_{ij}c_j^\lambda\frac{\partial c_i^\lambda}{\partial \mathbf{R}}. \tag{3.109}$$

Using (3.74),

$$\begin{aligned} H_{ij}\frac{\partial b_{ij}}{\partial \mathbf{R}} &= 2\sum_\lambda S_{ij}\epsilon^\lambda c_j^\lambda\frac{\partial c_i^\lambda}{\partial \mathbf{R}} \\ &= \sum_\lambda S_{ij}\epsilon^\lambda c_j^\lambda\frac{\partial}{\partial \mathbf{R}}\left(c_i^\lambda c_j^\lambda\right) \\ &= -\sum_\lambda \epsilon^\lambda c_i^\lambda c_j^\lambda\frac{\partial S_{ij}}{\partial \mathbf{R}}. \end{aligned} \tag{3.110}$$

Therefore, we do not need to evaluate $\dfrac{\partial b_{ij}}{\partial \mathbf{R}}$ in order to calculate the force on an atom.

Collecting the above equations together, we have the final expression for the force on an atom,

$$\begin{aligned}\frac{\partial E}{\partial \mathbf{R}} = \left[\frac{\partial T_{ij}}{\partial \mathbf{R}} + \frac{\partial V_{ij}^{ext}}{\partial \mathbf{R}}\right]b_{ij} + p_k^{(1)}b_{ij}\frac{\partial t_{ijk}}{\partial \mathbf{R}} - \frac{1}{2}d_k d_l\frac{\partial g_{kl}}{\partial \mathbf{R}} + d_l^{xc}(p_k^{(2)} - p_k^{(3)})\frac{\partial H_{kl}}{\partial \mathbf{R}} \\ +\frac{1}{2}e_k d_l^{xc}d_m^{xc}\frac{\partial U_{klm}}{\partial \mathbf{R}} - \sum_\lambda \epsilon^\lambda c_i^\lambda c_j^\lambda\frac{\partial S_{ij}}{\partial \mathbf{R}} + \frac{\partial E_{ion-ion}}{\partial \mathbf{R}},\end{aligned}$$

where the following vectors are defined,

$$p_k^{(1)} = d_k + \frac{3\alpha_k}{2\pi}p_k^{(3)}, \tag{3.111}$$

$$p_k^{(2)} = d_k^{xc}\exp(f_l)\left[\epsilon_{xc}{}'(m_{1k}) - \frac{s_0(s_0-1)\epsilon_{xc}(m_{1k})}{m_{1k}}\right], \tag{3.112}$$

$$p_k^{(3)} = p_k^{(4)}(H^{-1})_{kl} = \frac{\partial E_{xc}}{\partial d_k^{xc}}(H^{-1})_{kl}, \tag{3.113}$$

$$e_k = \frac{d_k^{xc}\epsilon_{xc}(m_{1k})\exp(f_l)s_0(s_0-1)}{m_{2k}}. \tag{3.114}$$

These vectors are in a convenient form for storage by the computer and use in the calculation of the forces on the atoms.

The analytical calculation of the forces acting on the atoms can lead to certain problems in the optimisation of the structure. For example, if an atom is located on the bond axis between two other atoms in a tetrahedral lattice, even if the total energy of the cluster would drop if the atom moved off the bond axis, there can be no force acting on an atom perpendicular to the bond. The atom would, therefore, remain on the bond axis. Various examples of this situation will be discussed as they arise in Chapters 5–8. In order to investigate whether, under such circumstances, the atom in question is at a minimum in the potential energy surface, the atom is displaced by a small amount (generally ≈ 0.05 Å) in the appropriate direction (perpendicular to the bond axis in the example given). There will then be a small component of the force acting on the atom in the required direction. Further optimisation of the structure will lead to the atom moving towards a minimum in the energy surface, including back to the original site if it was truly a minimum.

3.9 The Line Minimisation Subroutine

This subroutine minimises the total energy of the cluster with respect to the motion of the atoms along the search direction determined using (3.97). In the construction of the subroutine a balance has been struck between minimising the function more quickly than other more widespread and generalised numerical methods (for examples, see Press (1992)) whilst still being robust enough to successfully carry out the task in most cases. The subroutine is entitled CUBIC as the primary technique used for the minimisation is cubic interpolation or extrapolation.

The two points with the lowest energy are stored as x and w (the distance moved along the search direction). Also stored are the energies at these points, fx and fw, and the forces at these points, dx and dw. The points are ordered so that the lowest energy is stored in fx.

In the initiation of the minimisation, fx is set to the energy obtained from the minimisation along the previous search direction or, if this is the first minimisation iteration, the initial self consistent energy of the cluster. dx is set to the corresponding force and x to zero. A distance is moved along the search direction such that the maximum displacement of an atom is 0.2 a.u. The value of w is set to this distance, and the values of fw and dw are determined for this point. The procedure then followed by the CUBIC subroutine is shown in the flow chart on the following three pages.

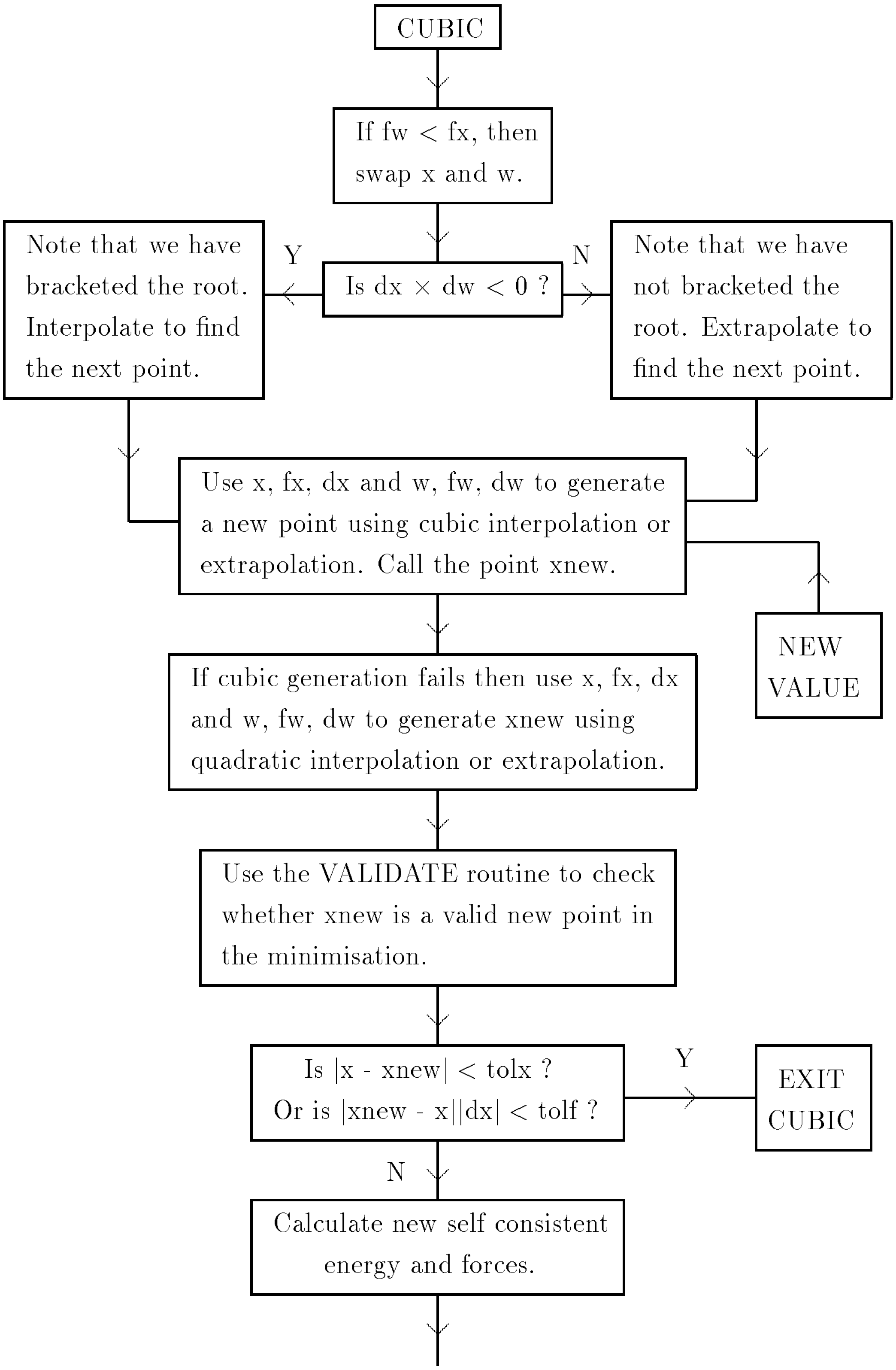
VALIDATE is a routine within CUBIC which is used to ensure that when xnew is generated its generation conforms to certain criteria. If xnew is validated then CUBIC continues.

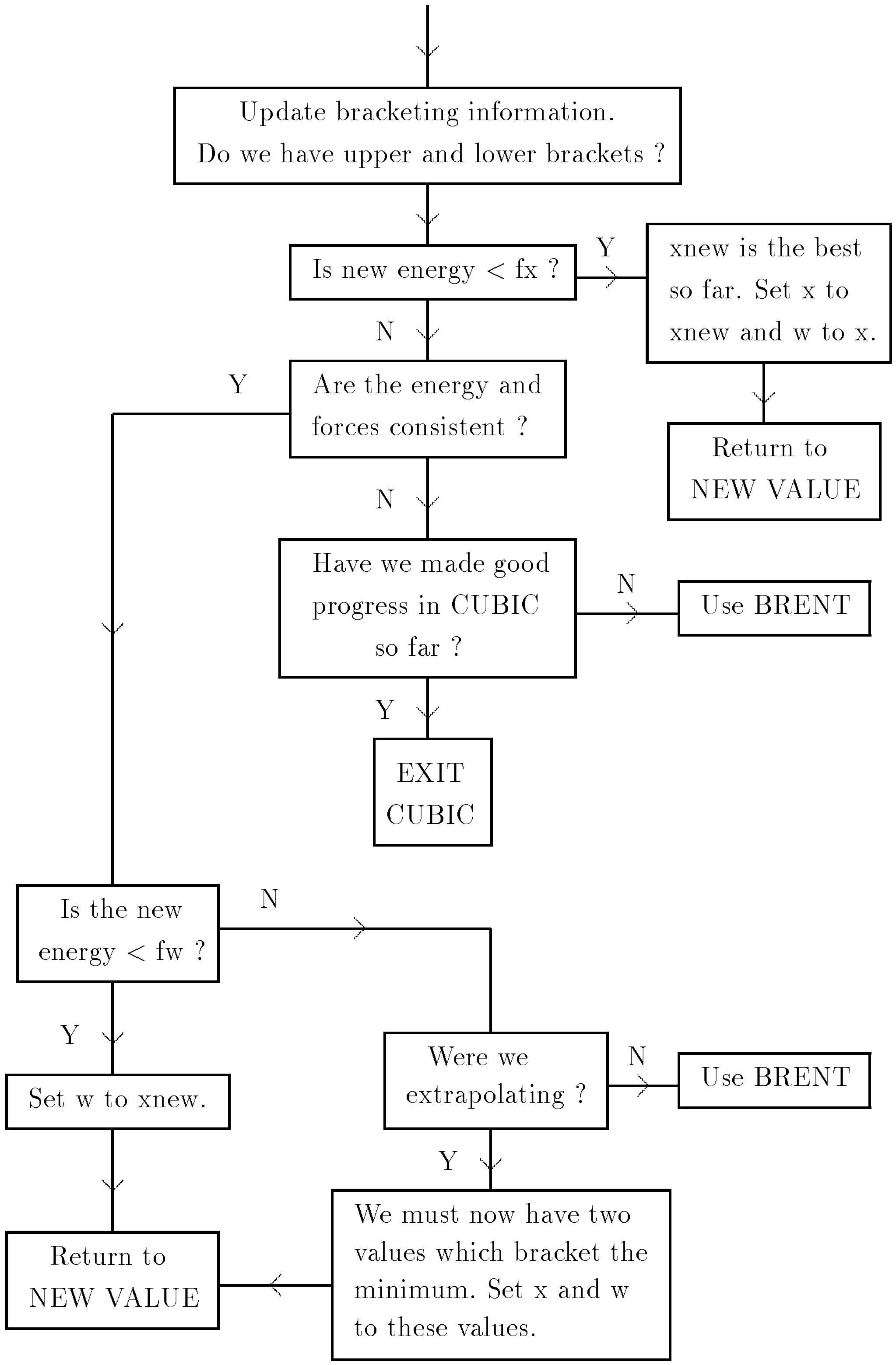
Possible end points of the CUBIC subroutine are as follows :

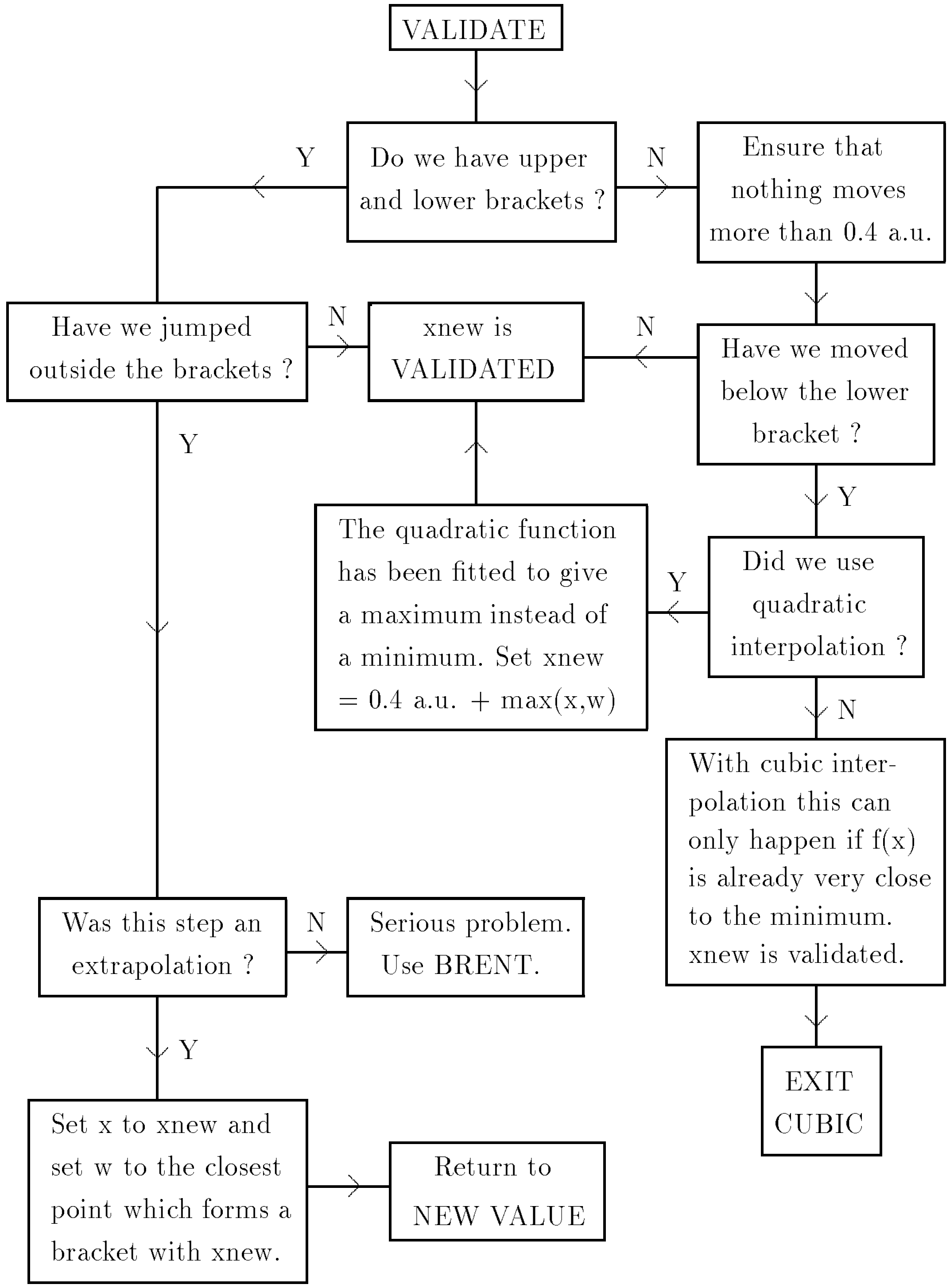
EXIT CUBIC – This denotes a successful exit from the subroutine.
Use BRENT – At this point another subroutine, BRENT, is jumped to which uses Brent's method for the minimisation (as described in Press (1992)). This method is slower but very robust and does not use the forces, thereby negating any problems due to inconsistencies between the forces and energies.

At certain points within CUBIC the subroutine jumps back to NEW VALUE. This is the point in the subroutine when a new point, xnew, is generated.

The two tolerances which are used in the subroutine in order to satisfy exit criteria, tolx and tolf, are variable and determined elsewhere in the program.


CUBIC
If fw < fx, then swap x and w.
Is dx × dw < 0 ?
Y
Note that we have bracketed the root. Interpolate to find the next point.
N
Note that we have not bracketed the root. Extrapolate to find the next point.
Use x, fx, dx and w, fw, dw to generate a new point using cubic interpolation or extrapolation. Call the point xnew.
NEW VALUE
If cubic generation fails then use x, fx, dx and w, fw, dw to generate xnew using quadratic interpolation or extrapolation.
Use the VALIDATE routine to check whether xnew is a valid new point in the minimisation.
Is |x - xnew| < tolx ?
Or is |xnew - x||dx| < tolf ?
Y
EXIT CUBIC
N
Calculate new self consistent energy and forces.


Update bracketing information.
Do we have upper and lower brackets ?
Is new energy < fx ?
Y
xnew is the best so far. Set x to xnew and w to x.
Return to NEW VALUE
N
Are the energy and forces consistent ?
Y
N
Have we made good progress in CUBIC so far ?
N
Use BRENT
Y
EXIT CUBIC
Is the new energy < fw ?
N
Y
Set w to xnew.
Were we extrapolating ?
N
Use BRENT
Y
We must now have two values which bracket the minimum. Set x and w to these values.
Return to NEW VALUE


VALIDATE
Do we have upper and lower brackets ?
Y
N
Ensure that nothing moves more than 0.4 a.u.
Have we jumped outside the brackets ?
N
xnew is VALIDATED
N
Have we moved below the lower bracket ?
Y
Y
The quadratic function has been fitted to give a maximum instead of a minimum. Set xnew = 0.4 a.u. + max(x,w)
Y
Did we use quadratic interpolation ?
N
With cubic interpolation this can only happen if f(x) is already very close to the minimum. xnew is validated.
Was this step an extrapolation ?
N
Serious problem. Use BRENT.
Y
EXIT CUBIC
Set x to xnew and set w to the closest point which forms a bracket with xnew.
Return to NEW VALUE

3.10 The Use of Symmetry Analysis

Incorporated into the self–consistent LDF program is a symmetry analysis subroutine. This determines the symmetry of the atomic cluster to within a predetermined tolerance. The analysis takes place by finding whether it is possible to map rotated and reflected cluster coordinates back onto the original coordinates and then assigning the symmetry group which corresponds to the possible mappings.

Once the symmetry has been determined, it is possible to minimise the total energy of the cluster with respect to the atomic positions with the constraint that the symmetry group of the cluster is retained. If a defect is proposed with a particular symmetry, then the imposition of this constraint is useful in order to find the equilibrium structure of the defect with that symmetry. When the optimised structure of the cluster has been found, the symmetry constraint can be removed and further optimisation carried out. If the proposed defect structure is not the most stable, then the atoms can then move to form a structure of lower total energy. Examples of this procedure occur in Chapters 5 and 6.

The determination of the symmetry of a cluster can also be used to increase the speed of calculation of the second derivatives of the total energy with respect to atomic position, $\frac{\partial^2 E^{tot}}{\partial \mathbf{R}_i \partial \mathbf{R}_j}$, the determination of which will be described in section 4.2.1. Only the minimum number of derivatives are then calculated and the symmetry mappings are used to generate the remaining values.

3.11 The Incorporation of Spin Polarisation

The formalism developed in sections 3.2 to 3.6 has been extended to include spin polarisation effects. These effects are very significant when dealing with certain atoms and defects. The charge density is split into spin up and spin down contributions,

$$n(\mathbf{r}) = n_{\uparrow}(\mathbf{r}) + n_{\downarrow}(\mathbf{r}). \tag{3.115}$$

The kinetic, Hartree and external potential energy terms are all dependent only on the total charge density so that,

$$T[n(\mathbf{r})] = T[n_{\uparrow}(\mathbf{r})] + T[n_{\downarrow}(\mathbf{r})] \qquad \text{etc.} \tag{3.116}$$

The single particle exchange–correlation energy, $\epsilon_{xc}(n)$, however, is explicitly dependent on the spin state and is replaced by $\epsilon_{xc}(n_{\uparrow}, n_{\downarrow})$. The spin polarised exchange–correlation energy is given by,

$$E_{xc} = \int [n_{\uparrow}(\mathbf{r})\epsilon_{xc}(n_{\uparrow}, n_{\downarrow}) + n_{\downarrow}(\mathbf{r})\epsilon_{xc}(n_{\downarrow}, n_{\uparrow})] \, d\mathbf{r}. \tag{3.117}$$

The Hamiltonian given by (3.6) for the spin up charge density is,

$$H_{ij}^{\uparrow} = T_{ij} + V_{ij}^{ext} + V_{ij}^{H} + \left(\mu_{ij}^{xc}\right)^{\uparrow}, \tag{3.118}$$

where,

$$\left(\mu_{ij}^{xc}\right)^{\uparrow} = \frac{\delta}{\delta n_{\uparrow}} [n_{\uparrow}(\mathbf{r})\epsilon_{xc}(n_{\uparrow}, n_{\downarrow}) + n_{\downarrow}(\mathbf{r})\epsilon_{xc}(n_{\downarrow}, n_{\uparrow})] \tag{3.119}$$

$$= \epsilon_{xc}(n_{\uparrow}, n_{\downarrow}) + n_{\uparrow} \frac{\delta\left[\epsilon_{xc}(n_{\uparrow}, n_{\downarrow})\right]}{\delta n_{\uparrow}} + n_{\downarrow} \frac{\delta\left[\epsilon_{xc}(n_{\downarrow}, n_{\uparrow})\right]}{\delta n_{\uparrow}}. \tag{3.120}$$

The spin up version of (3.74),

$$\sum_j H_{ij}^{\uparrow} c_{j\uparrow}^{\lambda} = \epsilon_{\uparrow}^{\lambda} \sum_j S_{ij} c_{j\uparrow}^{\lambda} \tag{3.121}$$

is solved to give the spin up wavefunction coefficients, $c_{j\uparrow}^{\lambda}$. The spin up version of (3.4) is constructed,

$$b_{ij}^{\uparrow} = \sum_{\lambda} f_{\lambda}^{\uparrow} c_{i\uparrow}^{\lambda} c_{j\uparrow}^{\lambda}, \tag{3.122}$$

and from this the coefficients of the intermediate fit to the spin up charge density for the Hartree energy are found,

$$c_{k\uparrow} = \left(g^{-1}\right)_{kl} t_{ijl} b_{ij}^{\uparrow}, \tag{3.123}$$

from the spin up equivalent of (3.21). The procedure from (3.118) to (3.123) is then carried out for the spin down charge density.

