Investigations on the Origin of Brown Colouration in Diamond

Luke Sebastian Hounsome

March 29, 2007

Submitted by Luke Sebastian Hounsome, to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Physics, March 2007.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

— Luke Hounsome

Oh, Diamond! Diamond! thou little knowest what mischief thou hast done!

Isaac Newton (1642 – 1727)

Abstract

This thesis documents theoretical *ab-initio* calculations undertaken with the purpose of discovering the source of brown colouration in diamond. Brown diamond has a unique absorption continuum across the bandgap, with no detectable onset at low energies. The source of colour is not known, but is vital for understanding the loss of colour which occurs during High-Pressure High-Temperature (HPHT) treatment. It can be surmised that the defect responsible for the colour introduces a broad spectrum of energy levels into the bandgap, allowing a range of electron transitions to occur.

Small vacancy clusters and chains are identified in diamond by Electron Paramagnetic Resonance (EPR) studies. Modelling of these defects reveals that the vacancy clusters are the more energetically favourable defects in all cases. In many of the defects the lowest energy spin state is $S=1$, which introduces bandgap levels due to unpaired electrons. The levels are clustered at the midgap and so will not induce a broad optical absorption.

Dislocations are a prevalent defect in natural diamond, with optically active types proposed as a source of brown colour. The colour removal would come from transformation to optically inactive types, with energy supplied by the HPHT treatment. Two dislocation types are studied, the 90° glide and the 90° shuffle types. The formation energy of the glide type is much lower that the shuffle type, indicating that it is a more stable defect. The shuffle type has dangling bonds which introduce absorption throughout the gap. However the density of shuffle dislocations required to match experimental observations is larger than the observed density in natural diamond.

Networks of π-bonds can form on the (111) surface of diamond, introducing π and π^* states into the bandgap. These states cause a wide range of electron transitions to be available, potentially leading to broad optical absorption. The (111) surface is recreated internally by removal of the αB double plane, leaving a vacancy disk. This gives a large area of π-bonds, and is found to be a stable defect, with an energy much lower than small vacancy clusters. The optical absorption of the disk is a close match for experimental measurements on type IIa diamond. There is also a close match to Raman frequencies detected in brown diamond.

Two different mechanisms for removal of the brown colour are proposed. In CVD diamond there is a large amount of hydrogen incorporated during the growth process which can move to the disk surface, passivating the π-bonds and removing any optical absorption. In natural diamond there is nothing to check the growth of the disk until it collapses into a stacking fault bounded by a dislocation loop. These different mechanisms could account for the different annealing temperatures required for natural and CVD diamond.

The optical absorption of brown type Ia diamond is characterised by additional absorption around 550 nm (2.25 eV), so the possibility of nitrogen interacting with the disk to modify the absorption is investigated. The most stable configuration for nitrogen at the disk surface is a nitrogen pair separated into a N-C-C-N arrangement. If the density of nitrogen at the surface is small enough to allow sizable chains of π-bonded carbon to form, then a broad absorption is induced, with a peak at 2.3 eV in excellent agreement with experiment. It can be surmised that vacancy disks are a strong candidate for the source of brown colour in both type IIa and type Ia diamond.

Acknowledgements

This thesis, and indeed all the work undertaken for my Ph.D. would not have been possible without the support of many people. I must firstly give my thanks to my supervisor, Prof. Bob Jones, who has guided and educated me over the last three years. His ability to recall any measured value and tell you in what paper it was published continues to impress. If it were not for his inspiration to publish and present my work I am sure my studies would hold much less value. I hope he feels that the response to the age-old question 'any progress', can now definitely be 'yes'!

As well as guidance from my supervisor, my studies have been made easier by the help of the rest of the AIMPRO group. Whether it be a deep discussion of formation of defects in semiconductors, useful scripts to perform tasks (Steve!) or just how to change my desktop background someone was always able to help. So thanks to Steve, James, Thomas, Naomi, Colin, Alexandra, Patrick, Malcolm and all the others.

Aside from my immediate colleagues I must thank my friends inside and outside the department for providing non work-related support. Steve H's dedication to tea-drinking always gave the opportunity for a useful break, and working in the Ram Bar took my mind off endless calculations. Nomads CC will give me many fond memories of summer evenings chasing a small red ball! In no particular order then, thanks Paul, John, Matt, Martin, Tristan, Leigh, Prof. G.P. Srivastava, Nathan, Pete V, Ian, P.C, Bex, Gobby, Helen, Claire, Tom and anyone I've overlooked.

Financial support for this work has been provided by the EPSRC and a CASE award from DTC. In addition Philip Martineau and David Fisher from DTC have been very helpful in providing experimental data and useful discussion.

Thanks must go to my family for their unending encouragement, and most importantly to my fiancé Catherine. She has always been there when I needed her, not least in proofreading all 180 pages! Completion of this thesis is only the second-most important event of this year. Thank you Catherine.

List of Publications

First author

- *Optical properties of vacancy related defects in diamond*
  L. S. Hounsome, R. Jones, P. M. Martineau, M. J. Shaw, P. R. Briddon, S. Öberg, A. T. Blumenau, and N. Fujita
  Physica Status Solidi (a) **202**, 2182 (2005).

- *Origin of brown colouration in diamond*
  L. S. Hounsome, R. Jones, P. M. Martineau, D. Fisher, M. J. Shaw, P. R. Briddon, and S. Öberg
  Physical Review B **73**, 125203 (2006).

- *Photoelastic constants in diamond and silicon*
  L. S. Hounsome, R. Jones, M. J. Shaw, P. R. Briddon
  Physica Status Solidi (a) **203**, 3088 (2006).

Co-author

- *Combined TEM and STEM study of the brown colouration of natural diamonds*
  R. Barnes, U. Bangert, P. M. Martineau, D. Fisher, R. Jones and **L. S. Hounsome**
  Journal of Physics: Conference Series **26**, 157 (2006).

- *Electron energy loss spectroscopic studies of brown diamonds*
  U. Bangert, R. Barnes, **L. S. Hounsome**, R. Jones, A. T. Blumenau, P. R. Briddon, M. J. Shaw, and S. Öberg
  Philosophical Magazine A **86**, 4757 (2006).

Contents

5 EELS of Bulk Diamond and Dislocations 99

6 (111) Plane Vacancy Disks 113

7 Photoelastic Constants in Diamond and Silicon 142

8 Conclusions 153

List of Tables

List of Figures

Chapter 1

Introduction

Carbon is a common element in nature, occurring naturally in three pure forms and in many compounds such as carbon dioxide (CO_2), or methane (CH_4). Its small size and affinity for bonding with other elements mean it is found in solids, liquids and gases. This prolificacy leads to the concept of *carbon-based life*, where all the elements that lifeforms require are supplied in carbon-containing molecules. Indeed, no lifeform has ever been observed that is not reliant on carbon in some way.

In nature, carbon is most commonly found in the allotropes of graphite, diamond and amorphous carbon. Graphite is the most abundant allotrope as it is the most thermodynamically stable *i.e.* it has the lowest energy per atom at low temperatures and pressures. Diamond is less stable, but once formed the barrier to transformation to graphite is prohibitively large, hence carbon can remain in the diamond form for millenia. Amorphous carbon, such as soot, is generally just formed of small graphite particles. Further solid allotropes may be artificially synthesised, notably the buckminsterfullerene C_{60} [1, 2] and carbon nanotubes [3, 4], but also C_{20}, C_{80} and others.

1.1 Properties of Diamond

The main properties of diamond are well-known to non-scientists, even if the physical causes are not. Diamond is extremely hard [5], with a value of 10 on the Moh scale, due to the strong covalent bonding between carbon atoms. The actual value of hardness depends on the crystal face, with a value of 137 GPa for the (100) face and a value of 167 GPa for the (111) face [6]. The bulk modulus is 442 GPa [7]. Diamond (especially when cut) has a distinct sparkle arising from its large value of refractive index ($\sim$2.4), and exhibits *fire*; the display of a range of colours arising from a dispersion value of $\sim$0.04.

Diamond's further unique properties are less well-known outside its field of research. It has a high thermal conductivity, which is unusual for electrical insulators. Diamond has no free electrons, unlike metals, but the strong covalent bonds and low atomic mass allow thermal energy to be transmitted quickly. This leads to a thermal conductivity of 895–2300 W m^{-1} K^{-1} at 300 K [8], depending on the purity. This value is to be compared to the low value of silicon at 148 W m^{-1} K^{-1} and of copper at 401 W m^{-1} K^{-1}, both at 300 K (although copper has a higher conductivity at very low temperatures [9]).

The value of the thermal conductivity depends on the diamond type (section 1.1.3) and the temperature [10]. Type I diamonds have the lowest thermal conductivity, which has been linked to the nitrogen content [11] as phonons can be scattered by point defects [12]. This proposal is consistent with the observed thermal conductivity for type IIa, which has the highest values. The conductivity is strongly dependent on temperature, with a maximum value for type IIa reached at $\sim$70 K. The magnitude at 70 K is approximately 9 times larger than at 300 K, and approximately 17 times larger than at 10 K [9].

1.1.1 Structure

The structure of all crystals can be described by the combination of a lattice and a basis. The lattice is a regular array of imaginary points in space, akin to a grid. The basis is the relation of the real atomic positions to the lattice points. The diamond structure (which is also the structure of silicon and germanium) consists of the Face Centred Cubic (FCC) lattice (figure 1.1), with lattice vectors of (1/2 1/2 0), (1/2 0 1/2) and (0 1/2 1/2), plus a two-atom basis of (0 0 0) and (1/4 1/4 1/4). The basis notation denotes that there is an atom at each lattice point, and one at 1/4 $\mathbf{x}$ + 1/4 $\mathbf{y}$ + 1/4 $\mathbf{z}$ relative to each lattice point, where $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ are the conventional unit cell vectors. This gives any atom 4 nearest neighbours and 12 next-nearest neighbours.

Figure 1.2 shows the full diamond structure, obtained when the two-atom basis is added to the FCC lattice. The value a_0 is the lattice constant, which has a value of 3.566 Å

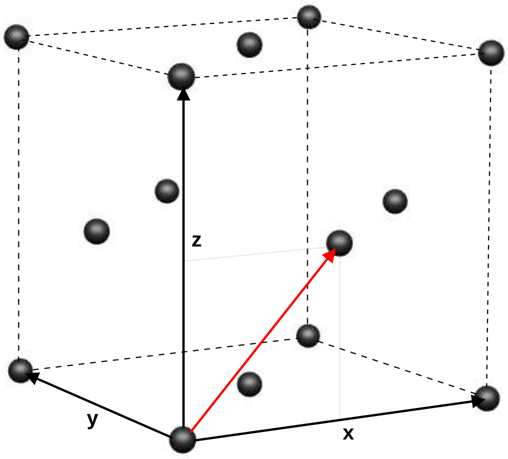


Figure 1.1: The Face Centred Cubic lattice, with lattice points represented by black spheres. One of the face-centred points is denoted with the vector relative to the corner point.

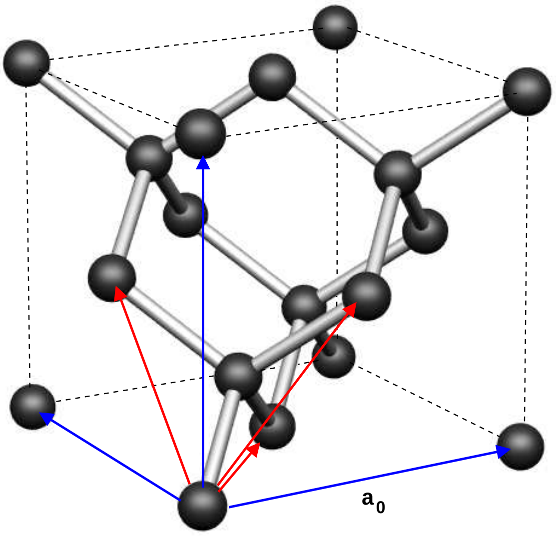


Figure 1.2: The diamond structure, with the basis atoms added to the FCC lattice. The conventional unit cell vectors are shown in blue, and the primitive unit cell vectors in red.

at 300 K [9]. With a (0 0 0) and (1/4 1/4 1/4) basis the distance between atoms is then given by $\sqrt{3}/4$ a_0, which yields a value of 1.544 Å. The blue arrows in the figure denote the vectors of the conventional unit cell of the FCC lattice, which is generally favoured over the primitive unit cell for calculations of density etc. due to its cubic structure. The atomic planes of diamond are referenced relative to these vectors. The primitive unit cell is described by the vectors: $A_1 = 1/2\hat{\mathbf{i}} + 1/2\hat{\mathbf{j}} + 0\hat{\mathbf{k}}$, $A_2 = 1/2\hat{\mathbf{i}} + 0\hat{\mathbf{j}} + 1/2\hat{\mathbf{k}}$ and $A_3 = 0\hat{\mathbf{i}} + 1/2\hat{\mathbf{j}} + 1/2\hat{\mathbf{k}}$. These represent the minimum volume that can be repeated (along those vectors) to create the diamond structure, and are shown in red in figure 1.2.

The conventional unit cell contains eight atoms in total; each face atom is shared between two cells and each corner between eight. The volume of the unit cell is a_0^3, which is 45.346 Å^3 or 45.346×10^{-30} m^3. Therefore there are 1.764×10^{29} atoms m^{-3}, or as it is more usually expressed, 1.764×10^{23} atoms cm^{-3}. This density is greater than germanium and silicon which have densities of ∼5×10^{22} atoms cm^{-3} [9].

1.1.2 Electronic structure

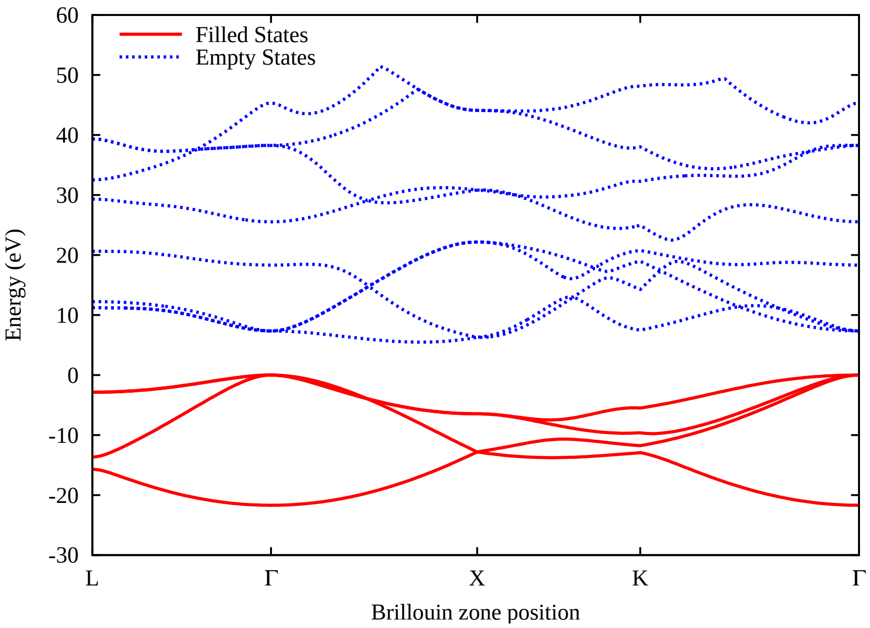


Figure 1.3: The electronic bandstructure of diamond, with the valence band top set to 0 eV. Conduction band energies are scaled up by 1.3 to compensate for the bandgap underestimation of LDA.

Diamond is a semiconductor with an indirect bandgap of 5.5 eV. The indirect gap is located between the Γ point at the Brillouin zone centre and a position 3/4 towards X [9] as seen in figure 1.3. The direct gap is at Γ and has a value of ∼7 eV. In figure 1.3, the calculated energies of the levels above the valence band top have been scaled by a factor 1.3 which brings both the direct and indirect gaps into agreement with experiment. Such a scaling is necessary due to the underestimation of the bandgap by the Local Density Approximation (LDA), described in section 2.4.2. The gap to transitions covers the IR-Vis-UV range, which is why pure diamond is colourless; colour is imparted by defects introducing gap levels, which allow photo-excited electron transitions. The indirect gap also means that phonon assistance is required for electronic transitions, limiting diamond's usefulness in electronic applications unless it is doped to modify the bandstructure, as described below.

As LDA is known to underestimate the bandgap, work has been undertaken to make corrections using the screened exchange method [13], described in detail in section 2.8. When the bandstructure of diamond is evaluated using this new method, the indirect gap is found to be 5.30 eV, compared to 4.27 eV without correction (LDA), and 5.5 eV experimentally. The direct gap using screened exchange is 6.60 eV, compared with 5.63 eV in LDA, and $\sim$7 eV experimentally. Using the previously mentioned 1.3 scaling factor for the LDA bandgap gives an indirect gap of 5.55 eV and a direct gap of 7.32 eV. While such a scaling corrects for the indirect gap it overestimates the direct gap, whereas the screened exchange obtains 96% accuracy for the indirect gap and 94% for the direct, which are more accurate relative to each other. A bandstructure using screened exchange correction is shown with an unscaled LDA bandstructure for comparison in figure 1.4.

As described above, the vast majority of diamond is electrically insulating, due to the lack of free electrons. However type IIb diamonds are p-type semiconducting due to the presence of boron impurities. Boron is a group 3 element, with one less electron than carbon, and so introduces an acceptor level close to the diamond valence band. It is energetically favourable for an electron from the valence band to occupy the acceptor level, which leaves a hole in the valence band, thus maintaining electrical neutrality. However, while the electron is spatially confined to the boron atom, the hole is free to travel through the diamond structure and thus there is an effective excess of positive charge carrying quasi-particles. Such type IIb diamonds are rare in nature and are highly valued due to their blue colour. Type IIb diamonds can be produced artificially by either Chemical Vapour Deposition (CVD) growth techniques [14], High-Pressure High-Temperature (HPHT) growth [15], or implantation [16].

N-type semiconducting diamond does not occur in nature, but is required for electronics applications; *e.g.* a transistor constructed from alternating layers of electron-conductive (n-type) or hole-conductive (p-type) material in an npn (or pnp) arrangement. Thus both

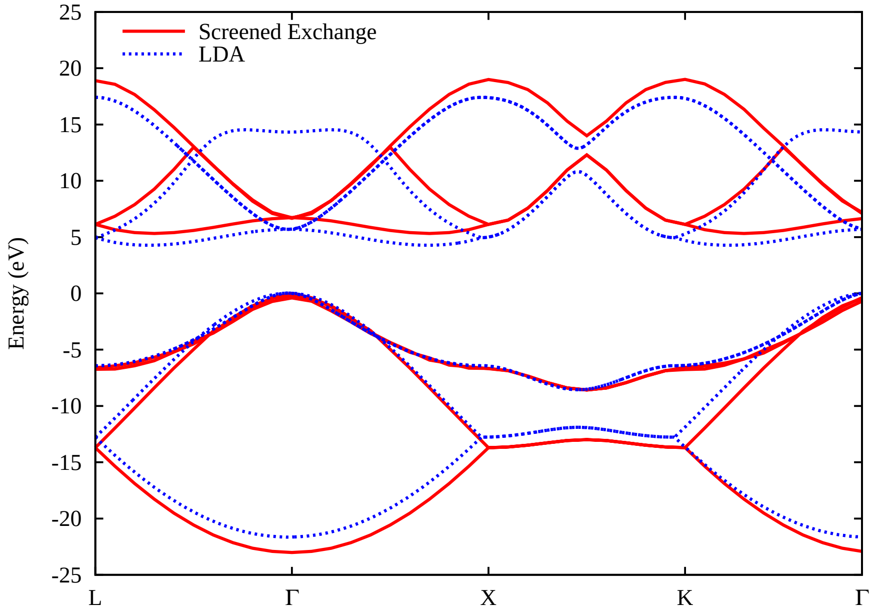


Figure 1.4: Bandstructure of bulk diamond calculated using both the screened exchange correction and with LDA. The screened exchange corrections recreates experimental results more accurately, without the need for any scaling. The valence band maxima are set to 0 eV.

experimental [17, 18] and theoretical [19, 20] work has focused on the growth of n-type diamond. N-type growth has been partially successful, using phosphorus as a dopant.

1.1.3 Classification

Diamond can be split into two types [21], which can then be further subdivided. Type I encompasses diamond that contains nitrogen in high concentrations, while type II diamond has little or no nitrogen. The various subdivisions are listed in table 1.1.

Type Ia diamonds are most commonly found, and type IIb the least, although pink type IIa diamonds are the rarest specimens. Artificial diamonds have the same category system applied, however certain growth methods tend to produce certain types. Nitrogen is usually added to the gas mixture in CVD growth and hence many of these diamonds are type Ib, although as technology improves, less nitrogen needs to be used. HPHT-grown diamonds usually have less inclusions when nitrogen is present.

Table 1.1: The types and subtypes of diamond, with their main defects and colours.

Type	Subtype	N concentration (atoms cm^{-3})	Defect centres	Colour
I	aA	$\sim 10^{20}$	A	Colourless Brown Pink
I	aB	$\sim 10^{20}$	N3, H3, B	Yellow Brown Pink
I	b	$\sim 10^{20}$	C	Intense Yellow Brown
II	a	$< 10^{17}$	Dislocations	Colourless Pink Red Brown
II	b	$< 10^{17}$	Boron	Blue Grey

Certain colours can be associated with certain types of diamond, as listed in table 1.1. This is due to the presence of defects which are associated with the colour, such as nitrogen defects in type I diamond. Green diamonds are produced through irradiation (either natural or artificial) to produce isolated vacancies, which induce absorption at 1.67 eV (740 nm) [22]. Hence the diamond type is less important.

1.2 Defects in Diamond

Defects in diamond can be of two types. Point defects consist of no more than a few atoms or vacancies, so are the order of 1 nm in size. Extended defects are much larger (at least in one direction) and may be hundreds of nanometres in extent, to the point where they are visible optically.

1.2.1 Point defects

Point defects in diamond can be either intrinsic or extrinsic. Intrinsic or self-defects, are misalignments of the diamond structure, creating vacancies or interstitials, and involve no foreign elements. Extrinsic defects involve foreign elements *e.g.* nitrogen, but many de-

Table 1.2: Point defects in diamond, and their known properties [23].

Defect	Structure	Symmetry	Optical Absorption	Donor /Acceptor
A centre	N-N pair	D_{3d}	$\sim$4 eV	(+/0) $E_c - 4$ eV
B centre	VN_4	T_d	-	-
C centre	N sub	C_{3v}	$>$2 eV	(+/0) $E_c - 1.7$ eV
N3 centre	VN_3	C_{3v}	3 eV	-
H3 centre	N-V-N	C_{2v}	2.463 eV	-
H2 centre	N-V-N^-	C_{2v}	1.257 eV	(+/0) $E_c - 2$ eV
VN centre	V-N	C_{3v}	2.156 eV	-
W15 centre	V-N^-	C_{3v}	1.945 eV	-
VN^+ centre	V-N^+			
VNH centre				
Boron	B sub	T_d	-	(0/-) $E_v + 0.37$ eV
Neutral vacancy	V^0	T_d	1.673 eV	(+/0) $E_c - 3.2$ eV
Negative vacancy	V^-	T_d	3.150 eV	(+/0) $E_c - 3.2$ eV
Divacancy	V-V	C_{2h}	2.543 eV	-
Interstitial	C_i	D_2	1.685; 4 eV	-

fects are a combination of both intrinsic and extrinsic. Depending on the electron configuration point defects may be detectable by Electron Paramagnetic Resonance (EPR) [23], or their electronic structure may produce photoluminescence (section 3.5). The main defects and their known properties are listed in table 1.2.

1.2.2 Extended defects

Extended defects in crystals can, like point defects, either be intrinsic or extrinsic. Unlike other semiconductors however, diamond generally only exhibits intrinsic defects in the form of dislocations, although the presence of nitrogen in carbon interstitial platelets on (100) planes has been detected [24]. Dislocations are present at densities of 10^8–10^9 cm^{-2} in natural diamond [25], but at much lower densities in CVD or HPHT-grown artificial diamond, often only 10^4 cm^{-2}. They generally exhibit photoluminescence in the blue or orange regions, which can be used to differentiate between natural and artificial stones [26].

Dislocations were first suggested in a material by Mügge and Ewinr in the late nineteenth century, in relation to metals [27]. It was interpreted that observed slip was related to the

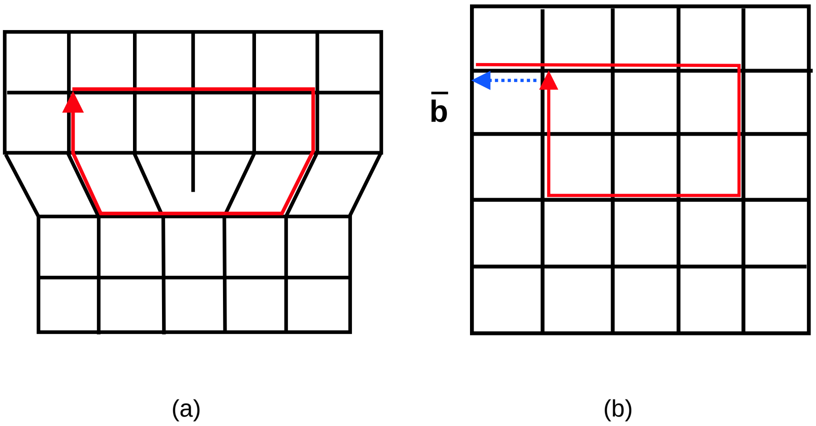


Figure 1.5: A perfect edge dislocation, with the Burgers circuit shown in (a) and the equivalent in a perfect lattice in (b). The closure of the circuit yields the Burgers vector **b** perpendicular to the line direction, which is out of the page.

shearing of one portion of a crystal with respect to another, upon a plane. A dislocation is defined by two vectors, the Burgers vector **b** and the line vector **t**. The line vector is directed along the dislocation core. The Burgers vector is defined in a more complex manner by imagining a closed circuit surrounding the dislocation of interest, this circuit must finish exactly where it starts but otherwise may consist of any combination of distances and directions. If this exact circuit is then reproduced in a section of perfect crystal structure, there will be an offset between the start and end points of the circuit, the direction and length of the vector required to close the circuit is then defined as the Burgers vector **b**. Examples of the loop required to construct the Burgers vector are shown in figures 1.5 and 1.6.

The edge and screw dislocations are the simplest to visualise and construct. The edge dislocation (figure 1.5) is created by the insertion (or removal) of a half-plane of atoms causing stress perpendicular to the inserted plane. The line direction **t** in this instance is perpendicular to the direction of **b**, so the edge dislocation is also called a 90°dislocation. The screw dislocation (figure 1.6) is caused by the offset of one region of the lattice with respect to another region along a plane, with the magnitude of the offset an integer multiple of the lattice constant. In contrast to the edge dislocation, it can be seen that **b** is parallel to **t**.

In reality, many dislocations are mixtures of the edge and screw types, with Burgers vector

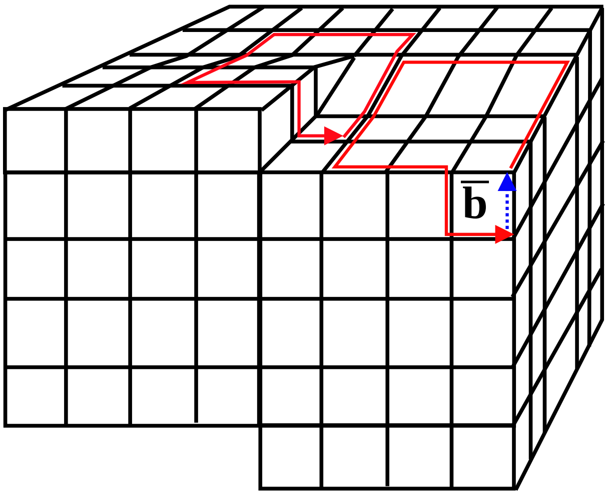


Figure 1.6: The perfect screw dislocation, with the Burgers vector **b** parallel to the line direction.

at *e.g.* 45°or 60°to the line direction. In diamond the main slip planes are the $\{111\}$ set of planes, and the dislocations usually lie in the $\langle 110 \rangle$ directions. The perfect dislocation in diamond has a Burgers vector $\frac{1}{2}\langle 110 \rangle$ allowing two basic types to form. The screw dislocation has a minimum Burgers vector of $\frac{1}{2}[1\bar{1}0]$, which is parallel to its line direction, whereas the 60° dislocation, which is a combination of screw and edge types, has a Burgers vector of $\frac{1}{2}[01\bar{1}]$.

The two atom basis of diamond effectively causes two planes to be formed in the structure, with unequal spacing. If a dislocation is formed by insertion of a half-plane of atoms (the 60° dislocation), then two further discrete types of dislocation can be defined, both with the same Burgers vector described above. The glide type of dislocation terminates between the narrowly spaced planes, as shown in 1.7, the extra half plane terminating on the $\frac{a}{12}\langle 111 \rangle$ plane. The shuffle type terminates on the widely spaced planes, the $\frac{a}{4}\langle 111 \rangle$ plane. It is notable that conversion between glide and shuffle types is achieved by insertion or removal of a half-plane of atoms [28].

The screw dislocation, described above, can also be classified as a shuffle or glide type, with a third type know as a mixed screw. In figure 1.7 the possibility that glide and shuffle planes run diagonally, as well as horizontally/vertically, is evident.

In his thesis A. T. Blumeanau shows several possibilities for core structures in diamond [29], reproduced in figures 1.8 and 1.9, with kind permission. These figures show a view along

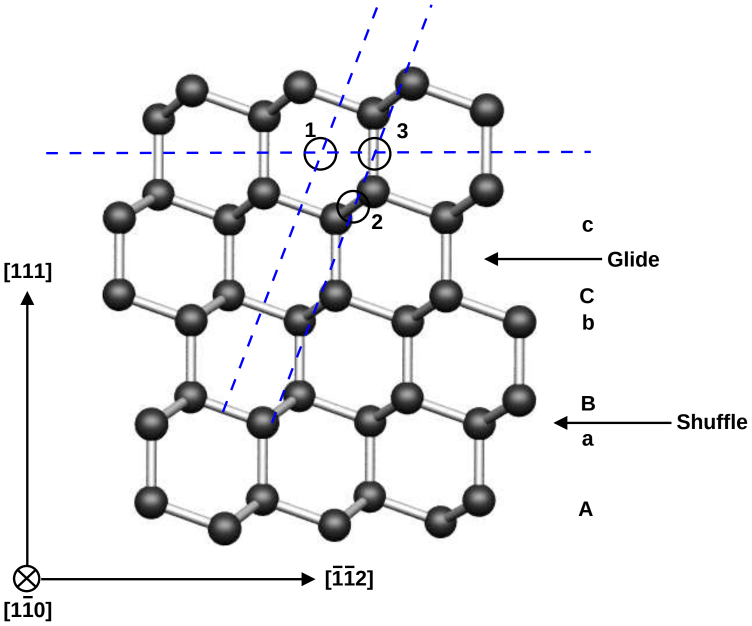
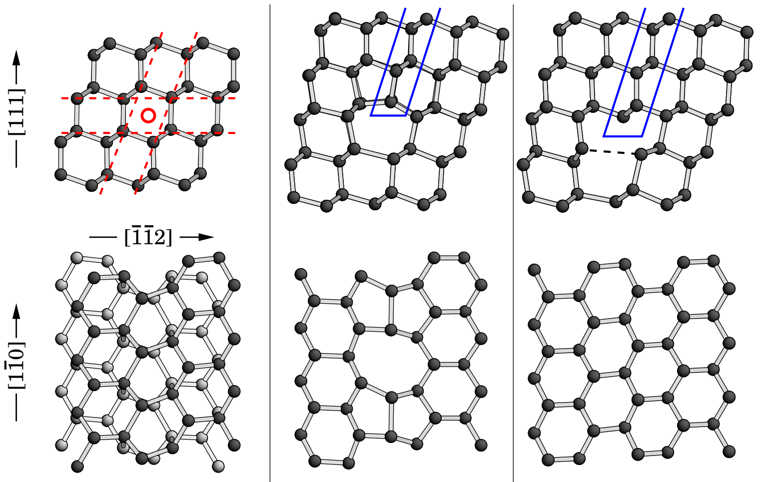


Figure 1.7: The possible dislocation configurations in diamond. (1) is the shuffle type, (2) is the glide type and (3) is mixed.

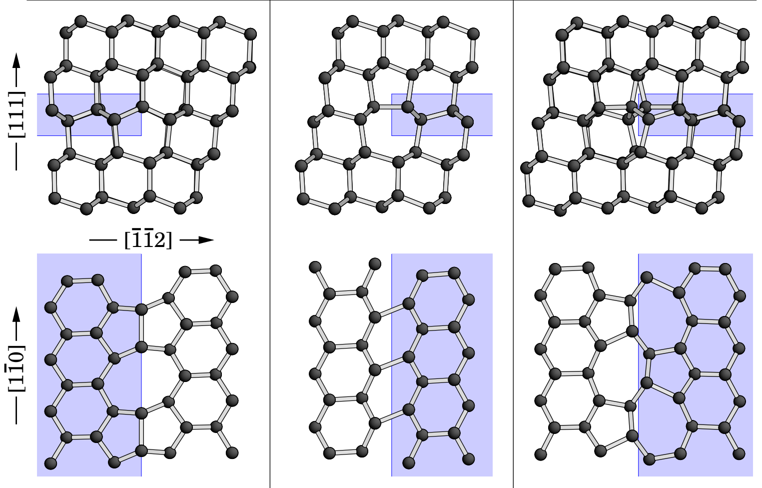
the dislocation line in the [1$\bar{1}$0] direction and a view of the [111] glide plane for each dislocation. Stacking fault regions accompanying the partials are shaded. In the case of the 90° partial the single and double period reconstructions are denoted in the labels by SP and DP respectively.

The insertion or removal of pairs of layers, *e.g.* Aa or Bb in figure 1.7, creates a stacking fault; a change in the stacking sequence of the crystal planes. Removal of a layer creates an intrinsic stacking fault, with the addition of a layer leading to an extrinsic stacking fault. Stacking faults are important when considering partial dislocations which, by definition, have Burgers vectors which are not translation vectors of the lattice. These dislocations necessarily border a two-dimensional defect, which is usually a stacking fault.

In reality perfectly straight dislocations are almost non-existent, they contain kinks and jogs. The potential-energy path that an atom would have to move through can be considered as a regular series of peaks and valleys, with each extreme separated by the Burgers vector **b** of the dislocation. This layout is known as the Peierls potential. This potential defines special low energy directions in which the dislocation can lie, the dislocation running in the minimum of the Peierls potential, with the crossing kinks as short as possible.



(a) shuffle screw (b) 60° glide (c) 60° shuffle

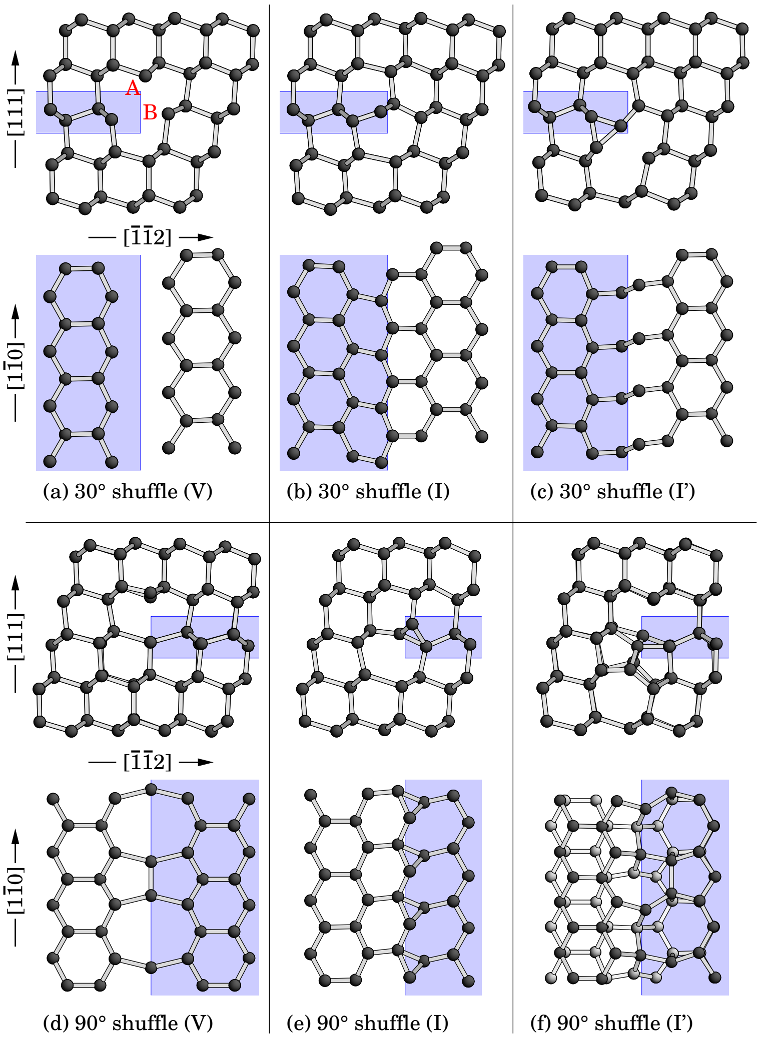


(d) 30° glide (e) 90° glide (SP) (f) 90° glide (DP)

Figure 1.8: Dislocation structures in diamond I. Reproduced from [29].

A kink is a step in the dislocation line which is fully contained within the glide plane, the plane of movement of a dislocation defined by the Burgers vector. Once a kink has crossed into a neighbouring valley it may cross back to the original valley to form a double kink. Subject to stresses the two kinks may then move apart, causing the entire dislocation line to move, as illustrated in figure 1.10. The movement of a dislocation without assistance and within the glide plane is known simply as glide.

Jogs can be considered as all breaks or steps in a dislocation line, but it is customary to use it to describe steps that are not contained in the glide plane, or move from one glide plane to another. Jogs in dislocations are prime places for the emission and absorption of point defects e.g vacancies and interstitials, and the effect of this is to cause movement of



(a) 30° shuffle (V)
(b) 30° shuffle (I)
(c) 30° shuffle (I')
(d) 90° shuffle (V)
(e) 90° shuffle (I)
(f) 90° shuffle (I')

Figure 1.9: Dislocation structures in diamond II. Reproduced from [29].

the dislocation. This is known as climb movement, it does not take place in the glide plane of the dislocation. The necessity of movement of point defects to achieve climb motion makes this process dependent on temperature, but it is a major mechanism in allowing dislocation lines to pass otherwise insurmountable obstacles such as foreign elements in the lattice. Many natural diamonds (and all brown ones) exhibit plastic deformation, which is associated with dislocation movement and thought to be the major source of vacancies in natural diamond.

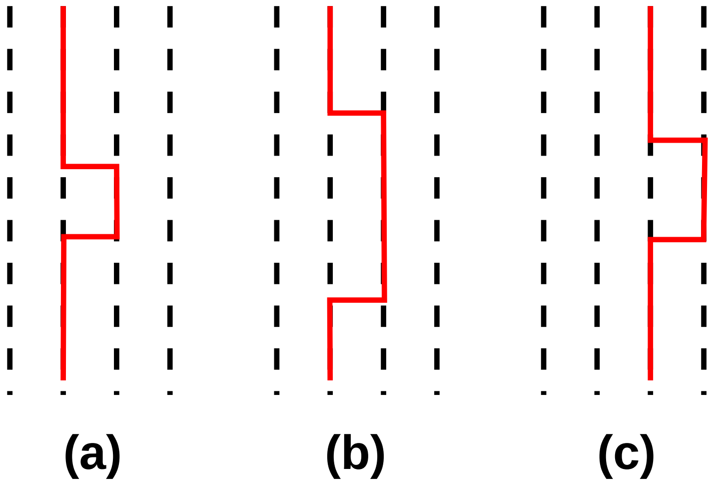


Figure 1.10: Kink movement causing dislocation glide. Initially a double kink causes the dislocation to move into an adjacent Peierls valley (a), the kinks separate causing more of the dislocation to cross the potential (b), finally the process continues with a new thermally-activated double kink (c). Peierls valleys are indicated with dashed lines.

1.3 Brown Colouration in Diamond

Diamond occurs in a wide range of colours, as well as the classic colourless type. Yellow stones are the most common of the 'fancy' diamonds, with blue, green, pink and red being less common. Brown coloured diamonds make up the majority of those mined [30], although compared to the other coloured diamonds they are rarely found in jewellery. Those stones that do make their way to the commercial market are often yellow-brown rather than a deep brown colour.

The origin of colour in diamond is known in some cases. Yellow diamonds contain nitrogen in various defects, with substitutional nitrogen (C-centre) causing colour on its own, and the A-centre, B-centre and NV defect combining to produce yellow hues. Blue colour is caused by substitutional boron defects, hence all blue diamonds are type IIb and p-type semiconducting. Green diamonds have been irradiated to produce single vacancy defects which have a zero phonon line (ZPL) at 1.67 eV [22], imparting the colour.

In the cases of red, pink and brown diamond the source of the colour is not known. All three colours occur in both type I and II diamond, increasing the difficulty of linking them to a specific impurity or defect. Due to its lack of commercial value, less research has been conducted on brown diamond compared to other colours, so little has been known until relatively recently. The catalyst for renewed interest in the origin of brown colour is the discovery that by subjecting a diamond to High-Pressure High-Temperature (HPHT) annealing treatment the colour can be quickly removed [26, 30, 31, 32, 33]. The result is a stone much increased in commercial value, either colourless or with a yellow colour, which is also difficult to differentiate from an untreated diamond. It is therefore desirable to understand both the origin of the brown colour and the process which occurs to remove it when undergoing treatment. The removal of the colour is illustrated in figure 1.11, which is taken from [34]. In their work a sample was grown by CVD, and then sectioned into four parts. One part was left untreated and the other three subjected to HPHT annealing of 1900° C for 1 hour, 2200° C for 1 hour and 2200° C for 10 hours. With the lower temperature anneal there is a strong yellow tinge from incorporated nitrogen, but this is removed at higher temperature.

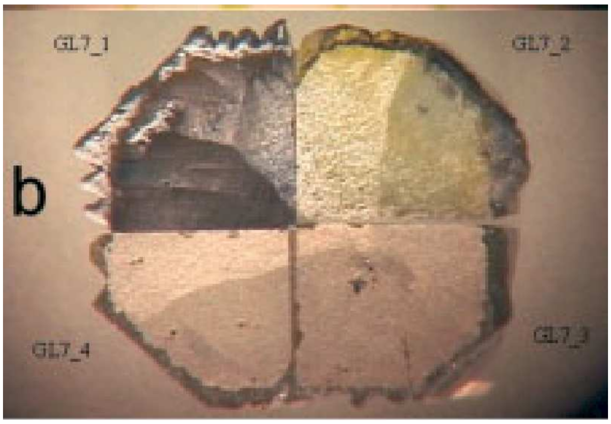


Figure 1.11: The removal of brown colour by HPHT treatment, taken from Charles *et al.* [34]. Proceeding clockwise from the top left are: (a) The as-grown diamond (b) Treated at 1900° C (c) Treated at 2200° C and (d) Treated at 2200° C for a longer time.

Two main factors have influenced the previously proposed origins of brown colour. Firstly, all brown diamonds have been subjected to plastic deformation during their formation [30], although not all plastically deformed diamonds are brown. It has been suggested that this deformation causes a large number of dislocations to be formed in the diamond, some of which are known to be optically active [28]. These dislocations could then be transformed into more energetically stable types which are optically inactive [29], with the energy provided by HPHT treatment. Such a transformation will release large numbers of vacancies or interstitials, which may or may not annihilate depending on other defects or impurities present [30].

The second area of interest is the presence of vacancies and vacancy-related defects in diamond. Positron Annihilation Spectroscopy (PAS) and Electron Paramagnetic Resonance (EPR) studies (described in sections 3.6 and 3.7) have both been undertaken to discover the nature of these defects. Two distinct PAS lifetimes are seen in brown diamonds, $\tau_2 \sim 130$ ps and $\tau_3 \sim 430$ ps [35, 36]. The shorter lifetime is also seen in γ-irradiated diamonds [37], but is not that of the monovacancy which is known to have a lifetime of 142 ps [38]. It is hypothesised that τ_2 is similar in volume to the monovacancy thus could be a modified vacancy structure, such as a vacancy-interstitial complex [37]. Modelling has predicted that, in a spherical arrangement, approximately 50 vacancies would be required to produce a signal of the order of 400 ps [36]. When subjected to annealing the τ_3 lifetime drops to ~ 350 ps, which would correspond to approximately 30 vacancies following the same spherical model.

EPR studies have focused on the small vacancy chains first proposed by Lomer *et al.* [39]. A range of centres denoted R5 – R12 and KUL11 – KUL15 (not necessarily inclusive) have been identified in natural [40] and implanted [41] diamond. The majority of these centres have C_{2v} or monoclinic-I symmetry which is compatible with a chain of vacancies lying in the $\langle 110 \rangle$ direction. In addition, the value of the spin-spin interaction matrix **D** allows the distance between unpaired spins, *i.e.* at the ends of a chain, to be calculated. The result is the assignment of several of these centres to $\langle 110 \rangle$ vacancy chains of 3 – 7 vacancies in size [40]. These centres are not necessarily linked directly to brown colouration, as the samples studied were not themselves brown, but are further evidence of the presence of a number of vacancy structures in diamond.

The absorption spectrum of brown diamond differs between type IIa and type Ia. Generally there is a continuum absorption from 0 eV until the indirect gap transitions start to take effect at $\sim$5.5 eV. This is unlike the typical absorption from a point defect, which often displays a distinct onset followed by vibronic bands *e.g.* N_s [42]. In type IIa the absorption is completely featureless with no onset, as shown on figure 1.12. This figure only displays visible-range and UV absorption, with figure 1.13 demonstrating the continuation of the absorption to low energies.

The absorption in the IR region ($<$ 1 eV) is continuous, but not featureless. There is a distinct peak around 0.25 eV (4960 nm) for all five samples, the origin of which is not known. It is clear that for a defect to be responsible for the brown colour in these samples it must induce into the diamond bandgap a range of levels, with allowed transitions from

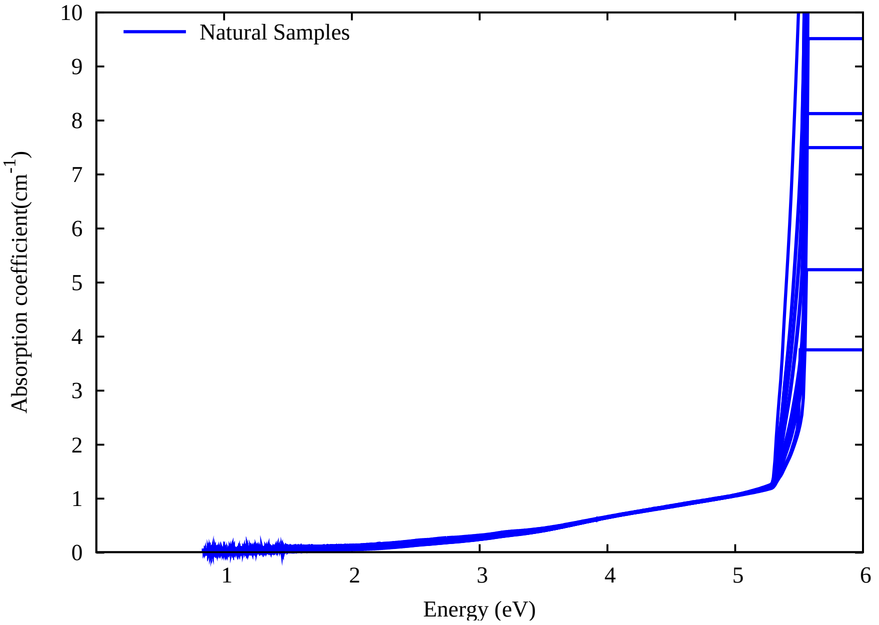


Figure 1.12: Visible and UV absorption of 15 natural type IIa brown diamonds, measured by the Diamond Trading Company (DTC).

