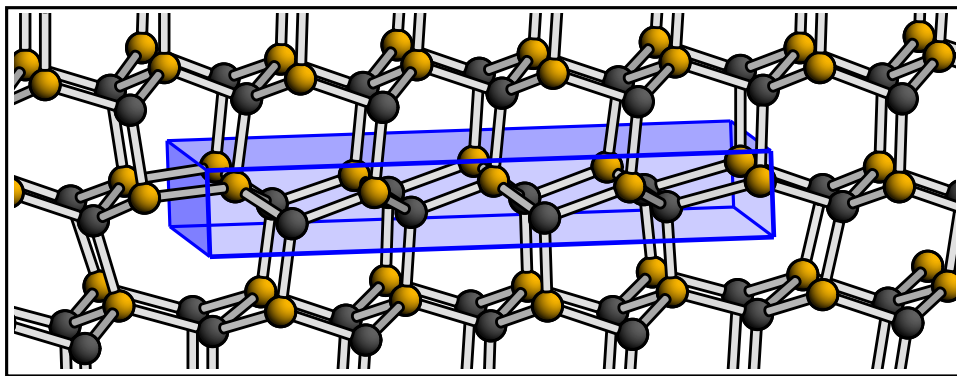# The Modelling of Dislocations in Semiconductor Crystals

Alexander Thorsten Blumenau

The Modelling of Dislocations in Semiconductor Crystals

Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
vorgelegt dem
Department Physik der Fakultät für Naturwissenschaften an
der Universität Paderborn

Alexander Thorsten Blumenau

Paderborn, 2002

Von der Fakultät für Naturwissenschaften der Universität Paderborn als Dissertation genehmigt.

Tag der Einreichung: 22. Oktober 2002
Tag der mündlichen Prüfung: 13. Dezember 2002

Promotionskommission

Vorsitzender	Prof. Dr. rer. nat. Artur Zrenner
Erstgutachter	Prof. Dr. rer. nat. Thomas Frauenheim
Zweitgutachter	Prof. Dr. rer. nat. Harald Overhof
Beisitzer	PD Dr. rer. nat. Siegmund Greulich-Weber

Abstract

This thesis studies dislocations in semiconductor crystals by means of theoretical modelling.
In particular, the work presented is focussed on dislocations in diamond and silicon carbide
and addresses technologically relevant effects.

A thorough description of dislocation effects covers a wide range of length scales and requires a rather precise modelling of the electronic structure of a small core region at one end
to long range elastic effects at the other end of the scale. This can hardly be achieved with
one method only. Therefore, it is one of the main objectives of this work, *to give a more complete description of dislocations in semiconductor crystals by combining different theoretical methods*:
Density functional theory (DFT) forms the basis for quantum mechanical atomistic calculations. Within DFT a pseudopotential approach is used to obtain electronic structures. A
more approximate and far less computationally expensive DFT-based tight-binding method
allows the prediction of core structures and energies embedded in larger models, representing a more extended region of the crystal. And finally, linear elasticity theory enables one to
describe long range elastic effects at almost no computational costs.

After introducing the methods to be used and presenting the basics of dislocation theory,
the dislocations of the $\{111\} \langle 110 \rangle$ slip system in diamond serve as an example of how
the different methods are combined in this work. Straight perfect as well as dissociated
dislocations are investigated. The resulting low energy core geometries can be used as input
coordinates for the simulation of high-resolution transmission electron microscopy images,
and the calculation of the electronic structure allows the modelling of electron energy-loss
spectra. Both are compared with experimental data.

Furthermore, the thermally activated glide motion of Shockley partial dislocations in diamond is modelled in a process of kink formation and migration.

Summarising, the results obtained support an annealing scenario for natural brown diamond involving a transition from shuffle to glide character at the dislocation core, which
might explain the decolouring observed under high-pressure, high-temperature treatment
— a process still not understood, but of utmost interest to the international diamond trade.

In modern silicon carbide technology a mechanism of recombination-enhanced dislocation glide is believed to play the key role in the observed disastrous degradation of bipolar
devices under forward bias. To shed some light on this mechanism, in this work the different
partial dislocations involved and their glide motion are modelled. The resulting electronic
structures and glide activation energies are then directly related to recent experimental observations.

Keywords

dislocations, diamond, silicon carbide, SiC, density functional theory, elasticity theory

PACS

61.72.Lk Linear defects: dislocations, disclinations
61.72.Bb Theories and models of crystal defects
61.72.-y Defects and impurities in crystals; microstructure
71.15.Nc Total energy and cohesive energy calculations
71.55.-i Impurity and defect levels
61.72.Ff Direct observation of dislocations and other defects

Kurzfassung

Diese Dissertation befaßt sich mit der theoretischen Modellierung von Versetzungen in Halbleiterkristallen. Ein besonderes Augenmerk gilt dabei den technologierelevanten Auswirkungen von Versetzungen in Diamant und Siliziumkarbid.

Eine umfassende Beschreibung der durch Versetzungen hervorgerufenen Effekte erstreckt sich über mehrere Größenordnungen auf der Längenskala — angefangen von der elektronischen Struktur einer kleinen Core-Region, bis hin zu langreichweitigen elastischen Effekten. Da dies alles mit einer einzigen Methode nicht in jeweils ausreichender Genauigkeit erreicht werden kann, ist es ein Ziel dieser Arbeit, *durch die Kombination unterschiedlicher theoretischer Methoden eine umfassendere Beschreibung von Versetzungen in Halbleiterkristallen zu erreichen*: Die Dichtefunktionaltheorie (DFT) bildet die Grundlage für atomistische quantenmechanische Rechnungen, in denen elektronische Strukturen in einem DFT-Pseudopotentialansatz bestimmt werden. Im Gegensatz dazu erfolgt die Vorhersage von atomaren Strukturen und deren Energien unter Verwendung einer DFT-basierten Tight-Binding Methode. Diese ist zwar approximativer, aber deshalb auch rechentechnisch wesentlich effizienter und erlaubt somit die Einbettung von Versetzungen in Modelle mit einer größeren Anzahl von Atomen. Die Beschreibung der langreichweitigen elastischen Effekte ist schließlich mit Hilfe linearer Elastizitätstheorie unter minimalem Rechenaufwand möglich.

Nach einer Einführung in die verwendeten Methoden und die Grundlagen der Versetzungslehre wird anhand des $\{111\} \langle 110 \rangle$-Versetzungssystems in Diamant die Kombination der verschiedenen Methoden demonstriert. Dabei werden sowohl geradlinige perfekte als auch dissoziierte Versetzungen betrachtet. In beiden Fällen erfolgt ein direkter Vergleich mit experimentellen Daten: Die resultierenden Core-Geometrien können zur Simulation von transmissions-elektronenmikroskopischen Bildern verwendet werden und die Berechnung der elektronischen Struktur ermöglicht die Modellierung von Elektronen-Energieverlust-Spektren.

Desweiteren wird das thermisch aktivierte Gleiten von Shockley-Partialversetzungen in Diamant durch einen Prozeß modelliert, dessen wesentliche Schritte durch die Bildung von Knicken ("kinks") in der Versetzunglinie und deren Migration gegeben sind.

Insgesamt unterstützen die erhaltenen Ergebnisse schließlich ein Ausheil-Szenario für braunen Naturdiamant, in dem Versetzungen mit "Shuffle"-Charakter in solche mit "Glide"-Charakter überführt werden. Ein solcher Mechanismus könnte die beobachtete Entfärbung brauner Diamanten unter extrem hohem Druck und hoher Temperatur erklären. Dieser bisher unverstandene Prozeß ist von größtem Interesse im internationalen Diamanthandel.

Ein aktuelles Problem der modernen Siliziumkarbid-Technologie stellt die beobachtbare katastrophale Degradation von unter Vorwärtsspannung betriebenen bipolaren Bauelementen dar. Diese ist vermutlich auf ein sogenanntes rekombinationsverstärktes Versetzungsgleiten zurückzuführen. Um diesen Mechanismus besser zu verstehen, werden in dieser Arbeit die verschiedenen beteiligten Partialversetzungen und ihre Gleitbewegung modelliert. Die resultierenden elektronischen Strukturen und Gleitaktivierungsenergien werden im Anschluß direkt mit aktuellen experimentellen Beobachtungen in Verbindung gebracht.

Schlagwörter

Versetzungen, Diamant, Siliziumkarbid, SiC, Dichtefunktionaltheorie, Elastizitätstheorie

Contents

List of Figures

List of Tables

Introduction

Real semiconductor crystals contain numerous defects which perturb the perfect lattice periodicity and influence its electrical and mechanical properties. Thus, these defects play an important role in semiconductor technology: Some defects have useful effects and are crucial for device design (e.g. doping to achieve n- or p-type conductivity), while others have a disastrous or destructive effect on devices and their performance. With regard to dimensionality, one has to distinguish between point-like defects (lattice vacancies, interstitial atoms or impurities), line defects (e.g. dislocations), planar defects (e.g. stacking faults or domain boundaries) and volume defects (e.g. voids or precipitates).

Since experiments alone often cannot yield information about the exact origin of defect-related effects or about how to suppress unwanted effects or promote those which are desirable, we need theoretical models and calculations to interpret the experimental data.

This thesis studies the atomic structures, energies and electronic properties of dislocations in semiconductor crystals by means of theoretical calculations and modelling. Whenever possible (and valuable) results are compared with experimental data. The prediction of the mechanical properties of dislocated crystals, however, is clearly outside the scope of this work.

Objectives

In the current literature on dislocations in semiconductors it is not uncommon to find investigations to be restricted to only one or two aspects of the defect. This very often stems from the wide range of length scales and precisions needed for a thorough description. The range spans from the very precise modelling of the electronic structure of a rather small core region at one end to the calculation of long range elastic lattice effects at the other end of the scale. For the former, quantum mechanical approaches are crucial — the latter however barely needs any quantum mechanics.

This wide range cannot be covered by one method only. Therefore it is one of the main objectives of this work, *to give a more complete description of dislocations in semiconductor crystals by means of combining different theoretical methods and discussing their overlap*: Density functional theory (DFT) forms the basis for quantum mechanical electronic and atomistic calculations. Within DFT a pseudopotential approach (implemented with a localised basis in the AIMPRO code [17]) is used to obtain precise electronic structures. A more approximate and far less computationally expensive

DFT-based tight-binding method (DFTB [18, 19]) allows the prediction of core structures and energies embedded in larger models, representing a more extended region of the crystal. Finally, linear elasticity theory enables us to describe long range elastic effects at almost no computational costs. In the course of this work, the afore mentioned combination of methods has proven very effective.

Outline

Chapter 1 introduces the different theoretical methods used in this work to model defects and dislocations in particular. Here we have to distinguish between DFT-based methods, namely the local density pseudopotential approach (AIMPRO) as well as DFT-based tight-binding (DFTB), and elasticity theory as a continuum theory. As mentioned, these methods represent different levels of approximation to the description of real semiconductor crystals and their defects.

Some readers might not be very familiar with the theory of dislocations. Thus, to provide the necessary background, the basic concepts of dislocation theory are introduced in Chapter 2. This includes selected results from linear elasticity theory which either help in the general understanding or will be useful in later chapters.

To prepare for the subsequent examples, Chapter 3 gives an overview of the low energy types of dislocations in tetrahedrally bonded semiconductors.

With Chapter 4 the introduction of methods and dislocation basics is followed by the first example: Dislocations in diamond. In terms of methodology this is surely the central and most important chapter, as it shows in detail how the three theoretical methods mentioned above are combined to describe the different aspects of a dislocation on different length scales. With the $\{111\} \langle 110 \rangle$ slip system in diamond as an example, the core structure, core energy, elastic long range energy terms, the dissociation into partial dislocations and dislocation glide motion are modelled. The resulting geometries can be used as input coordinates for the simulation of high-resolution transmission electron microscopy (HRTEM) images and the calculation of the electronic structure allows the modelling of electron energy-loss spectra (EELS) which can be compared with experimental data. The results obtained support a specific scenario explaining the high-pressure, high-temperature decolouring of brown diamond.

After this very detailed example of a homonuclear system Chapter 5 discusses dislocations in the compound semiconductor silicon carbide. In particular the role of Shockley partial dislocations and stacking faults in the technology-relevant degradation of silicon carbide PiN-diodes is examined.

The appendices mainly give additional considerations and calculations that would disturb the flow of the main text if included there. In Appendix A linear elasticity theory is applied explicitly to some simple dislocation-related problems. This should give the reader a feeling for the general procedure. The Appendices B and C are rather specific as they contain a brief presentation of the important experimental techniques — especially (HR)TEM and EELS — which yield information about the atomic and the electronic structure of dislocations in semiconductors and thus are directly related to the properties calculated in this work.

Chapter 1

Modelling the Crystal: Theories and Methods

This chapter will give a *brief* overview of the two DFT-based methods used in this work on the one hand, and linear elasticity theory on the other. As will be shown from Chapter 4 onwards, the combination of the three methods allows us to cover the description of dislocation properties on the whole length scale — ranging from below interatomic distances to the continuum elastic limit.

1.1   Density functional theory

To model a semiconductor crystal on the atomic scale, a wide variety of classical methods is known — ranging from simple ball and spring models to more sophisticated empirical interatomic potentials with two- or even three-body contributions [20, 21, 22, 23]. These methods are more or less successful in describing perfect and strained crystals of tetrahedrally bonded semiconductors. However, when it comes to defects with considerable deviations from the standard bonding configuration (full four-fold coordination of atoms and tetrahedral bond angles), then quantum mechanical effects, or effects involving charge transfer between atoms, might dominate. Very often this is the case in the very core region of point or line defects in semiconductors. Hence here the use of quantum mechanics is essential.

To describe a crystal in terms of non-relativistic and stationary quantum mechanics[1] with no further approximations would mean to solve the time independent Schrödinger equation for M atomic nuclei and N electrons. This is an impossible task if one considers the corresponding antisymmetric wavefunction, which depends on the $3M + 3N$ coordinates of all atoms and electrons. A first reduction of complexity is given by the *Born-Oppenheimer approximation* [24], where the motion of electrons and nuclei are decoupled — inspired by their large difference in masses. The remaining problem is now the determination of the electronic wavefunction for a given configuration of nuclei. Since the modelling of defects involves tens or even

---

[1] In this work relativistic effects are not taken into consideration explicitly, even spin-related effects are mostly ignored, as they play no crucial role in the discussed problems.

hundreds of atoms, the full many-body wavefunction is still by far too complex to be handled. The method of choice to further reduce the number of parameters is given with density functional theory (DFT), where the electron density is the central variable. This implies a reduction from the $3N$ coordinates of the electron wavefunction to only three coordinates (if we ignore spin-polarisation), as shown in the next section. The reader interested in a more detailed and more exact description may read the original work by Hohenberg and Kohn [25] and Kohn and Sham [26] or the rather mathematical treatment of Lieb [27], to give just a few examples.

1.1.1   Fundamental concepts and equations

In the investigation of defects in semiconductors, it is one central goal to identify the stable structures of the defect: Structures where the forces between atoms are in equilibrium. To achieve this, one needs to calculate the total energy of the model representing the defect. An equilibrium of forces is then given as a minimum on the energy surface.

The measurable quantities of a quantum mechanical ensemble are determined via a state function, which is a solution of the Schrödinger equation, e. g. the total energy $\mathcal{E}$ of a many-electron system is determined by its electronic wavefunction Ψ:

$$\mathcal{E}[\Psi] = \langle \Psi | \widehat{\mathbf{H}} | \Psi \rangle \tag{1.1}$$

Here $\widehat{\mathbf{H}}$ is the corresponding Hamiltonian and the right hand side represents a scalar product in Hilbert space. It is the basic idea of DFT, to introduce the electron density $n(\boldsymbol{r})$ as the central variable, and determine the measurable quantities in terms of this density:

$$n(\boldsymbol{r}) = \langle \Psi | \sum_i \delta(\boldsymbol{r} - \boldsymbol{r}_i') | \Psi \rangle \tag{1.2}$$

As Hohenberg and Kohn have shown, the total energy of the (non-degenerate) ground state of a quantum mechanical ensemble of electrons is a unique functional of its electron density [25]:

$$\mathcal{E} = \mathcal{E}[n(\boldsymbol{r})] \tag{1.3}$$

Therefore the mere knowledge of $n(\boldsymbol{r})$ is sufficient to determine the corresponding total energy.

Now we will turn towards the explicit form of $\mathcal{E}[n(\boldsymbol{r})]$. We separate $\mathcal{E}[n(\boldsymbol{r})]$ into different components as follows[2]:

$$\mathcal{E}[n] = T[n] + \frac{1}{2} \iint \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \, \mathrm{d}^3 r' \, \mathrm{d}^3 r + \int n(\boldsymbol{r}) v_{\mathrm{ext}}(\boldsymbol{r}) \, \mathrm{d}^3 r \tag{1.4}$$

Here the second term denotes the classical Coulomb-interaction of all electrons and the last term gives the energy arising from the external potential $v_{\mathrm{ext}}(\boldsymbol{r})$. The latter

---

[2] Quantities are given in atomic units, and the energy given in Hartree.

in addition depends on the atomic coordinates. Unfortunately the kinetic contribution $T[n]$ is generally unknown. However, in the specific case of N non-interacting electrons a simple solution can be given:

$$T_0[n] = \sum_{i=1}^{N} \int \phi_i^*(r) \left[-\frac{\nabla^2}{2} \right] \phi_i(r) \, d^3r \tag{1.5}$$

Here the density $n(r)$ has been expanded into normalised and orthogonal one-electron wavefunctions $\phi_i(r)$:

$$n(r) = \sum_{i=1}^{N} |\phi_i(r)|^2 \tag{1.6}$$

To make use of Eq. (1.5), Kohn and Sham suggested to expand $n(r)$ accordingly to Eq. (1.6) for the case of *interacting* electrons also [26]. This approximation means that for the system of interacting electrons the $\phi_i(r)$ are not true one-electron wavefunctions. In the following we shall call them *Kohn-Sham orbitals*. To still include the many-electron effects in $T[n]$, we have to add the so called *exchange-correlation functional $E_{xc}[n]$*[3]:

$$T[n] = T_0[n] + E_{xc}[n] \tag{1.7}$$

A variation of $\mathcal{E}[n]$ in Eq. (1.4) with $\delta\phi_i^*(r)$ and subject to the normalisation constraint delivers the Kohn-Sham orbitals $\phi_i(r)$, which yield a minimal total energy. We obtain the *Kohn-Sham equations* [26]:

$$\underbrace{\left[-\frac{\nabla^2}{2} + v_{\text{eff}}[n(r)] \right]}_{= \widehat{H}[n]} \phi_i(r) = \varepsilon_i \, \phi_i(r) \tag{1.8}$$

$$\text{with} \quad v_{\text{eff}}[n(r)] := v_{\text{ext}}(r) + \int \frac{n(r')}{|r - r'|} \, d^3r' + \frac{\delta E_{xc}[n(r)]}{\delta n(r)} \tag{1.9}$$

Here the ε_i are merely the Lagrange parameters of the normalisation constraint, and *not* one-electron energies[4]. Correspondingly the Kohn-Sham equations (Eq. (1.8)) are not independent one-electron equations — they are coupled via the effective potential $v_{\text{eff}}[n(r)]$, which depends on all (occupied) $\phi_i(r)$, and hence Eq. (1.8) can only be solved self-consistently.

The last term of the effective potential (Eq. (1.9)) is the *exchange-correlation potential $v_{xc}[n(r)]$*. Generally this potential is unknown. For special cases, however, like the homogeneous electron gas at high (or low) density, the explicit form has been evaluated. Assuming now that the exchange-correlation contribution to the total energy

---

[3] This formalism is not used uniformly in the literature: $E_{xc}[n]$ is sometimes excluded from $T[n]$ and included in Eq. (1.4) directly.

[4] Only in special cases, when the removal of electron j does not influence the remaining ε_i, then one may interpret ε_j as the one-electron energy of electron j (Koopmans' theorem [28]).

is varying only a little, one can use approximations for v_{xc} based on a homogeneous electron gas. The approximation used in this work is the well-approved local density approximation (LDA) of density functional theory. In this approximation, if expressing the exchange-correlation contribution via an integral, the integrand ϵ_{xc} depends only locally on the density $n(\boldsymbol{r})$:

$$E_{xc}[n] = \int n(\boldsymbol{r})\epsilon_{xc}[n(\boldsymbol{r})]\, \mathrm{d}^3 r \approx \int n(\boldsymbol{r})\epsilon_{xc}^{\mathrm{LDA}}(n(\boldsymbol{r}))\, \mathrm{d}^3 r \tag{1.10}$$

The main drawback of LDA is the underestimation of the forbidden electronic band gap of semiconductor crystals. Along with this, density functional theory describes the ground state in a given Hilbert space only. Therefore excited states can usually not be described properly anyway.

For a given spatial arrangement of nuclei one can now determine the ground state electron density $n(\boldsymbol{r})$ within the Born-Oppenheimer approximation by solving Eq. (1.8) self-consistently. Therefore the external potential $v_{\mathrm{ext}}(\boldsymbol{r})$ in Eq. (1.9) is given as the electrostatic potential of the nuclei. The total energy of all electrons in the potential of all nuclei can then be evaluated via Eq. (1.4). Adding the Coulomb repulsion E_{nuc} of the nuclei yields the total energy of the system consisting of electrons and nuclei:

$$E_{\mathrm{tot}}[n] = \mathcal{E}[n] + E_{\mathrm{nuc}}\left(\{\boldsymbol{R}_j\}\right) \tag{1.11}$$

$$\text{with} \quad E_{\mathrm{nuc}}\left(\{\boldsymbol{R}_j\}\right) = \frac{1}{2}\sum_{i,j\neq i}^{M}\frac{Z_i Z_j}{|\boldsymbol{R}_i - \boldsymbol{R}_j|} \tag{1.12}$$

Z_i is the atomic number of atom i and $\boldsymbol{R}_i$ its coordinates. With this energy at hand, one can search for the equilibrium of forces — the local minima of the energy surface. This structural optimisation is usually done with the help of algorithms involving atomic forces. The latter can be obtained via the Hellmann-Feynman theorem [29, 30]. For atom-centred basis sets the so called Pulay corrections have to be applied [31].

The next sections will give two examples for implementations of DFT. Both involve additional approximations to reduce the computational effort.

1.1.2   The pseudopotential approach

In this section, the pseudopotential approach to DFT — as implemented in the AIM-PRO computer code — will be introduced briefly. This introduction avoids discussing any of the arising problems like singularities in potentials, relativistic effects for heavier atoms and the transferability of pseudopotentials. Further details can be found in more specific literature on the AIMPRO code [17, 32] and the approach in general [33, 34, 35].

In practice, the bonding between atoms in solids is predominantly governed by their valence electrons. The core electrons of the closed inner shells basically behave like in an isolated atom and create a screening effective potential. Thus it seems a

promising idea to only consider the valence electrons when self-consistently calculating the electron density. The remaining core electrons as well as the electrostatic potential of the nuclei are included in a new effective potential — the so called *pseudopotential* $v^{\mathrm{ps}}(\boldsymbol{r})$, which represents the potential created by all ions. Then the overall effective potential in Eq. (1.8) becomes:

$$v_{\mathrm{eff}}[n^{\mathrm{v}}(\boldsymbol{r})] = v^{\mathrm{ps}}(\boldsymbol{r}) + \int \frac{n^{\mathrm{v}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\, \mathrm{d}^3 r' + v_{\mathrm{xc}}[n^{\mathrm{v}}(\boldsymbol{r})] \tag{1.13}$$

Here $n^{\mathrm{v}}(\boldsymbol{r})$ is the valence electron density. The pseudopotential $v^{\mathrm{ps}}(\boldsymbol{r})$ is given as the superposition of all single ionic pseudopotentials:

$$v^{\mathrm{ps}}(\boldsymbol{r}) = \sum_{j=1}^{M} v_j^{\mathrm{ps}}(\boldsymbol{r}-\boldsymbol{R}_j) \tag{1.14}$$

Here the $\boldsymbol{R}_j$ are the positions of the respective nuclei. To obtain the ionic pseudopotential $v_j^{\mathrm{ps}}(\boldsymbol{r})$ one starts with the effective all-electron potential $v_j^{\mathrm{eff}}(\boldsymbol{r})$ of a neutral atom. Stripping off the valence electron contributions yields:

$$v_j^{\mathrm{ps}}(\boldsymbol{r}) = v_j^{\mathrm{eff}}(\boldsymbol{r}) - \int \frac{n_j^{\mathrm{v}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\, \mathrm{d}^3 r' - v_{\mathrm{xc}}[n_j^{\mathrm{v}}(\boldsymbol{r})] \tag{1.15}$$

There are several recipes how to construct pseudopotentials in detail. In this work the norm-conserving pseudopotentials of Bachelet, Hamann and Schlüter [35] are used. Assuming a *frozen core*, the same pseudopotential can be used for every spatial distribution of atoms and different charge states.

The advantages of the pseudopotential approach are obvious: Comparing all-electron energies of similar systems might lead to relatively large errors, since two large numbers are subtracted. The problem is resolved if only the valence electrons contribute to the difference in energy. Another crucial advantage is the reduction of computational complexity. The number of electrons, and with this the number of orthogonal wavefunctions, to be included explicitly in the self-consistent calculation of the total energy is reduced dramatically if only valence electrons are considered. Hence the modelling of larger systems becomes possible[5].

The choice of basis set: For a practical implementation of this approach, one has to decide for a basis in terms of which the Kohn-Sham orbitals are expanded. In AIMPRO the choice is a basis set of *Cartesian Gaussian orbitals* centred at $\boldsymbol{R}_\nu$:

$$\psi_\nu(\boldsymbol{r}) = (x_1 - R_{\nu 1})^{n_{1\nu}} (x_2 - R_{\nu 2})^{n_{2\nu}} (x_3 - R_{\nu 3})^{n_{3\nu}} e^{-a_\nu (\boldsymbol{r}-\boldsymbol{R}_\nu)^2} \tag{1.16}$$

With specially selected and real $n_{i\nu}$ it is possible to describe atom-centred *s*-, *p*- and *d*-like functions. Expanding the Kohn-Sham orbitals $\phi_i(\boldsymbol{r})$ into Cartesian Gaussian orbitals yields:

$$\phi_i(\boldsymbol{r}) = \sum_\nu C_{\nu i} \psi_\nu(\boldsymbol{r}) \tag{1.17}$$

---

[5] At the time of writing this work, AIMPRO on parallel supercomputers was capable of conveniently treating models consisting of a few hundred atoms.

Inserting the expansions of the Kohn-Sham orbitals in Eq. (1.8), followed by left-side multiplication with $\psi_\nu(\boldsymbol{r})$ and integration over all space, reduces Eq. (1.8) to a set of algebraic *secular equations*:

$$\sum_\nu C_{\nu i}\left(H_{\mu\nu} - \varepsilon_i\, S_{\mu\nu}\right) = 0 \tag{1.18}$$

Here the $C_{\nu i}$ are the expansion coefficients and $H_{\mu\nu}$ and $S_{\mu\nu}$ the Hamiltonian and overlap matrix respectively:

$$H_{\mu\nu} = \int \psi_\mu^*(\boldsymbol{r})\left[-\frac{\nabla^2}{2} + v_{\text{eff}}[n^{\text{v}}(\boldsymbol{r})]\right]\psi_\nu(\boldsymbol{r})\, \text{d}^3 r \tag{1.19}$$

$$S_{\mu\nu} = \int \psi_\mu^*(\boldsymbol{r})\psi_\nu(\boldsymbol{r})\, \text{d}^3 r \tag{1.20}$$

The strong point of Gaussian basis functions is their localisation. Therefore considerably fewer basis functions are necessary to approximatively describe a localised wavefunction. Dislocations (line defects) usually yield wavefunctions localised in at least two dimensions. Since most models used in this work include a great amount of vacuum, the use of a localised basis is almost inevitable.

1.1.3 The density-functional based tight-binding approach

We will now introduce the Density-functional based tight-binding method (DFTB) — the second approach used in this work to approximately solve the Kohn-Sham equations (1.8). Just like in the last section, the description will be introductory and lack many details. For further insight the original work of Seifert et al. [36] and Porezag et al. [18, 37] or the review in [19] are recommended. A compact but nevertheless detailed overview including a sketch of the historical development can be found in [38].

To understand the basic idea of the DFTB method, one best starts with the explicit form of the total energy in DFT after introducing the Kohn-Sham orbitals. This is obtained easily by inserting Eq. (1.7) and (1.5) in Eq. (1.4):

$$\mathcal{E}[n] = \sum_i^{\text{occ}} n_i \int \phi_i^*(\boldsymbol{r})\left[-\frac{\nabla^2}{2} + \frac{1}{2}\int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\,\text{d}^3 r' + v_{\text{ext}}(\boldsymbol{r})\right]\phi_i(\boldsymbol{r})\,\text{d}^3 r + E_{\text{xc}}[n] \tag{1.21}$$

Here we additionally introduced normalised occupation numbers n_i following Janak [39] :

$$n(\boldsymbol{r}) = \sum_i^{\text{occ}} n_i|\phi_i(\boldsymbol{r})|^2\,; \quad N = \sum_i^{\text{occ}} n_i \tag{1.22}$$

These occupation numbers n_i contain the information how many electrons occupy the respective Kohn-Sham orbital. In a spin-restricted formalism as used here, the maximum occupation per orbital is 2.

The idea is now to simplify Eq. (1.21) by applying a Taylor-expansion around a reference density $n_0(\mathbf{r})$ and neglecting higher order contributions. Introducing $\Delta n(\mathbf{r}) = n(\mathbf{r}) - n_0(\mathbf{r})$ we obtain (Foulkes and Haydock [40]):

$$
\begin{aligned}
\mathcal{E}[n] \;=\; & \sum_i^{\text{occ}} n_i \int \phi_i^*(\mathbf{r}) \, \widehat{H}[n_0] \, \phi_i(\mathbf{r}) \, \mathrm{d}^3r \;-\; \frac{1}{2} \iint \frac{n_0(\mathbf{r}) n_0(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}^3r' \, \mathrm{d}^3r \\
& + E_{\text{xc}}[n_0] \;-\; \int v_{\text{xc}}[n_0(\mathbf{r})] n_0(\mathbf{r}) \, \mathrm{d}^3r \\
& + \frac{1}{2} \iint \left[\frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2 E_{\text{xc}}[n]}{\delta n^2} \right]_{n_0} \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') \, \mathrm{d}^3r' \, \mathrm{d}^3r \;+\; \mathcal{O}(\Delta n^3) \quad (1.23)
\end{aligned}
$$

$\widehat{H}[n]$ was defined in Eq. (1.8). The first and second lines give the zero-th order terms in Δn and the third line gives the second and third order. All first order contributions cancel.