3.12 The Simulation of Finite Temperature

The application of a 'finite temperature' to a system is used to aid the convergence of the charge density coefficients in finding a self–consistent total energy, by allowing the partial occupancy of the electronic levels. This overcomes the convergence problem described in section 3.7. The Kohn–Sham eigenvalues, ϵ_λ, are determined as previously described. The number of electrons, which is held constant, is defined to be,

$$N = \sum_\lambda \frac{1}{e^{(\epsilon_\lambda - \mu)/kT} + 1}. \tag{3.124}$$

This is solved for μ using a numerical routine, ZBRENT, which uses the Van Wijngaarden–Dekker–Brent method and is described in Press (1992). The occupancies of the electronic energy levels are then determined according to the formula,

$$f_\lambda = \frac{1}{e^{(\epsilon_\lambda - \mu)/kT} + 1}. \tag{3.125}$$

The contribution to the total energy of the Kohn–Sham eigenvalues is then given by $\sum_\lambda f_\lambda \epsilon_\lambda$. The elements of b_{ij} are calculated as usual using (3.4). Finally an entropy term is added to the total energy. The definition of the entropy of a thermodynamic state is,

$$S = k \ln \Omega, \tag{3.126}$$

where Ω is the number of microstates of a system consistent with a particular macrostate. In order to find an expression for S, we assume that there are N non–interacting defects in a solid, so that there are N levels with energy ϵ_λ. We put m_λ electrons into these identical levels, where,

$$m_\lambda = f_\lambda N. \tag{3.127}$$

The statistical weight, Ω_λ for each set of N λ'th levels is,

$$\Omega_\lambda = \frac{N!}{m_\lambda!(N - m_\lambda)!}. \tag{3.128}$$

Therefore, substituting (3.128) into (3.126) and summing over λ, the entropy contribution for all levels is,

$$S = k\sum_{\lambda}\ln\left[\frac{N!}{m_{\lambda}!(N-m_{\lambda})!}\right]. \tag{3.129}$$

Stirling's approximation for x! is, to two terms,

$$x! \approx x\ln x - x + \ldots, \tag{3.130}$$

and is valid for large x. As we are assuming a large N, using (3.130), (3.129) becomes,

$$\begin{aligned} S &= k\sum_{\lambda}\left[N\ln N - m_{\lambda}\ln m_{\lambda} - (N-m_{\lambda})\ln(N-m_{\lambda})\right] \\ &= k\sum_{\lambda}\ln\left[\frac{N^{N}}{m_{\lambda}^{m_{\lambda}}(N-m_{\lambda})^{(N-m_{\lambda})}}\right]. \end{aligned} \tag{3.131}$$

Dividing the numerator and denominator by N^N and using (3.127),

$$S = k\sum_{\lambda}\ln\left[\frac{1}{f_{\lambda}^{Nf_{\lambda}}(1-f_{\lambda})^{N(1-f_{\lambda})}}\right], \tag{3.132}$$

so that finally, on rearranging, we obtain,

$$S = -kN\sum_{\lambda}\left[f_{\lambda}\ln f_{\lambda} + (1-f_{\lambda})\ln(1-f_{\lambda})\right]. \tag{3.133}$$

As (3.133) is the entropy contribution for N defects, we set $N = 1$ to obtain the entropy contribution of the single defect that we are modelling. The total energy of the system is replaced by the total free energy,

$$F_{tot} = E_{tot} - TS \tag{3.134}$$

$$= E_{tot} + kT\sum_{\lambda}\left[f_{\lambda}\ln f_{\lambda} + (1-f_{\lambda})\ln(1-f_{\lambda})\right]. \tag{3.135}$$

This is not an attempt to carry out meaningful finite temperature calculations. As stated above, it is an aid to obtaining a self–consistent total energy in certain circumstances. The highest occupied/lowest unoccupied electronic levels can cross with a smooth change in the total energy, rather than the discontinuous change which occurs with fixed electronic level occupancy. The total number of electrons is constrained to be as requested. In the case of spin polarised calculations this constraint is applied to both the total number of spin up and total number of spin down electrons specified. This leads to there being a different chemical potential for the spin up electrons to that for the spin down electrons. This is obviously not correct, however as the procedure is only used as a tool in converging the total energy, this is not important.

3.13 The Initiation of the calculation

As previously stated, the only input required to model a cluster of atoms are the initial coordinates of the atoms and their atomic types. However, what should we set the initial values of the charge density coefficients to?

Firstly we consider the case of an isolated atom. The determination of the exponents and coefficients for the charge density fitting functions in (3.23) and (3.49) for an atom proceeds as follows. The Kohn–Sham equations are solved by direct numerical integration for an atom in an exact LDA calculation. The procedure carried out by BHS is then followed to generate the pseudopotential and pseudo–charge density for the atom. This procedure need only be carried out once for each type of atom (defined by the atomic number). Tabulated values of this spherically symmetric charge density are fitted to combinations of Gaussian functions. Finally the coefficients and exponents of the fit to the charge density are optimised and fixed at the values obtained. This fixing of the exponents has been found to produce no significant change from calculations carried out in which they were allowed to vary.

In a full cluster calculation the charge density is initially set to be a linear combination of the atomic charge densities. No charge is initially located at the bond centres if bond–centred fitting functions are included.

We also use the coefficients of the orbitals in the wavefunction expansion of an isolated atom. The Kohn–Sham equations are solved, using the formalism described in this chapter, for an atom in order to determine the wavefunction. The coefficients and exponents of the Gaussian orbitals used as the basis for the wavefunction expansion are then optimised so that the total energy of the atom is minimised.

The wavefunction basis on atoms can be reduced in order to reduce the size of the Hamiltonian to a size for which it is tractable to perform calculations. This is carried out by taking a new basis as the linear combination of certain Gaussian orbitals. In the work presented in this thesis this is always the case with the hydrogen atoms terminating the clusters used. The basis is also reduced on certain other atoms in the clusters that we have used. The atoms closest to the terminating hydrogen atoms have their basis less severely contracted than the surface hydrogens. These atoms are those which can be considered to be outside the central defect region and are usually at least at third neighbour distance from the centre of the defect. More specific details of which orbitals are combined in these linear combinations are provided in Chapters 5 to 9 in the descriptions of calculations performed. The coefficients used in the construction of these linear combinations of Gaussian orbitals are those optimised values generated as described above.

Chapter 4

The Determination of Vibrational Properties

4.1 Introduction

Using the theory developed in Chapter 3, we are now in a position where we can determine the equilibrium structure of a cluster of atoms. By finding the total energies of various proposed structures for a defect in a cluster we can deduce which would have the lowest energy and so be preferentially formed in the material. This procedure will be used in Chapters 5 and 6. Once we have determined the microscopic structure in this way, in most cases we wish to determine properties which we can compare with the experimentally determined values. In this thesis we have analysed the vibrational properties of defects and determined the energy barriers to diffusion of particular defects in order to achieve this goal. The approach used in determining the paths and energies for migration is discussed in Chapter 7 and essentially involves the constrained motion of atoms.

In this Chapter we concentrate on the vibrational modes and frequencies of the atomic structures, which we determine in the harmonic approximation.

In section 4.3 we show how the frequencies of localised vibrational modes of atoms can be found by constructing and diagonalising the dynamical matrix. We then derive an expression for the phonon density of states using a Green function approach in section 4.4.3, which must be used when the vibrational modes of a defect cannot be assumed to be confined to a local area of the lattice. Finally, in section 4.4.4, an expression for the induced dipole moment caused by the motion of the atoms when charge transfer is induced is derived. This should correspond to the strength of the vibrational interaction with electromagnetic radiation. Electromagnetic radiation is an important experimental probe and infra–red and optical absorption experiments have been extensively used in the study of defects in solid state physics. By determining characteristic vibrational frequencies for defects we thus generate information directly comparable with experimental work.

4.2 Atomic Motion in the Harmonic Approximation

In our determination of the vibrational properties of atoms we assume the harmonic approximation for the motion of an atom about its equilibrium site. This can be expressed mathematically by expanding the potential energy of the atoms, Φ, as a function of their displacement, $U_{k\alpha}$, from their equilibrium positions as a Taylor series to three terms,

$$\Phi = \Phi^{(0)} + \sum_{k\alpha} \Phi^{(1)}_{k\alpha} U_{k\alpha} + \frac{1}{2} \sum_{kl\alpha\beta} \Phi^{(2)}_{kl\alpha\beta} U_{k\alpha} U_{l\beta}, \tag{4.1}$$

where k labels the atom and α labels the Cartesian direction. The first term of the expansion, $\Phi^{(0)}$, is the potential energy of the atoms in the equilibrium structure and is an arbitrary constant which we set to zero. The coefficient of the second term is the derivative of the potential energy at the equilibrium site or the force on the atom at this site,

$$\Phi_{k\alpha}^{(1)} = \frac{\partial \Phi}{\partial U_{k\alpha}}. \tag{4.2}$$

The equilibrium position of an atom must be a minimum in the potential energy surface and, by definition, the force on the atom at this point must be zero, so that this term is also zero. This leaves only the third term, the harmonic term, to contribute, the coefficient of which is,

$$\Phi_{kl\alpha\beta}^{(2)} = \frac{\partial^2 \Phi}{\partial U_{k\alpha} \partial U_{l\beta}}. \tag{4.3}$$

The expression for the potential energy for the atoms in this approximation is, therefore,

$$\Phi = \frac{1}{2} \sum_{kl\alpha\beta} \Phi_{kl\alpha\beta}^{(2)} U_{k\alpha} U_{l\beta}. \tag{4.4}$$

The Hamiltonian for the system can then be constructed using (4.4),

$$H = \frac{1}{2} \sum_{k\alpha} m_k \left(\frac{\partial U_{k\alpha}}{\partial t} \right)^2 + \frac{1}{2} \sum_{kl\alpha\beta} \Phi_{kl\alpha\beta}^{(2)} U_{k\alpha} U_{l\beta}, \tag{4.5}$$

where m_k is the mass of the k'th atom. From (4.5) we obtain the equations of motion for the atoms,

$$m_k \frac{\partial^2 U_{k\alpha}}{\partial t^2} = - \sum_{l\beta} \Phi_{kl\alpha\beta} U_{l\beta}. \tag{4.6}$$

4.2.1 The Determination of the Double Derivatives

We to need evaluate $\dfrac{\partial^2 \Phi}{\partial U_{k\alpha} \partial U_{l\beta}}$ in order to determine the potential energy terms. In the determination of these double derivatives use is made of the analytically determined forces on the atoms. In this way, effectively only the numerical evaluation of the derivative of the forces needs to be carried out. The expression used for this is,

$$\frac{\partial^2 E_{tot}}{\partial \mathbf{R}_{k\alpha} \partial \mathbf{R}_{l\beta}} = - \frac{F_{l\beta}[\delta_\alpha] - F_{l\beta}[-\delta_\alpha]}{2\delta_\alpha}, \tag{4.7}$$

where $F_{l\beta}[\delta_\alpha]$ is the force on atom l in the direction β when atom k is displaced along the direction α by δ_α. The procedure followed is self–evident from (4.7). Atom k is displaced along a direction α by δ_α and then $-\delta_\alpha$, in each case the force on atom l is evaluated along each of the three Cartesian directions β. These values are used in (4.7) to determine the required double derivatives. It has been found that the determination of the double derivatives is numerically stable for values of δ_α between 0.05 a.u. and 0.1 a.u. The value of δ_α has been set at 0.05 a.u. for the work carried out in this thesis.

4.3 Vibrational Properties 1 – The Dynamical Matrix

The first step in solving (4.6) is the choice of a trial solution. We choose,

$$U_{k\alpha}(t) = \frac{1}{\sqrt{m_k}} U'_{k\alpha} \exp(-i\omega t), \tag{4.8}$$

where ω is the frequency of vibration. Substituting (4.8) into (4.6),

$$\sum_{l\beta} U'_{l\beta} \frac{1}{\sqrt{m_k m_l}} \Phi_{kl\alpha\beta} = U'_{k\alpha}\omega^2. \tag{4.9}$$

Rearranging, we obtain,

$$\sum_{l\beta} \left[\mathbf{D}_{kl\alpha\beta} - \delta_{\alpha\beta}\delta_{kl}\omega^2\right] U'_{l\beta} = 0, \tag{4.10}$$

where we have defined the dynamical matrix as,

$$\mathbf{D}_{kl\alpha\beta} = \frac{1}{\sqrt{m_k m_l}} \Phi_{kl\alpha\beta}. \tag{4.11}$$

We have thus derived an expression from which the vibrational frequencies of N vibrating atoms can be determined. In order to do this the $3N \times 3N$ dynamical matrix is constructed and then diagonalised. An assumption implicit in this procedure is that the vibrational frequencies correspond to localised vibrational modes in which only the vibration of the atoms for which the dynamical matrix is constructed is of any significance. Resonances, in which the vibrational mode has strong coupling to the rest of the lattice, are dealt with in section 4.4.

4.4 Vibrational Properties 2 – Green Function Analysis

4.4.1 The Perfect Lattice Green Function

Considering (4.10), we can write the two equivalent expressions,

$$(\mathbf{\Phi} - \mathbf{M}\omega^2)\mathbf{U} = 0, \tag{4.12}$$

and,

$$(\mathbf{D} - \mathbf{I}\omega^2)\mathbf{U}' = 0, \tag{4.13}$$

where $\mathbf{M}$ is a diagonal matrix, the elements of which are the masses, m_k, corresponding to the atoms and $\mathbf{I}$ is the identity matrix. The vectors $\mathbf{U}$ and $\mathbf{U}'$ are connected by,

$$\mathbf{U} = (\mathbf{M})^{-1/2}\mathbf{U}'. \tag{4.14}$$

Using (4.12) and (4.13) we have two possible definitions of the perfect lattice phonon Green function,

$$(\mathbf{\Phi} - \mathbf{M}\omega^2)\mathbf{G}_0 = -\mathbf{I}, \tag{4.15}$$

and,

$$(\mathbf{D} - \mathbf{I}\omega^2)\mathbf{G}_0' = -\mathbf{I}. \tag{4.16}$$

Firstly (4.13) is used to determine the eigenvectors, $\mathbf{U}'$, and eigenvalues, ω, for the dynamical matrix constructed for the perfect lattice of the material in which the defect has been modelled. In the case of diamond, the evaluation is carried out in reciprocal space using a two atom unit cell to model the bulk material. In moving from using a cluster to a supercell, the interatomic force constants $\mathbf{\Phi}_{kl\alpha\beta}$ must be labelled according to which atoms within the unit cell are involved and which unit cells those atoms are located in. This dependence is included by replacing $\mathbf{\Phi}_{kl\alpha\beta}$ by $\Phi_{\alpha\beta}\begin{pmatrix} p_k & p_l \\ k & l \end{pmatrix}$, where this refers to the k'th and l'th atoms in the p_k'th and p_l'th unit cells respectively. Due to the translational invariance of the perfect diamond lattice if we substitute $p_k + m$ for p_k and $p_l + m$ for p_l, then Φ is identical. Therefore, if we use $-p_k$ as our value of m, we can put one of the atoms into the zeroth unit cell,

$$\Phi_{\alpha\beta}\begin{pmatrix} p_k & p_l \\ k & l \end{pmatrix} = \Phi_{\alpha\beta}\begin{pmatrix} 0 & p \\ k & l \end{pmatrix}, \tag{4.17}$$

where $p = p_l - p_k$. Turning to the displacement amplitude, $U'_{l\beta}$, this must now also be labelled as referring to the l'th atom in the p_l'th unit cell, $U'_{\beta}\begin{pmatrix} p_l \\ l \end{pmatrix}$. Due to the periodicity of the lattice, we can apply Bloch's theorem which relates the displacements of equivalent atoms in different supercells by a phase factor. If the displacements of the atoms in other unit cells are related to those of the atoms in the zeroth unit cell, then,

$$U'_{\beta}\begin{pmatrix} p_l \\ l \end{pmatrix} = U'_{\beta}\begin{pmatrix} 0 \\ l \end{pmatrix} \mathrm{e}^{(i\mathbf{q}\cdot\mathbf{p})} = U'_{l\beta}\mathrm{e}^{(i\mathbf{q}\cdot\mathbf{p})}, \tag{4.18}$$

where $\mathbf{p}$ is the vector connecting equivalent atoms in the zeroth and p_l'th unit cells and $\mathbf{q}$ is the wavevector propagating through the crystal. Substituting (4.17) and (4.18) into (4.9),

$$\sum_{l\beta} \frac{1}{\sqrt{m_k m_l}} \Phi_{\alpha\beta}\begin{pmatrix} 0 & p \\ k & l \end{pmatrix} U'_{l\beta}\mathrm{e}^{(i\mathbf{q}\cdot\mathbf{p})} = \delta_{\alpha\beta}\delta_{kl}\omega_n^2 U'_{k\alpha}. \tag{4.19}$$

If we now define the dynamical matrix to be,

$$\mathbf{D}_{kl\alpha\beta}(\mathbf{q}) = \sum_{p} \frac{1}{\sqrt{m_k m_l}} \Phi_{\alpha\beta}\begin{pmatrix} 0 & p \\ k & l \end{pmatrix} \mathrm{e}^{(i\mathbf{q}\cdot\mathbf{p})}, \tag{4.20}$$

we obtain the equivalent equation to (4.10) in reciprocal space,

$$\sum_{l\beta} \left[\mathbf{D}_{kl\alpha\beta}(\mathbf{q}, n) - \delta_{\alpha\beta}\delta_{kl}\omega_n^2(\mathbf{q})\right] U'_{l\beta}(\mathbf{q}, n) = 0, \tag{4.21}$$

where n labels the phonon branch. The dynamical matrix for the supercell is diagonalised for each q–point to determine the eigenvectors and eigenvalues, $U'_{l\beta}(\mathbf{q}, n)$ and $\omega_n^2(\mathbf{q})$. This is carried out for a mesh of up to 50,000 q–points in the irreducible wedge of the Brillouin zone. The interatomic potential is assumed to be short ranged, so that only the interatomic force constants between first or second nearest neighbours are evaluated and used to construct the dynamical matrix.

This procedure calculates the eigenvectors corresponding to the $\mathbf{U}'$ in (4.13). Then (4.14) is used to convert the $\mathbf{U}'$ into $\mathbf{U}$. Finally, we construct the perfect lattice Green function from,

$$\mathbf{G}_0(k\alpha, l\beta, \omega, \mathbf{p}) = \frac{1}{N}\sum_{\mathbf{q}n}\left(\frac{U_{k\alpha}(\mathbf{q}, n)U^*_{l\beta}(\mathbf{q}, n)}{\omega^2 - \omega_n^2(\mathbf{q})}\right) \mathrm{e}^{(i\mathbf{q}\cdot\mathbf{p})}, \tag{4.22}$$

where N is the number of q–points used. The phase factor, $\mathrm{e}^{(i\mathbf{q}\cdot\mathbf{p})}$, is included in order to construct the elements of the Green function for atoms outside the zeroth unit cell. This equation corresponds to the definition of the Green function given by (4.15). Once the perfect lattice Green function has been calculated it can be used for any defects modelled in that particular material.

4.4.2 The Perturbed Lattice Green Function

In order to construct a matrix to represent the perturbation caused by the the presence of the defect in the lattice, we use the equivalent cluster to that containing the defect, but of the perfect material. The interatomic force constant matrix, $\mathbf{\Phi}^0_{kl\alpha\beta}$, is constructed for those atoms in the perfect cluster corresponding to the atoms in the cluster containing the defect for which $\mathbf{\Phi}_{kl\alpha\beta}$ was determined. The perturbation due to the change in force constants of the interatomic bonding and/or the change in masses of the atoms when the perfect lattice is replaced by that containing a defect is then given by,

$$\mathbf{V}_{kl\alpha\beta} = \mathbf{\Phi}_{kl\alpha\beta} - \mathbf{\Phi}^0_{kl\alpha\beta} - (\mathbf{M} - \mathbf{M}^0)\omega^2\delta_{kl}\delta_{\alpha\beta}, \tag{4.23}$$

where $\mathbf{M}^0$ and $\mathbf{M}$ denote the atomic mass matrices for the perfect lattice and that containing the defect respectively.

In order to determine the perturbed lattice Green function, $\mathbf{G}$, we must solve Dyson's equation,

$$\mathbf{G} = \mathbf{G}_0 + \mathbf{G}_0\mathbf{V}\mathbf{G}. \tag{4.24}$$

By rearranging (4.24) we obtain the formal solution,

$$\mathbf{G} = \left[\frac{1}{1 - \mathbf{G}_0\mathbf{V}}\right]\mathbf{G}_0, \tag{4.25}$$

which is used to determine the perturbed lattice Green function for the system.

4.4.3 The Vibrational Density of States

We now derive an expression for the local vibrational density of states of a defect. Firstly, we state the following definition of the delta function,

$$\delta(x) = -\frac{1}{\pi} \lim_{\Gamma\to 0} Im\left[\frac{1}{x + i\Gamma}\right] \tag{4.26}$$

$$= \frac{1}{\pi} \lim_{\Gamma\to 0} \frac{\Gamma}{x^2 + \Gamma^2}. \tag{4.27}$$

This clearly satisfies the requirement that $\delta(x) = 0$, for $x \neq 0$, and the factor of $\frac{1}{\pi}$ occurs due the property that $\int \delta(x)dx = 1$ and,

$$\int_{-\infty}^{\infty} \frac{\Gamma}{x^2 + \Gamma^2}dx = \pi. \tag{4.28}$$

The vibrational local density of states is given by the expression,

$$\rho(k\alpha, \omega) = \sum_n |U^n_{k\alpha}|^2\delta(\omega - \omega_n) \tag{4.29}$$

Consider the imaginary part of a diagonal element of $\mathbf{G}$,

$$Im[\mathbf{G}(k\alpha, k\alpha, \omega)] = Im \sum_n \frac{|U_{k\alpha}^n|^2}{\omega^2 - \omega_n^2}. \tag{4.30}$$

We now apply a 'broadening' to the frequency by replacing ω by $\omega + i\Gamma$, in order to remove the singularities which occur in (4.30) when $\omega = \omega_n$,

$$Im[\mathbf{G}(k\alpha, k\alpha, \omega)] = \lim_{\Gamma \to 0} \; Im \sum_n \frac{|U_{k\alpha}^n|^2}{(\omega + i\Gamma)^2 - \omega_n^2}. \tag{4.31}$$

Expanding the denominator in (4.31), ignoring the term Γ^2, multiplying top and bottom by the complex conjugate of the new denominator and taking the imaginary part,

$$Im[\mathbf{G}(k\alpha, k\alpha, \omega)] = \lim_{\Gamma \to 0} \; \sum_n |U_{k\alpha}^n|^2 \frac{1}{2\omega} \left(\frac{-\Gamma}{(\omega^2 - \omega_n^2)^2/4\omega^2 + \Gamma^2} \right). \tag{4.32}$$

Using the definition of the delta–function given in (4.27),

$$\begin{aligned} Im[\mathbf{G}(k\alpha, k\alpha, \omega)] &= \sum_n |U_{k\alpha}^n|^2 \frac{\pi}{2\omega} \delta \left(\frac{\omega^2 - \omega_n^2}{2\omega} \right) \\ &= \sum_n |U_{k\alpha}^n|^2 \frac{\pi}{2\omega} \delta \left(\omega - \omega_n \right). \end{aligned} \tag{4.33}$$

Substituting (4.29) into (4.33),

$$Im[\mathbf{G}(k\alpha, k\alpha, \omega)] = -\frac{\pi}{2\omega} \rho(k\alpha, \omega). \tag{4.34}$$

Therefore, our final expression for the vibrational local density of states is,

$$\rho(k\alpha, \omega) = -\frac{2\omega}{\pi} Im[\mathbf{G}(k\alpha, k\alpha, \omega)]. \tag{4.35}$$

It is apparent that we only require the diagonal elements of the perturbed lattice Green function in order to determine the local vibrational density of states.

4.4.4 The Induced Dipole Moment

In order to determine which of the vibrational modes associated with a defect would couple strongly with electromagnetic radiation, a common experimental probe, we derive an expression for the induced dipole moment of a defect. This is a more accurate guide to what can be compared with experimental observations than the vibrational density of states as if a vibrational mode does not interact with electromagnetic radiation, then it is effectively invisible to experiment.

The induced dipole moment is given by the expression,

$$d_\alpha(\omega) = \sum_{kn} q_k U^n_{k\alpha} \delta(\omega - \omega_n), \tag{4.36}$$

where q_k is the induced charge on the k'th atom. The interaction of electromagnetic radiation with an electric dipole is proportional to the square of the dipole moment. Squaring (4.36),

$$|d_\alpha(\omega)|^2 = \sum_{klnn'} q_k q_l U^n_{k\alpha} U^{n'*}_{l\alpha} \delta(\omega - \omega_n)\delta(\omega - \omega_{n'}). \tag{4.37}$$

As a non–zero value of $|d_\alpha(\omega)|^2$ can only occur if $n = n'$ (due to the two delta functions), (4.37) reduces to,

$$|d_\alpha(\omega)|^2 = \sum_{kln} q_k q_l U^n_{k\alpha} U^{n*}_{l\alpha} \delta(\omega - \omega_n). \tag{4.38}$$

Using (4.33) in (4.38), we obtain,

$$|d_\alpha(\omega)|^2 = -\frac{2\omega}{\pi} Im[q_k q_l G(k\alpha, l\alpha, \omega)], \tag{4.39}$$

from which we can calculate the induced dipole moment using the perturbed lattice Green function.