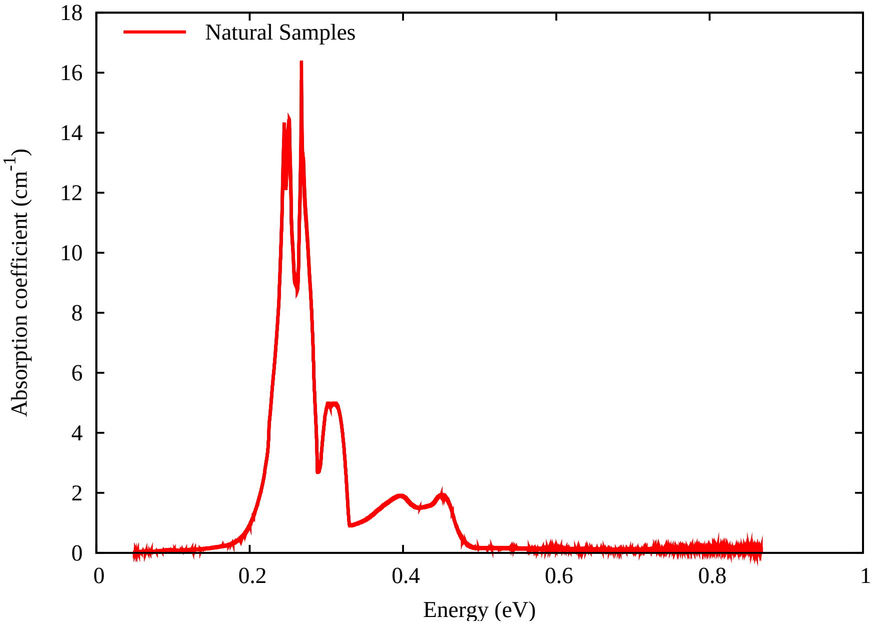


Figure 1.13: NIR and IR absorption of five type IIa diamonds, measured by DTC. These five are included in figure 1.12.

$\sim$0 eV up to the indirect gap of 5.5 eV.

CVD diamond often has nitrogen incorporated into the lattice, due to the growth conditions. This means that although nominally type IIa (nitrogen concentration is low overall) nitrogen features can sometimes be seen. Figure 1.14 shows absorption data from two different sets of CVD stones. In the lighter stones (red lines) the absorption profile is similar in magnitude and shape to the natural IIa diamonds of figure 1.12. However, in the darker stones (blue lines) the absorption magnitude is ∼5–10 times larger and there are distinct peaks/broad bands at 4.55 eV (272 nm) and 2.4 eV (516 nm). The 4.55 eV peak has been proposed as a nitrogen-boron complex [42], which is possible even if no borane is added to the growth gas as boron is near-impossible to remove from a system if it has ever been introduced. The 2.4 eV peak is very close to the 550 nm (2.25 eV) peak which is observed in type Ia brown diamond, and is quite conceivably the same defect.

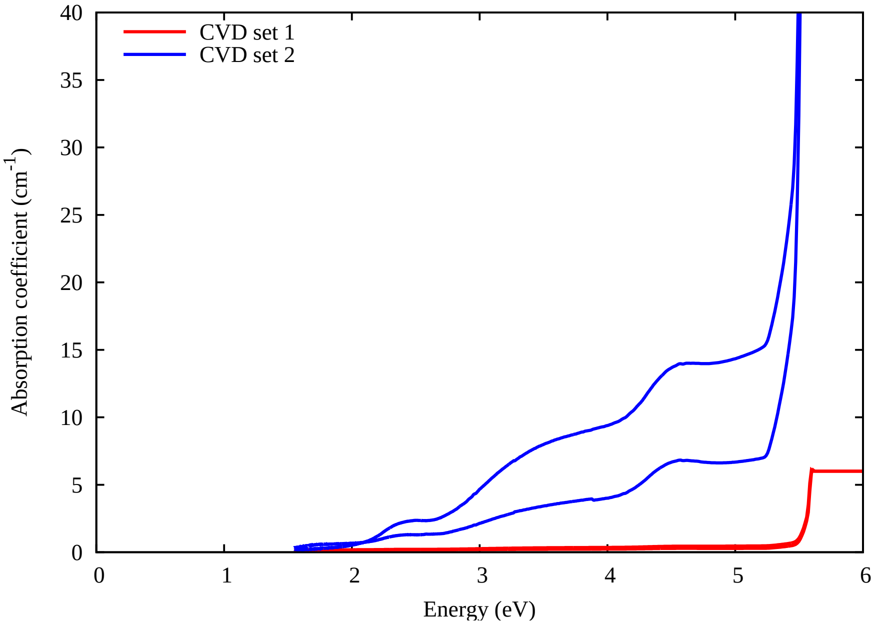


Figure 1.14: Vis/UV absorption of four type IIa CVD grown diamonds, measured by DTC.

A darker brown colour is associated with greater nitrogen inclusion and faster growth. This greater speed of growth may lead to disordered regions and defects being trapped before they can reform or dissipate, which can influence the colour. During CVD growth nitrogen is most often included as a substitutional defect. This is quite different from natural type Ia diamond, in which nitrogen is mainly found in aggregated defects. The lowest energy structure for nitrogen in diamond is predicted to be the B complex of four nitrogen atoms surrounding a vacancy [43].

Brown type Ia diamond displays a different absorption profile from brown type IIa diamond. The main continuum of absorption exists down to low energies with no distinct onset, however there are additional absorption features displayed. There is a peak at 2.25 eV (550 nm) [30] which is correlated with the A-centre concentration [44]. Brown type Ia diamonds sometimes display a band at 3.18 eV (390 nm), the cause of which is undetermined. Additionally, some brown diamonds display 'amber centres' which have an associated absorption peak at 0.52 eV (2384 nm) [45].

Figure 1.15 shows three type Ia brown diamonds measured by F. De Weerdt at Hoge Raad Voor Diamant (Diamond High Council), Antwerp. These samples have an absorption magnitude similar to the natural type IIa brown diamonds, but clearly display the 3.18 eV (390 nm) band. In addition, the absorption magnitude increases strongly from ∼4 eV towards the bandgap onset, this is unlike the type IIa absorption which remains low until near 5 eV, then increases very rapidly.

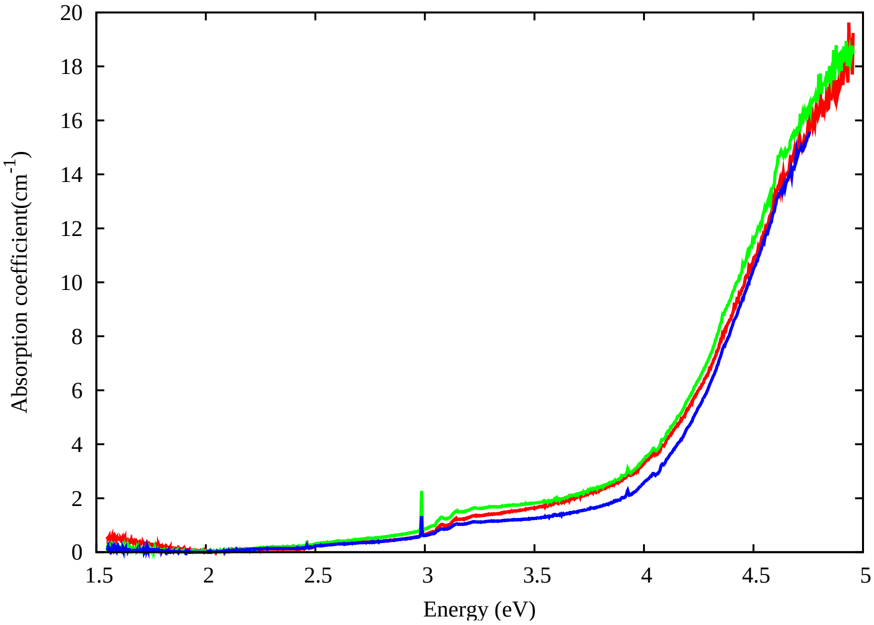


Figure 1.15: Vis/UV absorption of three type Ia brown natural diamonds, measured by F. De Weerdt.

There are significant differences between the annealing characteristics of the different kinds of diamond. In irradiated material [39] the absorption continuum disappears at 1400° C, and in the range 1400-1600°C for brown single-crystal CVD material [33]. For brown natural type IIa diamond however, temperatures above 2200°C are required to remove the absorption continuum [32]. This suggests that either the stabilities of the defects responsible for the absorption continuum in brown single-crystal CVD material and natural diamond are significantly different, or the mechanism for colour loss is different.

It is prudent to mention here the absorption profile of pink diamonds, as they bear a number of similarities to type Ia brown diamond. Pink diamonds can occur in all types, but in the vast majority of stones the pink colour has the same characteristics. The pink colour is arranged in striations throughout the crystal and has a distinct spectrum, which contains both the 2.25 eV (550 nm) and 3.18 eV (390 nm) peaks [44, 46, 47, 48]. This spectrum is shown in figure 1.16, with the infra-red and visible/ultra-violet spectra split. The split arises as two different spectrometers were used to take measurements, and they were not calibrated to the same levels. There is a photochromic link between the absorption bands in pink diamond. Illumination with light of more than $\sim$2.6 eV ($<$475 nm) will diminish the absorption bands concurrently by several cm^{-1}, while illumination at less than 2.3 eV ($>$530 nm) will restore the bands and subsequently enhance them. This behaviour suggests that the bands are associated with the same centre, perhaps as excited states. The energy dependence suggests a barrier of $\sim$2.25 eV for excitation [48].

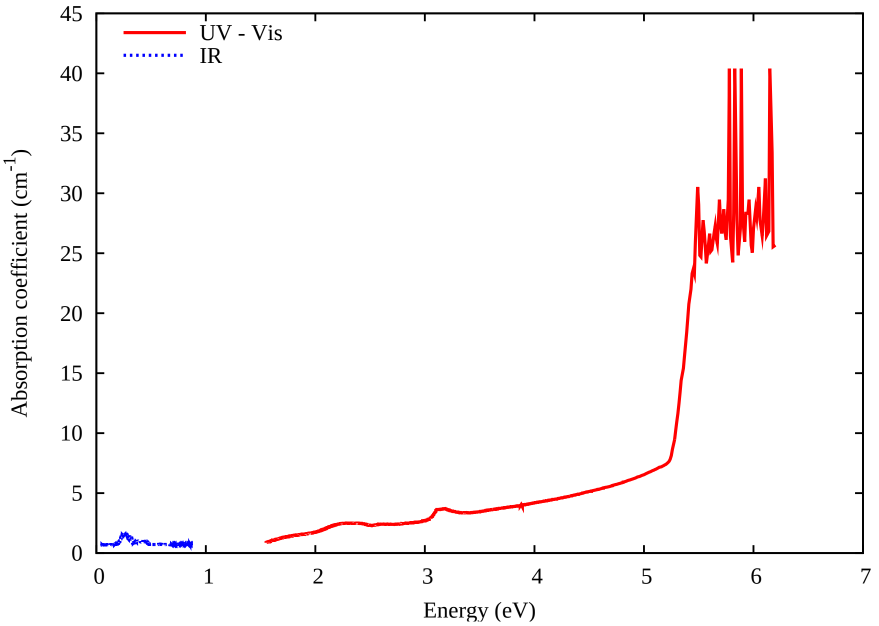


Figure 1.16: IR/Vis/UV absorption of a type IIa pink natural diamond, measured by DTC.

The overall absorption magnitude and shape is close to the brown diamond type IIa absorption, being only modified by the aforementioned peaks. Many of the fancy colours of diamond, especially yellow, can have a brown tinge, so it may be the case that the DTC sample is a dark pink approaching brown. The similarity between the pink diamond spectral defects and the type Ia brown spectrum suggests that the same or related defects may be responsible for both colours, perhaps the concentration or density of defects being the difference between the two cases.

1.4 Summary

Diamond is a material with several extreme values. Its crystal structure gives it the largest value of hardness in a natural material, and it has a larger thermal conductivity than many metals, mainly due to very strong covalent bonding between atoms. The electronic structure is governed by the 5.5 eV indirect bandgap which renders diamond transparent to visible light, allowing the high refractive index to impart diamond's characteristic sparkle. The large bandgap means that diamond needs to be doped to become electrically useful, which is an area of current research; although p-type semiconducting diamond does occur naturally.

Classification of diamond is based on defect content and configuration. The largest split is between diamond that contains significant amounts of nitrogen (type I) and that which has little or no nitrogen (type II). Type I diamond is further split into several subtypes depending on the configuration of nitrogenous defects within the structure, with type II diamond split into two groups; boron-containing stones and non-boron-containing stones.

Diamond contains a variety of defects, both point-like (small) and extended (large). Point defects are generally related to vacancies and impurities, with nitrogen being important due to its prevalence. Extended defects are most often dislocations; a region where the regular crystal structure is disrupted along a line. These defects can alter the electronic properties of the diamond to *e.g.* make it a useful semiconductor, or modify its colour. Some type I diamonds exhibit a yellow tinge caused by the presence of nitrogen defects.

The colour of diamond varies from colourless to the 'fancy' colours of yellow, blue, green, pink and red. Brown diamonds are common but few find their way to the jewellery market, for this reason little investigation on the nature of brown colour in diamonds has been undertaken. Developments in the HPHT treatment of brown diamond to render it colourless or fancy-coloured have spurred interest in brown diamond and its properties. The absorption spectrum of all types of brown diamond is broad and featureless, with a continuum down to near-zero energies. Type Ia brown diamond does have additional absorption features however, which may be linked to the nitrogen content. Several studies have identified key areas of investigation for brown diamond. Vacancies are present, according to PAS, possibly in large clusters, the lifetime of which are modified by annealing. Dislocations are prevalent in all natural diamonds, and may introduce absorption states in the bandgap, but are absent in CVD material which can also be brown. Whatever defect is responsible must be stable to high energies as evidenced by the temperatures required for removal of brown colour; 1400–1600° C for CVD material and $\sim$2000° C for natural diamond.

Chapter 2

Theory of Modelling Crystals

This chapter will describe the techniques used to study the properties of materials in the context of this work. It will address the topic in terms of conceptual ideas, mathematical framework and practical application. General properties of a cluster of atoms are assessed using the techniques of Density Functional Theory (DFT) and Density Functional Tight-Binding Theory (DFTB). Additionally, techniques to calculate specific properties such as vibrational modes or optical absorption are discussed.

2.1 Fundamental Concepts

A crystal can be considered as a collection of a large number of electrons and nuclei, of the order 10^{23} cm^{-3}. The energy of the crystal can be evaluated by solution of the Schrödinger equation:

$$\hat{\mathscr{H}}\Psi(\mathscr{R},r) = E\Psi(\mathscr{R},r) \tag{2.1}$$

With $\hat{\mathscr{H}}$ the Hamiltonian and $\mathscr{R}$ and *r* the collections of ions at $\mathbf{R_i}$ and electrons at $\mathbf{r_k}$ respectively. Here also $\mathbf{r_k}$ denotes both the electron coordinates and the spin. The Hamiltonian can then be expressed:

$$\hat{\mathscr{H}} = \hat{\mathscr{T}}_e + \hat{\mathscr{T}}_n + \hat{\mathscr{U}}_{e-e} + \hat{\mathscr{U}}_{e-n} + \hat{\mathscr{U}}_{n-n} \tag{2.2}$$

Where $\hat{\mathscr{T}}_e$ and $\hat{\mathscr{T}}_n$ are the kinetic energy operators for the electron and nuclei respectively. $\hat{\mathscr{U}}_{e-e}$, $\hat{\mathscr{U}}_{n-n}$ and $\hat{\mathscr{U}}_{e-n}$ yield the energy due to the electron only interactions, nuclei only interactions and the electron–nuclei interactions. These terms can be written more completely thus:

$$\hat{\mathscr{T}}_e = -\frac{\hbar^2}{2m}\sum_{k}^{N}\nabla_k^2 \tag{2.3}$$

with $\nabla_k = \nabla(\mathbf{r_k})$, for the k-th electron.

$$\hat{\mathscr{T}}_n = -\frac{\hbar^2}{2M}\sum_{i}^{N}M_i\nabla_i^2 \tag{2.4}$$

where M_i is the mass of the i-th nucleus and ∇_i follows as before. The interaction terms are given:

$$\hat{\mathscr{U}}_{e-e} = \frac{1}{4\pi\varepsilon_0}\frac{1}{2}\sum_{kl}{}' \frac{e^2}{|\mathbf{r_k} - \mathbf{r_l}|} \tag{2.5}$$

$$\hat{\mathscr{U}}_{n-n} = \frac{1}{4\pi\varepsilon_0}\frac{1}{2}\sum_{ij}{}' \frac{Z_i Z_j e^2}{|\mathbf{R_i} - \mathbf{R_j}|} \tag{2.6}$$

and

$$\hat{\mathscr{U}}_{e-n} = \frac{1}{4\pi\varepsilon_0}\sum_{kj}\frac{(-Z_j e^2)}{|\mathbf{r_k} - \mathbf{R_j}|} \tag{2.7}$$

In theory, knowledge of these terms allows a solution of the many–particle wavefunction of a crystal.

2.2 First Approximations

Solution of equation 2.1 using 2.2 is computationally impossible for the number of nuclei and electrons in a real crystal, therefore a range of approximations are applied to render the problem soluble.

Translational symmetry

Considering a crystal as a repetition of identical parts, Born-von Karman [49] periodic boundary conditions are applied and the wavefunction ψ at a point can then be described:

$$\psi(\mathbf{r}+N_i\mathbf{a}_i) = \psi(\mathbf{r}) \tag{2.8}$$

With N_i a set of integers (1,2,3) and $\mathbf{a}_i$ the related translation vectors in three-dimensional space. The potential of a crystal with translational periodicity can now be expressed:

$$U(\mathbf{r}+\mathbf{T}) = U(\mathbf{r}) \tag{2.9}$$

Where $\mathbf{T}$ is the direct space translation vector. Bloch's theorem [49] now states that the wave function of a particle with wave-vector **k** in such a crystal can be expressed:

$$\psi(\mathbf{r}+\mathbf{T}) = \mathrm{e}^{(\mathbf{k}\cdot\mathbf{T})}\psi(\mathbf{r}) \tag{2.10}$$

The system of atoms to be evaluated has now been reduced by $\sim 10^{23}$ orders of magnitude; the primitive unit cell of the system is now considered. However, even for just a few atoms the equations are still difficult to solve and further reductions in complexity must be found.

Adiabatic approximation

The large difference in mass of the nuclei and the electrons allows the motion of electrons to be described as reacting instantaneously to any motion of the nuclei; in effect the ionic nuclei are fixed to a particular configuration. The Born–Oppenheimer approximation [50] proposes that the electron wavefunction can be modulated by a function which is only dependent on the nuclei:

$$\eta(\mathscr{R},r) = \chi(\mathscr{R})\Psi(\mathscr{R},r) \tag{2.11}$$

where $\eta(\mathscr{R},r)$ is the total wavefunction, $\chi(\mathscr{R})$ is the ionic wavefunction and $\Psi(\mathscr{R},r)$ is the electronic wavefunction dependent on $\mathscr{R}$ in a continuous manner. This means that $\nabla\Psi(\mathscr{R},r)$ and $\nabla^2\Psi(\mathscr{R},r)$ can be assumed insignificant compared with $\nabla^2\chi(\mathscr{R})$, allowing the kinetic energy operator for the nuclei to be written:

$$\hat{\mathscr{T}}_n[\chi(\mathscr{R})\Psi(\mathscr{R},r)] \simeq \Psi(\mathscr{R},r)\hat{\mathscr{T}}_n\chi(\mathscr{R}) \tag{2.12}$$

Thus, $\hat{\mathscr{H}}\eta = E\eta$ can be expressed:

$$[\hat{\mathscr{T}}_n + \hat{\mathscr{U}}_{n-n} + E_e(\mathscr{R})]\chi(\mathscr{R}) = E_n\chi(\mathscr{R}) \tag{2.13}$$

the solution of $E_e(\mathscr{R})$ being provided by:

$$[\hat{\mathscr{T}}_e + \hat{\mathscr{U}}_{e-e} + \hat{\mathscr{U}}_{e-n}]\Psi(\mathscr{R},r) = E_e(\mathscr{R})\Psi(\mathscr{R},r) \tag{2.14}$$

To gain a solution for the nuclear eigenvalue equation 2.12, the electronic eigenvalues can be used as part of the potential energy:

$$\hat{\mathscr{H}}_e\Psi(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_N) = E_e\Psi(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_N) \tag{2.15}$$

for N electrons. This equation is still highly complicated if N is greater than 2, so further approximations are required, to reduce equation 2.14 to single–particle eigenvalue equations.

2.3 Wavefunction Theories

It can be recalled that $\Psi \equiv \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N)$ where $\mathbf{r}_k$ denotes both the spin and the position of the *k*–th electron. This cannot be solved exactly so described here are two methods to provide approximate solutions. The following must first be considered. The variational principle states that the energy of an approximate wavefunction Ψ for a system is never less than the ground state energy given by the true system wavefunction Ψ_0:

$$E = \left[\frac{\langle\Psi|\hat{\mathscr{H}}|\Psi\rangle}{\langle\Psi|\Psi\rangle}\right] \geq E_0 = \left[\frac{\langle\Psi_0|\hat{\mathscr{H}}|\Psi_0\rangle}{\langle\Psi_0|\Psi_0\rangle}\right] \tag{2.16}$$

This principle allows the quality of the approximated wavefunction to be evaluated by simply comparing energies, with the lowest energy expectation value denoting the closest match to the true system wavefunction. With regard to the techniques described below, the variational principle can be written:

$$\delta\left[\frac{\langle\Psi|\hat{\mathscr{H}}|\Psi\rangle}{\langle\Psi|\Psi\rangle}\right] = 0 \tag{2.17}$$

hence:

$$\delta[\langle\Psi|\hat{\mathscr{H}}|\Psi\rangle - \lambda\langle\Psi|\Psi\rangle] = 0 \tag{2.18}$$

where λ is a Lagrange multiplier.

2.3.1 Hartree's method

The previously described Schrödinger equation contains a term depending on the interaction of pairs of electrons, $\hat{\mathscr{U}}_{e-e}$. It is not possible to separate the variables in a second-order differential equation, so Hartree's method applies the approximation that each electron interacts with an averaged electron density, rather than individually with other electrons.

The electronic wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N)$ for a N–electron system can be expressed as a product of N orthonormal single–electron wavefunctions:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2), \ldots \psi_N(\mathbf{r}_N) = \Pi_i \psi_k(\mathbf{r}_k) \tag{2.19}$$

Then, $\langle\Psi|\Psi\rangle$ and $\langle\Psi|\hat{\mathscr{H}}|\Psi\rangle$ can be set:

$$\langle\Psi|\Psi\rangle = \int d^3r \int d^3r' \Psi^*(r')\Psi(r) = \prod_k \int d^3r_k \psi_k^*(\mathbf{r}_k)\psi_k(\mathbf{r}_k) \tag{2.20}$$

and:

$$\begin{aligned}\langle\Psi|\hat{\mathscr{H}}|\Psi\rangle &= \prod_k \int d^3r_k \psi_k^*(\mathbf{r}_k)\left[\frac{-\hbar^2}{2m}\nabla_k^2 + \mathscr{U}_{e-n}(\mathbf{r}_k)\right]\psi_k(\mathbf{r}_k) \\ &+ \frac{1}{4\pi\varepsilon_0}\frac{1}{2}\prod_{kl}{}' \int d^3_{r_k} \int d^3_{r_l} \frac{e^2}{|\mathbf{r}_k - \mathbf{r}_l|}|\psi_k(r_k)|^2|\psi_l(r_l)|^2\end{aligned} \tag{2.21}$$

The variational principle then reads:

$$\delta[\langle\Psi|\hat{\mathscr{H}}|\Psi\rangle - \varepsilon_k^H\langle\Psi|\Psi\rangle] = 0 \tag{2.22}$$

where $\varepsilon_k^H = \lambda$ *i.e.* a Lagrange multiplier, and ε_k^H are Hartree eigenvalues. The Hartree Hamiltonian can then be written:

$$\hat{\mathscr{H}}_H \psi_k(\mathbf{r}) = \varepsilon_k^H \psi_k \tag{2.23}$$

with:

$$\hat{\mathscr{H}}_H = \hat{\mathscr{T}}_e + \mathscr{U}_{e-n} + \mathscr{U}_H \tag{2.24}$$

explicitly:

$$\hat{\mathscr{H}}_H = \frac{-\hbar^2}{2m}\nabla^2 - \frac{1}{4\pi\varepsilon_0}\sum_{\mathbf{R}_j}\frac{Z_j e^2}{|\mathbf{r}-\mathbf{R}_j|} - \frac{e}{4\pi\varepsilon_0}\int d^3r^{'}\,\frac{\rho(\mathbf{r}^{'})}{|\mathbf{r}-\mathbf{r}^{'}|} \tag{2.25}$$

where $\mathbf{R}_j$ is an ionic position and $\rho(\mathbf{r}^{'})$ is the electronic charge density at $\mathbf{r}^{'}$. Which is given:

$$\rho(\mathbf{r}^{'}) = \sum_k \rho_k(\mathbf{r}^{'}) = -e\sum_k |\psi_k(\mathbf{r}^{'})|^2 \tag{2.26}$$

To use Hartree's method in practise, some approximate electron orbitals ψ_k are taken and used to solve all N equations (where N is still the number of electrons). A set of new orbitals $\psi_k^{'}$ are obtained, giving a better approximation for the orbitals which can be used as a new starting point for the process. When the $\psi_k^{'}$'s do not change, it is said that the *self-consistent field orbitals* have been obtained. The expectation value of energy can then be found from the Hamiltonian given in equation 2.24.

The starting wavefunction for the Hartree method is unphysical. This results in a final wavefunction which in not anti-symmetric. That is to say, interchanging the electron labels in the wavefunctions does not necessarily change the sign for the electrons. Without this property the function will not correctly describe the system, in fact in all fermion systems no two particles can be described by the same one-particle wavefunction. Thus the Hartree method is flawed in respect to this, the solution being given in the Hartree–Fock method.

2.3.2 Hartree–Fock theory

The solution to the problems of the Hartree method were devised independently by Fock [51, 52] and Slater [53]. One-electron functions are still used, but the total wavefunction for the system is an antisymmetrized sum of all the products which can be obtained by the interchange of electron labels. The antisymmetric Hartree–Fock wavefunction can be expressed as a Slater determinant:

$$\Psi_{HF}(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(\mathbf{r}_1) & \cdots & \psi_N(\mathbf{r}_1) \\ \vdots & \ddots & \vdots \\ \psi_1(\mathbf{r}_N) & \cdots & \psi_N(\mathbf{r}_N) \end{vmatrix} \tag{2.27}$$

Where, as before, $\mathbf{r}_i$ represents both space and spin. The Slater determinant is always antisymmetric as exchanging two single electron spin-orbitals changes $\Psi(r)$ by a factor of -1. Note also that if two identical orbitals are present $\Psi(r) = 0$.

The expectation value of the Hamiltonian then becomes:

$$\begin{aligned} \langle\Psi|\hat{\mathscr{H}}|\Psi\rangle &= \prod_k \int d^3r\psi_k^*(\mathbf{r})\left[\frac{-\hbar^2}{2m}\nabla^2 + \mathscr{U}_{e-n}(\mathbf{r})\right]\psi_k(\mathbf{r}) \\ &+ \frac{1}{4\pi\varepsilon_0}\frac{1}{2}\prod_{kl}\int d^3r\int d^3r'\,\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}|\psi_k(\mathbf{r})|^2|\psi_l(\mathbf{r})|^2 \\ &- \frac{1}{4\pi\varepsilon_0}\frac{1}{2}\prod_{kl}\int d^3r\int d^3r'\,\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}\psi_k^*(\mathbf{r})\psi_k(\mathbf{r}')\psi_l^*(\mathbf{r}')\psi_l(\mathbf{r}) \end{aligned} \tag{2.28}$$

where the second term of equation 2.28 is an electron repulsion term and the third term is an exchange interaction, that is, an *attractive* interaction between an electron at **r** and all other electrons with the same spin.

Hartree–Fock theory can be approached by a numerical technique, but this method is computationally intense. It is more usually approached by representing the electron orbitals ψ_k as a linear expansion of a set of basis functions χ, which can be expressed:

$$\psi_k(\mathbf{r}) = \sum_1^n C_{nk}\chi_k(\mathbf{r}) \tag{2.29}$$

The coefficients C_{nk} are optimized rather than the basis functions.

Application of the variational principle:

$$\frac{\delta}{\delta\psi_k^*(\mathbf{r}')}[\langle\Psi|\hat{\mathscr{H}}|\Psi\rangle - \varepsilon_k\langle\Psi|\Psi\rangle] = 0 \tag{2.30}$$

leads to the Hartree-Fock equations:

$$\hat{\mathscr{H}}_{HF}\,\psi_k(\mathbf{r}) = \varepsilon_k^{HF}\,\psi_k(\mathbf{r}) \tag{2.31}$$

Explicitly:

$$\hat{\mathscr{H}}_{HF}\,\psi_k(\mathbf{r}) = \hat{\mathscr{H}}_H\,\psi_k(\mathbf{r}) - \frac{1}{4\pi\varepsilon_0}\sum_l \int d^3r'\,\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}\psi_l(\mathbf{r})\psi_l^*(\mathbf{r}')\psi_k(\mathbf{r}) \tag{2.32}$$

where $\hat{\mathscr{H}}_H$ is the Hartree Hamiltonian. The last term in equation 2.32 is an attractive exchange operator, the form of which is generally intractable. Local exchange forms have been sought to make the equations soluble [54, 55, 56].

If the ψ_k's in equation 2.31 are now expanded into their basis functions the Hamiltonian becomes:

$$\hat{\mathscr{H}}_{HF}\sum_1^n C_{nk}\chi_k(\mathbf{r}) = \varepsilon_k^{HF}\sum_1^n C_{nk}\chi_k(\mathbf{r}) \tag{2.33}$$

Multiplying by $\chi_l(\mathbf{r})$ and integrating over **r**:

$$\sum_1^n C_{nk}\int \chi_l^*(\mathbf{r})\hat{\mathscr{H}}_{HF}\,\chi_k(\mathbf{r})d(\mathbf{r}) = \varepsilon_k\sum_1^n C_{nk}\int \chi_l^*(\mathbf{r})\chi_k(\mathbf{r})d(\mathbf{r}) \tag{2.34}$$

these integrals are now represented as:

$$(\mathbf{F})_{lk} = \int \chi_l^*(\mathbf{r})\hat{\mathscr{H}}_{HF}\chi_k(\mathbf{r})d(\mathbf{r}) \tag{2.35}$$

and:

$$(\mathbf{S})_{lk} = \int \chi_l^*(\mathbf{r})\chi_k(\mathbf{r})d(\mathbf{r}) \tag{2.36}$$

$(\mathbf{F})_{lk}$ is an element of an $n \times n$ Fock matrix **F**, and similarly $(\mathbf{S})_{lk}$ is an element of an $n \times n$ overlap matrix **S**. There are n such equations where n is the number of basis functions. Expressed in matrix form this is:

$$\mathbf{FC} = \mathbf{SC}\varepsilon \tag{2.37}$$

where ε is the expectation value for the energy. The matrix **C** which diagonalises **F** can be obtained by first finding a matrix **X** to diagonalise **S**:

$$\mathbf{X}^{\dagger}\mathbf{SX} = 1 \tag{2.38}$$

so that:

$$\mathbf{C}^{'\dagger}\mathbf{F}^{'}\mathbf{C}^{'} = \varepsilon \tag{2.39}$$

The elements of **F** is this case depend on the orbitals ψ_k, so must be solved through an iterative process. Additionally the Fock matrix elements contain a large number of two-electron integrals, which are complex to solve.

To try and overcome the difficulties the problem can be approached by describing the electron system using density rather than wavefunctions.

2.4 Density Functional Theories

Density functional schemes express the exchange-correlation potential not as a function of all electron orbitals, as in Hartree-Fock theory, but as a function only of the electron charge density at a point. The number of variables to compute is reduced from $3n$ – where n is the electron number – to just 3. The electronic charge density at a point $\mathbf{r}_1$ caused by all other electrons can be expressed:

$$\rho(\mathbf{r}_1) = \int d\mathbf{r}_2 \ldots \int d\mathbf{r}_N \Psi^*(\mathbf{r}_1 \ldots \mathbf{r}_N)\Psi(\mathbf{r}_1 \ldots \mathbf{r}_N) \tag{2.40}$$

2.4.1 The Thomas-Fermi-Dirac scheme

The Thomas-Fermi model [57, 58] assumed that the electrons move independently in an effective electrostatic potential determined by Poisson's equation. The kinetic energy is expressed in terms of the kinetic energy of a system of non-interacting electrons with slowly varying density. This model did not attempt to represent the exchange energy predicted by the Hartree-Fock technique, but this was later included by Dirac [59].

The neglect of electron correlation in the Thomas-Fermi-Dirac scheme, along with errors in the exchange energy and kinetic energy representation, cause this method to be too inaccurate for most situations.

2.4.2 The Hohenberg-Kohn-Sham scheme

The Hohenberg-Kohn-Sham scheme was formulated much later than the Thomas-Fermi-Dirac scheme. In essence the theory states that the Schrödinger equation in 2.13 can be reduced to a set of single-particle equations to be solved self-consistently. The Hohenberg-Kohn-Sham scheme is the method used in this work so will be described more fully here.

There are two basic theorems laid down for the Hohenberg-Kohn-Sham density functional theory, each of which are expressed here.

Density as the central variable

It was demonstrated by Hohenberg and Kohn that the external potential affecting a number of electrons N, is a function of the electronic density in the ground state – plus an additive constant [60]. Where the electronic density is:

$$n(\mathbf{r}) = N \int |\Psi(r')|^2 d^3 r' \tag{2.41}$$

Consequently, since the external potential fixes the Hamiltonian, the full many-particle ground state wavefunction is a function of $n(\mathbf{r})$. Then the function is defined:

$$F[n] = \langle \Psi | (\mathscr{T}_e + \mathscr{U}_{e-e}) | \Psi \rangle \tag{2.42}$$

The total energy of the system is given by:

$$E_e[\mathscr{U}_{e-n}, n] = F[n] + \int \mathscr{U}_{e-n}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \tag{2.43}$$

The potential $\mathscr{U}_{e-n}(\mathbf{r})$ is the external potential which defines the electron motion. $F[n]$ is an independent functional to account for electronic kinetic, electron correlation and exchange-correlation energies. The form of $F[n]$ is generally unknown.

Energy variational principle

Subject to the correct electronic density, $E_e[\mathscr{U}_{e-n}, n]$ is equal to the ground state energy E. The energy functional of $\Psi^{'}$, given by:

$$\varepsilon_v[\Psi^{'}] \equiv (\Psi^{'}, V\Psi^{'}) + (\Psi^{'}, (T+U)\Psi^{'}) \tag{2.44}$$

has a minimum at the correct ground state Ψ; relative to arbitrary variation of $\Psi^{'}$ where the number of particles N is kept constant. Applied to the ground state electronic energy:

$$E_e[\mathscr{U}_{ext},\rho] < E_e[\mathscr{U}_{ext},n] \tag{2.45}$$

with $n(\mathbf{r}) = \rho(\mathbf{r})$ as the correct ground state charge density.

Electronic energy functional

The kinetic energy is treated differently in the Hohenberg-Kohn-Sham theory compared to the Thomas-Fermi approach, in which it has only an approximate form. Kohn and Sham propose a different separation of the energy functional, so that the universal functional is expressed:

$$F[n] = T_0[n] + \frac{e^2}{8\pi\varepsilon_0}\int d\mathbf{r}\int d\mathbf{r}'\,\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}[n] \tag{2.46}$$

where $T_0[n]$ is the kinetic energy of non-interacting electrons with density $n(\mathbf{r})$. The integral terms are the classic mean-field Coulomb energy. $E_{xc}[n]$ is a functional of the density and is the many-body exchange-correlation energy functional; representing all corrections to the independent electron model. Using equations 2.43 and 2.46 the ground state energy of a many-electron system is:

$$E_e[\mathscr{U}_{ext},\rho] = T_0[\rho] + \int d\mathbf{r}\;\mathscr{U}_{ext}(\mathbf{r})\rho(r)\frac{e^2}{8\pi\varepsilon_0}\int d\mathbf{r}\int d\mathbf{r}'\,\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}[\rho] \tag{2.47}$$

Three major problems arise from evaluation of $E_e[\rho]$:

- The ground state density $\rho(\mathbf{r})$ must be evaluated self-consistently.
- $T_0[\rho]$ cannot be evaluated as there is no wavefunction information if only $\rho(\mathbf{r})$ is known.
- $E_{xc}[\rho]$ is only known for a few simple systems; such as a homogeneous electron gas. It must be represented in a simplified but accurate form.

The Kohn-Sham equations

Kohn and Sham's approach was to take the approach of Hohenberg and Kohn [60] and extend it to obtain a set of self-consistent equations including approximate exchange and correlation effects. These Kohn-Sham equations [61] require a knowledge of the chemical potential $\mu(\rho)$ of a homogeneous interacting electron gas as a function of the density ρ. It is defined as:

$$\mu = \frac{\delta T_0[\rho]}{\delta\rho(\mathbf{r})} + \mathcal{U}_{KS}[\rho] \tag{2.48}$$

where $\mathcal{U}_{KS}[\rho]$ is an effective Kohn-Sham potential of three parts, given as:

$$\mathcal{U}_{KS}[\rho] = \mathcal{U}_{ext}(\mathbf{r}) + \mathcal{U}_{H}(\rho) + \mathcal{U}_{xc}(\rho) \tag{2.49}$$

The Hartree potential is defined as:

$$\mathcal{U}_{H}(\rho) = \frac{e^2}{4\pi\varepsilon_0}\int d\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.50}$$

and the exchange-correlation potential is:

$$\mathcal{U}_{xc}(\rho) = \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})} \tag{2.51}$$

With these equations Kohn and Sham then showed that the total system is described by a set of one-electron Schrödinger equations:

$$\left[\frac{-\hbar^2}{2m}\nabla^2 + \frac{e^2}{4\pi\varepsilon_0}\int d\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})} + \mathcal{U}_{ext}(\mathbf{r})\right]\phi_j(\mathbf{r}) = \varepsilon_j\phi_j(\mathbf{r}) \tag{2.52}$$

where ϕ_j and ε_j are the single-particle wavefunctions and eigenvalues of a non-interacting system governed by the effective Kohn-Sham potential. The charge density is obtained by summing over the occupied states:

$$\rho(\mathbf{r}) = \sum_{j}^{occ} |\phi_j(\mathbf{r})|^2 \tag{2.53}$$

These equations 2.52 and 2.53 are the Kohn-Sham equations, which can be solved self-consistently to give the exact ground state charge density and total system energy to be obtained; the major strength of Density Functional Theory. The total energy is then:

$$E[n] = \sum_{j}^{occ} \varepsilon_j - \mathscr{U}_H(\rho) + E_{xc}(\rho) - \int d\mathbf{r}\ \mathscr{U}_{xc}(\rho)\rho(\mathbf{r}) \tag{2.54}$$

The Kohn-Sham eigenvalues are real but are not representative of the occupied electron energy states or excitation energies of the system; apart from the highest occupied Kohn-Sham eigenvalue which is equal to the physical chemical potential.

The Exchange-Correlation functional

The precise form of the exchange-correlation functional E_{xc} is unknown, but there are several approximations devised. The approximation suggested by Hohenberg and Kohn and Kohn and Sham, and used in this work, is the Local Density Approximation (LDA). It is assumed that the electronic charge density in the system corresponds to that of a homogeneous electron gas. An exact solution for this is known, if $\rho(\mathbf{r})$ is a sufficiently slowly varying quantity in space. The exchange-correlation energy under the LDA scheme can then be expressed:

$$E_{xc}^{LDA}[\rho] = \int \varepsilon_{xc}[\rho(\mathbf{r})]\rho(\mathbf{r})d(\mathbf{r}) \tag{2.55}$$

Here $\varepsilon_{xc}[\rho(\mathbf{r})]$ is the exchange-correlation energy per electron of a uniform electron gas of density $\rho(\mathbf{r})$. The exchange-correlation potential can then be written [62]:

$$V_{xc}^{LDA}(\mathbf{r}) = \frac{d}{d\rho}\varepsilon_{xc}(\rho(\mathbf{r}))\rho(\mathbf{r}) \sim \mu_{xc}(\rho(\mathbf{r})) \tag{2.56}$$

The exchange-correlation potential can further be developed by expressing $\mu_{xc}(\rho)$ in terms of the mean interelectronic spacing r_s, thus:

$$V_{xc}(\mathbf{r}) \sim \mu_{xc}[r_s] = \varepsilon_{xc} - \frac{r_s}{3}\frac{d\varepsilon_{xc}}{dr_s} \tag{2.57}$$

The requirement now is for an analytical expression of $\varepsilon_{xc}[r_s]$. There are several expressions devised by Kohn-Sham and G'asp'ar, Perdew and Zunger and Perdew and Wang among others [62].

The main shortfall of the LDA is to underestimate the bandgap, which in diamond is ~76 % of the true value. One reason for this is the LDA one-electron effective potential contains an incorrect interaction of each electron with itself [63]. It is necessary to consider this when estimating bandgap transitions, absorptions and emissions.

2.5 Application Using *AIMPRO*

In this work DFT calculations have been implemented using the **A**b **I**nitio **M**odelling **PRO**gram – AIMPRO [64]. To achieve satisfactory computational simplicity, further simplification of the equations are made – notably in the use of pseudopotentials and the basis set chosen. This is a valence electron approach to the modelling of solids, as is the tight-binding scheme; which is described later in this chapter.

2.5.1 Pseudopotentials

When solving the self-consistent Schrödinger equation not all electrons need to be considered; the interaction of any atom with other atoms and its general environment is mainly dependent on its valence electrons. The other electrons are closely bound to the ionic core and hence are known as the core electrons. The core electron potential can be considered inclusively with the ionic potential as one total potential acting upon the valence electrons. This is the pseudopotential, which has the benefit of reducing the number of functions required for all states – especially compared to the all-electron model – allowing larger systems to be studied.

There are two constraints to be applied when constructing pseudopotentials for a system [65]:

- The pseudopotential has exactly the same valence energy levels as the true atomic potentials; that is the s, p and d levels are the same. It follows that the lowest bound state solutions are the valence energy eigenvalues – there being no core levels.
- The pseudo-wavefunctions must be equal to the true wavefunctions outside a given core radius, which varies with atomic species.

Inside the core radius mentioned above, the pseudo-wavefunctions are not equal to the true valence wavefunctions, as these wavefunctions often vary rapidly. Instead, a mathematically straightforward function is chosen which still satisfies the above conditions. Throughout this work the pseudopotential of choice has been the Hartwigsen-Goedecker-Hutter (HGH) type [66].

2.5.2 Basis sets

The Kohn-Sham orbitals can be written as a product of a spin-function and a spatial function, which is expanded in terms of a basis set $\phi_i(\mathbf{r})$, so that:

$$\psi_k(\mathbf{r},s) = \chi_k(s)\sum_{1}^{n} c_{nk}\phi_n(\mathbf{r}) \tag{2.58}$$

where $\chi_k(s)$ is the spin-up wavefunction for $s = \frac{1}{2}$ and spin down for $s = -\frac{1}{2}$. AIMPRO uses Cartesian Gaussian orbitals for $\phi_i(\mathbf{r})$, the form of which depends on whether a full or contracted basis set is chosen.

In the contracted basis set 13 functions per atom are used, which consist of 2 s-type functions, 6 p-type functions and 5 d-type functions. The s and p functions are linear combinations of four Gaussians, and are pre-multiplied by x, y or z for p-type functions. The d-type functions are necessary for polarisation effects and consist of single Gaussian functions with premultipliers: $2z^2-x^2-y^2$, x^2-y^2, xy, xz and yz. The exponents of the underlying Gaussians and the contraction coefficients are optimized for each material under study.

The larger basis sets consist of at least 22 functions per atom, with the form:

$$\phi(x,y,z) = x^{n_x}y^{n_y}z^{n_z}e^{(-\alpha r^2)} \tag{2.59}$$

The s and p-type functions are constructed using three different α exponents, with $n_x + n_y + n_z \leq 1$. The d-type functions are formed using a further, different α and by setting $n_x + n_y + n_z \leq 2$. As before, the exponents are chosen to optimize the bulk energy of each particular system.

A more complex set of 28 functions can be formed by letting two of the four exponents have $n_x + n_y + n_z \leq 2$. This is sometimes necessary for calculation of optical observables as these often have a strong dependence on the d-type orbitals created in this basis set.

When the basis set has been constructed the equation to solve is reduced to:

$$H\sum_j \chi_\lambda(s)c_j^\lambda \phi_j(r) = E_\lambda \chi_\lambda(s)\sum_j c_j^\lambda \phi_j(r) \tag{2.60}$$

If this is multiplied through by $\chi_\lambda^*(s)\phi_i^*(r)$, integrated over r and then summed over s the equation is reduced to a set of matrix equations:

$$\sum_j (H_{ij} - E_\lambda S_{ij})c_j^\lambda = 0 \tag{2.61}$$

where H_{ij} and S_{ij} are the Hamiltonian and overlap matrix elements respectively. The output spin density can then be generated:

$$n_o(\mathbf{r},s) = \sum_{\lambda} |\chi_{\lambda}(s)\psi_{\lambda}(\mathbf{r})|^2 \tag{2.62}$$

This can be iterated to obtain self-consistency so that the self-consistent density is obtained:

$$n(\mathbf{r},s) = \sum_{k} c_k(s) g_k(\mathbf{r}) \tag{2.63}$$

The self-consistent density then allows the structural energy to be evaluated; and from this other properties and observables.

2.5.3 Brillouin zone sampling

To calculate physical properties of a material the charge density integrated over the Brillouin zone is required. To reduce complexity a set of special **k**-points can be taken, from which an average value for the intergrand function $f(\mathbf{k})$ is calculated. The average value is given by:

$$\bar{f} = \frac{\Omega}{(2\pi)^3} \int f(\mathbf{k}) d\mathbf{k} \approx \frac{1}{N_k} \sum_{i}^{N_k} f(\mathbf{k}_i) \tag{2.64}$$

where the volume of the Brillouin zone is given by $\frac{\Omega}{(2\pi)^3}$ and N_k is the number of special **k**-points.

This work has utilised the scheme of Monkhorst and Pack [67], where the special **k**-points are a grid of $I \times J \times K$ points in reciprocal space given by:

$$\mathbf{k}(i,j,k) = u_i \mathbf{g_1} + u_j \mathbf{g_2} + u_k \mathbf{g_3} \tag{2.65}$$

where $\mathbf{g_1}, \mathbf{g_2}$ and $\mathbf{g_3}$ are the reciprocal unit vectors and u_i, u_j and u_k are defined:

$$u_i = \frac{(2i - I - 1)}{2I}, \quad (i = 1, ...I) \tag{2.66}$$

$$u_j = \frac{(2j - J - 1)}{2J}, \quad (j = 1, ...J) \tag{2.67}$$

$$u_k = \frac{(2k - K - 1)}{2K}, \quad (k = 1, ...K) \tag{2.68}$$

It is one of the advantages of the scheme that energy convergence of the system can be ensured by increasing I,J and K to increase the accuracy of $f(\mathbf{k})$.

It is necessary to ensure that appropriate values of I,J and K are chosen for the system under study. There is an inverse relationship between the number of points chosen and the size of the cell in that direction. This is due to the sampling being taken across the Brillouin zone, in reciprocal space. A large real-space lattice is small in reciprocal space, requiring less sampling points.