Choosing a localised (atom-centred) set of basis functions, the second to fifth terms in Eq. (1.23) can be expressed as a sum of two-centre integrals[6]. Subsuming those terms as $E_{2\,\text{cent}}$ and neglecting three-centre and higher-order contributions yields the total energy in DFTB:

$$
\mathcal{E}[n] \;\approx\; \sum_i^{\text{occ}} n_i \int \phi_i^*(\mathbf{r}) \, \widehat{H}[n_0] \, \phi_i(\mathbf{r}) \, \mathrm{d}^3r \;+\; E_{2\,\text{cent}} \;=\; \mathcal{E}_{\text{DFTB}} \qquad (1.24)
$$

DFT calculations show, that the density of an atom surrounded by other atoms in a molecule or a solid is slightly compressed compared to the density of a free atom. Hence in DFTB the starting density $n_0(\mathbf{r})$ is expressed as a superposition of weakly confined neutral atoms, so called *pseudo-atoms*[7]. In all DFTB implementations used in this work, the respective pseudo-atomic wavefunctions are represented by Slater-type orbitals $\varphi_\nu(\mathbf{r})$.

Expanding the Kohn-Sham orbitals of Eq. (1.24) into the atom-centred Slater-type orbitals $\varphi_\nu(\mathbf{r})$ yields (compare Eq. (1.17)):

$$
\phi_i(\mathbf{r}) = \sum_\nu C_{\nu i} \, \varphi_\nu(\mathbf{r}) \qquad (1.25)
$$

This resembles an LCAO basis (linear combination of atomic orbitals). Further we approximate the two-centre contributions by the sum of short-ranged repulsive pair potentials v_{rep}^{ij}:

$$
E_{2\,\text{cent}} \;\approx\; \frac{1}{2} \sum_{i \neq j} v_{\text{rep}}^{ij} \left(|\mathbf{R}_i - \mathbf{R}_j| \right) \;=\; E_{\text{rep}} \left(\{\mathbf{R}_i\} \right) \qquad (1.26)
$$

---

[6] For the exchange-correlation terms this only holds approximately.

[7] These pseudo-atom densities result from self-consistent DFT calculations with an additional weak parabolic constriction potential.

Just as a note for those familiar with the internal details of the DFTB method: In this work the diamond C–C interaction includes the superposition of the *potentials* of the pseudo-atoms. These potentials are compressed by an additional term $(r/r_0)^2$ with a compression radius of $r_0 = 1.42$ Å. In the case of silicon carbide, however, it is the superposition of the *electron densities* of the pseudo-atoms which is used. Here the C compression radii are $r_0^{\text{wav}} = 2.7$ Å and $r_0^{\text{den}} = 7.0$ Å for the wavefunction and the density respectively. For Si the chosen radii are $r_0^{\text{wav}} = 3.3$ Å and $r_0^{\text{den}} = 6.7$ Å.

Inserting Eq. (1.25) and (1.26) in (1.24) gives the DFTB energy in terms of the expansion coefficients $C_{\nu i}$. Subsequent variation with respect to the expansion coefficients and subject to normalisation yields the *DFTB secular equations* (compare Eq. (1.18)):

$$\sum_{\nu} C_{\nu i} \left(H^0_{\mu\nu} - \varepsilon_i \, S_{\mu\nu} \right) = 0 \tag{1.27}$$

If $\boldsymbol{R}_{\nu}$ denotes the atom site the orbital is centred at, then the Hamiltonian and the overlap matrix are given as:

$$H_{\mu\nu} = \begin{cases} \varepsilon^{\mathrm{atom}}_{\mu} & : \quad \mu = \nu \\ \int \varphi^*_{\mu}(\boldsymbol{r}) \, \widehat{H}[n_0] \, \varphi_{\nu}(\boldsymbol{r}) \, \mathrm{d}^3 r & : \quad \boldsymbol{R}_{\nu} \neq \boldsymbol{R}_{\mu} \\ 0 & : \quad \mathrm{else} \end{cases} \tag{1.28}$$

$$S_{\mu\nu} = \int \varphi^*_{\mu}(\boldsymbol{r}) \varphi_{\nu}(\boldsymbol{r}) \, \mathrm{d}^3 r \tag{1.29}$$

Notice, that the diagonal elements of $H_{\mu\nu}$ are taken to be atomic orbital energies of a free atom, to ensure the right energies for dissociated atoms. Once the pseudo-atom orbitals have been calculated self-consistently in DFT, the non-diagonal elements of the Hamiltonian $H_{\mu\nu}$ and the overlap matrix $S_{\mu\nu}$ can be tabulated as a function of distance between the two centres. Thus both matrices can be calculated in advance. Once these tables have been generated, Eq. (1.27) can be solved straight away and non self-consistently for *any* given coordinates of the nuclei. The DFTB energy then becomes:

$$\mathcal{E}_{\mathrm{DFTB}} = \sum_i^{\mathrm{occ}} n_i \, \varepsilon_i + E_{\mathrm{rep}} \left(\{ \boldsymbol{R}_i \} \right) \tag{1.30}$$

Apart from the repulsive energy, the total energy is now determined. Assuming approximate transferability of the repulsive potentials $v^{ij}_{\mathrm{rep}}(|\boldsymbol{R}_i - \boldsymbol{R}_j|)$, they are obtained in comparison with self-consistent DFT calculations for selected reference systems (small molecules or infinite crystals): Varying one distance $|\boldsymbol{R}_i - \boldsymbol{R}_j|$ the respective repulsive potential is tabulated just like the matrix elements $H_{\mu\nu}$ and $S_{\mu\nu}$. In practice, the repulsive pair potentials also contain the respective ion-ion repulsion as given in Eq. (1.12). Further, the compensating effect of the repulsive potential allows one to use a minimal basis of Slater-type orbitals, reducing matrix size and thus speeding up all calculations.

As has been just shown, all DFT integrals can be calculated in advance. In the calculation of the eigenvalues ε_i and the total energy this saves a great deal of computational effort and speeds up the calculation dramatically. Such a speed up allows a structural relaxation of systems containing more atoms than all-electron DFT or even pseudopotential methods[8]. This advantage will prove crucial when it comes to comparing DFT-based atomistic calculations with the results of elasticity theory in Chapter 4 — a task, where models containing a relatively large number of atoms are needed.

---

[8] At the time of writing this work, DFTB on workstations was capable of conveniently relaxing models consisting of more than 700 atoms — more than in the DFT-pseudopotential approach on parallel supercomputers.

SCC-DFTB: For some systems with strong ionic bonding character the approximations of standard DFTB might fail. Especially the second order term in Eq. (1.23) becomes too large to be transferable and simply tabulated within the repulsive potential. The term has then to be treated self-consistently. This has been successfully achieved in the self-consistent-charge extension of DFTB (SCC-DFTB), where the charge transfer through $\Delta n(\boldsymbol{r})$ is approximated by spherical atom-centred charge fluctuations. For further details see [19].

In this work however, the self-consistent-charge extension is not used. Test calculations have shown, that even in silicon carbide the non-SCC treatment is sufficient to describe at least the atomic structure and energy of the dislocations investigated. As all subsequent electronic structure calculations are performed within AIMPRO, the DFTB calculations can be restricted to the non-SCC formalism. Given the size of the involved models (~ 500 atoms), this reduces the computational effort considerably.

1.2 Linear elasticity theory

While the preceding sections presented methods which allow an atomistic description of a semiconductor crystal as a system consisting of atoms and electrons, we will now turn towards continuum theory: Elasticity theory basically describes the relation between strain and stress fields, displacements and elastic strain energy in a continuous and elastic medium. This section starts with an *brief* introduction of the fundamental equations and definitions of elasticity theory. After explaining the contracted matrix notation and discussing the isotropic case, the Voigt average for anisotropic crystals is presented, giving merely the results and skipping any details. We basically follow the notation of Hirth and Lothe [41], who give a good review of those aspects of elasticity theory which are essential in dislocation theory. More complete texts on the whole subject are those of Love [42], Sokolnikoff [43], Nye [44] and Landau and Lifshitz [45].

1.2.1 Fundamental concepts and equations

Elasticity theory is a continuum theory: It describes an elastic solid by means of a continuous medium. Let us now assume a homogeneous medium, in other words, the mass density $\varrho(\mathbf{r})$ is constant in the case of zero strain. Let $\mathbf{r} = (x_1, x_2, x_3)$ be a vector in orthogonal Cartesian coordinates and $\sigma_{ij}(\mathbf{r})$ be the stresses in the medium. As shown in Fig. 1.1 for an infinitesimal volume element, $\sigma_{ij}(\mathbf{r})$ is the force in x_i direction per unit area on a plane normal to the x_j direction. In the following the dependence of variables on $\mathbf{r}$ will not be given explicitly. In mechanical equilibrium there is no net torque on the element and if we also rule out internal torques, then σ_{ij} is symmetric:
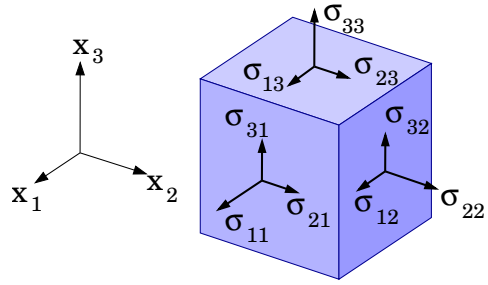
Figure 1.1: Stress distribution on an infinitesimal volume element.

$$\sigma_{ij} = \sigma_{ji} \tag{1.31}$$

Further, there is also no net force on the element. We obtain the *equilibrium equations* of classical elasticity written with the Einstein convention:

$$\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0 \qquad \forall i \tag{1.32}$$

Here $\mathbf{f}$ is the body force per unit volume. Such forces can arise, for example, from a charge distribution in the presence of an electric field.

A body which is acted upon by stress, deforms. Let now $\mathbf{u}$ be the displacement field after deformation, then we can define the *strains* ε_{ij} and the *stiff rotations* ω_{ij}:

$$\varepsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \tag{1.33}$$

$$\omega_{ij} = \frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \tag{1.34}$$

Both Eq. (1.33) and (1.34) are restricted to first order terms in the derivatives, assuming only small displacements and stresses. The factor $1/2$ in Eq. (1.33) is conventionally used for transformation purposes. This differs from the definition of the shear strain γ in engineering. Shear strain is given, when $i \neq j$:

$$\gamma_{ij} = 2\,\varepsilon_{ij} \qquad i \neq j \tag{1.35}$$

In linear theory (small distortions $\partial u_i / \partial x_j$) *Hooke's law* applies — we can assume a linear dependence of the stress on the deformation. Since the stiff rotations ω_{ij} cannot give rise to stresses, we obtain:

$$\sigma_{ij} = c_{ijkl}\,\varepsilon_{kl} \tag{1.36}$$

The c_{ijkl} are called *elastic constants*. Considering now the reversible deformation of a volume element under differential strain, one can determine the *strain energy density* by integration:

$$e_s = \frac{1}{2}\,c_{ijkl}\,\varepsilon_{ij}\,\varepsilon_{kl} \tag{1.37}$$

In principle, with Eq. (1.36) and (1.37), we are now able to calculate the elastic strain energy for a given strain. However, it should be noted, that here we did *not* discriminate between a fixed external and an embedded coordinate system, which deforms with the medium: Only the terms common to both descriptions are retained. For small displacements and stresses this is sufficient, for strongly distorted parts of a crystal however, the results of linear theory have to be questioned.

1.2.2 The contracted matrix notation

Let us turn back to Eq. (1.36) and the elastic constants. The symmetry of ε_{ij} imposes several symmetries on the tensor of the c_{ijkl} indicating there are only 21 independent elastic constants among the 81 elements of the tensor. Hence in matrix notation we can write the tensor as a symmetric (9×9) matrix representation $\{c_{ijkl}\}$ with 21 independent entries. We obtain for Eq. (1.36):

$$
\begin{bmatrix}
\sigma_{11} \\
\sigma_{22} \\
\sigma_{33} \\
\sigma_{23} \\
\sigma_{31} \\
\sigma_{12} \\
\sigma_{32} \\
\sigma_{13} \\
\sigma_{21}
\end{bmatrix}
=
\begin{bmatrix}
c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} & c_{14} & c_{15} & c_{16} \\
 & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} & c_{24} & c_{25} & c_{26} \\
 & & c_{33} & c_{34} & c_{35} & c_{36} & c_{34} & c_{35} & c_{36} \\
 & & & c_{44} & c_{45} & c_{46} & c_{44} & c_{45} & c_{46} \\
 & & & & c_{55} & c_{56} & c_{45} & c_{55} & c_{56} \\
 & & & & & c_{66} & c_{46} & c_{56} & c_{66} \\
 & & & & & & c_{44} & c_{45} & c_{46} \\
 & & & & & & & c_{55} & c_{56} \\
 & & & & & & & & c_{66}
\end{bmatrix}
\cdot
\begin{bmatrix}
\varepsilon_{11} \\
\varepsilon_{22} \\
\varepsilon_{33} \\
\varepsilon_{23} \\
\varepsilon_{31} \\
\varepsilon_{12} \\
\varepsilon_{32} \\
\varepsilon_{13} \\
\varepsilon_{21}
\end{bmatrix}
\tag{1.38}
$$

In Eq. (1.38) we abbreviated the first two and the last two indices of the c_{ijkl} into single indices running from 1 to 9. Thus, the indices m and n of the c_{mn} correspond to the index pairs ij and kl of the c_{ijkl} as follows:

$$
\begin{array}{ccccccccccc}
ij & \text{or} & kl & : & 11 & 22 & 33 & 23 & 31 & 12 & 32 & 13 & 21 \\
m & \text{or} & n & : & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{array}
\tag{1.39}
$$

Since the indices $7, 8, 9$ correspond to $4, 5, 6$ respectively, we can conveniently write Eq. (1.38) in a contracted (6×6) matrix notation as:

$$
\begin{bmatrix}
\sigma_{11} \\
\sigma_{22} \\
\sigma_{33} \\
\sigma_{23} \\
\sigma_{31} \\
\sigma_{12}
\end{bmatrix}
=
\begin{bmatrix}
c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\
 & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\
 & & c_{33} & c_{34} & c_{35} & c_{36} \\
 & & & c_{44} & c_{45} & c_{46} \\
 & & & & c_{55} & c_{56} \\
 & & & & & c_{66}
\end{bmatrix}
\cdot
\begin{bmatrix}
\varepsilon_{11} \\
\varepsilon_{22} \\
\varepsilon_{33} \\
\gamma_{23} \\
\gamma_{31} \\
\gamma_{12}
\end{bmatrix}
\tag{1.40}
$$

Notice that here we have to use the shear strains γ_{ij} as given in Eq. (1.35) instead of the ε_{ij}. This allows a symmetric matrix.

The results so far all presumed an anisotropic medium with no symmetry. In crystals however, symmetry can usually help to reduce the number of independent elastic constants dramatically. For cubic symmetry the c_{ijkl} have to be invariant to $90°$ rotations. A comparison of tensor entries yields:

$$
\{c_{nm}\}^{\text{cubic}} =
\begin{bmatrix}
c_{11} & c_{12} & c_{12} & 0 & 0 & 0 \\
 & c_{11} & c_{12} & 0 & 0 & 0 \\
 & & c_{11} & 0 & 0 & 0 \\
 & & & c_{44} & 0 & 0 \\
 & & & & c_{44} & 0 \\
 & & & & & c_{44}
\end{bmatrix}
\tag{1.41}
$$

Similarly we obtain for hexagonal symmetry with the third axis perpendicular to the basal plane [44]:

$$
\{c_{nm}\}^{\text{hexagonal}} =
\begin{bmatrix}
c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\
 & c_{11} & c_{13} & 0 & 0 & 0 \\
 & & c_{33} & 0 & 0 & 0 \\
 & & & c_{44} & 0 & 0 \\
 & & & & c_{44} & 0 \\
 & & & & & \frac{1}{2}(c_{11} - c_{12})
\end{bmatrix}
\tag{1.42}
$$

1.2.3   The link to isotropic elasticity theory

After the two examples of cubic and hexagonal symmetry, we will now discuss the case of an isotropic medium. In such a medium all possible directions are equivalent, so that the $\{c_{ijkl}\}$ are invariant to arbitrary rotations. A short calculation yields as a reduction from cubic symmetry in Eq. (1.41) to the isotropic case:

$$
c_{44} = \frac{1}{2}\left(c_{11} - c_{12}\right)
\tag{1.43}
$$

With this only two of the three constants are independent — the *shear modulus* μ and the *Lamé constant* λ:

$$\mu = c_{44} = \frac{1}{2}\left(c_{11} - c_{12}\right) \quad ; \quad \lambda = c_{12} \tag{1.44}$$

With these, Eq. (1.40) writes simply as:

$$
\begin{aligned}
\sigma_{11} &= (\lambda + 2\,\mu)\,\varepsilon_{11} + \lambda\,\varepsilon_{22} + \lambda\,\varepsilon_{33} & \sigma_{23} &= \mu\,\gamma_{23} \\
\sigma_{22} &= \lambda\,\varepsilon_{11} + (\lambda + 2\,\mu)\,\varepsilon_{22} + \lambda\,\varepsilon_{33} & \sigma_{31} &= \mu\,\gamma_{31} \\
\sigma_{33} &= \lambda\,\varepsilon_{11} + \lambda\,\varepsilon_{22} + (\lambda + 2\,\mu)\,\varepsilon_{33} & \sigma_{12} &= \mu\,\gamma_{12}
\end{aligned}
\tag{1.45}
$$

Further we can introduce two alternative elastic constants. The ratio of simple tensile stress to strain, the *Young's modulus E*, and the ratio of transverse contraction to elongation in simple tension, the *Poisson's ratio ν*, are commonly used elastic constants:

$$E = \frac{\mu(3\,\lambda + 2\,\mu)}{\mu + \lambda} \quad ; \quad \nu = \frac{\lambda}{2(\mu + \lambda)} \tag{1.46}$$

Finally, we can also express the *bulk modulus* in terms of the elastic constants. The bulk modulus gives the ratio between the applied pressure $p = -\frac{1}{3}\left(\sigma_{11} + \sigma_{22} + \sigma_{33}\right)$ and the resulting compression $-e = -\varepsilon_{11} - \varepsilon_{22} - \varepsilon_{33}$ of a crystal:

$$B = -\frac{p}{e} = \frac{1}{3}\left(3\,\lambda + 2\,\mu\right) = \frac{1}{3}\left(c_{11} + 2c_{12}\right) \tag{1.47}$$

1.2.4 The Voigt average of elastic constants

As shown in the last section, assuming an isotropic medium simplifies elasticity theory dramatically. Consequently, later in Chapter 2 this assumption allows the analytical solution of many dislocation related problems. Crystals usually do not show exactly isotropic behaviour, but fortunately, in many cases, they are approximately isotropic. Thus, we need a way of sensibly averaging the anisotropic elastic constants to approximate isotropic constants. There are two basic approaches to this: One can either average over the elastic constants themselves, or over their inverse, the elastic compliances. The former, the so called *Voigt average* [46], models a polycrystal in which every grain is in the same state of strain, whereas the latter, the so called *Reuss average* [47], models a situation where every grain has the same stress. Experience has shown, that in most cases, especially when describing local strain around a dislocation, the Voigt average is most appropriate [41][9]. The Voigt average can be determined by the use of invariants to arbitrary rotations: The trace of $\{c_{ijkl}\}$ and the strain energy density of homogeneous expansion. One obtains as the average isotropic elastic constants:

$$\mu = \frac{1}{30}\left(3\,c_{ijij} - c_{iijj}\right)$$

$$\lambda = \frac{1}{15}\left(2\,c_{iijj} - c_{ijij}\right) \tag{1.48}$$

---

[9] The Reuss average has proven valuable in the description of long-range internal stress fields.

Eq. (1.48) yields respectively for a crystal with cubic (Eq. (1.41)) and for a crystal with hexagonal (Eq. (1.42)) symmetry:

$$\mu^{\text{cubic}} = \frac{1}{5}\left(3\,c_{44} - c_{12} + c_{11}\right)$$

$$\lambda^{\text{cubic}} = \frac{1}{5}\left(-2\,c_{44} + 4\,c_{12} + c_{11}\right)$$

(1.49)

$$\mu^{\text{hexagonal}} = \frac{1}{30}\left(7\,c_{11} - 5\,c_{12} + 2\,c_{33} + 12\,c_{44} - 4\,c_{13}\right)$$

$$\lambda^{\text{hexagonal}} = \frac{1}{15}\left(c_{11} + c_{33} + 5\,c_{12} + 8\,c_{13} - 4\,c_{44}\right)$$

(1.50)

Chapter 2

An Introduction to Dislocation Theory

It is the main purpose of this chapter to introduce the concepts and definitions concerning dislocations. The reader not familiar with dislocation theory will be provided with the basic ideas and the necessary vocabulary. This includes selected results from linear elasticity theory whenever this supports the general understanding or will be useful for later chapters.

The introduction given here can of course *not* replace a thorough discussion as given for example by Hirth and Lothe [41].

2.1 The Burgers vector

If stress is applied to a crystal, it deforms. Let $u(r)$ be the displacement field describing the displacement of a small volume element from its original position after deformation. This implies a continuum description of the crystal as in Section 1.2. In an ideal crystal the closed integral circuit over $u(r)$ will always yield zero — even after arbitrary elastic deformation. Real crystals, however, are far from being ideal: Structural line defects with a line direction $\ell(r)$ may exist, where the integral circuit around the defect results in a non-zero vector:
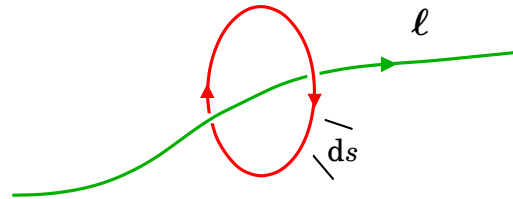
Figure 2.1: The Burgers integration circuit

$$b = \oint \frac{\partial u}{\partial s} \, \mathrm{d}s \qquad (2.1)$$

This defines the so-called *Burgers vector* b of the line defect. Line defects with a Burgers vector $b \neq 0$ are called *dislocations*. They form either during the growth process of the crystal, or through subsequent plastic deformation.

17
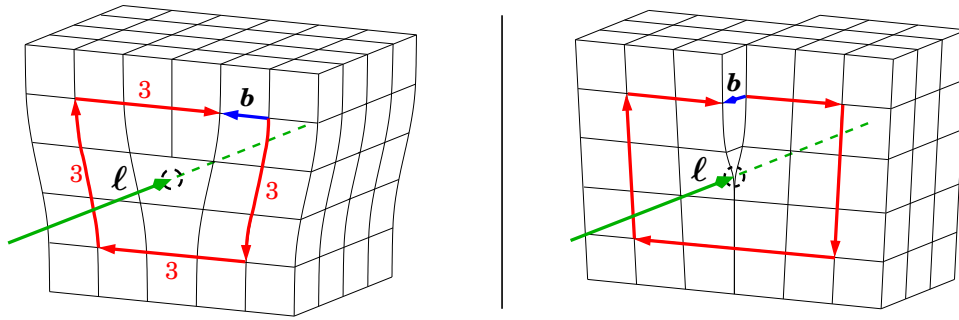
Figure 2.2: Edge and screw dislocation in a simple cubic lattice, only a small region of the lattice is drawn. *Left:* Edge dislocation. *Right:* Screw dislocation. In addition to the line direction ℓ and the Burgers vector b, a possible Burgers circuit for each dislocation is given. The closure fault of this circuit defines the Burgers vector b.

If we now abandon the continuous description of the crystal and take it as a periodic arrangement of discrete lattice sites[1], then it becomes obvious that b cannot be arbitrary but has to be somehow restricted by the structure of the lattice: If the Burgers integration circuit only encircles the dislocation and no other defects, and if the integration path runs through perfect (but strained) lattice only, then the Burgers vector b has to be a linear combination of the lattice translational vectors. Dislocations of this type are called *perfect dislocations* as opposed to *partial dislocations*, which are accompanied by a secondary structural lattice defect, e.g. a stacking fault. Partial dislocations will be discussed later in Chapter 4.

2.2 Edge and screw dislocations

All considerations in this section will be clarified and illustrated with the example of a simple cubic lattice. However, the conclusions drawn are not restricted to this lattice type. Assuming a simple cubic lattice, the local line direction $\ell(r)$ and a minimum Burgers vector b can be either parallel or perpendicular, resulting in two elementary dislocation types with minimal Burgers vector: The *edge dislocation* and the *screw dislocation*. Both are schematically depicted in Fig. 2.2 and can be characterised as follows:

The edge dislocation: Topologically speaking we obtain the edge dislocation by inserting a half-plane of lattice sites into the crystal. The dislocation line is the line where the half-plane terminates. As a result the Burgers vector is *perpendicular* to the line direction:

$$b_{\text{edge}} \cdot \ell(r) = 0 \qquad\qquad (2.2)$$

---

[1] With either one atom or a "basis" of several atoms at each lattice site. The lattice itself is defined by its translational symmetry given by translational vectors: A translation along one of these vectors maps the whole lattice onto itself. The crystal is then described as the periodic arrangement of the afore mentioned "basis" of atoms following the translational lattice vectors. A simple introduction into these basics of solid state physics can be found in the book of Kittel [48].

The screw dislocation: This dislocation type can be constructed by shearing one part of the crystal with respect to the other part at a half-plane only. The dislocation line is given as the line where the half-plane terminates. The Burgers vector is then *parallel* to the line direction:

$$\boldsymbol{b}_{\text{screw}} \times \boldsymbol{\ell}(\boldsymbol{r}) = 0 \tag{2.3}$$

Superpositions of both types resulting in mixed type dislocations are possible. Fig. 2.2 also shows the Burgers circuit in the discrete lattice. Note that the circuit — now of course in discrete steps from lattice site to lattice site — must be a circuit which would be closed in an ideal undislocated lattice, e.g. a parallelogram. The closure fault defines the Burgers vector $\boldsymbol{b}$. Usually $\boldsymbol{b}$ is given in units of the lattice, in internal and strained coordinates, thus $\boldsymbol{b}$ does not depend on the exact Burgers circuit path through more or less strained lattice: Both dislocations in Fig. 2.2 have a Burgers vector with the length of *one* minimal lattice translation, even though the length in external coordinates may vary due to varying strain.

2.3 Straight dislocations in linear elasticity theory

Linear elasticity theory as described in Section 1.2 proves to be a very useful tool to describe the long-range elastic strain effects of dislocations. In the following we will discuss the elastic strain energy of an isolated straight dislocation and the interaction forces between two straight dislocations. Curved dislocations are excluded since they play no important role in this work.

2.3.1 The elastic strain energy of a straight dislocation

The strain energy of an infinite straight dislocation in an otherwise perfect crystal can be calculated analytically using linear elasticity theory. As shown in Appendix A one writes the energy per unit length of a dislocation, contained in a cylinder of radius R and length L, as

$$\frac{E(R)}{L} = \frac{k(\beta)|\boldsymbol{b}|^2}{4\pi} \ln\left(\frac{R}{R_{\text{c}}}\right) + \frac{E_{\text{c}}}{L}, \quad R \geq R_{\text{c}}, \tag{2.4}$$

where β is the angle between $\boldsymbol{b}$ and the line direction $\boldsymbol{\ell}$ of the dislocation, R_{c} is the core radius and E_{c}/L the core energy per unit length. The energy factor $k(\beta)$ depends on β and the elastic properties of the material. Assuming an isotropic medium, $k(\beta)$ can be evaluated as

$$k(\beta) = \mu\left(\cos^2\beta + \frac{\sin^2\beta}{1-\nu}\right) \tag{2.5}$$

with μ being the shear modulus and ν the Poisson's ratio as defined in Eq. (1.44) and (1.46).

The logarithmic term in Eq. (2.4) diverges for both the limits $R \to 0$ and $R \to \infty$. Continuum theory supposes that the atomic displacements vary slowly over the dimensions of a unit cell and this breaks down at the core resulting in the divergence as $R \to 0$. Also, for R of the order of magnitude of interatomic distances, continuum theory cannot describe the discrete system correctly. As a result, below a certain radius Eq. (2.4) does not give a good description of the elastic energy: The core radius R_c is defined by the condition that (2.4) ceases to be applicable for radii $R < R_c$. In other words, the dislocation core is the minimum region which cannot be described by elasticity theory and therefore, discrete (atomistic) models are necessary to evaluate the core energy E_c. If the core energy is known, the elastic strain energy for $R \geq R_c$ can be plotted as shown in Fig. 2.3.
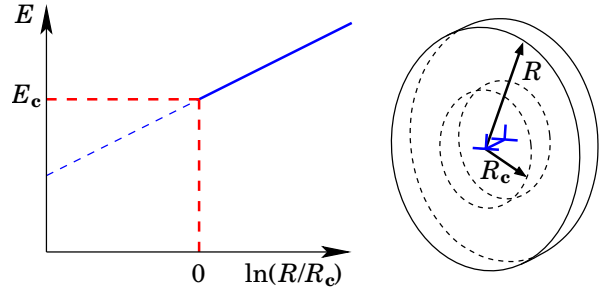
Figure 2.3: The elastic strain energy of a dislocation integrated in a surrounding cylinder. The *x*-axis is scaled logarithmically.

As already mentioned, Eq. (2.4) also diverges for $R \to \infty$. Consequently, in an infinite crystal we *cannot* evaluate a finite total elastic energy of a dislocation, but only its core energy, the energy factor $k(\beta)$ and core radius which together describe the variation of the strain energy with R.