In order to utilise (4.39) consideration must be given to what charges should be placed on the atoms. Two criteria have been followed in this work. The charges on equivalent atoms must be equal and the magnitude of the total charge must be zero. Taking the case of the substitutional boron atom which will be discussed in Chapter 5 as our example. A charge of $+q$ was placed on the boron atom and charges of $-\frac{q}{4}$ were placed on each of the four equivalent carbon nearest neighbours in the determination of the induced dipole moment.

Chapter 5

Substitutional Boron and the Boron–Hydrogen Complex

5.1 Introduction

Boron is one of the most common impurities introduced into diamond. As a dopant in diamond boron adds an acceptor level and it is used to make the p–type material. Research continues into its use as a dopant, especially by its incorporation in CVD grown thin diamond films (Malta 1993). There has been much work published on various properties of boron doped diamond, but not on the microscopic properties of individual boron substitutional defects.

Hydrogen is also a common impurity in diamond and much work has been carried out on interstitial hydrogen in the lattice, including a theoretical study using the *ab initio* technique employed in this thesis (Briddon 1988). The implantation of boron atoms into thin films causes widespread microscopic structural damage. One of the post–implantation annealing methods currently being investigated is the exposure of the material to a hydrogen plasma (Mori 1993). In this case especially, there would be a high concentration of both boron and hydrogen atoms in the lattice. Hydrogen diffusing through diamond has a singly occupied level in the band gap. In the presence of significant concentrations of boron atoms the hydrogen would lose its electron and become positively charged and so attracted towards the boron, eventually forming a defect complex. There has been no experimental or theoretical work published on the structure or vibrational properties of this possible defect complex to our knowledge, apart from a paper describing the results presented in this chapter (Breuer 1994).

In recent years extensive investigations have been carried out into the structures of the complexes formed by the passivation of shallow donors and acceptors by hydrogen. This process has therefore been well documented for many defects in semiconductors. Without exception, these have all been found to have C_{3v} symmetry. It has been found that when hydrogen forms complexes with acceptor impurities it prefers to sit approximately at the centre of one of the bonds between the acceptor and a host atom (Fig. 5.1). For example, this has been shown to be the case for the boron–hydrogen complex in silicon (Denteneer 1989) and the beryllium–hydrogen complex in GaAs (Briddon 1990a). In contrast, hydrogen atoms attach themselves in anti–bonding directions when forming complexes with donor impurities (Figs. 5.2 and 5.3). For example, in GaAs, a hydrogen atom bonds to a Si(Ga) in this position (Briddon 1990a) and in silicon it attaches itself in an anti–bonding direction to a silicon neighbour of a substitutional phosphorus atom (Denteneer 1990).

In this chapter we study the structure and vibrational properties of the point defect due to the substitutional boron atom in diamond. We then investigate the structure of the defect complex formed when the boron is passivated by hydrogen.

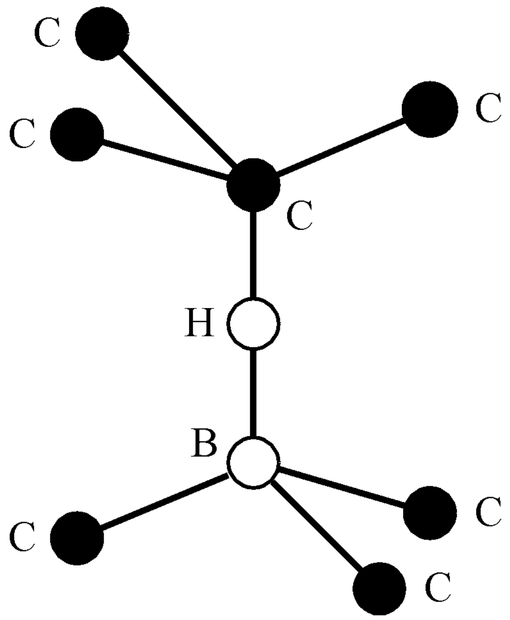


Figure 5.1: The hydrogen atom at the bond–centred site between a boron atom and a carbon atom, a possible structure for the boron–hydrogen complex in diamond.

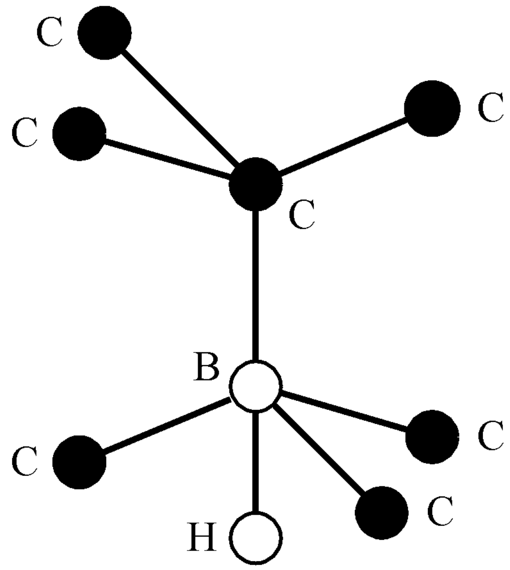


Figure 5.2: The hydrogen atom attached to the boron atom at the anti–bonding site, a possible structure for the boron–hydrogen complex in diamond.

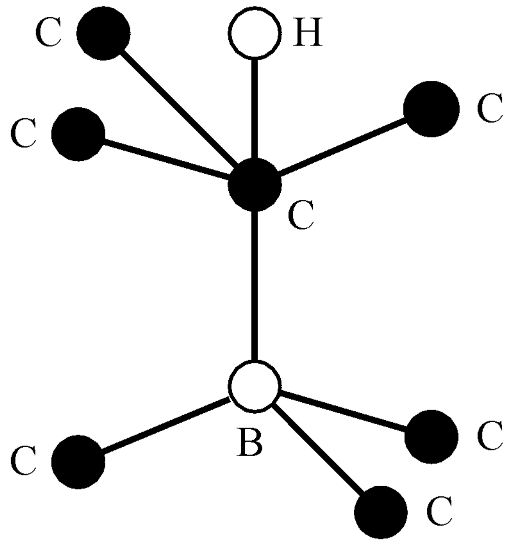


Figure 5.3: The hydrogen atom attached to a carbon nearest neighbour of the boron atom at the anti–bonding site, a possible structure for the boron–hydrogen complex in diamond.

5.2 The Clusters and Bases Used

The basic cluster that was used in this investigation is $C_{70}H_{60}B$ (Fig. 5.4) which comprises 131 atoms, including 60 terminating hydrogen atoms. The cluster has tetrahedral symmetry before any atomic motion is allowed, the boron atom is at the substitutional site at the centre of the cluster with the other atoms at the perfect diamond lattice positions. The defect has the terminating hydrogen atoms as 4th nearest neighbours. The pseudopotentials of Bachelet (1982) are used for the boron and carbon atoms and the bare coulombic potential is used for the hydrogens.

The charge density and wavefunction of the cluster are fitted using bases of Gaussian–type orbitals centred on the atoms and also on the centres of particular bonds as described in the following sections. The bases centred on the individual atoms and bond–centres are as follows.

For the charge density :

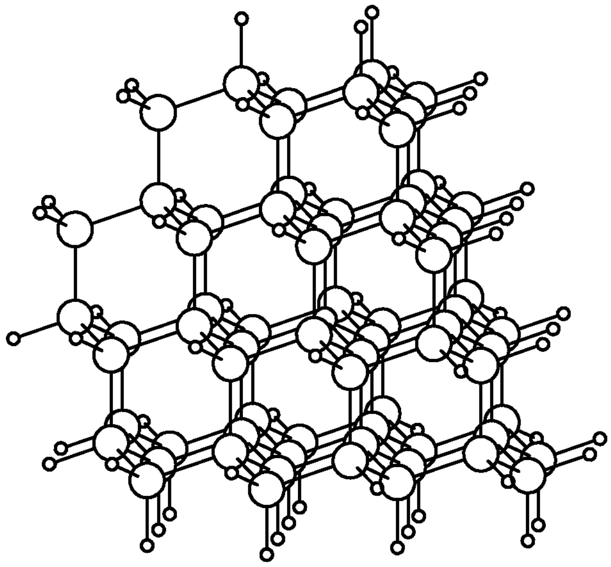


Figure 5.4: The 131 atom cluster $C_{70}H_{60}B$ used in the calculations. The hydrogens are the atoms with the smallest radii, at the surface of the cluster, and the boron atom is at the centre of the cluster with a smaller radius than that of the carbon atoms.

The boron atom has 5 orbitals of different widths.
The carbon atoms have 4 orbitals of different widths.
The terminating hydrogen atoms have 3 orbitals of different widths.
The passivating hydrogen atom has 4 orbitals of different widths.
The bond centres have 1 orbital.

For the wavefunction :

The boron atom has 16 orbitals (s,p_x,p_y and p_z type orbitals each of 4 different widths).
The full basis carbon atoms have 16 orbitals (s,p_x,p_y and p_z type orbitals each of 4 different widths).
The contracted basis carbon atoms have 4 orbitals (Each a linear combination of a particular type of orbital (s,p_x,p_y and p_z) each of which has 4 different

widths).

The terminating hydrogen atoms have 2 orbitals (A linear combination of the s–type orbitals of two different widths and a linear combination of the p–type orbitals (p_x,p_y and p_z) each of two different widths).

The passivating hydrogen atom has 16 orbitals (s,p_x,p_y and p_z type orbitals each of 4 different widths).

The bond centres have 4 orbitals (An s,p_x,p_y and p_z type orbital each of the same width).

The innermost 16 carbon atoms were in full basis (i.e. the first and second neighbours of the boron atom) and the remaining carbon atoms were in the contracted basis.

5.3 The Substitutional Boron Atom

We initially investigated the equilibrium structure of the substitutional boron defect. The 131 atom cluster described in section 5.2 was used with bond–centred fitting functions located on the 16 bonds between the innermost 17 atoms, comprising the boron atom and its first and second nearest neighbours. These same atoms were allowed to move in the minimisation of the total energy of the cluster. This defect was studied in the −1 charge state as would be observed experimentally either if the acceptor was ionised or in the presence of significant concentrations of other defects. In any case, it is well known that structural and vibrational properties of shallow impurities are not sensitively dependent on the charge state.

When the structure was relaxed the atoms surrounding the boron moved outwards, retaining the tetrahedral symmetry, with the boron atom remaining on–centre. The bonds to the first nearest neighbours of the boron increased by 3% and the bonds from those atoms to the next shell of carbon atoms increased by 1%. This corresponded to an outwards displacement of 0.05 Å for the first shell of atoms and an outwards displacement of 0.04 Å for the second shell of atoms.

In order to confirm that the substitutional boron atom is an on–site defect it was displaced from its central position by distances of up to 1 Å. These displacements were in the anti–bonding direction (towards the tetrahedral interstitial site) and along the $\langle 100 \rangle$ direction. In each case the boron was initially fixed at the displaced position whilst the surrounding 16 atoms were allowed to move until the equilibrium structure under this constraint was found. Then all 17 atoms, including the boron, were once again allowed to move, with the result that the boron atom always moved back to the on–centre position with the tetrahedral symmetry being restored. This implies that substitutional boron is an on–site defect as expected.

5.3.1 The Vibrational Properties of the Defect

The second derivatives of the total energy with respect to displacements of the boron and its four nearest neighbours were evaluated giving a 15×15 matrix of interatomic force constants, as described in section 4.2.1. The dynamical matrix was then constructed and diagonalised (as in section 4.3) in order to determine the vibrational frequencies associated with the motion of the boron atom, assuming that the vibrational modes are localised. The mass of a boron atom is less than that of a carbon atom. However, the longer bonds of the defect lead to the force constants for the boron–carbon bonds being 5% weaker than those for carbon–carbon bonds in pure diamond. We found that the resultant effect was that the frequency associated with the vibration of the substitutional boron was below the Raman frequency of diamond.

A Green function technique was therefore used to find the vibrational spectrum of this defect as a perturbation from a perfect 131 atom diamond cluster, as described in section 4.4. Second derivatives of the total energy with respect to displacements of the inner five atoms of the pure diamond cluster, $C_{71}H_{60}$ (with the boron atom in Fig. 5.4 replaced by a carbon atom), were determined to construct a 15×15 matrix of interatomic force constants for the perfect diamond lattice. The difference between this matrix and the 15×15 matrix for the relaxed substitutional boron then describes the perturbation caused by the defect. The perfect lattice Green function was constructed and Dyson's equation was solved in order to determine the Green function for the perturbed lattice. Using the relation derived in section 4.4.3 the vibrational density of states of the relaxed defect was calculated (Fig. 5.5(a)). The induced dipole moment

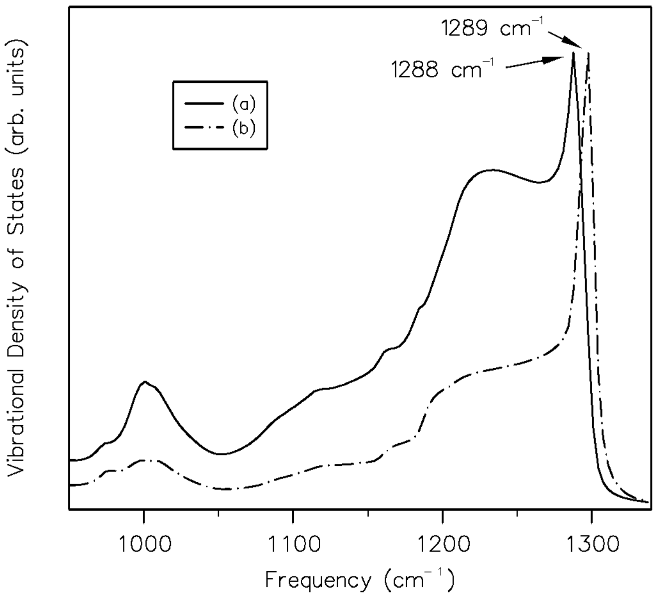


Figure 5.5: The vibrational density of states for the substitutional boron atom. The units are arbitrary and the results have been scaled so that the two main peaks have the same height. The results for the relaxed defect are shown in (a) and those for the mass defect approximation are shown in (b).

produced by the vibration of this defect was also calculated by following the method described in section 4.4.4 (Fig. 5.6(a)). This should correspond to the

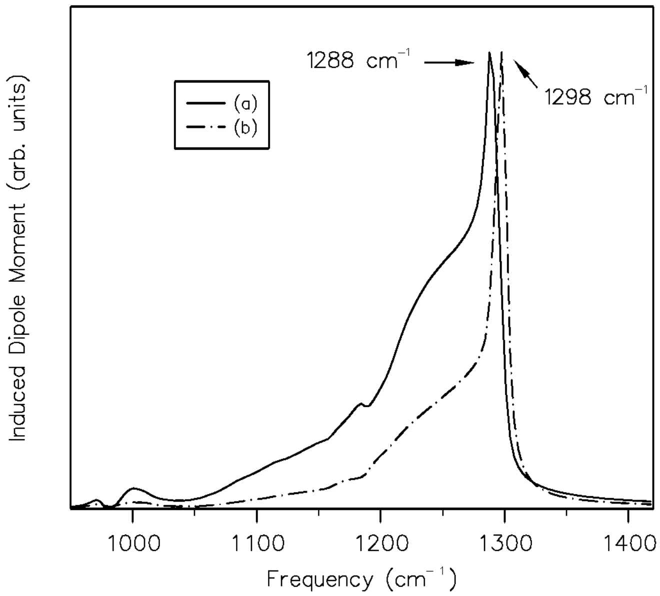


Figure 5.6: The induced dipole moment for the substitutional boron atom. The units are arbitrary and the results have been scaled so that the two main peaks have the same height. The results for the relaxed defect are shown in (a) and those for the mass defect approximation are shown in (b).

infra–red absorption spectrum of a diamond lattice containing the defect. The spectrum is overwhelmingly dominated by a peak at 1288 cm^{-1}. This is just below the Raman frequency at 1332 cm^{-1} so we would predict that substitutional boron gives rise to a resonance, not a local mode, even though boron is lighter than carbon.

Similar calculations were carried out with the boron treated as a mass defect only, without taking the change in interatomic couplings into account (Fig. 5.5(b) and 5.6(b)). The interatomic force constant matrix for the perfect diamond lattice was used, so that the only perturbation present was the change in mass of a carbon atom to the mass of a boron atom. It is seen that the shape of the absorption in Fig. 5.6(b) is somewhat different to that in Fig. 5.6(a) but that the position of the peak is correctly given by this well known approximation. This is to be expected as the structural relaxation around the substitutional boron was no more than 3%.

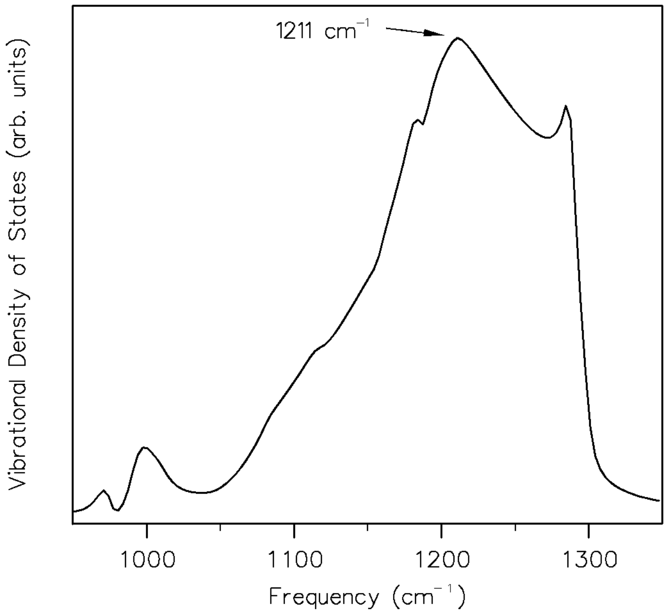


Figure 5.7: The vibrational density of states for the substitutional boron atom using a mass of 11 in order to investigate the isotopic shift.

The isotopic shift in the peak due to replacing ^{10}B by ^{11}B was investigated. The Green function analysis was carried out using a mass of 11 for the boron atom in the relaxed structure and the vibrational density of states was determined (Fig. 5.7). It was found that the peak originally at 1288 cm^{-1} was much reduced in magnitude so that the broad peak at 1211 cm^{-1} dominated the spectrum. This peak is around the same frequency as that for the vibrational density of states for pure diamond (1184 cm^{-1}), which was also calculated (Fig. 5.8). The fact that the vibrational density of states begins to resemble more that of bulk diamond is a reflection on the effectively small size of the perturbation from pure diamond of a substitutional ^{11}B atom.

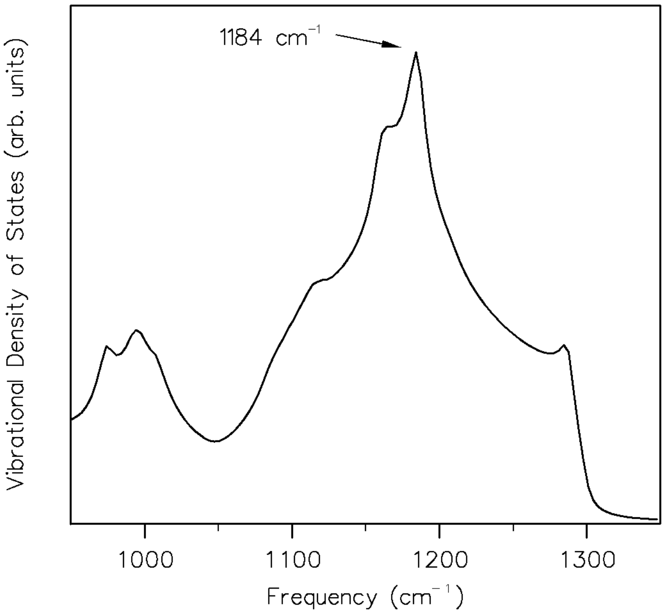


Figure 5.8: The vibrational density of states for the perfect diamond lattice.

5.4 The Boron–Hydrogen Complex

In order to study the boron–hydrogen complex a passivating hydrogen atom was added to the 131 atom cluster. Bond–centred fitting functions were located at the centre of the bonds between the innermost 17 atoms, not including one of the boron–carbon bonds and any bond involving the passivating hydrogen atom, for the initial investigations of the structure of the complex (15 in total). Bond–centred fitting functions were omitted from one of the boron–carbon bonds so that the hydrogen atom could be inserted into that bond without the risk of having an overcomplete basis. No fitting functions were located on bonds between the hydrogen and its nearest neighbours as at different sites there were different numbers of nearest neighbours and an equal total basis of fitting functions is required for each of the investigated structures in order to compare the total energies of the defects. Also, it was possible that the hydrogen atom would move a significant distance during unconstrained optimisation of the structure, so that the final nearest neighbours were unknown. The innermost 18 atoms were free to move in the study of this defect unless otherwise stated.

As discussed in the introduction, all previous acceptor–hydrogen complexes have been found to have C_{3v} symmetry and so the initial attempts to model the boron–hydrogen complex were based on structures with that symmetry. The C_{3v} symmetry axis in the cluster containing the defects is along a $\langle 111 \rangle$ direction. However, for clarity when discussing these defects in specific terms, this axis will be referred to as being aligned along the [111] direction and other vectors will be given in a similarly unique manner. It is, of course, implicit that these discussions hold true for any equivalent directions in a tetrahedral lattice.

The hydrogen atom was put in the centre of one of the four equivalent boron–carbon bonds (Fig. 5.1). A relaxation was performed with the constraint that the C_{3v} symmetry was preserved, so that the hydrogen was forced to remain on the bond axis, and the equilibrium structure of the cluster was found. The boron–hydrogen bond in this configuration was 1.2 Å and the carbon–hydrogen bond was 1.0 Å. Due to the forces acting on the atoms being calculated by an analytical method, it is impossible for there to be a component of the force acting on the hydrogen atom perpendicular to the bond axis on which it sits (as discussed in section 3.8.2). In order to impart a small force in this direction, the hydrogen atom was displaced perpendicular to the bond axis by 0.05 Å in the $[1\bar{1}0]$ direction with the other atoms left in their relaxed positions. When the atoms were again allowed to move the hydrogen atom moved completely out of the bond. It took up a position in virtually a [100] direction relative to the boron atom with a bond length of 1.2 Å (Fig 5.9). The boron itself was displaced in essentially the $[\bar{1}00]$ direction by 0.1 Å. The resultant energy drop in the structural relaxation from the relaxed bond–centred defect to the approximately [100] oriented defect was 0.7 eV.

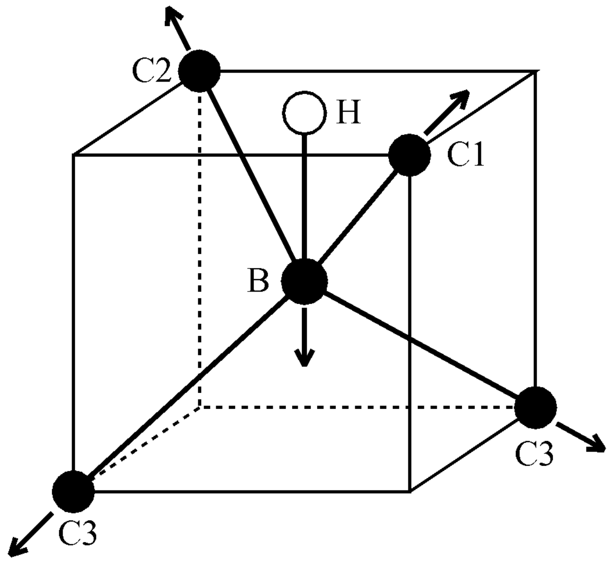


Figure 5.9: The $\langle 100 \rangle$ aligned boron–hydrogen defect and the first nearest neighbour carbon atoms, with labelling as used in Table. 5.1). The carbon atoms are labelled C1,C2 and C3, depending on the equivalence of the their positions in the structure. The arrows indicate the direction in which the important relaxations of the atoms take place.