Computational complexity can be reduced in high symmetry cells as certain **k**-points are equivalent. In this case redundant **k**-points are removed and a compensating weighting factor applied to the remaining equivalent **k**-point.

2.6 Calculation of Material Properties

This section will give overviews of several of the methods that are used in AIMPRO to calculate material properties. AIMPRO has the ability to calculate many properties, which is useful for comparison to experimental work. The methods discussed in this section are ones that have been used in this work.

The Kohn-Sham eigenstates are not electron-wavefunctions but a set of basis functions from which the charge density is expanded. The consequence of this is that DFT is a ground-state theory; it does not describe excited-state properties well.

2.6.1 Total energy

The force on an atom is given simply by:

$$\mathbf{F}_i = -\nabla_i E \tag{2.69}$$

It follows that the total energy of a system can be found by minimising the force on each atom, structurally optimizing the system. In this technique each atom is moved a small amount inducing an energy change δE and the new forces are then calculated. The atoms are then moved in accordance with these forces. This process can be repeated iteratively until the change in total energy and forces are deemed insignificant.

AIMPRO uses the Conjugate Gradient Algorithm (CGA) method to attempt system relaxation. In the case of M atoms in the cell, the force equation:

$$F[\rho, \{\tau\}] = 0 \tag{2.70}$$

can be solved for the equilibrium geometry $\{\tau_0\}$. The approach for solving this equation to to assume the charge density remains fixed while the atomic configuration τ is variable. In the case of an m^{th} iteration in relaxation, then the atomic configuration is $\{\tau^{(m)}\}$. A direction $\mathbf{d}^m$ is chosen, then a minimization in this direction undertaken; a solution t_0 is found to minimize:

$$E[\rho, \{\tau^{(m)} + \mathbf{d}^m t\}] \tag{2.71}$$

so that the geometry for the $(m+1)^{st}$ iteration can be set:

$$\{\tau^{(m+1)}\} = \{\tau^{(m)} + \mathbf{d}^m t_0\} \tag{2.72}$$

$E[\rho, \{\tau^{(m+1)}]$ and $F[\rho, \{\tau^{(m+1)}]$, the total energy and force can then be found. The process is iterated as described above.

The key aspects of the CGA method are the choice of direction **d** and the minimization to determine t.

Choice of d

If the initial choice for **d** of **d**$= F[\rho,\tau]$ is maintained for all relaxation integration, this is known as the steepest-descent technique. In the CGA method the $\{\mathbf{d}^{(m)}\}, m = 1, 2, \ldots, m$ directions are chosen to be 'conjugate', that is:

$$(\mathbf{d}^{(i)})^T C\mathbf{d}^{(j)} \propto \delta_{ij} \tag{2.73}$$

where $i < j$ and C is a symmetric matrix. The energy is related by:

$$E(\mathbf{x}) = a - b\mathbf{x} + \frac{1}{2\mathbf{x}^T C\mathbf{x}} + O(x^3) \tag{2.74}$$

with a and b constants of the energy quadratic, and **x** the distance along the energy curve.

Minimization

When the results for the energy $E(\tau)$ and the force $F(\tau)$ have been calculated for the m^{th} relaxation integration, the values $E(\tau+\sigma\mathbf{d})$ and $F(\tau+\sigma\mathbf{d})$ for a small increment σ can be found. These are fitted to a third-order (or greater) polynomial and an optimal value t_0 estimated to correspond to minimum energy in direction **d**.

The CGA approach is simple and effective but has one large factor to be taken into account. The minimum energy that is found for a particular starting atomic geometry is not guaranteed to be the global minimum. The atoms cannot be moved to escape from a local minimum, so the start positions must be chosen with care for complex structures, and ideally several start positions for optimization should be evaluated for the total energy.

2.6.2 Formation energy

Comparison of total energy when studying different defects can be misleading. It is often the case that there are different numbers of atoms in each cell, or that different elements are present. In this situation the formation energy is useful as it takes account of the number and types of atoms in the cell.

The chemical potential of a species s is represented by μ_s and is the derivative of the Gibbs free energy with respect to the number of atoms of that species. In the case of thermodynamic equilibrium μ_s is equal across the system. The Gibbs free energy is:

$$G = E + PV - TS \tag{2.75}$$

However, it is possible to neglect the second and third terms as PV is small for solid state calculations and the $-TS$ approximation is valid only at high temperatures. The formation energy is then given by:

$$E_f = E_T = q\mu_e - \sum_s n_s \mu_s \tag{2.76}$$

where E_T is the total energy of the system, q is the net positive charge on the system (with μ_e the fundamental electron charge) and n_s and μ_s are the number and chemical potential species of a species s. μ_e is usually taken as $E_F + E_v$, where E_F is the Fermi level and E_v is the highest occupied energy level.

The formation energy can be used to compare energy convergence of a defect in differently sized supercells, or to check accuracy with different basis sets or pseudopotentials. In addition the relative stability of defects can be compared, for example dislocations that contain different numbers of atoms, or varying sizes of vacancy defects. The possible transformation of defects at different temperatures can be inferred by looking at energies required to dissolve defects; for example by vacancy/interstitial migration.

2.6.3 Vibrational modes

Calculation of a defect's vibrational modes is highly important in relation to experimental results as they can be compared to infra-red, photoluminescent measurements and Raman spectra. As the vibrational modes are dependent on the structure of a defect and the bonding of the atoms an agreement between calculated frequencies and those observed in experiment is a strong indicator that the proposed defect structure is close to the real situation.

The vibrational modes for a system of N atoms are given by solution of the eigenvalue problem:

$$\mathbf{D} \cdot \mathbf{U} = \omega^2 \mathbf{U} \tag{2.77}$$

Where **D** is a dynamical matrix and the eigenvalues ω^2 are the squared frequencies associated with the normal modes **U**. The normal mode is a $3N$ dimensional vector describing the motion of all atoms for that mode. Thus the dynamical matrix **D** has size $3N \times 3N$, with each element:

$$D_{ab}(i,j) = \frac{1}{\sqrt{M_i M_j}} \frac{\partial^2 E}{\partial_{u_{ia}} \partial_{u_{jb}}} \tag{2.78}$$

here a and b are any of the Cartesian coordinates and u_{ia} and u_{jb} are the displacements of atoms with mass M_i and M_j respectively.

The energy double derivatives must now be found, starting from a relaxed cell. An atom i is moved a small amount ε in direction a, thus affecting the charge density and resulting in non-zero forces on the structure. The new force acting on atom j is in a direction b and is denoted $f_{bj}^{+}(a,i)$. Then atom i is moved by $-\varepsilon a$ giving rise to a new force on b of $f_{bj}^{-}(a,i)$. Then the second derivative on the energy is given to second order in ε by:

$$\frac{\partial^2 E}{\partial_{u_{ia}} \partial_{u_{jb}}} = \frac{f_{bj}^{+}(a,i) - f_{bj}^{-}(a,i)}{2\varepsilon} \tag{2.79}$$

The displacement ε is finite so the method includes a contribution from all even powers of ε; so some anharmocity is included. Thus the resultant frequencies are referred to as quasi-harmonic.

In practise only the atoms in or very near to a defect will have a significantly different potential from those in the bulk atom. The energy second derivative of the bulk atoms can be obtained using an empirical potential, in this work the Musgrave-Pople potential is used [68]. The use of this empirical potential reduces computational expense.

2.6.4 Mulliken bond population

Mulliken analysis allows the contribution of atomic bonds to a particular Kohn-Sham level to be calculated, by calculating the contribution from the basis functions. These results can then to be compared to experimental EPR measurements.

The contribution to a Kohn-Sham level λ from a basis function ϕ_i is given by:

$$p_\lambda(i) = \frac{1}{N_L} \sum_{(j,\mathbf{k})} c^{\mathbf{k}}_{\lambda i} S^{\mathbf{k}}_{ij} (c^{\mathbf{k}}_{\lambda j})^* \tag{2.80}$$

The overlap matrix elements $S^{\mathbf{k}}_{ij}$ can be given by:

$$S^{\mathbf{k}}_{ij} = \int B^*_{\mathbf{k}i} e^{-i\mathbf{k}\cdot\mathbf{r}} B_{\mathbf{k}j} d\mathbf{r} \tag{2.81}$$

where $B_{\mathbf{k}i}$ are basis functions, and $\sum_i p_\lambda(i) = 1$. The inclusion of basis functions in the equations allows contributions to bond population from specific orbitals to be found, so hybridisation of a state can be found.

2.6.5 Electron energy-loss

Defects often affect the electronic structure of a crystal. The removal of atoms in a vacancy defect, for example, or a dislocation can introduce dangling bonds which may result in deep band-gap states being introduced. Additionally, strain on bonds in the defect and surrounding crystal can alter the density of states outside the band-gap, or cause shallow gap states.

One of the most direct methods of measuring the electronic structure of extended defects is Electron Energy-Loss Spectroscopy (EELS), first proposed by Hillier and Baker [69]. In this technique an electron microscope passes high-energy electrons through a crystal, inducing electronic transitions between the valence and conduction bands, known as Low-Loss EELS. This is described more fully in section 3.2

There is a further area of measurement possible for bound atomic core state excitation, known as Core-Excitation EELS. This requires knowledge of the core atom energy levels, however these are removed through the use of pseudopotentials so cannot be modelled in AIMPRO.

The high-energy electrons which pass through the crystal interact with the field of polarisation caused by the electric field of the crystal's electrons. The quantity measured is the fraction of electrons scattered into a solid angle $d\Omega$ having lost energy between ΔE and $\Delta E + dE$. A differential cross-section $d\sigma^2/d\Omega dE$ can be related to the dielectric function by:

$$\frac{d\sigma^2}{d\Omega dE} \propto \frac{1}{q^2} Im \left\{ \frac{-1}{\varepsilon(\mathbf{q},\omega)} \right\} \tag{2.82}$$

Within the random phase approximation [70, 71], the macroscopic dielectric function is given as:

$$\varepsilon(\mathbf{q},\omega) = 1 + \left(\frac{e^2}{\varepsilon_0 \Omega}\right) \sum_{\mathbf{k}} \frac{|(\mathbf{r})_{\mathbf{k}v,\mathbf{k}+\mathbf{q}c}|^2}{E_{\mathbf{k}+\mathbf{q},c} - E_{\mathbf{k},v} - \hbar\omega} \tag{2.83}$$

Here $E_{\mathbf{k},c}$ and $E_{\mathbf{k},v}$ denote the conduction (empty) and valence (filled) bands and Ω is the volume of the crystal. The constant of proportionality linking the loss intensity with the inverse dielectric function depends weakly on energy, thus making an exact correspondence with experiment difficult. The scalar expression for the longitudinal dielectric function is, within the random phase approximation, the same as the expression for the transverse optical dielectric function [72]. The optical absorption coefficient, α, is related to the complex dielectric function by:

$$\alpha = \frac{4\pi}{\lambda}\left[\frac{-\varepsilon_1+\sqrt{\varepsilon_1^2+\varepsilon_2^2}}{2}\right]^{\frac{1}{2}} \tag{2.84}$$

where ε_1 and ε_2 are the real and imaginary parts of the dielectric function, as described below.

From a theoretical point of view, the imaginary part to $\varepsilon(\mathbf{q},\omega)$ and written as $\varepsilon_2(\mathbf{q},\omega)$, is first found and then its real part, $\varepsilon_1(\mathbf{q},\omega)$, is given by the Kramers-Kronig relation:

$$\varepsilon_1(\mathbf{q},\omega) = 1+\frac{2}{\pi}P\int_0^\infty \frac{\omega'\varepsilon_2(\mathbf{q},\omega')}{\omega'^2-\omega^2}d\omega' \tag{2.85}$$

A more general tensorial expression for the dielectric function appropriate for extended defects is:

$$\varepsilon_{i,j}(\mathbf{q},\omega) = \delta_{i,j}+\left(\frac{e^2}{\varepsilon_0\Omega}\right)\sum_{\mathbf{k}}\frac{(\mathbf{r}.\mathbf{i})_{\mathbf{k}v,\mathbf{k}+\mathbf{q}c}(\mathbf{r}.\mathbf{j})_{\mathbf{k}+\mathbf{q}c,\mathbf{k}v}}{E_{\mathbf{k}+\mathbf{q},c}-E_{\mathbf{k},v}-\hbar\omega} \tag{2.86}$$

In the case where the momentum loss $\hbar\mathbf{q}$ lies in the beam direction and for sufficiently high energies $q=\omega/v$ is be taken to be zero. The average of the diagonal components of the tensor is taken in order to compare with experimental results. This is related to the trace of the tensor and is independent of the orientation of the defect in the cell.

It is important to recognise that the expression for ε is limited and relies on perturbation theory and the neglect of local fields, plasmons and exciton effects. As a consequence the absorption of diamond in this theory begins strongly at the direct gap of 7 eV, as opposed to the 5.5 eV of the indirect gap. More involved theories of dielectric functions have been developed for semiconductors, to describe effects beyond the local density approximation [73], beyond the long-wavelength limit [74], and beyond the independent particle limit [74, 75], but they are currently too computationally intensive to be applied to anything other than bulk materials.

2.7 Density Functional Tight-Binding Theory

The Density-Functional Tight-Binding theory (DFTB) [76, 77] is a more approximate method of solving the Kohn-Sham equations, which has application for studying defects in larger cells, or to gain results quickly with less accuracy.

The DFT energy after introducing Kohn-Sham orbitals [29] is expressed as:

$$\varepsilon[n] = \sum_i^{occ} n_i \int \phi_i^*(\mathbf{r}) \left[-\frac{\nabla^2}{2} + \frac{1}{2} \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r' + \mathscr{U}_{ext}(\mathbf{r}) \right] \phi_i(\mathbf{r}) d^3r + E_{xc}[n] \tag{2.87}$$

where:

$$n(\mathbf{r}) = \sum_i^{occ} n_i |\phi_i(\mathbf{r})|^2, \quad N = \sum_i^{occ} n_i \tag{2.88}$$

The numbers n_i are the occupation numbers which describe how many electrons occupy each Kohn-Sham orbital.

The equations can be simplified by using a Taylor expansion around a reference density $n_0(\mathbf{r})$, then neglecting the higher order contributions. The energy is then [78]:

$$\begin{aligned} \varepsilon[n] &= \sum_i^{occ} n_i \int \phi_i^*(\mathbf{r}) \hat{\mathscr{H}}[n_0] \phi_i^{(}\mathbf{r}) d^3r - \frac{1}{2} \int\int \frac{n_0(\mathbf{r}) n_0(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r' d^3r \\ &+ E_{xc}[n_0] - \int \mathscr{U}_{xc}[n_0\mathbf{r}] n_0(\mathbf{r}) d^3r \\ &+ \frac{1}{2} \int\int \left[\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta^2 E_{xc}[n]}{\delta n^2} \right]_{n_0} \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') d^3r' d^3r \qquad (2.89) \\ & \qquad (2.90) \end{aligned}$$

The second to fifth terms here are expressible as the sum of two-centre integrals; which can be written as $E_{2_{cent}}$. If three-centre and higher contributions are neglected the total DFTB energy is:

$$\varepsilon[n] = \sum_i^{occ} n_i \int \phi_i^*(\mathbf{r})\hat{\mathscr{H}}[n_0]\phi_i^{(}\mathbf{r})d^3r + E_{2_{cent}} = \varepsilon_{DFTB} \tag{2.91}$$

In DFTB the starting density, $n_0(\mathbf{r})$, is expressed as a superposition of weakly confined neutral atoms; known as pseudo-atoms. The wavefunctions of the pseudo-atoms are represented by Slater orbitals, $\phi_\nu(\mathbf{r})$. The Kohn-Sham orbitals can then be expanded into atom centred orbitals [77], yielding:

$$\phi_i(\mathbf{r}) = \sum_\nu C_{\nu i}\phi_\nu(\mathbf{r}) \tag{2.92}$$

which is akin to a linear combination of atomic orbitals. The two-centre contributions can be approximated by a sum of short-range repulsive pair potentials $\mathscr{U}_{rep}^{ij}$:

$$E_{2_{cent}} \approx \frac{1}{2}\sum_{i\neq j} \mathscr{U}_{rep}^{ij}(|\mathbf{R}_i - \mathbf{R}_j|) = E_{rep}(\{\mathbf{R}_i\}) \tag{2.93}$$

The DFTB energy can then be written in terms of the expansion coefficients $C_{\nu i}$. This can then be varied with respect to the expansion coefficients and normalised to give the DFTB secular equations:

$$\sum_\nu C_{\nu i}(\mathscr{H}_{\mu\nu}^0 - \varepsilon_i S_{\mu\nu}) = 0 \tag{2.94}$$

If R_ν denotes the atom site the orbital is centred at, the Hamiltonian and overlap matrix are given by:

$$H_{\mu\nu} = \left\{ \begin{array}{ccc} \varepsilon_\mu^{atom} & : & \mu = \nu \\ \int \phi_\mu^*(\mathbf{r})\hat{\mathscr{H}}[n_0]\phi_\nu(\mathbf{r})d^3r & : & \mathbf{R}_\nu \neq \mathbf{R}_\mu \\ 0 & : & otherwise \end{array} \right. \tag{2.95}$$

and:

$$S_{\mu\nu} = \int \phi_{\mu}^{*}(\mathbf{r})\phi_{\nu}(\mathbf{r})d^{3}r \tag{2.96}$$

Once the pseudo-atom orbitals have been calculated self-consistently, the non-diagonal elements of $H_{\mu\nu}$ and $S_{\mu\nu}$ can be tabulated as a function distance between the two centres. The diagonal elements of $H_{\mu\nu}$ are taken as the atomic orbital energies of a free atom. It is clear that both matrices can be calculated in advance so that equation 2.95 can be solved directly and non self-consistently for any co-ordinates of the nuclei. In this case the DFTB energy becomes:

$$\varepsilon_{DFTB} = \sum_{i}^{occ} n_i \varepsilon_i + E_{rep}(\{\mathbf{R}\}) \tag{2.97}$$

with a repulsive energy E_{rep} and a repulsive potential $\mathscr{U}_{rep}^{ij}(|\mathbf{R}_i - \mathbf{R}_j|)$. Variation of distance allows tabulation of the repulsive potential akin to $H_{\mu\nu}$ and $S_{\mu\nu}$. The use of the repulsive potential in practise allows a minimal basis of Slater-type orbitals, reducing computational complexity.

The 'in-advance' calculations of DFTB reduce significantly the time taken to derive the Kohn-Sham eigenvalues and the total system energy. Although there is a penalty in accuracy.

2.8 Screened Exchange

The local density approximation (LDA) and Kohn-Sham approximations are known to underestimate the bandgap of semiconductors, leading to 0 eV bandgaps in some cases [79]. One technique which has been developed to compensate for this problem is the screened exchange technique. The technique is described in full in [13], and only a summary is provided here.

The Kohn-Sham bandgap underestimation occurs in the main because the zeroth Fourier component of the exchange potential changes discontinuously across the energy gap, *i.e.* $V_{\rm xc}(\mathbf{G})$ is undefined. Bylander *et al.* separate the exchange-correlation operator into two parts. One part can be treated exactly, the other has a weak ψ-dependent V_ψ which can be replaced by a density functional. The Hamiltonian then reads:

$$\hat{H}_{xc} = \hat{H}_{sx} + \frac{\partial E_x}{\partial \rho} - \frac{\partial E_{sx}}{\partial \rho} + \frac{\partial E_c}{\partial \rho} \qquad (2.98)$$

where $\hat{H}_{sx}$ is an exchange operator calculated with the screened Coulomb interaction $(e^{-K_s r})/r$. $E_x[\rho]$ is the LDA exchange energy-density functional in Rydberg units, $E_c[\rho]$ is one of the LDA correlation functionals, and $E_{sx}[\rho]$ is the LDA screened exchange density functional. $E_{sx}[\rho]$ is dependent on a factor γ, and it is possible to choose γ to be dependent on the electron density as a function of position $\gamma^2 \sim \rho^{-1/3}$, or not $\gamma^2 \sim \rho_0^{-1/3}$. The choice of γ to be independent from $\rho(\mathbf{r})$ is supported [13].

The LDA results in an exchange energy which is too small, and a correlation energy which is too large, when used for materials with a finite energy gap. A modified correlation functional is used:

$$\hat{E}_c[\rho] = WE_c[\rho]/(W + \Delta) \qquad (2.99)$$

Here $E_c[\rho]$ is a standard correlation density functional, W is the bandwidth and Δ is some average energy gap. In the free-electron-gas limit $\hat{E}_c[\rho] = E_c[\rho]$.

The technique was tested on silicon, where an energy gap of 1.32 eV was found. This compares to 0.44 eV using unmodified LDA and 1.17 eV for experiment. In sections 1.1 and 6.2.1 this theory is applied to AIMPRO and tested on diamond with very good results.

2.9 Summary

This chapter has described methods of solving the Schrödinger equation that reduce the complexity of the problem with various approximations. Comparison has been made of several techniques, and a special emphasis has been placed on the Density Functional Theory used in this work. In all the techniques large changes are made to the original energy equations to render them soluble. However, the results that are obtained have a high accuracy and have been proven to reproduce experimental measurements favourably.

The implementation of DFT using AIMPRO is detailed, and the concept of pseudopotentials and basis sets introduced. AIMPRO uses the Local Density Approximation to evaluate electron density, but this leads to bandgap underestimation. A method to correct this, the Screened Exchange technique, is briefly described.

The ability to compare theoretical predictions with experimental measurements is key to modern research. As new technology allows calculations to be performed on ever larger clusters of atoms, results should become more accurate and meaningful. To this end the remainder of this thesis will describe the results obtained in this work and compare them to experimental techniques and measurements.

Chapter 3

Experimental Techniques

The assignment of a particular experimental feature, such as a Raman line, to a defect is one area where experimental and theoretical work can combine particularly well. Theory can investigate the many possible configurations of a defect to find one that satisfies all experimental data, or may suggest new experiments to confirm previous assumptions.

This chapter will review experimental techniques that are directly applicable to the results of the theoretical work described subsequently.

3.1 Transmission Electron Microscopy

Transmission Electron Microscopy (TEM) is a highly important technique for studying defects in crystals. Its strength lies in the accessibility of its results, which can be analysed intuitively as they are the analogue of visible-light microscopy. It is easy for even a casual observer to observe and understand defects at a basic level.

3.1.1 Fundamental concepts

The electron microscope is similar in concept to the optical microscope, but uses the unique wave properties of electrons to overcome the resolution barrier of light with wavelength of 100's of nm.

The image from a point source is constrained by diffraction into a bright disc surrounded by dark and light halos [80]. The radius of this disc is $1.22F\lambda/d$, where F is the focal length of the lens, d is the diameter of the lens and λ the wavelength of the light. Rayleigh stated that if the maximum intensity arising from a point were to fall in the minimum of intensity from a separate point, then those two points are resolvable. This occurs when the separation is $1.22F\lambda/d$, and is known as the resolving power.

The relation between the diffraction-limited resolving power and the angle subtended by a point object at the objective lens, its apparent size, was given by Abbe in 1873 [81]. The relation states that the minimum resolvable separation in the object being observed is:

$$d_0 = 0.61\lambda/n\sin\alpha \tag{3.1}$$

where α is the half-angle subtended by the objective at the object, and n is the refractive index of the space between the object and the lens. Clearly, in air the minimum resolvable size is $\sim 0.6\lambda$, which for visible light of 600 nm is 3.6 μm.

Hence to increase the resolving power of the microscope λ must be reduced. Using ultraviolet light ($\lambda \sim$ 10nm) or X-rays ($\lambda <$ 1nm) would theoretically produce orders of magnitude improvement in resolution, but lenses for this range of electromagnetic radiation are difficult or impossible to produce. To achieve much greater resolution in microscopy the wave nature of fundamental particles must be exploited.

Wave nature of the electron

De Broglie showed that a moving electron can be regarded as being both particle-like and wave-like in its nature [82]. The wavelength of such a particle is expressed as $\lambda = h/mv$. Where h is Planck's constant and m and v are the mass and velocity of the particle respectively. A stream of moving electrons can then be considered as a beam of radiation in the same way as light. The other important factor to note is that by controlling the energy given to the electron the velocity (and mass at high energy) are altered and hence the resolution. As an example consider electrons accelerated through an electric field of 10 kV. The resultant energy of each electron is 10 keV, and using the relation: $pc = \sqrt{E^2 + 2Em_0c^2}$ the wavelength is 0.12 Å and the resolvable limit 7.2 pm which is less than the Bohr radius of the atom. Thus, in principle, it is possible to resolve individual atoms using a stream of highly energetic electrons.

3.1.2 Transmission Electron Microscopy in practice

Control of the electron beam

The generation, manipulation and use of electron beams can only occur in a vacuum, hence the TEM is a fully enclosed instrument evacuated to $< 10^{-4}$ mbar. In the simplest form the TEM is analogous to a simple optical microscope with two lenses - enabling a magnification of about 10,000 times [80]. Adding more lenses enables a larger magnification, so a modern high-resolution TEM may have three or more lenses as well as extra condensers for high accuracy incident electrons. This setup is shown schematically in figure 3.1.

Control of the stream of electrons occurs through the use of electric and magnetic fields. As electrons are charged particles and moving in a magnetic field they feel a force equal to $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, with B the magnetic field and q the charge on the electron. The lenses in a TEM consist of electromagnets which produce a variable field allowing focusing of the electron stream.

The actual resolution limit of a TEM is dominated by aberration from the lenses. If the 7 pm resolution calculated above is to be achieved with an electron wavelength of 0.12 Å then the beam width is 4.2 mm (assuming $F = 2$ mm). The lenses must control the wavefront over this distance with accuracy better than a quarter of the wavelength [83], that is, 0.3 Å. However, the shape of the lens is dictated by Maxwell's equations which means that the ideal, circularly symmetric lens cannot be formed in free space [84] and spherical aberration will always limit the resolution.

Scherzer, and later Feynman [85], realised that to correct spherical aberration a combination of octupole and quadropole fields was required. The octupole magnet has eight alternating poles, which produces a fourfold rotationally symmetric field. The field strength increase as the cube of the distance from the centre, which is necessary as spherical aberration is a third-order effect. In the octupole electrons are deflected away from the

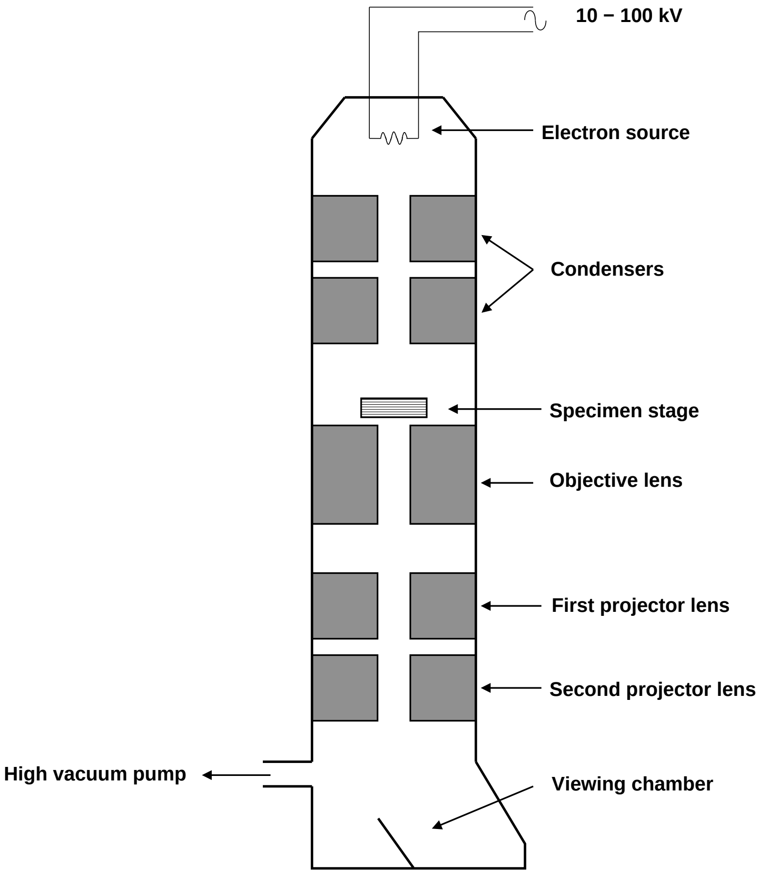


Figure 3.1: Schematic diagram of a high-resolution TEM.

centre in two perpendicular directions but towards it in between. This aberration is pre-corrected with a quadropole, which diverges the stream in one direction aligned to the divergent axis of the octupole (with resultant perpendicular convergence). In practice, two such quadropole-octupole combinations are used to correct in both perpendicular directions.

Sample preparation and imaging

Preparation of a specimen to be observed in TEM must be undertaken carefully, and satisfy several requirements:

- The sample must transmit a sufficient number of electrons and minimise energy loss (inelastic scattering). This limits the sample thickness, with a maximum of 100 nm generally being applicable.

- The sample must be able to withstand the energy of the incident electrons, as some fraction will be inelastically scattered or absorbed; with resultant energy exchange.
- Any preparatory treatment should not affect the sample on a level observable with the TEM.

Samples of crystalline defects (*e.g.* diamond) are often prepared by ion-beam milling or chemical etching. The disadvantage of ion-beam milling is that it may impart significant energies to atoms in the sample, which can then induce structural defects. An example is patches of graphite or non-diamond carbon on the surface of diamond thin films. It is necessary to exercise great care when preparing samples to ensure this does not occur. Samples are often prepared with the faces perpendicular to some particular crystallographic direction to facilitate identification of diffracted directions.

Electrons incident onto the sample can interact in a number of ways, with their final trajectories determined by the process involved. The two groups of interactions are elastic and inelastic, with zero and finite energy transfer respectively. These can then be subdivided into the groups of back-scattered and forward transmitted electrons. The number of elastic and inelastic electrons gives information on the variation of the atomic number throughout the sample. Elastic scattering is proportional to $Z^{4/3}$, inelastic scattering varies as $Z^{1/3}$ and hence the ratio is proportional to Z, the atomic number.

Transmitted electrons are focused onto a screen coated with a photo-emitting substance. If no sample is present then the image field is uniformly bright, when the sample is inserted darker areas indicate where electrons have been scattered by passing close to atoms. In a regular crystalline material bright and dark images are formed depending on whether the Bragg condition has been satisfied, so the sample is often mounted on a tilting platform to allow adjustment.

Depending on whether the primary or diffracted beam is used for imaging gives rise to the *bright-field* and *dark-field* conditions respectively. In a bright-field, the image is formed by electrons that have not interacted with the sample, so defects or atoms appear dark as electrons interacting with them have deflected away. It is useful to image a sample using the diffracted beam of electrons, as they have undergone interactions with the sample [80]. This is the dark-field condition, which has the advantage of showing low-contrast detail more strongly.

In chapter 5, dislocations in diamond are studied, with reference to the Burgers vector. TEM can accurately determine the Burgers vector and hence dislocation type using the condition that $\mathbf{g} \cdot \mathbf{b} = 0$, where $\mathbf{b}$ is the Burgers vector and $\mathbf{g}$ is the reciprocal lattice vector. So, by tilting the sample and observing the same dislocation with various diffraction vectors the angle of the Burgers vector to the line vector can be determined.

3.2 Electron Energy Loss Spectroscopy

Electron Energy Loss Spectroscopy (EELS) is a technique related to TEM, in that it uses principally the same equipment setup, and the same method of analysing a sample with a stream of high energy electrons. In fact, the two techniques can often be carried out in the same instrument. EELS makes use of the energy lost by electrons through inelastic collisions to determine the local electronic structure of bulk material or a defect.

3.2.1 Fundamental concepts

In terms of equipment, up until the collection of transmitted electrons the TEM and EELS instruments are the same. In EELS, a magnetic spectrometer is used to separate electrons of different energies for analysis, shown in figure 3.2. Electrons which have suffered the greatest energy loss are moving the slowest and hence are deflected most. To observe electrons having a particular energy a slit is placed in the beam of separated electrons, which only allows electrons which have been deflected to that angle to pass through. Electron numbers are counted by a scintillator or photomultiplier. In an alternative technique a lens focuses all paths of electrons with various energies onto an array of detectors, similar to a CCD device. Thus, all energy ranges can be sampled simultaneously.

The detected energy loss of the electrons is generally displayed relative to the incident beam energy. Therefore any electrons which have undergone only elastic collisions contribute to a peak at 0 eV, known as the zero-loss peak, which is seen in all EELS ex-

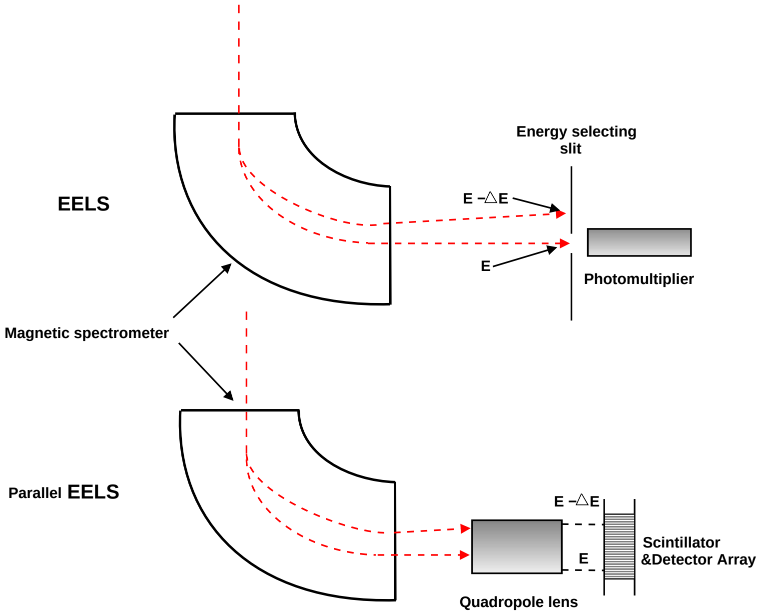


Figure 3.2: Diagram of serial and parallel methods of detecting energy loss of atoms after interaction with a sample [80]. The dashed lines indicate the paths of electrons with energy from E to $E-\Delta E$.

periments. The rest of the energy-loss spectrum can then be divided into two regions: the low-loss and core-loss regions. The core-loss region relates to energy lost by the electrons by transitions from the core to empty levels of the atoms in the sample, so has higher energy loss of the order 300 – 800 eV. This requires a detailed knowledge of the core electron wavefunctions, but these are eliminated in this work through the use of pseudopotentials, thus core-loss EELS is beyond the scope of this work and is not discussed further.

Low-loss EELS studies the energy losses occurring from the promotion of transitions from occupied valence states to conduction states. This region is generally below 50 eV and information can be deduced on the bandgap of a material as well as the valence and conduction bands. In addition this is where the plasmon peak of a material often occurs, which is observed experimentally at 33 eV in diamond [86], and can be predicted by the free electron theory:

$$\omega_p = \left(\frac{4\pi n_e e^2}{m}\right)^{1/2} \tag{3.2}$$

ω_p is the plasmon frequency, n_e is the electron density, e the electron charge and m the *effective* mass of the electron in diamond.

3.3 Optical Absorption

Optical absorption simply measures the ratio of transmitted to incident power of a material, for a variety of wavelengths. In the range 400 nm – 700 nm (3.09 eV – 1.77 eV) the absorption of the material will be directly visible as colour, but experiments will often extend in to the ultra-violet and infra-red regimes. The absorption spectrum of a material is unique and can quickly classify that material. Defects may have a well-known absorption ‘signature’ which can modify the host material absorption spectrum, sometimes with visible results, allowing identification of defects and an estimate of their concentration.

3.3.1 Fundamental concepts

When a beam of electromagnetic radiation passes through a material photons may interact with the material atoms in a number of ways. There are two processes which are relevant to absorption in the visible wavelength range. An incident photon may be absorbed causing an electron to jump from a ground to an excited state, such as a valance band to conduction band transition, leaving a hole. Alternatively the incident photon may give rise to a phonon or the local vibration of a defect. A dipole is required for one-phonon absorption which explains its absence in pure diamond, however two-phonon and three-phonon absorption in evident in the infra-red region.

The intensity of an electromagnetic beam is described by:

$$I = (h\nu)N \tag{3.3}$$

where N is the photon flux [87]. When the beam traverses a material the flux is reduced, by the processes described above. If photons are incident on a material of thickness dx, and the number of transmitted photons is $N - dN$, then the absorbed photons can be related to the thickness and incident flux by the relation:

$$dN = -\mu N dx \tag{3.4}$$

with a proportionality constant μ, known as the absorption coefficient. This constant is unique to any given material, and depends on the density of the material (as this dictates the number of atoms encountered by the incident photon beam). This relation can be integrated for a material of total thickness x and for incident and transmitted fluxes N_0 and N respectively. Therefore:

$$N = N_0 e^{-\mu x} \tag{3.5}$$

and from equation 3.3:

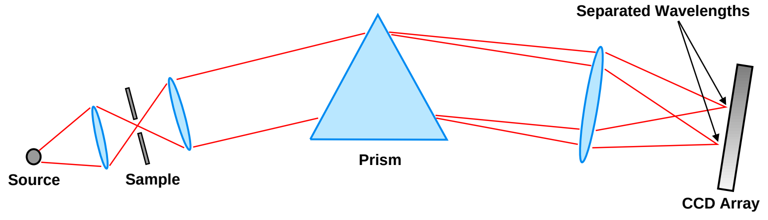
$$I = I_0 e^{-\mu x} \tag{3.6}$$

It is clear that the absorption increases with the thickness x and the absorption coefficient μ. It is important to note that the absorption coefficient is dependent on the incident photon energy (frequency), which allows the absorption spectrum to be determined.

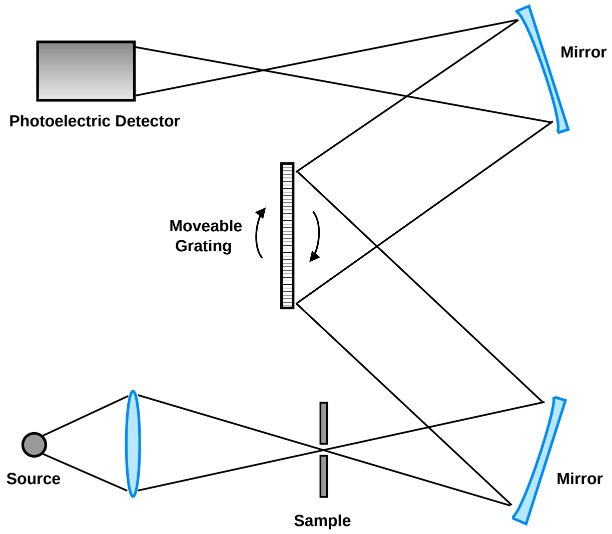
3.3.2 Optical Absorption in practice

Optical absorption appears straightforward in principle, with simply a photon source shining through a sample and photon detectors measuring the incident and transmitted flux at any given wavelength. In practice, modern experimental techniques use a variety of techniques to vary the incident wavelength and measure it accurately. The emergence of lasers as a suitable light source has depended on the ability to 'tune' the wavelength to any desired value.

The simplest instruments, Spectrographs and Monochromators, both rely on the separation of transmitted light into its components by a prism or diffraction grating, but differ in their detection of photon flux [88]. In a spectrograph a CCD array is used to sample all offset wavelengths simultaneously, whilst a monochromator uses a photoelectric recording of a selected interval. In the monochromator the prism or grating is rotated to focus the selected wavelength onto the detector, see figures 3.3(a) and 3.3(b).



(a)



(b)

Figure 3.3: (a) Prism spectrograph and (b) Grating monochromator. The method of selection of wavelength is interchangeable between the devices, the difference is defined by the detection technique.

Advanced instrumentation

A laser light source has the advantage of steady power (intensity), coherent light output and ease of manipulation with optical components. The main disadvantage which must be overcome is the single wavelength output, which can be modified by tuning a single-mode laser.

Single-mode lasers have an output wavelength determined by the size of the resonant cavity [88], $\lambda = 2nd/q$. Here n, d and q are the refractive index of the gain medium, the mirror separation and the quality factor of the cavity respectively. The output wavelength is then altered by variation of n or d. If the gain medium is a gas then n may be altered by changing the pressure. The mirror separation d can be accurately changed using a piezoelectric element as one of the mirror mounts. The variation of d alone presents problems with mode hopping which can be solved using a tilting etalon inside the resonance cavity, which is beyond the scope of this review but described in depth in [88].

3.4 Raman Spectroscopy

Raman spectroscopy makes use of the quantised rotational and vibrational energies of a molecule or solid to deduce its properties. These excitations in the system absorb or emit photons which causes a shift in energy of a light source incident on the system, which is the basis of the technique. Raman spectroscopy is very sensitive to molecular bonding and structure of a sample, which leads to a distinct spectral pattern for a material. However it is a weak effect so must be carefully measured.

3.4.1 Fundamental concepts

When a photon of light is incident on an atom or molecule, it may be energetic enough to induce an electronic transition, but if this is not the case it can be scattered in one of three ways [89]. Rayleigh scattering, where there is no energy loss, is of no interest in Raman spectroscopy. Inelastically scattered light can either lose or gain energy during interaction, giving rise to a shifted frequency in the emitted radiation. Such scattering of light (or radiation generally) with an accompanying change of frequency is known as the Raman effect [90]. The scattered radiation usually has different polarization characteristics from the incident radiation [91].

A system at energy E_1 can make a transition to a higher energy E_2 if supplied with energy from incident radiation. It can be visualised as the annihilation of the incident photon, with the simultaneous creation of a lower energy photon [91]; accompanied by the energetic transition described. The newly created photon then has a wavelength longer then the incident one. This is known as Stokes Raman scattering (figure 3.4). It is also possible for the incident radiation to induce the reverse transition; lowering the system energy from E_2 to E_1. The emitted radiation has a shorter wavelength, known as Anti-Stokes scattering.

The Raman effect may be described both classically and quantum mechanically, with an overview of each given here.

Classical approach

The classical approach regards the scattering molecule/system as a collection of atoms undergoing simple harmonic motion [89]. To describe the scattering types requires knowledge of the dipole moment μ_i induced by the incident radiation, which for small fields is proportional to the field strength:

$$\mu_i = \alpha\varepsilon \tag{3.7}$$

where ε is the field strength and α is the polarizibility of the molecule. The electric field generated by electromagnetic radiation is expressed:

$$\varepsilon = \varepsilon_0 \cos 2\pi t \nu_0 \tag{3.8}$$

with ε_0 the equilibrium field strength for the material and ν_0 the angular frequency (or wavenumber) of the radiation. The incident radiation will induce a fluctuating dipole with the same frequency in the system. If the dipole then emits radiation of the same frequency, there is no energy loss. This is Rayleigh scattering.

If the incident radiation interacts with the molecule the equations 3.7 and 3.8 combine to give:

$$\mu_i = \alpha\varepsilon_0 \cos 2\pi t \nu_0 \tag{3.9}$$

If a molecule undertakes simple harmonic vibrations with a frequency ν_0, its polarizibility changes during vibration and is expressed:

$$\alpha = \alpha_0 + \left(\frac{\partial \alpha}{\partial q_v}\right)_0 q_0 \cos 2\pi t \nu_0 \tag{3.10}$$

where q_v is a coordinate along the axis of vibration at time t. If this is then substituted into equation 3.9, the total dipole moment can be written:

$$\mu_i = \alpha_0 \varepsilon_0 \cos 2\pi \nu_0 t + \left(\frac{\partial \alpha}{\partial q_v}\right)_0 \frac{\varepsilon_0 q_0}{2} \times [\cos 2\pi(\nu_0 + \nu_v)t + \cos 2\pi(\nu_0 - \nu_v)t] \tag{3.11}$$

The first term describes the Raleigh scattering with the second and third terms describing the Stokes and anti-Stokes Raman scattering. It can been seen that the polarizibility of the system must change during vibration, for that vibration to be Raman active.

Quantum approach

The vibrational and rotational energies of a molecule are quantised according to the relationship:

$$E_v = h\nu(n + 1/2) \tag{3.12}$$

Where n is the vibrational quantum number, having values 0,1,2,3.... The levels can be visualised in the ideal model of figure 3.4. At room temperature most vibrations are in the ground state ($n = 0$), which makes anti-Stokes transitions much less likely and hence weaker in the scattering spectrum.

The transition from one energy level to another can be visualised using perturbation theory, with a perturbing wavefunction applied to the ground state which then produces a new wavefunction corresponding to the excited state.

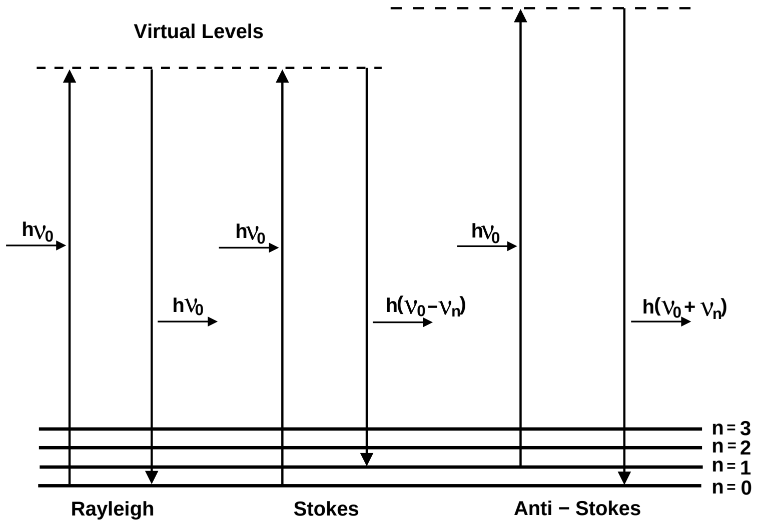


Figure 3.4: The vibrational energy level transitions which can be induced by a photon of energy $h\nu_0$.

3.4.2 Raman Spectroscopy in practice

The basic components needed to undertake Raman spectroscopy are a monochromatic light source, a system to collect the scattered photons and separate them into individual wavelengths (a monochromator) and a detector to measure intensity.

Monochromatic light is most easily obtained from a laser in either continuous wave or pulsed configuration. It can be performed at various excitation energies, however sample fluorescence is reduced by using near-infrared excitation. Fluorescence is undesirable as it can obscure the Raman signal. Red wavelength excitation is also less likely to cause photo-decomposition or thermal damage due to absorption.

The Raman scattered light is weak in magnitude compared to the elastically-scattered (Raleigh-scattered) light, so one of the critical tasks of the detection optics is to separate out these components. There are two distinct methods to achieve this. Dispersive Raman has become the most widely used method, consisting of a monochromator to separate out the Raleigh wavelengths with a grating to further separate the light into its constituent wavelengths (figure 3.5(a)). The spectrum is then detected by a CCD array. Fourier-Transform (FT) Raman utilises an interferometer to introduce a path difference between the source and the signal beams (figure 3.5(b)). The resultant interference pattern can then be used to reconstruct the Raman spectrum.

FT Raman generally uses a longer wavelength laser of 1064 nm which is necessary when analysing samples which fluoresce when exposed to visible wavelength radiation [89]. This wavelength laser has poor illumination, and FT Raman suffers from poor spatial resolution. Dispersive Raman spectroscopy can be used with several types of laser, allowing selection of an appropriate wavelength for the material to be analysed. The concept of the Raman Microscope, a non-destructive and non-contact analysis method, can be realised only with dispersive Raman.