2.3.2 The elastic interaction between two straight dislocations

Let us consider two straight and parallel dislocations A and B of arbitrary Burgers vectors $\boldsymbol{b}_A$ and $\boldsymbol{b}_B$ in an infinite and isotropic crystal. Then a somewhat involved calculation yields the elastic interaction energy per unit length:

$$\frac{E_{AB}(\boldsymbol{R})}{L} = -\frac{\mu}{2\pi}\left\{(\boldsymbol{b}_A \cdot \boldsymbol{\ell})(\boldsymbol{b}_B \cdot \boldsymbol{\ell}) + \frac{(\boldsymbol{b}_A \times \boldsymbol{\ell}) \cdot (\boldsymbol{b}_B \times \boldsymbol{\ell})}{1 - \nu}\right\}\ln\left(\frac{R}{R_0}\right)$$
$$-\frac{\mu}{2\pi(1-\nu)R^2}\left[(\boldsymbol{b}_A \times \boldsymbol{\ell}) \cdot \boldsymbol{R}\right]\left[(\boldsymbol{b}_B \times \boldsymbol{\ell}) \cdot \boldsymbol{R}\right] \tag{2.6}$$

Here $\boldsymbol{R}$ is the distance vector between the two dislocations and R its length. $\boldsymbol{\ell}$ gives the line direction ($|\boldsymbol{\ell}| = 1$). Eq. (2.6), which was first developed by Nabarro [49], allows the calculation of the interaction energy except for a constant shift $\propto \ln(R_0/\text{Å})$. Similar to the core energy in Eq. (2.4), this shift cannot be determined in linear elasticity theory. However, this shift does not influence the elastic force between the two dislocations. By differentiation we obtain the radial component of the interaction force per unit length:

$$\frac{F_{AB}(R)}{L} = -\frac{\partial}{\partial R}\left(\frac{E_{AB}(\boldsymbol{R})}{L}\right) = \frac{\mu}{2\pi R}\left\{(\boldsymbol{b}_A \cdot \boldsymbol{\ell})(\boldsymbol{b}_B \cdot \boldsymbol{\ell}) + \frac{(\boldsymbol{b}_A \times \boldsymbol{\ell}) \cdot (\boldsymbol{b}_B \times \boldsymbol{\ell})}{1 - \nu}\right\} \tag{2.7}$$

In analogy to this, the angular component can be derived by differentiation $-(1/R)\,\partial/\partial\vartheta$ perpendicular to $\boldsymbol{R}$. Eq. (2.7) will be useful in the calculation of equilibrium stacking fault widths in Section 4.4.1.

Figure 2.4: Dislocation glide in a simple cubic lattice: Lattice sites are shown as empty circles and the glide plane lies horizontally. The small arrows indicate the glide direction if one proceeds from the structure shown on the left to the one shown on the right.

Figure 2.5: Dislocation climb in a simple cubic lattice: Lattice sites are shown as empty circles and the glide plane lies horizontally. The small arrows indicate the climb direction if one proceeds from the structure shown on the left to the one shown on the right.

2.4 Dislocation motion

When talking of dislocation motion, we have to differentiate between two basic types: *Dislocation glide* and *dislocation climb*:

Dislocation glide: A dislocation in an otherwise perfect and infinite crystal will be in equilibrium and it will remain on its position. However, if either external forces act upon the crystal or if other defects are present, then this will generate a strain field which might result in a net force on the dislocation line and it will possibly move (see for example Eq. (2.7)). For dislocations with an edge component ($b \times \ell \neq 0$) this motion is restricted to the plane defined by the Burgers vector b and the line direction ℓ (Fig. 2.6). In this plane

Figure 2.6: The glide plane of a dislocation

the dislocation is able to move rather easily by breaking and forming new bonds as schematically depicted in Fig 2.4. This process is called dislocation glide and the plane accordingly *glide plane*. Pure screw dislocations, of course, do not have a featured glide plane in that sense.

Since during glide the number of lattice sites of the whole system remains constant, we speak of dislocation glide as *conservative motion*.

Dislocation climb: The second sort of dislocation motion is called dislocation climb. As shown in Fig. 2.5, during climb the dislocation moves from one glide plane to another. In the case of a pure edge dislocation with minimal Burgers vector, the direction of motion points exactly along the inserted half-plane: The half-plane either shrinks (upward climb, Fig. 2.5), or it expands (downward climb). Hence, as opposed to glide motion, for climb motion an absorption or creation of lattice sites is necessary. In a real crystal this can be achieved through the release or trapping of either interstitial atoms or lattice vacancies. Thus, climb does not require a driving elastic force acting upon the dislocation line.

Since dislocation climb involves a net change in the number of lattice sites, we speak of *non-conservative motion*.

2.5 The dissociation of dislocations

Looking at Fig. 2.2 it is obvious that a larger Burgers vector leads to larger lattice distortion in the region at and near the dislocation. Hence it seems plausible, that in terms of elastic strain energy only dislocations with the one or two smallest Burgers vectors are stable: A dislocation with a larger Burgers vector $\boldsymbol{b}$ might *dissociate* into dislocations A and B with smaller vectors $\boldsymbol{b}_A$ and $\boldsymbol{b}_B$. For topological reasons in this process the overall Burgers vector has to be conserved:

$$\boldsymbol{b} = \boldsymbol{b}_A + \boldsymbol{b}_B \tag{2.8}$$

Fig. 2.7 shows the dissociation of an edge dislocation with a Burgers vector of two times a minimal lattice translation in a simple cubic lattice. A comparison of energies using Eq. (2.4) gives a simple criterion for dissociation:

The generalised Frank criterion for dissociation: *A perfect dislocation with a Burgers vector $\boldsymbol{b}$ will dissociate into two perfect dislocations A and B if*

$$k(\beta)\,|\boldsymbol{b}|^2 > k(\beta_A)\,|\boldsymbol{b}_A|^2 + k(\beta_B)\,|\boldsymbol{b}_B|^2 \; . \tag{2.9}$$

For the original criterion F. C. Frank assumed an energy proportional to $|\boldsymbol{b}|^2$ only and therefore neglected the variation with β Hirth and Lothe [41].
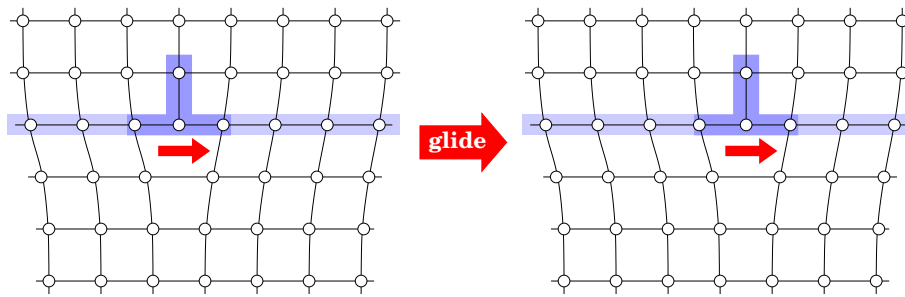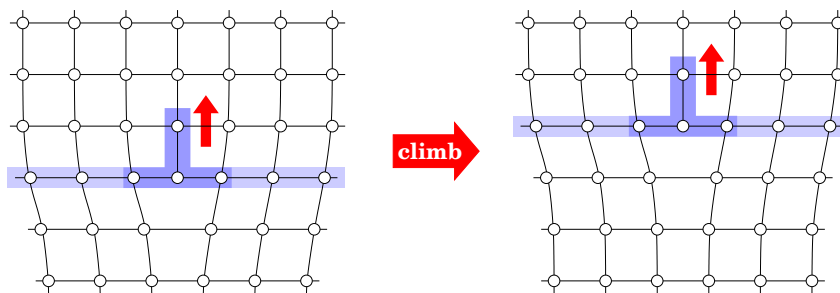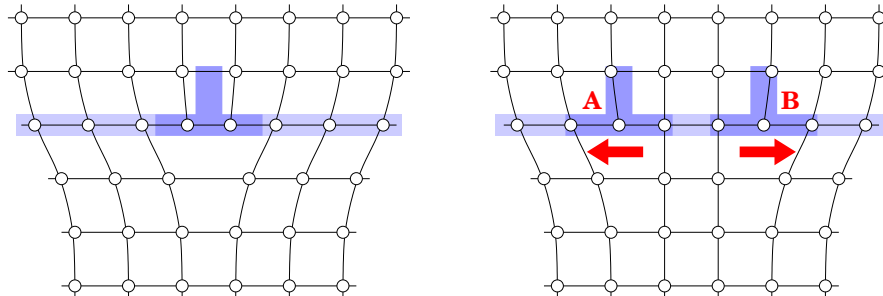

Figure 2.7: Dislocation dissociation in a simple cubic lattice: Lattice sites are shown as empty circles and the glide plane lies horizontally. The arrows indicate the direction of motion for the two resulting dislocations.

2.6 The influence of lattice periodicity: Kinks and jogs

Whilst the last sections dealt with straight perfect dislocations, we will now turn towards those deviations from the straight line, that represent a lateral offset. In the description of those so-called *kinks* and *jogs* the influence of lattice periodicity on the energy of a dislocation plays a major role.

2.6.1 The periodic displacement potential of a crystal

Plastically shearing a crystal on a rational plane[2] involves the stretching, breaking and forming of new bonds accordingly to the periodicity in this plane along the direction of shear. This means, microscopically the shear translation *x* will depend periodically on the applied shear stress σ and not linearly. The classical approach to this problem was that of Frenkel [50], who assumed a sinusoidal periodicity of the displacement potential energy, yielding a sinusoidal expression for the shear stress:

$$\sigma = \sigma_0 \sin \left(\frac{2\pi x}{T} \right) \tag{2.10}$$

Here *T* is the length of the minimal lattice translation vector in the plane. It is obvious, that a dislocation line moving in the same plane will experience a similar displacement potential that reflects the lattice periodicity. Starting from this idea, one can develop a dislocation model that includes lattice periodicity. The first and phenomenological approach is known as the *Frenkel-Kontorova model* [51, 52, 53], where a one-dimensional array of spring-connected balls lying on a periodic substrate serves as an analogue. However, this approach merely gives a qualitative description.

The *Peierls-Nabarro model* [54, 55] gives a more formal solution for the displacement potential. In this model we speak of a dislocation resting in a *Peierls valley*. In motion it has to periodically overcome the barriers (*Peierls hills*) between the valleys. An approximation for dislocation core energies results from this model. However, since in this work lattice periodicity will be treated atomistically, it is unnecessary to introduce the formalism of Peierls and Nabarro. A good overview is given in [41].

2.6.2 Dislocation kinks

Let us consider a dislocation, where segments have been displaced within the glide plane, keeping the same line direction ℓ. Stable low energy configurations will always be those, where each of the segments lies in a Peierls valley of the glide plane, as depicted in Fig. 2.8. The connections between the segments — those parts, where the dislocation line crosses a Peierls hill — are called dislocation kinks. A pair consisting of a right kink (RK) and a left kink (LK) might be formed on a straight dislocation by thermal fluctuations in the crystal: After the pair the dislocation runs in the same Peierls valley again. In fact it is

---

[2] A rational plane is a plane defined by two non parallel lattice translational vectors.
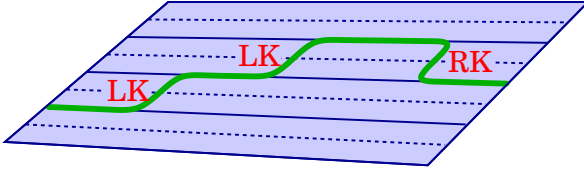
Figure 2.8: A kinked dislocation in the glide plane. Straight solid lines in the plane are Peierls valleys and dashed lines are hills. LK and RK denote left kinks and right kinks respectively.

assumed that for small forces[3] on the dislocation line, glide involves the thermal creation and migration of kinks. For their migration along the dislocation one has to imagine a similar variation of the Peierls potential — now however *along* the dislocation line and not perpendicular to it. Thus, glide velocity is controlled by the *kink formation energy* and the *kink migration barrier*. In semiconductors, where for atomistic processes like kink formation and migration quantum mechanical effects play an important role, the Peierls-Nabarro model and elasticity theory are not sufficient to calculate the related energies. Details on how this problem can be solved using DFT-based approaches will be given later in Section 4.5. However, here we can still calculate the elastic interaction energy of kinks: If we assume an infinite dislocation with a kink pair where the kink segments of the dislocation are perpendicular to the line direction ℓ, then the elastic kink–kink interaction energy is given as [41]:
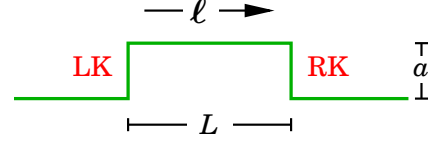
Figure 2.9: A dislocation kink pair.

$$E_{\text{LK,RK}}(L) = -\frac{\mu\,a^2}{8\pi\,L\,(1-\nu)}\left\{|\boldsymbol{b}\cdot\boldsymbol{\ell}|^2\,(1+\nu) + |\boldsymbol{b}\times\boldsymbol{\ell}|^2\,(1-2\nu)\right\} \tag{2.11}$$

Fig. 2.9 shows the corresponding situation. L is the kink–kink separation. The kink height a is given by the distance between the Peierls valleys, or in other words, by the lattice translational symmetry in the glide plane. The overall formation energy of the kink pair now can be written as:

$$E_{\text{pair}}(L) = E_{\text{f}}(\text{LK}) + E_{\text{f}}(\text{RK}) + E_{\text{LK,RK}}(L) \tag{2.12}$$

Here E_{f} denotes the formation energy of the respective single kink. Eq. (2.12) will play an important role in the calculation of the kink formation energies in Section 4.5. By differentiation of Eq. (2.11) we obtain the attractive force between kinks:

$$\begin{aligned} F_{\text{LK,RK}}(L) &= -\frac{\text{d}}{\text{d}L}\,E_{\text{LK,RK}}(L) \\ &= -\frac{\mu\,a^2}{8\pi\,L^2\,(1-\nu)}\left\{|\boldsymbol{b}\cdot\boldsymbol{\ell}|^2\,(1+\nu) + |\boldsymbol{b}\times\boldsymbol{\ell}|^2\,(1-2\nu)\right\} \end{aligned} \tag{2.13}$$

Note that the elastic interaction force between the two kinks falls off more rapidly with their separation L ($\propto 1/L^2$) than the elastic force between two parallel straight dislocations ($\propto 1/R$) in Eq. (2.7), where R denotes the separation of the dislocations.

---

[3] Small forces in this context are forces that are not large enough to lift the dislocation as a whole over the Peierls hill.
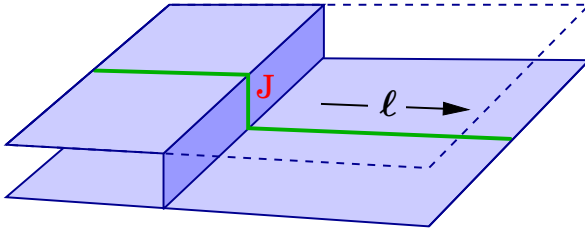
Figure 2.10: A dislocation jog. The upper of the two glide planes is only drawn half, so the underlying plane with the dislocation line can be seen. J denotes the position of the jog.

2.6.3 Dislocation jogs

Similarly to displacing segments of the dislocation within the glide plane, one can imagine the same normal to it. The connections between the displaced segments are called dislocation jogs. They are an analogue to dislocation kinks in the preceding section. Fig. 2.10 shows a single jog between two adjacent glide planes. Such a jog is also called unit jog, as opposed to a superjog, which spans across several glide planes. As kinks correspond to glide motion, jogs correspond to climb motion. This implies that jogs will neither form, nor migrate along the dislocation line by mere thermal activation. They are sessile and can only move along the dislocation line by trapping or releasing interstitial atoms or vacancies.

Analogous to the kink–kink interaction one can evaluate the elastic interaction energy of a pair consisting of an up jog (UJ) and a down jog (DJ) in a geometry similar to that of the kink pair in the last section [41]:

$$E_{\text{UJ,DJ}}(L) = -\frac{\mu a^2}{8\pi L (1 - \nu)} \left\{ |\boldsymbol{b} \cdot \boldsymbol{\ell}|^2 (1 + \nu) + |\boldsymbol{b} \times \boldsymbol{\ell}|^2 \right\} \tag{2.14}$$

Here for a unit jog a depends on the glide plane separation. For pure screw dislocations ($\boldsymbol{b} \times \boldsymbol{\ell} = 0$) with Eq. (2.14) we obtain the same result as for the kink–kink interaction energy in Eq. (2.11). This reflects the fact that screw dislocations do not have a featured glide plane: For a screw dislocation jog with respect to one glide plane one always finds another glide plane where this jog resides in. With respect to the latter plane the jog has to be considered a kink.

Chapter 3

Dislocations in Tetrahedrally Bonded Semiconductors

To prepare for the analysis of the examples of dislocations in various tetrahedrally bonded semiconductors, this chapter will give an overview of the low energy types of dislocations in the respective crystal structures. After discussing perfect and partial dislocations belonging to the crystal's main slip system, further types are introduced briefly.

Following Eq. (2.4), low energy dislocation structures are those with a minimum Burgers vector. Further, in Section 2.1 we learnt, that the Burgers vector of a perfect dislocation is always a linear combination of the lattice translational vectors. The same holds for the dislocation line direction. Thus, if the lattice structure is known, one can already specify possible low energy types of dislocations.

The prevalent bonding configuration in semiconductor crystals is that of each atom being bonded to four immediate neighbour atoms. The electronic configuration of each atom is then approximately given as an sp^3 hybrid. Ideally, the four neighbours are located at the vertices of a tetrahedron around the central atom, forming bond angles of $109.5°$. With this bonding several periodic structures are possible. The two most common crystal structures are those of cubic symmetry and a hexagonal structure known as $2H$. Both can be described as a lattice with a two-atom basis, or equivalently, as two intersecting lattices [48]. In Fig. 3.1 the structures are depicted for compound semiconductors, where the basis consists of two atoms of different species. In both cases the lattice is that of a closed-packed structure: A face-centred cubic (fcc) lattice and a hexagonal closed-packed (hcp) lattice[1]. The characteristic parameters of the respective structure are its lattice constants. The cubic cell is characterised by just one lattice constant a_0. Ideally, one lattice constant would be sufficient to describe the hexagonal cell as well. In real crystals however, variations in bond lengths and angles lead to deviations from the ideal ratio 1.633 between cell height and edge length of the basal hexagon. Thus the structure is defined by two lattice constants a_0 and c_0 (see. Fig. 3.1).

---

[1]In a hard sphere model of the lattice, where each sphere is in direct contact with its neighbours, both fcc and hcp have the maximum possible space-filling (74%). For comparison: The space-filling of body-centred cubic (bcc) lattices is 68%.
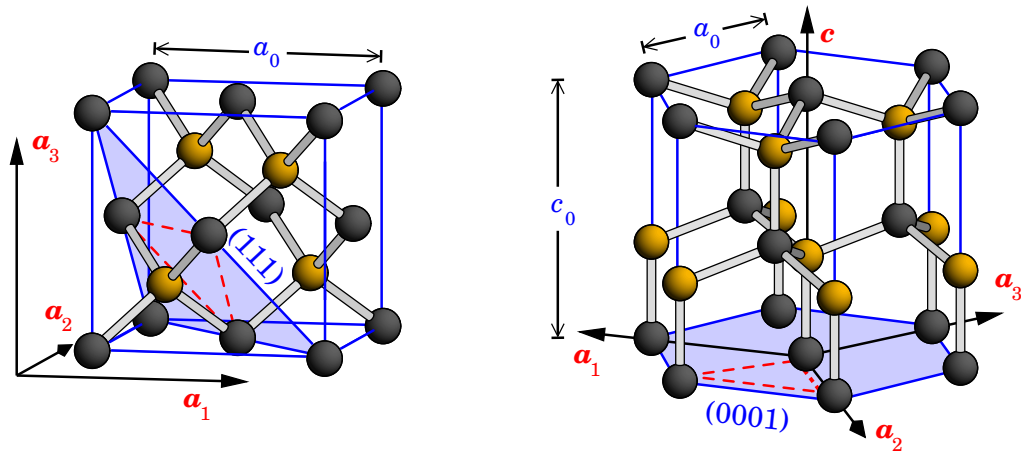
Figure 3.1: Unit cells of tetrahedrally bonded semiconductors. *Left:* The cubic structure, also known as the zincblende structure. It can be constructed as a face-centred cubic lattice with a two-atom basis. The lattice constant a_0 is given as the edge of the cubic cell. The three perpendicular basis vectors are labelled a_1 to a_3. *Right:* The hexagonal $2H$ structure, also known as the wurtzite structure. It can be constructed as a hexagonal closed-packed lattice with a two-atom basis. The two lattice constants are given as one edge of the basal hexagon (a_0) and the height of the unit cell (c_0). The basis vectors are labelled a_1 to a_3 and c. The a_i lie in one plane and form angles of $120°$. a_3 is conventionally introduced for reasons of symmetry in the description.
In both unit cells a plane with the highest density of lattice sites is drawn: (111) and (0001) for cubic and hexagonal symmetry respectively. The minimum lattice translations within these planes are shown as dashed lines.

3.1   Crystal slip systems and perfect dislocations

In the process of plastic deformation one part of the crystal might be sheared macroscopically with respect to the other part. This so-called *crystal slip* usually occurs on specific crystallographic planes only. This can be explained microscopically: Crystal slip involves the formation and propagation of dislocations [41]. Those dislocations which have minimum Burgers vectors are easiest to form. Since a Burgers vector is always given as a lattice translation, dislocations are preferentially formed on planes containing the minimum lattice translations. These planes are the preferred slip planes. A slip plane contains both the Burgers vector and the line direction of the generated dislocations. Hence the plane of crystal slip is identical with the glide plane of the dislocations involved in the process. Let us now investigate the preferred slip systems of the two crystal lattices introduced above.

3.1.1   The main slip systems

The $\{111\}\langle110\rangle$ slip system in fcc lattices: In the cubic structure the planes of highest lattice site density are the $\{111\}$ planes. Fig. 3.1 (left) shows one specific plane of that family. These planes contain the minimum lattice translations and are thus preferred slip planes. The slip system is defined by the plane and the direction of slip, conventionally written as $\{111\}\langle110\rangle$. For each plane we have three possible

directions. Specifically in the (111) plane the minimum lattice translations (Burgers vectors) are $\frac{1}{2}[1\bar{1}0]$, $\frac{1}{2}[\bar{1}01]$ and $\frac{1}{2}[01\bar{1}]$. They are drawn as dashed lines in Fig. 3.1 (left).

Since the line direction, at least locally, is also given by lattice translations, the afore mentioned directions yield two basic dislocation types with minimum Burgers vector:

- The screw dislocation with line direction and Burgers vector parallel, e.g. $\boldsymbol{\ell} = [1\bar{1}0]$ and $\boldsymbol{b}_{\mathrm{s}} = \frac{1}{2}[1\bar{1}0]$.

- The 60° dislocation, where line direction and Burgers vector form an angle of 60°, e.g. $\boldsymbol{\ell} = [1\bar{1}0]$ and $\boldsymbol{b}_{60} = \frac{1}{2}[0\bar{1}1]$.

The $\{0001\}\,\langle 11\bar{2}0\rangle$ slip system in hcp lattices: Whilst in the fcc system there is a whole family of planes with the highest lattice site density, in the hcp structure there is only one, the (0001) or *basal plane*, as shown in Fig. 3.1 (right). The slip system is of the $\{0001\}\,\langle 11\bar{2}0\rangle$ type and the minimum lattice translations are given as $\frac{1}{3}[11\bar{2}0]$, $\frac{1}{3}[1\bar{2}10]$ and $\frac{1}{3}[\bar{2}110]$. They are drawn as dashed lines in Fig. 3.1 (right). The two basic dislocation types with minimum Burgers vector are:

- The screw dislocation with line direction and Burgers vector parallel, e.g. $\boldsymbol{\ell} = [11\bar{2}0]$ and $\boldsymbol{b}_{\mathrm{s}} = \frac{1}{3}[11\bar{2}0]$.

- The 60° dislocation, where line direction and Burgers vector form an angle of 60°, e.g. $\boldsymbol{\ell} = [11\bar{2}0]$ and $\boldsymbol{b}_{60} = \frac{1}{3}[2\bar{1}\bar{1}0]$.

3.1.2 Glide and shuffle structures

In a lattice with a one-atom basis the Burgers vector and the line direction uniquely define the dislocation type. However as mentioned before, the crystal structure of tetrahedrally bonded semiconductors is that of a lattice with a two-atom basis. A direct consequence is, that dislocations with a featured glide plane exist in *two* geometrically different types: As illustrated in Section 2.2, the edge component of a dislocation involves an extra half-plane of lattice sites on one side of the glide plane. With a two-atom basis, for one specific glide plane the extra half-plane can terminate on either of the two atoms, resulting in two geometrically different types — the *glide* and the *shuffle* dislocation. Fig 3.2 shows the atomistic situation for the perfect 60° dislocation in cubic material: If the half-plane terminates between two closely-separated (111) planes we obtain the glide structure, if it terminates between widely separated planes we obtain the shuffle structure. The transition between shuffle and glide is a climb process and thus involves the trapping or release of interstitial atoms or vacancies (compare with Fig. 2.5 in Section 2.4). As we will see in later chapters, the glide structure is usually the energetically more stable structure.

Further, in compound semiconductors AB with two atom species A and B, we have to distinguish the atom species of the terminating line of atoms of the extra half-plane. Conventionally the respective structures are called α and β dislocations. Hence depending on the combination of glide and shuffle with α and β there are 4 different types of 60° dislocations.
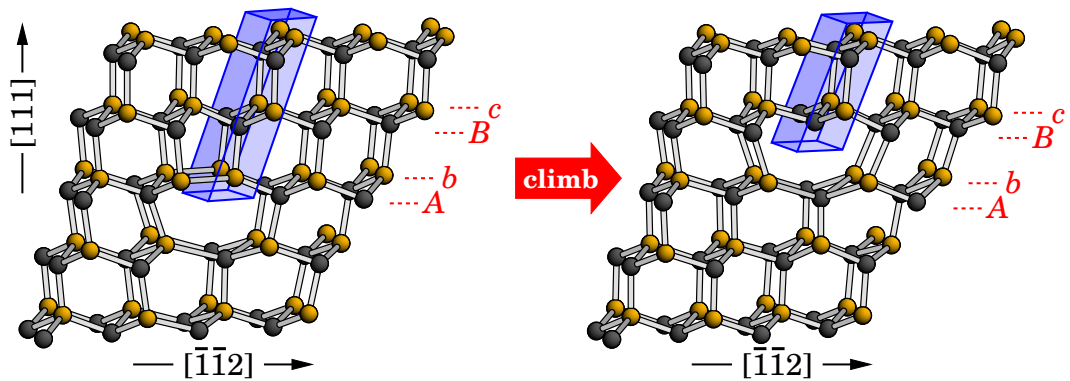
Figure 3.2: The glide and the shuffle structure of the perfect 60° dislocation in cubic material. Only two $(1\bar{1}0)$ layers along the dislocation line are shown for each structure. The extra half-plane is enclosed by a shaded box. Depending on where the half-plane terminates, we obtain the glide or the shuffle structure. The transition between both is given by a climb process as in Fig. 2.5. *Left:* The 60° glide dislocation. Here the half-plane terminates between two closely-spaced (111) planes *A* and *b*. *Right:* The 60° shuffle dislocation. Here the half-plane terminates between two widely-spaced (111) planes *b* and *B*.
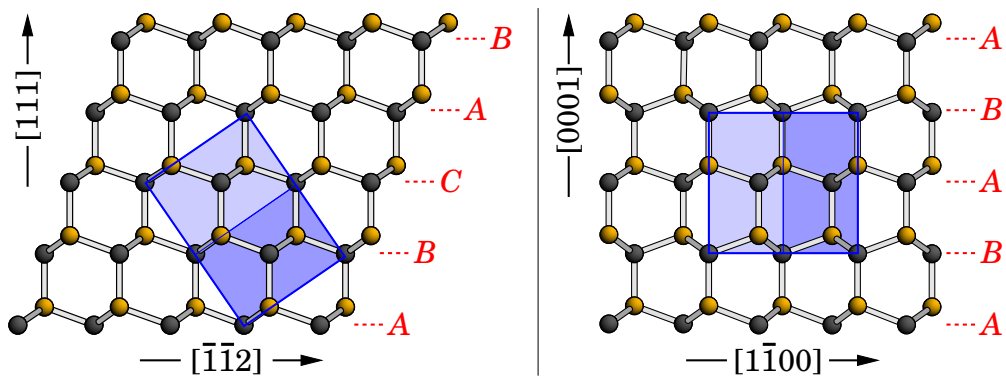
Figure 3.3: The Stacking sequence in the cubic and hexagonal lattice. In both structures the respective unit cell from Fig. 3.1 is drawn. *Left:* The cubic structure. The view is along $[1\bar{1}0]$ and the (111) slip planes lie horizontally. *Right:* The hexagonal structure. The view is along $[11\bar{2}0]$ and the (0001) basal slip planes lie horizontally.
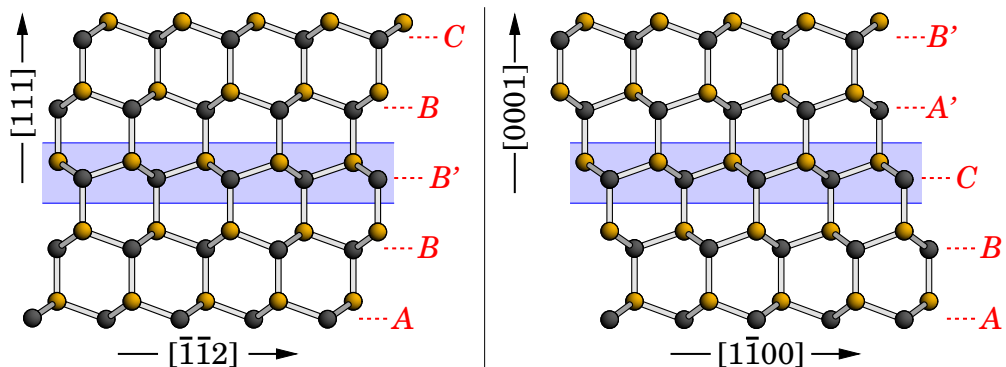
Figure 3.4: Intrinsic stacking faults in the cubic and hexagonal lattice. In both structures the stacking fault plane is shaded. *Left:* An intrinsic stacking fault in the cubic structure. *Right:* An intrinsic stacking fault in the hexagonal structure.