The hydrogen atom was also put at the anti–bonding site closest to the boron atom (Fig. 5.2). On relaxation with the symmetry constrained to remain C_{3v} the equilibrium structure was determined. This had a boron–hydrogen bond length of 1.1 Å. Due to the analytical calculation of the forces, the hydrogen atom was again displaced perpendicular to the C_{3v} axis by 0.05 Å in the $[1\bar{1}0]$ direction with the other atoms left in their relaxed positions. On further relaxation the hydrogen again moved to essentially a [100] displacement from the boron. The energy drop in this case was 1.2 eV.

Finally, the hydrogen atom was put at the anti–bonding site closest to a carbon nearest neighbour of the boron atom (Fig. 5.3). The relaxed structure determined with C_{3v} symmetry had an energy 0.6 eV higher than that of the approximately [100] oriented defect. The carbon–hydrogen bond length in this case was 1.1 Å. On the displacement of the hydrogen atom by 0.05 Å perpendicular to the C_{3v} symmetry axis and further relaxation the hydrogen atom moved to a site close to the centre of one of the bonds between the carbon nearest neighbour of the boron and a carbon in the next shell of atoms out. This can be explained as being due to the preference of an isolated hydrogen atom in diamond to move to a site of high charge density in the diamond lattice. This bond–centre is one of three equivalent high charge density sites nearest to the hydrogen at the anti–bonding site. Briddon (1988) demonstrated that the most stable site for a hydrogen atom in pure diamond was at a carbon–carbon bond centre. It is expected that in the real material the hydrogen would diffuse to the negatively charged boron atom by moving over potential barriers. The significant point is that this anti–bonding site has a higher energy than the approximately [100] oriented site and is unstable.

All three of the above constrained structural relaxations were repeated without any bond–centred fitting functions. It was found that the bonds involving the hydrogen atom were unchanged to 0.01 Å and that the total energy relative to the approximately [100] oriented defect (also relaxed without bond–centred fitting functions) changed by at most 0.02 eV.

The actual displacement of the hydrogen from the substitutional site was along the [611] direction, with the boron also displaced from the substitutional site along the [$\bar{6}$11] direction. This structure, the exact [100] aligned defect, equivalent structures and others formed by slight perturbations in the position of the hydrogen atom were then investigated. Significant motion of the hydrogen atom was not expected, as this structure is essentially the equilibrium one, so bond centred fitting functions were located between all the innermost 17 atoms. In the case of the hydrogen atom, the fitting functions were located between it and its three nearest neighbours, giving a total of 19 bonds with fitting functions. The boron–hydrogen defect was aligned along the [6$\bar{1}\bar{1}$] direction which was expected to be equivalent to the [611] aligned defect, and on relaxation produced a structure with the same energy. The exact [100] aligned defect structure was relaxed, and retained its symmetry, again due to the analytical determination of the forces The defect was also aligned along the $[61\bar{1}]$ and equivalent $[6\bar{1}1]$ directions, which were displaced from the [100] aligned defect in a direction perpendicular to the [611] and equivalent $[6\bar{1}\bar{1}]$ structures. On allowing the atoms to move these two structures relaxed to the [100] aligned defect. The structure of both the $\langle 100 \rangle$ and $\langle 611 \rangle$ aligned structures is given in Table 5.1, with reference to Fig. 5.9.

	$\langle 100 \rangle$	$\langle 611 \rangle$
Bond Lengths		
B–H	1.19 Å	1.20 Å
B–C1	1.68 Å	1.62 Å
B–C2	1.68 Å	1.77 Å
B–C3	1.54 Å	1.54 Å
Bond Angles		
C1–B–C2	104.2°	103.6°
C3–B–C3	115.7°	116.1°
C1–B–H	52.1°	58.0°
C2–B–H	52.1°	45.6°

Table 5.1: The bond lengths and angles corresponding to the structures of the relaxed $\langle 100 \rangle$ and $\langle 611 \rangle$ oriented boron–hydrogen complex structures as shown in Fig. 5.9.

It therefore appeared that small displacements of the hydrogen atom from along the [100] direction in the [011] and $[0\bar{1}\bar{1}]$ directions gave a small, but in fact physically insignificant, drop in the total energy. Any perturbations of the hydrogen relaxed to the [100] or [611] (or equivalent $[6\bar{1}\bar{1}]$) aligned structures.

The calculations showed that the energy surface around the structure which minimised the total energy had smaller scale detail. In order to investigate the energy change as a function of displacement of the hydrogen along the [011] and $[0\bar{1}\bar{1}]$ directions, the hydrogen atom was fixed at various fractions along the vector connecting its position in the [100] and [611] or $[6\bar{1}\bar{1}]$ structures. The other 17 innermost atoms were allowed to move under the constraint that the hydrogen remained fixed at this site. The energy of each relaxed structure relative to the [100] aligned boron–hydrogen defect was determined and is shown in Fig. 5.10. As displacements perpendicular to those shown in

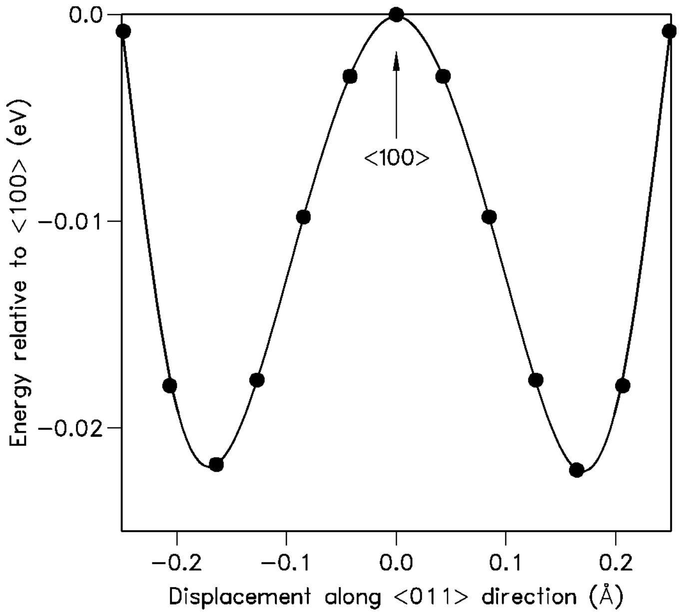


Figure 5.10: The small–scale variation of the energy of the boron–hydrogen defect with displacement of the hydrogen atom along the [011] and $[0\bar{1}\bar{1}]$ directions from the exact $\langle 100\rangle$ aligned defect. The energies are relative to that with the hydrogen aligned along the $\langle 100\rangle$.

Fig. 5.10 relaxed to the [100] structure, in 3D the $\langle 100\rangle$ direction alignment of the defect must be a saddle point.

The energy difference between the maxima and minima in Fig. 5.10 is only ≈ 0.02 eV which is smaller than the estimated accuracy of the energy surface. In any case, whether this is an artifact of the calculation or not is irrelevant as the energy changes involved are smaller than the zero point energy of the defect. The observed physical system is expected to be the hydrogen moving on a flat–bottomed 'valley' in the energy surface centred on the $\langle 100\rangle$ displacement from the boron and aligned along the corresponding $\langle 011\rangle$ and $\langle 0\bar{1}\bar{1}\rangle$ directions. The purpose of describing the details of the energy surface is to illustrate the difficulty in determining the vibrational properties of the defect with atomic motion having such a high degree of anharmonicity.

5.4.1 The Vibrational Properties of the Defect

We now consider the vibrational properties of the boron–hydrogen complex. Three modes are expected to be principally associated with the motion of the hydrogen atom. The stretching of the boron–hydrogen bond should have the highest frequency of these, and there should be two modes associated with the perpendicular vibration of the atom along the flat–bottomed energy surface, known as 'wagging' modes. The determination of the vibrational frequencies is made more complicated by the shape of the energy surface because the harmonic approximation cannot be used for the 'wagging' modes. However, motion of the hydrogen atom along the $\langle 100\rangle$ or $\langle 611\rangle$ directions is almost decoupled from motion in a perpendicular direction. The frequency of the mode associated with the stretching of the boron–hydrogen bond could therefore be determined independently of the 'wagging' modes.

The second derivatives of the total energy with respect to displacements of the innermost six atoms were evaluated. The dynamical matrix was then constructed and diagonalised, as described in 4.3, in order to determine the vibrational frequency associated with the stretching of the boron-hydrogen bond. The result obtained was 2540 cm^{-1} for the structure of minimum total energy, i.e. the $\langle 611 \rangle$ aligned defect. This is far lower than a typical carbon–hydrogen bond–stretching mode and indicates that the hydrogen atom in this complex is only weakly bonded to the boron atom. When the hydrogen atom was forced to lie exactly along the $\langle 100 \rangle$ direction, the frequency increased slightly to 2655 cm^{-1}. The Green function technique was also used for this defect to find whether any other localised vibrational modes are present after the passivation of the substitutional boron. The results from the Green function analysis showed that no other modes appeared above the Raman frequency due to the addition of the hydrogen atom. The boron–hydrogen bond stretching mode was therefore the only localised vibrational mode associated with this defect. The flat shape of the energy surface for motion of the hydrogen perpendicular to this suggests that the 'wagging' modes should have anomolously low frequencies.

5.5 Conclusions

The substitutional boron atom in diamond remains on site when relaxed from different starting positions. There are no localised vibrational modes associated with this defect even though boron is lighter than carbon. Instead it gives rise to a resonance at 1288 cm^{-1}. This frequency was also determined by taking the mass defect approximation. The value obtained was 1298 cm^{-1} which shows that the approximation would be acceptable in this situation. On replacing ^{10}B by ^{11}B the dominant vibrational peak became that centred on 1211 cm^{-1}.

The structure of the boron–hydrogen complex which minimised the total energy of the cluster was aligned along essentially the $\langle 100 \rangle$ direction. The boron–hydrogen bond stretch mode has a frequency estimated to be between 2540 cm^{-1} and 2655 cm^{-1}, depending on the position of the hydrogen in a flat–bottomed minimum in the energy surface. The anharmonicity of this energy surface makes it impractical to determine any vibrational 'wagging' modes involving the hydrogen atom. In any case, this strong boron–hydrogen bond stretching mode gives the most characteristic frequency associated with the defect and should be experimentally observable in diamonds containing sufficient concentrations of boron and hydrogen. The results of a Green function analysis show that there are no other localised vibrational modes apart from that associated with the hydrogen atom.

Chapter 6

The Native Defects in Diamond

6.1 Introduction

In this chapter we investigate the structure and important properties of the most simple native defects in diamond, the vacancy and the self interstitial. There has been work published over the last two decades on native defects in diamond and associated properties, however, this is just a fraction of that published on the same area in silicon. There are still questions involving these intrinsic defects, especially that of the position which would be occupied by a self interstitial in the lattice. Our work is compared with previous results found using theoretical techniques of various levels of sophistication and experimental findings.

The presence of the self interstitial in diamond is often overlooked, although the irradiation of semiconductors causes extensive damage to the lattice and leads to the directly connected formation of both vacancies and self interstitial defects. This irradiation can be either as the standard technique of irradiating semiconductors by high energy electrons prior to annealing and the subsequent investigation of the material using infra–red or optical radiation or by the nature of the environment in which the diamond is placed. Examples of this latter case are the use of diamond films in detectors in particle physics experiments, or as radiation dosimeters in medical physics, as mentioned in Chapter 1. In Chapter 7 we investigate self diffusion in diamond and the determination of the stability of possible self interstitial structures is a necessary step in the estimation of the migration energy of a self interstitial.

6.1.1 The Atomic Bases Used

Unless otherwise stated the bases used on the atoms in the work presented in this chapter are as follows.

For the charge density :

The carbon atoms have 4 orbitals of different widths.
The terminating hydrogen atoms have 3 orbitals of different widths.

For the wavefunction :

The full basis carbon atoms have 16 orbitals (s,p_x,p_y and p_z type orbitals each of 4 different widths).
The contracted basis carbon atoms have 4 orbitals (Each a linear combination of a particular type of orbital (s,p_x,p_y and p_z) each of which has 4 different widths).
The terminating hydrogen atoms have 2 orbitals (A linear combination of the s–type orbitals of two different widths and a linear combination of the p–type orbitals (p_x,p_y and p_z) each of two different widths).

In some of the calculations a larger basis is used on certain carbon atoms in order to check the convergence of a result with basis size. This larger basis will be referred to as the 'C8_8' basis and has the following numbers of orbitals.

For the charge density :

The carbon atoms have 8 orbitals of different widths.

For the wavefunction :

The full basis carbon atoms have 32 orbitals (s,p_x,p_y and p_z type orbitals each of 8 different widths).

6.2 The Self Interstitial

6.2.1 Previous Investigations

Despite the fact that theoretical studies of the self interstitial in diamond have been carried out using increasingly sophisticated methods over the years, experimentally little has been deduced regarding the defect. There have been no infra–red or optical absorptions attributed to interstitial carbon atoms and no EPR or similar techniques have deduced any properties which could definitely be linked with these defects. Only in the area of annealing stages have proposals been made as to the identification of the interstitial as a diffusing species, as will be discussed in Chapter 7.

Basically, five possible self interstitial defect structures in tetrahedral semiconductors have been proposed and investigated by previous workers. These are the tetrahedral, hexagonal, bond–centred, $\langle 100 \rangle$ and $\langle 110 \rangle$ split interstitials, as shown in Figs. 6.1 to 6.4. Initial feelings towards which site an interstitial atom would occupy in the diamond lattice were that the tetrahedral or hexagonal sites would be favoured. This conclusion was based on the fact that these positions had the greatest separation from the surrounding atoms in the perfect lattice. Consequently the minimum amount of lattice distortion would be necessary to accommodate the interstitial atom. Theoretical

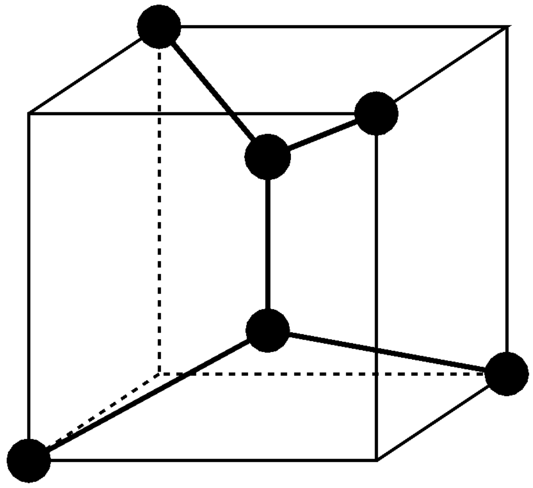


Figure 6.1: The structure of the $\langle 100 \rangle$ split interstitial defect. The bond between the 2 central atoms is in the crystallographic $\langle 100 \rangle$ direction.

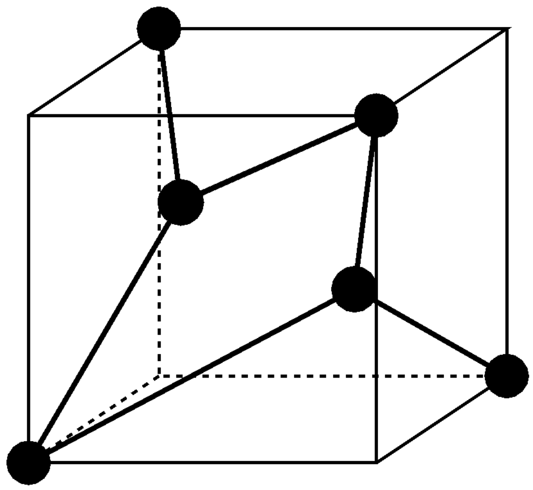


Figure 6.2: The structure of the $\langle 110 \rangle$ split interstitial defect. The 2 atoms within the cube are displaced from each other along the crystallographic $\langle 110 \rangle$ direction.

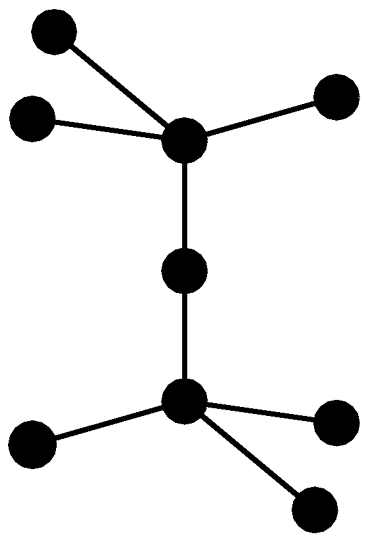


Figure 6.3: The structure of the bond–centred interstitial defect.

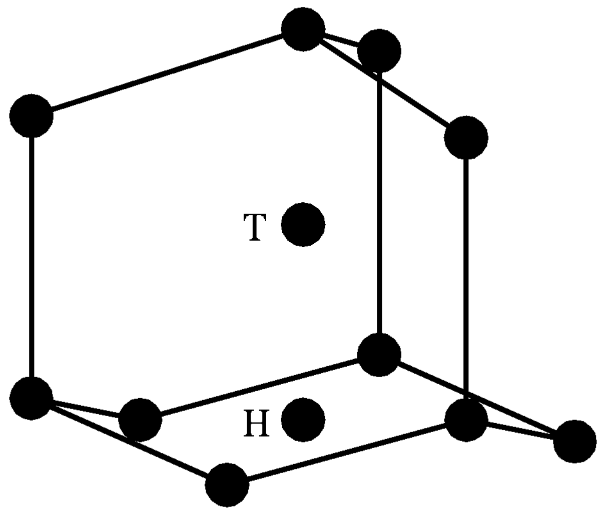


Figure 6.4: The 10 atom cage in a tetrahedral lattice with the tetrahedral interstitial site (T) at the centre and the hexagonal site (H) equidistant from the 6 atoms surrounding it.

calculations later pointed towards the possibility that other structures, such as the bond–centred and $\langle 100 \rangle$ and $\langle 110 \rangle$ split interstitials, could be lower in energy. We now outline the findings of some of the more significant theoretical contributions to this area over the last couple of decades.

The first investigations on diamond were carried out using semi–empirical methods with experimental parameter input. Weigel (1973) carried out calculations by an LCAO molecular orbitals technique using extended Hückel theory and clusters of up to 48 atoms, with the surface bonds left as unsaturated dangling bonds. Calculations were performed using both no relaxations of the surrounding lattice and very limited manual displacements of the atoms in the cluster. The order of the energies of the defects (lowest energy first) was found to be $\langle 100 \rangle$ split interstitial, bond–centred, $\langle 111 \rangle$ split interstitial, $\langle 110 \rangle$ split interstitial, hexagonal and tetrahedral. The $\langle 111 \rangle$ split interstitial is an improbable structure which was not considered by later work. The hexagonal and tetrahedral sites were found not only to be significantly higher in energy than the other sites (which were all within 3 eV of each other), but also not to be sites of potential energy minima. A charge state dependence was noted for the energy difference between the $\langle 100 \rangle$ split interstitial and bond–centred defects, with the bond–centred site becoming the most stable for the highest charge state considered (+2).

Mainwood (1978a) used a CNDO (Complete Neglect of Differential Overlap) method in clusters of up to 38 atoms with effective atoms saturating the surface bonds. Valence–bond forces were used to relate the first nearest neighbour motion to the more distant neighbours. It was reported that the $\langle 100 \rangle$ split interstitial had the lowest energy with the hexagonal interstitial 1.5 eV above and with the tetrahedral, bond–centred and $\langle 110 \rangle$ split interstitials all determined to be significantly higher in energy. The formation energy of the $\langle 100 \rangle$ split interstitial was calculated to be 2.8 eV.

More recent work has used first principles (*ab initio*) techniques which do not require experimental input and do not have the severe approximations used by the older methods. The first application of these methods was by Bernholc (1988) who modelled these defects using a 16 atom supercell with the wave function expanded in terms of plane waves with a cut–off energy of 14.2 Ry and 25.8 Ry for those added via perturbation theory. The formation energies of the interstitials were found to be as follows, bond–centred (15.8 eV), $\langle 100 \rangle$ split interstitial (16.7 eV), tetrahedral (23.6 eV). The supercell used only contained 16 atoms and so the very high defect concentration produced could have given rise to strong electronic defect–defect interactions producing changes specific to these calculations. Also, the relaxations of the neighbouring atoms in such a small supercell will probably be affected by the close proximity of the neighbouring defects (the first nearest neighbours were reported to move outwards by 15% and 20% for the $\langle 100 \rangle$ split interstitial and bond–centred defects respectively). The $\langle 110 \rangle$ split interstitial was not considered in this work.

A more recent *ab initio* study (Chadi 1992) applied the same approach as Bernholc (1988) to the self interstitial in silicon. A 32 atom supercell was used with a cut–off frequency of 7.0 Ry for the plane wave basis. It was reported that the $\langle 110 \rangle$ split interstitial was the most energetically favourable structure for the neutral, +1, –1 and +2 charge states, with the $\langle 100 \rangle$ split interstitial, bond–centred, tetrahedral and hexagonal structures also being studied. However, in the neutral and $\pm$ 1 charge states, all the structures had energies within 1 eV of each other.

Taking the previous work into account, we felt that a comprehensive *ab initio* study of all the proposed self–interstitial sites in diamond using large clusters would be a worthwhile addition to this area of research.

6.2.2 Calculational Details

The investigation of each of the possible self interstitial structures was carried out in the -1, 0, $+1$ and $+2$ charge states. All the interstitial structures were investigated using a 163 atom cluster ($C_{93}H_{70}$) (Fig. 6.5). This cluster

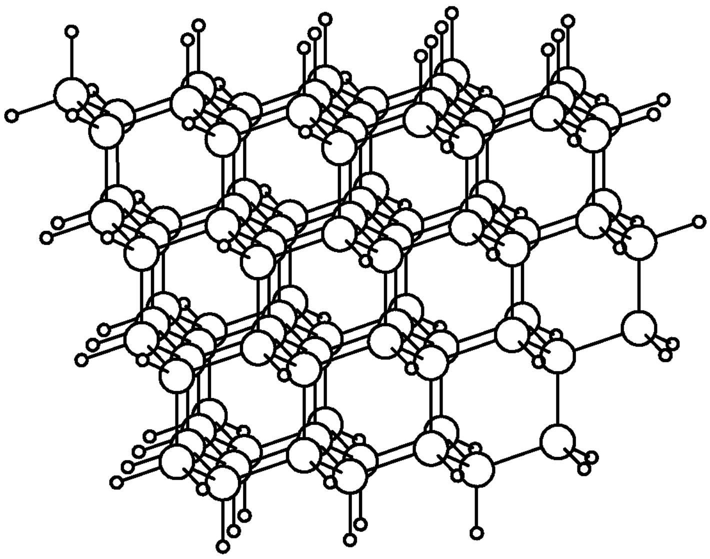


Figure 6.5: The 163 atom cluster ($C_{93}H_{70}$) used in the calculations. The hydrogen atoms are those with the smaller radii at the surface of the cluster.

has C_{3v} symmetry and contains a substitutional, tetrahedral, hexagonal and bond–centred interstitial site which are all distant from the surface of the cluster, making it suitable for the comparison of the energies of the different defect structures considered. The minimum distance from the defect sites to the hydrogen atoms at the surface is fourth nearest neighbour distance and the sites for each defect are within two atomic spacings of the centre of the cluster. With the addition of another carbon atom to form the interstitial defects, the innermost 30 carbon atoms were in full basis and these same atoms were allowed to move in the optimisation of the structure.

6.2.3 The $\langle 100 \rangle$ Split Interstitial

An atom was added to the 163 atom cluster to form a 164 atom cluster with a split interstitial centred on the substitutional site and aligned along the $\langle 100 \rangle$ direction (Fig. 6.1) for each charge state. The atoms were equivalently displaced from the substitutional site with a distance of 1.3 Å between them. The structure was allowed to relax and it was found that the $\langle 100 \rangle$ split interstitial was stable in all charge states. The two central atoms in the defect were also initially made inequivalent by displacing the atoms by different amounts along the $\langle 100 \rangle$ direction from the substitutional site. The magnitude of the difference was 0.2 Å. The final relaxed structure always had the two atoms equivalently displaced from the substitutional site. The bond lengths for the atoms forming and surrounding the relaxed defect structure were quite insensitive to the charge state. The structure of this defect will be investigated more fully in section 6.2.7 along with a detailed study of the possible Jahn–Teller distortions of the defect. The bond lengths determined for the relaxed defect in the 163 atom cluster agree with those determined in section 6.2.7 (Table 6.2), for the defect studied in a 131 atom cluster, to 0.02 Å.