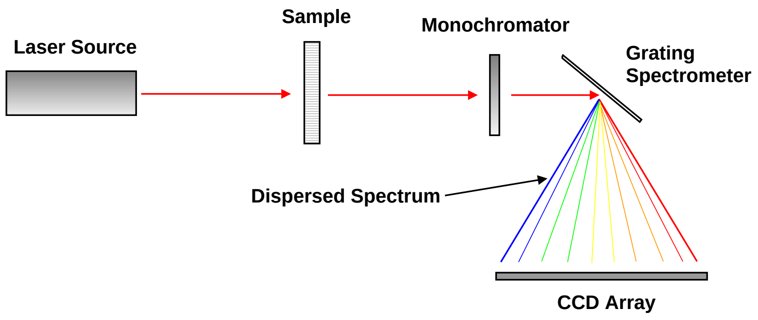
3.5 Photoluminescence Spectroscopy

Photoluminescence (PL) is the spontaneous emission of light from a material when optically excited. PL spectroscopy utilises this phenomena to investigate the electronic levels of a material, and can classify bandgap and defect states as well as thermally activated processes [92]. It has the advantage of being non-contact and non-destructive of any sample material.

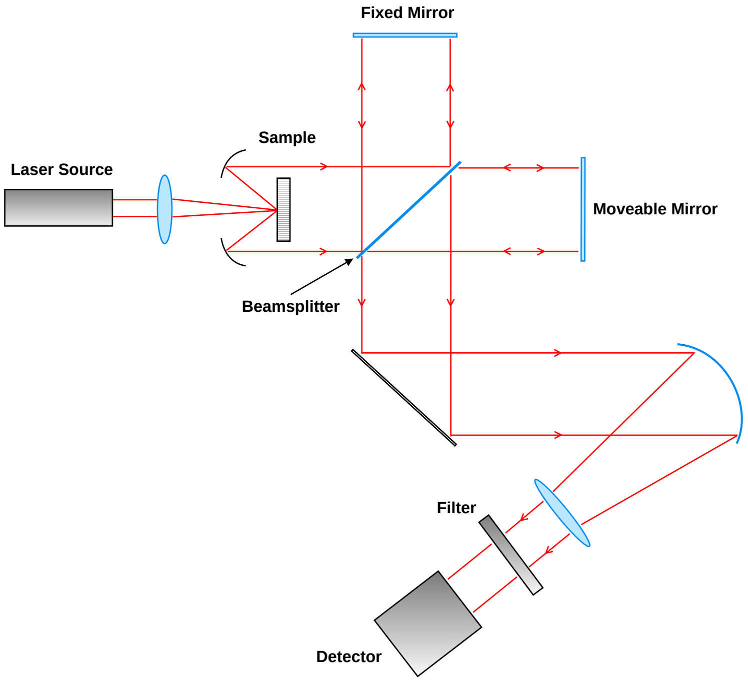
3.5.1 Fundamental concepts

When a photon of sufficient energy is incident upon an electron in a material, it may cause that electron to be excited to a higher energy level. The excited electron can then directly or indirectly relax back to its original state, emitting a photon in the process. In the case of a direct-gap intrinsic semiconductor, the emission energy is the bandgap and hence the frequency of the emitted light is dictated by $E_{gap} = h\nu_0 = h2\pi/\lambda$. If the material is doped, and hence has a shifted Fermi level then the excited electron may move down to a lower energy level before emission, and hence emit with a lower frequency (longer wavelength) as shown in figure 3.6. Here $h\nu_1$ will be equal to $\mathrm{E_{gap}}$ in the intrinsic case and $\mathrm{E_{reduced}}$ in the doped case.

There are modifications to these equations depending on phonon assistance and exciton effects. Phonons allow transitions to take place between different positions in the Brillouin zone, adding momentum when there is an indirect bandgap. This is the case for diamond.



(a)



(b)

Figure 3.5: (a) Dispersive Raman with a grating spectrometer and (b) Fourier Transform Raman using a Michelson-Morely interferometer.

In this situation the incoming photon must have sufficient energy to both excite the electron across the gap, and create a phonon to assist the transition across **k**-space. Thus the energy gained by the excited electron is $h\nu_0 - h\Omega$ where Ω is the phonon frequency. When the electron relaxes through PL, it must sacrifice more energy to create an second phonon with equal and opposite momentum to the first (thus satisfying conservation of momentum), so the total possible energy of the emitted photon is $h\nu_1 = h\nu_0 - 2(h\Omega)$ as

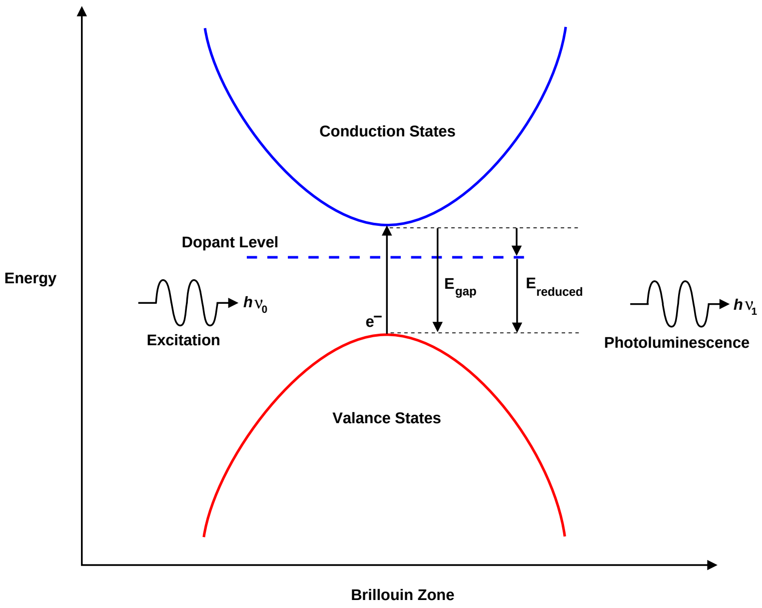


Figure 3.6: Schematic of Photoluminescence in a direct-gap semiconductor.

shown in figure 3.7.

If there are defects with localised energy levels in the gap, then an electron or hole may become trapped there. The PL emission is then characteristic of that defect. If both the ground and the excited states are in the gap there will be a sharp line in the emission spectrum.

When an electron is excited to the conduction band it leaves behind a hole of opposite electric charge. These opposite charges are attracted to each other by the Coulomb force, forming a quasiparticle known as an exciton. The exciton can move throughout the material without feeling the effects of other electrons around it, known as screening. There is an energy difference between the exciton and an unbound electron and hole, due to the binding force. This results in a lesser amount of energy being available for photon emission when recombination occurs. The modified emission energy is given by $h\nu_1 = E_{gap} - E_{exciton}$ in the case of direct transitions, and by $h\nu_1 = E_{gap} - E_{exciton} - 2(h\Omega)$ for an indirect transition. The reduced energy in exciton transitions can be visualised as quasistates near the conduction band, shown in figure 3.8. A material must be very pure for exciton emission to be prominent, so this can be used as a test of purity.

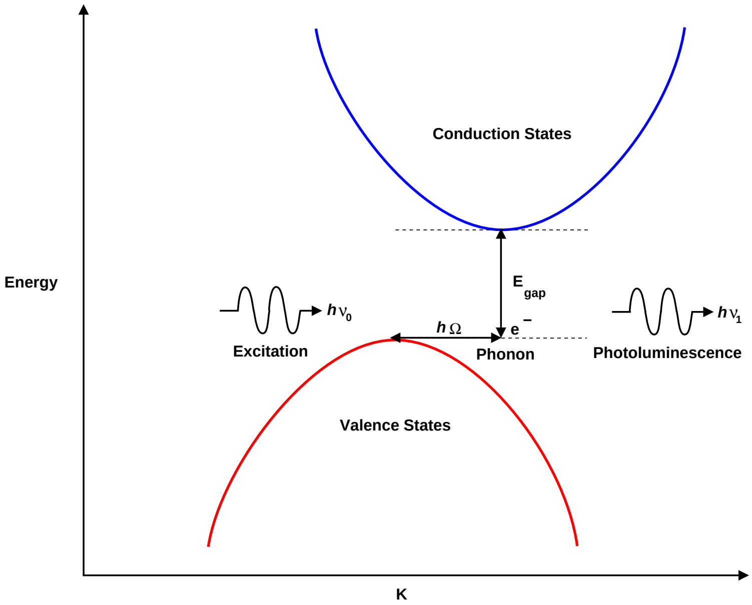


Figure 3.7: Phonon-assisted PL in a indirect-gap semiconductor. The conservation of momentum is achieved by the emission of phonons in opposite directions during excitation and emission.

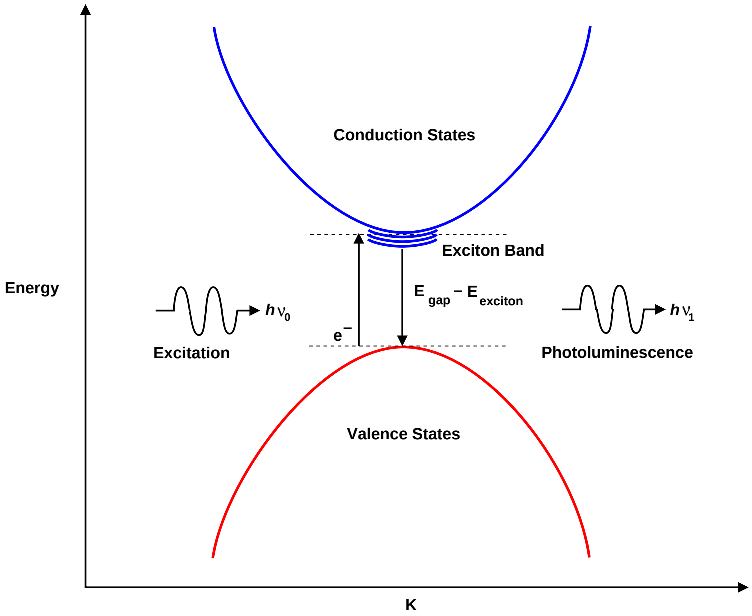


Figure 3.8: Exciton effects in Photoluminescence. The reduced energy due to electron-hole binding can be visualised as a lowering the conduction band.

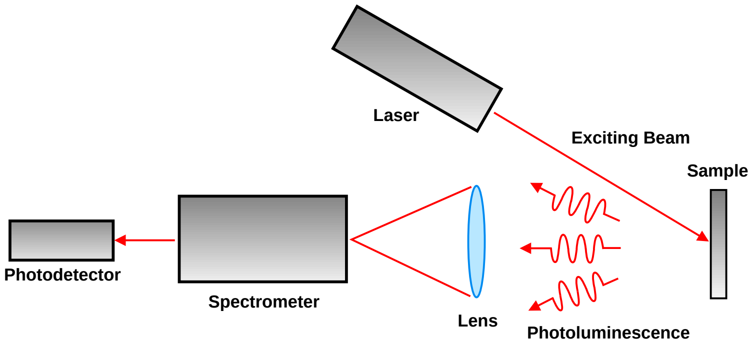


Figure 3.9: Generalised experimental setup for measurement of Photoluminescence.

3.5.2 Photoluminescence Spectroscopy in practice

Depending on the number and position of dopant related levels in the material there may be a wide range of frequencies emitted from any electron recombination. Clearly then, there may also be a range of excitation energies required. The requirement for the experimental setup is thus similar to the optical absorption in section 3.3, with a tunable laser to deliver the range of energies, and a spectrometer to separate the emitted wavelengths for analysis. The emitted photons may be spatially dispersed by virtue of the location of the emission site in the sample as well as scattering from surrounding material. This dispersion requires the positioning of a lens to direct emission onto the detector as shown in figure 3.9. When exciting a material at room temperature bandgap emission can mask luminescence from defect states. The experimental setup may make provision to cool the sample to below 100 K to allow defect-related luminescence to be observed.

3.6 Positron Annihilation Spectroscopy

Positron Annihilation Spectroscopy (PAS) can be used to deduce information on vacancy related defects in crystals. It utilises the positron, which is the electron anti-particle, by measuring its lifetime of interaction in a crystal with the existing electrons. Vacancy related defects have a lower electron density and thus show up as having a longer positron lifetime than the bulk material.

3.6.1 Fundamental concepts

PAS is based on the electromagnetic interaction between positrons (e^+) and electrons (e^-), which makes possible annihilation of the two particles, with the resultant energy transferred to gamma radiation. The principle reaction is the two-photon annihilation [93]:

$$e^+ + e^- \rightarrow \gamma_1 + \gamma_2 \tag{3.13}$$

the positron e^+ annihilates with an electron e^- to produce two photons, which emit in directions differing by π radians *i.e* opposite, with an additional directional shift θ:

$$\Delta\theta \approx p_{-,T}/m_0 c \tag{3.14}$$

$p_{-,T}$ is the transverse momentum component of the annihilating electron. The annihilation processes are characterised with an annihilation rate λ which is related to the positron lifetime by $\lambda = \tau^{-1}$. The annihilation rate is proportional to the effective electron density n_e and the classical electron radius r_c:

$$\lambda \approx \pi r_c^2 c n_e \tag{3.15}$$

In a metal the conduction electron density may be of order 10^{-29} m^{-3}, which gives an expected positron lifetime of $\sim$ 600 ps. Experimentally, values of $\tau \sim$ 200 ps are measured, due to electron density enhancement at the positron site [93]. This occurs when the Coulomb force between the positrons and electrons causes mobile conduction electrons to move towards the positron site and increase the effective density.

As positrons are injected into the material, the majority interact with electrons in the bulk and give rise to a dominant bulk lifetime τ_b [38]. If monovacancies or larger vacancy clusters are present, then some fraction of the injected positrons will become trapped at them. These vacancy related defects have a lower electron density than the bulk, so according to equation 3.15 the lifetime will be longer. In any given sample a range of lifetimes τ_1, $\tau_2 \ldots$ etc. can be extracted from the total lifetime spectrum obtained over

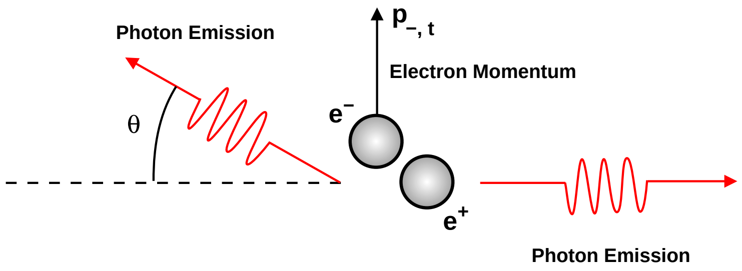


Figure 3.10: Positron annihilation reaction schematic.

many annihilation events. These lifetimes can then be related to individual defect groups in the sample.

The trapping rate of the positrons, κ is proportional to the vacancy concentration according to:

$$\kappa = \mu C_v \tag{3.16}$$

where μ is the absolute specific trapping rate, and C_v the absolute vacancy concentration. The trapping rate is a central parameter but cannot be directly determined by analyses of the lifetime spectra [94], rather it must be calculated from experimentally obtained parameters:

$$\kappa = \frac{I_2(1/\tau_b - 1/\tau_2)}{(1 - I_2)} \tag{3.17}$$

with τ_2 the lifetime for positrons trapped at vacancies, and I_2 its intensity.

3.6.2 Positron Annihilation Spectroscopy in practice

The usual source of positrons for PAS experiments is a radioisotope emitting β^+ radiation. The energy range of the emitted positrons is typically of order 0.1 – 1 MeV [93], which gives a penetration depth of about 10 – 100 μm. Sources that simultaneously emit gamma radiation are particularly useful, as the delay in detection between gamma

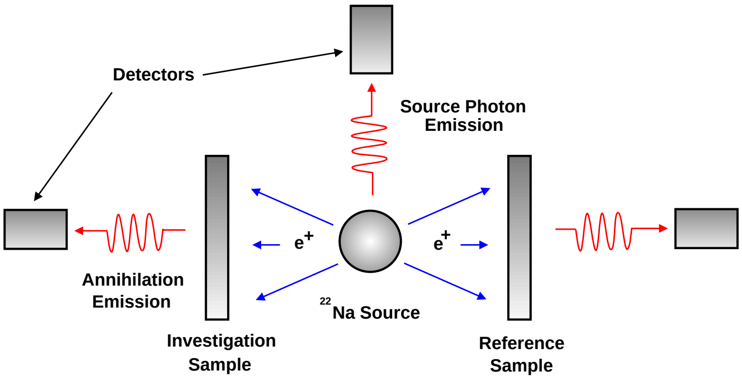


Figure 3.11: PAS experimental setup. The double sample setup to provide a known background count is illustrated.

emission from decay, and that from positron annihilation, gives the positron lifetime.

To ensure that all positrons emitted from the source are annihilated in a sample, the source is placed between two sample plates [35]. If two identical samples are not available one can be replaced with a sample for which the properties are known, in the case of diamond this could be a CVD or HPHT grown sample. The final lifetime spectrum then contains a contribution from the sample of interest and the source contribution, which is subtracted. Figure 3.11 demonstrates the sample setup.

3.7 Electron Paramagnetic Resonance

Electron Paramagnetic Resonance (EPR), also known as Electron Spin Resonance (ESR) or Electron Magnetic Resonance (EMR), is technique utilising the reaction of unpaired electrons to an applied magnetic field, the Zeeman effect. It has the capability to identify the atomic species, structure and symmetry of many defects in a wide range of materials.

3.7.1 Fundamental concepts

The electron has a magnetic moment which, when subjected to an external field $\mathbf{B}_0$, can align itself parallel or antiparallel to that field. These alignments differ in energy, with the parallel state lower, hence the levels are split. This splitting is known as the Zeeman

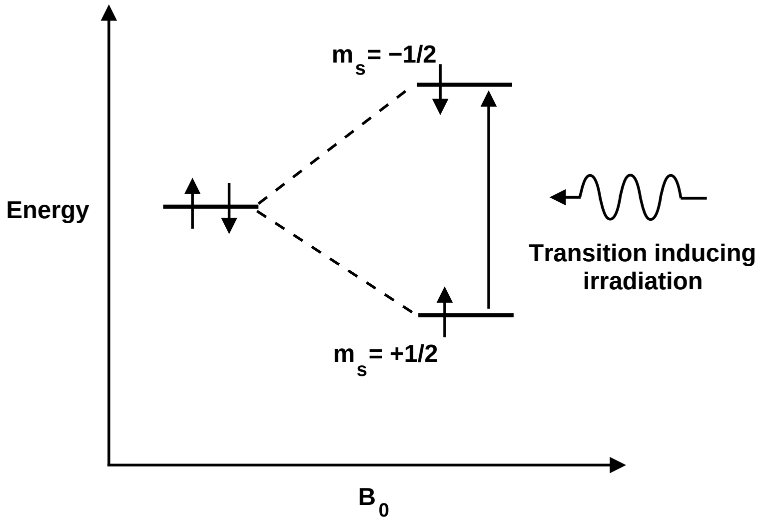


Figure 3.12: Zeeman splitting for a single electron. Incident radiation may induce a transition between the parallel and anti-parallel spin states.

effect, and the energy between the two states is given by:

$$\Delta E = E(m_s = 1/2) - E(m_s = -1/2) = \frac{g_e \beta_e B_0}{h} \tag{3.18}$$

where $m_s = \pm 1/2$ is the spin quantum number for each spin state, g_e is the gyromagnetic ratio of the electron and β_e denotes the Bohr magneton.

If electromagnetic radiation is applied to the electron, and the photon energy matches the energy difference between the states, then a transition between the states can be induced. The spin is flipped from one orientation to the other, and the resonance condition is fulfilled. The electron splitting and spin states transition are illustrated in figure 3.12.

The single electron case is simple but can be applied to more complex cases such as transition metal ions. In most other cases additional magnetic or electrical fields in the material or the presence of more than one electron complicate the spectra. This will lead to further energy level splitting. Interaction of the electron spin with the external magnetic field yields an EPR spectrum with one single line which may be broadened due to g_e anisotropy. If nuclei with a spin quantum number of $I > 0$ are in the same region, the electron spin will experience a local magnetic field built up by these nuclei. This is due to the magnetic moment μ associated with the intrinsic spin angular momentum L, and is called the hyperfine interaction. The nuclear spin is quantised in a magnetic field so, like the electron spin, undergoes a nuclear Zeeman splitting, expressed by:

$$E_n = -\frac{g_n \beta_n l B_0}{h} \tag{3.19}$$

where g_n is the nuclear g-factor, and β_n the nuclear magneton. For a given nuclear quantum number l, there will be a splitting into $(2l+1)$ energy levels, each characterised by a nuclear magnetic spin quantum number $m_i = l, l-1, \ldots, -l$ The hyperfine interaction depends on the orientation of the nuclear magnetic moment with respect to the external magnetic field, and results in a splitting of each electron state into $(2l+1)$ levels.

In practice, EPR observes a large number of centres throughout a material, which are statistically described by the Boltzmann distribution:

$$\frac{n_{M_j+1}}{n_{M_j}} = \exp\left(-\frac{\Delta E}{k_b T}\right) \tag{3.20}$$

where n_{M_j+1} is the number of centres in level M_j.

3.7.2 Electron Paramagnetic Resonance in practice

Transitions from one energy level to another occur at certain combinations of the external magnetic field $\mathbf{B}_0$ and the irradiation frequency v. It is experimentally simpler to maintain a steady irradiation and to sweep through a range of magnetic field strengths, rather than vice-versa. The excitation frequency is usually in one of four bands; the S-band (2-4 GHz), X-band (8-10 GHz), Q-band ($\sim$35 GHz) or W-band ($\sim$90 GHz). The swept magnetic field varies in the range of 100 – 300$\times$10^{-3} T, for example in the X-band a value of g_e =2 corresponds to 320 mT.

It is the value of g_e which describes the centre's electronic structure. If the value of g_e measured at a resonance condition differs from the free electron value (g_e = 2.0023), then the electron has lost or gained angular momentum through spin-orbit coupling. The

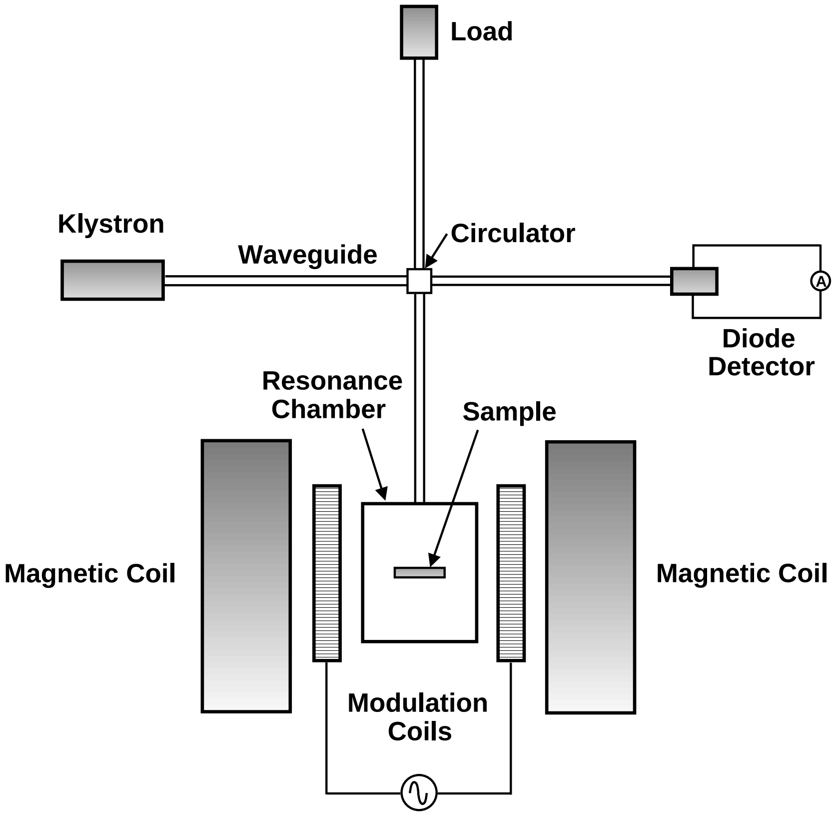


Figure 3.13: Experimental setup for EPR/ESR experiments. The resonance chamber may be cooled to low temperatures to reduce noise

change in angular momentum is anisotropic and varies with orientation to the external field, yielding information about the atomic orbitals containing the unpaired electron.

The sample to be studied is placed in the resonance chamber as shown in figure 3.13, and the microwave radiation transferred via a waveguide (which can be as simple as a brass tube). The illuminating power is low, typically 200 mW with a 2 MHz bandwidth. The resonance chamber is designed to absorb the incident power completely, hence any absorption of radiation by the sample will detune the chamber and radiation will be reflected back up the waveguide to the diode detector. The measurement of this radiation as a function of magnetic field yields the EPR spectrum. The signal-to-noise ratio is improved by applying an amplitude modulation to the magnetic field, typically at a frequency of 100 kHz.

As well as the constant wave approach described above, EPR can be conducted in a pulse configuration. In this arrangement a high-power microwave pulse at a single frequency ν, and fixed magnetic field $\mathbf{B}_0$, excites a large range of frequencies in the sample. Fourier transform leads to an excitation range for the pulse of $\nu\pm$1/2 (t_p), which typically could be 100 MHz for a 10 ns pulse. Reduction of the pulse time increases the excitation range, however due to technical limitations it is difficult to excite the whole spectrum by a single EPR pulse.

Continuous wave EPR has the advantage of higher sensitivity, plus spectra can be recorded at room temperature unlike pulse EPR. Pulse EPR does allow additional information about weakly coupled nuclei and relaxation properties of the spin system to be gathered, via manipulation of spin with a sequence of microwave pulses. A combination of both techniques is necessary for full details of a paramagnetic centre to emerge.

3.8 Summary

There are a range of experimental techniques that can be applied to diamond, or any other material. It is possible to use them to deduce mechanical, electronic and optical properties, or to discover point and extended defects. The main importance of experimental techniques to theoretical modelling work is to compare known results to those calculated by theory, for example the absorption spectrum of bulk diamond, to gauge the accuracy of the model. In turn this accuracy will provide the degree of confidence in the model's prediction of unknown quantities.

The main techniques relevant to this work are those that deal with the optical properties of a material, which can be modelled as described in section 2.6.5. Therefore most reference will be made to the Electron Energy Loss spectroscopy (EELS), on which collaboration with experimental work [95] has been undertaken, and to optical absorption studies. In addition, the predictions of this work have been studied using Transmission Electron Microscopy [96].

Chapter 4

Small Vacancy Clusters

4.1 Introduction

Elements of this work have been published in *Physica Status Solidi (a)* **202**, 2182 (2005).

Electron paramagnetic resonance (EPR) studies of multivacancy centres have been conducted on chemically pure type IIa diamond which has been irradiated and subsequently annealed [39, 41]. Eight defects are detected following anneals between 700 and 1400° C. These are labelled O1, R5, R6, R7, R8, R9, R10 and R11. The first three possess spin $S =$1 and R7 is described as a $S =$3/2 centre [39]. For the remainder of centres it was uncertain whether they possessed spin 1 or 3/2. More recent work indicates a $S =$1 for R10 and R11 [41], and also for R7 [40]. Centres O1,R5,R6,R10 and R11 possess C_{2v} symmetry, and are thought to be composed of vacancy chains along $\langle 110\rangle$ [39, 40]. Centres R7 and R8 along with a new centre R7a have been tentatively linked with vacancy clusters possessing C_{2v} (R8/R7a) and mon-I (R7) symmetry [40]. This is further to the trigonal R12 defect with a $\langle 111\rangle$ axis previously proposed [39].

In the case of longer vacancy chains, V$_n$, $n >$2, it is possible that the spin-density does not lie on interior atoms having dangling bonds but on the dangling bonds associated with the end atoms of the chain. If this is the case then the fine structure term D in the $S =$1 multi-vacancy centres can be used to determine the length of the chain. The interaction of two spin dipoles, separated by $(n-2)a_0/\sqrt{2}$ in V$_n$, at the ends of a $[110]$ vacancy chain is $D = 3/2(g^2\beta^2 r^{-3})$. Such an analysis indicates that O1, R7, R8, R6, R10 and R11 are due to $\langle 110\rangle$ chains of 4, 4, 4-5, 5, 6 and 7 vacancies respectively [39]. It is important to reiterate that the spin designation of R8 - R11 is not certain, but is assumed to be 1 for this analysis. If the centres were to have spin 3/2 then this assignment may be invalid. More recent work [41, 40] assigns R5, O1, R6, R10 and R11 to chains of 3, 4, 5, 6 and 7 vacancies and an additional EPR centre (KUL11) to a chain of eight vacancies. However, this analysis also assumes that the spin is localized at the chain ends and not on dangling bonds along its length.

The optical properties of annealed irradiated type IIa diamonds have been reported [39]. With heat treatments of 600° C there is a loss of the GR1 peak and a correlated growth of TH5. After the 700° C anneal there is the loss of TH5 and the growth of a continuum. The absorption spectrum simply scales the spectrum seen in unirradiated type IIa diamond and may signify graphitic or disordered regions. A comparison of the absorption produced by electron and neutron irradiation followed by heat treatments at 450° C suggests that the latter produces a set of thermally stable defects giving rise to a continuum extending right up to the band edge of diamond. These defects were attributed to disordered carbon network produced by end-of-range neutron damage.

The initial irradiation rendered the diamond opaque but an anneal beyond 1050° C revealed an absorption edge starting at 1.5 eV and cutting off at 2.2 eV with lesser features at around 1 eV. The introduction rates of the EPR centres following a 1000° C anneal are very low. For a dose of almost 10^{20} cm^{-2} electrons with energy 2 MeV, only about 10^{18} cm^{-3} O1 defects are created.

It could be expected that the centres associated with vacancy chains would appear sequentially. The dissolution of the monovacancy and divacancy would occur at similar temperatures so there would be a clear start-point for the formation of vacancy chains. With the increase of time there would be a steady appearance of larger chains. In the experiments referenced the annealing time was in the order of 1-2 hours [41], so that all possible centres would have formed by the time of observation. Hence the centres appear to form simultaneously.

A study of possible vacancy defects is necessary to examine the validity of the assignment of these EPR centres. In addition, the prevalence of vacancy defects in brown diamond means that they must be investigated to see if they could be acting as the source of brown colour.

4.2 Structures Modelled

Vacancy chains and clusters are created by the removal of atoms from the bulk diamond structure. Chains of vacancies were created by the removal of atoms in the $\langle 110\rangle$ direction as shown in figure 4.1. Candidate vacancy clusters, where not all vacancies lie along [110], were investigated from models minimizing the number of dangling bonds and are shown schematically in figure 4.2. Here the atoms are labelled numerically in order of removal. Such an approach has been used previously to describe multi-vacancy structures in silicon [97, 98].

Vacancy clusters up to 14 vacancies and chains of up to four vacancies were inserted into cubic unit cells of 216 atoms. As well as AIMPRO studies DFTB calculations were performed by N. Fujita (Universität Paderborn/Exeter University). Initial investigations for V_3 chains were carried out in cells of 64 atoms using both AIMPRO and DFTB. The formation energies in AIMPRO were calculated with a contracted basis set of 13 functions, and found to be 12.908 eV in a 64 atom cell and 13.129 eV in a 216 atom cell; a difference of 1.7%. Calculations with DFTB showed the formation energy in a 64 atom cell to be 12.607 eV, and 14.175 eV in the 216 atom cell; a difference of 11%. The large differences in formation energy with cell size indicate that a 216 atom cell is the minimum size that should be used for reliable results. DFTB calculations in 512 and 1000 atom cells show only marginal changes in formation energy from a 216 atom cell, supporting the use of 216 atoms cells for modelling. A 2^3 set of Monkhorst-Pack (MP) [67] $\mathbf{k}$-points was used to sample the Brillouin zone.

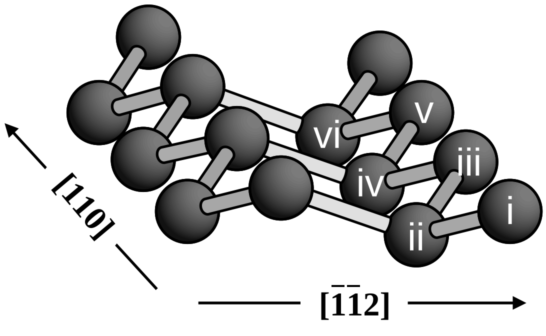


Figure 4.1: Diagram of order of removal of atoms to create the $\langle 110\rangle$ vacancy chain.

Longer vacancy chains of five and six vacancies were also placed in cells of 216 atoms,

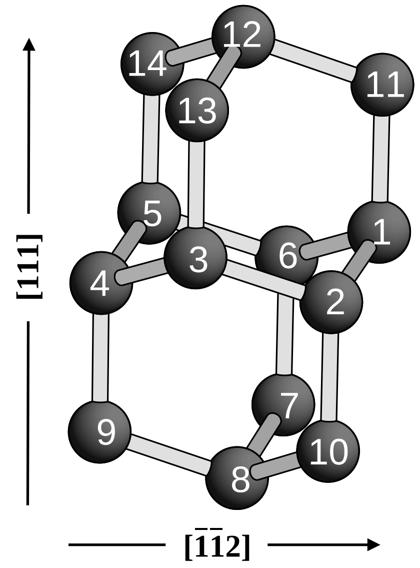


Figure 4.2: Diagram of order of removal of atoms to create vacancy clusters of up to 14 vacancies.

but these are non-cubic. These cells were longer in the [110] direction to ensure the terminating atoms were not too close to one another upon repetition. A vacancy chain of infinite length was investigated using a thin slab of diamond, whose surface normal is parallel to [110] and contains 140 atoms. This approximates the energy of a long loop of vacancies.

It has previously been reported how a distortion to C_{2_h} symmetry for the divacancy could be energetically favourable if the reduction in one-electron energies caused by such a distortion exceeds the additional strain energy [99]. The vacancy chains studied here have been subjected to inward and outward distortion to C_{2_h} (even numbers) and C_{2_v} (odd numbers) symmetry, and the results compared to undistorted models.

4.3 Results

4.3.1 Energies

Figure 4.3 shows the calculated formation energies. The formation energy per vacancy decreases with n due an increasing binding in the larger defects. However for the vacancy clusters, as found for other semiconductors, there are sets of 'magic' numbers showing exceptional stability. These occur at 6, 10 and 14 vacancies and are the same numbers as found in silicon [98] where they are attributed to bond reconstruction and hence a minimization in the number of dangling bonds. The shape of these clusters is shown in figure 4.2. V_6 in silicon has been linked with a defect found in neutron irradiated material and detected in luminescence [100]. The bond reconstruction explanation for the stability of these magic numbered vacancy clusters holds only for the 6 vacancy cluster, as explained below.

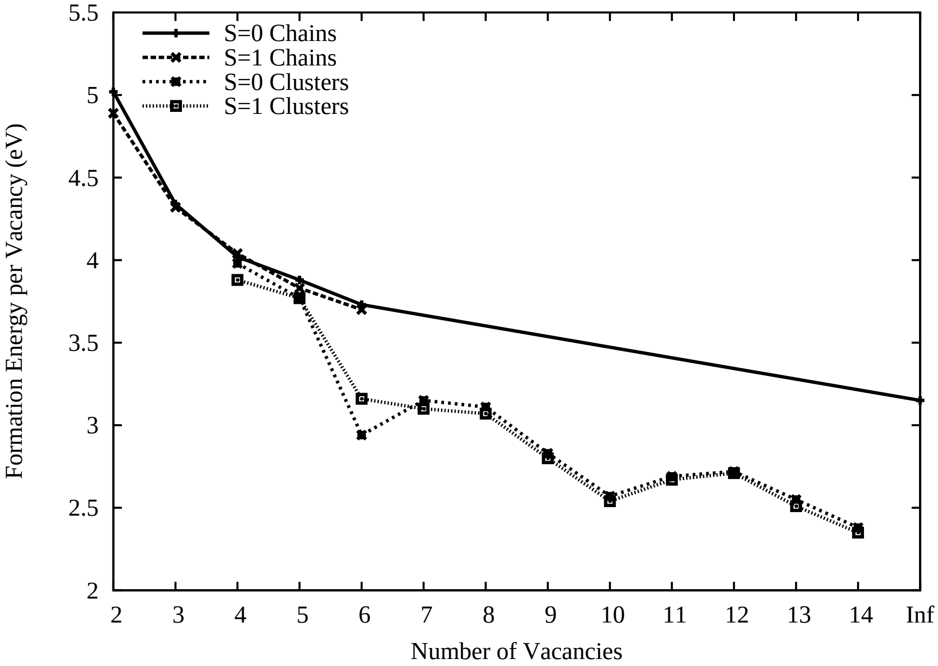


Figure 4.3: Formation energies of vacancy clusters calculated with AIMPRO.

An alternative structure for the V_4 vacancy cluster in silicon has been suggested by Makhov *et al.* [101]. Here, one, two or three self-interstitials are added to the six vacancy model of Staab, with the extra atoms being added between atoms neighbouring a vacancy. These interstitials eliminate dangling bonds and lower the energy. However, for diamond the formation energy of this structure for V_4 is 4.12 eV per vacancy which is greater than the Staab model of V_4 of 3.88 eV per vacancy. A model of 11 vacancies in an arrangement which differs from this work has been proposed for the R7 centre [40]. This model satisfies the apparent monoclinic-I symmetry of the R7 centre. It was found that the formation energy per vacancy was higher at 2.88 eV compared to 2.67 eV for the Staab model - both defects being in the $S=1$ state. This can be explained by the number of dangling bonds caused by removal of atoms in each case. In the Staab model for 11

Table 4.1: Formation energies, in eV, of defects with n vacancies calculated using AIMPRO. The lowest energy spin state is indicated.

Defect	E_{form} Lowest	E_{form}/n Lowest	Spin Lowest	E_{form} Highest	E_{form}/n Highest	Spin Highest
Divacancy	9.89	4.94	1	10.11	5.06	0
3 vacancy chain	13.13	4.37	0	13.15	4.38	1
4 vacancy chain	16.21	4.04	0	16.23	4.06	1
5 vacancy chain	19.19	3.83	1	19.44	3.88	0
6 vacancy chain	22.19	3.70	1	22.38	3.73	0
4 vacancy cluster	15.54	3.88	0	15.90	3.98	1
5 vacancy cluster	18.85	3.77	1	18.86	3.77	0
6 vacancy cluster	17.64	2.94	0	18.94	3.16	1
7 vacancy cluster	21.73	3.10	1	22.02	3.15	0
8 vacancy cluster	24.59	3.07	1	24.85	3.11	0
9 vacancy cluster	25.16	2.80	1	25.51	2.83	0
10 vacancy cluster	25.41	2.54	1	25.69	2.57	0
11 vacancy cluster	29.14	2.67	1	29.58	2.69	0
12 vacancy cluster	32.02	2.71	1	32.67	2.72	0
13 vacancy cluster	32.68	2.51	1	33.22	2.55	0
14 vacancy cluster	32.93	2.35	1	33.34	2.38	0
∞ vacancy chain	12.58	3.15	0	-	-	-

vacancies, 18 dangling bonds are left, but 19 are formed by the model for R7.

For both ⟨110⟩ vacancy chains and vacancy clusters, spin-polarized calculations were also performed to determine the most favourable spin state. These energies are also given in table 4.1, where it is seen that there are relatively small differences in the energies of $S=1$ and $S=0$ defects. It is noticeable that the V_4 vacancy chain is not as stable as the V_4 cluster. This is the same in silicon [98].

It should be noted that the prediction of the three and four vacancy chains to be at their lowest energy in a spin state $S=0$ is inconsistent with their apparent detection by EPR [41]. Detection in EPR relies on S being non-zero, so the divacancy and chains of five/six vacancies should be detectable, as should all clusters bar those of four and six vacancies. This discrepancy is unresolved as yet, but there are two main possibilities. The formation energies (per vacancy) of many of the defects are very close, but it is possible that the output error of AIMPRO is larger that the energy differences. This would lead to misidentification of the favourable spin states of the defects. It is also possible that there has been an incorrect interpretation of EPR data, and hence assignment of signals to structures.

Table 4.2: Formation energies, in eV, of chain defects of n vacancies calculated using DFTB. Calculations performed by N. Fujita at Universität Paderborn.

Defect	E_{form} 216 atom	E_{form}/n 216 atom	E_{form} 512 atom	E_{form}/n 512 atom
Divacancy	10.87	5.43	11.06	5.53
3 vacancy chain	14.18	4.73	14.41	4.80
4 vacancy chain	17.61	4.40	17.63	4.40
5 vacancy chain	-	-	20.77	4.15
6 vacancy chain	-	-	23.84	3.97
7 vacancy chain	-	-	27.58	3.94

Table 4.3: Formation energies, in eV, of cluster defects of n vacancies calculated using DFTB. Calculations performed by N. Fujita at Universität Paderborn.

Defect	E_{form} 216 atom	E_{form}/n 216 atom	E_{form} 512 atom	E_{form}/n 512 atom	E_{form} 1000 atom	E_{form}/n 1000 atom
4 vacancy cluster	16.78	4.19	16.90	4.23	16.92	4.23
5 vacancy cluster	18.50	3.70	18.56	3.71	19.43	3.89
6 vacancy cluster	18.79	3.13	18.87	3.14	18.86	3.14
7 vacancy cluster	23.02	3.29	23.16	3.31	23.17	3.31
8 vacancy cluster	24.07	3.01	24.12	3.01	24.11	3.01
9 vacancy cluster	25.66	2.85	25.21	2.80	25.21	2.80
10 vacancy cluster	25.09	2.51	25.16	2.52	25.16	2.52
11 vacancy cluster	29.75	2.71	29.63	2.69	29.64	2.70
12 vacancy cluster	31.21	2.60	31.30	2.61	31.31	2.61
13 vacancy cluster	32.31	2.49	32.42	2.49	32.43	2.50
14 vacancy cluster	-	-	32.45	2.32	32.44	2.32

These energies are to be compared to the values calculated for the same systems in DFTB, shown in tables 4.2 and 4.3. DFTB also finds that the clusters of 6, 10 and 14 vacancies are the most stable, and the cluster is more stable than the chain configuration for four vacancies. The energy values for differing size unit cells support the use of 216 atom cells to model the defects.

4.3.2 Structure

The two atoms forming the ends of the $\langle 110 \rangle$ chains possess dangling bonds and these relax back onto the plane of their neighbours thus adopting sp^2 character. This relaxation increases with the chain length. The separation of the atoms at the chain ends increases by 0.13, 0.142, 0.152, 0.165 and 0.190 Å for the divacancy and V_3 – V_6 chains respectively.

The atoms which surround the finite chains of vacancies along its length are paired with other atoms which have also been left with dangling bonds; the unrelaxed separation of these atoms is 2.499 Å. The magnitude and direction of movement under relaxation varies between the divacancy, V_3 and V_4 chains. The divacancy shows a symmetric movement outwards of 0.102 Å for the surrounding atoms. This is in contrast to the V_3 and V_4 chains where the surrounding atoms move both towards and apart from each other. In the V_3 chain two pairs move 0.094 Å apart while the other moves 0.069 Å together. In V_4 four pairs of atoms are repelled by 0.1 Å with the remaining two moving 0.119 Å closer to each other.

The ideal infinite vacancy chain along $\langle 110 \rangle$ has four atoms surrounding the line of removed atoms, arranged so that each atom with a dangling bond is 2.52 Å away from another. Upon relaxation, these atoms move closer together and form a reconstructed bond of length 2.16 Å.

When six atoms in a chair configuration are removed to form an ideal hexavacancy cluster, the distance between pairs of atoms with dangling bonds is 2.49 Å. Upon structural optimization two pairs of these atoms move closer together to form a very weak bond of length 2.30 Å. The others move apart with a separation of 2.63 Å. The neighbours to the hexavacancy lying along $\{111\}$ are pulled closer to the hexavacancy resulting in a bond length of 1.83 Å, corresponding to approximately 19% strain. It can be seen in figure 4.2 that the 10 and 14 vacancy clusters consist of closed ‘cages’ of vacancies adjoining the hexavacancy. Formation energy calculations predict that these are especially stable defects. Atoms with dangling bonds in the 10 vacancy defect exhibit a larger movement than those in the six vacancy cluster. The resultant separation is 2.26 Å, apart from two atoms which remain unpaired and move away from the vacancy group by 0.26 Å. To allow this movement to occur several bonds are stretched to 1.81 Å, a strain of just less than 18%. As a larger amount of atoms are removed for the 14 vacancy cluster, the separation of atoms with dangling bonds is increased. There are just four pairs of atoms that display signs of rebonding occurring, moving closer together with a resultant separation of 2.32 Å. Atoms surrounding the vacancy move inwards with a range of lengths of movement. The bonds of four of the surrounding atoms are stretched, but to a lesser degree that that seen in the 6 and 10 vacancy clusters; the bonds being stretched to 1.77 Å, a strain of 15%.

As reported above, the predicted stable cluster sizes are 6,10 and 14 vacancies, but only the 6 vacancy cluster exhibits full bond reconstruction. The reason for this is clear if the structure of the 10 and 14 vacancy clusters is considered, both being based on the 6 vacancy ring. The 6 vacancy ring has atoms with dangling bonds close enough together to allow reconstruction, but to make larger clusters, some of these must be removed. Thus there are atoms which are too far apart to reconstruct and the appearance of filled bandgap energy levels is linked to this. For the 10 vacancy cluster, those atoms which rebonded in the 6 vacancy ring and still have a close dangling bond become even closer, with resultant separation of 2.26 Å. To allow optimization of energy several bonds of atoms surrounding the vacancies have to stretch to 1.81 Å, slightly shorter than in the 6 vacancy ring. The 14 vacancy cluster has all the atoms which were bonding partners for atoms surrounding the 6 vacancy ring removed. Therefore it is not obvious by simply analysing the optimized structure where rebonding – if any – occurs. It is possible to find further explanation for the stability of 6,10 and 14 vacancy clusters by considering the number of dangling bonds formed by removal of each atom. The formation of 7 and 11 vacancy clusters by removal of an atom causes 2 more dangling bonds, and removal of any atom adjacent to the 14 vacancy cluster creates 3 further dangling bonds. However, the creation of 6,10 and 14 vacancy clusters from smaller clusters does not require the creation of further dangling bonds.

4.3.3 Electronic properties

Assignment of centres to the family of $\langle 110\rangle$ vacancy chains has assumed that the spin density is localized on the atoms terminating the chain. It is possible to deduce the wavefunction localization by performing a Mulliken analysis on the optimized structures, and this has been calculated for all the $\langle 110\rangle$ vacancy chain defects. In no instance is it observed that the spin density is exclusively localized on the terminating atoms of the vacancy chains.

The three vacancy chain has the spin density located on three pairs of atoms, separated by 2.45 and 2.65 Å. In the four vacancy chains, just two pairs of atoms have the spin located on them, both having an atom separation of 2.64 Å. This trend of pairs of atoms differs for the five vacancy chain, here three atoms have spin density upon them. One of these three is a terminating atom of the chain, and the other two are nearest neighbours to the vacancy at the opposite end of the chain, with a separation of 2.63 Å. The six vacancy chain exhibits spin density located on the terminating atoms of the chain but also on four further atoms, which are arranged in nearest neighbour pairs to the end vacancies of the chain. The pairs are at opposite ends of the chain, with an atom separation of 2.56 Å. Assumptions in other work [41, 40] about the spin density being located at the chain end is unsupported by this analysis, as is the assignment of defects based on this assumption.

In the consideration of a defect's effect upon the optical properties, it is desirable to know what transitions are possible between energy levels as this will indicate what optical absorption takes place. Previous measurements of optical absorption on natural brown diamond have shown the existence of a broad and largely featureless absorption band, with the absorption coefficient increasing with photon energy [30, 44]. It is expected that if the multi-vacancy defects studied here are related to brown colour we will see a range of energy levels throughout the band-gaps of these defects. This range of energy levels allows a broad distribution of transitions and hence absorption. The Kohn-Sham levels for the defects considered here are shown in figures 4.4–4.8. The Kohn-Sham levels are scaled by a factor 1.31 to compensate for the underestimation of the bandgap by LDA and are taken at the gamma point. In the case of defects in an $S=1$ spin state, two spectra are produced.