3.2 Crystal stacking and partial dislocations

Looking only at one of the two sub-lattices, the preferred slip planes mentioned in the last section are also the planes stacked with the widest separation. The cubic and the hexagonal structure can not only be distinguished by their distinct unit cells, but alternatively also by their respective stacking sequence of these planes. As can be seen in Fig. 3.3, the stacking sequence is $\cdots ABC|ABC\cdots$ in cubic[2] and $\cdots AB|AB|AB\cdots$ in hexagonal material[3]. Faults in this stacking sequence are very common planar defects. If the normal sequence is maintained on both sides of the fault, then we speak of an *intrinsic stacking fault* (ISF), otherwise of an *extrinsic fault*. In this work only intrinsic faults are considered. Fig. 3.4 shows an ISF in cubic and in hexagonal material. In the ISF plane one type of atoms is displaced with respect to the bulk plane in normal stacking. To illustrate this, we compare plane C of the cubic crystal structure (Fig. 3.3 (left)) with the respective plane B' in the cubic stacking fault (Fig. 3.4 (left)). Fig. 3.5 shows the atom positions of the displaced species in the ISF plane and in the corresponding bulk plane. The offset is $\boldsymbol{o} = \frac{1}{6}[1\bar{2}1]$ or equivalently $\frac{1}{6}[11\bar{2}]$ and $\frac{1}{6}[\bar{2}11]$. The situation is exactly the same in the hexagonal structure. Actually one only has to relabel the crystallographic directions $[1\bar{1}0]$ and $[\bar{1}\bar{1}2]$ in Fig. 3.5 with $[11\bar{2}0]$ and $[1\bar{1}00]$ respectively, since the stacking fault lies in the basal plane. In hexagonal coordinates we obtain an offset of $\boldsymbol{o} = \frac{1}{3}[10\bar{1}0]$ or equivalent.
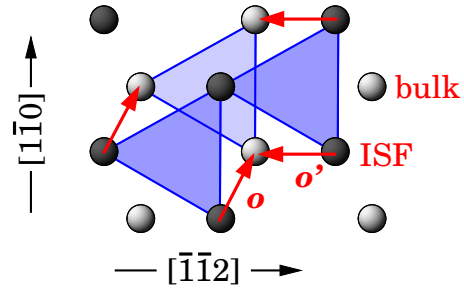
Figure 3.5: The offset between bulk and faulted region (ISF) for the sublattice of one atom species.

Let us now assume a dislocation, which is not surrounded by ideal lattice only, but which borders an intrinsic stacking fault in its glide plane. This type of dislocation is called *Shockley partial dislocation* [41]. The minimum Burgers vector of such a dislocation is given as the offset vector $\boldsymbol{o}$ between the faulted and unfaulted region of the glide plane.

Fig. 3.6 (left) shows the resulting principal Burgers vectors. The two partials are:

- The 30° Shockley partial. E.g. in cubic material: $\boldsymbol{\ell} = [1\bar{1}0]$ and $\boldsymbol{b}_{30} = \frac{1}{6}[1\bar{2}1]$ and in hexagonal material: $\boldsymbol{\ell} = [11\bar{2}0]$ and $\boldsymbol{b}_{30} = \frac{1}{3}[10\bar{1}0]$.

- The 90° Shockley partial. E.g. in cubic material: $\boldsymbol{\ell} = [1\bar{1}0]$ and $\boldsymbol{b}_{90} = \frac{1}{6}[\bar{1}\bar{1}2]$ and in hexagonal material: $\boldsymbol{\ell} = [11\bar{2}0]$ and $\boldsymbol{b}_{90} = \frac{1}{3}[1\bar{1}00]$.

---

[2] Therefore the cubic structure is also alternatively referred to as the 3*C* structure, since the same three planes are periodically repeated in the [111] direction. It should be noted, that we attach a collective label (A, B, C) to actually a pair of planes — one plane per sub-lattice. To be precise one would have to label the sub-lattices separately as done in Fig. 3.2. This yields for the cubic structure $\cdots aAbBcC|aAbBcC\cdots$.

[3] Analogously to the cubic structure, the hexagonal structure presented here is referred to as 2*H*. Also other hexagonal structures are possible, for example 4*H* and 6*H*, with a stacking sequence of $\cdots ABAC|ABAC\cdots$ and $\cdots ABCACB|ABCACB\cdots$ respectively. 4*H* and 6*H* are very common polytypes of silicon carbide crystals.
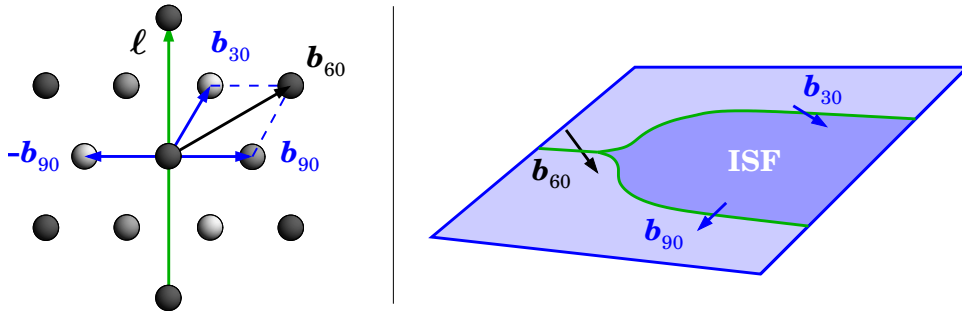
Figure 3.6: The dissociation of the perfect 60° dislocation. *Left:* Burgers vector of the perfect 60° dislocation and of the 30° and 90° Shockley partials. The vectors of the partial dislocations are given as the offset vectors between faulted and unfaulted region of the glide plane (compare Fig. 3.5). We obtain $-\boldsymbol{b}_{90}$ or $\boldsymbol{b}_{90}$ depending on which side of the dislocation the faulted region is located. The perfect 60° dislocation can be decomposed into the two Shockley partials: $\boldsymbol{b}_{60} = \boldsymbol{b}_{30} + \boldsymbol{b}_{90}$. *Right:* Schematic sketch of the corresponding dissociation reaction in the glide plane.

As shown in Fig. 3.6 (left), the Burgers vector of the perfect 60° dislocation is given as the sum of the two Shockley partials: $\boldsymbol{b}_{60} = \boldsymbol{b}_{30} + \boldsymbol{b}_{90}$. The Burgers vectors of the Shockley partials are smaller than that of the perfect dislocation: The difference is large enough, to ensure that the Frank criterion (2.9) is usually satisfied. Hence a dissociation is very likely if the energy of the stacking fault is not too large. The dissociation reaction is given as:

$$\boldsymbol{b}_{60} \quad \longrightarrow \quad \boldsymbol{b}_{30} + \mathrm{ISF} + \boldsymbol{b}_{90} \tag{3.1}$$

The geometrical situation is shown schematically in Fig. 3.6 (right): The two partials are separated by an intrinsic stacking fault between them. Their equilibrium distance is determined through the energy of the stacking fault and the partial–partial interaction force, usually given as the elastic interaction as in Eq.(2.7). Fig. 3.7 gives an impression of the microscopic arrangement of the two partials and the stacking fault in cubic a crystal. The dissociation distance shown is rather small: In real semiconductor crystals the two partials are very often dissociated with widths of 10 or more lattice constants, depending on the stacking fault energy.

Analogously the screw dislocation can dissociate into two 30° partials enclosing an intrinsic stacking fault:

$$\boldsymbol{b}_{\mathrm{s}} \quad \longrightarrow \quad \boldsymbol{b}_{30} + \mathrm{ISF} + \boldsymbol{b}_{-30} \tag{3.2}$$

Here $\boldsymbol{b}_{-30}$ is given as $\frac{1}{6}[2\bar{1}\bar{1}]$ for cubic and $\frac{1}{3}[\bar{1}2\bar{1}0]$ for hexagonal material, assuming ℓ, $\boldsymbol{b}_{\mathrm{s}}$ and $\boldsymbol{b}_{30}$ as introduced above. In a compound the two 30° partials are of opposite termination (α and β type) to maintain stoichiometry.

It has to be noted, that in this chapter all atomic dislocation structures were deduced from displacing the lattice according to the respective combination of Burgers vector and line direction. They do not represent real structures! The atomic structures close to the dislocation line — the so called *core structures* — might reconstruct and vary with material. Examples will be given in the following chapters.
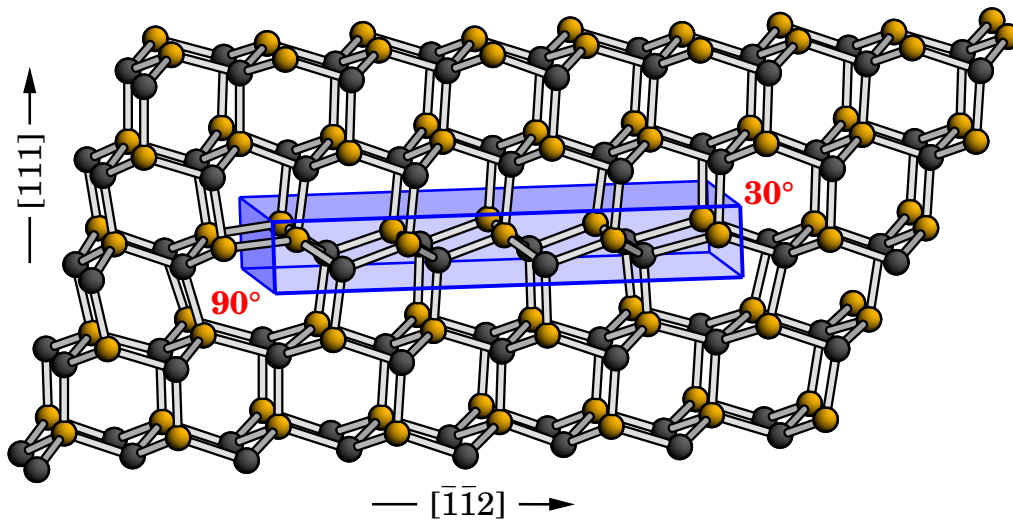
Figure 3.7: The dissociation of the perfect 60° glide dislocation seen atomistically. Only two (1$\bar{1}$0) layers along the dislocation line are shown. The intrinsic stacking fault separating the 90° and 30° Shockley partials is enclosed by a shaded box. The very core region of each partial is located at the respective end of the stacking fault.

3.3 Further classes of dislocations

All dislocations discussed so far can be generated through plastic deformation of the crystal and belong to a low energy slip system. In real crystals however, other classes of dislocations might be present. The most common are *misfit dislocations*, *threading dislocations* and *grain boundary dislocations*:

Misfit dislocations: Semiconductor crystals are very often grown on a substrate with a considerable lattice mismatch. Misfit dislocations compensate this mismatch and are basically the termination lines of additional lattice planes. They are very often similar or equivalent in structure to dislocations of the crystal's main slip system.

Threading dislocations: During crystal growth, unevenness of the substrate surface or the collision of growth islands might lead to the nucleation of dislocations with line direction parallel to the direction of growth. These dislocations thread the epilayer and are very common in hexagonal semiconductors like hexagonal GaN [56]. The two basic types are the *threading screw* (ts) and the *threading edge dislocation* (te), both with a line direction of $\boldsymbol{\ell} = [0001]$ and Burgers vectors $\boldsymbol{b}_{\text{ts}} = [0001]$ and $\boldsymbol{b}_{\text{te}} = \frac{1}{3}[1\bar{2}10]$ respectively.

Grain boundary dislocations: These dislocations occur at the boundary between two tilted crystal grains in a polycrystal. They compensate for the tilt angle between the two grains. They are very often similar or equivalent in structure to dislocations of the crystal's main slip system or to threading dislocations.

Chapter 4

Dislocations in Diamond

After a short introduction to diamond as a material and to the experimental evidence for dislocations in diamond, this chapter demonstrates, how the methods described in Chapter 1 can be combined to describe the different aspects of dislocations in semiconductors. With the 30°glide partial dislocation as a first example, the chosen model geometry (a supercell-cluster hybrid) and its convergence with size are tested in comparison with elasticity theory. Subsequently the low energy core structures of the remaining predominant perfect and partial dislocations of the $\{111\} \langle 110 \rangle$ slip system in diamond are determined and their core energies and the long range elastic energy contributions are calculated. This is then followed by a discussion of possible dissociation reactions and dislocation glide of the partials is considered. The latter requires to model kink formation and migration processes on the atomistic level.

In the last part of this chapter it is shown how the calculated core geometries can be used as input coordinates for HRTEM simulations, while the calculation of the electronic structures allows the modelling of EEL spectra. Thus a direct link to experiments can be established.

4.1   Introduction and background

Diamond exhibits some interesting physical properties, like extreme hardness, high heat conductivity, high refractive index and a wide indirect electronic band gap (5.49 eV)[57], which make it an interesting material not only as a gemstone but also for technological applications. The development of several techniques to produce synthetic diamond at rather low costs allows its technological application: Taking advantage of its high hardness, synthetic diamond powder is preferentially used for the abrasive coating of cutting tools.

Conventional growth techniques for single crystal semiconductors usually involve the melting of the material (e.g. the Czochralski method) — this unfortunately fails for diamond due to its high melting point at the high pressure needed to avoid graphitisation. However, chemical vapour deposition (CVD) growth of diamond produces thin films, which might be used for semiconductor devices [58]. The wide band gap would make diamond an especially attractive semiconductor in optoelectronics or for high-temperature high-power devices. *p*-type doping of dia-

mond can be achieved by boron implantation, but *n*-type doping yet remains an unsolved problem, hindering the widespread application of diamond as a semiconductor[1]. Nevertheless, CVD diamond films are currently being used in high-power microwave devices, as heat spreaders in power transistors and as optical infrared-transmissive windows — to name just a few applications. Recent developments look promising for high voltage devices. ABB and De Beers Industrial Diamonds (UK) report single-crystal CVD diamond with carrier lifetimes > 2 μs, carrier mobilities around 4000 cm^2/Vs and 4 kV diodes with a breakdown voltage > 4 MV/cm [58, 60].

4.1.1   Experimental evidence for dislocations in diamond

In general, natural gem-quality type Ia diamonds, where nitrogen is present in an aggregated form, (see Dyer et al. [61] for the classification scheme) are almost dislocation-free [57]. Chemically pure type IIa, however, has been reported to contain a dislocation density of up to 10^7 cm^{-2} and boron containing type IIb also proves to be dislocation-rich [62, 63, 64]. Further, CVD-grown polycrystalline diamond reveals high densities of dislocations — sometimes up to 10^{12} cm^{-2}. Some originate at the substrate-interface and propagate through the thin film [65], but others lie at or near grain boundaries [66]. As in many other fcc semiconductors, the slip system in diamond is of the $\{111\}$ $\langle 110 \rangle$ type. This means that perfect dislocations lying on $\{111\}$ planes possess $\langle 110 \rangle$ Burgers vectors and line directions [41, 67, 68, 69] (see also Fig. 4.1). Details on this slip system have been given in Chapter 3.

Using weak-beam electron microscopy, Pirouz et al. [69] found that dislocations in type IIa diamond are dissociated into glide partials separated by an intrinsic stacking fault ribbon of width $25 - 42$ Å. Further, many extended dislocation nodes and dipoles consisting of $60°$ dislocations were observed. This is similar to the properties of deformed Si and Ge.

Further it is known, that dislocations in diamond influence the electrical and optical properties of the material: In undoped CVD diamond and natural type-II diamond dislocations have been correlated with the so-called *A*-band cathodoluminescence at $2.8 - 2.9$ eV [62, 63, 70, 71]. This luminescence is polarised with the electrical field parallel to the dislocation line [62, 63] and is associated with under-coordinated carbon atoms [71].

4.1.2   HPHT treatment — a threat to the international gem trade

More recently, it has also been observed that colour changes produced in natural brown diamonds by high-pressure, high-temperature annealing (HPHT) are linked to plastic deformations and dislocations [72]. Even complete decolouring of natural brown diamond by HPHT treatment has proven possible and the technique is applied in gem industry (General Electric and Lazare Kaplan): The majority of diamonds is brown and useless as gem stones. If treated however, they can be sold

---

[1] However, *n*-type doping seems to improve over recent years. Nowadays phosphorous *n*-type doping at least seems to be 100% reproducible [59] with increasing carrier lifetimes.
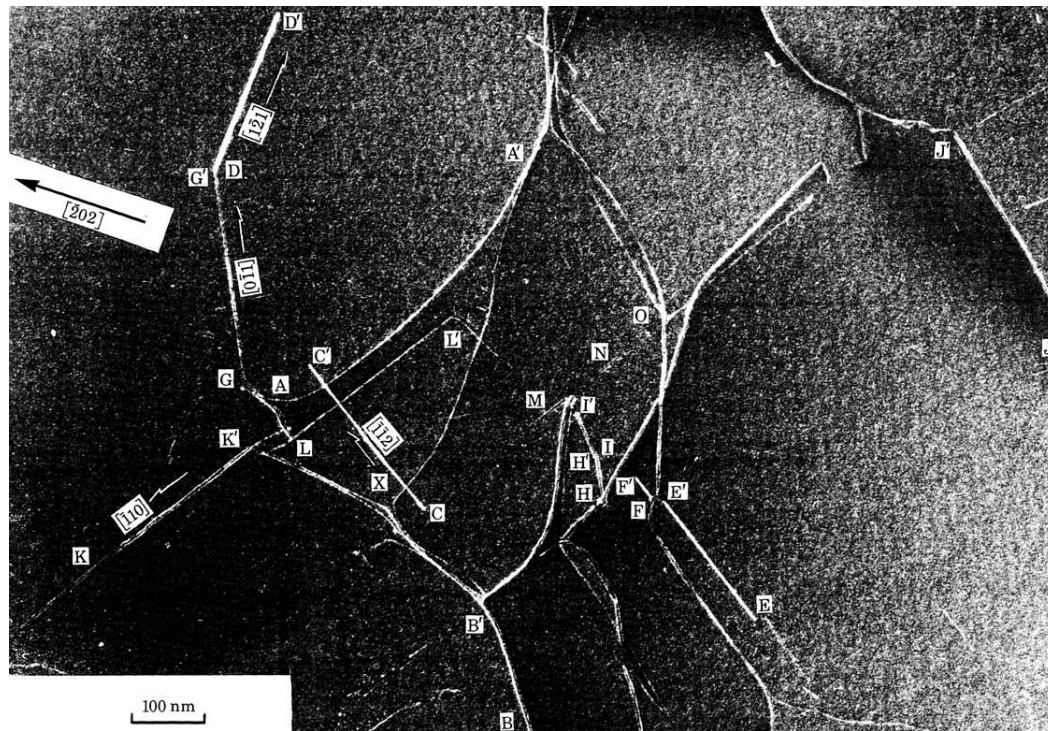
Figure 4.1: A weak-beam image of a typical dislocation distribution in diamond (diffraction vector $g = [\bar{2}02]$ indicated). To give some examples of the dislocations shown: AA' and BB' are glide dislocations; GG' and HH' are dipoles consisting of $60°$ dislocations and M, N and O are faulted dipoles (dissociated dipoles consisting of 4 partials). For detailed information on this image see Ref. [69]. (Reproduced from Ref. [69] with kind permission from the authors.)

as a cheaper alternative. Since the detection of HPHT treatment is far from trivial, dubious companies now try to sell treated diamonds as natural untreated stones. If they succeed, their profit is enormous, as is the loss for the major diamond selling companies[2]. Understandably, research to find ways of easy detection and thus eliminate this commercial threat to the gem trade is part of the major companies' gem defensive work [74, 75].

4.1.3 Earlier theoretical work

In contrast to silicon, relatively little theoretical work has been carried out on dislocations in diamond. Early studies by Nandedkar and Narayan [76] involved the usage of classical empirical potentials while recent *ab-initio* studies were restricted to the $90°$ partial glide dislocation [77, 78, 79]. Also work related to hydrogen and solitons at dislocations [80, 81, 82] and to graphitisation at core structures [83] considered the $90°$ partial only. *Hence a systematic investigation of the core structures, energies, mobility and electrical activity by means of DFT-based calculations was still missing.*

---

[2] HPHT costs are around 800 US $ per stone, yielding colourless diamonds worth an average value of 15,000 US $. However, not all treated diamonds appear to be colourless, IIa-type stones might show pink colouring and Ia green or even fancy yellow. The untreated natural equivalent of the latter can be sold for up to 100,000 US $ per carat [73]. Thus, treated stones also threaten the top end of the market.

4.2   The atomic scale modelling of dislocations

As discussed in Section 2.3, continuum elasticity theory is unable to address properties of the core where the strain diverges. Instead, the core energy and structure of a dislocation requires an atomistic treatment. As we will see later in this section, the modelling of the core often involves a considerable number of atoms — making a fully self-consistent treatment by means of the DFT-pseudopotential approach (AIMPRO, Section 1.1.2) computationally very expensive. Hence here the method of choice is the DFTB method described in Section 1.1.3, which is capable of conveniently performing structural relaxations of more than 700 atoms on a workstation or even over 1500 on state-of-the-art supercomputers. To verify the correct representation of dislocation core structures by DFTB, a range of small dislocated clusters and supercells containing around 200 atoms have been modelled with both methods, DFTB and AIMPRO. Even though these small clusters and supercells were not large enough to model a dislocation embedded in otherwise perfect bulk material, they are sufficient to compare the representation of typical bonding situations in dislocation cores between different methods. In this comparison DFTB has proven capable to model core structures comparably to AIMPRO — both approaches yield very similar lowest energy structures and deviations in bond lengths and angles are well below 2 %

As shown in Chapter 2, the determination of energies and forces in continuum elasticity theory always involves the elastic constants of the respective material. To now sensibly compare atomistic results with elasticity theory, one has to apply the elastic constants corresponding to the respective atomistic method and not the experimental ones. Hence the knowledge of the elastic constants within the DFTB method is crucial. The three independent elastic constants of diamond in a conventional cubic unit cell were found by suitably deforming the cell and calculating the respective total energies. The integration over the Brillouin zone was accomplished using a Monkhorst-Pack-optimised set of $3 \times 3 \times 3$ k-points [84]. The resulting elastic constants are given in Table 4.1. These are about 40 GPa larger than the experimental ones giving errors ranging from 3.7 % for c_{11} to 38 % for c_{12}. The calculated bulk modulus exceeds the experimental one by 10 %[3].

---

[3] Just for comparison, AIMPRO gives $B = 466$ GPa, which is 5 % too large.

Table 4.1: Some elastic properties of diamond: Comparison between calculated and experimental data. The first three columns give the independent elastic constants c_{ij} of diamond. The resulting shear modulus μ, the Lamé constant λ, the Poisson's ratio ν and the bulk modulus B are calculated as Voigt averages following Eq. (1.49),(1.46), and (1.47). All values are given in GPa, except for ν, which is dimensionless.

	c_{11}	c_{12}	c_{44}	μ^{Voigt}	λ^{Voigt}	ν^{Voigt}	$B = \frac{1}{3}(c_{11} + 2c_{12})$
Exp.[a]	1076	125	577	536	84	0.068	442
DFTB[b]	**1116**	**172**	**608**	**554**	**118**	**0.088**	**487**

[a]Experimental data by McSkimin and Bond [85] (ultrasonic waves) and Grimsditch and Ramdas [86] (Brillouin scattering). Within the accuracy given here, both report the same values.

[b]Values obtained using the DFTB method.

Some remarks on precision and errors: The large relative deviation for c_{12} seems considerable, and in fact it is this error which gives rise to the rather large error in the bulk modulus and the Poisson's ratio. One has to keep in mind though, that the quantities essential in dislocation related calculations are not B or ν, but μ, $1 \pm \nu$ and $1 - 2\nu$ (see all expressions for energies and forces in Chapter 2). The respective deviations are at maximum only 3.4 %. Consequently, the observed large errors for the bulk modulus and the Poisson's ratio have no adverse effect on the precision of the calculations performed in this work.

4.2.1 The supercell-cluster hybrid as the model of choice

Having shown the capability of the DFTB method to describe the significant elastic properties of diamond with acceptable accuracy, one now has to decide about the way how to model dislocations with this method. Generally there are two main approaches used to model defects in semiconductors atomistically: The *cluster* and the *supercell* approach. For the case of dislocations this means that either an atom cluster containing a single dislocation is considered or a supercell containing a dislocation multipole. The long range elastic effects are treated differently in each case but neither approach treats these effects rigorously, as will be discussed below.

The cluster approach: Here a single straight dislocation is inserted into an atomic cluster. As a result, the dislocation line intersects with the surface of the cluster. Fig. 4.2 depicts a cluster, whose surfaces are cut as crystallographic planes. For small clusters it is advisable to impose the displacements expected from elasticity theory for a dislocation in an infinite crystal at the surface of the finite cluster. In other words one then has to embed the cluster in a continuous medium, which represents the surrounding crystal environment. However, if the cluster is

Figure 4.2: Schematic sketch of a dislocation in a cluster. A and B denote the points of dislocation–surface intersection.

large enough, the difference with a freely relaxed surface is negligible [87].

In DFT-based methods where the electronic energy is explicitly included, the cluster surfaces are usually terminated with hydrogen atoms to avoid electronic gap states associated with dangling bonds or surface reconstructions, which might induce additional strain in the cluster core.

The main advantage of the cluster model is the possibility to study a *single and isolated* dislocation, but its principal disadvantage stems from the part of the dislocation line which intersects the surface. The dislocation core structure will be distorted near the surface, where it loses its periodicity. Thus, the cluster has to be considerably extended in the direction of the dislocation line. Experience from modelling dislocations in Si shows that clusters of around 700 atoms with approximately 6 repeat distances along the dislocation line are necessary, as applied in [2].

The supercell approach: Here a dislo-
cation dipole or multipole is placed in a
supercell. The periodic boundary con-
ditions require an overall Burgers vec-
tor of $b_{tot} = 0$. Thus, the most sim-
ple configuration is a dipole with two
dislocations of opposite Burgers vector.
Fig. 4.3 depicts an example supercell
containing a dipole, indicating its peri-
odically repeated images in one dimen-
sion as an example. In the case of a

Figure 4.3: Schematic sketch of a dislocation
dipole in a supercell.

dipole, the volume per single dislocation is only half that of the cell. Since DFT
calculations are nowadays limited to a few hundred atoms, usually the two disloca-
tions are separated by a relatively small distance. A result of this is a considerable
dislocation–dislocation interaction within and across the cell boundaries to neigh-
bouring cells and in compound semiconductors also a considerable charge transfer
might occur. The elastic interaction can be described by a procedure similar to an
Ewald sum, but the resulting expressions are not easy to handle [88]. Blase et al. [79]
perform this summation by taking advantage of an analytical solution found for tilt
boundaries. However, the same authors conclude that the core reconstruction of a
dislocation depends on the stress state. This implies that the stress of the selected
supercell geometry influences the core structure and hence possibly the core energy.

Seeing the advantages and disadvantages of both the cluster and the supercell, it is
a rather straightforward idea to combine the advantages of both in a *supercell-cluster
hybrid* approach — minimising the problems associated with either of the preceding.

The supercell-cluster hybrid approach:
Here the dislocation is placed in a model
which is periodic along the dislocation line,
however, it is non-periodic with a hydrogen-
terminated surface perpendicular to the line
direction [9, 89]. This allows to maintain the
'natural' dislocation periodicity as a line defect
and at the same time avoid the interactions
between dislocations in different cells. The
single dislocation in the hybrid model is sur-
rounded by twice the amount of bulk crystal

Figure 4.4: Schematic sketch of a dis-
location in a supercell-cluster hybrid
model.

compared to the pure supercell containing a dipole as in Fig. 4.3 — assuming both
models are of the same volume and neglecting surface effects. In other words, in
the hybrid model it is by far easier to give a good representation of the surrounding
bulk crystal. The latter in combination with the capability of modelling a single and
isolated dislocation (avoiding dislocation–dislocation interaction) makes the hybrid
model the ideal choice to describe line defects[4].

---

[4] The hybrid approach — just as the pure cluster approach — is however not suitable for DFT
methods based on plane wave basis sets, since these methods have considerable problems dealing with
the large amount of vacuum surrounding the model: The then necessary high energy-cutoff makes the
cluster and hybrid models computationally very demanding. Thus whenever plane wave methods are
applied to describe dislocations, the pure supercell is still a very popular model, regardless of all its

Table 4.2: The calculated intrinsic stacking fault energy γ in diamond and selected experimental data.

	Exp.[a]	Exp.[b]	Exp.[c]	DFTB[d]
γ (mJ/m^2)	285 ± 40	279 ± 41	295 ± 30	**293**

[a]Data from weak-beam electron microscopy [90] .

[b]Estimated from dissociated glide dislocations in weak-beam electron microscopy [69].

[c]Estimated from extended nodes in weak-beam electron microscopy [69].

[d]Calculated with the DFTB method in a supercell-cluster hybrid.

The supercell-cluster hybrids used in this work are usually of double-period length having an approximate radius of three or more lattice constants. As we will see in the next section, this is sufficient to describe the energetics and structure of the dislocations examined here. To check if these models are further capable of describing the intrinsic stacking fault that accompanies partial dislocations, we calculate its energy in a supercell-cluster hybrid. Here the stacking fault energy is given as the difference in total energy between a supercell-cluster hybrid containing the stacking fault and a bulk-like hybrid with the same number of C and H atoms. The two structures are geometrically optimised using a conjugate gradient algorithm until all forces are well below 5×10^{-3} eV/Å. As shown in Table 4.2, the method gives results well within the experimental errors.

4.2.2 The elastic energy as a size-convergence criterion

Whenever modelling defects in semiconductors on an atomistic level, one has to test whether the chosen model geometry is suitable and if its size — or in other words the number of atoms considered — is sufficient to describe the defect as embedded in an otherwise perfect crystal lattice. In particular for dislocations, which are extended structural defects of the lattice with a considerable strain field, a sufficient model size is crucial if the surface of the model is relaxed freely[5]. This convergence in size can be checked in comparison with elasticity theory. As we know, the elastic energy contained in a cylinder centred on the dislocation core shows logarithmic behaviour in the continuum limit. Per unit length we obtain (see Section 2.3.1, Eq. (2.4)):

$$\frac{E(R)}{L} = \frac{k(\beta)|\boldsymbol{b}|^2}{4\pi} \ln\left(\frac{R}{R_c}\right) + \frac{E_c}{L}, \quad R \geq R_c \tag{4.1}$$

Unfortunately we cannot directly compare this expression with the total energy obtained by DFT-based calculations, since this total energy contains all contributions from the interaction of the atomic nuclei and electrons. Elasticity theory, however, only gives the difference of energy with unstrained bulk material — the elastic energy. Hence the DFT total energy can only be compared with the elastic energy if the total energy of the same group of atoms in perfect material is subtracted. The DFTB method allows an easy determination of the total energy projected onto a single

---

deficiencies mentioned above.