The energy of the $\langle 100 \rangle$ split interstitial will be compared to that of the other possible interstitial structures studied in the following sections. An important point to note is that the total energies of other interstitial configurations relative to the $\langle 100 \rangle$ split interstitial (Table 6.1) are relative to the lowest possible energy structure of the defect in a particular charge state, i.e. a Jahn–Teller distortion is included if it lowers the total energy. It will be found in section 6.2.7 that this is indeed the case for the neutral, +1 and –1 charge states of the $\langle 100 \rangle$ split interstitial.

To check the effect of basis size, the relaxation was repeated for the neutral defect with a C8_8 basis on the two atoms forming the defect and their four nearest neighbours. The bond lengths changed by a maximum of 0.007 Å.

6.2.4 The Tetrahedral and Hexagonal Interstitials

An atom was added to the 163 atom cluster at the appropriate site to form a tetrahedral and an hexagonal interstitial defect (Fig. 6.4) for each charge state. Initially, the structure was allowed to relax with the interstitial atom constrained to remain at the appropriate symmetry site. This gave the best chance for the surrounding atomic network to relax to accommodate and thus to stabilise the interstitial. After this initial relaxation, this constraint was removed so that the interstitial atom was also allowed to move. On further relaxation of the structure, the interstitial atom migrated to one of the surrounding 'cage' atoms of Fig. 6.4, where it formed a relaxed $\langle 100 \rangle$ split interstitial defect. This occured in all cases. In the neutral charge state, the total energy of this relaxed $\langle 100 \rangle$ split interstitial which was formed with one of the nearest neighbours of the substitutional site (as used in sections 6.2.3 and 6.2.6) differs from the total energy of the $\langle 100 \rangle$ split interstitial centred on the substitutional site by only 0.002 eV.

An argument can be constructed to explain the instability of the tetrahedral interstitial defect on the basis of possible symmetry breaking distortions. The multiplet structure of the neutral ideal tetrahedral interstitial with the surrounding atoms at the perfect diamond lattice positions is somewhat similar to that of the neutral vacancy, which will be considered in section 6.3.3. Three degenerate one–electron levels of symmetry t_2 appear in the band gap into which two electrons must be placed. This configuration is known to be unstable with respect to a Jahn–Teller distortion for the neutral vacancy but, in that case, the relaxation is limited by the constraints of the surrounding lattice. That is not the case here, and a much larger relaxation can occur, possibly the reason for the migration of the interstitial atom to form a $\langle 100 \rangle$ split interstitial defect. In a similar way, the negatively charged interstitial can be compared to the negatively charged vacancy. For this defect, the T_d

symmetry ${}^4A_2$ state is the ground state, as will be discussed in section 6.3.4. However, a distorted excited state with spin $S = \frac{1}{2}$ of the negative vacancy lies only 0.3 eV higher in energy than the ${}^4A_2$ state with the distortion again being limited by the lattice. In the case of the negatively charged interstitial atom at the tetrahedral site, the lack of resistance to distortion could again lead to the off–site motion of the interstitial atom.

In any case, what has been established is that the tetrahedral and hexagonal interstitial sites are not at positions of minima in the potential energy surface. The energy of the relaxed tetrahedral interstitial structure, with the interstitial atom forced to remain on site, is greater than that of the $\langle 100 \rangle$ split interstitial by 7–10 eV, depending on the charge state. Thus, the energy gain in the migration of this defect to form a $\langle 100 \rangle$ split interstitial is considerable.

6.2.5 The Bond–Centred Interstitial

The interstitial atom was located at the centre of a carbon–carbon bond (Fig. 6.3) along the C_{3v} symmetry axis of the 163 atom cluster for each charge state. The structure was allowed to relax and in each case the interstitial atom remained on the bond axis. In each charge state the structure obtained had the bonds from the bond–centred atom to its two nearest neighbours at about 80% of the perfect diamond lattice bond length. The 2nd and 3rd nearest neighbours were displaced outwards in such a way that their bonds only differed in length from those in bulk diamond by at most 3%.

As discussed in section 3.8.2, the fact that the interstitial atom remained on the bond axis could have been due to the forces on the atoms being determined exactly using an analytical formula. There is no possibility of a component of the force acting on the atom in a direction perpendicular to the bond. To determine whether the apparent stability of the defect was an artifact of the calculation or not, the interstitial was displaced perpendicular to the bond

Charge State	–1	0	+1	+2
Energy of bond–centred interstitial (eV)	+4.45	+4.05	+3.04	+2.32
Energy of $\langle 110 \rangle$ split interstitial (eV)	—	+2.53	—	+1.85

Table 6.1: The energy of the bond–centred interstitial and $\langle 110 \rangle$ split interstitial defects relative to the $\langle 100 \rangle$ split interstitial defect for various charge states.

axis in each case by 0.05 Å. The result of further relaxation was the same for all charge states. The interstitial atom moved to form a $\langle 100 \rangle$ split interstitial defect with one of the two neighbouring atoms in the original bond–centred defect.

In addition to the demonstrated instability of this site, for all the charge states the energy of the relaxed bond–centred interstitial was found to be higher than that of the correspondingly charged $\langle 100 \rangle$ split interstitial (Table 6.1). The increase of relative energy with the addition of electrons is consistent with one electron energy levels high in the band gap, as we found for this defect.

6.2.6 The $\langle 110 \rangle$ Split Interstitial

An atom was added to the 163 atom cluster to form a 164 atom cluster with a split interstitial centred on the substitutional site and aligned along the $\langle 110 \rangle$ direction (Fig. 6.2) for each charge state. The atoms were equivalently displaced from the substitutional site with a distance of 1.3 Å between them. When the structure was relaxed from this configuration it was found that in the neutral and +2 charge states the structure remained a split interstitial aligned along the $\langle 110 \rangle$ direction. However, for both the +1 and –1 charge states the defect reorientated itself into a $\langle 100 \rangle$ split interstitial centred on the same site.

In order to test the stability of the defect in the neutral and +2 charge states the 2 atoms forming the split interstitial were displaced from their equilibrium positions. Both the atoms in the neutral defect were displaced, firstly by 0.005 Å and then by 0.025 Å. The direction of displacement was $[0\bar{1}0]$ for the atom located along $[\bar{1}10]$ from the substitutional site and [010] for the atom located along $[1\bar{1}0]$. The reason for the particular displacements is that they are in the direction most likely to precipitate a reorientation of the defect into a $\langle 100 \rangle$ split interstitial. This does indeed occur for the 0.025 Å displacement. The atoms relax back to the $\langle 110 \rangle$ split interstitial structure for the 0.005 Å displacement. Displacements in the same directions were carried out for the +2 defect with magnitudes of 0.05 Å, 0.025 Å and 0.005 Å for each atom. In each case the atoms reorientated themselves into a $\langle 100 \rangle$ split interstitial.

These displacements are considerably smaller than those associated with the zero point motion of the defect, which is neglected in our calculation, and we can conclude that $\langle 110 \rangle$ split interstitials are not stable in diamond. In any case, the energies of the relaxed neutral and +2 $\langle 110 \rangle$ split interstitials were higher than those of the correspondingly charged $\langle 100 \rangle$ split interstitials (Table 6.1).

6.2.7 A Further Study of the $\langle 100 \rangle$ Split Interstitial

As the only stable structure for the self interstitial in diamond has been found to be that of the $\langle 100 \rangle$ split interstitial, from the results of sections 6.2.3 to 6.2.6, this defect is now investigated in further detail. In this study the defect was modelled in a 131 atom cluster ($C_{71}H_{60}$) (as was used in Fig. 5.4, but of pure diamond). This cluster was chosen in preference to the 163 atom cluster used in the previous sections as only the substitutional site is required to model this defect. The criterion of the distance of the defect from the surface of the cluster being large is still satisfied, being fourth nearest neighbour distance.

Also, the cluster has tetrahedral symmetry (i.e. that of the lattice) about a central substitutional site, so that there is no possibility of minor perturbations of the defect structure due to the surrounding cluster initially having a symmetry lower than that of the bulk diamond lattice. With the addition of the interstitial atom, the central 18 atoms were in full basis and all the carbon atoms were allowed to move.

The central carbon atom in the 131 atom cluster was replaced by two carbon atoms to form a split interstitial centred on the substitutional site and aligned along the $\langle 100 \rangle$ direction (Fig. 6.6), with a distance of 1.3 Å between the atoms. This starting structure has D_{2d} symmetry. Two one–electron en-

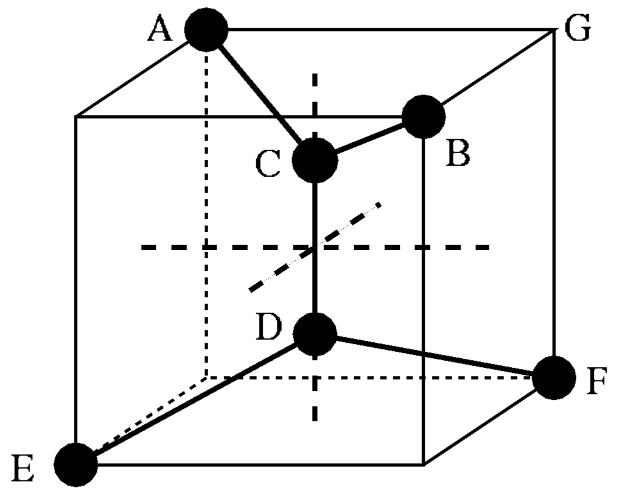


Figure 6.6: The $\langle 100 \rangle$ split interstitial defect, with labelling of atoms as used in the text. The dashed lines show the direction of the 3 perpendicular 2–fold rotation axes necessary for D_2 and D_{2d} symmetry.

ergy levels of symmetry e appeared towards the top of the band–gap. The wave functions of these levels are basically $p-$orbitals centred on the two carbon atoms forming the interstitial pair (atoms C and D in Fig. 6.6). These orbitals point in the direction perpendicular to the plane containing the three neighbours of each atom. In general, one would expect bonding and antibonding combinations of these to be formed leading to a splitting of the two states, as occurs in the H_2 molecule. This does not happen here as the two defect states are orthogonal and so do not mix.

In the neutral charge state two electrons are placed into these two degenerate orbitals. The multiplet structure of this defect is therefore found from $e \otimes e$ evaluated in D_{2d} symmetry. This gives rise to ${}^1A_1$, ${}^3A_2$, ${}^1B_1$ and ${}^1B_2$. If we label the two one electron states as ϕ_a and ϕ_b and the spin states as $\uparrow$ and $\downarrow$, the multiplets are :

$$\begin{aligned}
{}^1A_1 &: [\phi_a(1)\phi_a(2) + \phi_b(1)\phi_b(2)]\,[\uparrow\downarrow - \downarrow\uparrow] \\
{}^1B_2 &: [\phi_a(1)\phi_a(2) - \phi_b(1)\phi_b(2)]\,[\uparrow\downarrow + \downarrow\uparrow] \\
{}^1B_1 &: [\phi_a(1)\phi_b(2) + \phi_b(1)\phi_a(2)]\,[\uparrow\downarrow - \downarrow\uparrow] \\
{}^3A_2 &: [\phi_a(1)\phi_b(2) - \phi_b(1)\phi_a(2)] \left\{ \begin{array}{c} [\uparrow\uparrow] \\ [\uparrow\downarrow + \downarrow\uparrow] \\ [\downarrow\downarrow] \end{array} \right\}
\end{aligned}$$

Density functional theory was originally presented as a method for determining ground state properties. However, a detailed study of methods by which multiplet splittings can be obtained from density functional theory has been carried out (von Barth 1979). There it is shown that if a one electron configuration contributes to only a single multiplet state, then the local density functional approximation to the energy of that configuration can be a reasonable estimate of the energy of the multiplet. The multiplet calculations which follow use the local spin density approximation, implemented as described in section 3.11.

The above criterion is clearly satisfied, for example, for the one electron configuration in which electrons are placed with parallel spins into each of the two states, i.e. $|a\uparrow\rangle|b\uparrow\rangle$. The energy of this is an estimate of that of the

$^3A_2$ state. We first used this configuration in the 132 atom cluster described above and relaxed the central atoms to get the atomic structure and energy of the $S = 1$ state, which has D_{2d} symmetry. With this atomic structure, the energies of the different multiplets can then be found by following von Barth. The atoms are fixed in the positions obtained for the $^3A_2$ state and the electronic level filling is constrained to be of the required form for a multiplet of interest.

The energy of the mixed state $|a \uparrow\rangle|b \downarrow\rangle$ is the average of the $^1B_1$ and $^3A_2$ multiplets so that the energy of $^1B_1$ can also be determined. Similarly the energy of $|a \uparrow\rangle|a \downarrow\rangle$ is the average of the energies of $^1A_1$ and $^1B_2$. It is not possible to find the energies of these separately, as we do not have any additional distinct one–electron configurations (i.e. we have more multiplets than distinct configurations). As a result, we have an underdetermined set of equations for the evaluation of the multiplet energies. This is the opposite problem to that encountered by von Barth who had an overdetermined set of equations. The result of these calculations is that we find the energy of $^1B_1$ to be 0.55 eV lower than that of $^3A_2$, whereas the average of the energies of $^1A_1$ and $^1B_2$ lies 1 eV higher than $^3A_2$. Therefore, our calculations assign $^1B_1$ to be the ground state of the D_{2d} symmetry $\langle 100\rangle$ split interstitial. This suggests that Hund's rule fails to get the correct ground state in this case.

We now consider the possible Jahn–Teller distortion of this defect. Such a distortion is possible in the neutral or singly charged states in which the one–electron e states are partially occupied. In the doubly charged states, +2 or –2, the two one–electron states are either completely empty or completely occupied so that a Jahn–Teller distortion is not to be expected.

To investigate this the 2 atoms of the $\langle 100\rangle$ split interstitial were made inequivalent by displacing the atoms in the relaxed structures by up to 0.2 Å. Further relaxation was then allowed. In only one of the charge states, the +2, did the defect recover D_{2d} symmetry. In the neutral, +1 and –1 charge states the final structure had D_2 symmetry. D_2 symmetry requires 3 perpendicular 2–fold rotation axes. These are shown in Fig. 6.6 and basically connect the centres of the opposite faces of the cube. D_{2d} symmetry also requires 2 vertical mirror planes. These are shown in Fig. 6.7 and cut the top and bottom faces of the cube in Fig. 6.6 diagonally. The structural relaxation therefore resulted

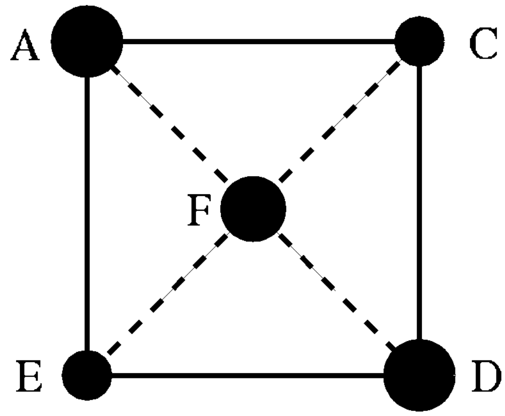


Figure 6.7: The top face of the cube in Fig. 6.6 containing the ⟨100⟩ split interstitial defect. The dashed lines show where the vertical mirror planes perpendicular to the top face cut that face.

in the loss of the 2 vertical mirror planes. The reason for this is that the two previously equivalent sides of the cube, AG and GB, in Fig. 6.6, become inequivalent. In all cases the sides of the cube remain perpendicular to each other so that the cube strictly becomes a rectangular parallelepiped. For the relaxed neutral defect with D_2 symmetry the sides had lengths AG=1.73 Å, GB=1.98 Å and GF=2.19 Å, with these distances varying by a maximum of 0.08 Å for the +1 and –1 charge states with the same symmetry. These are to be compared with AG=GB=1.87Å and GF=2.19Å for the neutral structure with D_{2d} symmetry, with a maximum variation of 0.07Å for the other charge states, and AG=GB=GF=1.77Å for the perfect lattice. The bond lengths obtained for the various charge states and symmetries of the defect are given in Table 6.2. It is interesting to note that the bond lengths are insensitive to

Charge State	neutral	neutral	-1	$+1$	$+2$
Symmetry	D_{2d}	D_2	D_2	D_2	D_{2d}
C–D	1.27 Å	1.25 Å	1.27 Å	1.26 Å	1.26 Å
A–C	1.40 Å	1.40 Å	1.38 Å	1.40 Å	1.40 Å
B–C	1.40 Å	1.40 Å	1.38 Å	1.40 Å	1.40 Å
D–E	1.40 Å	1.40 Å	1.38 Å	1.40 Å	1.40 Å
D–F	1.40 Å	1.40 Å	1.38 Å	1.40 Å	1.40 Å

Table 6.2: The bond lengths corresponding to the relaxed $\langle 100 \rangle$ split interstitial structures in all the charge states and symmetries studied. The bond labelling refers to that used for the atoms in Fig. 6.6.

the charge state, varying by at most 0.02 Å. Also, the difference between the bond lengths for the neutral defect with D_{2d} symmetry and with the Jahn–Teller distorted D_2 symmetry is only 0.02 Å. There is a corresponding loss of electronic level degeneracy with the symmetry breaking structural relaxation, with the one–electron e state splitting into b_2 and b_3 levels, with the lowest being doubly occupied for the neutral charge state. The same level splitting occurs for the $+1$ and -1 charge states. The difference in energy between the defect structures with D_{2d} and D_2 symmetry is the Jahn–Teller energy for the system. This was calculated to be 0.6 eV for the neutral defect.

Finally, we determined the vibrational properties of the neutral defect with D_2 symmetry. The second derivatives of energy with respect to displacements of the two atoms forming the interstitial together with their four nearest neighbours were found (i.e. all of the atoms in Fig. 6.6), giving an 18×18 matrix of interatomic force constants. From this the dynamical matrix was constructed and diagonalised, as described in section 4.3, and the vibrational frequencies found. The highest frequency vibrational mode, which is associated with the stretching of the central carbon–carbon bond, is an A mode with frequency 1984 cm^{-1}. The next two highest frequency modes are B modes associated with the motion of atoms C and D in Fig. 6.6 perpendicular to the $\langle 100 \rangle$ bond, but within the plane of their three neighbours. These had frequencies of 1560 and 1595 cm^{-1}. It was deduced that the A mode is infra–red inactive, but that the B modes are both infra–red active.

6.3 The Vacancy

6.3.1 Previous Investigations

The fact that far more experimental work has been carried out on the vacancy than on the interstitial in diamond is due to readily identified absorption bands for the two main charge states of the defect. It is accepted that the GR1 absorption band is attributable to transitions at the vacancy in the neutral charge state and that the source of the ND1 band is transitions at the vacancy in its negative charge state. Due to the large amount of both experimental and theoretical published work on the vacancy in diamond, in this section we shall refer to only some of the more significant contributions over the years.

Experimental investigations by Clark (1973) into the colour centres of diamond using optical absorption and uniaxial stress provided significant evidence in support of the conclusion for the neutral charge state. It had already been believed for many years before Clark's work that the colour induced in diamond following irradiation was due to vacancies or interstitials produced by that procedure. This was essentially a common sense deduction, as high energy radiation should eject atoms from lattice sites leaving a vacancy and an interstitial atom. Theoretical modelling of an unrelaxed vacancy had been carried out using a modified molecular orbital technique and group theory considerations (Coulson 1957) which showed that transitions at the neutral vacancy would be consistent with the GR1 band. Theoretical work was later carried out by Bachelet (1981) using a self–consistent Green function technique to analyse the unrelaxed neutral vacancy. There has been much material published on optically excited transitions at the neutral vacancy over the years so that experimentally the defect is well documented, for example Walker (1974) and Davies (1974,1981).

Experimental work by Davies (1977) led to the suggestion that the ND1 band was associated with an A to T transition at the negative vacancy, rather than the previously suggested substitutional nitrogen atom. Isoya (1992) used EPR (Electron Paramagnetic Resonance) and ENDOR (Electron Nuclear Double Resonance) in order to identify the negative vacancy and determine that the defect has an effective spin of $\frac{3}{2}$. This supports the $^4A_2$ ground state proposed for the defect and so leads to the conclusion that many body effects are important in diamond, as a Jahn–Teller distorted structure is not the ground state for this charge state of the vacancy as it is in silicon.

A more complete theoretical analysis of both the neutral and negative vacancies was undertaken by Mainwood (1978b) using a CNDO method. The relaxation of atoms surrounding the vacancy was taken into account, and the energy of the Jahn–Teller distortion for the neutral vacancy was predicted to be 0.46 eV. An experimental estimation of the Jahn–Teller energy was made by Davies (1982), the result being 0.2 eV.

6.3.2 Details of the Calculation

The 131 atom cluster used in section 6.2.7 was again used for the vacancy. A carbon atom was removed from the central substitutional site of the cluster leaving an unrelaxed vacancy with tetrahedral symmetry. The sides of the cube in Fig. 6.8 had lengths AE=BE=CE=1.77 Å for this structure. The central

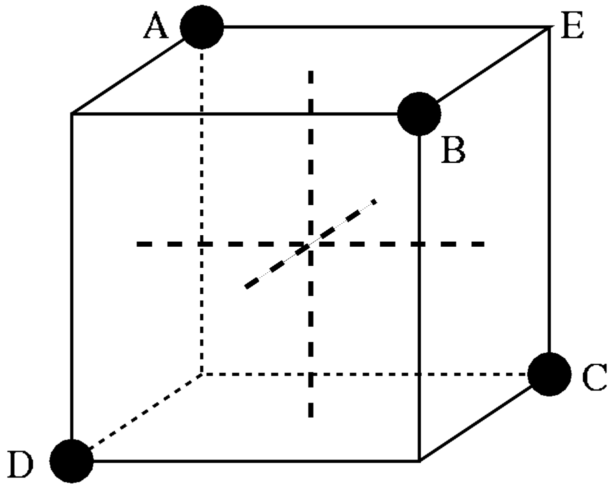


Figure 6.8: The vacancy with labelling of atoms used in the text. The dashed lines show the direction of the 3 perpendicular 2–fold rotation axes necessary for D_2 and D_{2d} symmetry.

16 carbon atoms were treated in full basis, and all 70 carbon atoms were allowed to move during structural optimisation, unless otherwise specified. The symmetry dictates that the four carbon atoms surrounding the vacancy are initially equivalent. This ideal, unrelaxed vacancy has four dangling sp^3 hybrids pointing inwards, and in tetrahedral symmetry these give rise to a one electron state of a_1 symmetry which falls just above the top of the valence band and three degenerate states of symmetry t_2 which occur deep in the band gap.

6.3.3 The Neutral Vacancy

In the neutral charge state, four electrons must be placed into the one–electron states. In the ground state two are placed into the a_1 state and two into the t_2 one–electron states, giving the configuration $a_1^2 t_2^2$.

Initially, the three t_2 states were occupied equally with an effective occupancy of $\frac{2}{3}$ and the structure was allowed to relax. With this rather artificial filling, tetrahedral symmetry was maintained and the four carbon atoms immediately surrounding the vacancy moved radially outwards by 0.2 Å. The sides of the cube in Fig. 6.8 had lengths AE=BE=CE=1.89 Å after relaxation. The bonds from the first nearest neighbours of the vacancy to the second nearest neighbours and from the second to the third were shorter than those in the bulk by no more than 2 %. This motion corresponds to an A mode relaxation. The relaxation is in the opposite direction to the results obtained by Mainwood (1978b) using a CNDO technique, who found that the atoms surrounding the vacancy relaxed inwards, but that a very weak metastable structure with outwards relaxation could be possible with a specific choice of basis. Also, two molecular orbital extended Hückel theory calculations predicted the opposite results that inwards (Astier 1978) and outwards (Messmer 1973) relaxations of the atoms surrounding the vacancy produced the minimum energy structure. To check our results thoroughly, the four atoms surrounding the vacancy were fixed at various radial displacements from their positions in the perfect lattice. The remaining 66 carbon atoms were all allowed to relax in order to determine the total energy of the cluster under this constraint. This was carried out for inwards displacements of up to 0.5 Å and outwards displacements of up to 0.6 Å. The energy plotted as a function of the fixed displacements of the four neighbours of the vacancy is shown in Fig. 6.9, and shows only one energy minimum, that centred on an outwards displacement of 0.2 Å.