It is seen in figure 4.4 that the six vacancy cluster has no occupied bandgap states, indicating that it is an $S=0$ defect and should exhibit reconstruction of any dangling bonds. The divacancy is more stable in a $S=1$ state. This can be understood from the energy levels, shown in figures 4.4 and 4.6, and described in more detail below. The C_{2h} divacancy symmetry leads to a configuration occupying levels in the lower half of the band gap.

The divacancy having C_{2h} symmetry leads to both filled and empty levels in the band gap. The lower is occupied with two electrons. The optical absorption band TH5 linked with divacancies has a threshold at 2.45 eV. The Kohn-Sham levels for the lower energy $S=1$ state are shown in figure 4.6, and it can be seen that two lowest transitions are of 0.96 and 2.14 eV. The other defects have a range of threshold absorption energies, from 0.10 eV for V_{12} to 2.28eV for V_6.

It is interesting to note that the band-structures of the various defects studied are similar, despite the greatly differing structures. All defects studied showed several deep band-gap energy states, which were grouped together in, or below, the middle of the band-gap. It is seen in figures 4.4 – 4.8 that as the clusters get larger the groups of empty energy states become nearer to E_c, although there are still shallow acceptor levels. The exact number of states varies between structures, depending on the number of dangling bonds left by the removal of atoms. For all defects possessing terminal dangling bonds the deep gap levels are occupied. However, for the infinite chain and hexavacancy where there are no terminal dangling bonds the occupied states are resonant with bulk states.

The 6 vacancy cluster shape allows a high reconstruction of the dangling bonds [102]. This means that there are no unpaired electrons and hence this defect has a spin of $S = 0$. Conversely, the other defects studied have non-zero spin, due to the presence of unpaired electrons. These unpaired electrons arise in the other clusters and finite length $\langle 110 \rangle$ chains due to the structure not allowing dangling bonds to reconstruct. These give rise to filled band-gap energy levels, which in turn dictate allowable energy transitions.

4.3.4 Optical properties

The small range of band-gap energy levels displayed by the multivacancy defects is expected to give rise to an absorption peak, at an energy below that of the absorption onset at $\sim$5.5 eV from the diamond indirect band-gap.

Studies of the optical transitions have been undertaken by modelling the Electron Energy-Loss Spectroscopy (EELS) spectrum. This is described in detail in section 2.6.5. This technique only takes into account direct transitions at any particular **k**-point, so the absorption rise at 5.5 eV is not reproduced. The energies at which transitions occur must be scaled due to the DFT bandgap underestimation. Transitions are subjected to a Gaussian broadening of 0.8 eV.

The expected absorption peak is clearly seen for selected defects in figure 4.9.

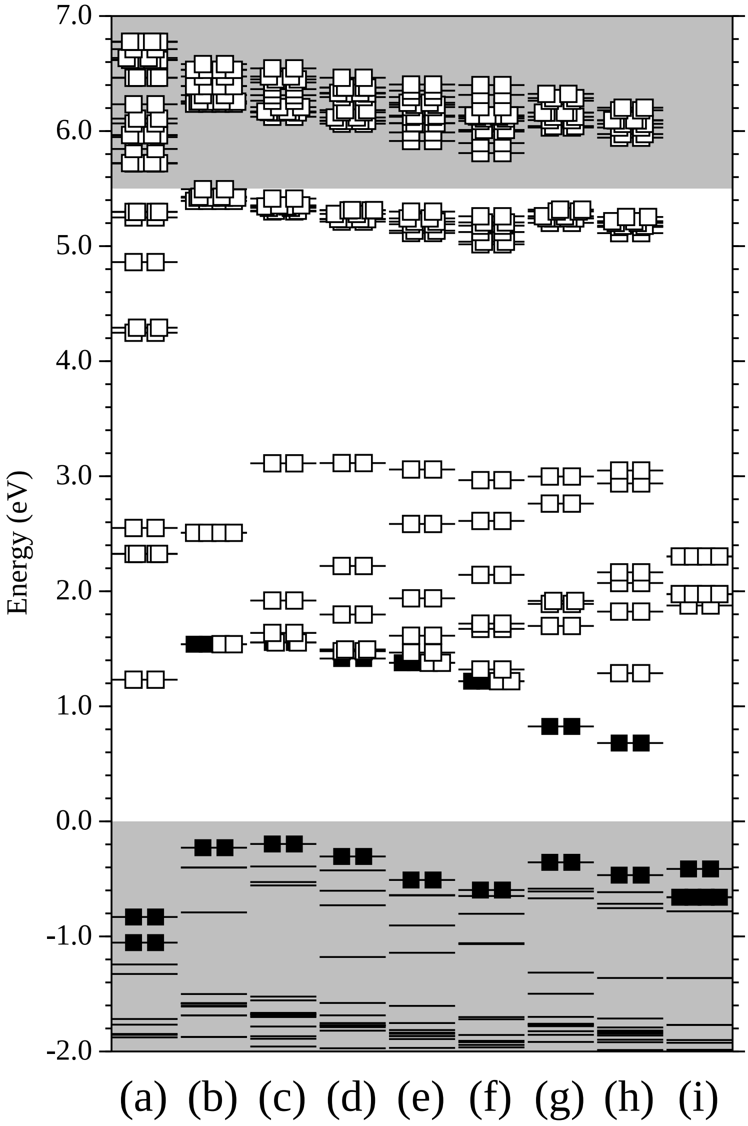


Figure 4.4: The Kohn-Sham levels (scaled) of various multivacancy defects, taken at Γ. (a) Infinite vacancy chain (b) Divacancy (c) 3 vacancy chain (d) 4 vacancy chain (e) 5 vacancy chain (f) 6 vacancy chain (g) 4 vacancy cluster (h) 5 vacancy cluster (i) 6 vacancy cluster. The filled boxes denote occupied states and the empty boxes denote empty states. The shaded regions indicate the position of the bands in bulk diamond.

4.3.5 π-bonded carbon in multivacancies

The $\{111\}$ surface in diamond can re-construct into the Pandey π-bonded chain formation [103], which has been shown to be energetically favourable [104] This surface structure gives rise to π and π^* bandgap levels and is described in section 6. The energy

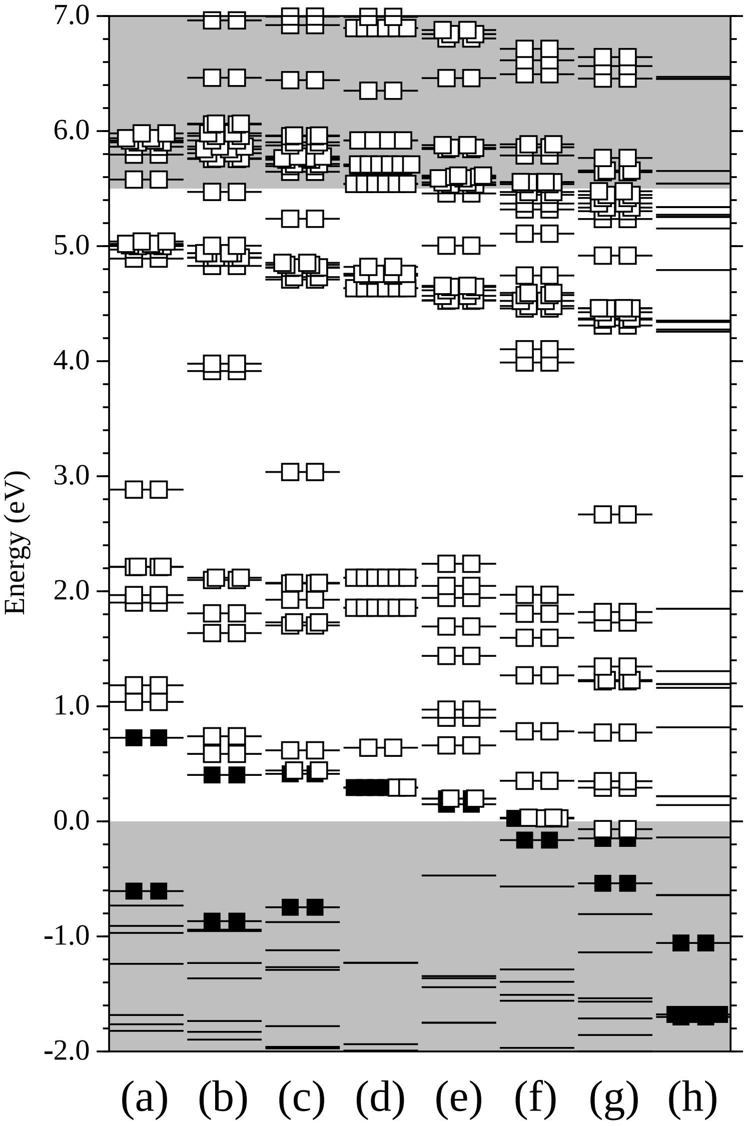


Figure 4.5: Further Kohn-Sham levels of multivacancy defects, taken at Γ. (a) 7 vacancy cluster (b) 8 vacancy cluster (c) 9 vacancy cluster (d) 10 vacancy cluster (e) 11 vacancy cluster (f) 12 vacancy cluster (g) 13 vacancy cluster (h) 14 vacancy cluster. The filled boxes denote occupied states and the empty boxes denote empty states. The shaded regions indicate the position of the bands in bulk diamond.

transitions between these levels have a range from zero to the bandgap energy [104]. This indicates that the optical absorption properties of π-bonded carbon may be comparable to that observed in brown diamond [34]. If multivacancy defects contained such types of carbon then their experimental detection in natural brown diamond [35] could be associated with the colouration.

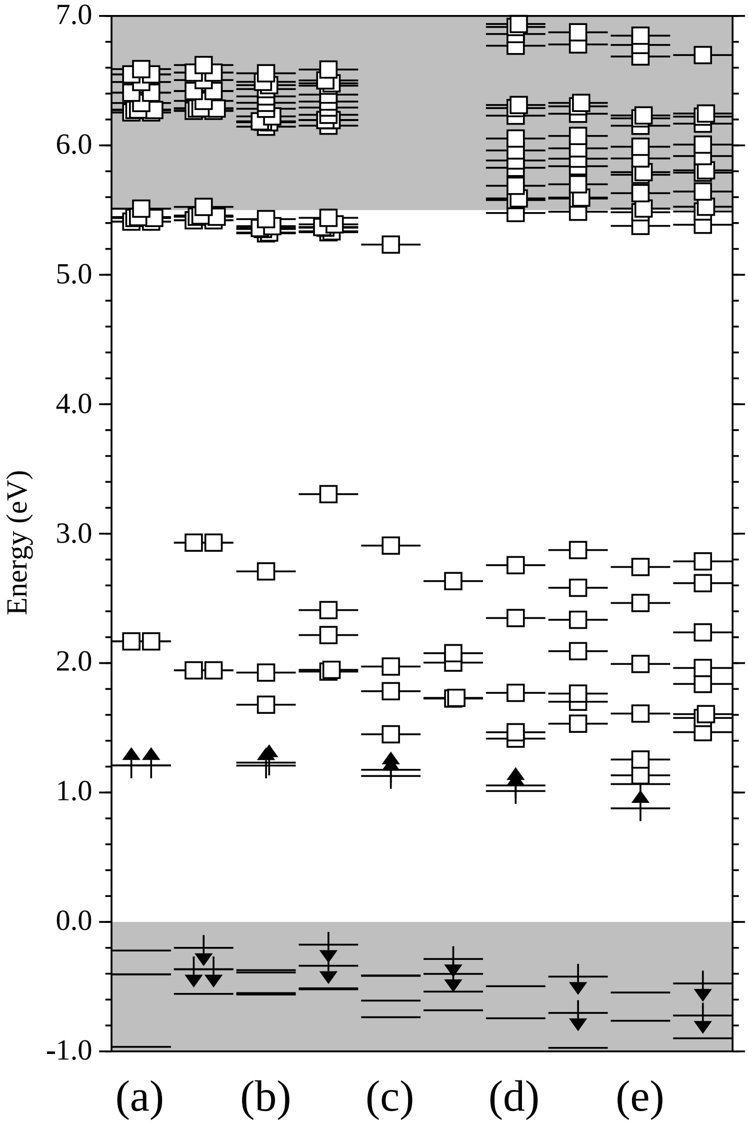


Figure 4.6: The Kohn-Sham levels (scaled) of multivacancy chains in $S = 1$ spin state, taken at Γ. The spin-up and spin-down electrons are separated for each defect. (a) Divacancy (b) 3 vacancy chain (c) 4 vacancy chain (d) 5 vacancy chain (e) 6 vacancy chain. The arrows denote occupied spin up/down levels and the empty boxes denote empty levels. The shaded regions indicate the position of the bands in bulk diamond.

Models were constructed for the $\langle 110 \rangle$ vacancy chains to attempt to introduce sp^2 bonded carbon atoms, while still retaining the symmetry of the defect. When analysed the total formation energies were considerably higher than those of the original models. Two adjacent chains of vacancies in the $\langle 110 \rangle$ direction were removed from the 140 atom cell used in the infinite vacancy chain, to attempt to introduce π-bonding. As with the other studies, the formation energy was larger.

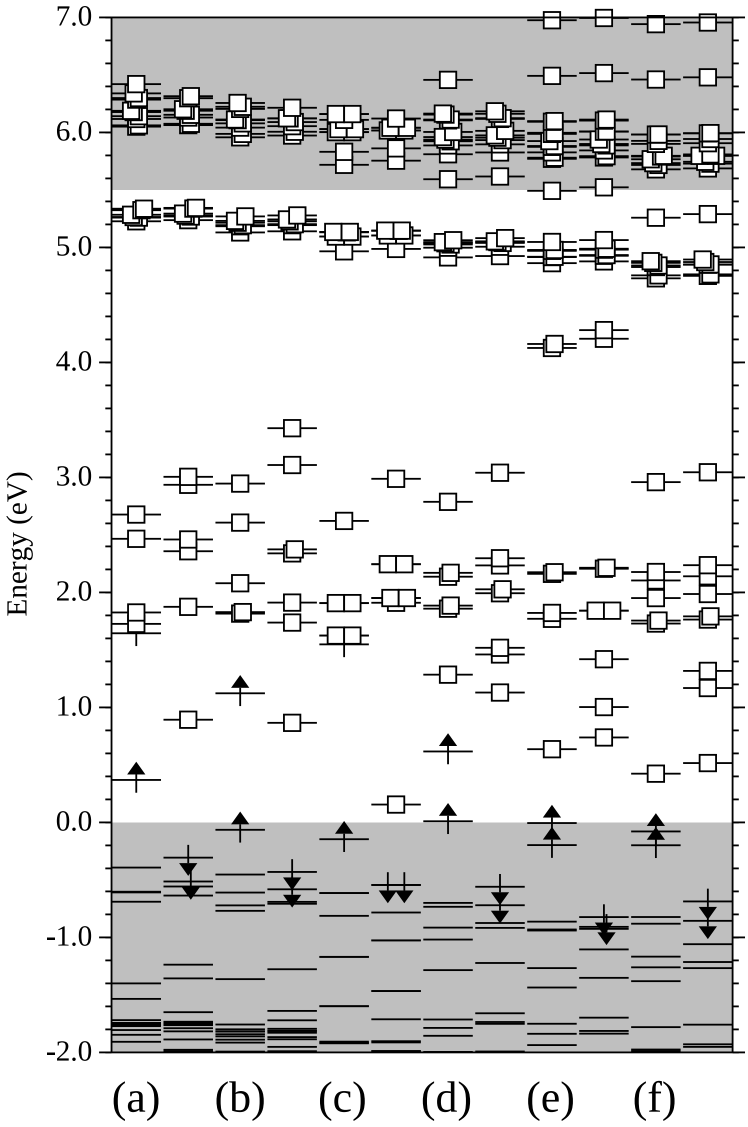


Figure 4.7: Kohn-Sham levels (scaled) of multivacancy clusters in $S = 1$ spin state, taken at Γ. Spin-up and spin-down electrons are separate. (a) 4 vacancy (b) 5 vacancy (c) 6 vacancy (d) 7 vacancy (e) 8 vacancy (f) 9 vacancy. The arrows denote occupied spin up/down levels and the empty boxes denote empty levels. The shaded regions indicate the position of the bands in bulk diamond.

It is possible to introduce π-bonding into diamond by removing every third atom in the $\langle 110 \rangle$ direction, the dangling bonds left on the pairs of adjacent atoms should be able to re-construct to give them sp^2 nature. Analysis of such a defect was made in a 188 atom unit cell, similar in construct to the cell used to model the infinite $\langle 110 \rangle$ vacancy chain. Here though, the cell was elongated in the $\langle 110 \rangle$ direction to allow correct reproduction of the defect when the cell is repeated during analysis. The formation energy of this

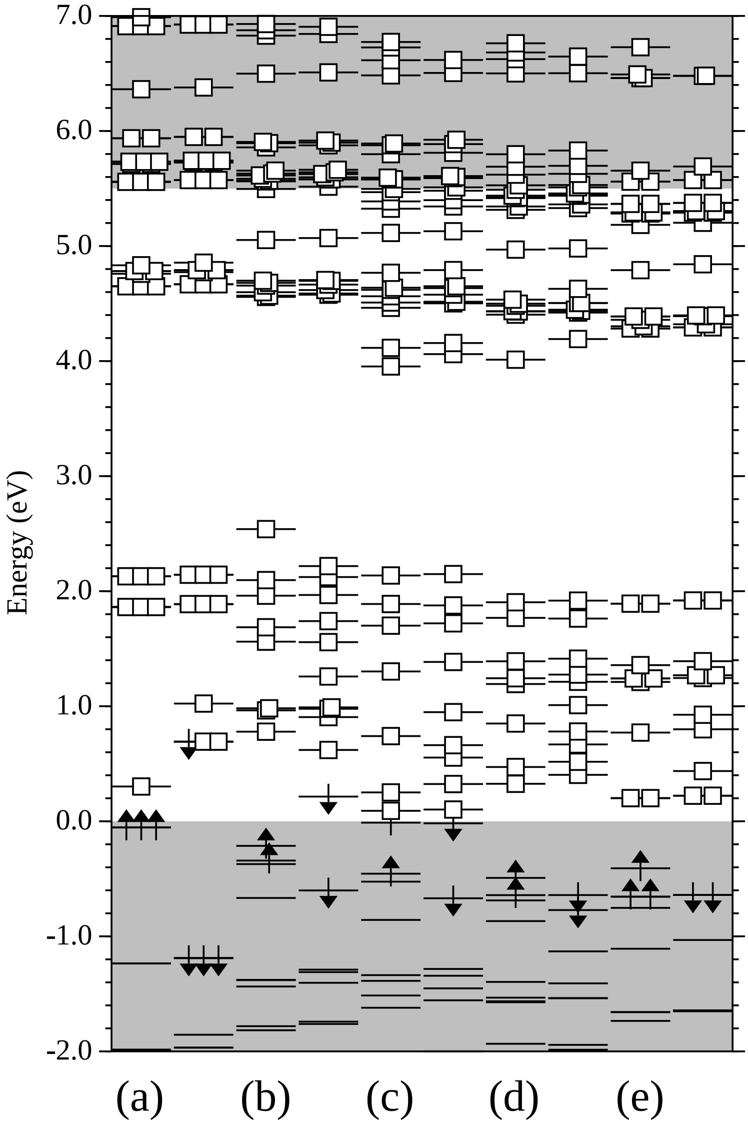


Figure 4.8: Further Kohn-Sham levels of multivacancy clusters in $S=1$ spin state, taken at Γ. Spin-up and spin-down electrons are separate. (a) 10 vacancy (b) 11 vacancy (c) 12 vacancy (d) 13 vacancy (e) 14 vacancy. The arrows denote occupied spin up/down levels and the empty boxes denote empty levels. The shaded regions indicate the position of the bands in bulk diamond.

defect was found to be 20.85 eV, which equates to 6.95 eV per vacancy; over double the formation energy of the infinite vacancy chain. The results of these preliminary studies suggest that π-bonded carbon cannot be a feature of $\langle 110 \rangle$ vacancy chains or clusters in diamond.

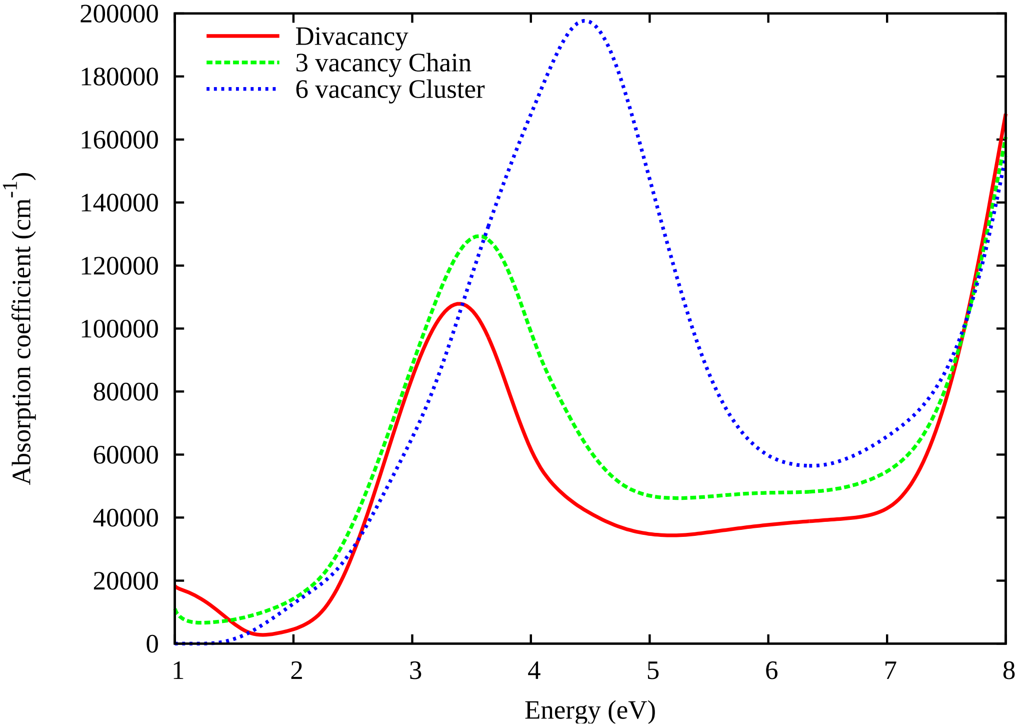


Figure 4.9: Absorption spectra for selected multivacancy defects. The absorption peak caused by the group of levels in the midgap is evident. Energy values are scaled by 1.3 to account for bandgap underestimation.

4.4 Conclusion

Studies were carried out on ⟨110⟩ orientated vacancy chains and clusters of vacancies. It is seen that for four, five and six vacancies the cluster formation is more energetically favourable than the ⟨110⟩ chain, and it can be speculated that this will be the case for larger numbers of vacancies also. The most stable sizes were the 6, 10 and 14 vacancy clusters, which all consist of closed structures minimising the number of dangling bonds and thus reducing the energy.

Calculated absorption onset for the divacancy is 2.14 eV, this is in fair agreement with the assignment of the $S = 1$ C_{2h} R4/W6 centre to this defect, and the absorption band at 2.54 eV.

The energies of defects calculated indicate that there will not be a range of optical absorption, but rather a distinct peak of absorption in the 3.5 – 4.5 eV range. Positron annihilation studies on brown diamond before and after annealing have shown the presence, and then loss, of vacancy defects [35]. This has lead to suggestions of a link between multivacancy structures and brown diamond. The absorption spectrum of brown diamond is broad and featureless, with an approximate square or cube dependence [34] on energy. It is clear that the defects studied are not able to account for such an absorption spectrum, so it must be concluded that small multivacancy structures are not the cause of brown colouration in diamond.

Chapter 5

EELS of Bulk Diamond and Dislocations

5.1 Introduction

This work has been published in part in *Philosophical Magazine* **86**, 4757 (2006).

As described in section 3.2, low-loss Electron Energy Loss Spectroscopy (EELS) offers the possibility of revealing the local electronic structure of extended defects in semiconductors or insulators with $\sim$10 nm spatial resolution and with energy resolution down to 0.3 eV. Diamond is an ideal prototype material to investigate by low-loss EELS as the band gap of 5.5 eV is large enough to minimise the influence of the zero-loss peak. Any defect having levels in the bandgap is of immediate interest given the importance of colouration in diamond for the gem industry.

Low-loss EELS involves measuring the loss of energy of a electron beam in promoting transitions between occupied and empty levels where the energy separation is less than about $\sim$50 eV. Core-loss EELS is due to transitions between core levels and empty levels, and is much harder to investigate by modern theoretical methods. It requires the detailed knowledge of the core wavefunctions which are not available to calculations using pseudopotentials as the core electrons are removed.

It has been shown [96, 105] that brown diamond exhibits a different EEL spectrum to colourless diamond. Low-loss studies do not show a clear relationship between dislocation cores and enhanced density-of-states (JDOS) which is a measure of available transitions, but rather an increase in sp^2 content in the vicinity of dislocations and stacking faults. This suggests the investigation of extended defects as a possible source of brown colour is required. An increase of the EEL signal in the energy region around 6 eV at dislocations in brown diamond has been noted, compared to the colourless variety [106]. These investigations have been extended to regions remote from dislocations [96] and it is observed that there is a generally higher ‘background’ in states giving rise to intensity at 6 eV in brown diamond than in colourless diamond. This is linked to sp^2 related defects.

5.2 Bulk Diamond

5.2.1 Calculated results

In order to understand the change, if any, that defects have on the EELS spectrum of diamond, a complete calculation of the properties of pure bulk diamond is necessary. Diamond is modelled in a minimal two-atom cell, repeated along the FCC lattice vectors, with the dielectric function calculated as described in section 2.6.5. The variation of the real part of the dielectric function with the number of Brillouin zone sampling points is shown in figure 5.1. The variation with the number of basis set functions and the polynomial broadening of data points is illustrated in figures 5.2 and 5.3 respectively. From these figures it can be seen that a Monkhorst-Pack grid of at least 20^3 points is required, with 22 functions in the basis set. A polynomial broadening of 0.8 eV is used to smooth the points into a easily readable curve.

Figure 5.4 shows the real and imaginary parts of the dielectric function of diamond compared with experimental data [107, 108]. The experimental spectrum for ε_2 shows a threshold around the direct band gap at 7 eV and a peak at 12 eV due to band-band transitions. It then falls slowly with increasing energy. Unusually, and in contrast with other semiconductors, diamond does not exhibit additional sharp peaks in this region, due to its bandstructure.

The theoretical curve displays a threshold at 5.4 eV, and a peak around 12 eV, which are shifted from the experimental values due to the underestimation of the gap and lower

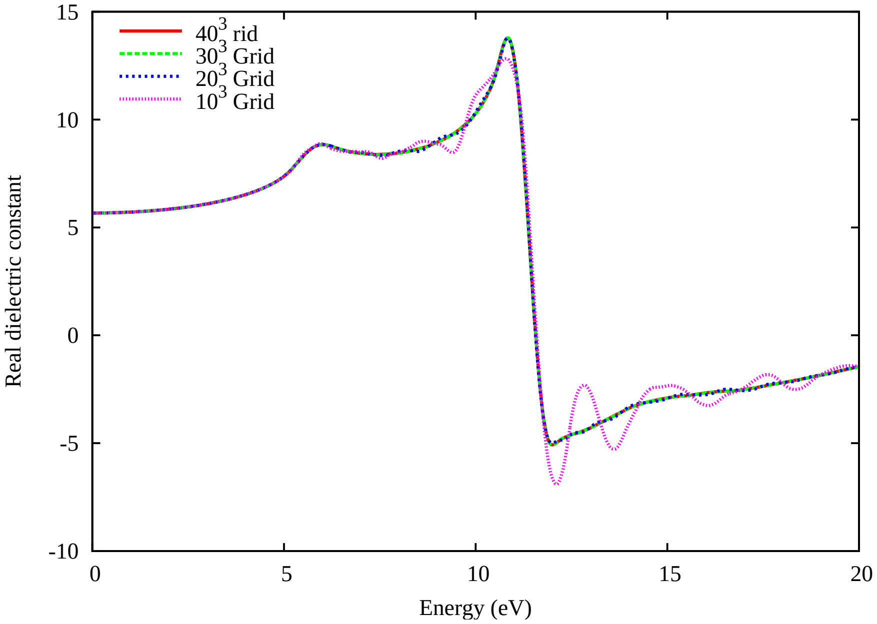


Figure 5.1: Variation of the real part of the dielectric function of bulk diamond with the Monkhorst-Pack sampling grid. 22 functions were used in the basis set, with a broadening of 0.8 eV.

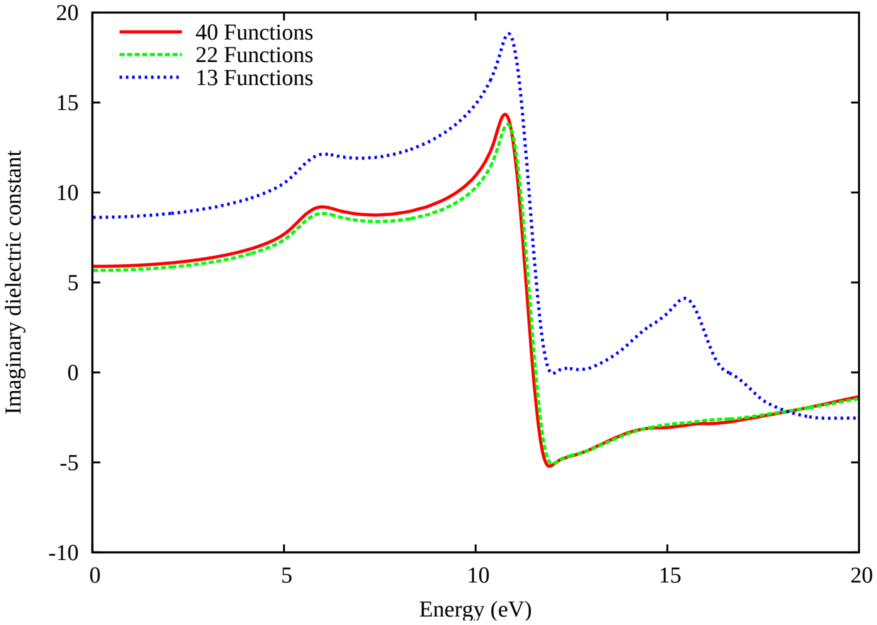


Figure 5.2: Variation of the real part of the dielectric function of bulk diamond with the number of functions in the basis set. A 40^3 grid of sampling points was used, with a broadening of 0.8 eV.

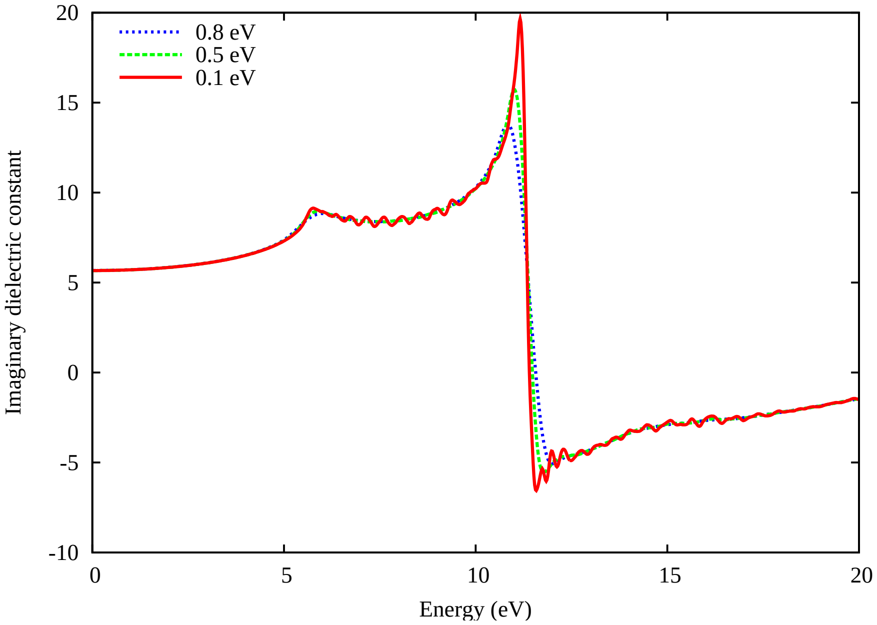


Figure 5.3: Variation of the real part of the dielectric function of bulk diamond with the polynomial broadening applied. A 40^3 grid of sampling points was used, with 22 functions in the basis set.

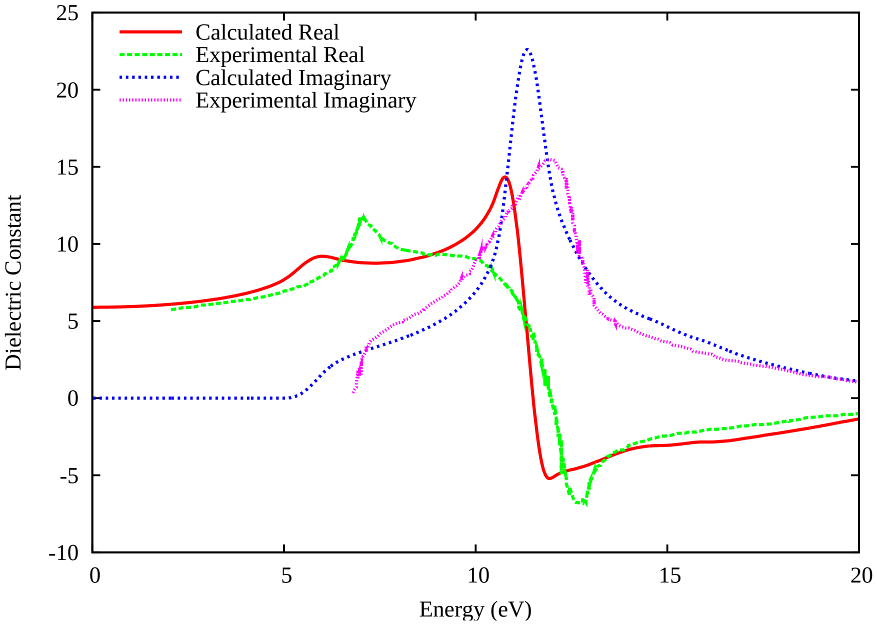


Figure 5.4: Real and imaginary parts of the calculated dielectric function for bulk diamond compared to experiment. Experimental values from [107].

conduction bands by LDA. It has been described when modelling diamond using AIMPRO how both the direct and indirect gap can be bought into agreement with experimental values, by scaling the levels from the valence band top upwards by a factor of 1.3. This technique has been applied to the defect related gap absorption in chapter 6. The method of scaling energy levels is quite distinct from a rigid energy shift, also known as a 'scissors' shift. Here all levels above the valence band top have the same value shift added, but this has the effect of changing the relative position of gap levels and hence related energy transitions. Due to this it is felt unsuitable for this work.

The real part of the dielectric function is also shown in figure 5.4 where it is compared with experiment. Note that in the long wave-length limit, the calculated value of $\varepsilon_1 =$5.89 agrees well with the experimental value of the static dielectric function of 5.82 [107]. The experimental peak at 7 eV is much sharper than the calculated one due to the neglect of excitonic effects. However the energy where $\varepsilon_2 =$0 is in good agreement.

The calculated and experimental absorption coefficients are shown in figure 5.5. The magnitude of the absorption agrees well with experiment but the threshold around 7 eV clearly reflects the underestimation of the direct gap by the theory. The peak in absorption at $\sim$12.5 eV is related to the vanishing of ε_1 and is reproduced in the theoretical spectrum. However, the sharp rise of the absorption near the threshold is not reproduced due to the lack of excitons in the theory, and the absorption is $\sim$13% too large in the region beyond the peak.

5.3 Experimental Results

Figure 5.6 shows the experimental EEL spectrum for diamond, measured by U. Bangert (University of Manchester). The apparent electron energy loss in the bandgap energy region is due to background noise. This is normally accounted for by a subtraction method, but this has not been completed here. In addition the x-axis range commences at 1 eV as there is a 0 eV singularity which arises from elastically scattered electrons, the so-called zero-loss peak. The peak around 35 eV is due to plasmons with frequency $\sqrt{(ne^2/\varepsilon_0 m)}$ where the valence electrons with density n behave collectively. The calculated EEL spectrum for bulk diamond shown in figure 5.7 shows excellent agreement with experimental data. The calculated peak occurs at 35.6 eV, with the experimental peak at 33.5 eV. No

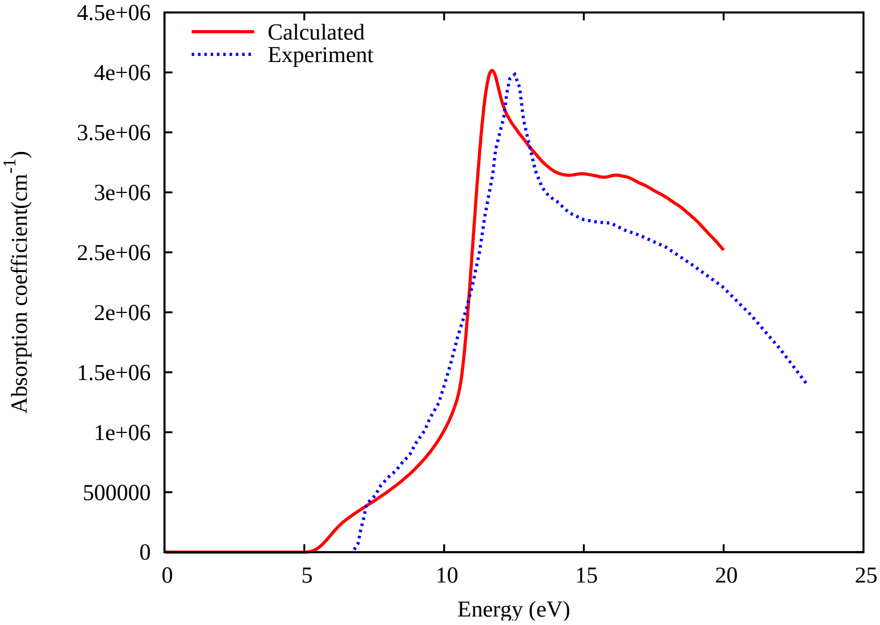


Figure 5.5: Calculated and experimental absorption coefficient of bulk diamond. Experimental values from [107].

compensation for LDA bandgap underestimation is used as the free-electron approximation becomes more accurate at these higher energies.

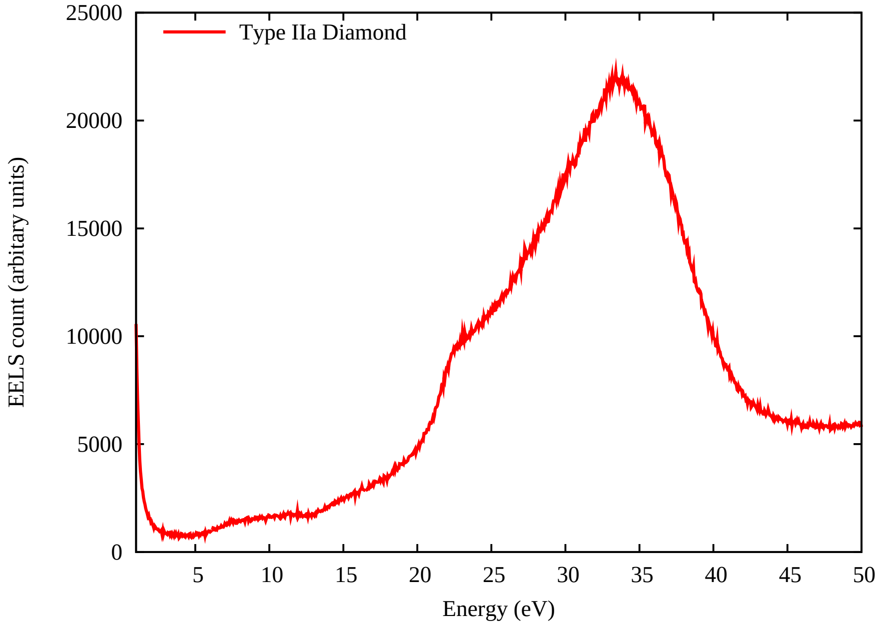


Figure 5.6: Experimental EELS of colourless diamond. The bandgap-energy counts arise from background noise which is not corrected for.

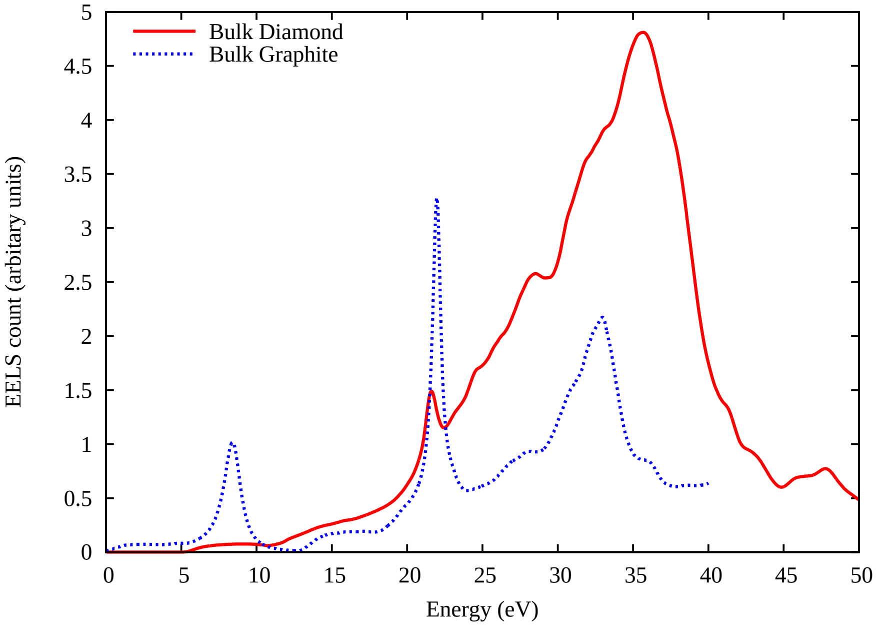


Figure 5.7: Calculated low-loss EEL spectra for bulk diamond and graphite.

The EEL spectrum for graphite has also been calculated and is shown with the calculated diamond spectrum in figure 5.7. It is clear that graphite has energy transitions down to zero, which is to be expected from its semi-metallic nature. There are peaks centred about 8 eV and 33 eV which agree well with those seen experimentally around 7 eV and 31 eV [109, 110].

If the raw EEL spectra of brown and colourless diamond is compared to measurements on a graphitic region of a brown type IIa diamond (figure 5.8), an increase in EELS count around 7 eV is seen for the brown diamond, with a much stronger increase for the graphitic region. This correlation indicates that the defect causing the brown colour must have some graphite-like component *i.e.* dangling bonds or sp^2 bonded carbon, which allows many electrons to be free to interact optically.

5.4 Dislocations

Dislocations are found in all types of natural diamond, with densities of up to $\sim 10^9$ cm^{-2}. They have been modelled previously [111], and some were found to have a bandstructure which could be consistent with a broad absorption. The absorption and relative stability

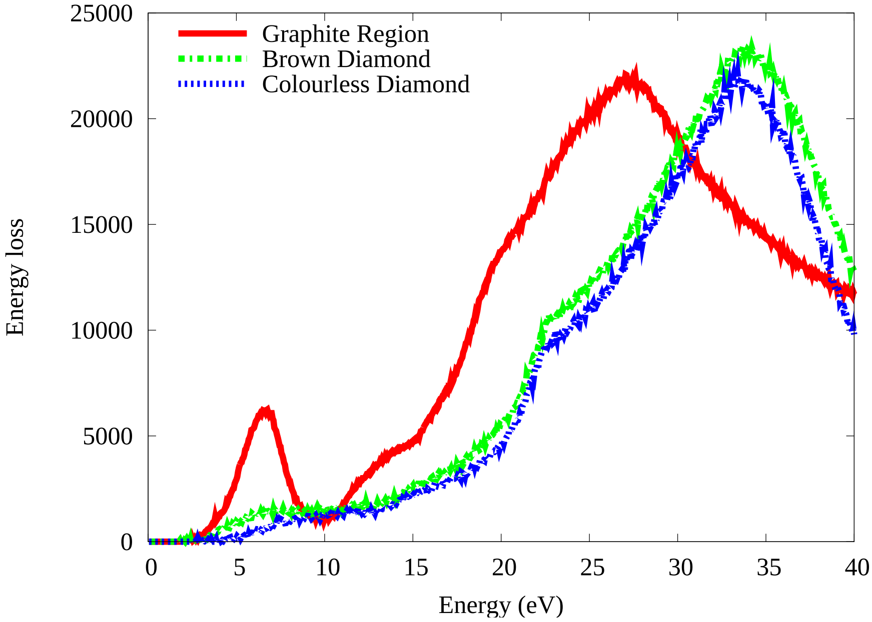


Figure 5.8: Raw EEL spectra of colourless diamond, brown diamond at fringes of graphitized region and within graphitic region. Measured by U. Bangert (University of Manchester).

of such dislocations is now described for two types of partial dislocation. These are the 90° glide (single period) and 90° shuffle (vacancy type) dislocations [27]. Both lie along $[1\bar{1}0]$ and are periodic with a repeat distance of $a/\sqrt{2}$. Both types border intrinsic stacking faults. The glide partial has only reconstructed bonds which do not introduce states into the gap. However, the shuffle dislocation has a line of dangling bonds and can be formed from a glide partial by removing a line of atoms at the core of the latter.

5.4.1 90° Glide dislocation

The 90° glide (single period) dislocation, hereafter referred to as the glide dislocation, is modelled in a cell of 440 atoms with a pair of glide partials inserted. The core construction is shown in figure 5.9. The pairing of partials is necessary as each dislocation introduces a Burgers vector which must be matched by another with equal magnitude but opposite direction. This requires the second dislocation to be inverted with respect to the first.

The reconstructed core bonds of the glide dislocation have a length of 1.65 Å which is 7% longer than those of bulk diamond. The formation energies of the two types of

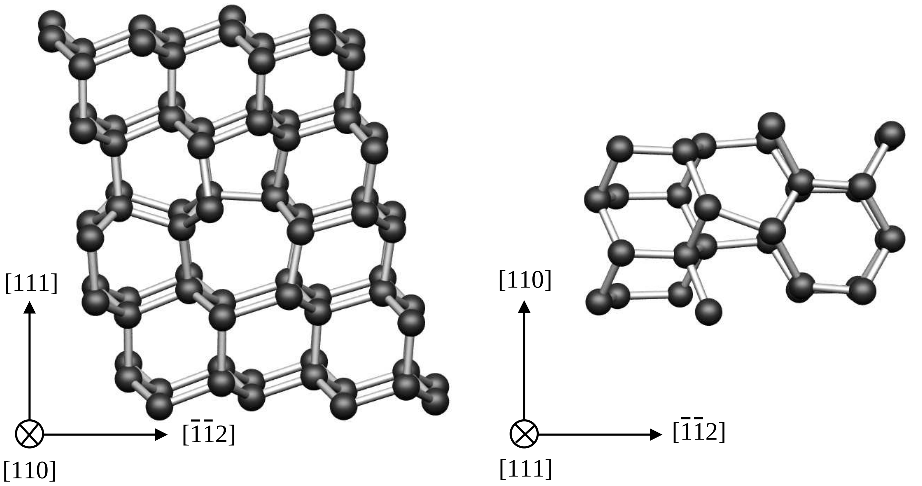


Figure 5.9: Structure of the 90° glide (single period) type dislocation. The reconstruction across the core can be seen, leading to a reduction in energy and eliminating bandgap states.

dislocations are calculated as follows. Defining core atoms as atoms that are no longer tetrahedrally bonded in the diamond structure, then there are four core atoms per repeat length ($a_0/\sqrt{2}$) for the 90° glide dislocation, and four core atoms per repeat length for the 90° shuffle type. The formation energy of the glide dislocation is then found to be 9.03 eV per core atom.