  [5] As mentioned earlier, a different approach is to embed the model in a bulk crystal by imposing the displacements expected from continuum elasticity theory at the model's surface.
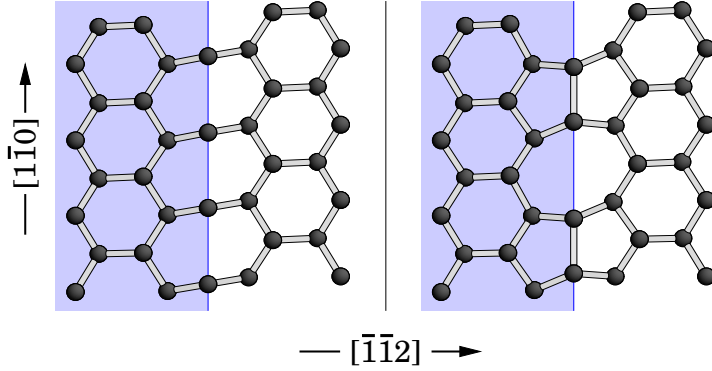
Figure 4.5: The core reconstruction of the 30° glide partial projected into the (111) glide plane. The stacking fault area is shaded. *Left:* The unreconstructed core structure. *Right:* The reconstructed core structure. The reconstruction bonds are about 18 % stretched compared to bulk diamond.

atom i. The formation energy[6] projected onto atom i is then defined as the difference between the total energy E^i_{tot} and the total energy of the same type of atom in perfect material E^{bulk}_{tot}:

$$E^i_f = E^i_{tot} - E^{bulk}_{tot} \tag{4.2}$$

To compare now the spatial distribution E^i_f of the formation energy in a dislocated supercell-cluster hybrid with the elastic energy as given in Eq. (4.1), one has to sum over all atoms in a cylinder $\mathcal{C}(R, L)$ of radius R and length L following the corresponding geometry. This procedure yields the radial formation energy per unit length[7]:

$$\frac{E_f(R)}{L} = \frac{1}{L} \sum_{\mathcal{C}(R,L)} E^i_f - \gamma R\ \delta_{ISF} \tag{4.3}$$

The second term on the right hand side is the stacking fault energy, which has to be subtracted for models containing a Shockley partial dislocation. For such models δ_{ISF} is defined to be 1 and 0 else.

The 30° glide partial in diamond will now serve as an example to illustrate the consistence between elastic energy and the DFT-based radial formation energy. In principle, the low energy core structure of the 30° glide partial in diamond is similar to that shown in Fig. 3.7. However, the unreconstructed single-period core structure is only metastable. As can be seen in Fig. 4.5, the three-fold coordinated core atoms

---

[6] If we consider a compound semiconductor crystal with two or more atomic species, then the chemical potentials of each species might have to be taken into account, depending on the conditions under which the crystal was grown.

[7] In principle, all DFT-based methods allow the definition of a spatially resolved density of total energy $\varepsilon_{tot}(\mathbf{r})$ This is obvious for the electronic energy. But also the long range terms (Ewald sum of the charges) can in principle be divided onto the respective charge positions (δ-function-like for the nuclei [91]). Starting from this, we can define a formation energy density by subtracting the total energy density of the respective bulk material:

$$\varepsilon_f(\mathbf{r}) = \varepsilon_{tot}(\mathbf{r}) - \varepsilon^{bulk}_{tot}$$

The sum in Eq. (4.3) then becomes an integral. Alternatively, it is also possible to project this energy density onto atom positions using Voronoi polyhedra [92].
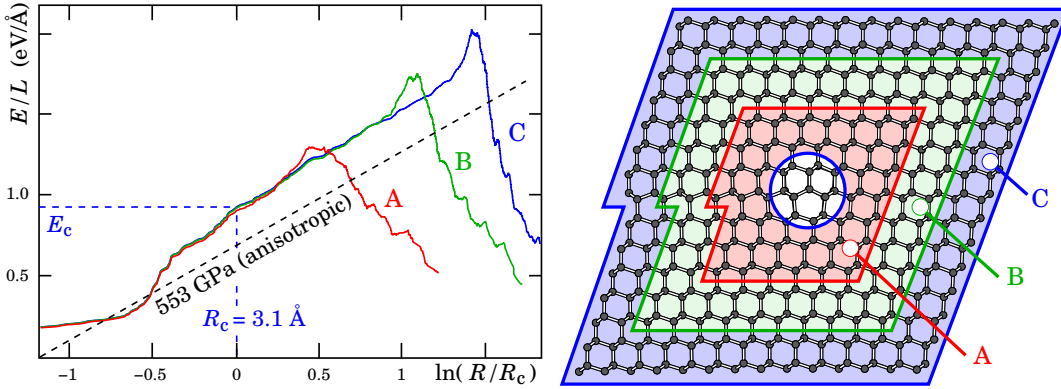
Figure 4.6: The radial formation energy of the 30° glide partial for hybrid models of different size. *Left:* $E_f(R)/L$ as obtained with Eq. (4.3) for hybrid model **A** (182 carbon atoms), **B** (462 carbon atoms) and **C** (870 carbon atoms). The gradient of the dashed line corresponds to the elastic energy in Eq. (4.1) with $k(\beta) = 553$ GPa, a result obtained in *anisotropic* elasticity theory. The core radius R_c and energy E_c can be obtained directly from the graph. See text for a detailed description. *Right:* Section through the three corresponding hybrid models $((1\bar{1}0)$ plane). The hydrogen termination at the surface is not shown. The inner white area marks the core region ($R_c = 3.1$ Å). The subsequently surrounding areas give the dimensions of the three respective models.

can form bonded pairs along the dislocation line, leaving each core atom now four-fold coordinated. This reconstructed core is found approximately 0.6 eV per formed atom-pair lower in energy than the unreconstructed core.

Just as the stacking fault energy in the last section, $E_{\text{tot}}^{\text{bulk}}$ is determined in a bulk-like model of the same shape and approximate size as the dislocated model. Fig. 4.6 shows $E_f(R)/L$ for hybrid models of three different sizes (**A** 182, **B** 462 and **C** 870 carbon atoms) plotted against $\ln(R)$. All three models are of double-period length along the dislocation line ([$1\bar{1}0$] direction). All three graphs show approximately linear behaviour from a core radius $R_c = 3.1$ Å onwards. The corresponding core energy is $E_c/L = 0.88$ eV/Å. Of course the core radius is only vaguely defined and it is somewhat arbitrary what one calls "approximately linear". The gradient of the linear part should be given as the pre-logarithmic factor in Eq. (4.1). Isotropic elasticity theory yields $k(\beta)=567$ GPa and anisotropic theory as described in Ref. [93] and [41] gives $k(\beta)=553$ GPa using the DFTB elastic constants from Table 4.1. The corresponding gradient is drawn in Fig. 4.6 and the correspondence seems to be reasonably good. The oscillations in all curves are an effect of the discrete sum and they should disappear in the limit $R \to \infty$.

From a certain radius onwards, the energy in each of the graphs increases rather rapidly and non-linearly. This can be explained by the radius getting too close to the surface of the respective model, where the strained bulk lattice is not represented sufficiently since surface effects dominate. For even larger radii the integration cylinder is not completely contained in the model and strain contributions from the model's edges lead to a sudden break-down of the formation energy. Hence one can directly conclude from Fig. 4.6, that the maximum radius where the models give a good representation of the surrounding strained bulk lattice — at least in terms of elasticity — is about 4.2, 8.4 and 12.4 Å for model **A**, **B** and **C** respectively. In other

words, in principle the smallest model with 182 carbon atoms is able to describe the dislocation as embedded in an otherwise perfect crystal. However, from curve **A** alone this would be hard to see: The nearly correct gradient between R_c and 4.2 Å could be a random coincidence. As a consequence and to be on the safe side for dislocations with larger R_c, for all further calculations supercell-cluster hybrid models similar in size to **B** (462 carbon atoms) will be used.

Some remarks on the effects of the k-point sampling: The structural calculations were performed at the Γ-point only, but models with double-period length and a $1 \times 1 \times 2$ or a $1 \times 1 \times 4$ k-point scheme gave similar energies and structures. We found less than 4 % difference in bond length for reconstruction bonds, less than 1 % for bulk-like bonds and up to 8 % difference in formation energy depending on the k-point sampling. Especially the latter may seem considerably large, however for dislocations of the same type but with a different core structure, the deviation seems to be an approximately constant shift. Also, compared to the large differences in core energies for the different dislocations modelled in this work, the k-point scheme related differences are negligible. Generally however, k-point sampling is an important issue and its effects always have to be tested.

4.3   Core structures and energies

Hitherto the 30° glide partial served as an example to illustrate the procedure used in this work to model the core structure and energy of dislocations. This procedure will now be applied to find the low energy core structures of the predominant perfect and partial dislocations of the $\{111\} \langle 110 \rangle$ slip system in diamond. Therefore the structures and their corresponding core energies and the long range elastic energy contributions are calculated.

4.3.1   Core structures

The minimum Burgers vectors b and line directions ℓ for the $\{111\} \langle 110 \rangle$ slip system in cubic material have been given in Chapter 3. With those known, one can construct atomistic models of the corresponding dislocations by displacing the atoms of an appropriate supercell-cluster hybrid according to b and ℓ. If required, an additional half-plane of atoms or a stacking fault has to be introduced. The so obtained dislocation core structures however, are nothing more but a mere first guess. As shown in the last section, it can be energetically favourable if the core reconstructs by forming new bonds and possibly breaking existing ones. Since core reconstructions can be rather complicated, it is not guaranteed to find the overall lowest energy structures by simply relaxing *one* starting structure by means of a conjugate gradient algorithm. The reconstruction found then might just be one of many local minima of the corresponding energy-surface. Thus, in this work usually several different starting structures for a given combination of b and ℓ are structurally optimised. The low energy core structures obtained are shown in Figs. 4.7 and 4.8. In the following, each structure will be described and discussed briefly.
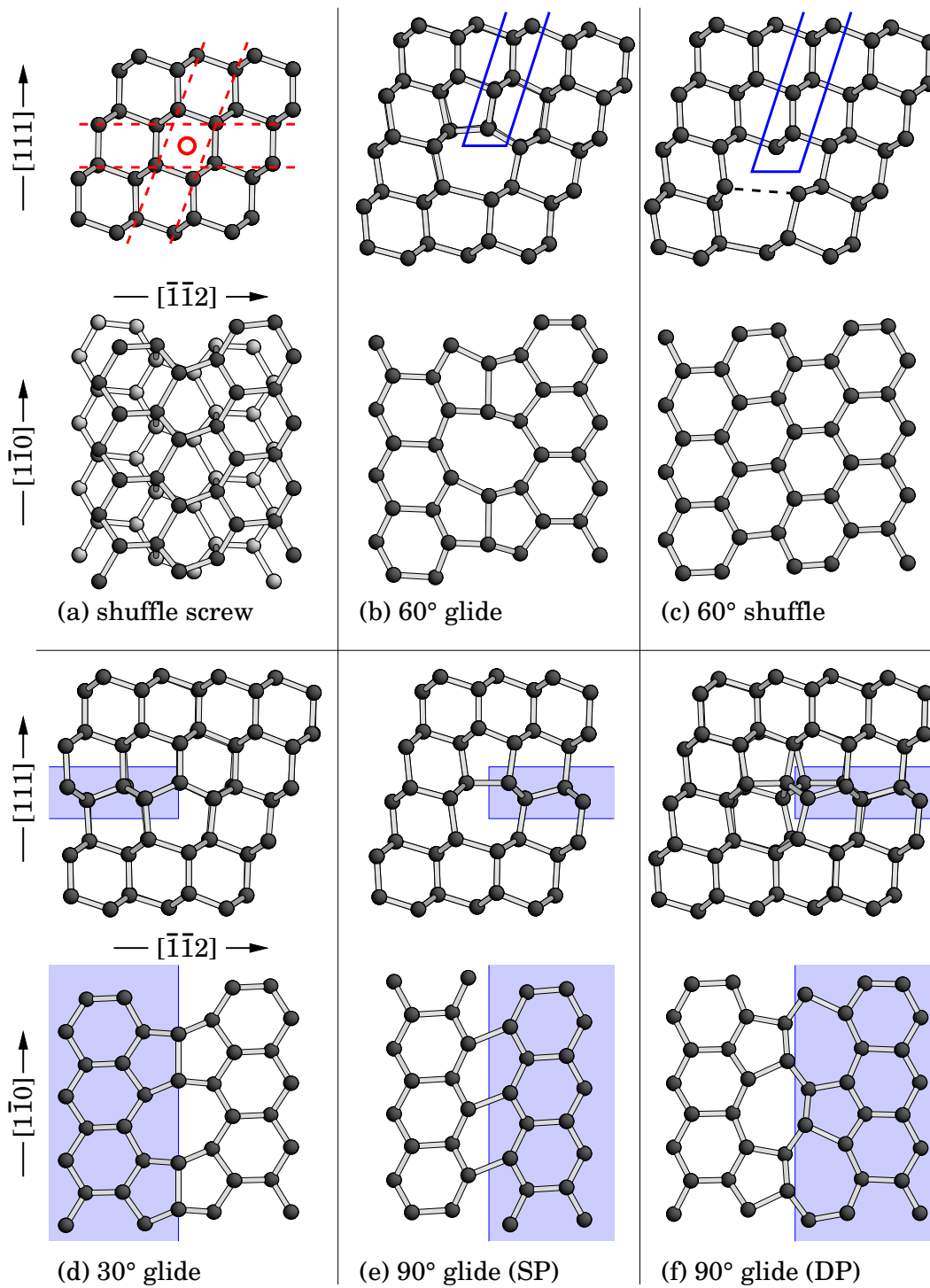
Figure 4.7: Dislocation core structures of the $\{111\}\,\langle 110\rangle$ slip system in diamond. The relaxed core structures of (a) the $[1\bar{1}0]$ shuffle screw dislocation, (b),(c) the undissociated 60° dislocation and (d) – (f) the glide set of Shockley partials are shown. For each structure, the upper figure shows the view along the dislocation line projected into the $(1\bar{1}0)$ plane, and the lower figure shows the (111) glide plane. In (a) two (111) planes are shown, one in dark and one in light grey. The region of the intrinsic stacking fault accompanying the partials is shaded in the respective structures. In case of the 90° partial SP and DP denote the single-period and double-period core reconstruction.
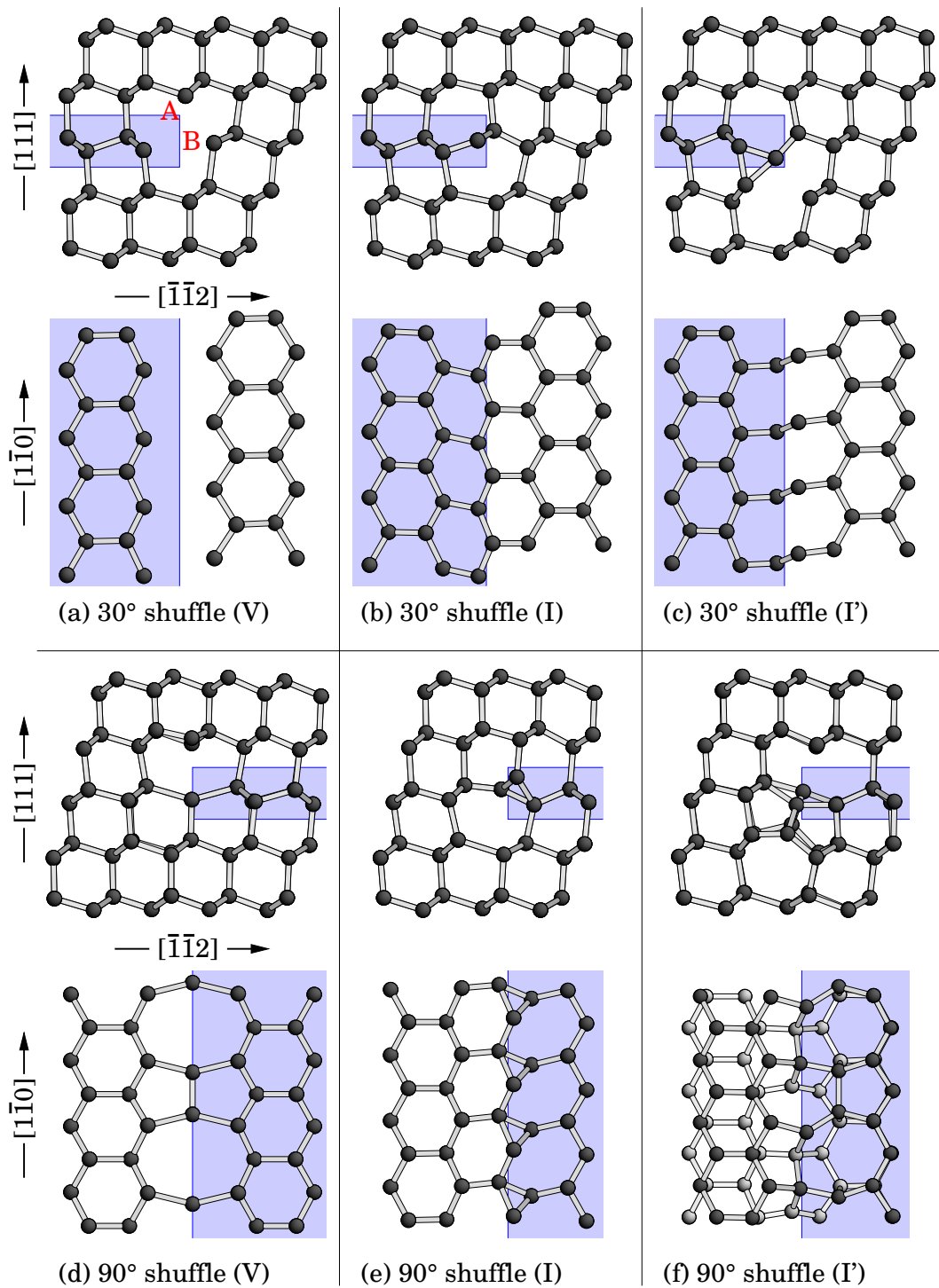
Figure 4.8: Dislocation core structures of the {111} ⟨110⟩ slip system in diamond. The relaxed core structures of (a) – (c) the 30° shuffle partial and (d) – (f) the 90° shuffle partial are shown. For each structure, the upper figure shows the view along the dislocation line projected into the $(1\bar{1}0)$ plane, and the lower figure shows the (111) glide plane. For both the 30° and 90° partial several stable core structures are possible. We distinguish between the vacancy structure (V), the interstitial structure (I) and an alternative interstitial structure (I'). The region of the intrinsic stacking fault accompanying the partials is shaded. In (f) the glide plane atoms are depicted in dark grey, and the underlying plane in light grey.

The $b = \frac{1}{2}[1\bar{1}0]$ screw dislocation (Fig. 4.7 (a)) : This dislocation shows the structure one would expect from locally shearing the crystal between two {111} glide planes. Some of the core bonds approximately parallel to the dislocation line are considerably weakened giving 17 % elongation due to the high stress and large core displacements.

For the structure shown in Fig. 4.7 (a) the core line (drawn as a small circle in the $[1\bar{1}0]$ projection) lies between two sets of two widely spaced planes (drawn as dashed lines)[8]. Consequently the structure given here can be considered a shuffle structure. However the pure screw is a special case, since it has no featured glide plane. Dislocation motion from shuffle to glide and vice versa preserves the number of lattice sites — no interstitial or vacancy emission or trapping is necessary. Also mixed shuffle-glide cores are possible. Hence the structure given here is just an example and more details on the screw dislocation will be given in Section 4.3.3.

The 60° dislocation (Fig. 4.7 (b,c)) : This dislocation is obtained by inserting (or removing) a {111} half-plane of atoms. Fig. 4.7 shows the half-plane for the shuffle and the glide cases framed with lines.

The relaxed core structure of the glide dislocation shows bond reconstruction of the terminating atoms of the half-plane with neighbouring atoms. Reconstruction bonds along the dislocation line impose a double-period structure.

The shuffle dislocation does not reconstruct. The terminating atoms possess dangling bonds normal to the glide plane which give rise to electronic gap states, as will be shown later. Further the local stress "below" the dislocation line stretches one bond extremely by 80 % (dashed line), hence breaking it.

The glide set of Shockley partials (Fig. 4.7 (d–f)) : In the glide set of partial dislocations each dislocation is accompanied by a stacking fault.

The *30° glide partial* reconstructs forming a line of bonded atom pairs. The reconstruction bonds are 18 % stretched compared to bulk diamond. The structure is double-periodic and has been described in the last section already.

For the *90° glide partial* there are two principal structures: A single-period (SP) and a double-period (DP) reconstruction. The SP structure results simply from forming bonds in the glide plane connecting the stacking fault region with the bulk region. These bonds are 13 % stretched. The DP structure can be obtained from the SP structure by introducing alternating kinks. Here the deviations from the bulk bond length is slightly less. The DP structure was first proposed for silicon by Bennetto et al. [94]. All atoms in both the SP and the DP structure are fully four-fold coordinated. This is in contrast to the ethene-like cores proposed recently [83]. Since the latter structures are higher in energy, here they are not discussed further[9].

---

[8] In the case of a pure screw dislocation no inserted half-plane is present. However comparing the lattice displacement associated with a screw as given by elasticity theory with the atom coordinates of the relaxed atomistic model, one can still determine the localisation of the core line.

[9] However, they could still exist at raised temperatures and then might assist graphite formation.

The shuffle set of 30° partials (Fig. 4.8 (a–c)):   To form the shuffle set of 30° partials
one can either add or remove a row of atoms along the dislocation line. In contrast
with the 60° dislocation, adding and removing a line of atoms leads to different
structures as the stacking fault is displaced with respect to the termination of the
inserted half-plane.

Upon the removal of one row of atoms from Fig. 4.7 (d), one obtains a *vacancy struc-
ture* (V) shown in (a). Unlike that found for silicon[10], no *strong* bond reconstruction
occurs at the core but the distance between atoms **A** and **B** is reduced from 2.52 Å to
2.08 Å.

(b) and (c) are distinct interstitial structures formed by the addition of a row of atoms
to the core (I). In (b) the additional row of atoms shows almost planar sp^2 bonding
geometry while (c) shows strong bond angle distortions, but is at least metastable.
The stress induced by adding the row of atoms pushes the surrounding lattice apart
and breaks bonds.

The shuffle set of 90° partials (Fig. 4.8 (d–f)):   The 90° shuffle set is also formed
by climb of the glide partial. In a similar way to the 30° shuffle it possesses both
vacancy and interstitial forms.

The vacancy structure (d) appears to be symmetric in the glide plane and a line of
bonded dimers is formed, leaving the dislocation with a double-period[11]. The bonds
of the dimers to neighbouring atoms in the glide plane, however, are weak and 22 %
stretched, and one row of dangling bonds remains unreconstructed.

We find two different metastable interstitial-like core structures. In the first one (e)
all atoms are fully four-fold coordinated, with appreciable bond length and angle
distortion, while the second (f) contains many dangling bonds. The main elements
in the latter structure are four membered [01$\bar{1}$] zig-zag chains forming a double-
period core.

4.3.2   Core energies

Having all the structures of the last section relaxed to their minimum in total energy,
it is now possible to obtain their core radii and energies via the procedure described
in Section 4.2.2. The results are given in Table 4.3, retaining the sequence of struc-
tures as in Figs. 4.7 and 4.8. The energy factors $k(\beta)$ have been evaluated in a linear
fitting procedure for $\ln(R/R_c) > 0$. For comparison, the energy factors resulting
from linear isotropic as well as anisotropic elasticity theory are also given — both
those based on the elastic constants as obtained by the DFTB method and those

---

[10] This structure, or a small segment of it, has been suggested as a candidate for the K1/K2 EPR
signal detected in plastically deformed Si [95]. More recently in Si, the single vacancy and a row of
vacancies at the 30° partial have been considered theoretically by Lehto and Öberg [96] and by Justo
et al. [97]. These authors found a strong rebonding at least for the single vacancy which enabled all
dangling bonds to be eliminated.
[11] An interesting aspect of this structure is its symmetry: If the corresponding glide structure
(Fig. 4.7 (e)) moves as a whole via applied stress, then the dislocation will proceed via a saddle point
with the same approximate symmetry as the vacancy structure (d) itself. Thus, the vacancies seem to
have pinned the dislocation into what is normally a saddle point structure.

Table 4.3: The calculated energy factors $k(\beta)$, core radii R_c and core energies E_c of the dislocations of the $\{111\}\langle110\rangle$ slip system in diamond. The corresponding structures can be found in Fig. 4.7 and 4.8 in the same sequence. To facilitate comparison between different dislocations with different core radii, the core energy E'_c corresponding to a radius of 6 Å is introduced. In some cases — mainly the interstitial structures of the 30° and 90° partials — the fluctuations in $E_f(R)$ were rather large, leading to errors of around 20 % in the fitting process. Affected values are marked with an asterisk.

		shuffle screw	60° glide	60° shuffle
$k(\beta)^a$	(GPa)	582	570	591
$k(\beta)$ isotropic^b	(GPa)	554 (536)	593 (566)	593 (566)
$k(\beta)$ anisotropic^c	(GPa)	536 (524)	587 (561)	587 (561)
R_c	(Å)	4.3	4.1	4.2
E_c/L	(eV/Å)	3.32	2.13	2.60
$E'_c(R = 6 \text{ Å})/L$	(eV/Å)	**3.94**	**2.55**	**3.18**

		30° glide	90° glide (SP)	90° glide (DP)
$k(\beta)^a$	(GPa)	573	601	599
$k(\beta)$ isotropic^b	(GPa)	567 (546)	607 (576)	607 (576)
$k(\beta)$ anisotropic^c	(GPa)	553 (536)	603 (574)	603 (574)
R_c	(Å)	3.1	3.0	3.3
E_c/L	(eV/Å)	0.88	1.32	1.18
$E'_c(R = 6 \text{ Å})/L$	(eV/Å)	**1.23**	**1.74**	**1.57**

		30° shuffle (V)	30° shuffle (I)	30° shuffle (I′)
$k(\beta)^a$	(GPa)	637	605	762*
$k(\beta)$ isotropic^b	(GPa)	567 (546)	567 (546)	567 (546)
$k(\beta)$ anisotropic^c	(GPa)	553 (536)	553 (536)	553 (536)
R_c	(Å)	6.0	5.0	5.5
E_c/L	(eV/Å)	2.55	3.31	4.32
$E'_c(R = 6 \text{ Å})/L$	(eV/Å)	**2.55**	**3.43**	**4.39**

		90° shuffle (V)	90° shuffle (I)	90° shuffle (I′)
$k(\beta)^a$	(GPa)	642*	602*	640*
$k(\beta)$ isotropic^b	(GPa)	607 (576)	607 (576)	607 (576)
$k(\beta)$ anisotropic^c	(GPa)	603 (574)	603 (574)	603 (574)
R_c	(Å)	4.4*	4.1*	4.5*
E_c/L	(eV/Å)	1.89	3.69*	2.67*
$E'_c(R = 6 \text{ Å})/L$	(eV/Å)	**2.09**	**3.94***	**2.86***

^{a}Fit to $E_f(R)$ vs. $\ln(R/R_c)$ plot for $R \geq R_c$ following Eq. (4.1).

^{b}Evaluated using Eq. (2.5). The first number is given by the elastic constants obtained in the DFTB method, the number in brackets by the experimental constants (for both see Table 4.1).

^{c}Evaluated using linear *anisotropic* elasticity theory as suggested in [93] and [41].

based on the experimental constants. Some of the interstitial-type shuffle partial dislocations show a rather uneven linear part leading to larger errors in the determination of the energy factors, the core radii and sometimes even the core energies. These are marked with an asterisk. Models of larger diameter might help to increase the accuracy.

As mentioned earlier, the determination of the core radii is a bit arbitrary. Hence one should not draw strong conclusions from R_c. However, some general trends are obvious: All the undissociated structures have similar core radii of about 4 Å. Especially in the case of the Shockley partials, the shuffle set appears to have considerably larger core radii compared to the glide set. This should be kept in mind since the reliability of the calculations is lower the larger the core radius gets.

A direct comparison of core energies only makes sense between dislocations of the same angle β between Burgers vector and line direction. Further, since the core radius R_c usually varies with different dislocations, one should not compare $E_c = E(R_c)$ but $E_c' = E(R')$ for an arbitrary radius $R' \geq R_c$ common to all dislocations. Here the choice is $R' = 6$ Å.

Since the screw dislocation will be discussed later separately, this section focuses on the $60°$ dislocations and their dissociation products, the $30°$ and $90°$ partials. The next paragraphs give the general trends.

Glide and shuffle dislocations: As a first result the undissociated $60°$ dislocation appears to be most stable in the glide structure: The shuffle structure is ~ 630 meV/Å higher in energy. Due to the variety of partial shuffle structures, one cannot give a characteristic energy difference, however here the lowest energy structures are also always of the glide type[12].

Single- and double-period reconstructions: One of the few problems concerning dislocation core structures in diamond which has been treated from a theoretical point of view before is the relative stability of the SP and DP structures of the $90°$ glide partial: Nunes et al. [78] find the DP core to be 235 meV/Å lower in energy in DFT calculations. Similarly, the low-stress quadrupole calculations of Blase et al. [79] yield 169 – 198 meV/Å. This is all in good agreement with the ~ 170 meV/Å difference found in this work by means of the DFTB method and supercell-cluster hybrid models[13]. This difference in energy between the SP and the DP structure is rather small, suggesting that possibly both structures co-exist.