The outwards relaxation is understandable from a chemical bonding point of view. It leads to the atoms surrounding the vacancy moving more into the plane of their three nearest neighbours, and the bonding becoming more like sp^2 than sp^3, with the former being preferred by carbon atoms.

Clearly, the one electron configuration t_2^2 is unstable with respect to a Jahn–Teller distortion. In order to investigate the structural lowering of symmetry,

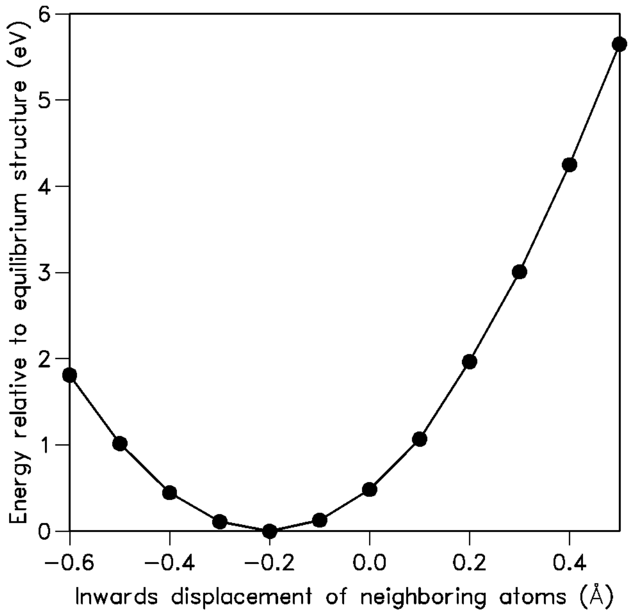


Figure 6.9: The energy of a relaxed neutral vacancy in eV as a function of the fixed inward radial displacement of the four surrounding atoms relative to their position in the perfect lattice. The energy is relative to the equilibrium vacancy structure with an outwards radial displacement of 0.2 Å.

displacements of 0.1 Å were applied to two of the atoms surrounding the vacancy to break the tetrahedral symmetry. The structure was then allowed to relax with no constraints placed on the filling of the electronic orbitals. It was found that one of the cubic axes in Fig. 6.8 became inequivalent to the other two. The symmetry changed from T_d to D_{2d}, corresponding to an E mode relaxation. With reference to Fig. 6.8, the sides of the cube for the relaxed neutral defect with D_{2d} symmetry remain perpendicular to one another and have lengths AE=BE=1.78 Å and CE=2.00 Å. The effect of this distortion is to split the t_2 one electron states into a doubly occupied a state and an unoccupied e state. The Jahn–Teller energy associated with the relaxation of the neutral vacancy was estimated from the difference between the total energies of the relaxed T_d and D_{2d} symmetry defects. It was calculated to be 0.36 eV, somewhat larger than an approximate experimental value of 0.2 eV (Davies 1982), but smaller than previous theoretical estimates of 0.46 eV (Mainwood 1978b) and 0.6 eV (Stoneham 1977).

In fact, the distortion is known to be dynamic for optical experiments carried out at room temperature, so that the optical absorption is analysed in terms of multiplets of a defect with effective T_d symmetry. The multiplet structure of this effectively tetrahedral centre has been most effectively treated using empirical methods following Coulson (1957) and work on related models is continuing (Lowther 1993). It is established that the ground state of the tetrahedrally symmetric vacancy is ^1E with a vibronic state only just higher in energy by a few meV (Lannoo 1968). The best known signature of this defect is the GR1 band with a zero phonon line at 1.673 eV. This is believed to be associated with a transition from the ground ^1E state to $^1T_2$. Unfortunately, it is not possible to determine the energy of the ground state ^1E using density functional theory as (following the discussion in section 6.2.7) only a weighted value of the energies of the $^1A_1$ and ^1E states can be found. Therefore, no multiplet calculations were carried out on the neutral vacancy.

6.3.4 The Negative Vacancy

For the negative vacancy, five electrons must be placed into the one–electron states. The lowest energy states of this defect with tetrahedral symmetry will correspond to the one–electron configuration $a_1^2 t_2^3$.

It is known experimentally (Isoya 1992) that the ground state of the negative vacancy is the many body state $^4A_2$ to which only a single one–electron configuration contributes, the one in which electrons are placed into each of the degenerate t_2 states with parallel spins. A spin polarised calculation was therefore performed with the electronic level filling constrained to be of that form. For this electronic configuration, each of the four atoms immediately surrounding the vacancy moves radially outwards by 0.2 Å on relaxation. The sides of the cube in Fig. 6.8 then have lengths AE=BE=CE=1.89 Å for the relaxed tetrahedral negative vacancy. This is almost identical to the case of the neutral vacancy in which each of the t_2 states is equally occupied.

As discussed in 6.2.7 it is possible to estimate the energies of certain multiplets using density functional theory. With this in mind further spin polarised calculations were carried out with the positions of the atoms fixed at those for the tetrahedral negative vacancy in the $^4A_2$ state, as determined above. The filling of the electronic levels was constrained to be that of a multiplet of interest and the energy determined. The differences in energy should then be analogous to electronic transitions measured in optical experiments as the electronic transition would occur and then any structural relaxations would take place over a significantly longer relative time scale. This is effectively in accordance with the Franck–Condon principle.

It is accepted that the ND1 optical band seen at 3.149 eV is associated with a transition from the $^4A_2$ state to the $^4T_1$ state. In the non–interacting limit, only a single configuration contributes to $^4T_1$, namely $a_1t_2^4$, in which the electrons are placed in the orbitals as $\xi_\uparrow\xi_\downarrow\eta_\uparrow\zeta_\uparrow$, where ξ, η and ζ label the 3 one electron states of symmetry t_2. The energy of this multiplet can be evaluated using density functional theory and the difference in energy between the two multiplets, and hence the energy of the ND1 band, was calculated to be 3.3 eV.

The energies of two other multiplets were determined, both with the single particle configuration $a_2t_2^3$. The electrons were placed in the orbitals $\xi_\uparrow\eta_\uparrow\zeta_\downarrow$ which corresponds to the state ^2E and, also, the electrons were placed in the orbitals $\xi_\uparrow\xi_\downarrow\eta_\uparrow$ which gives an energy equal to the average of the states $^2T_1$ and $^2T_2$. Hence we found the ^2E state to be 0.3 eV above the $^4A_2$ ground state and the average of the $^2T_1$ and $^2T_2$ multiplets to be at 1.6 eV above $^4A_2$. This placing of ^2E is considerably lower than in empirical methods (Lowther 1994) in which it is placed between 1.4 eV and 1.9 eV above ${}^4A_2$, depending on the choice of parameters. Our placing of the average of ${}^2T_1$ and ${}^2T_2$ multiplets is in reasonable agreement with the empirical calculations. As optical transitions between ${}^4A_2$ and these states are forbidden, experiment has not yet established their ordering or positioning.

Finally in this section, we consider possible Jahn–Teller distortions of this defect. The structural symmetry of the vacancy was firstly constrained to be that of a possible symmetry breaking distortion of the tetrahedrally symmetric defect (C_{2v}, C_{3v}, D_2 and D_{2d}). Structural relaxations using spin polarisation were then carried out with the appropriate electronic occupancies for each case. The three electrons originally placed into the three degenerate t_2 one–electron levels in the T_d symmetry vacancy are placed into the one–electron levels which arise in each structure from the splitting of the t_2 levels. This procedure is carried out for all plausible non–equivalent cases. For three of the atomic symmetries it was necessary to consider only one possible electronic configuration. These were $e^2(\uparrow)a_1(\downarrow)$ for the C_{3v} symmetry structure, $b_1(\uparrow)b_2(\uparrow)b_3(\downarrow)$ for the D_2 symmetry structure and $e^2(\uparrow)b_2(\downarrow)$ for the D_{2d} symmetry structure. For the C_{2v} symmetry structure it was necessary to consider five possible electronic configurations; $b_1(\uparrow)a_1(\uparrow)b_1(\downarrow)$, $b_1(\uparrow)a_1(\uparrow)b_2(\downarrow)$, $b_1(\uparrow)a_1(\uparrow)a_1(\downarrow)$, $b_1(\uparrow)b_2(\uparrow)a_1(\downarrow)$ and $b_1(\uparrow)b_2(\uparrow)b_1(\downarrow)$. The configuration $b_1(\uparrow)a_1(\uparrow)b_2(\downarrow)$ gave the relaxed C_{2v} symmetry structure with the lowest energy. Using this total energy for the C_{2v} case, we found that the energies of the relaxed distorted structures above the ground state T_d symmetry structure were 0.24 eV for the C_{3v} case, 0.4 eV for the C_{2v} case, 0.75 eV for the D_2 case and 0.89 eV for the D_{2d} case. Therefore, the $S = \frac{1}{2}$ Jahn–Teller distorted state has a higher energy than the ${}^4A_2$ state. This conclusion is supported by experiment in which it is known that the ground state of the negative vacancy has $S = \frac{3}{2}$ (Isoya 1992). Thus it is seen that in this case Hund's rule does in fact predict the correct ground state.

It is interesting to note that the energy splitting between the ground $S = \frac{3}{2}$ and excited $S = \frac{1}{2}$ states of the defect will in fact be given rigorously by density functional theory. This is because both of these states are the ground–electronic states of their respective structures (as defined by the positions of the atomic nuclei). This situation is illustrated by Fig. 6.10.

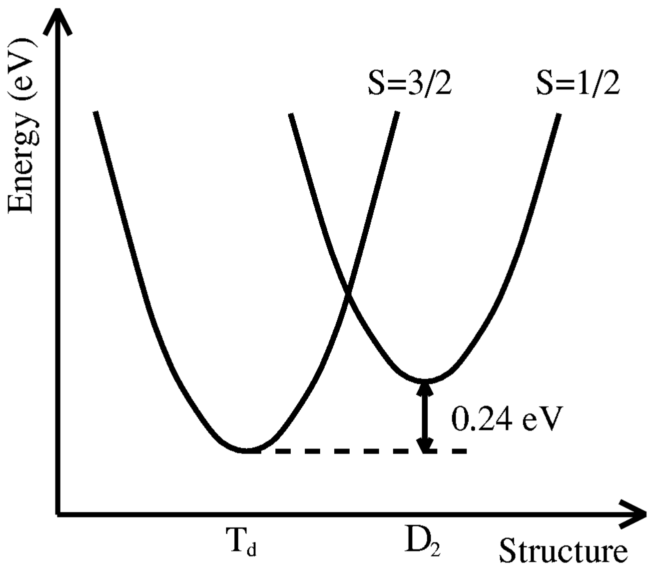


Figure 6.10: The ground state (T_d) and Jahn–Teller distorted (D_2) excited state of the negative vacancy.

The fact that the undistorted $^4A_2$ state has a lower energy than the Jahn–Teller distorted state again marks diamond out as different from silicon in that many–body correlation effects dominate the energy gain from the Jahn–Teller distortion in determining the ground state. This provides a more demanding challenge to theory.

6.4 The Formation Energy of a Vacancy–$\langle 100 \rangle$ Split Interstitial Pair

6.4.1 Previous Investigations

Finally, in considering the static properties of native defects, we estimate the energy required to form a vacancy–$\langle 100 \rangle$ split interstitial pair. Early theoretical approaches to the determination of the energy required to displace an atom from a lattice site in tetrahedral semiconductors, forming a vacancy–interstitial pair, were described and used by Kohn (1954) and Baüerlein (1959, 1962, 1963) yielding values in diamond of 24–30 eV. Experimental attempts at deducing a value for this process have produced such diverse values as 80 eV (Clark 1961) and 35 $\pm$ 5 eV (Bourgoin 1976). Also, in a more recent theoretical calculation, Bernholc (1988) deduced the separate formation energies of the vacancy and various interstitial configurations.

6.4.2 Details of the Calculation

Two approaches were used in order to estimate the pair formation energy.

In the first a 304 atom cluster ($C_{190}H_{114}$) (Fig. 6.11) with C_{3v} symmetry was used. This cluster contains sites for the vacancy and $\langle 100 \rangle$ split interstitial defects at fourth neighbour distance from the surface of the cluster and also with the defects at sixth nearest neighbour distance from each other.

The total energy of the perfect 304 atom cluster was determined with all the carbon atoms in contracted basis. The positions of the 16 atoms surrounding the vacancy and the 18 atoms surrounding and including the $\langle 100 \rangle$ split interstitial from previously relaxed clusters were then inserted into the cluster at the appropriate sites to create a relaxed vacancy–$\langle 100 \rangle$ split interstitial pair. The total energy of this cluster was then determined. This procedure

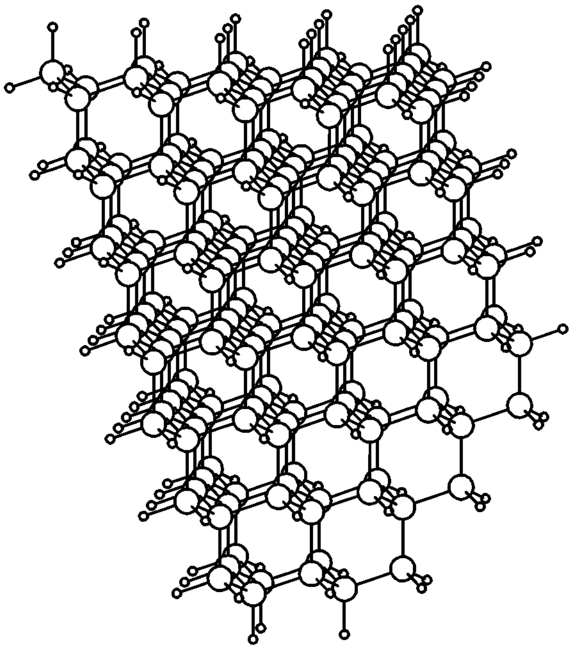


Figure 6.11: The cluster $C_{190}H_{114}$ used in the calculation of the energy required to form a vacancy–$\langle 100 \rangle$ split interstitial pair. The hydrogen atoms are those with the smaller radii at the surface of the cluster.

was repeated with 37 atoms in full basis. These atoms comprised those forming the defects and directly surrounding them. The difference between the total energies for the perfect and defect containing clusters for the contracted basis calculation was 25.2 eV. The difference between the total energies for the calculation with 37 atoms in full basis was 19.7 eV.

In the second approach used, the 163 atom cluster ($C_{93}H_{70}$) (Fig. 6.5) was used. The total energy of the perfect 163 atom cluster, with the innermost 29 carbon atoms in full basis, was determined, E_{pure}. Then an atom was removed from the substitutional site nearest the centre of the 163 atom cluster in order to create a vacancy. With the removal of the atom, the innermost 28 atoms were in full basis and these same atoms were allowed to move. The structure was allowed to relax in order to obtain the total energy of the cluster containing a relaxed neutral vacancy, E_{vac}. Using the total energy of the relaxed $\langle 100 \rangle$ split interstitial, E_{100}, as calculated in section 6.2.3, the formation energy of the vacancy–$\langle 100 \rangle$ split interstitial pair is calculated using,

$$E_{vac-int} = E_{vac} + E_{100} - 2E_{pure}. \tag{6.1}$$

The result obtained via this approach was 18.1 eV.

These values provide estimates of the energy required to form a relaxed vacancy–$\langle 100 \rangle$ split interstitial pair at a separation such that the relaxation of the atoms surrounding each defect does not significantly perturb the other defect. Our values can be compared with the sum of the formation energies of the neutral vacancy and $\langle 100 \rangle$ split interstitial calculated theoretically by Bernholc (1988) to be 23.9 eV and the crude experimental estimate of 35 eV (Bourgoin 1976). One possible reason why we obtain a smaller value than Bernholc is that in our cluster, both defects were not created in the neutral charge state, instead charge was able to transfer between them. That was not the case in Bernholc's calculation in which the formation energies were calculated separately. Also, as commented on in section 6.2.1, the size of the supercells used in Bernholc's calculations were small (only 16 atoms for the interstitial and 32 for the vacancy).

6.5 Conclusions

We have studied possible structures for an interstitial carbon atom in the diamond lattice and find that the $\langle 100 \rangle$ split interstitial defect structure is the only truly stable configuration for an interstitial atom in the neutral, -1, $+1$ and $+2$ charge states. This defect has D_{2d} symmetry and gives rise to two degenerate one electron energy levels of symmetry e in the band gap, into which two electrons are placed in the neutral charge state. We find the ground state of this structure to be ${}^1B_1$. However, a Jahn–Teller distortion is also possible in the neutral or singly charged states which splits the two one–electron levels and reduces the symmetry to D_2. The energy gain from this distortion is 0.6 eV for the neutral $\langle 100 \rangle$ split interstitial. The highest frequency vibrational modes of the neutral defect with D_2 symmetry were deduced to be an infra–red inactive A mode at 1984 cm^{-1}, followed by two infra–red active modes of B symmetry at 1560 and 1595 cm^{-1}. The hexagonal, tetrahedral, bond–centred and $\langle 110 \rangle$ split interstitial structures are all shown to be unstable. The interstitial atoms move to form $\langle 100 \rangle$ split interstitial defect structures in all cases when the atoms are relaxed. This result gives very strong evidence for the global stability of the $\langle 100 \rangle$ structure as in these relaxations atoms moved considerable distances and yet no other structure was formed. One possible reason for the stability of the $\langle 100 \rangle$ split interstitial is the preference for sp^2 bonding over sp^3 bonding that carbon atoms have, unlike other elemental semiconductors such as silicon.

We have studied the vacancy in diamond in the neutral and –1 charge states. For the neutral vacancy we find the four atoms surrounding the vacancy relax radially outwards from their positions in the perfect lattice structure by 0.2 Å, a different conclusion to some previous work. This defect can also undergo a Jahn–Teller relaxation which reduces the symmetry from T_d to D_{2d}. For optical experiments carried out at room temperature this is a dynamic effect. The Jahn–Teller energy was calculated to be 0.36 eV.

For the negatively charged vacancy we find the ground state to be ${}^4A_2$, in agreement with experiment. The atomic structure of this ground state retains the T_d symmetry of the lattice and the four atoms surrounding the vacancy relax radially outwards by 0.2 Å. A well known signature of this defect is the ND1 optical absorption at 3.149 eV, which is believed to occur between the ground state and an excited state ${}^4T_1$. We find the energy of this transition to be 3.3 eV, supporting this assignment. Clearly a Jahn–Teller distortion of this defect is possible. We find that the lowest energy distorted structure, which has spin $S = \frac{1}{2}$ and C_{3v} symmetry, has an energy 0.24 eV higher than than the undistorted structure, showing the dominance of many body effects over the energy gain from structural relaxation.

Finally, we estimate the formation energy for a vacancy–$\langle 100 \rangle$ split interstitial pair to be between 18 and 20 eV.

Chapter 7

Self Diffusion in Diamond

7.1 Introduction

The diffusion of carbon atoms in bulk diamond is an important process which could have implications in thin film synthetic diamond for defect aggregate production by subsurface atomic migration. It is known that vacancies enhance nitrogen aggregation and various nitrogen, nitrogen–vacancy, self–interstitial and vacancy aggregate structures have been reported or proposed which could affect the electrical properties of the diamond. In the difficult quest to produce n–type diamond, which could be used in the electronics industry, a knowledge of any defects altering these properties, or involved in some way in the creation of other defects which have this effect, is highly useful. Also, a knowledge of the diffusion processes involving the vacancy and self interstitial is important in considerations of the annealing of materials following irradiation, examples of which were given in section 6.1.

The determination, both experimentally and theoretically, of diffusion barrier energies is a difficult task. In the experimental case the diffusing species must be identified as well as finding an estimate of the energy with which it migrates. In the theoretical case the diffusion path, for which the lowest barrier energy to diffusion exists, must be found. It can be difficult to be certain that the saddle–point for diffusion has actually been elucidated. There have been many suggestions in work carried out in the field of self diffusion in diamond that the self interstitial is mobile at low temperatures. Some effects have been noted which could be attributed to this process, but no strong evidence has been presented linking experimental results solely to an interstitial structure. The lack of theoretical input into this area could be due to the fact that there are still conflicting opinions on the stability of the various possible structures for the self interstitial (this problem was addressed in Chapter 6). This multiplicity of possible structures leads to an even more complex set of possible paths that the diffusing atom can take.

Recently, comprehensive experimental work has been carried out on vacancy diffusion in diamond. The work in the first publication (Lawson 1992) was expanded upon by mainly the same authors (Davies 1992). Electron irradiated diamonds of type IaA, IaB, Ib and IIa were annealed at various temperatures. Use was made of the GR1 and ND1 absorption bands, associated with the neutral and negative vacancies respectively, which decreased as the vacancies migrated to trapping points (or as another species migrated to the vacancies). In the virtually nitrogen free type IIa diamond only the GR1 band is present. During annealing there was an initial fast decay of the GR1 band on which no measurements were taken. It was stated that this could have been due to self interstitial migration leading to vacancy–interstitial recombination. From the measurements taken after this initial ignored process had taken place it was deduced that the migrating species had a migration energy of 2.4 ± 0.3 eV. Proposing that the vacancy was the migrating species, both first and second order kinetic processes were taken into account in analysing the GR1 decay. However, the trapping points in type IIa diamond are unknown. During the annealing of types Ia and Ib diamond it was concluded that the only trapping points for vacancy migration were nitrogen defects. In type IaA diamond the destruction of the GR1 and ND1 bands corresponded to the growth of the H3 band. This latter band is associated with a defect formed from the trapping of a vacancy at an A centre nitrogen defect (i.e. two adjacent substitutional nitrogen atoms). The migration energy, which was found to be 2.3 $\pm$ 0.2 eV, was therefore attributed to the vacancy. As this is consistent with the migration energy found from annealing type IIa diamond, in which there is no ND1 band, it was concluded that it must be the neutral vacancy which migrates to the trapping point in type IIa material. It was stated that in type IaA diamond direct correlation of the GR1 loss with H3 growth was observed, but that the same was not true of the ND1 loss. The results on the graph used to illustrate this point in the article appeared to be inconclusive. It was also stated that "We do not expect different charge states of a defect to have the same migration energy". These points were then used to draw the conclusion that the neutral vacancy is the migrating species in type IaA diamond also, and that the negative vacancy must have a higher migration energy. A charge transfer process was then described by which a negative vacancy can become a neutral vacancy and then diffuse with the assumed lower migration energy of this defect. It was found that the charge transfer process would have to reversible. In an analogous way to the case of type IaA diamond, in type IaB diamond the destruction of the GR1 and ND1 bands was related to the growth of the H4 band. The H4 band is associated with the centre formed from the trapping of a vacancy at a B centre nitrogen aggregate, which consists of four nitrogen atoms. The migration energy was found to be 2.2 $\pm$ 0.2 eV and again attributed to the neutral vacancy. In both IaA and IaB diamond types two different rates were found for the decay of the GR1 and ND1 bands and for the growth of the H3 and H4 bands. Both had the same migration energy and evidence was presented for most of those vacancies being trapped more quickly being initially in more highly perturbed environments in the lattice than those vacancies which migrate to a trapping point more slowly. Perturbation of the lattice due to proximity of the vacancy to a nitrogen atom was proposed as being consistent with the results presented.

A theoretical estimate of the diffusion energy of the vacancy was made by Bernholc (1988) (the details of the calculational procedure are outlined in Chapter 6). The energy determined was the difference between the energy of a vacancy at an atomic site and an atom midway between two vacant, adjacent atomic sites. The position of this atom was therefore midway along what would be a carbon–carbon bond in the perfect lattice and on the bond axis. The two structures were relaxed, but without the known Jahn–Teller symmetry breaking distortion of the neutral vacancy being taken into account and without the diffusing atom being allowed to move off–site. The energy was found to be 1.7–1.9 eV, depending on the charge state, with the lower energies corresponding to the more highly charged defects. Lawson (1992) and Davies (1992) both quoted this result as 2.1 ± 0.1 eV for the migration energy of the neutral vacancy after the addition of the Jahn–Teller energy as determined experimentally by Davies (1982).

In this chapter we study the diffusion paths of the vacancy and carbon interstitial in diamond and estimate the energy barriers to the migration of these defects.