Figure 5.10 shows the bandstructure of the glide dislocation. The glide partial has an almost totally clear bandgap with no electrons available for optical interactions, indicating that optical absorption will not occur until in excess of 4 eV.

Analysis is now extended to the EEL and absorption spectra. The large size of the cell means the electon wavefunction is less dispersed and fewer **k**-points are needed for accurate sampling. Here 216 points are used compared to nearly 6000 for bulk diamond in a 2 atom cell. A broadening of 0.8 eV is used, and the EELS count is scaled for clarity by a factor of 1×10^7. The resultant spectra are shown in figure 5.11. It is clear that the glide dislocation has little bandgap absorption or EELS count, as expected from the lack of any shallow or deep bandgap levels. The absorption onset is around 4 eV, as predicted from the energy gap.

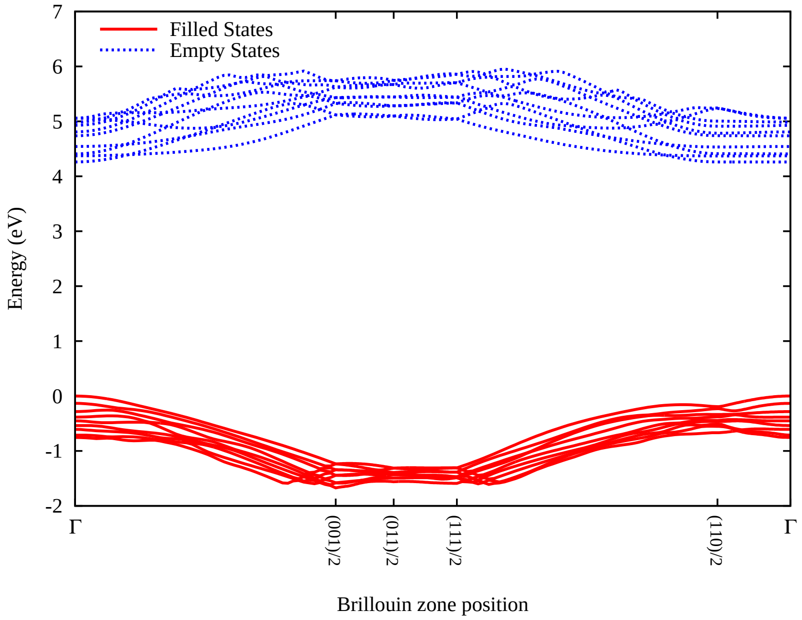


Figure 5.10: Bandstructure of the 90° glide type dislocation. The clear bandgap indicates that no absorption is expected until ∼4 eV. Empty states have had their energies multiplied by 1.3, and the valence band top is set to 0 eV.

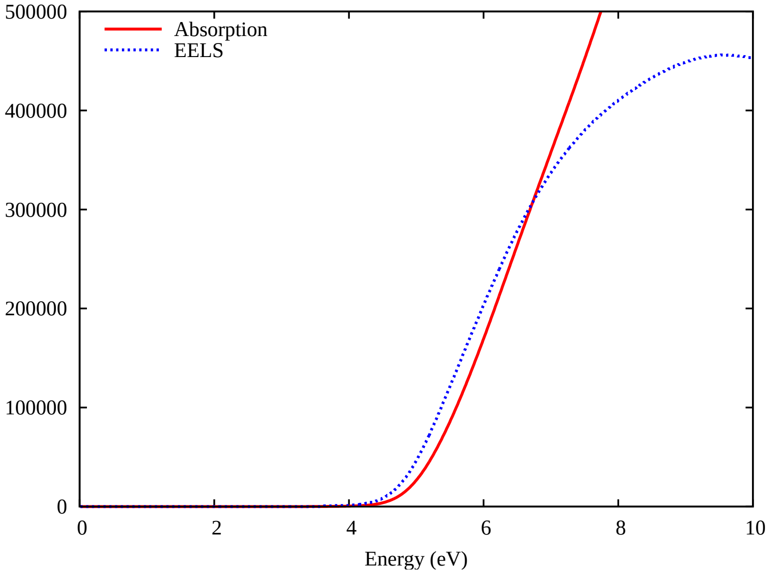


Figure 5.11: EELS spectrum and absorption of 90° glide type dislocation, using unscaled energy values. As expected from the bandstructure there is no bandgap EELS count or absorption.

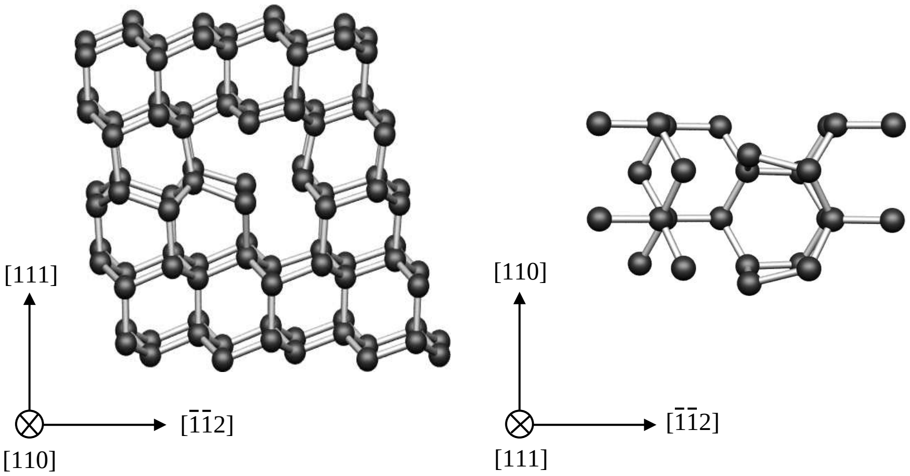


Figure 5.12: Structure of the 90° shuffle (vacancy) type dislocation. Here there are dangling bonds along the dislocation core leading to available bandgap transitions.

5.4.2 90° Shuffle dislocation

The 90° shuffle (vacancy type) dislocation, hereafter referred to as the shuffle dislocation, is modelled in a cell of 448 atoms. As with the glide dislocation model a pair of dislocations is required, but here the shuffle dislocation is paired with a glide type. The advantage of using a mixed pair of types is that since the glide partial does not lead to dangling bonds, the electronic levels lying in the gap will only arise from the shuffle partial. Figure 5.12 shows the core structure of the shuffle dislocation, with the unreconstructed core atoms leading to dangling bonds.

In contrast to the glide dislocation, there is an asymmetric separation across the core of the shuffle dislocation, between one atom having sp^2 co-ordination and one with only two bonds. The separations are 1.80 Å or 16% elongation, and 1.77 Å or 15% elongation. Therefore many dangling bonds are present, leading to a high formation energy and to a wide range of optical transitions. As before, there are four core atoms for the shuffle dislocation, with the formation energy found to be 16.03 eV per core atom, much higher than the glide partial. This demonstrates the increased stability for the glide dislocation, indicating that many of the dislocations in diamond are likely to be of the glide type. However, the shuffle type could arise as a grown-in defect or when vacancies are present and are trapped at the dislocation core.

The bandstructure of the shuffle type dislocation shown in figure 5.13 reveals a half-

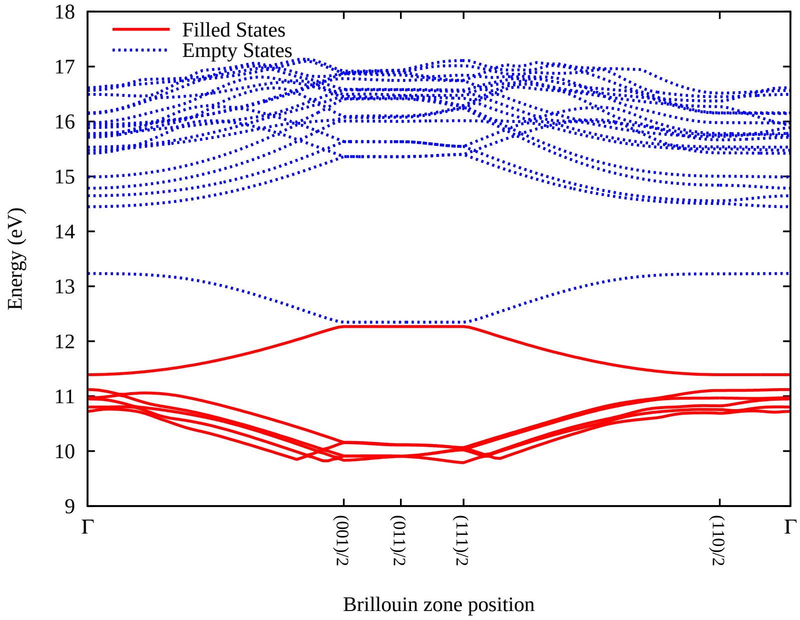


Figure 5.13: Bandstructure of the 90° shuffle type dislocation. There is a range of transitions from 0 to ∼2.5 eV, so a broad low-energy peak in EELS/absorption is expected. Energy levels above the valence band top at 0 eV have been multiplied by 1.3

filled band of states in the gap. This implies a range of optical transitions, but with a possible gap at higher energies. This would not provide a featureless absorption but would be stronger at lower energies. Along with the reduced stability compared to the glide type, this would suggest that the shuffle dislocation also does not have a link to brown colouration in diamond.

The EEL and optical absorption spectra of the shuffle dislocation are shown in figure 5.14, with the EEL spectrum scaled by a factor of 2.5×10^6. Bandgap absorption is evident but it does not correspond to the experimentally observed absorption of brown diamond, displaying a dip in absorption around 1 eV and a strong onset that is below the bandgap edge, even with scaling. The EEL spectrum also shows the predicted low-energy peak due to a concentration of levels in the lower bandgap.

A further check as to whether shuffle dislocations can be responsible for brown colour in diamond is to calculate the expected number of core atoms required to produce an absorption magnitude comparable to that seen experimentally. There are four core atoms per cell which contribute to the absorption magnitude in a cell of 438 atoms, *i.e.* 0.009%. The model has an absorption coefficient of 2475 cm^{-1} at 2.5 eV, compared to ∼1 cm^{-1}

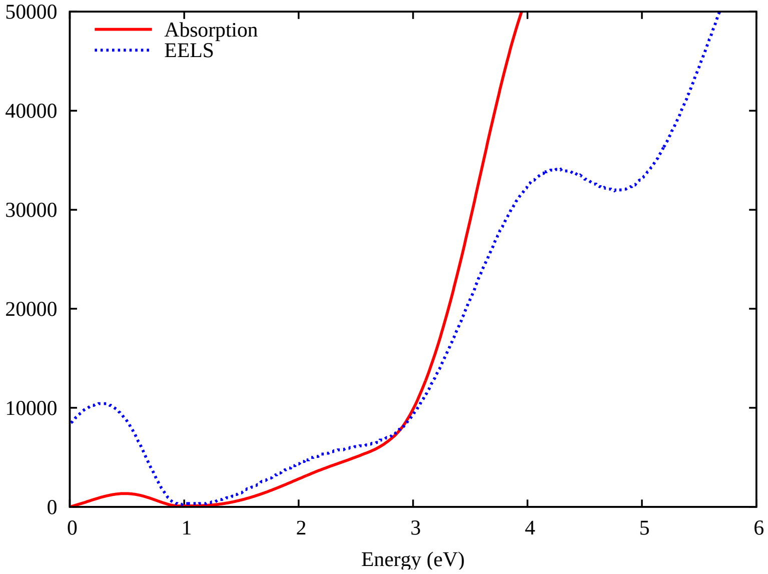


Figure 5.14: EELS spectrum and absorption of 90° shuffle type dislocation EELS spectrum. The energy values are unscaled. There is clear bandgap absorption but it does not correspond to the the characteristic pattern seen in brown diamond.

in most natural brown diamonds. Thus the fractional core atom density must be scaled to 3.6×10^{-6} to account for this, which corresponds to an atomic density of 6.49×10^{17}. The maximum density of dislocations observed in natural diamond is $\sim 10^9$ cm^{-2}, which corresponds to 1 core atom in about 10^8 carbon atoms, or an atomic density of 1×10^{15}. Thus it is clear that there are insufficient core atoms to account for observed absorption even if all dislocations were of the shuffle type, which seems unlikely. In CVD diamond the dislocation density is much lower, around 10^4 cm^{-2}, which makes the assignment of brown colour to dislocations in this material even more difficult to accept.

5.5 Conclusion

The EEL and absorption spectra have been calculated for both bulk diamond and for defects. The modelling requirements have been explored for bulk diamond, in terms of the sampling grid and the wavefunction bases needed. Excellent agreement for bulk diamond has been demonstrated for the real and imaginary parts of the dielectric function, and the EELS spectrum. Graphite has also been modelled to demonstrate that the enhanced EELS count around 7 eV in brown diamond can be linked to sp^2 bonded carbon.

Two different dislocations have been modelled to assess their link to brown diamond. The 90° glide type dislocation is found to be optically inactive through most of the bandgap energy range, due to core atom reconstruction of dangling bonds. The 90° shuffle type dislocation has a core of dangling bonds which induces absorption across the bandgap, however there are several reasons why these dislocations cannot be responsible for the brown colour of diamond. Firstly the shuffle dislocation is much higher in energy that the glide dislocation, meaning it is less likely to naturally occur. Secondly the absorption profile does not match experiment. Most critically however, it has been shown that the dislocation density required by this model to account for the experimentally observed magnitude far exceeds the dislocation densities of both natural and CVD diamond.

Chapter 6

(111) Plane Vacancy Disks

6.1 Introduction

This work has been published in *Physical Review B* **73**, 125203 (2006).

It has been shown in chapters 4 and 5 how small point defects and dislocations cannot be linked to the origin of brown colour in diamond. The presence of vacancies is known from PAS experiments on brown, colourless and transformed diamonds [30, 35, 112], but a single multi-vacancy cluster cannot explain the brown colouration as it is expected to induce a peak in absorption. On the other hand, an ensemble of clusters of different sizes may not exhibit a sharp peak but would be expected to give an absorption pattern which varied from sample to sample reflecting a different relative concentration of the clusters. Moreover, during the anneal, the less stable clusters would dissolve leaving the more stable ones. The latter would lead to an absorption spectrum with structure contrary to observations.

Carbon atoms in diamond prefer to form four covalent bonds in a tetrahedral fashion, as described in chapter 1. If this is not possible then a double bond may be formed between two adjacent carbon atoms. Such a bond will introduce π and π* states in to the bandgap, as observed in graphite [113], and in disordered carbon at grain boundaries [114]. Such π-bonds can form both on the (001)-(2$\times$1) surface [115] and the (111)-(2$\times$1) surface of diamond [103]. If the π-bonds are delocalized, *i.e.* they are connected to many other π-bonds, then the π and π* states in the bandgap will be dispersed, varying in separation at different points in the Brillouin zone. This is clearly seen in the bandstructures for the (001)-(2×1) surface in figure 6.1, and the (111)-(2×1) surface in figure 6.2.

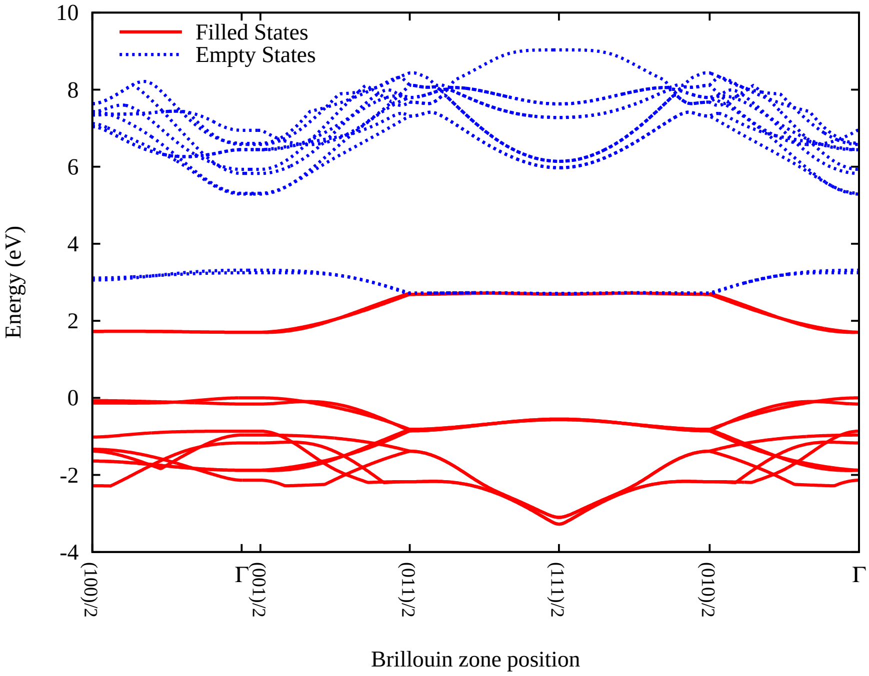


Figure 6.1: The bandstructure of the (001)-(2×1) surface of diamond, showing the occupied π and unoccupied π* states in the gap. There is little dispersion, as the π-bonded atoms are not connected. Energies are scaled by 1.3 to compensate for the bandgap underestimation by LDA.

The dispersed bandgap levels allow a range of transitions to occur, which in principle will lead to an optical absorption across a wide range. The occupied and empty levels meet in the bandgap, leading to a semi-metallic nature, which indicates transitions will start close to 0 eV, *i.e.* there will be no threshold of absorption.

If a high density of π-bonded atoms are to be incorporated into diamond then the most promising method is the creation of surfaces on the (111) set of planes inside the material as small defects, dislocations and grain boundaries cannot support an extended network of such bonds. A (111) surface can be created by the removal of two planes of atoms, as shown in figure 6.3. These lie on αB planes of the normal $A\alpha B\beta C\gamma$ (111) stacking sequence of diamond and their removal leads to dangling bonds lying parallel to [111] on each internal surface. Subsequently all dangling bonds are eliminated through the Pandey reconstruction [103] which leads to lines of π-bonds lying along $[1\bar{1}0]$. The length of the π-bond is 1.426 Å, a similar length to that found in graphite. The disk of vacancies

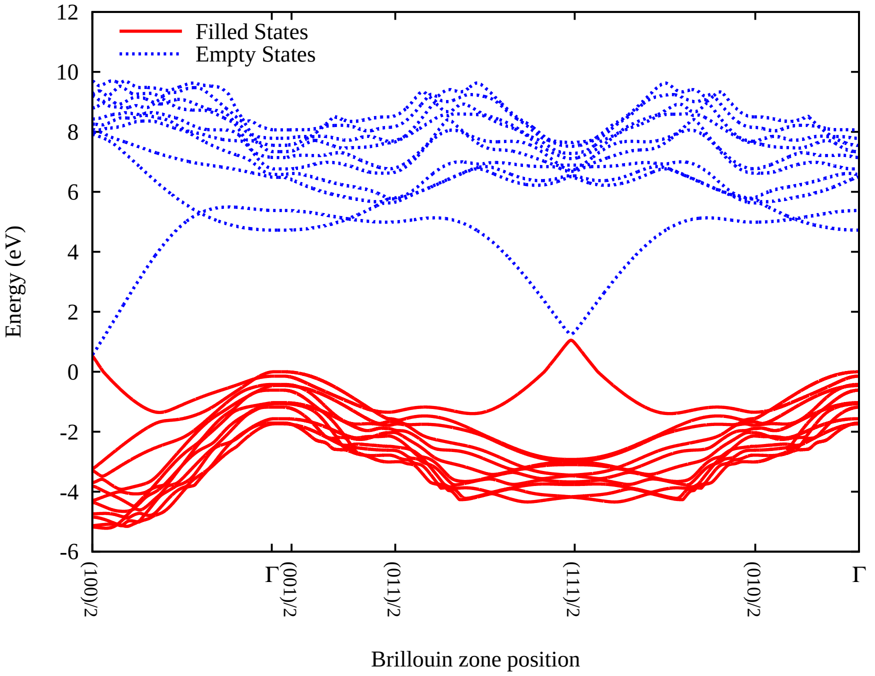


Figure 6.2: The bandstructure of the (111)-(2×1) surface of diamond, showing the occupied π and unoccupied π* states in the gap. There is strong dispersion of these levels, due to the long chains of π-bonded atoms. Energies are scaled by 1.3 to compensate for the bandgap underestimation by LDA.

which is subsequently created should display a bandstructure similar to that of the plain (111)-(2×1) surface of diamond, as will be described in the following sections.

6.2 Simple Vacancy Disks

6.2.1 Structures and energies

Vacancy disks are modelled in 40 and 80 atom cells. The separation between the disks in different cells is then 20 and 40 atomic layers respectively. The use of two differently sized cells helps determine the interaction range of the disks and will also allow the absorption density of vacancy disk defects to be deduced. The cells are structurally optimized using a $2 \times 2 \times 1$ grid of Monkhorst-Pack **k**-points.

The formation energy per vacancy is 1.46 eV in both the 40 atom and 80 atom cells,

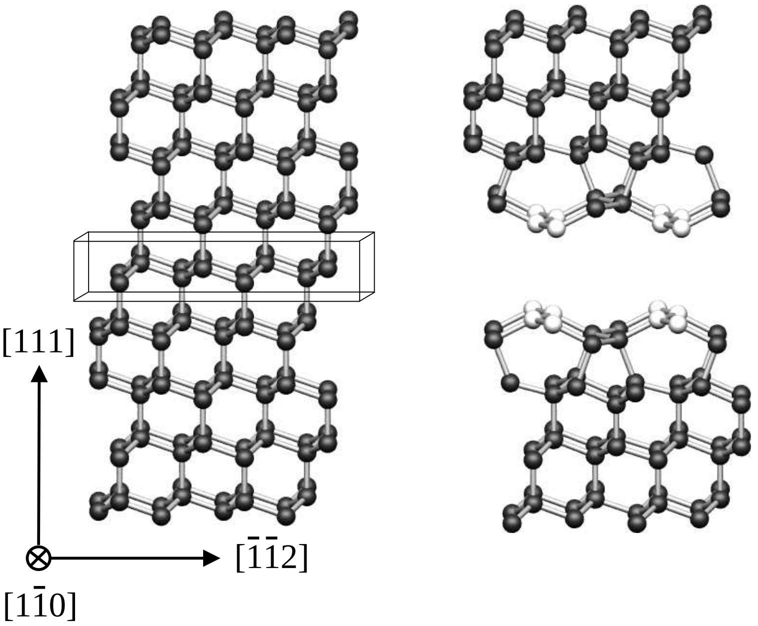


Figure 6.3: Illustration of the removal of a $\{111\}$ double plane (boxed) from bulk diamond and the subsequent surface rebonding leading to a vacancy disk with chains of π-bonded atoms along $\langle 1\bar{1}0\rangle$. The π-bonded atoms are shown in white for clarity.

demonstrating that a separation of 20 layers is sufficient for modelling purposes. The formation energy of a disk with (111)-(1×1) unreconstructed surfaces is 2.21 eV per vacancy. These energies are to be compared with 5.96 eV for an isolated vacancy studied using a Quantum Monte Carlo technique [116], and 7.17 eV for the vacancy using AIMPRO. The most stable multi-vacancy cluster found previously is a cluster of 14 vacancies with formation energy 2.35 eV per vacancy [117]. Comparison of these formation energies allows prediction of which is more likely to occur. It is clear that, given the ability to overcome any migration barrier, it is preferable for vacancies to form a disk lying on the (111) plane over any other defect.

The band structure, shown in figure 6.4, demonstrates that the band gap is completely filled with states and that the top of the occupied π-band is degenerate with the bottom of the empty π^*-band. The similarity to the bandstructure of the (111)-(2×1) is clear, but here there are two sets of bandgap levels arising from the dual surfaces. It can be anticipated from the arrangement of levels that the absorption will be continuous without any threshold.

The bandstructure can also be evaluated using the screened exchange technique de-

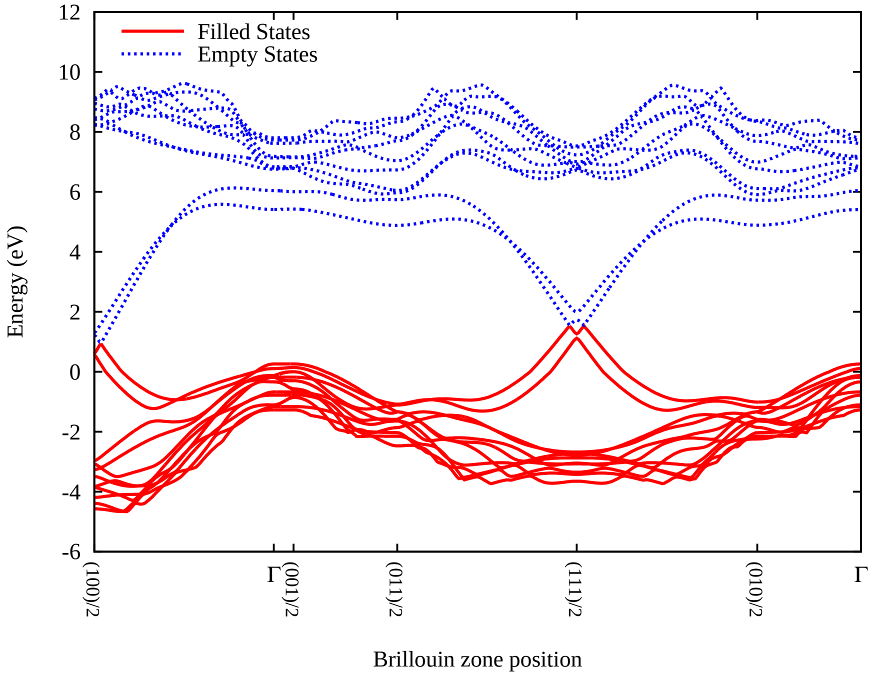


Figure 6.4: The bandstructure of a vacancy disk in a 40 atom cell. Note the absence of any pronounced gap in the spectrum suggesting a featureless absorption spectrum. The centre of the Brillouin zone is marked Γ, and other **k**-points denoted with their positions relative to the primitive unit vectors of the Brillouin zone of the cell. Energies are scaled by 1.3 to compensate for the bandgap underestimation by LDA.

scribed in section 2.8. Figure 6.5 compares the LDA values, which are scaled by 1.3, with those produced by the screened exchange functional. When the valence maxima are set to 0 eV for both methods there is very good agreement for valence and conduction states, with a fair agreement for the π/π^* crossing point. As the screened exchange functional has been shown to reproduce the bulk diamond bandstructure correctly, it can be assumed to be accurate in this case. This agreement justifies the use of the 1.3 scaling factor in the bandgap energy range. It is still computationally difficult to use screened exchange for large cells or dielectric analysis, hence no comparison can be made for optical absorption.

Due to its small cross-sectional area, the vacancy disk can be expected to have a PAS lifetime comparable to the monovacancy (140 ps). Calculations perfomed by J-M. Mäki show the disk to have a lifetime of 271 ps, which is similar to a cluster of 14 vacancies. Although larger than expected, there is a clear discrepancy between the observed ∼400 ps lifetime and that produced by the vacancy disk.

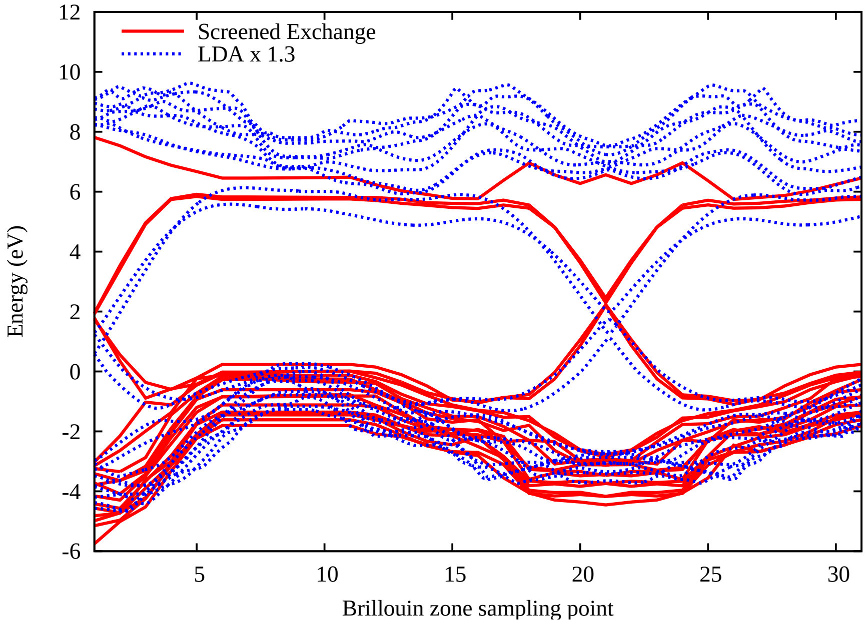


Figure 6.5: The bandstructure of a vacancy disk in a 40 atom cell using LDA and screened exchange. The LDA energy values are scaled by 1.3 above the VBM, which is set to 0 eV. Good agreement for valence and conduction bands is evident.

6.2.2 EELS and optical absorption

Initially the diagonal components of the dielectric tensor for the planar defect along the [111], [$1\bar{1}0$], [$2\bar{1}\bar{1}$] directions are calculated. The largest component lies in the (111) plane as is the case for graphite. The average values of the dielectric function and the corresponding optical absorption coefficient are calculated.

The EELS spectrum of the vacancy disk is shown in 6.6, compared to the spectrum of bulk diamond. The effect of the filled and empty bandgap states on the magnitude of the EELS signal is clear. Above ∼8 eV the two curves are essentially the same as the transitions here are mainly due to the bulk, but below this point the bandgap transitions dominate. It is especially significant that there is an increase in signal for the disk around 6 eV, as this agrees with experimental findings on brown diamond [95] These transitions maintain the EELS magnitude down to 0 eV. It is described below how the calculation technique can lead to a non-physical peak in the 0 – 1 eV range, which is evident for the vacancy disk. Hence the 0 – 1.5 eV region is omitted from the graph for clarity of viewing.

Figures 6.7 and 6.8 show the calculated optical absorption compared to experimental

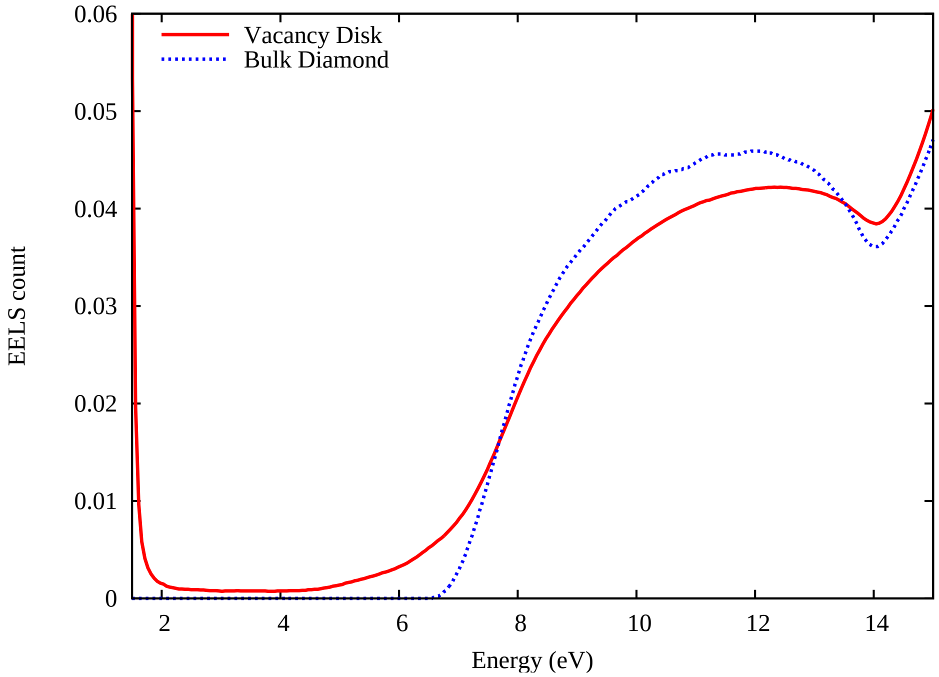


Figure 6.6: The calculated EELS spectrum of the (111) vacancy disk in an 80 atom cell, with the EELS spectrum of bulk diamond shown for comparison. It is clear that there are electron transitions for the vacancy disk far below those for bulk diamond, which indicates bandgap absorption. The 0 – 1.5 eV region is omitted as described in the text. Spectra are scaled on the energy axis to account for LDA bandgap underestimation.

data, in standard and log-log plots. The absorption is featureless in the range 1 to 5 eV as indeed is the experimental spectrum. The energy dependence of the absorption is very similar to the more transparent CVD-1 sample in the mid-gap region. Supplementary absorption in CVD-4 is due to additional defects which are likely to be nitrogen related, due to its presence in such samples. The expression for the dielectric function above takes only direct electron-hole transitions into account. Hence the large rise in absorption at 5.5 eV for diamond, due to indirect excitations and exciton effects, is not reproduced. In addition, the semi-metallic character of the infinite vacancy disk results in an absorption coefficient diverging at low energies [118]. A finite disk can be expected to possess a small gap between occupied π-bonding and empty π^* states leading to an absorption coefficient which vanishes with energy. However the vacancy disk leads to an absorption spectrum consistent with experiment on CVD-1, although the natural diamond may have additional centres.

An important finding is that the absorption coefficient calculated in the 80 atom cell is roughly a factor of half that found in the 40 atom cell where the density of vacancies

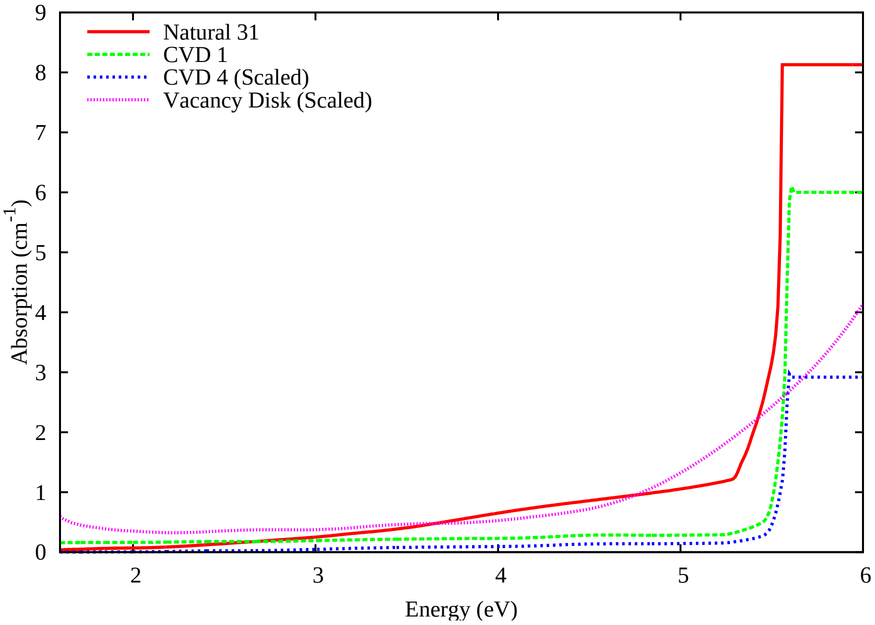


Figure 6.7: Absorption of natural and CVD-grown brown single-crystal diamond compared to absorption of (111) vacancy disk in an 80 atom cell. The absorption of the vacancy disk has been multiplied by a factor of 0.00025, and the CVD-4 sample by 0.01 for clarity.

is twice as large. Thus the absorption coefficient scales with this fraction. This allows calibration of the absorption with the sp^2/sp^3 fraction or vacancy density. An absorption coefficient at 2.5 eV of 0.05 cm^{-1} for a vacancy concentration of $\sim$ 1 ppm or 1.52$\times$10^{17} cm^{-3} is found. This is comparable with an extrapolation of experimental data on poor quality CVD material [119]. Thus one can deduce that there would be about 10^{17} vacancies in the CVD-1 diamond whose absorption is shown in figure 6.8 if all the absorption were due to disks.

The Raman active mode of bulk diamond is 1332 cm^{-1}, and of graphite is $\sim$1600 cm^{-1} [120]. Amorphous carbon displays shifted peaks, or a mixture of the two. In nanocrystalline diamond a 1490 cm^{-1} band has been observed and linked to diamond undergoing structural modification due to high pressure [121], which is likely at grain boundaries where sp^2 bonding can occur. It is to be expected that π-bonds could lead to vibrational modes in the 1400 – 1600 cm^{-1} region. A Raman active mode at 1494 cm^{-1} is calculated, which is close to one observed around 1540 cm^{-1} in CVD material [26]. Similar modes must be present to some extent in brown diamond if the vacancy disk model is correct.

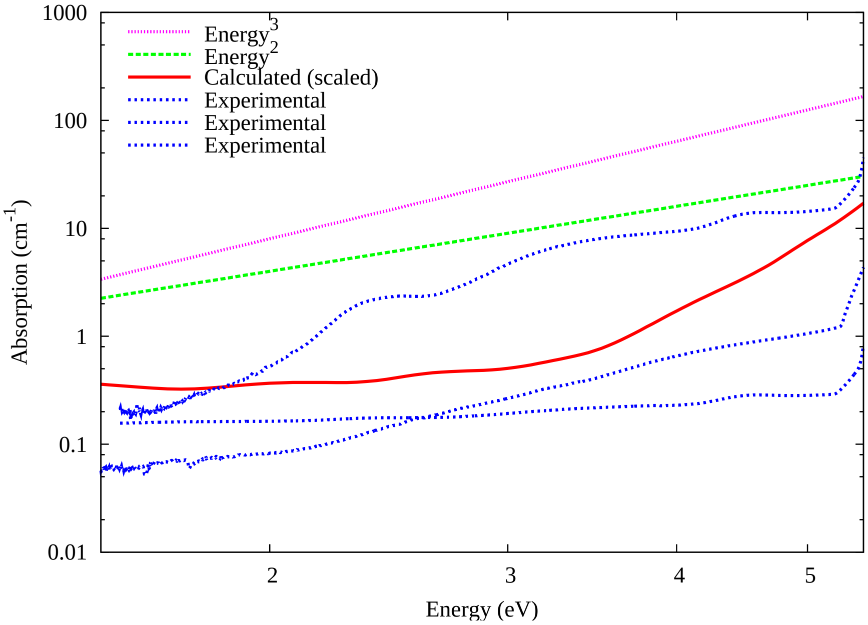


Figure 6.8: Comparison of the calculated absorption coefficient for a (111) vacancy disk with experimental data from natural and CVD-grown brown diamonds. The vacancy disk absorption is calculated in an 80-atom cell, and scaled by a factor of 2.5×10^{-4}. The lines E^2 and E^3 indicate the energy dependence of the absorption.

6.2.3 Electronic properties

It is possible to consider the (111) plane vacancy disk as a charge trap. There are a large amount of sp^2 bonds which are free to interact with mobile electrons, if such an interaction reduces the system energy, then a charge trap may be formed. This is calculated by comparing the energy change of the vacancy disk when an electron is added with the energy change of a known defect, in this case substitutional boron.

An 80 atom unit cell containing a vacancy disk was structurally optimized with one electron added, to mimic charge transfer. The energy of the neutral vacancy disk is taken as zero, and the extra energy gained by the negative negative disk is then 14.42 eV. In order to calculate any energy saving the substitutional boron defect was modelled in neutral and singly negative charge states. The neutral defect was found to have C_{3v} symmetry and the negative case T_d symmetry, with an increase in energy of 13.58 eV for the negative defect. Taking the energy gain for the negative disk and subtracting the boron energy gain yields 0.837 eV, which is added to the acceptor level of substitutional boron ($E_v + 0.37$ eV [122]) resulting in an acceptor level for the negative vacancy disk of

$E_v + 1.137$ eV.

The energy saved by charge transfer depends on the defect from which an electron is supplied. In the case of the A-centre with a donor level at $E_c - 4$ eV [122, 123] the saving is 0.293 eV.

If vacancy disks do act as a charge trap they could be a critical defect in the field of diamond electronics, as they will have a significant effect on the conduction properties of any diamond that they are present in. Recent results from the EPR group at the University of Warwick indicate that the density of acceptors in CVD diamond is 10 times greater than the density of known point defects [124]. This large margin means that there must be some unidentified defect responsible for the acceptor level.

6.2.4 Removal of colour

There are significant differences between the annealing characteristics of the different types of diamond. The absorption continuum disappears at 1400° C for irradiated material [39] and in the range 1400-1600°C for brown single-crystal CVD material [33]. For brown natural type IIa diamond however, temperatures above 2200°C are required to remove the absorption continuum [32]. This suggests that either the stabilities of the defects responsible for the absorption continuum in brown single-crystal CVD material and natural diamond are significantly different, or the mechanism for its loss are different. In this section the latter argument is examined, as it has been postulated that the vacancy disk model can explain the brown colour in both natural and CVD diamond.

Recent experiments [34] show that the transformation of brown to colourless CVD brown diamond is accompanied by the growth of a broad C-H band around 2900 cm^{-1}. This frequency is close to the C-H stretch mode on a (111) surface, which occurs at 2838 cm^{-1} [125]. The question arises as to whether hydrogen could passivate the optical activity of the disk.

The vacancy disk is modelled in the same 40 atom configuration previously described, but with 4 hydrogen atoms added; a number equivalent to the number of vacancies. This cell was structurally optimized with two surface arrangements, the (111)-(2×1) surface (the Pandey chain) and the (111)-(1×1) surface, which is equivalent to the bulk (111) plane. As both systems contain the same number of each species of atoms they can be compared directly, with the (111)-(1×1) surface configuration being 3.08 eV lower in energy.

Hydrogen has also been modelled on the other major crystallographic planes, with the relative energies shown in table 6.1. As there are differing numbers of carbon and hydrogen atoms in the cells a comparison is made of the energy per carbon-hydrogen bond at the surface. The formation energy of the cell is calculated, and then divided by the number of carbon-hydrogen bonds. The lowest energy configuration here is (111)-(1×1):H which releases energy during the bonding process. This supports the view that mobile hydrogen will become trapped by vacancy disks.

Table 6.1: Relative energies of hydrogen at diamond surfaces.

Surface	Formation energy per C-H bond (eV)
(111)-(2×1):H	0.41
(111)-(1×1):H	-0.36
(110)-(1×1):2H	-0.29
(100)-(2×1):H	0.21
(100)-(1×1):H	0.21 (relaxes to (100)-(2×1):H)
(100)-(1×1):2H	0.55

Hydrogen requires two electrons in its S shell to be stably bonded, hence hydrogen forms H_2 naturally. If hydrogen is near the (111) disk surface it will effectively gain an extra electron by sharing one from the π-bonded carbon in the Pandey chains. This means that the carbon is now four-fold co-ordinated and there are no free electrons for optical or electronic interaction. Figure 6.9 shows the bandstructure of a (111)-(1×1):H plane vacancy disk with hydrogen saturation at the surface. It is evident from the figure that all absorption up to an energy value of 4.35 eV is eliminated, which covers the visible spectrum and would render the diamond colourless.

Calculation of the vibrational modes for the hydrogenated (111)-(1×1):H plane disk indicates that there are modes localized on the surface hydrogen atoms at 2923, 2927, 2973 and 2976 cm^{-1}. These values are in excellent agreement with those observed when transforming CVD diamond from brown to colourless with HPHT annealing [34]. This supports the proposition that hydrogen passivation of disk surface bonds is the method of removal of colour in CVD diamond.

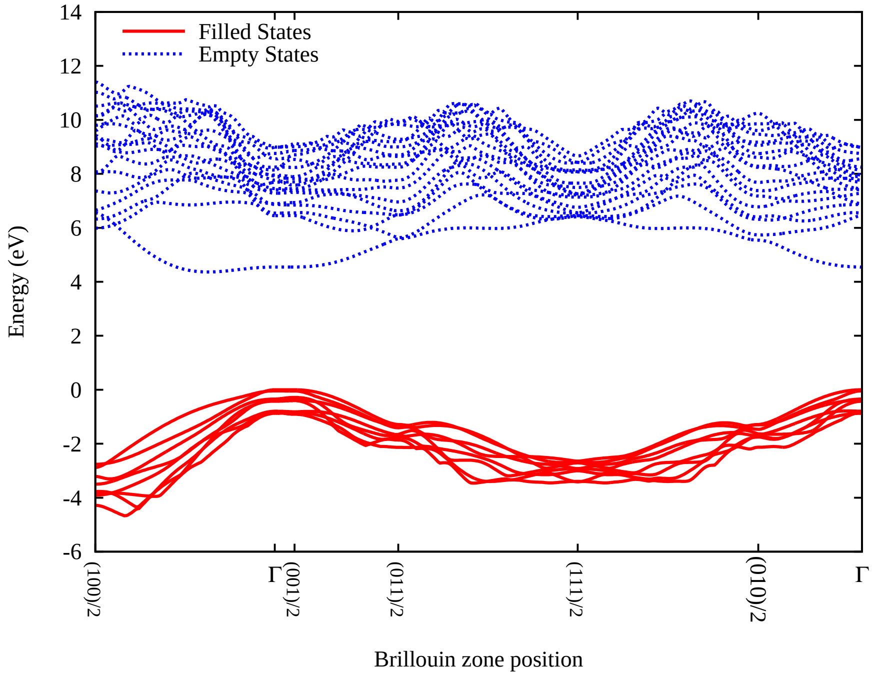


Figure 6.9: Bandstructure of the (111) vacancy disk with hydrogen termination. The bandgap is cleared of states and hence optical absorption in that energy range is removed. Energy states are scaled by 1.3 to compensate for LDA underestimation.

Natural type IIa diamond is very pure and contains little hydrogen, so here another process must account for the majority of the colour removal. Monovacancies and divacancies are mobile at temperatures around 700° C [126, 127] and 900° C [128] respectively, hence under annealing conditions they can move through the crystal and become trapped by existing disks - causing them to grow. There is nothing to check the growth of the disks until they become unstable against collapse. During collapse one of the surfaces will be displaced towards the other by $a(110)/2$, eliminating the fault but forming a perfect dislocation loop with a Burgers vector b equal to this displacement. The energy of such a collapsed disk of radius r, according to elasticity theory [27] is:

$$\sim \frac{\mu b^2 r}{2(1-\nu)} \left(\ln(4r/b) - 1 \right) \tag{6.1}$$

here μ and ν are the shear modulus and Poisson ratio of diamond. Plotting the surface area energy of the disk versus the energy from equation 6.1 leads to the dislocation loop being more stable than the disk for loops of radius greater than about 12 Å which contain about 200 vacancies. This is shown in figure 6.10. This process for removal of the colour would require much higher temperatures to remove the colour in a reasonable space of time, hence the higher annealing temperature for natural diamonds [31, 32].

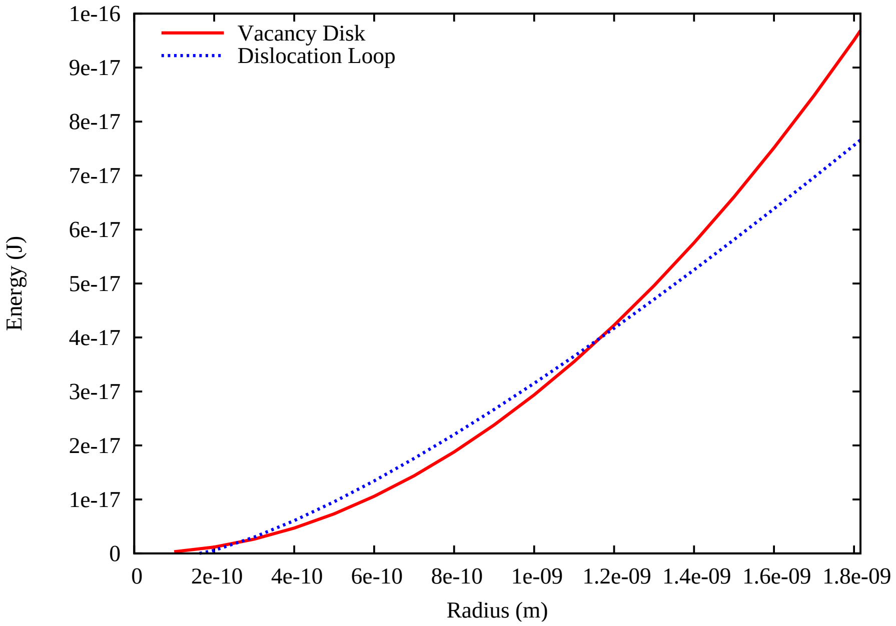


Figure 6.10: Comparison of the energy of a vacancy disk to a dislocation loop as described in the text. The crossing of the curves indicates the maximum stable loop radius to be just less than 12 Å.