Vacancy and interstitial shuffle structures: Within the shuffle set of $30°$ partials as well as within the shuffle set of $90°$ partials, the vacancy structures are clearly

---

[12] The core energy of the $60°$ perfect glide dislocation is slightly lower than that of a pair of $90°$and $30°$glide partial dislocations by 0.25 eV/Å suggesting that dissociation would be prohibited. However, the contribution of the long range stress field, to be discussed in Section 4.4, reverses this.

[13] Comparing the different results obtained by different methods and models, one can discover a trend here: The smaller the stress induced by the dislocation–dislocation interaction within the particular model geometry, the smaller is the energy difference between SP and DP. In Ref. [79] the stress is already minimised by optimising a quadrupole in a supercell, giving a smaller difference than found in [78]. And finally, supercell-cluster hybrid models with no dislocation–dislocation interaction give the smallest difference (this work). However, one should keep in mind, that the approximations of DFT used in the three approaches are rather different, so this trend could still be an artificial effect.

Table 4.4: A comparison of the DFTB results for the core radius and energy of the 90° partial dislocation (SP) with independent classical potential and LDA-pseudopotential calculations. r_0 is the radius for which in Eq. (4.1) E_c becomes zero. The core energy for LDAPP had to be extrapolated to $R_c = 3$ Å to facilitate comparison.

		TP[a]	LDAPP[b]	DFTB[c]
R_c	(Å)	3.0	—	**3.0**
r_0	(Å)	0.004	0.4	**0.38**
$E_c'(R_c = 3\text{ Å})$	(eV/Å)	4.69	1.26	**1.32**

[a]Work by Nandedkar and Narayan [76] using Tersoff potentials in a cylindrical model.

[b]Work by Blase et al. [79] using LDA-pseudopotentials.

[c]Calculated with the DFTB method (see Table 4.3).

favoured by approximately 900 and 800 meV/Å respectively. This is a direct result of the high strain induced by the additional row of atoms in the case of the interstitial-type partials.

Comparison with independent calculations: For a comparison of the core radii and core energies with independent calculations we are restricted to the single-period reconstruction of the 90° partial, which has been discussed in the literature. Results from different methods are shown in Table 4.4: Tersoff potential (TP)[76], LDA-pseudopotentials (LDAPP)[79] and DFTB. In Ref. [79] (LDAPP) a different definition of the core radius is applied — it is taken as the radius r_0 where in Eq. (4.1) E_c becomes zero. For details see also Appendix A. Since R_c as used in this work cannot be determined from their published data, Table 4.4 gives the extrapolated alternative core radius r_0 for the TP and DFTB results instead. Further, $E_c'(R_c)$ had to be extrapolated for LDAPP. As can be seen, the core radius and energy obtained in DFTB are in very good agreement with the LDAPP results[14].

The core radius R_c of the TP calculation is exactly the same as found in this work. For the corresponding core energy however, Tersoff potentials yield an extremely large value, exceeding DFTB and LDAPP by more than 300 %. This reflects the difficulties in constructing an atomic potential to describe the energy of a heavily distorted dislocation core.

---

[14] However, the LDAPP calculations in Ref. [79] are computationally by far more expensive. And also, being *plane wave* pseudopotential calculations, the used model was a pure supercell. Therefore the effort to minimise and subtract the artificial dislocation–dislocation interaction inherent to the super-cell was apparently rather large. This demonstrates that the appropriate choice of a good combination of method and model is often more important than using the most exact approximation for the energy.

4.3.3   The screw dislocation — a special case

Hitherto the detailed discussion of the pure screw dislocation has been avoided and only the very simple structure of the shuffle screw was briefly introduced. The pure screw is a special case, since it has no featured glide plane and there is no such thing as dislocation climb. But still, depending on the location of the dislocation core line — between widely or between closely separated {111} planes — one can distinguish between "shuffle" and "glide" structures in a similar way as in the case of the perfect 60° dislocation. This core line is defined by the radial displacement field being centred at it. When constructing a pure screw dislocation, first this origin has to be chosen, and then the surrounding atoms are displaced with respect to their positions in a perfect undislocated lattice[15].

For a $[1\bar{1}0]$ screw dislocation there are two sets of {111} planes containing its line direction. Hence in terms of shuffle and glide, there are three unique possible types as illustrated in Fig. 4.9 (Left): Shuffle, glide and mixed. The first two have shuffle or glide character respectively with respect to both the horizontal and the inclined set of {111} planes, whereas the mixed type has shuffle character with respect to one and glide character with respect to the other set of {111} planes.

Core structures: The relaxed structure of the shuffle screw dislocation can be found in Fig. 4.7 (a). The bond angles and lengths near the core are stretched but all bonds are bulk-like and all atoms remain four-fold coordinated. This is a result of the origin of the displacement field not intersecting with any bonds. This is very different for the glide screw and for the mixed screw, where the origin lies exactly on a row of bonds (see Fig. 4.9 (Left)). As a result, the two neighbouring atoms connected by one of those bonds are displaced with respect to each other by half the

---

[15] In cylindrical coordinates with $\boldsymbol{r} = (r, \vartheta, z)$ centred on the dislocation line, the displacement field is given as $\boldsymbol{u}(\boldsymbol{r}) = (0, 0, |\boldsymbol{b}_\mathrm{s}|\vartheta/(2\pi))$. Varying ϑ from 0 to 2π in a Burgers circuit then yields the correct Burgers vector, in cylindrical coordinates $\boldsymbol{b}_\mathrm{s} = (0, 0, |\boldsymbol{b}_\mathrm{s}|)$.
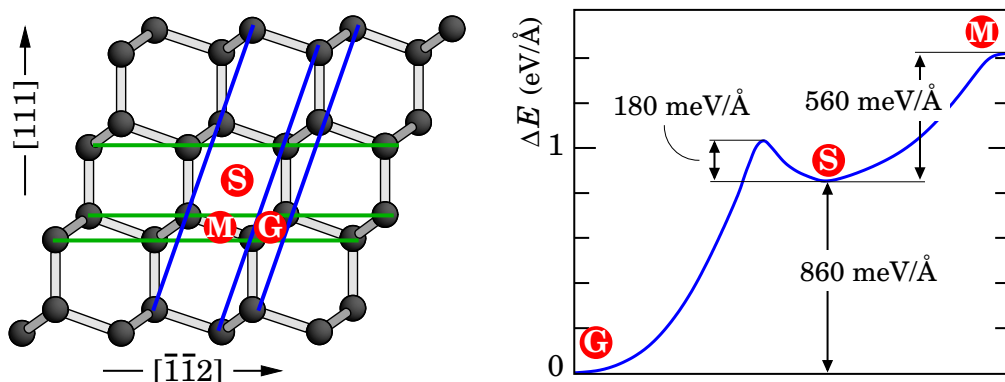
Figure 4.9: The three unique types of the $\frac{1}{2}[1\bar{1}0]$ screw dislocation: Shuffle, glide and mixed (labelled **S**, **G** and **M** respectively). *Left:* The core positions projected into the $(1\bar{1}0)$ plane. The two sets of {111} planes containing the line direction are indicated. *Right:* Relative core energies of the three different types and barriers between them. The *x*-axis gives an arbitrary coordinate for the transformation **G** → **S** → **M**.

Figure 4.10: The $\frac{1}{2}[1\bar{1}0]$ glide screw dislocation (**G** in Fig. 4.9). Atoms **A** and **B** (and equivalent atoms along $[1\bar{1}0]$) are only three-fold coordinated. *Left:* The view projected into the $(1\bar{1}0)$ plane. *Right:* The view projected into the (111) plane.
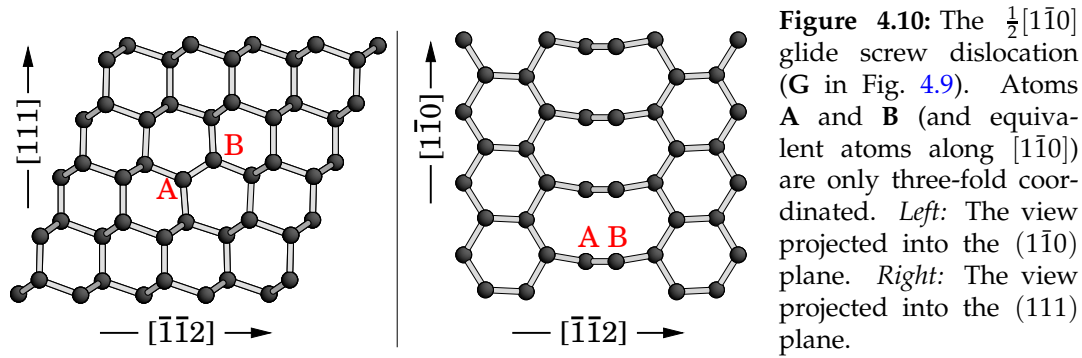
Burgers vector along $[1\bar{1}0]$.

In case of the glide screw this leads to a planar sp^2 bonding situation for both atoms. Subsequent structural optimisation pushes the two atoms apart, resulting in an appreciable C–C bond length. Fig. 4.10 shows the relaxed structure, which is basically that of two 30° partial glide dislocations in the same Peierls valley. However there is no bond reconstruction along the dislocation line.

For the mixed-type screw dislocation the two atoms closest to the origin of the displacement are widely separated in the $[1\bar{1}0]$ projection. Thus, after displacement their distance is large enough to break the connecting bond. Subsequent relaxation does *not* lead to a bond reconstruction. The mixed structure was suggested by Koizumi et al. [98], and for III-V compound semiconductors the same authors claim it to be more stable than the shuffle structure.

Energetics and stability: To model the transformation from $\mathbf{G} \rightarrow \mathbf{S} \rightarrow \mathbf{M}$, the atoms of a bulk hybrid model were displaced according to the process described earlier in this section. Varying the origin of the displacement field between **G** and **S** or **S** and **M** allows one to scan for barriers between start and final position. The structure is then relaxed for each point of the scan grid[16].

Fig. 4.9 (Right) shows the relative core energies of the glide, the shuffle and the mixed type. The glide screw appears to be the lowest energy structure. The shuffle screw has a $\sim$ 860 meV/Å higher core energy, is however at least metastable with a barrier of $\sim$ 180 meV/Å. Finally the mixed type exceeds the shuffle type by $\sim$ 560 meV/Å. The latter structure is unstable and represents a saddle point between two adjacent sites of the shuffle core.

---

[16] To prevent the origin of the displacement field moving towards a stable or metastable position during relaxation, the two atoms adjacent to **G** or respectively **M** are constrained parallel to the $(1\bar{1}0)$ plane. The curve in Fig. 4.9 (Right) represents about 20 grid points for $\mathbf{G} \rightarrow \mathbf{S}$ and 10 grid points for $\mathbf{S} \rightarrow \mathbf{M}$. Details on how to model barriers for dislocation motion will be given in Section 4.5.

4.4   The dissociation of dislocations in diamond

In Section 3.2 the two basic dissociation reactions of the perfect $\frac{1}{2}[1\bar{1}0]$ screw and the 60° dislocation have been introduced. Furthermore, shuffle and glide set dislocations with an edge component can be transformed into each other by adding (or removing) one row of atoms along the dislocation line (climb). A perfect glide dislocation can always dissociate into two glide partials. Hence if a perfect shuffle dislocation dissociates, the extra row of atoms involved can go with either partial. Therefore, the following *conservative* (not involving net absorption or emission of vacancies or interstitials) dissociation reactions are allowed:

$\frac{1}{2}[1\bar{1}0]$ screw $\longrightarrow$ 30° glide   +   ISF   +   30° glide

$\frac{1}{2}[1\bar{1}0]$ screw $\longrightarrow$ 30° shuffle +   ISF   +   30° shuffle

60° glide $\longrightarrow$ 30° glide   +   ISF   +   90° glide

60° glide $\longrightarrow$ 30° shuffle +   ISF   +   90° shuffle

60° shuffle $\longrightarrow$ 30° glide   +   ISF   +   90° shuffle

60° shuffle $\longrightarrow$ 30° shuffle +   ISF   +   90° glide
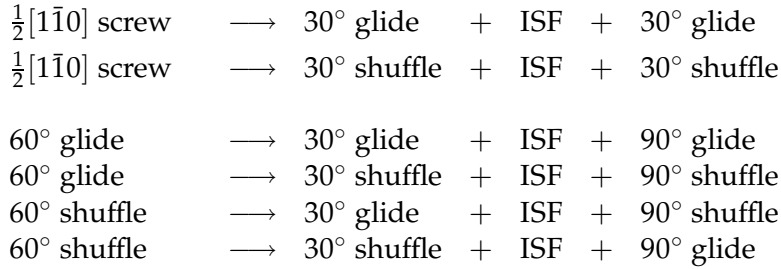
With the core energies of the participating partial dislocations at hand (Table 4.3), one can distinguish which combinations of partial dislocations are energetically favoured and which are not. Based on this it is possible to rule out certain dissociation reactions. The following reactions are energetically favoured[17]:

$\frac{1}{2}[1\bar{1}0]$ screw $\longrightarrow$ 30° glide   +   ISF   +   30° glide

60° glide $\longrightarrow$ 30° glide   +   ISF   +   90° glide (DP/SP)

60° shuffle $\longrightarrow$ 30° glide   +   ISF   +   90° shuffle (V)
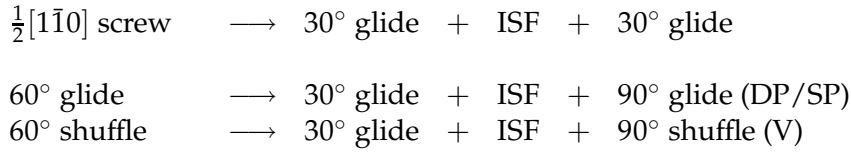
At this point it still remains unclear *if* dissociation occurs or not. For all reactions the Frank criterion (Eq. 2.9) holds. This criterion assumes a dissociation with almost infinite separation, so that contributions from dislocation–dislocation interaction can be neglected. The dissociation into partials, however, involves the creation of an intrinsic stacking fault between the two partials. Hence an almost infinite separation is prohibited by the then almost infinite stacking fault energy. At a finite partial–partial separation, however, *all* energy contributions have to be taken into account, including the partial–partial interaction. As a consequence, the Frank criterion is not sufficient to decide in favour of dissociation into partials.

In the next two sections this problem will be addressed by elasticity theory and atomistic modelling.

---

[17] The dissociation of the screw into two 30° shuffle partials is 2.6 eV/Å higher in energy than the dissociation into two glide partials (twice the energy difference between 30° shuffle (V) and 30° glide). Similarly, the 60° glide dislocation will dissociate into two glide partials. Comparing the sum of core energies of a 30° glide and the lowest energy 90° shuffle with that of the lowest energy 30° shuffle and 90° glide favours the first by $\sim$ 800 meV/Å. Here elasticity theory yields at least qualitatively a similar result: Assuming the vacancy (both lowest energy shuffle partials are vacancy structures) to be a centre of isotropic contraction or expansion, then the vacancy only interacts with the hydrostatic component of the internal stress, which for a straight dislocation is produced by its edge component only [41]. Thus, a vacancy will be preferentially drawn to the partial with the Burgers vector with the largest edge component — in this case the 90° partial dislocation.
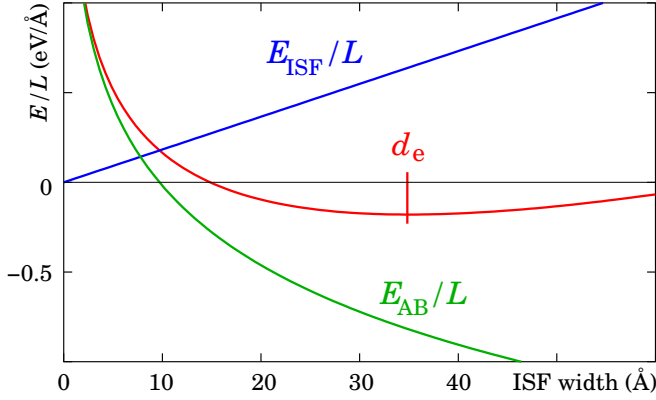
Figure 4.11: The competing energy contributions for dissociation into partials: The stacking fault energy $E_{\mathrm{ISF}}/L = \gamma R$ and the elastic partial–partial interaction $E_{\mathrm{AB}}/L \propto \ln(R/R_0)$ as given in Eq. (2.6) for a 90° and a 30° partial. The sum of both has a minimum at the equilibrium partial separation d_{e} as given in Eq. (4.6). In elasticity theory E_{AB} is known except for a constant offset.

4.4.1 The equilibrium separation of partials

When it comes to the dissociation of perfect dislocations into Shockley partials, then the two competing energy contributions in the elastic limit are the stacking fault energy and the elastic partial–partial interaction energy. The first grows proportionally with the partial–partial separation R (energy per unit length: $E_{\mathrm{ISF}}/L = \gamma R$) and the interaction energy lessens logarithmically with R (Eq. (2.6)). Fig. 4.11 shows a preliminary energy balance: Unfortunately the offset $\propto \ln(R_0)$ of the elastic interaction energy is unknown and consequently at this stage a complete energy balance is still impossible. In equilibrium however, the force $F_{\mathrm{ISF}}/L = -\gamma$ resulting from the stacking fault and the partial–partial repulsion F_{AB}/L, as given in Eq. (2.7), cancel each other and one easily obtains the equilibrium partial separation:

$$d_{\mathrm{e}} = \frac{\mu}{2\pi\gamma} \left\{ (\boldsymbol{b}_{\mathrm{A}} \cdot \boldsymbol{\ell})\,(\boldsymbol{b}_{\mathrm{B}} \cdot \boldsymbol{\ell}) + \frac{(\boldsymbol{b}_{\mathrm{A}} \times \boldsymbol{\ell}) \cdot (\boldsymbol{b}_{\mathrm{B}} \times \boldsymbol{\ell})}{1-\nu} \right\} \tag{4.4}$$

Thus the dissociation of a $\frac{1}{2}[1\bar{1}0]$ screw or a perfect 60° dislocation leads to the following equilibrium stacking fault widths:

$$d_{\mathrm{e}}(\text{screw}) = \frac{\mu}{8\pi\gamma} \left(3 - \frac{1}{1-\nu} \right) |\boldsymbol{b}_{30}|^2 \tag{4.5}$$

$$d_{\mathrm{e}}(60°) = \frac{\mu}{4\pi\gamma} \frac{1}{1-\nu} |\boldsymbol{b}_{30}|\,|\boldsymbol{b}_{90}| \tag{4.6}$$

Table 4.5: The calculated stacking fault widths for the dissociated $\frac{1}{2}[1\bar{1}0]$ screw and 60° dislocation in diamond given by Eq. (4.5,4.6).

		Experimental data[a]	DFTB[b]
$d_{\mathrm{e}}(\text{screw})$	(Å)	30.5	**30.4**
$d_{\mathrm{e}}(60°)$	(Å)	33.9	**35.0**

[a]Calculated using the experimental elastic constants and the arithmetic mean of the experimental stacking fault energies (see Tables 4.1 and 4.2).

[b]Calculated with the DFTB elastic constants and stacking fault energy (see Tables 4.1 and 4.2).
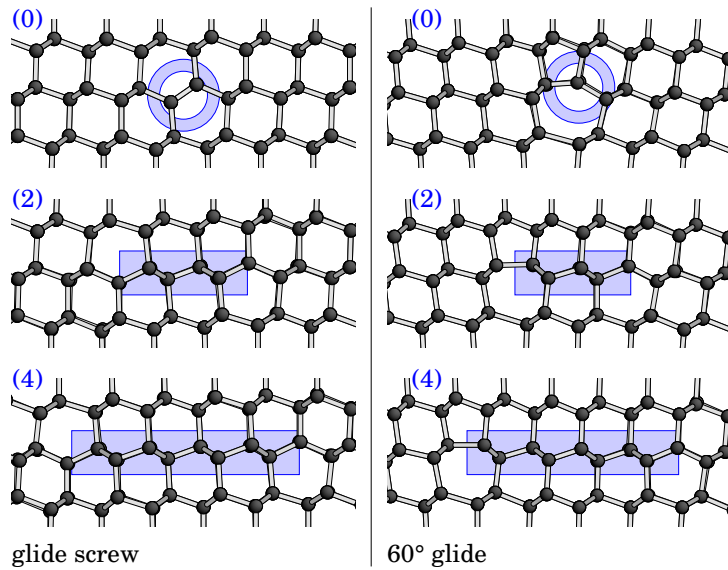
Figure 4.12: The first stages of dissociation of the glide screw and the 60° glide dislocation. The view is projected into the $(1\bar{1}0)$ plane. The top structure labelled **(0)** shows the respective undissociated perfect dislocation. The stacking fault in the second **(2)** and fourth **(4)** dissociation step is shaded. *Left:* The dissociation of the $\frac{1}{2}[1\bar{1}0]$ glide screw into two 30° Shockley partials. *Right:* The dissociation of the 60° glide dislocation into a 90° (SP) and a 30° Shockley partial.

Table 4.5 shows the resulting widths obtained with the DFTB stacking fault energy and elastic constants as well as with the averaged experimental values. The widths vary between 30 and 35 Å. This is an order of magnitude larger than the core radii, giving confidence of being in the elastic limit, where the stacking fault energy and the elastic partial–partial interaction are the dominating contributions. Consistently, dissociation in the $\{111\}\langle 110\rangle$ slip system of diamond has been reported in weak-beam electron microscopy [69]. The dissociation widths vary from 25 to 42 Å. The large variation in widths is a reminder that pinning of dislocations by point defects may be important in preventing the equilibrium stacking fault width being attained. Also, the energy minimum at the equilibrium distance is rather flat, as can be seen in Fig. 4.11.

Even though there is experimental proof for dissociation, from a theory point of view it can still not be conclusively demonstrated whether dissociation is energetically favoured overall, since the offset in the interaction energy is unknown in elasticity theory: The possibility persists that even though d_{e} is a local minimum in energy, the undissociated dislocation is even lower. Hence atomistic calculations are called for to model the first steps of dissociation.

4.4.2 Modelling the first stages of dissociation atomistically

Dissociation of a perfect dislocation into two partials bordering a stacking fault means creating a defect whose core is considerably extended within the glide plane. Thus, describing such a defect requires larger hybrid models than describing just one isolated perfect or partial dislocation. The models used here for the dissociation of the $\frac{1}{2}[1\bar{1}0]$ glide screw and the 60° glide dislocation contain about 600 carbon atoms. Without the core regions of the partials getting too close to the hybrid's surface, this size allows dissociation up to the fourth step — which corresponds to a stacking fault width of 8.75 Å. Fig. 4.12 shows the relaxed geometries for the
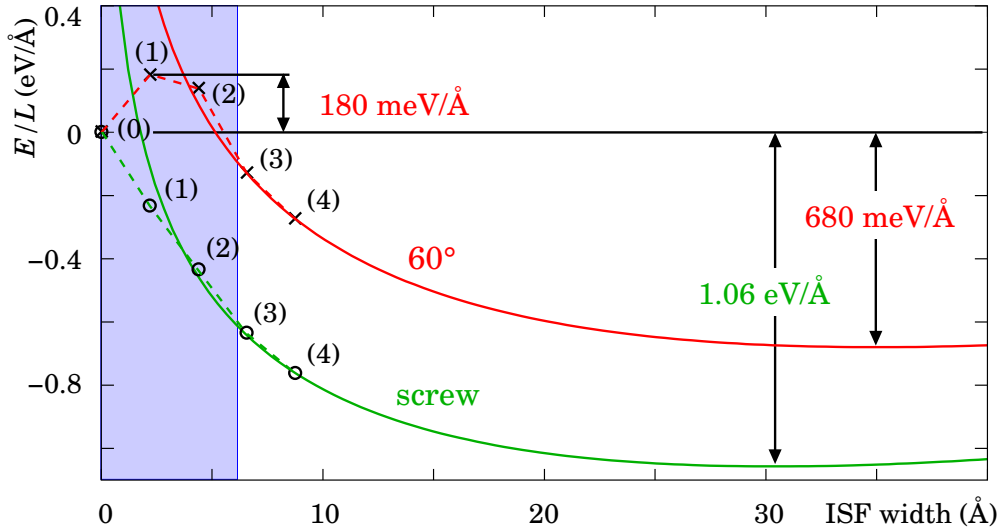
Figure 4.13: The dissociation energy of the glide screw and the 60° glide dislocation. The relative atomistic DFTB energies of the first four dissociation steps and of the undissociated dislocation are labelled **(1)** to **(4)** and **(0)** respectively. Zero energy is set to the undissociated dislocations. The solid lines represent the sum of the stacking fault energy and the elastic interaction energy as given in continuum theory (Eq. (2.6)). Shading below ~6 Å indicates the region where the two core radii of the respective partial dislocations overlap — the region where continuum elasticity theory fails.

respective undissociated dislocation and the second as well as the fourth dissociation step. The corresponding discrete relative energies are presented in Fig. 4.13. Now, with the atomistic results at hand, the unknown energy offset $\propto \ln(R_0/\text{Å})$ for the elastic interaction energy E_{AB}/L in Eq. (2.6) can be found easily by adjusting $E_{AB}/L + E_{ISF}/L$ to the atomistic DFTB energy of the fourth dissociation step. The procedure yields $R_0 = 1.7$ Å and $R_0 = 4.4$ Å for the screw and the 60° dislocation respectively. In Fig. 4.13 the resulting continuous energies are drawn as solid lines for both dislocation types. One can observe that the atomistic calculation and continuum theory still agree for the third stage of dissociation. For smaller stacking fault widths in the region of overlapping dislocation core radii, however, the deviation becomes obvious and finally for $R \to 0$ the continuum result diverges.

Both for the screw and for the 60° dislocation the final dissociation energy at equilibrium separation between the partials is clearly below the energy of the undissociated dislocation (-1.06 eV/Å and -0.680 eV/Å respectively). Hence also from a theory point of view, dissociation into partial dislocations is strongly favoured[18]. A striking difference between the dissociation of the two types of dislocations is the presence of an energy barrier of ~ 180 meV/Å to initiate dissociation of the 60° dislocation, whereas for the screw there is no such barrier. This barrier should have a considerable effect, since it is approximately 5 Å wide and cannot be overcome in a single step. This might explain the presence of undissociated 60° dislocations as observed in weak-beam electron microscopy (compare Fig. 4.1 and Ref. [69]).

---

[18] In the case of the dissociation of the 60° glide dislocation only the single-period 90° partial is considered here. The energy of the respective systems containing a double-period 90° instead would be approximately 170 meV/Å lower.
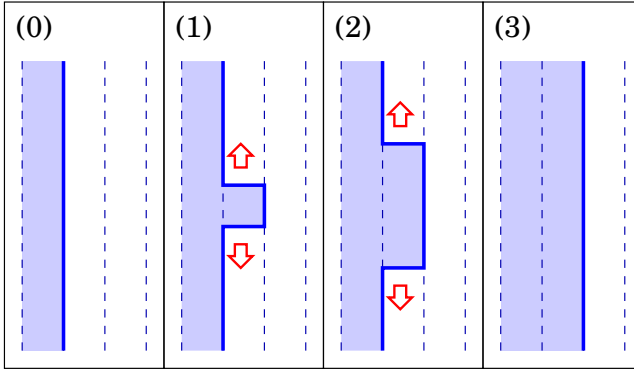
Figure 4.14: Dislocation glide of partials by kink formation and migration. The Peierls valleys are drawn as dashed lines. **(0)** shows the straight dislocation line in one valley, in **(1)** a kink pair has formed, whose single kinks then migrate along the dislocation line **(2)**. Finally, in **(3)**, the shown segment of the dislocation has moved to the next Peierls valley.

4.5   Kinked Shockley partials and dislocation glide

All energies given in the last section in the context of dissociation into partial dislocations are energies in the Peierls valleys only. Thus, each dissociation step requires to overcome the intermediate Peierls hill. Unless at very high temperature or under very high stress, the partial dislocations will not move over such hill as a whole, but piecewise. This motion of the glide partials will be modelled in this section.

4.5.1   Dislocation glide by kink formation and migration

Glide motion arises from stress acting on the dislocation core. This stress can result from external forces on the crystal or, as discussed in the last section, from the interaction with neighbouring dislocations. It leads to the formation of kinks at the dislocation line. When the stress is insufficient to overcome the Peierls barrier, kinks must be generated by a thermal process and motion occurs by their migration along the dislocation line. Formation and migration then controls the dislocation velocity in the absence of pinning centres or other obstacles. Fig. 4.14 schematically shows the process. Since kinks on a dislocation line can only be created in pairs, their density in thermodynamic equilibrium is controlled by the double-kink formation energy $2E_f$, whereas their rate of motion depends on the kink migration barrier W_m. In the case when strong obstacles like point defects and impurities are absent the activation energy for the glide motion of short dislocation segments Q is given as the sum of $2E_f$ and W_m [19]. One obtains for the glide velocity [41]:

$$v_{disl} \propto e^{-Q/(kT)} \qquad \text{with} \qquad Q = 2E_f + W_m \tag{4.7}$$

For long dislocation segments however, the activation energy is controlled by the sum of the single-kink formation energy E_f and the kink migration barrier. The exact expressions and their rather lengthy derivations shall not be given here, since in this work they are only of marginal interest. For more details see for example Ref. [41].

---

[19] If the dislocation is pinned by obstacles and subsequently released, then it is the corresponding pinning energy, which controls the velocity [99, 100].

In the next two sections the two crucial parameters E_f and W_m of dislocation glide will be modelled atomistically for the two glide partials involved in the dissociation processes discussed earlier. An investigation of reconstruction phase shifts in the context of dislocation kinks is excluded here and only the lowest energy structures are introduced [20].