7.2 The Clusters and Bases Used

There were two clusters used in the investigation of diffusion paths for the vacancy and $\langle 100 \rangle$ split interstitial. These comprised 86 atoms ($C_{44}H_{42}$) (Fig. 7.1) and 176 atoms ($C_{98}H_{78}$) (Fig. 7.2). Both clusters were centred on the middle of a carbon–carbon bond and had D_{3d} symmetry, with this bond being aligned along the principal symmetry axis. These clusters therefore had two equivalent substitutional sites nearest to their centres and the diffusion of the appropriate defect was studied between these two sites. The

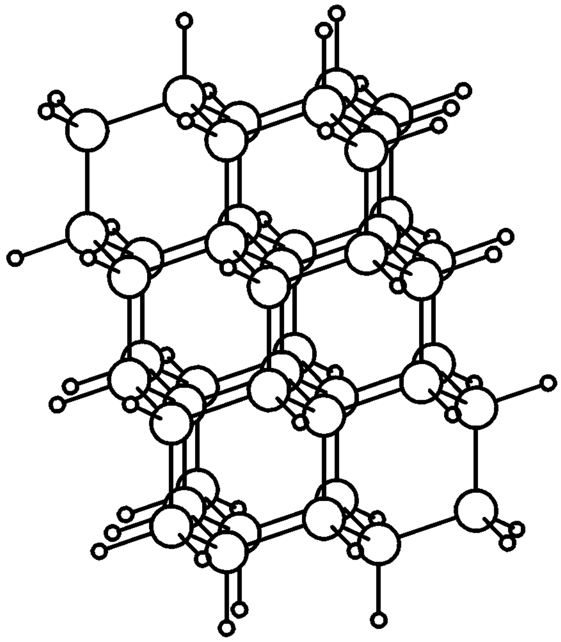


Figure 7.1: The 86 atom cluster $C_{44}H_{42}$ used in the investigation of diffusion.

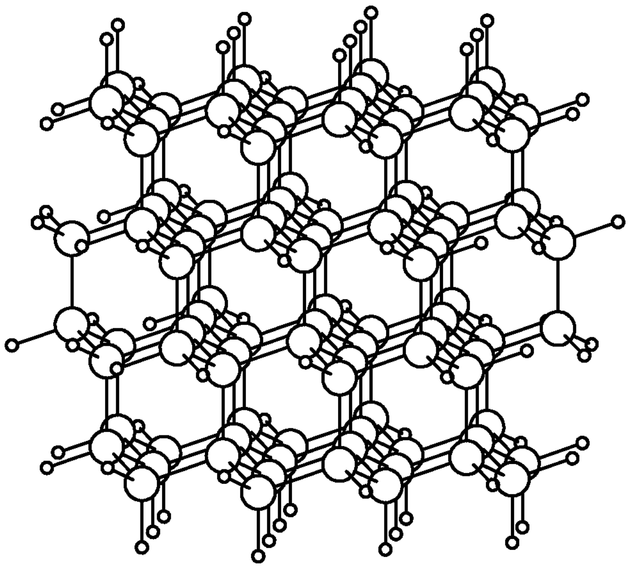


Figure 7.2: The 176 atom cluster $C_{98}H_{78}$ used in the investigation of diffusion.

bases used for the carbon and hydrogen atoms were as described in section 6.1.1.

7.3 Self Interstitial Diffusion

As the $\langle 100 \rangle$ split interstitial was found to be the only stable structure for an interstitial carbon atom in diamond, as shown in Chapter 6, it was only for this model of the interstitial that diffusion was investigated.

In the discussion of the diffusion paths which follows Figs. 7.3 and 7.4 show the important atoms involved in the diffusion path analysis, with labelling of the atomic sites. The atoms are shown in the perfect tetrahedral diamond lattice positions and cubic axes are used. The centre of the cluster and the origin of the Cartesian coordinate axes is at point E. The bond AF is oriented along the D_{3d} symmetry axis of the cluster, which is along the [111] direction. The initial positions of the atoms are given in Table 7.1.

Atom	x (Å)	y (Å)	z (Å)
A	-.44	-.44	-.44
B	-1.33	.44	-1.33
C	-1.33	-1.33	.44
D	.44	-1.33	-1.33
F	.44	.44	.44
G	1.33	1.33	-.44
H	1.33	-.44	1.33
I	-.44	1.33	1.33

Table 7.1: The positions of the atoms in the perfect lattice, as labelled in Figs. 7.3 and 7.4. The coordinate axes are aligned as described in the text.

The approximate path of diffusion for this defect is not easy to decide upon. Clearly, the lowest energy diffusion path will be the one which involves the stretching and breaking of the smallest number of bonds. After consideration, it was decided that a strong candidate was the path in which the diffusing interstitial atom passes through a bridge bonded structure between sites A and F. In this model an initial [001] oriented defect becomes oriented along either [010] or [100] after diffusion (these two end structures being equivalent).

The initial study of the diffusion path was carried out using the 86 atom cluster. With the addition of the interstitial atom, the innermost 9 atoms were in full basis and the innermost 27 atoms were allowed to move. Firstly a carbon atom was added to the cluster, to create the [001] and [010] oriented defects, at sites A and F respectively in Fig. 7.3. These two structures were relaxed to provide the end point coordinates for diffusion. The following prescription was then used to determine the path of minimum energy between the two interstitial sites.

1) Using the atomic positions thus obtained, the vectors of displacements connecting the positions of the equivalent atoms in the two relaxed end–point structures were determined, i.e. $\mathbf{v}_i = \mathbf{R}_i^A - \mathbf{R}_i^F$, where $\mathbf{R}_i^A$ is the position of atom i for the relaxed [001] split interstitial at site A, and $\mathbf{R}_i^F$ is the position of atom i for the relaxed [010] split interstitial at site F.

2) All the atoms were moved the same fraction, α, along these vectors giving the positions of the atoms at an intermediate stage in the diffusion process. This was carried out for a series of values of α ranging between 0 and 0.5.

3) For each of these structures the atoms surrounding and including the defect were again allowed to move to minimise the energy, but subject to the constraint that the vectors of further atomic displacements should be perpen-

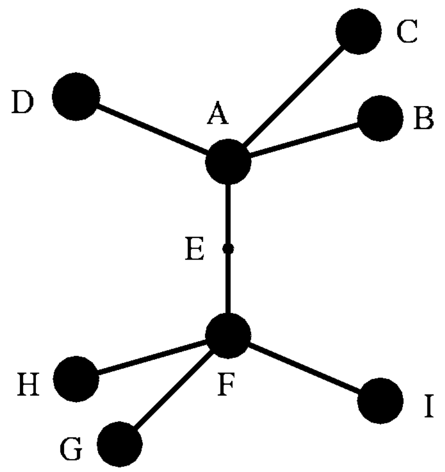


Figure 7.3: The atoms involved in the investigation of diffusion, with labelling as referred to in the text.

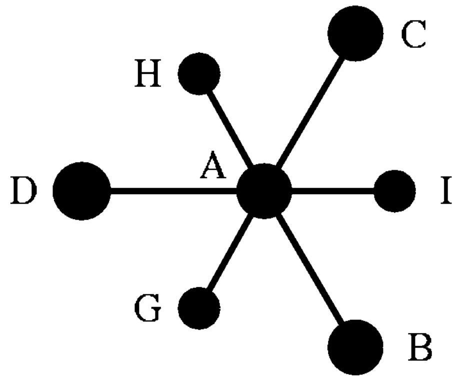


Figure 7.4: The atoms involved in the investigation of diffusion as viewed along the [111] direction.

dicular to the vectors $\mathbf{v}_i$ connecting the end–points defined above.

This procedure defines an energy profile for the diffusion process as a function of α and the maximum value of this gives the energy barrier to diffusion. By symmetry, the energies obtained for structures generated using values of α between 0.5 and 1.0 would be the same as for those generated using values of α between 0.5 and 0.

This approach was then repeated using the 176 atom cluster which, with the addition of the interstitial atom, had the innermost 27 atoms in full basis and the innermost 27 atoms were allowed to relax. The positions of the end–points of diffusion determined using this cluster are given in Tables 7.2 and 7.3, where atom P is the added interstitial atom.

Atom	x (Å)	y (Å)	z (Å)
P	-.43	-.43	.20
A	-.44	-.43	-1.05
B	-1.33	.54	-1.51
C	-1.32	-1.43	.64
D	.43	-1.44	-1.51
F	.44	.56	.66
G	1.31	1.34	-.42
H	1.37	-.42	1.41
I	-.44	1.43	1.44

Table 7.2: The positions of the atoms as labelled in Figs. 7.3 and 7.4 following the addition of an interstitial atom P to form a [001] oriented split interstitial centred on atomic site A and the subsequent relaxation of the structure in the 176 atom cluster.

Atom	x (Å)	y (Å)	z (Å)
P	.43	-.20	.43
A	-.56	-.66	-.44
B	-1.34	.42	-1.31
C	-1.42	-1.44	.44
D	.42	-1.41	-1.37
F	.43	1.06	.44
G	1.44	1.51	-.43
H	1.43	-.64	1.32
I	-.54	1.51	1.33

Table 7.3: The positions of the atoms as labelled in Figs. 7.3 and 7.4 following the addition of an interstitial atom P to form a [010] oriented split interstitial centred on atomic site F and the subsequent relaxation of the structure in the 176 atom cluster.

The results produced for the diffusion path in the two clusters were qualitatively the same, with small quantitative differences for the diffusion barrier energy. The energies for the relaxed structures corresponding to the various values of α in the 176 atom cluster are given in Table 7.4. It was therefore found that the saddle point occurs for $\alpha = 0.5$, midway between the two atomic sites. At this point the diffusing atom is displaced from point E by 0.62 Å in essentially the $[0\bar{1}1]$ direction. The atoms originally at the sites A and F are displaced from the perfect lattice positions by (0.08, -0.07, -0.40)Å and (-0.10, 0.45, 0.03)Å respectively. The atoms originally at sites B,C,D,G,H and I are displaced from the perfect lattice positions by no more than 0.2 Å. The positions of the atoms at the saddle point determined using the 176 atom cluster are given in Table 7.5.

α	Energy (eV)
0.5	1.67
0.45	1.68
0.4	1.67
0.3	1.24
0.2	0.62
0.1	0.17

Table 7.4: The energy for each relaxed point along the diffusion path, denoted by α, relative to the energy of the relaxed end–point structures.

Atom	x (Å)	y (Å)	z (Å)
P	.04	-.41	.46
A	-.36	-.51	-.84
B	-1.33	.44	-1.41
C	-1.27	-1.30	.63
D	.48	-1.42	-1.49
F	.34	.89	.47
G	1.33	1.42	-.45
H	1.28	-.66	1.32
I	-.48	1.50	1.42

Table 7.5: The positions of the atoms as labelled in Figs. 7.3 and 7.4, with the addition of an interstitial atom P, in the relaxed saddle point structure as determined using the 176 atom cluster.

The barrier for diffusion of the neutral $\langle 100 \rangle$ split interstitial along this path was then taken to be the difference in total energy between the relaxed end point structures and the relaxed saddle point structure. This was found to be 1.75 eV for the 86 atom cluster and 1.67 eV for the 176 atom cluster. We believe this to be the most reasonable path, but cannot rule out a lower energy mechanism. Rigorously it can be stated that we have determined an upper bound on the diffusion energy.

7.4 Vacancy Diffusion

The 86 atom cluster (Fig. 7.1) was used to investigate vacancy diffusion. With the removal of a carbon atom to form the vacancy, the innermost 25 atoms were in full basis and all the carbon atoms were allowed to move. All calculations on vacancy diffusion were carried out for both the neutral and –1 charge states. Referring to Fig. 7.3, one of the two equivalent atoms A and F in the perfect 86 atom cluster was removed and the 43 innermost atoms were allowed to move so that the total energy of the equilibrium structure was found. This produced the end–point structure for vacancy diffusion.

The approach first used to find the energy barrier for the diffusion of the vacancy was that used by Bernholc (1988). The assumption made in this case is that the path of diffusion involves an atom moving from a site neighbouring the vacancy, along the vector **AF**, to the vacancy site whilst all the surrounding atoms relax at each step. The saddle point was assumed to be midway between the substitutional sites. To investigate this, atoms A and F were replaced by one atom at point E. This atom was fixed at that position and the surrounding 42 atoms were relaxed to find the total energy of the saddle point of vacancy diffusion in this approach. The outwards relaxation of the surrounding atoms retained the D_{3d} symmetry and had a maximum magnitude of 0.13 Å. The difference in the total energies of the cluster containing the relaxed end point and that containing the relaxed saddle point was then determined. The value that was obtained was 3.3 eV using this approach for the neutral vacancy and 2.9 eV for the negatively charged vacancy.

A more complete investigation of the diffusion path was then carried out to find whether a lower energy saddle point existed. In this study, the two central atoms, A and F, were again replaced by a single atom which was placed at a point a certain fraction between the substitutional sites A and F, $\beta\mathbf{R}_A + (1-\beta)\mathbf{R}_F$, say, where $0 \leq \beta \leq 1$. This atom was then given a symmetry breaking displacement in the plane perpendicular to the vector **AF** of 0.15 Å in the [101] direction. The structure was then allowed to relax for each position, subject to the constraint that the diffusing atom remained in this plane, whereas all other atoms were allowed to move with no constraints imposed upon them. This is a far more general relaxation than that described in the previous paragraph as the diffusing atom can move off the [111] symmetry axis. This was repeated for a variety of fractional moves, β, and the total energy determined as a function of β relative to the relaxed end point structure. The results are given in Table 7.6 for the neutral vacancy and Table 7.7 for the negative vacancy.

With reference to Figs. 7.3 and 7.4 an accurate picture of the diffusion path can be visualised. By symmetry, the total energies for the diffusing atom in each (111) plane between points E and F along the second half of the diffusion path should be the same as those for the corresponding plane in the first half (between A and E), which are calculated here. The highest energy position for the atom between atomic sites A and E is hence the energy barrier for diffusion. We find that this point does in fact occur for $\beta = 0.5$, but at this saddle point the diffusing atom is 0.4 Å off the [111] axis for both charge states. In this way, the diffusion energy was found to be 2.8 eV for the neutral vacancy and 2.5 eV for the -1 vacancy.

The position of the diffusing atom can be written as the sum of two vec-

β	Energy (eV)	d_x (Å)	d_y (Å)	d_z (Å)	d (Å)
0.00	0.00	0.00	0.00	0.00	0.00
0.13	0.29	0.07	-0.04	-0.04	0.09
0.21	0.83	0.12	-0.06	-0.06	0.15
0.27	1.27	0.20	-0.10	-0.10	0.18
0.33	1.68	0.27	-0.13	-0.13	0.24
0.39	2.09	0.31	-0.16	-0.16	0.38
0.41	2.30	0.32	-0.16	-0.16	0.39
0.44	2.50	0.32	-0.16	-0.16	0.39
0.47	2.67	0.32	-0.16	-0.16	0.39
0.48	2.74	0.31	-0.16	-0.15	0.38
0.50	2.80	0.22	-0.29	0.07	0.37

Table 7.6: The energy change associated with the diffusion of a neutral vacancy. This is given as a function of β, which is defined in the text. Also given are the off axis displacement vector (d_x, d_y, d_z) and its magnitude, d. The energy is given in eV and is relative to the energy of a relaxed neutral vacancy.

tors, one being $\beta\mathbf{R}_A + (1-\beta)\mathbf{R}_F$, the movement of the atom along the [111] direction, and the other being the perpendicular displacement of the atom off the symmetry axis. As can be seen from Tables 7.6 and 7.7, the displacement is almost exactly along $[2\bar{1}\bar{1}]$ for $0 \leq \beta \leq 0.45$. For values of β between 0.45 and 0.55, the diffusing atom rotates around the [111] symmetry axis by 60°.

This diffusion path can be understood by considering the chemical bonding present. As the diffusing atom moves between points A and E it remains closest to, and equidistant from, atoms B and C until it is within 0.05 Å of the midpoint. By doing this only one bond, to atom D, is significantly lengthened,

β	Energy (eV)	d_x (Å)	d_y (Å)	d_z (Å)	d (Å)
0.00	0.00	0.00	0.00	0.00	0.00
0.12	0.29	0.07	-0.03	-0.03	0.08
0.21	0.80	0.08	-0.04	-0.04	0.10
0.27	1.21	0.13	-0.07	-0.07	0.16
0.33	1.57	0.27	-0.13	-0.13	0.32
0.38	1.85	0.31	-0.16	-0.16	0.38
0.41	2.00	0.32	-0.16	-0.16	0.39
0.44	2.16	0.32	-0.16	-0.16	0.39
0.47	2.33	0.31	-0.16	-0.15	0.39
0.48	2.40	0.31	-0.16	-0.16	0.38
0.50	2.47	0.24	-0.27	0.03	0.36

Table 7.7: The energy change associated with the diffusion of a negative vacancy. This is given as a function of β, which is defined in the text. Also given are the off axis displacement vector (d_x, d_y, d_z) and its magnitude. The energy is given in eV and is relative to the energy of a relaxed negative vacancy.

whereas atoms B and C move in such a way as to maintain a reasonable bond–length between themselves and the diffusing atom. This movement reduces the energy cost of this displacement. The path taken for $0.45 \leq \beta \leq 0.55$ is governed by the need for the diffusing atom to move through 60° in the (111) plane in order to follow a path to atomic site F whilst being again closest to, and equidistant from, two atoms. With reference to Fig. 7.4, if initially the atom travelled between points A and E with atoms B and C most closely bonded to it for most of the distance, then between points E and F it would move so that it was closest to either atoms I and G or atoms I and H. There are 3 equivalent diffusion paths on either side of the midpoint, each remaining closest to a different pair of atoms. These neighbouring atoms, of course, undergo large displacements. Atoms B and C move in the same direction as the diffusing atom, whilst atom D relaxes back towards the three atoms it is still attached to, changing its bonding towards sp^2 (as favoured by carbon atoms over sp^3) for a large part of the diffusion path, giving an additional gain of energy.

Finally, spin polarised calculations were carried out for the end point and saddle point of negative vacancy diffusion. The use of the local spin density is necessary to correctly model this charge state of the vacancy. All other details of the calculation were as for the spin averaged work described above. The end point calculation was performed with spin $S = \frac{3}{2}$, which is accepted as the spin state of the ground state of the negative vacancy (which was also verified in Chapter 6). The saddle point was assumed to be in the plane midway between the two atomic sites, as was found for both charge states in the spin averaged case. Two saddle point calculations were carried out with spin $S = \frac{3}{2}$ and $S = \frac{1}{2}$. The barrier energies were thereby found to be 3.4 eV for the $S = \frac{1}{2}$ case and 4.3 eV for the $S = \frac{3}{2}$ case. At the saddle point the diffusing atom is displaced off the [111] axis by 0.3 Å in both spin states.

7.5 Conclusions

We have calculated an upper bound for the migration energy of the $\langle 100 \rangle$ split interstitial to be 1.7 eV. There are no definite numerical results with which we can compare the value we have determined for self interstitial diffusion. It is, however, significantly lower than that for vacancy diffusion, leading us to conclude that the self interstitial should be mobile at lower temperatures than the vacancy. The unexplained and ignored fast process which took place on the annealing of type IIa diamond in the work carried out by Davies (1992) and

Lawson (1992) could indeed have been due to more highly mobile self interstitials. One of the findings of Chapter 6 was that the $\langle 100 \rangle$ split interstitial in its Jahn–Teller distorted structure with D_2 symmetry has a highest frequency vibrational mode that is infra–red inactive, but lower frequency modes which are infra–red active. In recent experimental investigations (Davies 1992, Lawson 1992) of vacancy diffusion energetics the characteristic optically active modes of particular defects were used to deduce the identity of the diffusing defect. None have yet been linked experimentally to the $\langle 100 \rangle$ split interstitial which could be used to measure changes in the concentration of the defect if it diffused to a trapping point which destroyed its vibrational properties. Bernholc (1988) dismissed the importance of the self interstitial in self diffusion due to it having a higher formation energy than the vacancy. However, the irradiation of diamond causing the ejection of an atom from a lattice site must produce equal concentrations of vacancies and interstitials. It is not known which would be the most effective trapping centres for self interstitial diffusion. Such trapping points as dislocations, interstitial aggregations or substitutional nitrogen atoms have been proposed in the past as well as possible straightforward recombination with a vacancy or the diffusion of the atom to the the surface. It is highly probable that the importance of the diffusion of this defect depends on where it eventually migrates to.

We have calculated the energy barrier to the diffusion of the vacancy to be 2.8 eV for the neutral vacancy and 3.4 eV for the negative vacancy. The value for neutral vacancy diffusion can be compared favourably with the experimentally determined value of 2.3 ± 0.3 eV (Davies 1992, Lawson 1992). This experimental value was attributed to the diffusion of the neutral vacancy, and it was deduced that the diffusion energy of the neutral vacancy is less than that of the negative vacancy, as discussed in section 7.1. Our result agrees with this conclusion, and disagrees qualitatively with the theoretical result (Bernholc 1988) that found the vacancy diffusion energy to be 1.7–1.9 eV, with the lower energies for the higher charge states and the highest for the tetrahedral neutral vacancy. Although we agree with the experimental work as to the ordering of diffusion energies with respect to charge state, we feel that aspects of the experimental evidence on which this deduction is based are inconclusive. Therefore, we believe that further and more complete experimental work would be valuable in confirming that neutral vacancy diffusion has a significantly lower energy barrier than the negative vacancy and not simply a very similar energy barrier.

Finally, it is tempting to compare our results with work carried out over three decades ago by Palmer (1961), although no conclusions could be drawn at that time as to the nature of the migrating species. By the annealing of one type IIa diamond Palmer found two activation energies for the destruction of the neutral vacancy centres of 1.62 eV and 2.35 eV. It is difficult to assess the accuracy of this work as only one type IIa diamond was annealed. However, it is interesting to note that the lower value is in good agreement with our value for self interstitial diffusion and the higher value is reasonably consistent with our lowest value for vacancy diffusion.

Chapter 8

The Reorientation Energy of Substitutional Nitrogen

8.1 Introduction

Nitrogen is the dominant impurity in natural diamond, and this is generally also the case in synthetic diamond. The dominant impurity in those diamonds classified type Ib has been shown to be the isolated neutral substitutional nitrogen atom. This centre is ESR active, giving rise to the P1 paramagnetic centre (Smith 1959) and exhibits infra–red absorption at 1130 cm^{-1} and 1344 cm^{-1} (Collins 1982). Many theoretical studies have been carried out on this defect. Early approaches used extended Hückel theory (Messmer 1973), Green functions (Brand 1977) and CNDO (Mainwood 1979). More recently, *ab initio* studies have been carried out using an unrelaxed Green function technique (Bachelet 1981), the supercell (Kajihara 1991) and cluster (Jones 1992b, Briddon 1993) approaches. The work of Jones and Briddon produced vibrational frequencies which agree with experiment and solved the one unresolved problem associated with the defect, that of the source of the 'anomolous' vibrational mode at 1344 cm^{-1}. Both the theoretical and experimental work published on this defect centre provides overwhelming evidence that the microscopic structure of the isolated substitutional nitrogen is as proposed by Smith (1959). This structure is shown in Fig. 8.1 and the bond lengths and angles, as determined by Jones and Briddon, are given in Table 8.1. This calculation determined that there were three equivalent carbon–nitrogen bonds (N–C2) and one extended bond (N–C3) which is 28% longer than the perfect diamond lattice carbon–carbon bond. Other estimates of the length of this extended bond include 5–10% (Smith 1959), 5–10% (Mainwood 1979), 36% (Ammerlaan 1981) and 25% (Kajihara 1991).

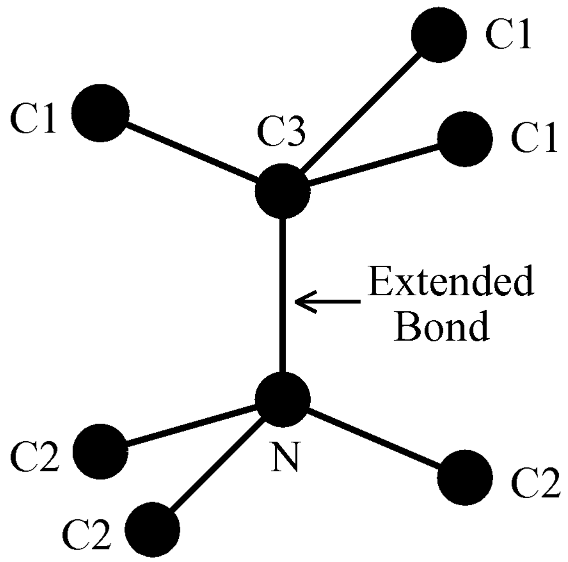


Figure 8.1: The structure of the neutral substitutional nitrogen in diamond. The nitrogen atom is labelled N and the carbon atoms are labelled C1,C2 and C3 depending on their location in the structure. Identical labelling implies equivalence of position. The N–C3 bond is along the $\langle 111 \rangle$ direction and is significantly longer than the carbon–carbon bond in diamond. This bond is also the principle symmetry axis of the defect, which has C_{3v} symmetry.