6.3 Nitrogen at Vacancy Disks

The previous section considered type II diamond, which contains little nitrogen. Brown coloured diamonds exist in all types, and as type I diamond contains nitrogen to densities in excess of 1 part-per-million (ppm) the interaction of the vacancy disks with existing nitrogen defects is of high importance.

As described in chapter 1 the optical absorption of type I brown diamond is quite different from type II brown diamond. Type I diamond exhibits extra absorption peaks, notably at $\sim$2.25 eV (550 nm) [44], over the background continuum. It is known that nitrogen in diamond can form several centres; the A-centre consisting of an N-N pair, the B-centre consisting of four nitrogen atoms surrounding a vacancy and the N3-centre of three nitrogen neighbouring a vacancy being the main defects [123, 22]. However, none of these defects are known to produce an absorption band at 2.25 eV (table 1.2). It has been observed that electron-irradiation of type I diamond with subsequent annealing creates an unknown defect with absorption in the 536 and 575 nm areas (2.31 and 2.15 eV) [129]. This is attributed to the H4 centre dissociating to form other defects.

It is known from CVD growth studies that nitrogen disrupts growth to a greater extent when the growth direction is (111), due to easier incorporation on these surfaces [130, 131]. Ab-initio studies suggest that nitrogen incorporated into Pandey chains on the (111) surface are less energetically favourable [132], but this may differ for a disk. This section will examine the energetically favourable configurations in which nitrogen may be able to exist at the surface of a (111) vacancy disk, and the optical and electronic properties which arise from this.

Initially a study of point vacancy-nitrogen defects is made. This allows the lowest energy defect to be identified. The chemical potential of nitrogen in diamond can be determined from this low-energy defect, allowing further formation and binding energies to be calculated. The defects are all modelled in 216-atom cubic supercells, with a $2\times2\times2$ Monkhorst-Pack sampling grid.

Table 6.2 shows the relative formation energies of the defects studied with energies of other defects for comparison. The B-centre formed of four nitrogen atoms surrounding a vacancy is found to be lowest in energy, in agreement with previous theoretical work [43]. The lower formation energy of the N3, A, B, C and H3 centres indicates that these are more likely to be found in diamond containing nitrogen. Measurements by DTC of brown type Ia diamond show low concentrations of H3 (N-V-N) and H2 (N-V-N^-) centres. When such diamonds are subjected to HPHT treatment the concentration of these defects, as well as N3, is seen to increase [30]. This indicates that free nitrogen and vacancies are released from a source under these conditions.

The defect concentration based on these energies is able to account for the defect concentration observed in treated brown type Ia diamond, but is in contrast to the population in irradiated material. It is likely that irradiated brown diamond is in a strong non-equilibrium state, which would change the relative concentrations of any defects.

Table 6.2: The formation energies of various nitrogen-vacancy defects, compared to the B-centre.

Defect	Label	Relative Formation Energy (eV)
VN_4	B-centre	0
VN_3	N3	1.8
N-N	A-centre	1.9
N	C-centre	3.1
N-V-N	H3	3.5
VN	-	5.6
V_2N_4	H4	6.1
V	GR1	7
V-V	TH5	9.9

6.3.1 Structures and energies

The vacancy disks are created in unit cells consisting of a total of 80, 120 or 160 atoms depending on the configuration of nitrogen and the surface and in the bulk. Removal of the αB (111) planes allows Pandey chain reconstruction to lower the surface energy [103], as shown in figure 6.3 with the π-bonded atoms marked in white.

The binding energy of a single nitrogen (C-centre) and the nitrogen pair (A-centre), to the vacancy disk, are calculated using the formation energies of the individual defects:

$$E_{binding} = (E_A^{form} + E_B^{form}) - E_{AB}^{form} \tag{6.2}$$

where E_x^{form} is the formation energy of defects A and B separately, and together AB. The formation energy of the C-centre is found to be 3.12 eV, and of the A-centre to be 1.9 eV. These energies are calculated using a chemical potential for nitrogen derived from the energy of the B centre, which is the most favourable configuration for nitrogen in diamond. In comparison the formation energy for the vacancy disk is 1.46 eV/vacancy [133]. Knowing these values the formation energies of the combined systems can be calculated. The formation energy of the C-centre at eight atomic layers from the surface, in a 120 atom cell, is 3.25 eV, The A-centre at the same distance has a formation energy of 1.93 eV. These values lead to a binding energy of the C-centre to the disk of 1.0 eV and a value for the A-centre of 2.35 eV.

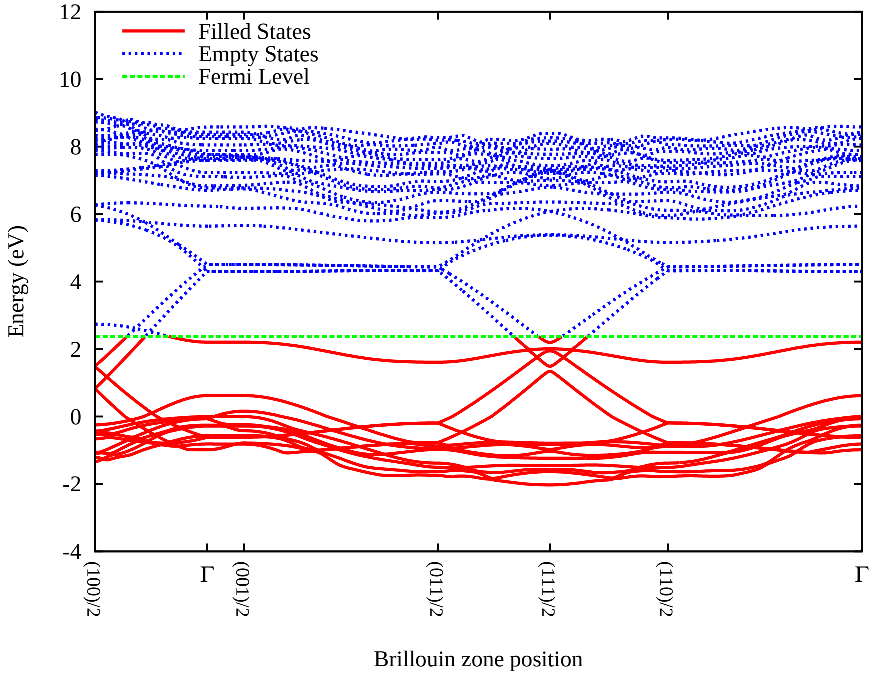


Figure 6.11: Bandstructure of an A-centre near to a vacancy disk. The centre is 6 atomic layers from the disk surface.

The bandstructures of an A-centre and a C-centre near the disk are displayed in figures 6.11 and 6.12. Single defects have been placed at 6 and 7 atomic layers from the disk surface for the A and C centres. In both cases it can be seen that the π and π^* states induced by the disk surfaces are still evident, and cross to give a continuous range of energy transitions. The gap level associated with the nitrogen defect in each case is superimposed over the disk bandstructure. In the case of the A-centre it is noticable that the nitrogen-related level is mostly below the Fermi level, and hence filled. This suggests that charge transfer to the disk is unlikely, even though calculations indicate that this would lower the system energy. The level associated with the C-centre is much closer to the Fermi level and crosses it at several points. This suggest that electron transfer here is more feasible.

The binding energies of further bulk vacancy-nitrogen defects have been calculated for comparison, with the binding energy of two single vacancies to form a divacancy being 4.45 eV. The binding energy of a vacancy to a nitrogen is 4.69 eV and the binding energy of a vacancy to an A-centre to form the H3 centre is 5.53 eV.

If an A-centre migrates to the surface of a (111) vacancy disk then it may arrange itself

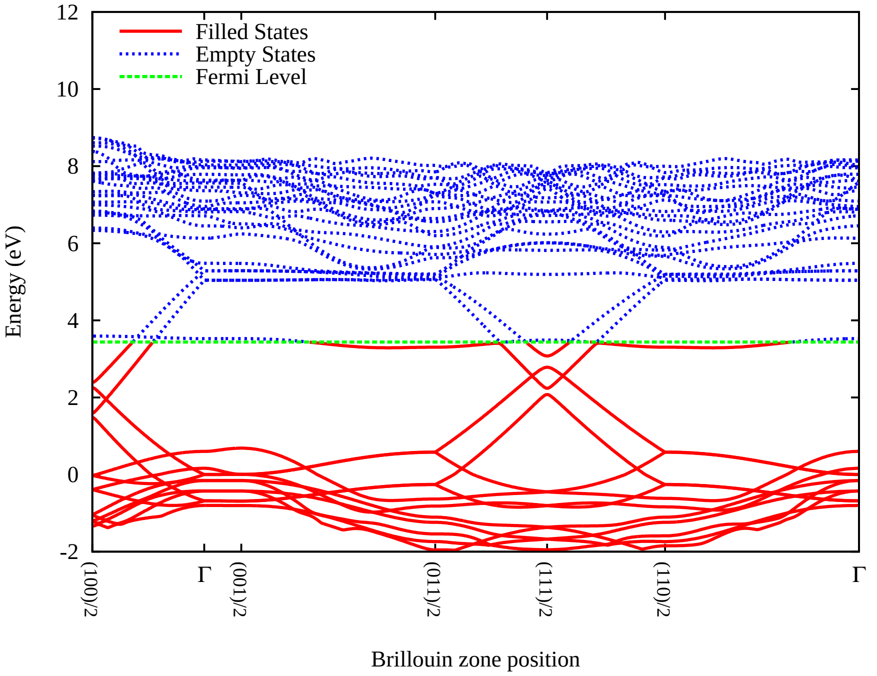


Figure 6.12: Bandstructure of a C-centre near to a vacancy disk. The centre is 7 atomic layers from the disk surface.

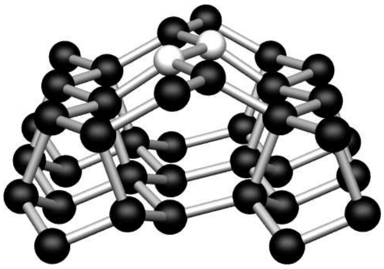


Figure 6.13: Nitrogen pair at the (111) plane vacancy disk surface.

into several configurations, which are shown in figures 6.13 – 6.17. Here the nitrogen atoms are coloured white for clarity. The presence of the nitrogen in the Pandey chains causes disruption to the usual 1.42 Å bond length of the carbon. The symmetry of the system is broken due to varying bond lengths. The nitrogen pair (figure 6.13) has a separation of 1.47 Å. Nitrogen has 1.41 – 1.44 Å bond lengths with neighbouring carbon atoms. Carbon-carbon bonds next to the nitrogen have 1.40 – 1.47 Å bond lengths. This disruption of the bond lengths will affect the optical absorption properties of the Pandey chain.

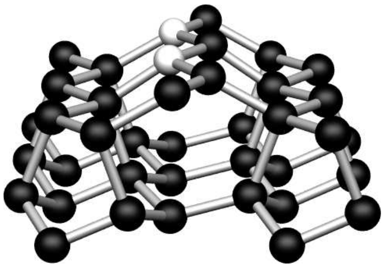


Figure 6.14: Nitrogen pair separated by one carbon atom, at the (111) plane vacancy disk surface.

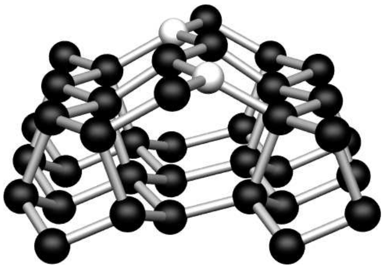


Figure 6.15: Nitrogen pair separated by two carbon atoms, at the (111) plane vacancy disk surface.

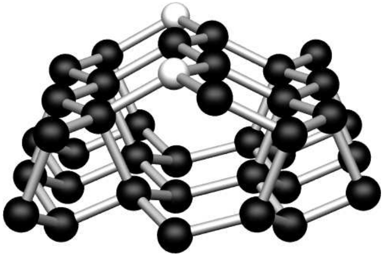


Figure 6.16: Nitrogen pair separated by three carbon atoms, at the (111) plane vacancy disk surface.

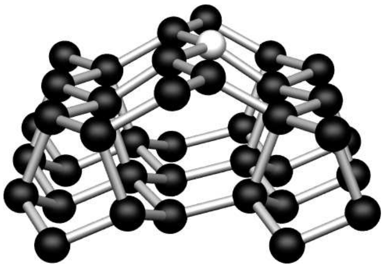


Figure 6.17: Isolated nitrogen at the (111) plane vacancy disk surface.

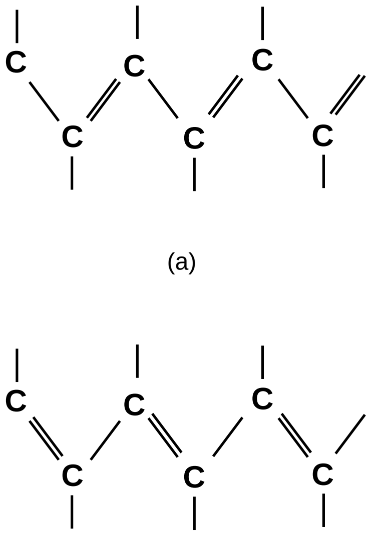
Table 6.3: Total energies of nitrogen configurations at the disk surface, compared to the nitrogen A-centre.

Configuration	Energy difference (eV)	Energy per nitrogen (eV)
N-N	-4.709	-1.177
N-C-N	-5.177	-1.294
N-C-C-N	-6.310	-1.577
N-3C-N	-5.707	-1.427
N	-2.067	-1.033

The favourable configuration for nitrogen atoms at a disk surface is deduced by simply comparing the total energy of each defect in a supercell with the supercell containing the A-centre several layers from the disk. The cells containing isolated nitrogen consist of 118 carbon + 2 nitrogen atoms and the N-3C-N cells have 156 carbon + 4 nitrogen. All other cells consist of 116 carbon + 4 nitrogen. Table 6.3 displays the energies of the various configurations of nitrogen compared to the A-centre.

It is shown that the most stable configuration for nitrogen to be in at the disk surface is for two nitrogen atoms to be separated by two carbon atoms. This configuration has the most regular symmetry with carbon-carbon bonds of 1.40 Å length, and nitrogen-carbon bonds of 1.44 Å length. Using figures 6.19 – 6.23, the ratio of π-bonds to single-bonds along the chain when nitrogen is included can be compared. The ratio of 2:6 π-bonds to single-bonds is the same for the N-N and N-C-C-N configurations, but the energy is much lower in the latter case. The nitrogen pair has many varying-length bonds and hence low symmetry, which may be reflected in its higher energy. This can also be understood by thinking of the π-bonded chain as a resonance between the two reconstructions shown in figure 6.18. The effect of N is to disrupt the bonding leading to a lone pair on each N atom together with three N-C bonds. Replacing -C=C- by -N-N- is less stable than replacing -C=C-C=C- by -N-C=C-N-. This shows that A centres will split apart at the surface of the disk.

It has been suggested by M. I. Heggie [134] that each nitrogen atom disrupts the system of π-bonds around it, weakening the three neighbouring bonds (but principally the two bonds lying on the surface). At the first neighbour positions this disruption is a minimum, but the nitrogen atoms repel due to a mixture of electrostatics and filling of anti-bonding orbitals with the extra two electrons. At the third neighbour positions the disruption of the π system is least because the carbon atoms having weakened bonds due to nitrogen



(a)

(b)

Figure 6.18: Alternative reconstructions of a π-bonded Pandey chain. The horizontal axis is $[1\bar{1}0]$ (see figure 6.3) and the vertical axis is along $[\bar{1}\bar{1}2]$. Only the π-bonded atoms are shown.

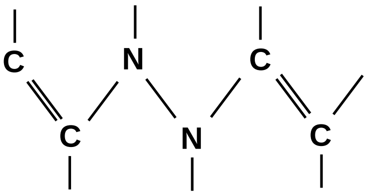


Figure 6.19: Bonding configuration of a nitrogen pair at the (111) plane vacancy disk surface.

now neighbour each other and can enhance their bond strength with each other. This can be summarised as the nitrogen atoms creating a π radical on the neighbouring carbon atoms. In NCCN these radicals neighbour each other and form a stronger π bond.

The bonding configuration for nitrogen in Pandey chain is shown in figures 6.19 – 6.23. Single and double bonds are denoted by single and double lines, and those carbon atoms which have a dangling bond are circled. The cell repetition is such that there will be only a short length of Pandey chain between nitrogen sites. This high nitrogen density may result in higher energy due to interference between the nitrogen defects, which is investigated in larger cells. Cells of 240 and 360 atoms are created by removing the nitrogen atoms to

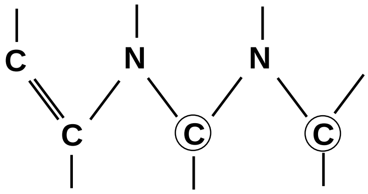


Figure 6.20: Bonding configuration of a nitrogen pair separated by one carbon atom, at the (111) plane vacancy disk surface.

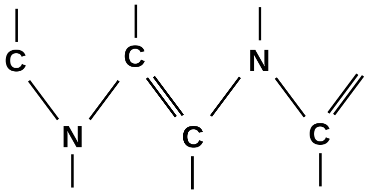


Figure 6.21: Bonding configuration of a nitrogen pair separated by two carbon atoms, at the (111) plane vacancy disk surface.

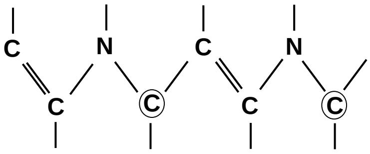


Figure 6.22: Bonding configuration of a nitrogen pair separated by three carbon atoms, at the (111) plane vacancy disk surface.

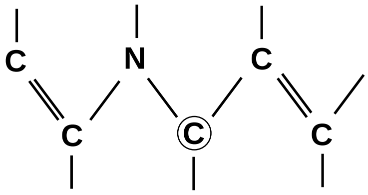


Figure 6.23: Bonding configuration of isolated nitrogen at the (111) plane vacancy disk surface.

create a cell of only carbon, then appending this cell onto the original one to create a cell with a longer chain of carbon atoms between the nitrogen atoms. For the 240 atom cell, both the N-N and N-C-C-N configuration are tested. It is found that the N-C-C-N system is 0.17 eV lower in energy when there is a larger gap between the nitrogen atoms. The

N-N system is also lower in energy, when this configuration is tested, by 0.1 eV. These lowering in energies can be attributed to the lower disruption of the Pandey chains and reduced interaction between nitrogen atoms.

The bandstructure of the N-C-C-N configuration in the 120 atom cell is shown in figure 6.24. Compared to the simple (111) vacancy disk [133], the semi-metallic bandstructure is lost, indicating a absorption onset will occur in the region of 2 eV. A broad continuum absorption like that seen experimentally in brown diamond is not to be expected from such a bandstructure.

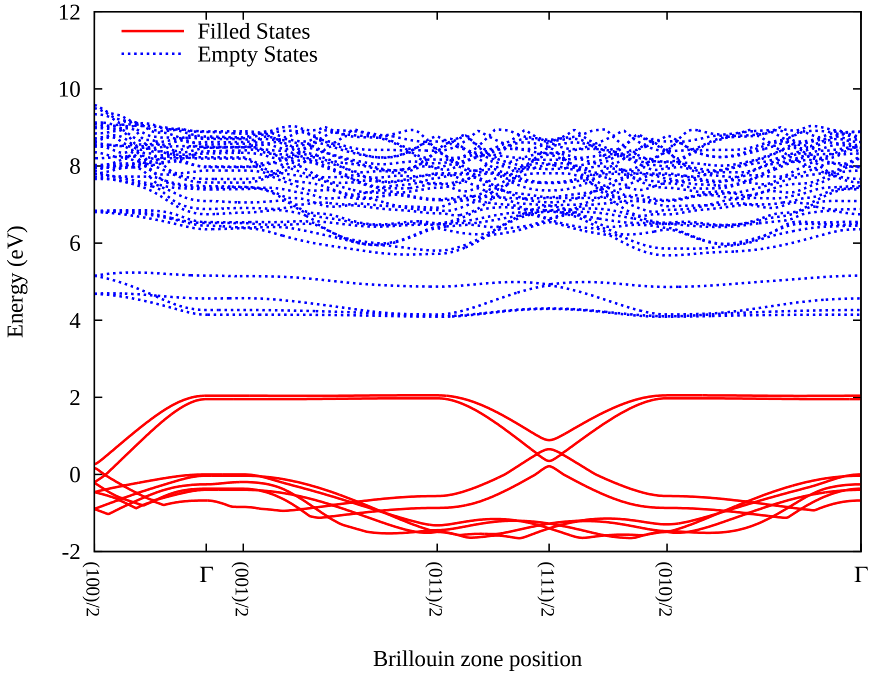


Figure 6.24: Bandstructure of the N-C-C-N defect at the surface of a (111) plane vacancy disk. This is modelled in a 120 atom cell with a small separation between the nitrogen atoms. This introduces a gap of ∼2 eV into the bandstructure. The top of the valence band is set to 0 eV, and states above it have their energies scaled by 1.3.

The bandstructure of the N-C-C-N configuration in a 240 atom cell is shown in figure 6.25. With a larger separation of the nitrogen atoms the gap in levels is narrowed to ∼1 eV, so it is expected that a much lower absorption onset will occur. This demonstrates a significant interaction between N-C-C-N defects, but in reality the density of nitrogen at the disk surface would be low enough that almost no defects will be at all close to any other.

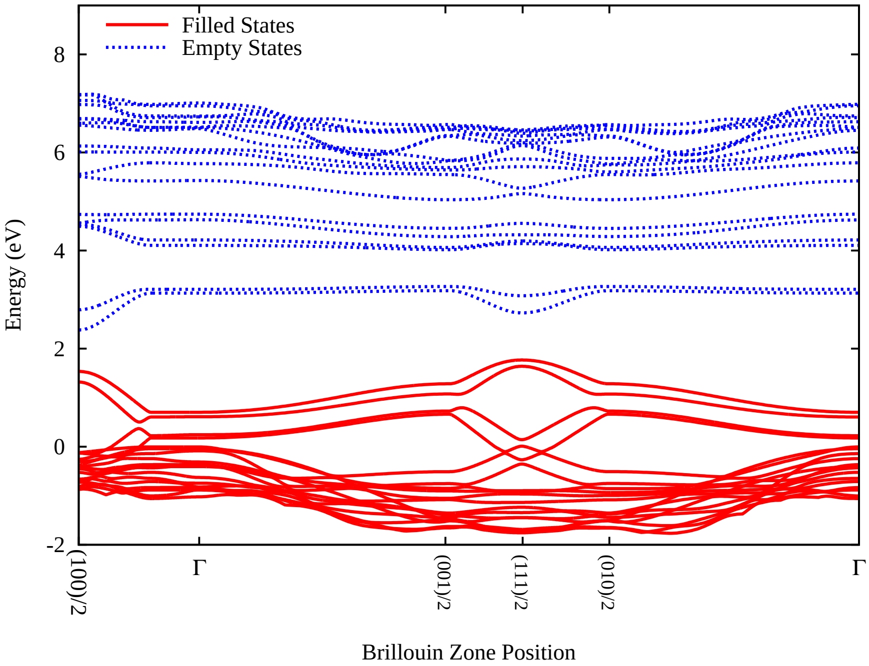


Figure 6.25: Bandstructure of the N-C-C-N defect at the surface of a (111) plane vacancy disk. This is modelled in a 240 atom cell with an increased separation between the nitrogen atoms. The gap in levels is reduced to ~1 eV. The top of the valence band is set to 0 eV, and states above it have their energies scaled by 1.3.

Using the 120 atom supercell, the nitrogen atoms are removed from one surface and the bandstructure re-calculated with half of the Pandey chains undisturbed. This approximates a lower density of N-C-C-N defects, in a system that is computationally more simple. As described above, this is a lower energy system due to less interaction between nitrogen atoms. The bandstructure of such a system is a combination of the semi-metallic structure of an uninterrupted Pandey chain combined with that of the nitrogen decorated disk. The bandstructure is shown in figure 6.26, with the same scaling as described previously. Such a semi-metallic bandstructure may be expected to return an absorption continuum, like that seen in brown diamond.

EELS and optical absorption

The absorption spectrum for the high-density N-C-C-N configuration is shown in figure 6.27, with a 0.5 eV polynomial broadening applied, and a scaling factor of 1.3 applied to the energy to compensate for LDA bandgap underestimation. The absorption displays

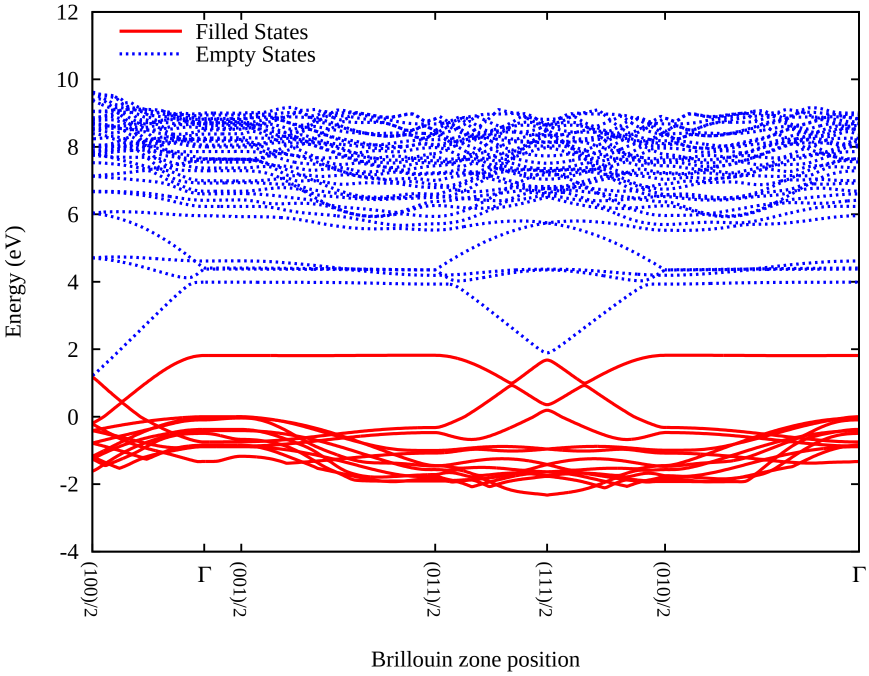


Figure 6.26: Bandstructure of the N-C-C-N defect with a 50% coverage on the (111) plane vacancy disk. The combination of π-bonded chains and nitrogen defects has given a modfied but still semi-metallic system. The top of the valence band is set to 0 eV, and states above it have their energies scaled by 1.3.

an onset at $\sim$1.7 eV (729 nm) and a peak at 2.5 eV (495 nm). Type I brown diamond shows a peak at 2.25 eV (500 nm) which is not inconsistent with the calculated peak, but brown and pink diamonds do not exhibit the absorption onset seen here, rather they have a finite absorption to low energies. The precise value of the onset cannot be predicted due to the broadening applied.

The density of nitrogen to carbon modelled for the absorption in figure 6.27 is high. There are four nitrogen atoms and 116 carbon atoms in the cell, giving a density of 3.45 % which is equivalent to 5.8×10^{21} cm^{-3} of nitrogen atoms. However, the calculated absorption can be scaled at 4 eV from 16 358 cm^{-1} to a value agreeing with one typical for brown type Ia diamonds. A more realistic absorption magnitude is 4 cm^{-1} at 4 eV (310 nm), and using this one obtains a revised nitrogen density of 1.5×10^{17} cm^{-3}. Though this assumes most nitrogen is at the disk surface, which is unrealistic.

As described above, the concentration of nitrogen at the disk surface can be reduced by half, leading to a bandstructure which has a wider range of available transitions (fig-

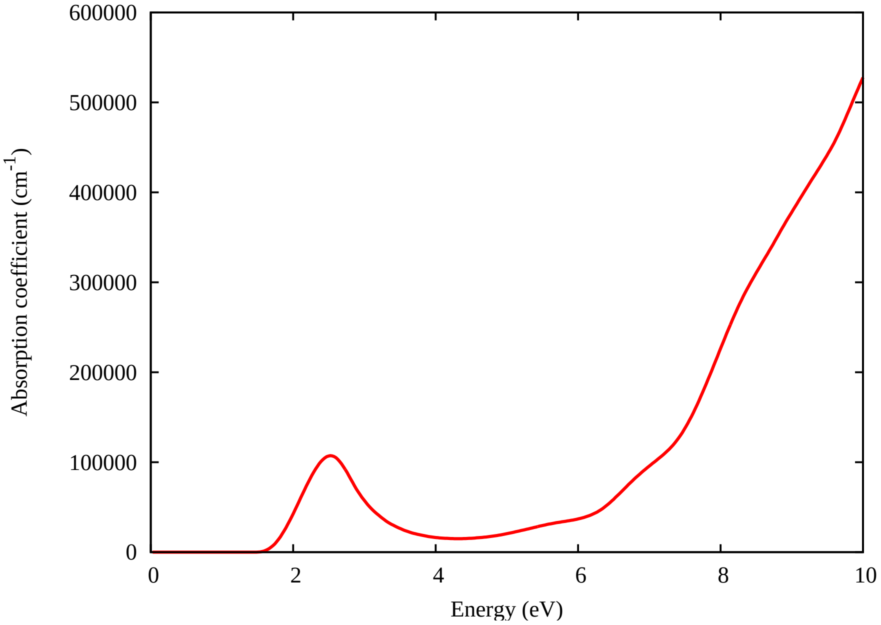


Figure 6.27: Optical absorption of the N-C-C-N defect at the surface of a (111) plane vacancy disk. This is modelled in the 120 atom cell with a high density of nitrogen. The energy values are scaled by a factor 1.3.

ure 6.26). The absorption profile for this defect is shown in figure 6.28. The higher-energy peak has now shifted to 2.4 eV (516 nm) and is 59 % of the magnitude when at full density. The absorption threshold has been replaced by a continuum absorption down to 0 eV, which is more alike that observed experimentally in brown diamond. Additionally there is now a peak in absorption at 0.7 eV (1770 nm), which is close to that observed in amber diamonds [45].

It is possible to get an idea of the density of nitrogen required for the observed absorption by scaling the calculated magnitude to match the experimental values, as described above. This requires density be ∼4000 times lower. The modelled nitrogen density is 2/118 or 1.69 %, this equates to 2.87×10^{21} cm^{-3}, which is reduced to 7×10^{17} cm^{-3} upon scaling. The fraction of the total nitrogen available to migrate to the disk surface must be considered, if it is 10 % then a total density of 7×10^{18} cm^{-3} is required, for 1 % it is 7×10^{19} cm^{-3} and at 0.1 % it is 7×10^{20} cm^{-3}, which is just less than 1000 parts per million (ppm). Nitrogen is known to be present in diamond at densities up to, and just over, 100 ppm [11] so these values are not unreasonable assuming 1 % of nitrogen is at a disk surface.

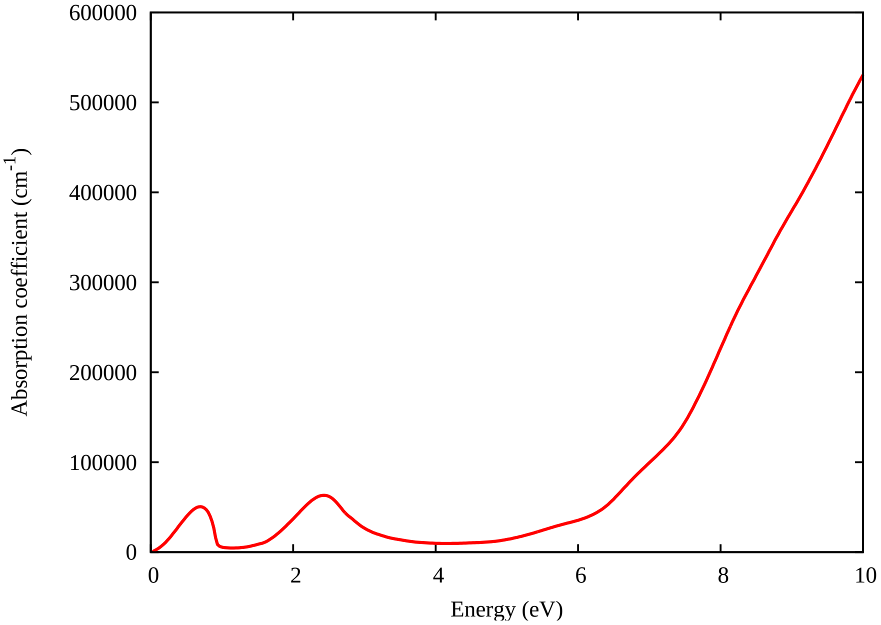


Figure 6.28: Optical absorption of the N-C-C-N defect at the surface of a (111) plane vacancy disk, at half the initial density. This is calculated in a 120 atom cell, with one surface free of nitrogen. The energy values are scaled by a factor 1.3.

The model described here fails to account for the 3.17 eV (390 nm) band exhibited in pink diamond as described in section 1.3. However it accurately reproduces the absorption spectrum of type Ia brown diamond observed experimentally. Both brown and pink diamonds have a similar magnitude of absorption, which will correlate to defect density, and hence the difference between the two may be more than just density of colour centres. It appears that another defect may be involved, which could be related to the ability of the disk to act as a charge trap.

6.3.2 Boron interaction with vacancy disks

In order to provide further comparison to these results, and for completeness, single boron has been modelled at the vacancy disk. In this model boron replaces one carbon atom in each chain. The binding energy is determined by modelling two single boron atoms away from the disk and is found to be 1.44 eV. This binding energy is slightly more than a single nitrogen which has a binding energy of 1.0 eV. If we look at the bandstructure of a single boron atom in a bulk configuration but close to the disk, we see that the Fermi level lies below the crossing of the π-$\pi*$ levels. The bandstructure is shown in figure 6.29.

The Fermi level position is such that rather than charge transfer to the disk, the reverse is expected to occur. Electrons from the disk could passivate the boron acceptor level, leaving an excess of holes in the π-bonded chains.

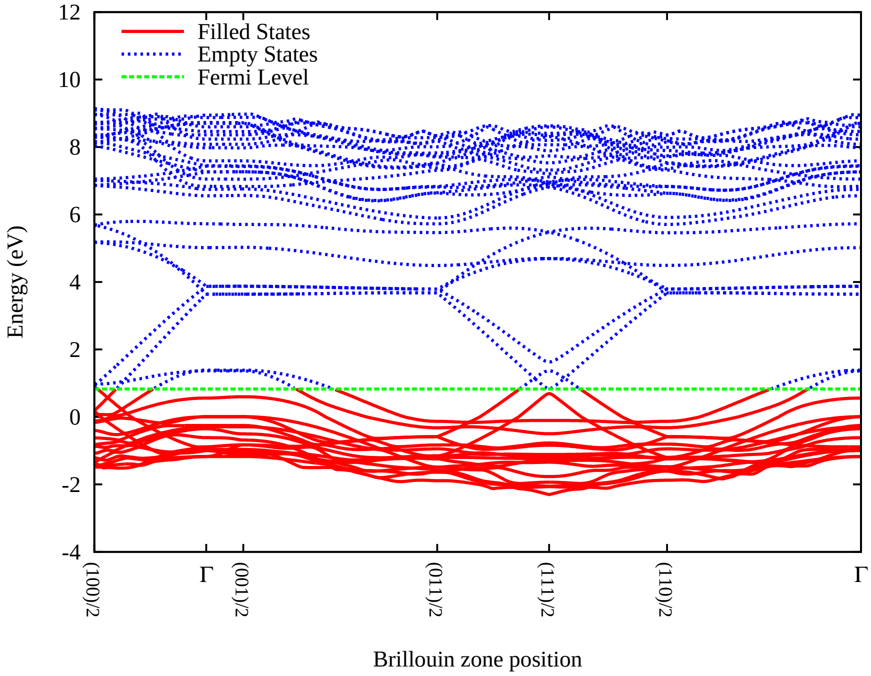


Figure 6.29: Bandstructure of single boron in bulk configuration, but near the surface of a (111) plane vacancy disk. The Fermi level lies below the crossing of the π-$\pi*$ levels, suggesting that no electron transfer occurs. The top of the valence band is set to 0 eV, and states above it have their energies scaled by 1.3.

As with the nitrogen case, the boron disrupts the π-bonding configuration, which perturbs the bandstructure. However, as boron has fewer electrons, states lower in the gap are unoccupied. This means that there is only a small gap available for transitions, evident in figure 6.30.

It is possible that the energy of substitutional boron could be lowered by moving to a position nearer to the disk. Here, instead of being surrounded by bulk diamond it could bond to sp^2 carbon in the Pandey chains. It has been shown for boron in silicon that the second-nearest neighbour site to a vacancy is energetically favourable [135]. This would require further work to clarify.

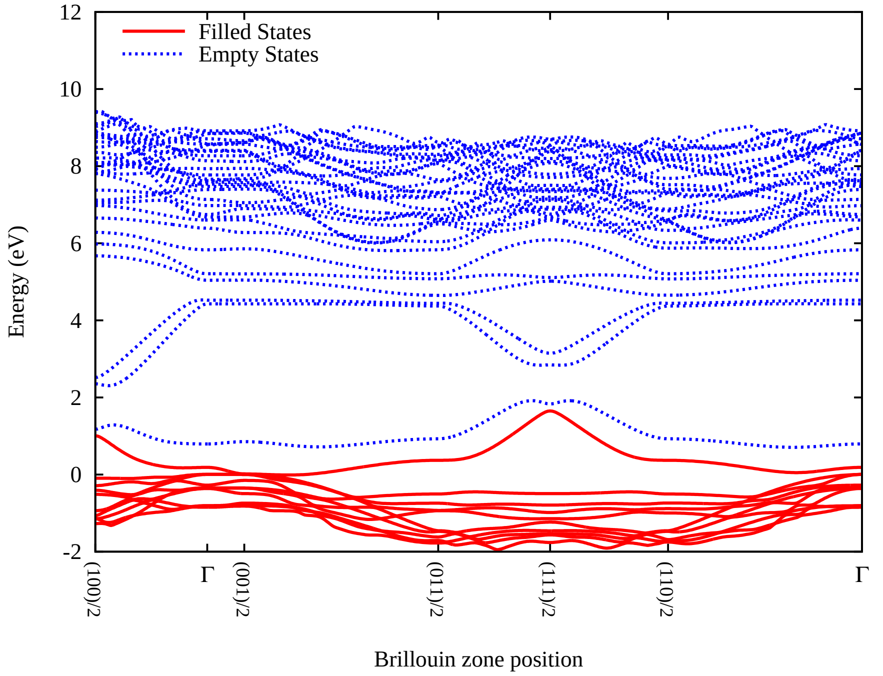


Figure 6.30: Bandstructure of single boron at the surface of a (111) plane vacancy disk. A small gap is introduced, but it is much less than the gap in the nitrogen-related defects. The top of the valence band is set to 0 eV, and states above it have their energies scaled by 1.3.

6.4 Conclusion

It has been shown that a disk of vacancies lying on the (111) plane is the most stable configuration for large vacancy clusters in diamond. The stability of the disk is due to the elimination of dangling bonds through the formation of π-bonded chains similar to the (2×1) reconstruction of the (111) surface.

The band structure of the (111) vacancy disk is quite distinct from that of the ideal 2×1 (001) diamond surface, where again a reconstruction leads to π-bonding. However, in this case the π-bonds do not overlap appreciably, and lead to only narrow π bands and π^* bands lying within the gap, of width less than 2 eV and separated by $\sim$1.5 eV [136]. To achieve very broad π and π^* bands, separated by a narrow gap, it is necessary for the π-bonds to be connected together in, for example, chains or as in a graphene sheet. Only in this case, will the optical properties approach those of brown diamond. Disks lying on (110) planes also do not exhibit the broad π and π^* bands which completely fill the gap and hence would not show the characteristic absorption of brown diamond.

The absorption coefficient of the disk exhibits the same featureless absorption profile as seen experimentally for brown diamond. Moreover, the highest local vibrational mode of the disk at 1496 cm^{-1} is close to a band detected experimentally at 1540 cm^{-1} in brown CVD diamond. Vacancy disks in the (111) plane are identified as a likely candidate for the origin of the brown colouration in diamond. In natural diamond, the experimental absorption bands vary more dramatically with energy which may reflect the presence of other vacancy centres. It may be that octahedral voids of vacancies having (111) surfaces could also explain the brown colour and these could also account for the very long 400 ps lifetime detected by positron annihilation in brown diamond [35, 112].

Two mechanisms for the loss of the brown colouration upon annealing are investigated. In CVD diamond, the surface of the disk is passivated with hydrogen, possibly explaining the growth of vibrational bands around 2900 cm^{-1} detected after the transformation of brown CVD diamond [34]. In natural type IIa diamond, the collapse of a large disk, and the formation of an optically inactive dislocation, is found to be energetically favourable for disks containing more than about 200 vacancies. Such a mechanism might account for the transformation of brown natural diamond, although it is unclear whether the disks are growing and then become unstable, or the activation barrier to the collapse and the formation of a dislocation loop are overcome at temperatures around 2000° C.

The interaction of nitrogen (specifically the A-centre) with a disk of vacancies lying on the (111) plane has been evaluated. The most stable configuration for a pair of nitrogen atoms at the disk surface is for a carbon pair to be between them, disrupting the Pandey chain. This leads to a gap in the bandstructure and an optical absorption onset at $\sim$1.7 eV. At lower densities of nitrogen the continuum absorption observed in type Ia brown diamond returns, with absorption bands at 2.4 eV (516 nm) and 0.7 eV (1770 nm), which are close matches to experiment. The total nitrogen density required is estimated at $\sim 10^{20}$ cm^{-3}, which is a realistic figure. This model is proposed as a possible cause of brown colour in type Ia diamond.

The model fails to completely explain the origin of pink colour in diamond, as no band at 3.17 eV (390 nm) is produced, but could be part of a system of defects which combine to produce the pink colour.

Chapter 7

Photoelastic Constants in Diamond and Silicon

7.1 Introduction

This work has been published in *Physica Status Solidi (a)* **203**, 3088 (2006).

When a strain is applied to a crystal there is a corresponding change in the refractive index, arising from a change in the dielectric function [137]. The photoelastic tensors relate the change of the inverse dielectric function with the applied strain. The presence of a dislocation or extended defect in a crystal causes a local strain field which in turn can cause a nominally isotropic material to become anisotropic and induce birefringence in light passing through cross-polarizers [138]. The birefringence due to dislocations in silicon has been observed [139] and modelled theoretically [140]. However, these calculations require the photoelastic tensor which is not well understood in diamond, compared to silicon [141].

The change in the inverse dielectric function due to strain is given by:

$$\delta\left(\varepsilon^{-1}\right)_{ij} = p_{ijkl}\mu_{kl} \approx -\left(\frac{1}{\varepsilon_b^2}\right)\delta\varepsilon_{ij} \tag{7.1}$$

where p_{ijkl} is the fourth rank photoelastic tensor, μ_{kl} is the strain tensor and ε_b is the value of the dielectric function in unstrained material. Adopting a contracted notation where 11 → 1, 22 → 2, 33 → 3 and 23,32 → 4, p_{ijkl} reduces to a symmetric matrix with three non-zero elements for silicon and diamond. These are p_{11},p_{12} and p_{44} and they can be found individually and in combination using various strains. For hydrostatic strain:

$$-\left(\frac{1}{\varepsilon_b^2}\right)\delta\varepsilon_{ij} = \left(\frac{\delta a}{a}\right)\begin{pmatrix} p_{11}+2p_{12} & 0 & 0 \\ & p_{11}+2p_{12} & 0 \\ 0 & 0 & p_{11}+2p_{12} \end{pmatrix} \tag{7.2}$$

where a is the lattice constant, so $\delta a/a$ is the percentage strain. Similarly for uniaxial strain along [001]:

$$-\left(\frac{1}{\varepsilon_b^2}\right)\delta\varepsilon_{ij} = \left(\frac{\delta a}{a}\right)\begin{pmatrix} p_{12} & 0 & 0 \\ 0 & p_{12} & 0 \\ 0 & 0 & p_{11} \end{pmatrix} \tag{7.3}$$

For uniaxial strain along [111], instead of a change in lattice parameter as a measure of strain, the change in a length l is considered:

$$-\left(\frac{1}{\varepsilon_b^2}\right)\delta\varepsilon_{ij} = \left(\frac{\delta l}{3l}\right)\begin{pmatrix} p_{11}+2p_{12} & 2p_{44} & 2p_{44} \\ 2p_{44} & p_{11}+2p_{12} & 2p_{44} \\ 2p_{44} & 2p_{44} & p_{11}+2p_{12} \end{pmatrix} \tag{7.4}$$

These three strains are thus sufficient to determine p_{11}, p_{12} and p_{44}.

7.2 Results

The photoelastic constants are calculated in two-atom unit cells, with a high Monkhorst-Pack sampling of 8^3 **k**-points required. The unit cells are taken and the lattice vectors changed to apply strain in the [001] and [111] directions and hydrostatically. The values of strain applied are ±0.5% for hydrostatic and [001] strain, and ±0.25% for strain along [111]. The values of p_{11}, p_{12} and p_{44} are found for an average of these compressive and tensile strains.

7.2.1 Silicon

Initially, values of p_{11}, p_{12} and p_{44} are calculated for silicon. This demonstrates the reliability of the method described, as the photoelastic constants for silicon are better studied [142, 143, 144, 145]. Previous *ab-initio* studies on the dielectric response of silicon [142] have reviewed the available experimental data. In only one case [143] were all the constants of interest found experimentally, so reference will only be made to the results from that work.

The use of a rigid shift to compensate for the LDA underestimation of the bandgap was tested. This shift affects all levels above the valence band top, to the same extent. It was used with some success by Levine *et al.* [142], and is often referred to as a 'scissors shift'. The scissors shift becomes important for defective crystals where states are introduced into the band gap and lead to additional optical transitions. To correct the AIMPRO bandgap requires a shift of 0.52 eV. Such a shift changes the dielectric constant so evaluation is made of the changes to the photoelastic constants when a scissors operator is used.

Table 7.1: Effect of scissors shift on the real part of the calculated dielectric constant ε_b, and the photoelastic constants for silicon. Experimental and previous *ab-initio* work is shown for comparison. All values refer to a wavelength of 3542 nm (0.35 eV).

	ε_b	$p_{11}+2p_{12}$	$p_{11}-p_{12}$	p_{44}
AIMPRO				
0 eV Scissor shift	13.670	-0.058	-0.109	-0.051
0.52 eV Scissor shift	12.195	-0.041	-0.116	-0.054
0.7 eV Scissor shift	11.764	-0.039	-0.119	-0.055
0.9 eV Scissor shift	11.327	-0.034	-0.121	-0.055
Experiment [143]	11.830	-0.054	-0.112	-0.051
Ab-initio [142]				
0.9 eV Scissor shift	10.900	-0.062	-0.118	-0.051

The photoelastic constants have been found using the calculated values of ε_b. It is shown in table 7.1 that increasing the scissor shift to compensate for the bandgap underestimation gives a better agreement for the real part of the dielectric constant with experiment. In contrast, the agreement for the low frequency values of the photoelastic constants becomes worse. An averaged combination of tensile and compressive strain is used for accuracy.