4.5.2 The 90° glide partial

As discussed earlier in Section 2.6.2, kinks in a dislocation interact with each other. Therefore, to minimise the effects of the interaction of the kink-pair itself with its periodic images along the dislocation line, in this section the supercell-cluster hybrid model is now abandoned in favour of pure cluster models. For the 90° glide partial (SP) a $C_{420}H_{214}$ model is used, which extends 6 lattice translations along the dislocation line. This geometry conveniently allows to introduce a double-kink with a kink–kink separation of two lattice translations ($\sqrt{2}\,a_0$), where each kink is separated $\sqrt{2}\,a_0$ from the surface. Fig. 4.15 shows
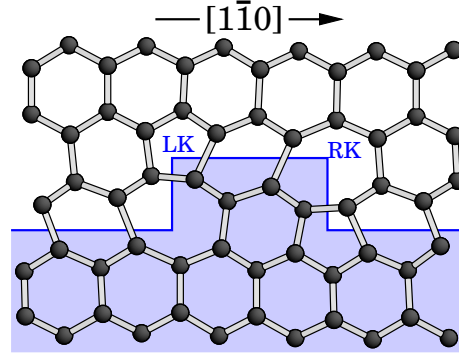

Figure 4.15: A double-kink in the 90° partial (SP). The stacking fault is shaded.

a section of the (111) glide plane containing the relaxed and reconstructed double-kink. All atoms are four-fold coordinated and within the (111) glide plane, in other words neglecting the different stacking on either side of the dislocation line, the structures of the left kink (LK) and the right kink (RK) are identical. The formation energy of the kink pair with separation $L = \sqrt{2}\,a_0$ is obtained as the difference in total DFTB energy between the model containing the kinked dislocation and a cluster of the same stoichiometry and geometry containing the unkinked dislocation segment. The calculation yields the formation energy $E_{\text{pair}}(L = \sqrt{2}\,a_0) = 1.00$ eV.

Assuming the formation energy of a single left kink and a single right kink to be the same ($E_f(\text{LK}) \approx E_f(\text{RK}) \approx E_f$) and further including the energy of the faulted region generated alongside with the kinks, Eq. (2.12) reads:

$$2E_f = E_{\text{pair}}(L) - E_{\text{LK,RK}}(L) - aL\gamma \tag{4.8}$$

Here a is the kink height and $E_{\text{LK,RK}}(L)$ the elastic kink–kink interaction energy of the 90° partial (Eq. (2.11)):

$$E_{\text{LK,RK}}(L) = -\frac{\mu\,a^2}{8\pi\,L}\frac{1-2\nu}{1-\nu}|b_{90}|^2 \tag{4.9}$$

With the kink height $a = \sqrt{3/8}\,a_0$ given as the Peierls valley distance, one obtains a single-kink formation energy of $E_f = 520$ meV — for the chosen kink–kink separation, elastic interaction and the contribution of the stacking fault nearly cancel each other.

---

[20]The discussion of the structures of kinks and *phase shifting defects* as given in Refs. [101, 102] for silicon can be directly applied to diamond as well. The lowest energy structures are basically the same.
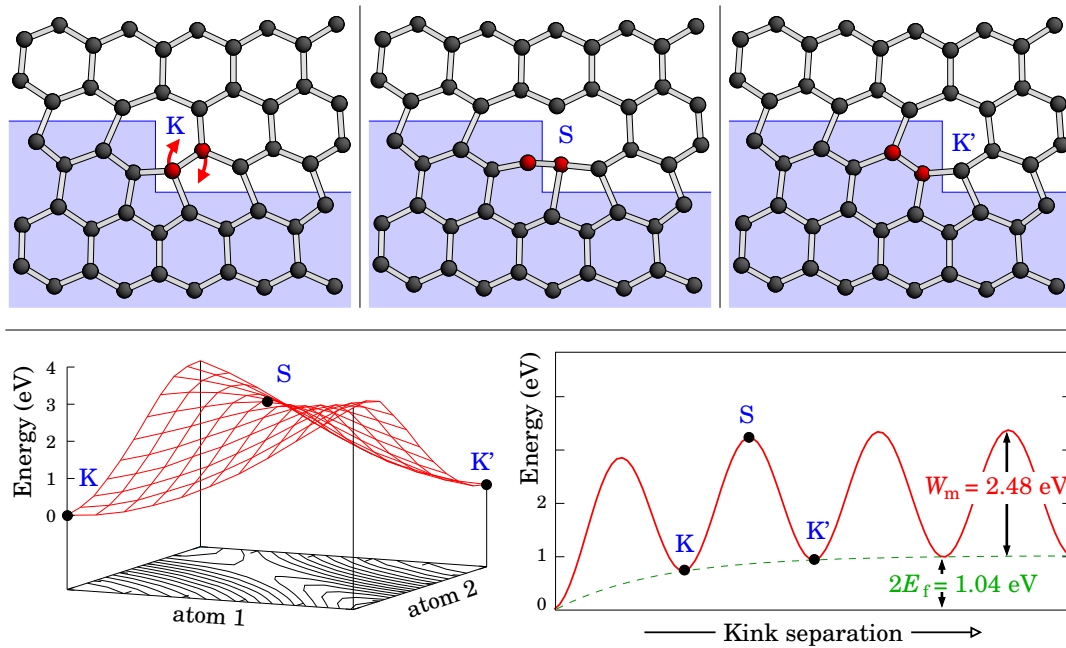
Figure 4.16: Kink migration at the $90°$ (SP) glide partial. *Upper panel:* The three main stages of one elementary kink migration step. The relaxed structures of the starting kink **K**, the saddle point **S** and the migrated kink **K'** are shown projected into the glide plane. The faulted region is shaded and arrows indicate the motion of the two atoms involved.
Lower left panel: The energy surface of the corresponding process leading from **K** via **S** to **K'**. The two parameters are defined via the two atoms involved (see text and Fig. 4.17).
Lower right panel: Schematic representation of the energy of the glide process. As shown in Fig. 4.14, a kink pair is formed and subsequent migration of the two kinks enlarges their separation. **K**, **S** and **K'** are labelled for an arbitrary migration step. The dashed line connecting the minima represents the formation energy of the kink pair. The energy contribution of the expanding stacking fault is not included in the graph.

To model kink migration along the dislocation line is more complicated. The elementary migration step has to be parameterised so that the minimum energy path leading from the starting to the final structure can be found and the barrier determined. Fortunately, all kink migration steps discussed in this work involve mainly the motion of two atoms, which break bonds with neighbouring atoms and form new bonds in the migration process. All other atoms remain at their respective lattice sites. For a right kink at the $90°$ (SP) partial this is illustrated in Fig. 4.16 (upper panel). The parametrisation of the motion (or diffusion) of one of the involved atoms is explained in Fig. 4.17. Varying the two parameters independently, yields a two-dimensional energy surface as given in Fig. 4.16 (lower left panel)[21]. This procedure was applied to a cluster containing a dislocation segment with a single right kink, as well to one with a single left kink. The migration barrier appears to be $W_m = 2.48$ eV for both.

---

[21] The energy surface shown was obtained at 10×10 points in the two-dimensional parameter space by relaxing the whole structure subject to constraining the two primary atoms to lie each in a plane perpendicular to the connecting line between their respective initial and final positions. In the vicinity of **S**, the parameter mesh was refined by a factor of 10.
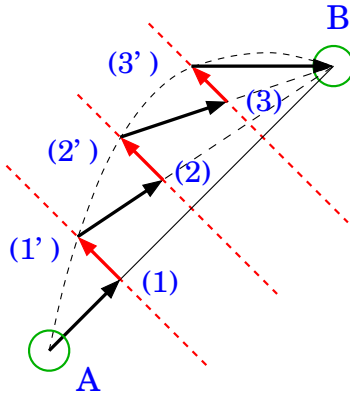
Figure 4.17: The parametrisation of a diffusing atom. **A** and **B** are the positions of the moving atom in the relaxed starting and relaxed final structure respectively. The movement of each atom is modelled as follows: Starting at **A**, the atom is displaced towards **B**. The resulting position is **(1)**. The whole model is relaxed constraining the moving atom to a plane perpendicular to the connecting line between **A** and **B**. The resulting position is **(1′)**. This process is repeated until **B** is reached. The structural configurations corresponding to **(1′)**, **(2′)**, **(3′)**, …then lie on a low energy path from **A** to **B**. And the movement of one atom can be parameterised by its distance from **A** projected onto the connecting line between **A** and **B**.

With the kink migration barrier as well as the formation energy now at hand, the activation energy in the kink migration process evaluates to $Q_{90} \approx 3.5$ eV for short dislocation segments following Eq. (4.7). The energy corresponding to the whole glide process of the 90° (SP) glide partial is schematically shown in Fig. 4.16 (lower right panel) for the first four kink migration steps. There the dashed line connecting the minima represents the formation energy of the kink pair and the solid line the energy of the minimum energy path for glide motion. The first few minima are considerably lower due to the attractive kink–kink interaction. Since the elementary processes are the same for kink migration along the 90° (DP) glide partial, a similar barrier and formation energy can be expected.

4.5.3 The 30° glide partial

After having gone in detail through the simple case of kink formation and migration along the 90° (SP) glide partial, let us now discuss the the equivalent processes at the 30° glide partial. With the latter partial adopting a double-period core reconstruction, the situation is far more diverse. Depending on the kink position relative to the 30° partial reconstruction bonds two different structures for the left kink as well as for the right kink are possible. These different structures will be labelled LK1, LK2 and RK1 and RK2 for the two left and right kinks respectively. Fig. 4.18 shows sections of the (111) glide plane containing relaxed and reconstructed double-kinks. Like in the case of the 90° partial, for each kink all atoms are four-fold coordinated. To obtain the kink formation energies, clusters similar to those for the 90° partial are used. However, unlike at the 90° partial, the structures of left and right kinks are very different, and therefore a similar formation energy cannot be expected. Since in a cluster the formation energy of a single-kink cannot be determined, it is only possible to obtain the sum of two formation energies similar to Eq. (4.8):

$$E_{\mathrm{f}}(\mathrm{LK}X) + E_{\mathrm{f}}(\mathrm{RK}Y) = E_{\mathrm{pair}}(X, Y, L) - E_{\mathrm{LK,RK}}(L) - aL\gamma \tag{4.10}$$

X and Y can be 1 or 2 (compare Fig. 4.18). For the 30° partial the elastic kink–kink interaction energy evaluates as (Eq. (2.11)):

$$E_{\mathrm{LK,RK}}(L) = -\frac{\mu a^2}{32\pi L} \frac{4+\nu}{1-\nu} |\boldsymbol{b}_{30}|^2 \tag{4.11}$$
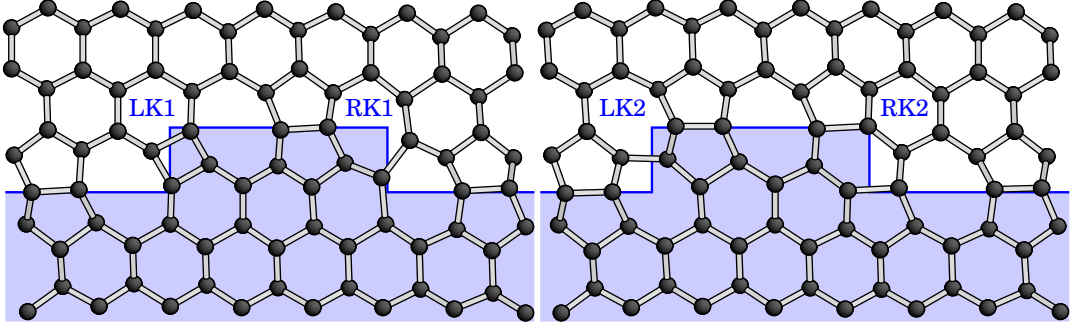
Figure 4.18: Double-kinks in the 30° partial. The stacking fault is shaded. *Left:* Left kink and right kink of type 1 separated by three lattice translations. *Right:* Left kink and right kink of type 2 separated by three lattice translations.

After calculating the formation energies of all possible four combinations (LK*X*,RK*Y*) comparison yields:

- $E_f(LK1) + E_f(RK1) = 3.00$ eV

- $E_f(LK2) = E_f(LK1) + 2.18$ eV

- $E_f(RK2) = E_f(RK1) + 0.78$ eV

LK1 and RK1 are the lowest energy kinks[22]. Thus in the kink migration process it is an LK1-RK1 double-kink that will be preferentially formed at the initial stage. However, subsequent migration inevitably involves the two high energy kinks as intermediate structures. As a consequence, four migration barriers have to be determined. The corresponding migration steps LK2 → LK1 → LK2' and RK2 → RK1 → RK2' are shown in Fig. 4.19. As one can see, LK2 → LK1 is just the reverse process of LK1 → LK2' since all structures involved are symmetric within the glide plane if the different stacking on either side of the dislocation is ignored. The same holds for RK2 → RK1 and RK1 → RK2'. As expected, the two barriers for left kink migration are found identical within 1 % error and so are the two right kink barriers:

- $W_m(LK) = 3.49$ eV

- $W_m(RK) = 2.66$ eV

Fig. 4.19 (lower panel) shows the resulting energy of the glide process at the 30° partial for a moving right and left kink separately. In principle, all which has been said for the 90° partial applies here as well. However having two different barriers for left and right kink migration respectively results in a modified expression for the partial velocity. Eq. (4.7) is replaced by [102]:

$$v_{disl} \propto e^{-\frac{E_f(LK1)+E_f(RK1)}{kT}} \left[e^{-\frac{W_m(LK)}{kT}} + e^{-\frac{W_m(RK)}{kT}} \right] \tag{4.12}$$

---

[22] There is no apparent reason why LK2 is so much higher in energy. Nunes et al. [102] encounter a similar phenomenon for the right kink in silicon, and attribute it to the medium-range behaviour of the associated strain field.
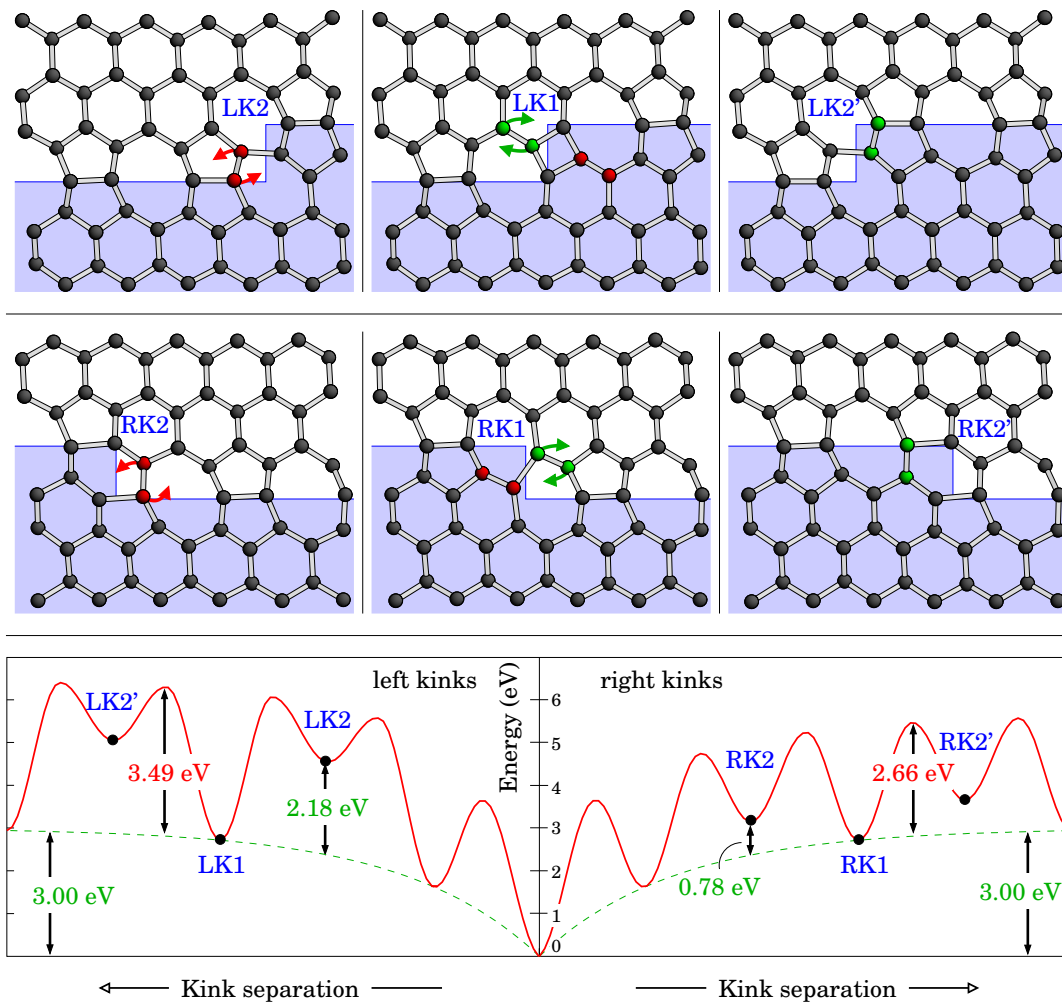
Figure 4.19: Kink migration at the 30° glide partial. *Upper panel:* The elementary kink migration steps of the left kink LK2 → LK1 → LK2'. The relaxed structures of a starting kink **LK2**, and the subsequent kinks **LK1** and **LK2'** are shown projected into the glide plane. The saddle point structures are not given. The faulted region is shaded and arrows indicate the motion of the two involved atoms. *Middle panel:* The elementary kink migration steps of the right kink RK2 → RK1 → RK2'. *Lower panel:* Schematic representation of the energy of the glide process. As shown in Fig. 4.14, a kink pair is formed and subsequent migration of the two kinks enlarges their separation. Since for the 30° partial the migration barriers differ for left and right kinks, the corresponding processes are shown separately. The dashed line connecting the minima represents the formation energy of the respective lowest energy kink pair. The energy contribution of the expanding stacking fault is not included.

Due to its considerably lower migration barrier, the right kink is more mobile than the left kink and dominates the partial velocity. However, unlike in the case of the 90° partial, where the migration barriers clearly control the velocity, here the rather large kink formation energy plays an almost equal role.

As can be seen easily now, the 90° partial appears to be the by far more mobile partial dislocation, and hence it can be expected to move faster under applied stress and raised temperatures. Overall, the activation energies found are rather large and under low pressure glide motion will only be observed at high temperatures.

Comparison with independent theoretical work Similar atomistic calculations of kink formation energies and mobilities have been reported in Ref. [102] for silicon only, and thus cannot be compared directly. The structures found for silicon are very similar to those found for diamond in this work. However, formation energies and barriers appear to be considerably lower in silicon than they are in diamond, which reflects diamond being of course the by far harder material. Also, for silicon the left kinks appear to be the more mobile species at the 30° partial. It should be noted, that all modelling in Ref. [102] utilises supercells, however, the resulting kink–kink interaction terms are apparently *ignored*.

For the 90° partial in diamond theoretical results for the activation barrier Q_{90} vary between 2.9 eV for a soliton mechanism as described in Ref. [80] and 3.3 eV in Ref. [77] — both lower than the 3.5 eV found here.

In this work, as well as in all the theoretical work cited above, the effects of point defects as obstacles to dislocation motion are ignored. However, point defects might have a major influence on glide motion [99, 100]. This has been discussed for silicon in Refs. [2, 3].

4.6 Electron microscopy — a first link to experiments

So far only the theoretical aspects and in particular the modelling of dislocations have been discussed. Results were mainly compared with other independent calculations. This is of course legitimate, but the capabilities of the models and approaches used and the quality of the results can only be judged in comparison with experiments. Therefore this section will establish a first link to experimental findings beyond the mere proof of the presence of dislocations in diamond.

As already mentioned in the previous sections, 60° and screw dislocations in diamond and their dissociation have been observed early on. A typical weak-beam TEM image was given in Fig. 4.1 (1983, [69]). These early studies cannot yield any information concerning the atomic core structure of dislocations. However, electron microscopy has improved since and high resolution imaging nowadays allows to visualise the atomic structure of dislocations in diamond to some extent (Appendix B).

Fig. 4.20 depicts a high-resolution transmission electron microscopy (HRTEM) image of a dissociated 60° dislocation in diamond. This, as all further HRTEM imaging in this section and also the corresponding image simulations and part of the analysis were performed by Bert Willems and Oleg Lebedev (EMAT, University of Antwerp). Assuming both dislocations in Fig. 4.20 are minimum Burgers vector dislocations, the projected Burgers vector allows an easy identification of the left one as a 90° and the right as a 30° Shockley partial. However, the contrast pattern does not identify single atoms. Also an exact localisation of the two dislocation cores is impossible. The latter is possibly a result of the dislocation not being straight through the whole thickness of the layer, but kinked back and forth. This then gives a rather diffuse image of the core region. The intrinsic stacking fault in between the two partials can be identified clearly though. The stacking fault is at least 35 Å wide, but might be considerably larger. Given the rather flat energy minimum for the predicted equilibrium width of 35 Å in Fig. 4.13, such variation seems reasonable — especially
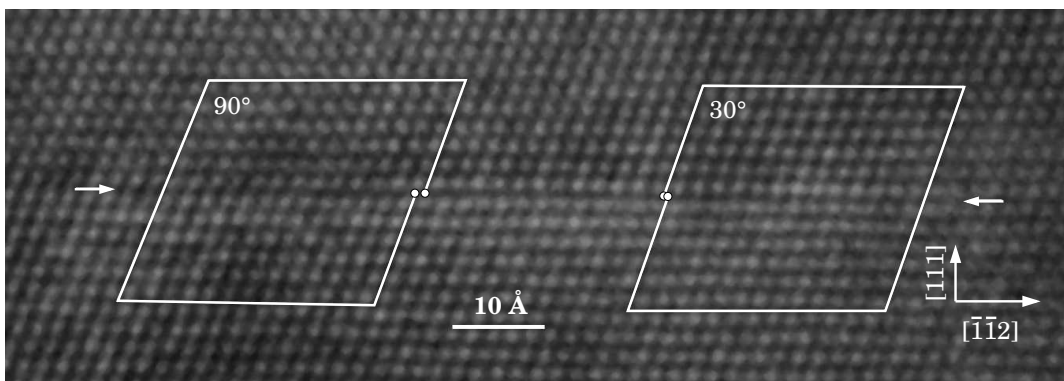
Figure 4.20: Cross sectional HRTEM image of a dissociated 60° dislocation in natural brown type IIa diamond. The two Burgers circuits around the 90°(left) and 30° (right) partial are drawn projected into the $(1\bar{1}0)$ plane, yielding a projected Burgers vector of $\frac{1}{6}[\bar{1}\bar{1}2]$ and $\frac{1}{12}[\bar{1}\bar{1}2]$ respectively. The glide plane is marked by two small arrows. (HRTEM performed by Bert Willems, EMAT, University of Antwerp.)

since preceding sections predict considerable barriers between two adjacent Peierls valleys.

To relate the observed contrast patterns to the atom positions calculated for various dislocation core structures in this work, TEM image simulation is a useful tool. Some details on the underlying calculations are given in Appendix B.2. Fig. 4.21 shows a simulated HRTEM contrast pattern for a given defocus and sample thickness. The dislocation is the 30° glide partial as shown in Fig. 4.7 (d). The input atom coordinates for the simulations are the relaxed atom positions of the corresponding hybrid model. Simulated images like these can now be compared with experimental images to identify atomic core structures.
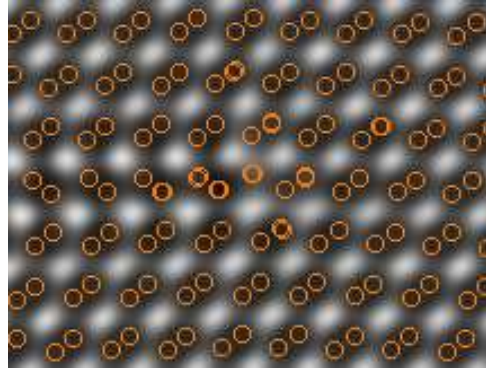
Figure 4.21: Simulated HRTEM image of the 30° glide partial with a defocus $\Delta f = -40$ nm and a sample thickness of 4 nm. Atom positions are indicated as circles.

Fig. 4.22 gives the simulated images for the 60° dislocation and all low energy Shock-

| (a) 60° glide | (b) 60° shuffle | (c) 30° glide |

| (d) 90° glide (SP) | (e) 90° glide (DP) | (f) 90° shuffle (V) |
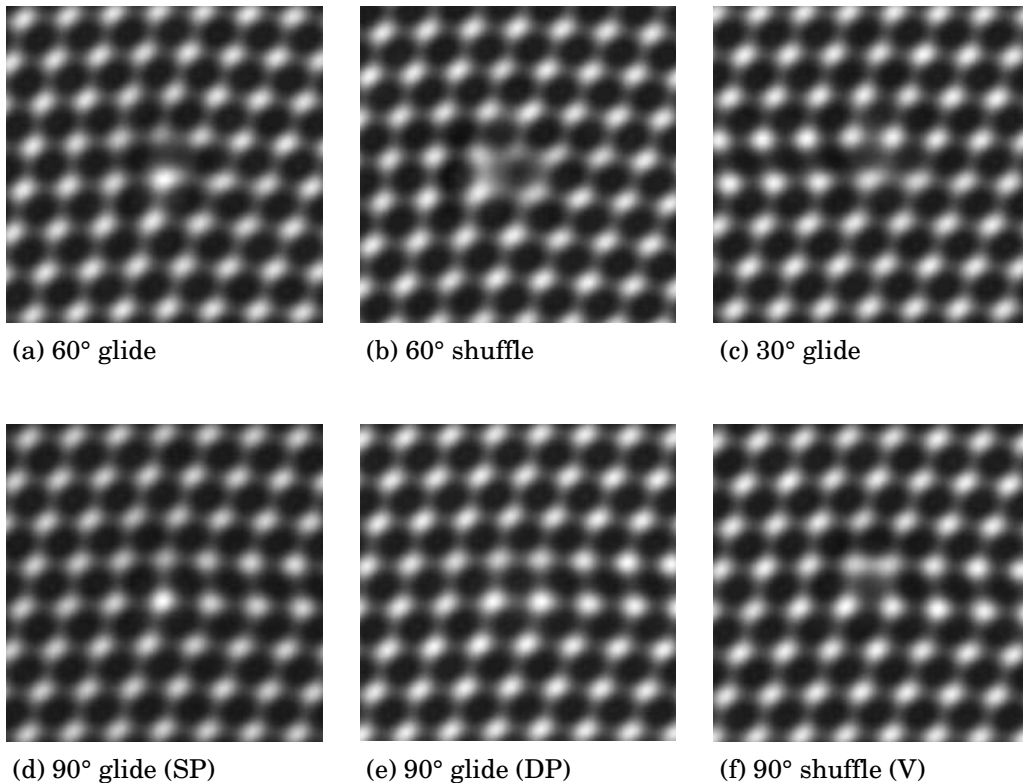
Figure 4.22: Simulated HRTEM images of low energy dislocation core structures with a defocus $\Delta f = -40$ nm and a sample thickness of 4 nm. The input atom coordinates for the simulations are the relaxed atom positions of the respective hybrid model. Fig. 4.7 and 4.8 show the corresponding atomic structures in the same orientation. Each image is approximately centred on the dislocation core line and the (111) glide plane lies horizontally.
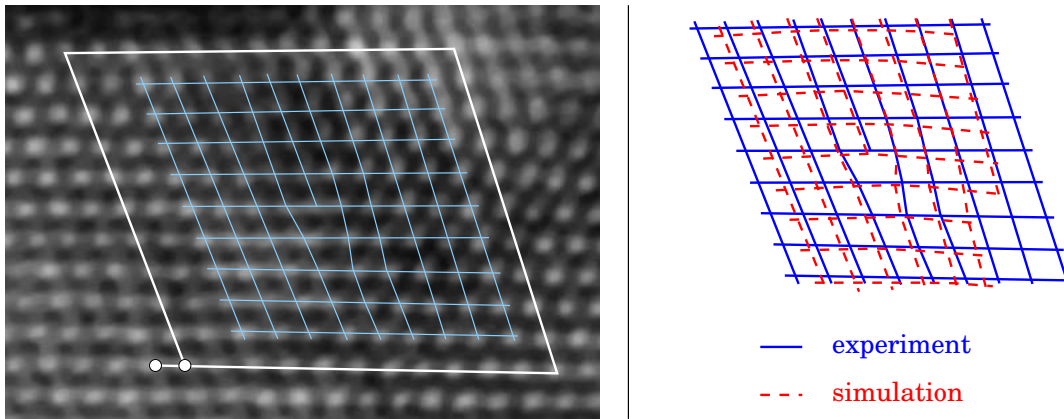
Figure 4.23: Comparison of the experimental and the simulated HRTEM image of an undissociated 60° dislocation. *Left:* Experimental image. The Burgers circuit shown yields a projected Burgers vector of $\frac{1}{4}[\bar{1}\bar{1}2]$ identifying the dislocation. The overlaying mesh of lines approximately connects the points of highest intensity in the contrast pattern. *Right:* The mesh of the experimental image superimposed with the corresponding mesh of the simulated image of a 60° glide dislocation (Fig. 4.22 (a)). The latter is mirrored to facilitate comparison. The small offset between the two meshes is intentional, to allow both to be seen clearly. As mentioned in the text, the main deviation between the simulated image and the experimental one arises from the artificial bending of the (111) plane in the hybrid models applied in this work. (HRTEM performed by Bert Willems, EMAT, University of Antwerp.)

ley partials[23]. The glide and shuffle type of the 60° dislocation appear to be rather distinct in their HRTEM image. Also the vacancy shuffle structure of the 90° partial can be easily distinguished from the two glide structures. However, the difference between the single and double-period reconstruction of the 90° glide partial is miniscule — at least they can hardly be distinguished by eye. An effect common to all simulated images is the bending of the (111) plane. The larger the edge component of the dislocation, the larger the effect. This bending is a result of the limited size of the hybrid models: The inserted material in the upper half (with respect to [111]) of the model leads to an artificial expansion, which would not occur if the dislocation was embedded in an infinite crystal. This further complicates the comparison with experimental HRTEM images[24].

In Fig. 4.23 the experimental and simulated cross sectional HRTEM images of a 60° dislocation are compared. Having the largest edge component of all investigated dislocations, the hybrid model of the undissociated 60° dislocation shows the

---

[23] Comparing Fig. 4.22 with Fig. 4.21 shows us how easily the human eye and the subsequent neural data-processing are deceived: Without the atom positions being shown, one might easily assume the bright spots to be atoms or atom pairs and the large dark areas to be the empty [1$\bar{1}$0] channels. However, the real situation for the given defocus and specimen thickness is exactly the opposite. This demonstrates the importance of image simulations when it comes to interpreting experimental HRTEM images.