In a tetrahedral lattice there are four equivalent nitrogen–carbon bonds along $\langle 111 \rangle$ directions associated with a substitutional nitrogen which could be the unique bond. An EPR study by Ammerlaan (1981) investigated the energy barrier to the reorientation of the unique bond to one of the other three carbon neighbours. Using the hyperfine splitting of the P1 EPR signal, the relative populations of the defect orientations were monitored. In thermal equilibrium the defects should align in equal numbers along the four degenerate orientations. Ammerlaan applied uniaxial stress to type Ib diamond along the $[0\bar{1}1]$ direction in order to lift the four–fold degeneracy and make two equivalent orientations higher in energy than two other equivalent orientations. This causes the stress induced alignment of the defect along the lower energy directions in accordance with Boltzmann's law. The stress was then removed and the isothermal annealing of the stress aligned defects in returning to a random alignment based on four–fold degeneracy was studied. It was found that at high temperatures (197–1230 K) the reorientation obeys Arrhenius law with first order kinetics and has a thermal activation energy of 0.7 eV. At low temperatures (78–197 K) the reorientation energy was higher than expected and an argument in terms of the quantum mechanical tunneling of the atoms through a potential barrier of height 0.7 eV was presented as an explanation. Two earlier experimental estimates of the reorientation energy, both at temperatures exceeding 600 K, by Shul'man (1967) and Loubser (1967), yielded the results 0.7 eV and 0.76 eV respectively

In this chapter we determine the energy barrier to the reorientation of the extended bond in the neutral substitutional nitrogen defect in diamond.

8.2 Calculational Details

The cluster that we have used in this work is $C_{70}H_{60}N$ (Fig. 8.2) which comprises 131 atoms, including 60 terminating hydrogen atoms and a central substitutional nitrogen atom. The cluster has tetrahedral symmetry before any

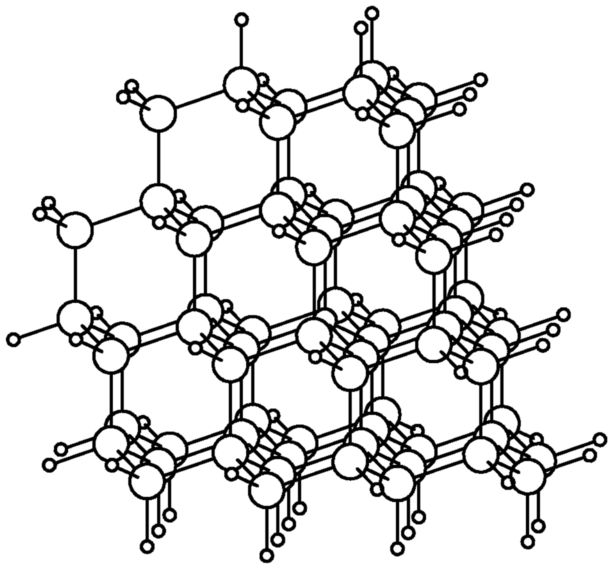


Figure 8.2: The cluster $C_{70}H_{60}N$ used in the calculations. The hydrogens are the atoms with the smallest radii, at the surface of the cluster, and the nitrogen atom is at the centre of the cluster with a smaller radius than that of the carbon atoms.

atomic motion is allowed, with the atoms at the perfect diamond lattice positions. The defect has the terminating hydrogen atoms as 4th nearest neighbours. The pseudopotentials of Bachelet (1982) are used for the nitrogen and carbon atoms and the bare coulombic potential is used for the hydrogens.

The charge density and wavefunction of the cluster are fitted using bases of Gaussian–type orbitals centred on the atoms and also on the centres of the bonds connecting the nitrogen to its four nearest neighbour carbon atoms and these carbon atoms to the 12 atoms in the next shell out. The innermost 16 carbon atoms are in full basis and the remaining carbon atoms are in the contracted basis. The bases centred on the individual atoms and bond–centres are as follows.

For the charge density :

The nitrogen atom has 10 orbitals of different widths.

The carbon atoms have 4 orbitals of different widths.

The terminating hydrogen atoms have 3 orbitals of different widths.

The bond centres have 1 orbital.

For the wavefunction :

The nitrogen atom has 32 orbitals (s,p_x,p_y and p_z type orbitals each of 8 different widths).

The full basis carbon atoms have 16 orbitals (s,p_x,p_y and p_z type orbitals each of 4 different widths).

The contracted basis carbon atoms have 4 orbitals (Each a linear combination of a particular type of orbital (s,p_x,p_y and p_z) each of which has 4 different widths).

The terminating hydrogen atoms have 2 orbitals (A linear combination of the s–type orbitals of two different widths and a linear combination of the p–type orbitals (p_x,p_y and p_z) each of two different widths).

The bond centres have 4 orbitals (An s,p_x,p_y and p_z type orbital each of the same width).

The calculations use the local spin density approximation in order to take into account any possible spin polarisation effects. The innermost 71 atoms were allowed to move in each case in order to determine the equilibrium structure, these comprised all the carbon atoms and the nitrogen atom.

8.3 The Substitutional Nitrogen Atom

We initially determined the equilibrium structure of the neutral substitutional nitrogen. A full analysis of this defect will not be presented here as this defect has been previously studied by the same *ab initio* method that we use in this thesis (Jones 1992b, Briddon 1993) and this involved a comprehensive discussion of the extension of the carbon–nitrogen bond. We will, however, give the quantitative results that we obtained for the structure and energy levels of the defect as this study uses a larger cluster (131 atoms instead of 86), spin polarisation, and more atoms were allowed to move in the optimisation of the structure.

	Briddon	This Work
Bond Lengths		
N–C3	1.97 Å	2.05 Å
C3–C1	1.50 Å	1.48 Å
N–C2	1.47 Å	1.46 Å
Bond Angles		
C3–C1–C3	117°	117.2°
C2–N–C2	115°	114.5°

Table 8.1: The bond lengths and angles corresponding to the structure of the relaxed neutral substitutional nitrogen in diamond, as shown in Fig. 8.1. The results of a previous *ab initio* study using the same formalism and an 86 atom cluster by Jones (1992b) and Briddon (1993) are shown for comparison with the present calculations, which were carried out in a 131 atom cluster.

One of the neighbouring carbon atoms of the nitrogen was moved radially outwards by 0.2 Å in order to induce the extended bond between the nitrogen and one carbon atom to form. On relaxation the structure formed was that given in Table 8.1 (with reference to Fig. 8.1). An energy level associated with the defect appeared in the band gap at 1.9 eV below the conduction band edge. In the formation of the extended bond the carbon and nitrogen moved away from each other along the $\langle 111 \rangle$ direction. The carbon moved 0.3 Å off–site and the nitrogen moved 0.2 Å off–site (with reference to the atomic positions in the perfect diamond lattice). This bond lengthening was repeated with another carbon neighbour so that we had the positions of the end–points for the reorientation of the the extended bond of the nitrogen atom.

8.4 The Determination of the Reorientation Energy

In order to investigate the motion of the atoms in the defect between their positions at the two end–points, we used the approach described in section 7.3 for the diffusion of the self interstitial in diamond. The energies of the structures following constrained atomic relaxation are given in Table 8.2 as a function of the fraction along the vector connecting the end–point structures, α. At the halfway point, $\alpha = 0.5$, the nitrogen atom has moved back onto the substitutional site at the centre of the cluster. The bonds to the first nearest neighbour carbon atoms are 1.54 Å and those from the first to the second nearest neighbour carbons are 1.55 Å, which are basically equal to those in the perfect diamond lattice (1.54 Å). The angles between the innermost 17 atoms varies between 108.5° and 109.5°, which is essentially that of a tetrahedral structure (all angles 109.5°). Therefore the mid–point for reorientation of the extended nitrogen–carbon bond appears to be that with the atoms in the perfect diamond structure with the nitrogen atom on the substitutional site.

α	Energy (eV)
0.1	0.04
0.2	0.18
0.3	0.40
0.35	0.54
0.375	0.61
0.4	0.68
0.425	0.68
0.45	0.67
0.475	0.66
0.5	0.66

Table 8.2: The energy for each relaxed point along the diffusion path, denoted by α, relative to the relaxed end–point structures.

This implies that both the nitrogen and its unique carbon neighbour must surmount an energy barrier to regain the perfect lattice positions before the bond between any of the four neighbouring carbons and the nitrogen can again extend to complete the reorientation process.

8.5 Conclusions

The significant difference between the structural results that we obtained here for the neutral substitutional nitrogen defect and those of Jones (1992b) and Briddon (1993) was that the carbon neighbour with the extended bond relaxed further back into the plane of its three neighbouring carbon atoms. The magnitude of radial relaxation that we obtained was 0.3 Å (Briddon obtained 0.2 Å). This makes the nature of the carbon–carbon bonding even closer to

sp^2, as opposed to the sp^3 bonding in diamond, which is a stronger type of interatomic bond for carbon atoms and so is expected to be favoured. This greater outwards relaxation might be a consequence of allowing a larger number of atoms to relax or a many–body spin polarisation effect. The important bond lengths N–C2 and C3–C1 (Fig. 8.1), which determine the vibrational frequencies, change by only 0.01 Å and 0.02 Å respectively. Therefore the vibrational frequencies are expected to change very little from those determined by Jones and Briddon for the defect.

The reorientation energy of the one extended nitrogen–carbon bond present in the equilibrium structure of the neutral substitutional nitrogen atom in diamond has been calculated to be 0.7 eV. This is in excellent agreement with an experimental estimate by Ammerlaan (1981) of 0.7 eV. At the mid–point of the reorientation of the extended nitrogen–carbon bond the structure has tetrahedral symmetry with all four of the nearest neighbour carbon atoms of the nitrogen equivalent. The barrier to reorientation of the bond is the energy required to reach this configuration. It is then equally probable that any of the four nitrogen–carbon bonds will extend to become the unique bond.

Chapter 9

Conclusions

In this thesis we have presented the results of work carried out in modelling defects in diamond using an *ab initio* approach which can be applied to many other solids and to molecules. The conclusions and discussions pertinent to the work carried out in Chapters 5–9 were included in those chapters. In this chapter we shall discuss general areas of significance in the implementation of the methods that we use. We shall also address extensions and adaptations to the formalism which have recently been developed and implemented.

Calculations based on the formalism derived in Chapter 3 have been carried out for several years. The details of the original methodology were published in 1988 by Jones. A comprehensive presentation of the theoretical approach and some applications was given by Briddon (1990b). Since that time the computational procedures used to implement our formalism have been continually improved and new areas of theory have been appended to the computer code developed to perform *ab initio* calculations. The result has been a considerable decrease in the time taken to calculate a self–consistent total energy and also to minimise the total energy of the cluster with respect to the positions of the atoms. This has lead to clusters such as the 164 atom cluster used in Chapter 6 and the 178 atom cluster used in Chapter 7 regularly being used in order to model defects. The number of atoms for which calculations can be performed in a practical time has effectively doubled in three years. The largest number of atoms that we have now modelled is 304 in work which is discussed in Chapter 6.

A significant step in the increased speed of optimisation of the structure of a cluster has been the switch from numerical to analytical evaluation of the forces acting on the atoms. This has also considerably reduced the time taken for the calculation of the double derivatives of the total energy with respect to the displacement of the atoms. When this increase in speed was combined with the added symmetry analysis routine described in Chapter 3 the complete *ab initio* evaluation of the dynamical matrix for the determination of vibrational frequencies became an easily accomplished task. A new line minimisation routine which is used at each step of the conjugate gradient minimisation of the total energy during structural optimisation has been developed which makes use of these analytical forces. Cubic interpolation is the primary tool used for the one parameter minimisation, and it is in general quite a robust minimisation scheme which is quicker than the previous scheme used. Under certain conditions, especially near the absolute minimum of the total energy, instabilities can be encountered. Then the more stable (but slower) Brent's method minimisation subroutine is reverted to. The number of self–consistent energies which must be evaluated during the minimisation is also (usually) less using this new algorithm, further reducing the time it takes to determine an equilibrium atomic structure.

The development of the formalism to use a spin polarised charge density instead of a spin averaged charge density has allowed the study of spin polarised defects. Spin polarised calculations use the local spin density approximation in constructing separate charge densities for spin up and spin down electrons. The study of the multiplets of the negative vacancy and the self interstitial in Chapter 6 depend on this. In fact, without spin polarisation the ground state of the negative vacancy is not given correctly by our calculations.

The determination of diffusion paths and barrier energies is an avenue of research which has recently been explored. A method for the implementation of such studies was described in Chapter 7 and used in work reported in Chapters 7 and 8. This approach has yielded results for migration energies which compare favourably with experimental estimates. The four diffusing species examined in this thesis (the neutral and negative vacancies, the self interstitial and the local motion of the substitutional nitrogen atom) have all essentially been single atoms. The determination of diffusion energetics involving two diffusing atoms is currently being studied for the case of the interchange of a substitutional nitrogen atom with a neighbouring carbon atom in diamond. It has proved more difficult to efficiently elucidate the migration path of minimum energy in this case. Further development of approaches to the study of the simultaneous diffusion of more than one atom may be needed before this class of diffusion is as readily analysed as that for a single diffusing species.

The eminent suitability of our approach to studying molecules, as it uses a localised orbital expansion in atomic clusters, has been utilised recently in calculations performed on C_{60} by Faulhaber (1993a), Estreicher (1992) and Jones (1992a) and on the novel organic superconductors bis(ethylenedithio)-tetrathiafulvalene (BEDT-TTF) and tetrathiafulvalene (TTF) by Faulhaber (1993b). The fact that our approach is tailor–made for molecules, unlike the plane wave expansion in supercells used by many other researchers performing *ab initio* computer modelling, leaves us well placed to take advantage of the current interest in such molecular structures as fullerene–type molecules by contributing theoretical input into that field of research.

The basis expansion of the wavefunction has been modified to include d–type Gaussian orbitals. This makes possible the study of materials or defects involving transition metals. Calculations have already been performed on Nickel impurities in diamond and Zinc impurities in InAs. The modelling of atoms containing d-orbitals is computationally much more intensive than those containing only s- and p-orbitals, so that it may be some time before the goal of modelling defects in such materials as II–VI compounds is realised. In this case approximately half the atoms would be required to have a significantly larger basis for expanding the wavefunction than in those calculations performed for diamond, silicon and GaAs, for example. This goal is, however, very worthwhile as II–VI and related materials should be increasingly important in future technology.

A supercell approach has recently been implemented using the same formalism presented in this thesis. Both the wavefunction and the intermediate fit to the charge density are still expanded in terms of localised orbitals, so that the benefits of being able to have a variable basis on the atoms modelled are still present. The supercell approach would eliminate the problems associated with cluster termination inherent in modelling certain materials using the cluster approach. In the modelling of III–V or II–VI compounds, if a non–stoichiometric cluster of atoms is used then the cluster will be charged. This can lead to a charge of, for example, ± 9 for the tetrahedral 131 atom cluster used in some of the work presented in this thesis. Stoichiometric clusters of atoms which have been generated to produce a neutral cluster have the problem that there is a large dipole moment across the cluster due to atoms of a particular type (and charge) being aligned along opposite surfaces. The problem is even more acute in the modelling of more highly ionic materials, when the nature of the bonding of the surface atoms to what would be the next layer of atoms out causes an additional problem. The termination of the cluster using point charges has been investigated as a possible solution. However, as the issue of termination does not arise when using a supercell, this approach seems to offer a straightforward solution to the above difficulties.

Finally, an application of our calculations currently being investigated is the possibility of evaluating the hyperfine couplings of the atoms studied. This requires the determination of the true spin polarised charge density at the nucleus and two approaches are being considered to this end. The atomic charge density could be integrated outwards from the centre of the nucleus, with the boundary condition that at the core radius of the pseudopotential it equals the charge density of the atom at that point that is produced by the complete cluster calculation. Alternatively the pseudopotential calculation could be replaced by an all electron calculation for the system.

To summarise, recent efforts in adapting our formalism to overcome problems and implement useful and desirable theory have lead to an increase in the number of materials which can in principle be treated by our approach. We also now have the ability to calculate additional properties which can be compared with experiment. This increases the power of our method in analysing defect structures. With the continuing decrease of the cost of computing power, the use of *ab initio* calculations in modelling increasingly more computationally demanding atomic systems seems assured.

References

Ammerlaan C. A. J., Inst. Phys. Conf. Ser. No. **59**, 81 (1981)

Astier M. and Bourgoin J. C., "International Conference on Radiation Effects in Semiconductor", Nice (1978).

Bachelet G. B., Baraff G. A. and Schluter M., Phys. Rev. B **24**, 4736 (1981).

Bachelet G. B., Hamann D. R. and Schluter M., Phys. Rev. B **26**, 4199 (1982).

Baüerlein R., Z. Naturforsch A **14**, 1069 (1959).

Baüerlein R., Proceedings of the International School of Physics "Enrico Fermi", Academic, New York, 358 (1962).

Baüerlein R., Z. Phys. **176**, 498 (1963).

Bernholc J., Antonelli A., Del Sole T. M., Bar–Yam Y. and Pantelides S. T., Phys. Rev. B **61**, 2689 (1988).

Bourgoin J. C. and Massarani B., Phys. Rev. B **14**, 3690 (1976).

Brand S. and Jaros M., Solid State Comm. **21**, 875 (1977).

Brent "Algorithms for Minimisation Without Derivatives ", Prentice-Hall, Chapter 5 (1973).

Breuer S. J. and Briddon P. R., Phys. Rev. B **49**, 10332 (1994).

Briddon P. R., Jones R. and Lister R. M. S., J. Phys. C **21**, L1027 (1988).

Briddon P. R. and Jones R., Phys. Rev. Lett. **64**, 2535 (1990a).

Briddon P. R., PhD thesis, University of Exeter (1990b).

Briddon P. R. and Jones R., Physica B **185**, 179 (1993).

Brommer K. D., Needels M., Larson B. E. and Joannopoulos J. D., Phys. Rev. Lett. **68**, 1355 (1992).

Ceperley D. M., and Alder B. J., Phys. Rev. Lett. **45**, 566 (1980).

Chadi D. J., Phys. Rev. B **46**, 9400 (1992).

Christiansen P. A., Lee Y. S. and Pitzer K. S., J. Chem. Phys. **71**, 4445 (1979).

Clark C. D., Kemmey P. and Mitchell E. W. J., Discuss, Faraday Soc **31**, 96 (1961).

Clark C. D. and Walker J., Proc. R. Soc. Lond. A **334**, 241 (1973).

Cohen M. L., Science **261**, 307 (1993).

Collins A. T., Stanley M. and Woods G. S., Phil. Mag. A **46**, 77 (1982).

Coulson C. A. and Kearsley M. J., Proc. R. Soc. London Ser. A **241**, 433 (1957).

Davies G., Proc. R. Soc. Lond. **338**, 359 (1974).

Davies G., Nature **269**, 498 (1977).

Davies G., J. Phys. C **14**, L391 (1981).

Davies G., J. Phys. C **15**, L149 (1982).

Davies G., Lawson S. C., Collins A. T., Mainwood A., Sharp S. J., Phys. Rev. B **46**, 13157 (1992).

Denteneer P. J. H., Van de Walle C. G., and Pantelides S. T., Phys. Rev. B **39**, 10809 (1989).

Denteneer P. J. H., Van de Walle C. G., and Pantelides S. T., Phys. Rev. B **41**, 3885 (1990).

Estreicher S. K., Latham C. D., Heggie M. I., Jones R., Oberg S., Chem. Phys. Lett. **196**, 311 (1992).

Faulhaber J. C. R., Ko D. Y. K., Briddon P. R., Phys. Rev. B **48**, 661 (1993a).

Faulhaber J. C. R., Ko D. Y. K., Briddon P. R., Synthetic Metals **60**, 227

(1993b).

Hartree D. R., Proc. Camb. Phil. Soc. **24**, 89, 111, 426 (1928).

Hohenberg P. and Kohn W., Phys. Rev. **136**, B864 (1964).

Isoya J. *et al*, Phys. Rev. B **45**, 1436 (1992).

Jones R., J. Phys. C **21**, 5735 (1988).

Jones R., Latham C. D., Heggie M. I., Torres V. J. B., Öberg S., Estreicher S. K., Phil. Mag. Lett. **65**, 291 (1992a).

Jones R., Briddon P. R. and Öberg S., Phil. Mag. Lett. **66**, 67 (1992b).

Kajihara S. A., Antonelli A., Bernholc J. and Car R., Phys. Rev. Lett. **66**, 2010 (1991).

Kerker G. P., J. Phys. C. **13**, L189 (1980).

Kahn L. R. and Goddard III W. A., J. Chem. Phys. **56**, 2685 (1972).

Kohn W., Phys. Rev. **94**, 1409 (1954).

Kohn W. and Sham L. J., Phys. Rev. **140**, A1133 (1965).

Lannoo M. and Stoneham A. M., J. Phys. Chem. Solids **29**, 1987 (1968).

Lawson S. C., Davies G., Collins A. T. and Mainwood A., J. Phys. Cond. Matter bf 4, L125 (1992).

Liu A. Y. and Cohen M. L., Science **245**, 841 (1989).

Liu A. Y. and Cohen M. L., Phys. Rev. B **41**, 10727 (1990).

Loubser J. H. N. and van Ryneveld W. P., Brit. J. Appl. Phys. **18**, 1029 (1967).

Lowther J. E., Phys. Rev. B **48**, 11592 (1993).

Lowther J. E. and van Wyk J. A., accepted for publication in Phys. Rev. B. (1994).

Lunqvist S. and March N. H., editors, "The Theory of the Inhomogeneous Electron Gas", Plenum, New York (1983).

Mainwood A., Larkins F. P. and Stoneham A. M., Solid State Electronics **21**, 1431 (1978a).

Mainwood A., J. Phys. C **11**, 2703 (1978b).

Mainwood A., J. Phys. C **12**, 2543 (1979).

Malta D. M., Vonwindheim J. A., Fox B. A., Appl. Phys. Lett. **62**, 2926 (1993).

Melius C. F. and Goddard III W. A., Phys. Rev. A **10**, 1528 (1974).

Messmer R. P. and Watkins G. D., Phys. Rev. B **7**, 2568 (1973).

Mori Y. *et al*, Japanese J. of Appl. Phys. **32**, 601 (1993).

Niu C., Lu Y. Z. and Lieber C. M., Science **261**, 334 (1993).

Palmer D. W., PhD thesis, University of Reading (1961).

Polak E. "Computational Methods in Optimisation", Academic Press, New York (1971).

Phillips J. C. and Kleinman L., Phys. Rev. **116**, 287 (1959).

Press W. H., Flannery B. P., Teukolsky S. A., Vetterling W. T., "Numerical Recipes", Cambridge University Press, Cambridge (1992).

Redondo A., Goddard III W. A. and McGill T. C., Phys. Rev B. **15**, 5038 (1977).

Sambe H. and Felton R. H., J. Chem. Phys. **61**, 3862 (1975).

Shul'man L. A., Zaritskii I. M. and Podzyarei G. A., Sov. Phys. – Solid State **8**, 1842 (1967).

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

Slater J. C., "The Self–Consistent Field for Molecules and Solids", McGraw–Hill, New York (1974).

Smith W. V., Sorokin P. P., Gelles L. L. and Lasher G. J., Phys. Rev. **115**, 1546 (1959).

Stoneham A. M., Solid State Comm. **21**, 339 (1977).

Topiol S. Zunger A. and Ratner M. A., Chem. Phys. Lett. **49**, 367 (1977).

von Barth U., Phys. Rev. B **20**, 1693 (1979).

Walker J., Vermeulen L. A. and Clark C. D., Proc. R. Soc. Lond. A **341**, 253 (1974).

Weigel C., Peak D., Corbett J. W., Watkins G. D. and Messmer R. P., Phys.

Rev. B **8**, 2906 (1973).

Zunger A., J. Vac. Sci. Technol. **16**, 1337 (1979).