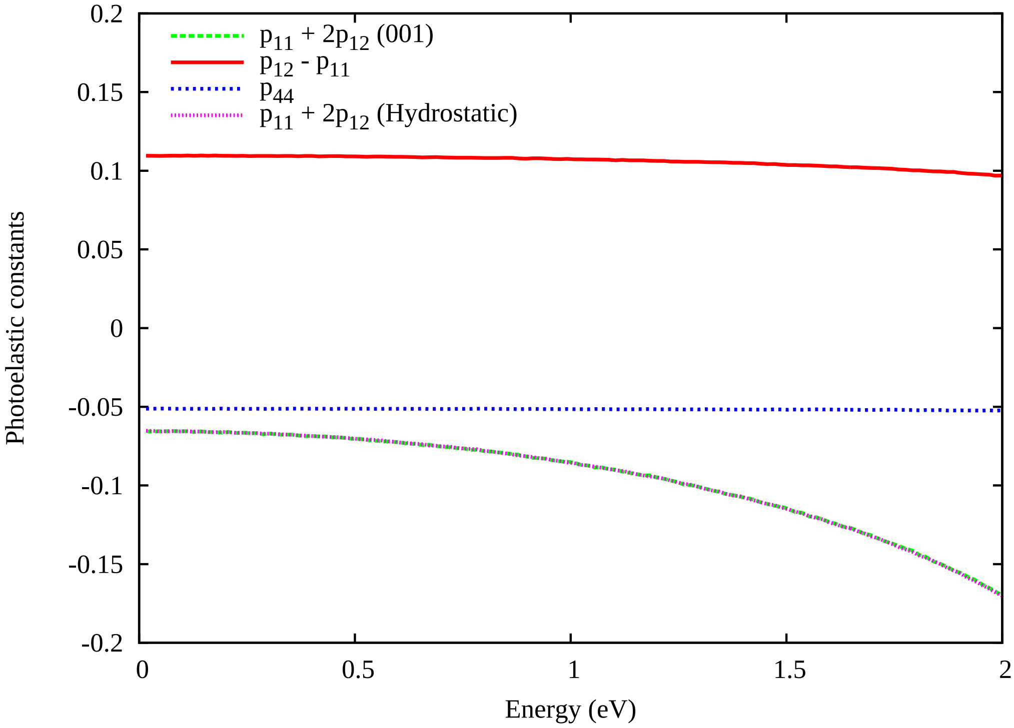


Figure 7.1: Photoelastic constants for silicon calculated from the average of compressive and tensile strain.

Figure 7.1 shows the variation of the calculated photoelastic constants with frequency for silicon. No scissors shift is applied, as described above. Comparison with experimental values [143] is shown in figure 7.2, and with previous *ab-initio* calculations [142] in figure 7.3. The values of $p_{11} - p_{12}$ are slightly larger than experiment by about 3% but the low dispersion with frequency is in good agreement. Excellent agreement is shown for p_{44}. The calculated values for $p_{11} + 2p_{12}$ are in good agreement with experiment at ~ 0.5 eV but disperse quite strongly with increasing frequency. As stated above, this is due to the small band gap and is lessened when a shift is used.

The calculated constants for silicon display relatively little dispersion except for the calculated values of $p_{11} + 2p_{12}$. If a scissors shift of 0.9 eV is used, $p_{11} + 2p_{12}$ becomes less sensitive to frequency but the overall values are increased. As described in the introduction, $p_{11} + 2p_{12}$ can be derived from hydrostatic strain or from strain imposed along [001]. Both methods lead to the same value of $p_{11} + 2p_{12}$.

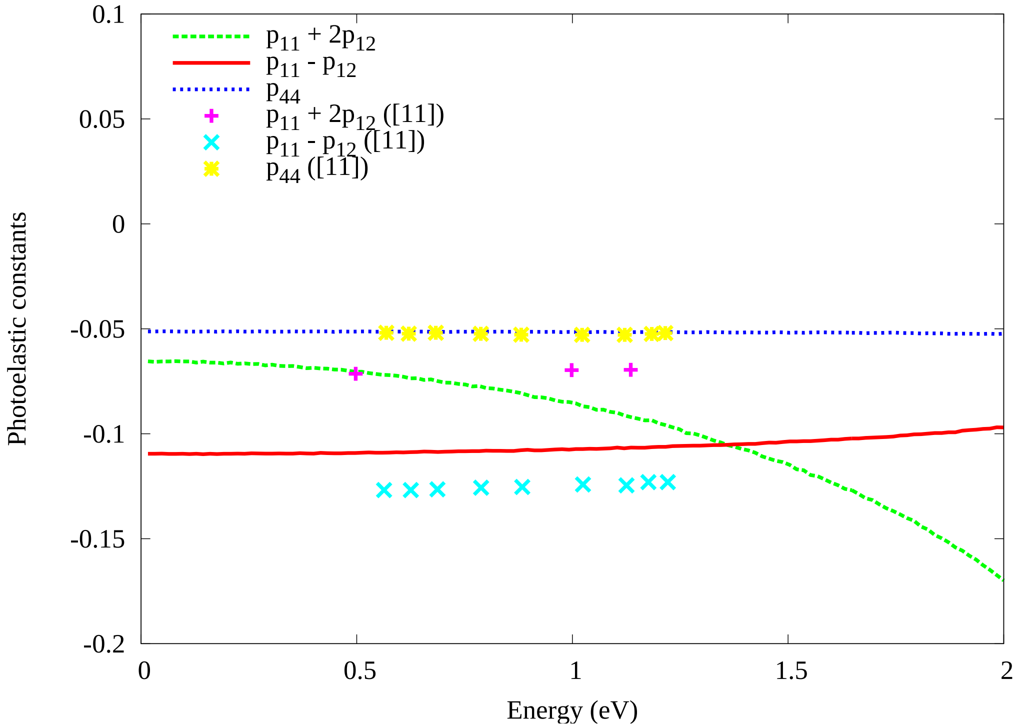


Figure 7.2: Calculated photoelastic constants for silicon using a zero scissors shift, compared to experimental data from [143].

7.2.2 Diamond

The values of the photoelastic constants for diamond in the literature differ widely and have been reviewed previously [141], where there is very little information on the variation of the constants with wavelength. Grimsditch *et al.* [141] evaluate one set of results as being ”most reliable”, these results are from [146]. For these reasons it is of interest to present a set of calculated values.

Table 7.2 shows the variation of the real part of the dielectric constant and photoelastic constants with scissor shift. A shift of 1.33 eV is required to correct the bandgap for diamond. As described for silicon the increase in shift causes the calculated photoelastic constants to diverge from the experimental results. It is notable that in the case of diamond, when the bandgap is corrected by a scissors shift of 1.33 eV the value of the dielectric constant ε_b is just less than 95% of the experimental value. In silicon, the required shift of 0.52 eV gives an ε_b of 103% of experiment. These results highlight the inconsistency of using a scissors shift, which cannot be relied upon to improve both the

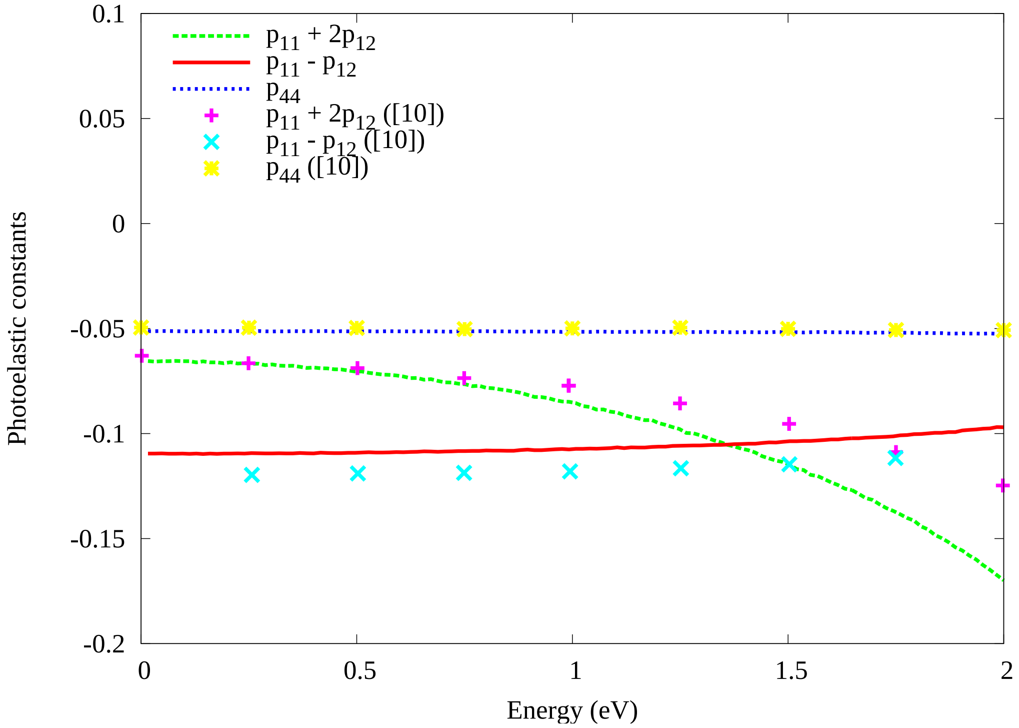


Figure 7.3: Calculated photoelastic constants for silicon using a zero scissors shift, compared to previous *ab-initio* work from [142].

dielectric constant and photoelastic constants for a given material.

Table 7.2: Effect of scissors shift on the real part of the calculated dielectric constant ε_b, and the calculated photoelastic constants for diamond. Experimental work is shown for comparison.

	Frequency (eV)	ε_b	$p_{11}+2p_{12}$	$p_{11}-p_{12}$	p_{44}
AIMPRO					
0 eV Scissor shift	2.27	6.172	-0.160	-0.335	-0.171
0.78 eV Scissor shift	2.27	5.755	-0.124	-0.331	-0.156
1.0 eV Scissor shift	2.27	5.662	-0.121	-0.331	-0.156
1.33 eV Scissor shift	2.27	5.526	-0.122	-0.330	-0.155
Experiment [107, 146]	2.27	5.819	-0.164	-0.292	-0.172
Experiment [147]	2.27	-	+0.12	-0.668	-0.162
Experiment [147, 141]	2.27	-	-0.1	-0.668	-0.162
Experiment [148]	2.30	-	-0.160	-0.292	-0.172

It is evident that $p_{11}-p_{12}$, is underestimated by approximately 14%, but $p_{11}+2p_{12}$ and p_{44} show very good agreement, with less than 3% and 1% difference respectively. Figure 7.4 shows the experimental data for diamond along with the calculated values for the photoelastic constants found for one frequency. The figure is in three parts for clarity.

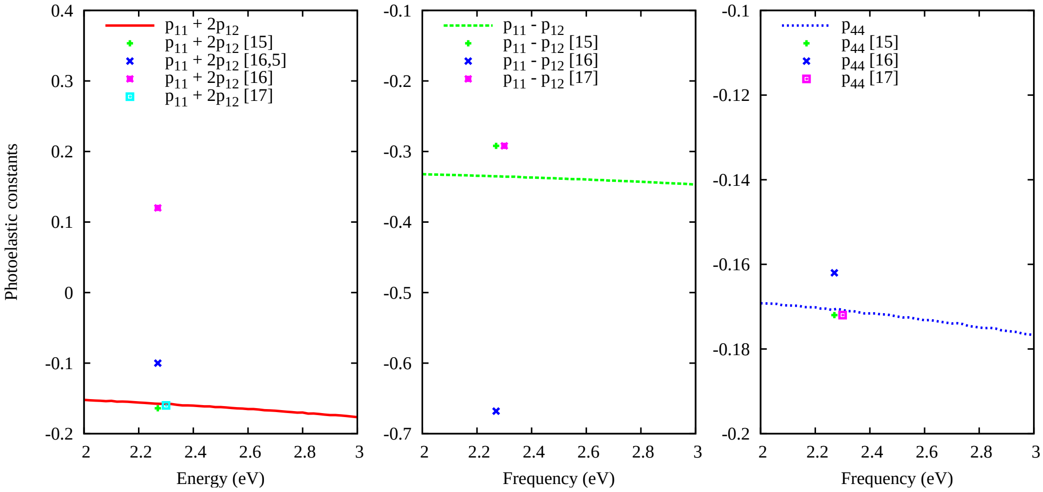


Figure 7.4: Photoelastic constants for diamond calculated with a zero value scissors shift. Experimental data is shown for comparison, taken from [146, 147, 148].

7.3 Strain of Dislocations

In a tetragonal solid such as diamond, the velocity of light along the three principal directions ([100], [110] and [111]) is different. Irradiating with polarized light then leads to a shift in the two polarized parts. This shift induces a phase difference at the exit point, which leads to birefringence.

If the sample to be analysed is placed between a polarizer and analyser which are perpendicular, the intensity of the transmitted light is given by [140]:

$$T = a^2 \sin^2(2\gamma) \sin^2(\delta/2) \tag{7.5}$$

where δ is the phase difference in the two principal directions of the crystal plane, a is the amplitude of the incident light and γ is the angle between the incident light vector and a principal direction. The relationship between these directions is shown more clearly in figure 7.5.

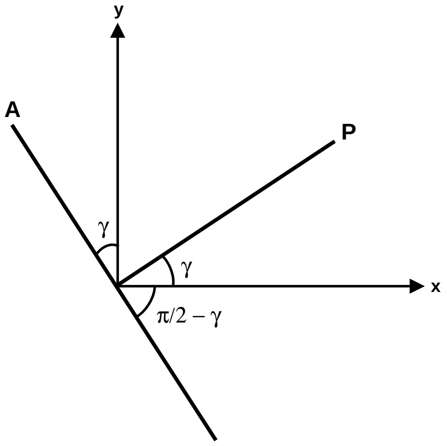


Figure 7.5: Relationship between the principal directions of the crystal and the polarizer/analyser directions P and A respectively.

If a dislocation is present in a crystal it will cause a strain effect in the area surrounding it, which can be characterised using the equations:

$$\mu'_{11} = \frac{A}{r^2}[x_1 - (1-2\nu)x_2] \tag{7.6}$$

$$\mu'_{22} = -\frac{A}{r^2}[x_1 + (1-2\nu)x_2] \tag{7.7}$$

$$\theta = \frac{1}{2}\tanh\left(\frac{x_2^2 - x_1^2}{2x_1x_2}\right) \tag{7.8}$$

where μ'_{11} and μ'_{22} are the strain values along the principal directions. θ can then be defined as the angle between the principal axis and the slip direction. The prefactor A is itself expandable to:

$$A = b/[4\pi(1-\nu)] \tag{7.9}$$

with b the Burgers vector magnitude. The angle γ between the principal x axis and the polarizer direction can now be redefined as $\gamma = \beta + \theta$ as shown in figure 7.6.

It can thus be shown [140] that the transmitted intensity can be expressed:

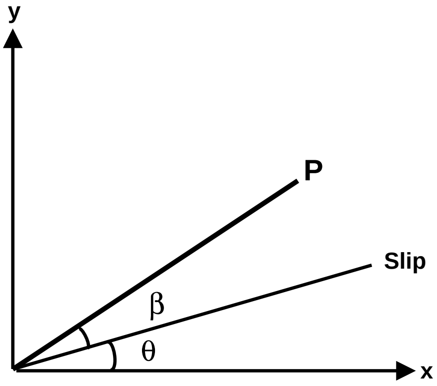


Figure 7.6: Relationship between the principal directions of the crystal and the polarizer direction. The slip plane of the dislocations is marked.

$$T(r,\psi) = (B^2/r^2)\cos^2(2\phi)\cos^2(\phi - B) \tag{7.10}$$

Where $\phi = \gamma + \pi/4$. The contours of constant intensity are given by the polar equation:

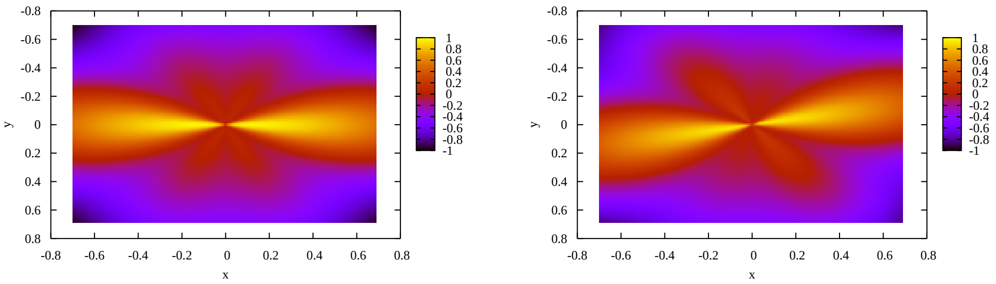
$$r^2 = (B^2/T)\cos^2(2\phi)\cos^2(\phi - B) \tag{7.11}$$

These contours are plotted in figures 7.7(a) – 7.7(g), for several different angles between the polarizer and the slip direction (β). It is clear that when β = 0° and 90° the plots are the same, and these plots are in exact agreement with previous work in silicon [140].

The ratio of the transmitted intensity to the incident intensity is defined:

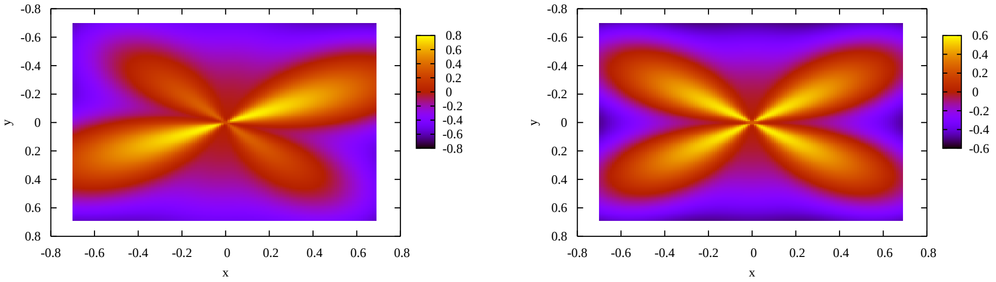
$$T_r \sim A^2C^2 4\pi t^2/\lambda^2 r^2 \tag{7.12}$$

Where t is the thickness of the material; in this case a typical thickness of 1 mm is assumed. C is defined as $p_{11} - 2p_{12}$, and is found from the values in table 7.2 to be -0.392, hence C^2 is 0.153. The magnitude of the Burgers vector for a pure edge dislocation in diamond is 2.5 Å hence A has a value of 2.48$\times$ 10^{-11} m. Inserting these values into equation 7.12 give the radial dependence of the strain:



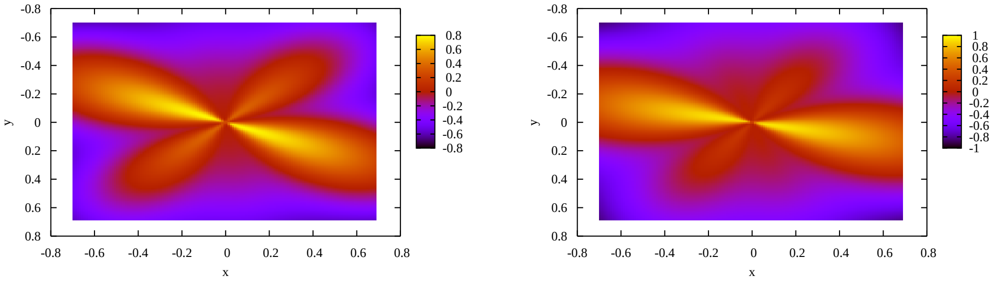
(a) β = 0°

(b) β = 15°



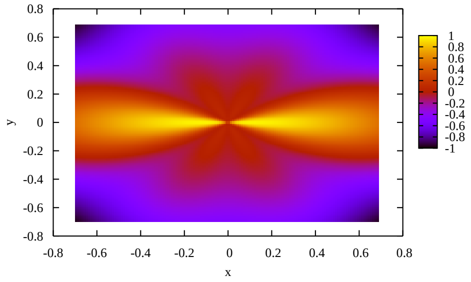
(c) β = 30°

(d) β = 45°



(e) β = 60°

(f) β = 75°



(g) β = 90°

Figure 7.7: Contours of constant diffraction intensity from equation 7.11, plotted with varying angles between the polarizer and the slip direction (β).

$$T_r \sim 1.81 \times 10^{-16}/r^2 \tag{7.13}$$

Thus at a radius of 1 μm from the dislocation core, the strain is 1.809$\times$ 10^{-4} the value at the centre - about 0.2%.

7.4 Conclusions

The photoelastic constants for diamond and silicon have been calculated using density functional theory applied to strained unit cells. Experimental data for silicon is accurately reproduced with absolute values and variation with frequency matched for almost all the constants. It is has been shown that the scissors shift technique of correcting the LDA underestimation of the bandgap does not uniformly improve the calculated photoelastic constants. The agreement for silicon gives confidence that the diamond values will accurately predict experimental results.

Experimental data for diamond is limited. However, the values calculated here agree well with one particular experimental set [142], which has previously been regarded as 'most reliable' [141]. In addition, the calculated results agree with the sign of the constants in the experimental set, confirming that $p_{11} + 2p_{12}$ should be a negative value.

It has been shown how the strain of dislocations can induce birefringence in polarized light, with the intensity caused by a pure edge dislocation used as an example. This strain falls off as r^2, to a low value at $\sim$1 μm from the dislocation core.

Chapter 8

Conclusions

Diamond is a scientifically interesting material with unique electrical, mechanical and thermal properties. These properties make diamond both cosmetically attractive and potentially important in next-generation electrical devices. Many features of diamond have been previously studied, such as point defects and dislocations, and these were overviewed in chapter 1. This thesis has focused on the optical properties of defects in an effort to come closer to understanding the origin of brown colour in diamond. Brown diamond has a unique absorption continuum across the bandgap, with no detectable onset at low energies. The source of colour is not known, but is vital for understanding the loss of colour which occurs during High-Pressure High-Temperature (HPHT) treatment.

The theoretical method used (AIMPRO) has been used on a range of materials, and is proven to be a good match for experimental results. A review of the inherent theory and its method of application is presented in chapter 2. In addition, a summary of relevant experimental techniques is given in chapter 3, as results from these techniques are referred to throughout this work. Full understanding of any scientific problem is made easier through a combination of theoretical and experimental work, hence as much of this work as possible is linked to known results.

8.1 Small Vacancy Clusters

Electron paramagnetic resonance (EPR) studies on diamond have revealed a range of small vacancy defects, which are identified as clusters and chains along $\langle 110 \rangle$. Due to their prevalence, a study is made of the properties of such defects. Vacancy chains of up to six vacancies are modelled, and vacancy clusters of up to 14 vacancies. The clusters are constructed using a bond-counting model previously demonstrated for clusters in silicon. DFTB theory (chapter 2) is used to check convergence of energy with cell size, with a cell of 216 atoms required to accurately model such defects.

Analysis of the formation energies of these defects in both $S=0$ and $S=1$ states indicates that the vacancy clusters are more stable than vacancy chains in all cases. In addition, the clusters of 6, 10 and 14 vacancies are extra-stable, due to their closed structure, a property also predicted for silicon. The 14 vacancy cluster has the lowest overall formation energy of 2.35 eV per vacancy. Apart from the four and six vacancy clusters, and the three and four vacancy chains, all defects are in the $S=1$ state for lowest energy, with unpaired electrons. These unpaired electrons will be expected to introduce bandgap energy levels, and an analysis of the Kohn-Sham energy states for each defect confirms this.

With the introduction of bandgap states a peak in absorption is observed around 3.5 – 4.5 eV range, with the energy of the peak maximum increasing with cluster size. This is unlike the absorption of brown diamond, so small vacancy clusters are discounted as a possibility for the source of brown colour, even as a combination. Initial work on the possibility of including π-bonded carbon in small vacancy defects indicates that this will not be feasible.

8.2 EELS of Dislocations

Dislocations are a prevalent defect in natural diamond, with a density of $\sim 10^9$ cm^{-2} observed by TEM. However the density in artificial CVD diamonds is five orders of magnitude lower. Previous *ab-initio* studies of dislocations in diamond predict that some types will have deep gap levels which could introduce optical absorption. The density of dislocations which is observed in HPHT treated brown (now colourless) diamond is not markedly different from untreated samples, so it can be concluded that few dislocations are created or destroyed during the process. If this is the case then the proposal that dislocations giving rise to absorption are converting to optically inactive types is considered.

The dielectric function, Electron Energy Loss (EEL) and absorption spectra are calculated for bulk diamond as a test of the technique and to verify if scaling is required to compensate for LDA bandgap underestimation. A good agreement is demonstrated for peaks in EEL and dielectric spectra. The static value of the real part of the dielectric function shows excellent agreement, with a calculated value of $\varepsilon_1 =$5.89 compared to 5.82 experimentally. The value where ε_2 falls to 0 is not so closely matched, with a value of 5.29 eV calculated, compared to 6.81 eV experimentally. This mismatch can be attributed to the broadening applied to the calculated points, which will lower the threshold of onset. In the case of the EEL spectra for diamond and graphite the main peaks at 35 eV (diamond) and 22 eV (graphite) are a very good match for experiment.

Two dislocation types are studied, the 90° glide and the 90° shuffle types. There is a potential that the shuffle type could convert to the glide type if supplied with sufficient energy. The formation energy of the glide type is much lower that the shuffle type, indicating that such a process could take place, but that it can be expected that there will be more glide type dislocations present. The glide type is optically inactive due to its reconstructed core atoms, with no absorption below $\sim$4 eV. In contrast the shuffle type has dangling bonds which introduce absorption throughout the gap, but the spectrum is not like that seen experimentally in brown diamond. A density of shuffle dislocations of $\sim 10^{11}$ cm^{-1} is calculated to be required to account for brown colour, however natural diamonds have a dislocation density of $\sim 10^9$ cm^{-1} and CVD diamonds even less. This eliminates dislocations as a possible source of brown colouration, but they may still have an important part to play as a source of vacancies or as defect traps.

8.3 (111) Plane Vacancy Disks

If a double bond is formed between two carbon atoms, it will introduce π and π^* states into the bandgap, hence is known as a π-bond. Previous *ab-initio* studies of diamond surfaces show that networks of π-bonds can form on the (111) surface, and the delocalization of the electrons over the network gives rise to dispersion of the π and π^* states. This dispersion causes a wide range of electron transitions to be available, potentially leading to broad optical absorption.

The (111) surface can be recreated internally by removal of the αB double plane in the (111) plane, with subsequent surface reconstruction into a Pandey chain configuration. Such a vacancy disk has a formation energy per vacancy of 1.46 eV, much lower that any of the vacancy clusters or chains previously modelled. This indicates that the disk will be a stable structure. The bandstructure of the disk displays the dispersed π and π^* states, as expected. These states give rise to a continuum bandgap absorption which is very like the experimentally observed brown diamond absorption, although our method does not take into account indirect transitions so the 5.5 eV onset is not reproduced. The density of vacancies required to match experimental magnitude is $\sim 10^{17}$ cm^{-3}. The Raman frequency of the disk is calculated to be 1494 cm^{-1}, which compares well with the 1540 cm^{-1} frequency observed in some brown diamonds.

If the (111) plane vacancy disk is taken to be the source of colour in brown diamond, a method for removal of colour must also be proposed. There is a major difference in the temperature required during HPHT treatment for natural and CVD material, which a temperature of $\sim$2100 $^\circ$ C required for the former but the lower temperature of $\sim$1500 $^\circ$ C needed for CVD material. This suggests that two different mechanisms are responsible. In CVD diamond there is a large amount of hydrogen from the growth process which can move to the disk surface, passivating the π-bonds and removing any optical absorption. In natural diamond there is nothing to check the growth of the disk until it will collapse into a stacking fault bounded by a dislocation loop. This is predicted to happen at a radius of 12 Å.

Type Ia diamond contains nitrogen in aggregated forms such as the V-N defect, N-N pair (A centre) and V-N_4 (B centre). The optical absorption of brown type Ia diamond is characterised by additional absorption around 550 nm (2.25 eV), so the possibility of nitrogen interacting with the disk to modify the absorption is investigated. The most stable configuration for nitrogen to be in at the disk surface is for a nitrogen pair to separate into a N-C-C-N arrangement. The bandstructure of this system is perturbed by the nitrogen, with an energy gap of 2 eV being opened up. This is attributed to the nitrogen being too close to one another, which was investigated with larger cells. These studies proved that the N-C-C-N defects needed to be sufficiently separated to allow a chain of $>$10 carbon atoms to form. Analysis of the absorption induced by the nitrogen perturbation reveals a peak at 2.3 eV, in excellent agreement with experiment. The total nitrogen density required to produce an absorption magnitude matching experiment is $\sim 10^{20}$ cm^{-3}.

8.4 Photoelastic Constants of Diamond

When strain is introduced into diamond, it changes the dielectric function both parallel and perpendicular to the direction of strain. This change is characterised by the photoelastic constants p_{11}, p_{12} and p_{44}. The constants are derived by inducing strain through variation of the lattice constants in the cell, and recalculating the dielectric function along x, y and z.

This procedure is first undertaken for bulk silicon, as there is more experimental data for comparison. It is also an opportunity to test the efficiacy of using a rigid ‘scissor’ shift to correct for LDA bandgap underestimation, as used in other *ab-initio* work. It is found that use of such a shift improves the values of the dielectric function somewhat, but that the values of the photoelastic constants have their best match with experimental data when no shift is used. Very good agreement for frequency dependence is also found.

The results from the silicon work give confidence in our values of the photoelastic constants for diamond. Experimental data for diamond is much more limited, and some values are markedly different from others. Good agreement is found for absolute values and frequency dependence with one particular set of results, the results also support the assignment of a negative value to $p_{11} + p_{12}$.

Strain around a dislocation has been shown to induce birefringence in incident polarised light. This birefringence can thus give an indication of the magnitude of the strain. The value is calculated for a pure edge dislocation in diamond and the radial dependence of the strain found to be $T_r \sim 1.81 \times 10^{-16}/r^2$.

8.5 Continuation Work

This work has investigated several ideas relating to brown colour in diamond. The concept of π-bonded carbon in vacancy disks is proposed as a colour source. However, there are further aspects of this work to be investigated to gain a complete understanding, and extensions to the work to give insight into new areas are also possible. Suggestions for further topics are elucidated upon below.

- **Colour removal in type Ia brown diamond** In chapter 6 two methods for removal of brown colour in type IIa diamond are proposed. In type Ia diamond, it is proposed that nitrogen is present at a disk surface, which may alter the colour removal process. In particular the stability of the disk and the formation of nitrogen-vacancy defects are of interest. As nitrogen disrupts the chains of π-bonded carbon atoms and introduces further free electrons, it can be thought to lower the disk stability. This needs to be checked against experimental data, and perhaps migration barrier calculations performed. If the disks are disintegrating then it is likely that nitrogen-vacancy defects may form. The stable configurations and expected defect densities can be calculated for comparison to experimental data on treated type Ia brown diamond.
- **π-bonded carbon in voids** One problem with the vacancy disk model is that it fails to account for the $\sim$400 ps lifetime seen in PAS. This lifetime extrapolates to a cluster of $\sim$50 vacancies if roughly spherical, although could be as much as 100 vacancies. The narrow cross-section of the disk gives it a lifetime comparable to a small vacancy cluster ($\sim$14 vacancies). It is possible to create an void by removing atoms in an octahedral pattern, which leaves a vacancy cluster with surfaces in the $\{111\}$ set of surfaces. These surfaces can conceivably form π-bonded chains, which can be expected to have a similar broad absorption as a vacancy disk. Careful investigation of the formation and energetics of such a void are required, but would require large cells in excess of 1000 atoms. This could only be approached though a more approximate method such as DFTB (chapter 2).
- **Vacancy disk generation** The formation of vacancy disks (or indeed voids) requires the creation of large numbers of vacancies, probably in areas of high density rather than an even distribution. There is also an observed link between brown colour and strain, with the brown colour stronger in areas that have reduced strain.

There is a suggestion that movement of dislocations (glide and climb) will reduce strain and release vacancies. As dislocations are often grouped together, the released vacancies will be close to each other allowing them to form vacancy defects. The plausibility of this proposal must be considered. The generation process in CVD material will be different as it has low strain and dislocation concentration, so another method must be postulated here.

- **Charge transfer to disk** Preliminary investigations on the vacancy disk model indicate the system energy can be lowered by transfer of an electron from a nearby defect to the disk. This may have implications for diamond electronics, as it will affect conduction and charge states on other defects. Investigations must be made on the density of electrons that a disk may hold, and the required defect density to support such an arrangement. It will be interesting to compare these results to experimental EPR work which detects high acceptor levels in nitrogen-doped CVD diamond. The source of these acceptor levels is as yet unknown.

In summary, there remain several areas of potential investigation related to brown diamond. *Ab-initio* modelling techniques are ideally placed to consolidate available experimental data, and help explain observations. As new developments in high-performance computing occur, and software is improved, ever more realistic models can be investigated and new areas of research become accessible. Continued research into the brown diamond problem, and diamond in general, is sure to lead to new and exciting discoveries and applications.

Bibliography

[1] H. Kroto *et al.*, Nature **318**, 162 (1985).

[2] R. Haufler *et al.*, J. Phys. C **94**, 8634 (1990).

[3] A. Thess *et al.*, Science **273**, 483 (1996).

[4] T. Ebbesen and P. Ajayan, Nature **358**, 220 (1992).

[5] L. Azaroff, *Introduction to solids* (McGraw-Hill, London, 1960).

[6] V. Blank *et al.*, J. Mater. Res. **12**, 3109 (1997).

[7] H. J. McSkimin and P. Andreatch Jr, J. Appl. Phys. **43**, 2944 (1972).

[8] *CRC handbook of chemistry and physics*, 75th ed., edited by D. R. Lide (CRC Press, London, 1995).

[9] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley-Interscience, New York, 1981).

[10] S. Barman and G. Srivastava, Phys. Rev. B **77**, 073301 (2006).

[11] W. Kaiser and W. L. Bond, Phys. Rev. **115**, 857 (1959).

[12] I. Pomeranchuk, J. Phys. U.S.S.R **6**, 237 (1942).

[13] D. Bylander and L. Kleinman, Phys. Rev. B **41**, 7868 (1990).

[14] J. Cifre, J. Puigdollers, M. Polo, and J. Esteve, Diamond and Related Mater. **3**, 628 (1994).

[15] R. Burns *et al.*, Diamond and Related Mater. **8**, 1433 (1999).

[16] G. Braunstein and R. Kalish, J. Appl. Phys. **54**, 2106 (1983).

[17] I. Sakaguchi *et al.*, Phys. Rev. B **60**, R2139 (1999).

[18] K. Okano *et al.*, Appl. Phys. A **51**, 344 (1990).

[19] S. A. Kajihara, A. Antonelli, J. Bernholc, and R. Car, Phys. Rev. Lett. **66**, 2010 (1991).

[20] S. J. Sque, R. Jones, J. P. Goss, and P. R. Briddon, Phys. Rev. Lett. **92**, 017402 (2004).

[21] R. Robertson, J. Fox, and A. Martin, Phil. Trans. R. Soc. A **232**, 463 (1934).

[22] J. Walker, Rep. Prog. Phys. **42**, 1605 (1979).

[23] K. Iakoubovskii and A. Stesmans, J. Phys.: Condens. Matter **14**, R467 (2002).

[24] S. D. Berger and S. J. Pennycook, Nature **298**, 635 (1982).

[25] B. Willems, L. Nistor, C. Ghica, and G. Van Tendeloo, 2005, presented at SBDD-X, LUC Belgium, February 2005.

[26] P. Martineau *et al.*, Gems Gemol. **40**, 2 (2004).

[27] J. P. Hirth and J. Lothe, *Theory of dislocations*, *Materials Science and Engineering* (McGraw-Hill, New York, 1968).

[28] A. T. Blumenau *et al.*, Phys. Rev. B **65**, 205205 (2002).

[29] A. T. Blumenau, Ph.D. thesis, Department Physik der Fakultät für Naturwissenschaften, Universität Paderborn, Germany, 2002.

[30] A. T. Collins, H. Kanda, and H. Kitawaki, Diamond and Related Mater. **9**, 113 (2000).

[31] D. Fisher and R. A. Spits, Gems Gemol. **36**, 42 (2000).

[32] R. Burns, D. Fisher, and R. Spitz, Patent Application WO 01/72406, April 04 2001.

[33] D. Twitchen, P. Martineau, and G. Scarsbrook, Patent Application WO 2004/022821, March 18 2004.

[34] S. Charles *et al.*, Phys. Status Solidi **201**, 2473 (2004).

[35] V. Avalos and S. Dannefaer, Physica B **340**, 76 (2003).

[36] K. Saarinen, 2005, presented at Diamond 2006, Oxford, July 2005.

[37] V. Avalos and S. Dannefaer, Diamond and Related Mater. **14**, 155 (2005).

[38] A. Pu, T. Bretagnon, D. Kerr, and S. Dannefaer, Diamond and Related Mater. **9**, 1450 (2000).

[39] J. N. Lomer and A. M. A. Wild, Radiation Effects **17**, 37 (1973).

[40] K. Iakoubovskii and A. Stesmans, Phys. Status Solidi A **201**, 2509 (2004).

[41] K. Iakoubovskii and A. Stesmans, Phys. Rev. B **66**, 045406 (2002).

[42] K. Iakoubovskii and G. Adriaenssens, J. Phys.: Condens. Matter **12**, L77 (2000).

[43] J. P. Goss *et al.*, Phys. Rev. B **67**, 165208 (2003).

[44] F. De Weerdt and J. Van Royen, Diamond and Related Mater. **10**, 474 (2001).

[45] L. Massi *et al.*, Diamond and Related Mater. **14**, 1623 (2005).

[46] A. Collins, J. Gemm. **18**, 37 (1982).

[47] F. Raal, Proc. Phys. Soc. **71**, 846 (1958).

[48] D. Fisher, 2006, private communication.

[49] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, international ed. (Saunders College, Philadelphia, 1976).

[50] M. Born and J. R. Oppenheimer, Ann. Phys. **84**, 457 (1927).

[51] V. Fock, Z. Phys. **61**, 126 (1930).

[52] V. Fock, Z. Phys. **62**, 795 (1930).

[53] J. C. Slater, Phys. Rev. **35**, 210 (1930).

[54] J. C. Slater, Phys. Rev. **81**, 385 (1951).

[55] J. C. Slater, Phys. Rev. **82**, 538 (1951).

[56] R. Gáspár, Acta Phys. Acad. Sci. Hung **3**, 263 (1954).

[57] L. H. Thomas, Proc. Cambridge Philosophical Soc. **23**, 542 (1927).

[58] E. Fermi, Z. Phys. **48**, 73 (1928).

[59] P. Dirac, Proc. Cambridge Philosophical Soc. **26**, 376 (1930).

[60] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).

[61] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[62] G. Srivastava, *Theoretical Modelling of Semiconductor Surfaces*, international ed. (World Scientific, Singapore, 1999).

[63] M. R. Norman and J. P. Perdew, Phys. Rev. B **28**, 2135 (1983).

[64] R. Jones *et al.*, in *Special defects in semiconducting materials*, Vol. 71 of *Solid State Phenomena*, edited by R. P. Agarwala (Scitech Publications Ltd., Zuerich-Uetikon, Switzerland, 2000), Chap. The interaction of hydrogen with deep level defects in silicon, pp. 173–248.

[65] R. Jones, B. Coomer, and P. Briddon, J. Phys.: Condens. Matter **16**, 2643 (2004).

[66] C. Hartwigsen, S. Goedecker, and J. Hutter, Phys. Rev. B **58**, 3641 (1998).

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

[68] M. Musgrave and J. Pople, Proc. R. Soc. **268**, 474 (1962).

[69] J. Hillier and R. Baker, J. Appl. Phys. **15**, 663 (1944).

[70] H. Ehrenreich and M. Cohen, Phys. Rev. **115**, 786 (1959).

[71] S. Adler, Phys. Rev. **126**, 413 (1962).

[72] P. Nozières and D. Pines, Phys. Rev. **113**, 1254 (1959).

[73] V. Olevano, M. Palummo, G. Onida, and R. Del Sole, Phys. Rev. B **60**, 14224 (1999).

[74] W. R. L. Lambrecht and S. N. Rashkeev, Phys. Status Solidi B **217**, 599 (2000).

[75] L. X. Benedict *et al.*, Solid State Commun. **112**, 129 (1999).

[76] G. Seifert, H. Eschrig, and W. Bieger, Zeitschrift fur Physikalische Chemie Neue Folge **267**, 529 (1986).

[77] D. Porezag *et al.*, Phys. Rev. B **51**, 12947 (1995).

[78] W. Foulkes and R. Haydock, Phys. Rev. B **39**, 12520 (1989).

[79] Y.-T. Shen, D. Bylander, and L. Kleinman, Phys. Rev. B **36**, 3465 (1987).

[80] I. Watt, *The principles and practise of electron microscopy*, 2nd ed. (Camridge University Press, Cambridge, 1997).

[81] E. Abbe, Archiv fur mikroskopische Anatomie **9**, 413 (1873).

[82] L. de Broglie, Phil. Mag. **47**, 446 (1924).

[83] P. Nellist, Physics World **18**, 24 (2005).

[84] O. Scherzer, J. Appl. Phys. **20**, 20 (1949).

[85] R. Feynman, Eng. and Sci. **22**, (1960).

[86] P. Kovarik, E. Bourdon, and R. Prince, Phys. Rev. B **48**, 12123 (1993).

[87] R. Weidner and R. Sells, *Elementary Modern Physics* (Allyn and Bacon, London, 1980).

[88] W. Demtröder, *Laser spectroscopy: Basic Concepts and Instrumentation* (Springer, International, 1998).

[89] D. Gardiner and P. Graves, *Practical Raman Spectroscopy*, 1st ed. (Springer Verlag, Berlin, 1989).

[90] V. Raman and K. Krishnan, Nature **121**, 501 (1928).

[91] D. Long, *Raman Spectroscopy*, international ed. (McGraw-Hill, International, 1977).

[92] T. Gfroerer, in *Encyclopedia of Analytical Chemisty*, edited by R. Meyers (Wiley, Chichester, 2000), Chap. Photoluminescence in Analysis of Surface and Interfaces.

[93] I. Procházka, Mater. Struct. **8**, 55 (2001).

[94] V. Avalos and S. Dannefaer, Phys. Rev. B **58**, 1331 (1998).

[95] U. Bangert *et al.*, Phil. Mag. **86**, 4757 (2006).

[96] R. Barnes *et al.*, Journal Phys. Conf. Series **26**, 157 (2006).

[97] D. J. Chadi and K. J. Chang, Phys. Rev. B **38**, 1523 (1988).

[98] T. E. M. Staab *et al.*, Phys. Rev. B **65**, 115210 (2002).

[99] B. J. Coomer *et al.*, Physica B **273–274**, 520 (1999).

[100] B. Hourahine *et al.*, Phys. Rev. B **61**, 12594 (2000).

[101] D. Makhov and L. Lewis, Phys. Rev. Lett. **92**, 5504 (2004).

[102] S. K. Estreicher, J. L. Hastings, and P. A. Fedders, Appl. Phys. Lett. **70**, 432 (1997).

[103] K. Pandey, Phys. Rev. B **25**, 4338 (1982).

[104] A. Scholze, W. G. Schmidt, and F. Bechstedt, Phys. Rev. B **53**, 13725 (1996).

[105] A. Kolodzie, Ph.D. thesis, Magdalene College, University of Cambridge, 2003.

[106] A. Kolodzie, M. Murfitt, and A. Bleloch, Institute Phys. Conf. Series **168**, 247 (2001).

[107] H. Philipp and E. Taft, Phys. Rev. **127**, 159 (1962).

[108] R. Roberts and W. Walker, Phys. Rev. **161**, 730 (1967).

[109] S. Logothetidis, J. Petelas, and S. Ves, J. Appl. Phys. **79**, 1040 (1996).

[110] J. Fink *et al.*, Phys. Rev. B **30**, 4713 (1984).

[111] C. J. Fall *et al.*, Phys. Rev. B **65**, 205206 (2002).

[112] J.-M. Mäki *et al.*, 2005, presented at SBDD-X, LUC Belgium, February 2005.

[113] J. Slonczewski and P. Weiss, Phys. Rev. **109**, 272 (1958).

[114] O. Williams, Semicond. Sci. Technol. **21**, R49 (2006).

[115] S. Sque, R. Jones, and P. Briddon, Phys. Rev. B **73**, 085313 (2006).

[116] R. Hood, P. Kent, R. Needs, and P. Briddon, Phys. Rev. Lett. **91**, 6403 (2003).

[117] L. Hounsome *et al.*, Phys. Status Solidi A **202**, 2182 (2005).

[118] T. Pedersen, Phys. Rev. B **67**, 113106 (2003).

[119] V. Vorlíček *et al.*, Diamond and Related Mater. **6**, 704 (1997).

[120] M. Hanfland, H. Beister, and K. Syassen, Phys. Rev. B **39**, 12598 (1989).

[121] L. Nistor *et al.*, Diamond and Related Mater. **6**, 159 (1997).

[122] A. Collins, J. Phys.: Condens. Matter **14**, 3743 (2002).

[123] I. Kiflawi, A. Mainwood, H. Kanda, and D. Fisher, Phys. Rev. B **54**, 16719 (1996).

[124] R. Cruddace, 2006, private communication.

[125] R. Chin *et al.*, Phys. Rev. B **52**, 5985 (1995).

[126] A. Collins, Diamond and Related Mater. **9**, 417 (2000).

[127] G. Davies *et al.*, Phys. Rev. B **46**, 13157 (1992).

[128] A. Pu, V. Avalos, and S. Dannefaer, Diamond and Related Mater. **10**, 585 (2001).

[129] A. Collins and L. Cheng-Han, J. Phys.: Condens. Matter **14**, L467 (2002).

[130] C. Cao, L. Giling, and P. Alkemade, Diamond and Related Mater. **4**, 775 (1995).

[131] C. Cao *et al.*, J. Appl. Phys. **79**, 1357 (1996).

[132] S. Han, J. Ihm, S. Louie, and M. Cohen, Phys. Rev. Lett. **80**, 995 (1998).

[133] L. Hounsome *et al.*, Phys. Rev. B **73**, 125203 (2006).

[134] M. Heggie, 2006, private communication.

[135] J. Adey *et al.*, Phys. Rev. B **71**, 165211 (2005).

[136] S. Sque, R. Jones, and P. Briddon, Phys. Status Solidi A **202**, 2091 (2005).

[137] J. Nye, *Physical properties of crystals: their representation by tensors and matrices* (Clarendon, Oxford, 1957).

[138] S. Mendelson, J. Appl. Phys. **32**, 1999 (1961).

[139] W. Bond and J. Andrus, Phys. Rev. **101**, 1211 (1956).

[140] R. Bullough, Phys. Rev. **110**, 620 (1958).

[141] M. Grimsditch and A. Ramdas, Phys. Rev. B **11**, 3139 (1975).

[142] Z. Levine *et al.*, Phys. Rev. B **45**, 4131 (1992).

[143] D. Biegelsen, Phys. Rev. Lett. **32**, 1196 (1974).

[144] C. Higginbotham, M. Cardona, and F. Pollak, Phys. Rev. Lett. **184**, 821 (1969).

[145] M. Chandrasekhar, M. Grimsditch, and M. Cardona, Phys. Rev. B **18**, 4301 (1978).

[146] W. Schneider, Ph.D. thesis, Materials Research Institute, Pennsylvania State University, United States of America, 1970.

[147] G. Ramachandran, Proc. Indian Acad. Sci. **32**, 171 (1950).

[148] R. Denning, A. Giardini, E. Poindexter, and C. Slawson, Acta Metall. **42**, 556 (1957).