[24]To reduce this effect by enlarging the diameter of the hybrid models would result in a large increase in computational cost. The alternative — to abandon the free relaxation of the hybrid's surface — is a non-trivial problem: A rigid surface, even with the surface atom positions calculated in elasticity theory, would not allow core structure specific volume expansion. With the core that strongly confined, the resulting core energies turn out to be unreasonably large. Hence these two most simple approaches to improve the geometrical boundary conditions are not really feasible.
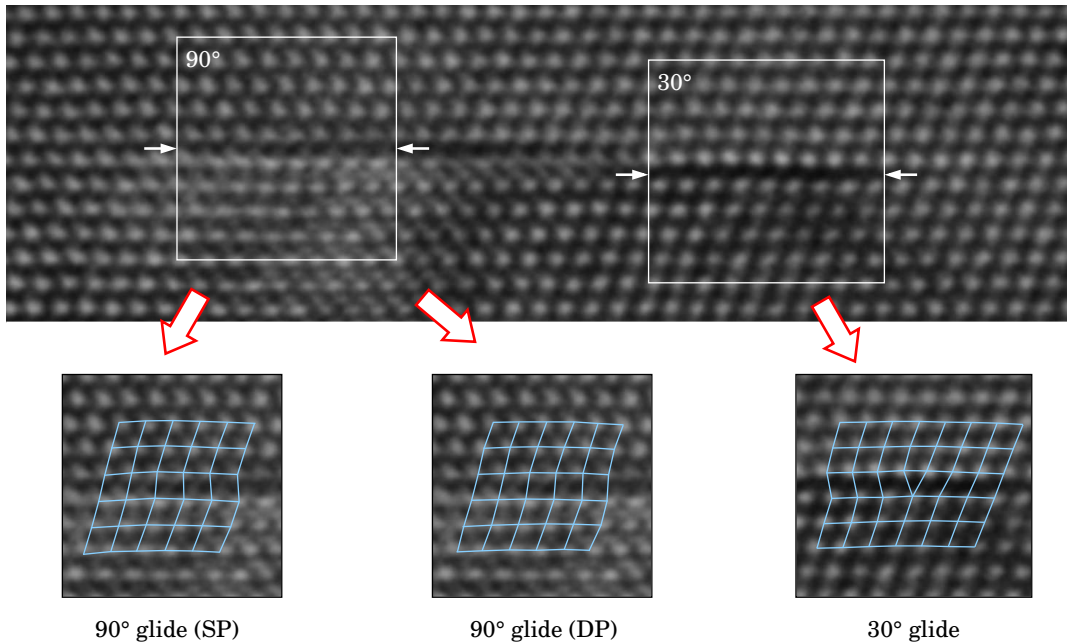
90° glide (SP)   90° glide (DP)   30° glide

Figure 4.24: Comparison of the experimental and the simulated HRTEM image of a dissociated 60° dislocation. *Top:* The experimental image. The two regions containing the 90° and 30° Shockley partial (left and right respectively) are framed in white. Apparently the two partials do not lie on the same glide plane — the two glide planes are indicated by small arrows. *Bottom:* The mesh derived from the simulated images of the two types of 90° glide partials and the 30° glide partial is shown superimposed on the respective region of the experimental image. (HRTEM performed by Bert Willems, EMAT, University of Antwerp.)

strongest artificial bending of the (111) plane, leading to considerable deviation. However, keeping this in mind, at least locally the agreement is reasonable. Similarly, in Fig. 4.24 a 90° and 30° partial are shown. Here the agreement is much better, especially for the 30° glide partial. Unfortunately the contrast pattern does not allow a rigorous decision over the periodicity of the core reconstruction of the 90° partial. However comparison with the simulated single-period structure shows slightly less deviation. Thus in this specific case the SP reconstruction seems more likely. One should keep in mind however, that the diffuse contrast of the 90° partial might result from the partial not being straight, but heavily kinked. Then the question of periodicity becomes obsolete anyway, since the DP structure can be simply seen as a periodically kinked SP structure (Fig. 4.7).

Conclusions: The cited experimental work proves the existence of some of the predicted dislocation types in the investigated specimens of natural HPHT-treated brown type IIa diamond.

At least within the rather crude approach to compare simulated and observed HRTEM images, the observed core structures are in reasonable agreement with those calculated in the earlier sections.

Also, an undissociated 60° dislocation could be identified, which is in agreement with the predicted barrier to its dissociation as shown in Fig. 4.13.

4.7   Electronic structure calculations and electron energy-loss

So far all calculated quantities have been exclusively related to dislocation structures, their interaction and their core energies. Except for test and reference calculations, in this context DFTB and continuum elasticity theory were the methods predominantly applied. However, as most other types of defects, dislocations also effect the electronic structure of the crystal in their vicinity. In this section, these effects will be investigated within the DFT-pseudopotential approach as implemented in AIMPRO (see Section 1.1.2).

In DFT the electronic structure of the ground state is given as the spatial distribution of the electron density, and the corresponding Kohn-Sham eigenvalues (the ε_i in Eq. (1.8)). In a periodic crystal the latter is usually best described as a band structure with k-dependent (quasi-)continuous eigenvalues. As mentioned before, pure diamond has a relatively large (5.49 eV) and indirect band gap (the valence band maximum (VBM) and the conduction band minimum (CBM) are associated with different k-points)[25]. It should be noted, that DFT in LDA usually *underestimates* the band gap: With the models used in this work AIMPRO yields a band gap 1.48 eV too small for diamond.

Dislocations are line defects, periodic only along their line, but not in the two perpendicular dimensions. Hence the electronic structure in a region of the crystal containing a dislocation is that of a one-dimensional periodic arrangement, parameterised by an only one-dimensional k-space. The band structure of the surrounding bulk crystal in that region is therefore projected onto the corresponding direction in its three-dimensional k-space[26]. Therefore, when comparing a dislocated region with a bulk region of the crystal, we have to compare the respective projected band structures[27].

There are three distinct major origins of dislocation related changes in the projected band structure:

Dangling bonds: Some of the dislocation structures discussed in previous sections appeared to have dangling bonds — resulting from atoms in an electronic sp^3-hybrid configuration, bonded to only three other neighbour atoms, instead of four. Such dangling bonds might result in states deep within the band gap. These states are usually strongly localised in the vicinity of the dangling bonds.

---

[25] In semiconductors and insulators the eigenvalues of the occupied Kohn-Sham eigenstates are separated from those of the unoccupied states by a forbidden energy gap, wherein no states are present. In terms of band structures we speak of the *valence band* (occupied) the *conduction band* (unoccupied states) and the *band gap* in between the two.

   The concept of band structures as well as its application to cubic or hexagonal crystals shall not be revised further. The subject is covered in most textbooks on quantum mechanics and a brief introduction is given in [48].

[26] This is similar to the situation at surfaces, where the three-dimensional k-space is reduced to a two-dimensional space.

[27] Of course strictly speaking a band structure is not defined for a finite region of the crystal, but only for an infinite crystal. However, for large enough regions, and assuming the spatial coordinates to be "good" quantum numbers at least on the large scale, one can approximately speak of a "local" band structure and local defect-induced changes to it.

Strained bonds and different bonding configurations: In the very core region of most dislocations, bond angles and lengths deviate considerably from those found in the bulk crystal. This is usually the case for reconstruction bonds, but also otherwise "bulk-like" bonds at the core are often considerably stretched or compressed due to the strain field. Also non-sp^3 hybridisations might occur.

All these lead to changes in the density of states below the VBM and above the CBM. However, if the strain is large enough, then localised gap states, usually close to the band edges, are possible.

The medium range strain field: All dislocations are surrounded by a strain field. The strain results in a local reduction of the band gap in the vicinity of the dislocation core. This effect leads to carrier confinement and might yield localised transitions near the dislocation, even if no states in the gap are present [103].

Within the approach used in this work, the band gap reduction due to the strain field can be observed to some extent, however it cannot be quantified and thus will not be discussed further.

4.7.1   The computational approach

To obtain a good representation of the Kohn-Sham eigenstate spectra associated with the dislocation core structures discussed in preceding sections, the DFTB method in its minimal basis approach is not sufficient: The valence band and defect states close to it are represented reasonably well, conduction band states however deviate considerably from more accurate DFT-pseudopotential calculations (AIM-PRO)[28]. The conduction bands play an important role in electron energy-loss spectroscopy (EELS), which will be discussed later in this section. Hence the use of DFT-pseudopotential calculations is essential at this point.

To obtain the projected band structure corresponding to one of the various low energy dislocation core structures in Section 4.3, the outer shells of atoms are removed from the relaxed structure and the new surfaces are terminated with hydrogen. The models thus obtained are still of the supercell-cluster hybrid type, however smaller in diameter to allow treatment within the computationally more expensive approach: The number of atoms per model now ranges between 60 and 160 atoms depending on the core radius and the periodicity of the structure[29]. These structures are then relaxed with the AIMPRO code. In all calculations the electronic wavefunctions are expanded in a set of s, p, and d atom-centred Gaussian orbitals with 28 variational degrees of freedom per atom and the Brillouin zone integration during self-consistency is performed using a set of between 2 and 4 Monkhorst-Pack reduced k-points along the axis in k-space [84], which corresponds to the dislocation's periodicity.

---

[28] But still one has to keep in mind that the description of the conduction band states suffers from DFT, at least in the form applied in this work, being a ground-state theory.

[29] This might sound rather small, and for the calculation of core energies it would be too small indeed (compare Fig. 4.6), however for all investigated cases the actual core structures were still represented well after relaxation of the smaller hybrid models.

To then obtain the electronic band structure of the relaxed structure, the Kohn-Sham eigenstates are calculated at approximately 25 different k-points along the same axis.

Electron energy-loss spectroscopy and its simulation: The experimental technique yielding the most direct information on the electronic band structure is electron energy-loss spectroscopy. One has to distinguish between *low-loss EELS*, where the signal results from the scattering of high-energy primary electrons with secondary valence electrons which undergo a transition to empty conduction or gap states, and *core-excitation EELS*, which results from electronic transitions between deep atomic core states and empty gap or conduction band states. For the former, the signal obtained is proportional to $\text{Im}(-1/\varepsilon(E))$ where $\varepsilon = \varepsilon_1 + i\varepsilon_2$ as the dielectric function and E is the energy loss. Core-excitation EEL spectra are related to an angular momentum projected local density of states at the respective atom. In the case of the s-like diamond K edge this will be the p-projected local density of states. More details on both the experimental as well as the modelling aspects of EELS can be found in Appendix C.

Some remarks on the effects of boundary conditions: In this work EEL spectra are calculated exclusively in supercell-cluster hybrid models. However, modelling in pure supercells might yield different results since the boundary conditions influence the electronic structure. This has been tested for dislocations in gallium nitride [11], and differences in the EEL spectra due to the different boundary conditions seem to occur only at high energies well above the band edge.

The theoretical results presented in the following have been obtained in collaboration with Caspar J. Fall (University of Exeter, UK), who implemented the EELS functionality into the AIMPRO code and subsequently performed the EELS simulations.

4.7.2   Calculated band structures and EEL spectra

The electronic structure calculations turn out to be computationally rather expensive, since a considerable number of k-points is required for the simulated EEL spectra: Depending on periodicity and model size, for the EELS calculations in this section 50–150 k-points have been used. On this account not all core structures as given in Fig. 4.7 and 4.8 are investigated, but only a selection of structures, which will either exist predominantly or show some interesting features.

Fig. 4.25 shows the calculated band structures, low-loss EEL spectra and K-edge core excitation EEL spectra for the perfect 60° glide and shuffle dislocations. The results for the lowest energy Shockley partials of the $\{111\}\langle 110\rangle$ slip system, namely the 30° and 90° (DP) glide partials and the vacancy structure of the 90° shuffle partial, are given in Fig. 4.26. The length of the k-axis reciprocally depends on the periodicity length along $[1\bar{1}0]$ of the core reconstruction. Therefore, being the only single-period structure, the 60° shuffle dislocation appears with twice the length in k-space. Both low-loss and core EELS at the dislocation core are shown in comparison with spectra calculated in bulk supercells. As shown in Appendix C, the calculated bulk spectra agree rather well with experiments.
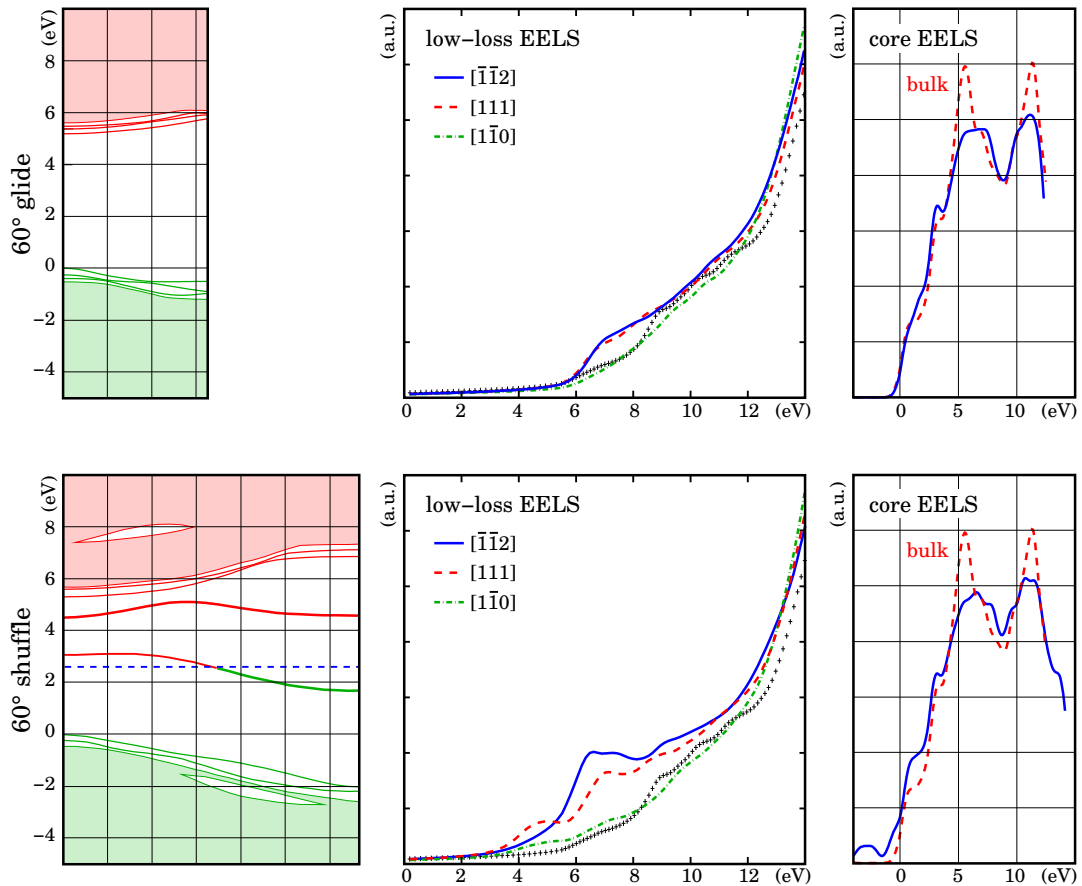
Figure 4.25: The projected band structures and simulated EEL spectra of the perfect 60° dislocation. In the band structures (left column) only the bands close to the gap are drawn, the rest of the valence and conduction band region is shaded. The origin of the Brillouin zone is at the far left of each band structure. The conduction band is shifted upwards by 1.48 eV to match the experimental gap. States within the gap are shifted proportionally to their distance in energy from the valence band top. For the low-loss EEL spectra (middle column) the results for different beam orientations are shown. The line of plus signs gives the bulk spectrum. The *K*-edge core EEL spectra (right column) are displayed with a Gaussian broadening of 0.8 eV and 8–10 inequivalent atoms neighbouring the dislocation core have been taken into account. The dashed line gives the bulk spectrum. *Upper panel:* The 60° glide dislocation. The band structure is folded once (reducing the Brillouin zone), since the structure is period-doubled (See Fig. 4.7 (b)). *Lower panel:* The 60° shuffle dislocation. In the band structure the Fermi level, up to which the bands are occupied, is shown as a dashed line. The dislocation is assumed to be uncharged.

Glide dislocations: All glide dislocations show no evidence of deep gap states. This directly stems from the reconstructed nature of the cores which allow all atoms to be four-fold coordinated with bond angles and lengths not too far from bulk material. Consequently, only small changes in the corresponding simulated EEL spectra are observed: Only for beam directions perpendicular to the dislocation line ($[\bar{1}\bar{1}2]$ and $[111]$) is the low-loss spectrum slightly enhanced in the 6–8 eV region.
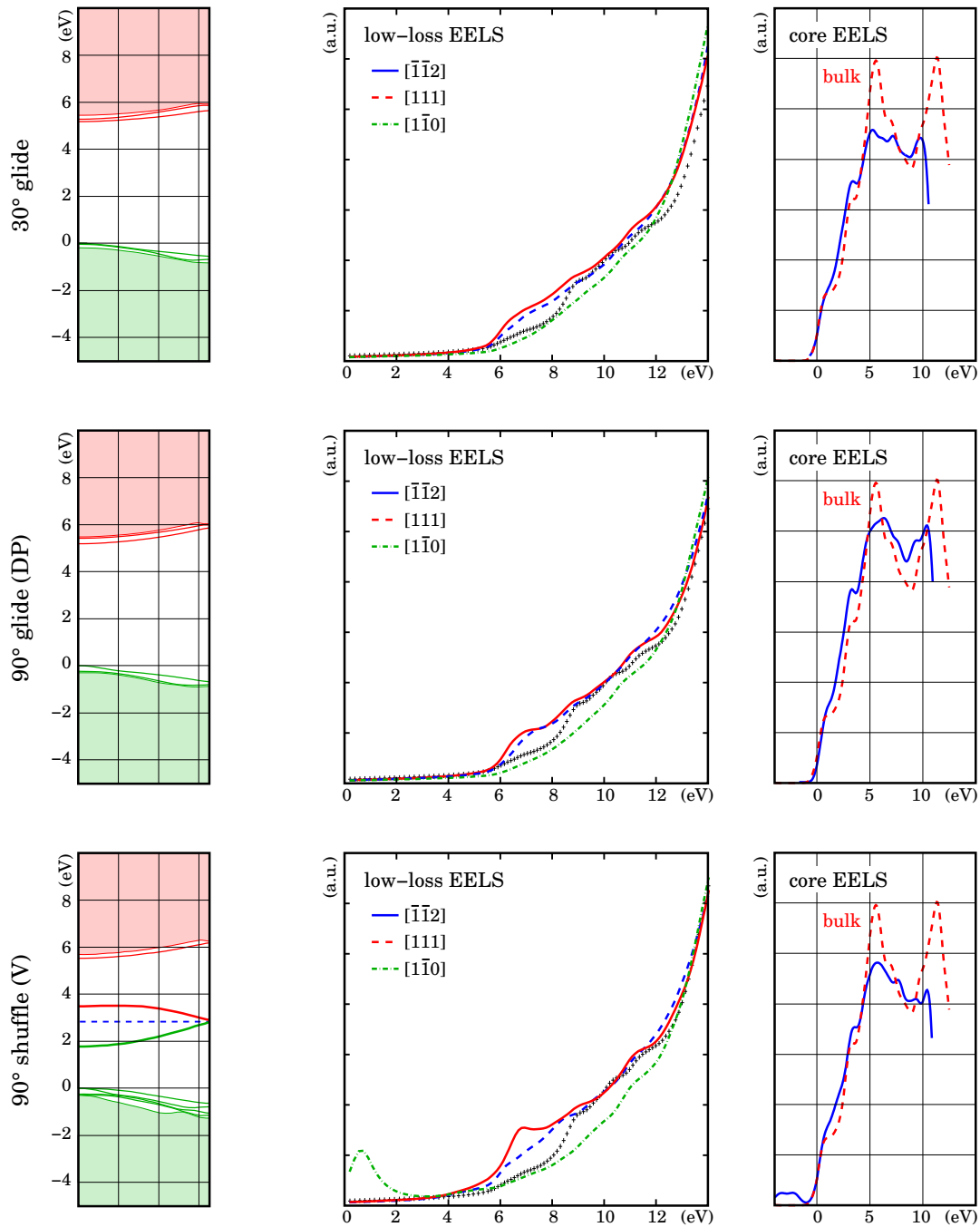
Figure 4.26: The projected band structures and simulated EEL spectra of the low energy partial dislocations. All band structures are folded once (reducing the Brillouin zone), since the structures are period-doubled (See Fig. 4.8). For more details see also caption of Fig. 4.25. *Upper panel:* The 30° glide partial dislocation. *Middle panel:* The 90° (DP) glide dislocation. *Lower panel:* The 90° (V) shuffle dislocation. In the band structure the Fermi level, up to which the bands are occupied, is shown as a dashed line. The dislocation is assumed to be uncharged.

Shuffle dislocations: As can be seen in Fig. 4.7 (c) and 4.8 (d), the two investigated shuffle dislocations each possess a row of three-fold coordinated atoms. The resulting dangling bonds protrude along $[\bar{1}\bar{1}\bar{1}]$ into the wide $[1\bar{1}0]$ core channel. In both cases this results in a band at mid-gap position with similar dispersion. For an uncharged dislocation this band will be only half occupied. At the 90° shuffle (V) structure the band is folded, leading to two gap states per k-point. By that it clearly shows its single-period character — the double periodicity stems from the reconstruction bonds along the dislocation line only, which has hardly any influence on the dangling bonds[30].

The presence of electronic bands around mid-gap is consistent with the observed energy of band-*A* emission at 2.8–2.9 eV, which has been correlated with threefold coordinated atoms in extended defects [71]. The 60° shuffle core additionally contains a stretched bond that contributes to an additional gap state that is higher in energy.

For the subsequent EELS calculations the material is assumed to be *p*-type, leaving all gap states empty. Under these conditions the 60° shuffle dislocation shows strongly enhanced low-loss EEL absorption in the 3–12 eV range for beam orientations perpendicular to the dislocation line. Similarly, for the same beam orientation, the vacancy structure of the 90° shuffle partial is found with enhanced absorption in the 6–9 eV range. Additional strong absorption at the latter occurs in the 0–2 eV range for electron beams parallel to the dislocation.

In core-excitation EELS on these dislocations the empty gap states create supplementary absorption peaks below the conduction band minimum. This might account for previous experimental findings [104].

4.7.3   Experimental EELS

Experimentally, EEL spectra can be taken with high spatial resolution, which allows to distinguish between spectra taken at dislocation core regions and those taken in dislocation-free areas of the crystal. This section will give an example of EELS being applied to dislocations in diamond.

All spectra shown are the result of experimental work performed by A. Gutiérrez-Sosa and U. Bangert at UMIST in Manchester (UK), A. E. Mora and J. W. Steeds at the University of Bristol (UK) and J. E. Butler at the Naval Research Laboratory, Washington DC (USA).

The investigated CVD-grown polycrystalline diamond sample was boron doped with an estimated boron concentration of 10^{21} cm^{-3}. Electron energy-loss spectroscopy was conducted in a scanning transmission electron microscope (STEM), operated at 100 kV. More details on sample preparation and the EELS experiments are presented in Ref. [6].

---

[30] To give some more details: The gap states of the 90° shuffle (V) partial are formed from *p*-like orbitals of the threefold coordinated carbon atoms. As mentioned, because of the double-period reconstruction, the band structure is folded and there are two gap states for each wave vector. In the energetically lower gap state at $k = (0,0,0)$, the *p*-like orbitals are repeated along the core with a phase change of π between successive atoms, while in the energetically higher gap state they are all in phase.
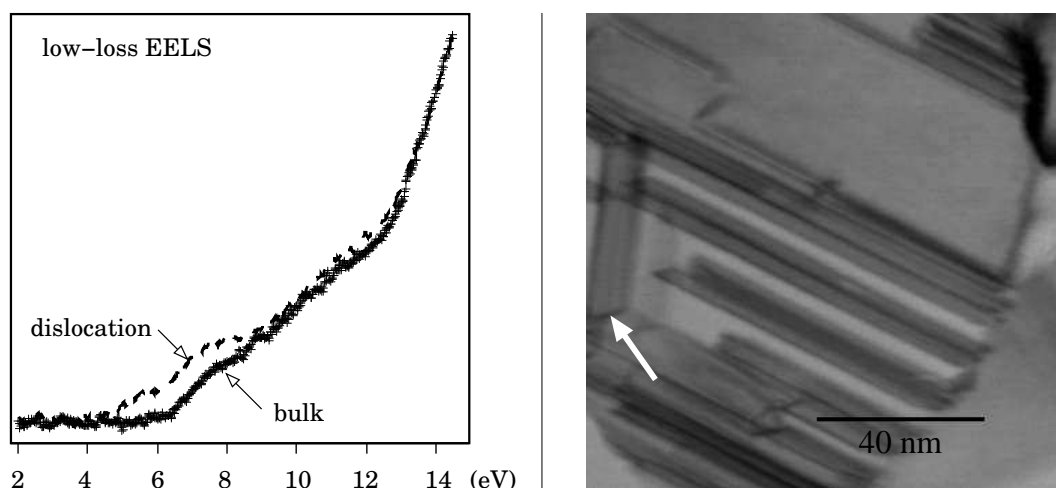
Figure 4.27: Experimental low-loss EELS and TEM on dislocations in CVD diamond. *Left:* Experimental low-loss EEL spectra for bulk (plus signs) and oblique partial dislocations (dashed line) with an electron beam along $\langle 111 \rangle$. Each graph shown results from several tens of EEL spectra taken. *Right:* TEM image of CVD diamond. The arrow marks the dislocation whose EEL spectrum is shown on the left, presumably a 30° glide partial of a dissociated screw.
(All experiments including sample-growth and preparation were performed by A. Gutiérrez-Sosa and U. Bangert at UMIST in Manchester (UK), A. E. Mora and J. W. Steeds at the University of Bristol (UK) and J. E. Butler at the Naval Research Laboratory, Washington DC (USA))

Plan view TEM of the investigated sample showed grain clusters up to a diameter of 0.2 mm with common $\langle 110 \rangle$ or $\langle 112 \rangle$ growth directions. Each such cluster contained up to ten individual grains and the dislocations were predominantly orientated along $\langle 110 \rangle$ directions. Fig. 4.27 shows the spectra obtained on and off such a dislocation. Unfortunately the grain size was too small to allow any determination of the Burgers vector. However, considering their origin and formation they are most likely to be dissociated screw dislocations [66]. In other words their experimental spectra have to be compared with that of the 30° glide partial in Fig. 4.26. In the experiment, the dislocation was oriented neither parallel nor perpendicular to the beam, but oblique. Hence its spectrum does not directly correspond to either of the calculated spectra. However, the calculated spectrum with a beam orientation of [111] perpendicular to the dislocation line can be taken as an approximate representation of the experimental situation: Both the experimental spectrum as the calculated spectrum clearly show an increase in the 6–8 eV region compared to the respective bulk spectrum. Also the relative increase around 7 eV is common to both spectra.

Further spectra taken with a beam orientation along the dislocation line appear to be less intense and rather featureless — a trend which can also be observed in all theoretical spectra.

4.7.4   Electronic structure calculations — conclusions

The calculations have shown that, at least among the low energy dislocations of the $\{111\}\,\langle 110\rangle$ slip system in diamond, and assuming the dislocations to be undecorated by point defects, only the shuffle set of dislocations appears to induce deep electronic states into the band gap. This suggests that the origin of the band-*A* luminescence as observed in natural type-II diamond is either due to point defects bound to dislocation lines, such as impurities or jogs, or to shuffle segments.

As a consequence of the absence of deep gap levels associated with the glide set of dislocations, only the calculated EELS spectra of shuffle dislocations show major differences in comparison with the respective bulk spectrum. Still, it was possible to correlate the features of experimentally observed low-loss EEL spectra in polycrystalline CVD diamond with those of $30°$ glide dislocations. This supports the earlier experimental presumption of dissociated screw dislocations being the predominant species near grain boundaries. In the quoted experimental work on polycrystalline CVD, no indication of shuffle segments could be found.

Earlier core EELS performed by Bruley and Batson [104] on natural type-II diamond shows a supplementary absorption below the conduction band, which is associated with dislocations. This agrees well with the spectra calculated for both the $60°$ shuffle dislocation and the $90°$ shuffle (V) partial, suggesting shuffle segments to be present in their sample.

4.8   Summary and conclusions (diamond)

In this chapter, the low energy dislocations of the $\{111\} \langle 110 \rangle$ slip system in diamond have been modelled in an approach combining DFT-based atomistic calculations with linear isotropic (and to some extent anisotropic) elasticity theory. In this approach elasticity theory has proven invaluable to describe the long range elastic strain effects associated with dislocations and also served as a good convergence criterion for the size of the atomistic models used.

The atomistic modelling — in particular the DFTB method — allowed a convenient determination of dislocation core energies and of the energy offset in the dissociation of dislocations. In elasticity theory the latter properties are both unknown. Wherever there was overlap between the elasticity theory description of dislocations and their atomistic modelling, the agreement was very good: The atomistic calculations were converged to within the elastic limit.

The predicted low energy core structures and dissociation distances match well with those observed experimentally in (HR)TEM. In particular to compare the calculated structures with high-resolution micrographs, the application of image simulations based on the calculated coordinates was crucial.

Based on the predicted core structures it was then possible to calculate the corresponding electronic structures in a pseudopotential approach (AIMPRO). The subsequent simulation of both core and low-loss EEL spectra allowed a direct link to recent experiments to be established.

For the two probably predominant partial dislocations — the $30°$ and the $90°$ glide partials — the kinetics of thermally activated glide motion have also been modelled, assuming a kink formation and migration mechanism.

Comparison with earlier theoretical work: There are relatively few independent DFT-based calculations on dislocations in diamond and they are limited to a small subset of the problems treated in the work presented here. Where comparison is possible the difference between this work and those earlier calculations is small — around 6% for glide activation energies and below 5% for dislocation core energies and their relative differences. This, along with the good agreement with experimental work gives confidence, that the chosen approach of combining two different approximations to DFT with elasticity theory for the long range effects, is well adapted for the problem. However, being a first step towards a multiscale approach, it is by far less computationally expensive than pure DFT-pseudopotential calculations and covers are larger variety of length scales and associated properties.

Further, the chosen supercell-cluster hybrid model has proven very successful to model core structures and energies as well as the electronic properties of straight perfect and partial dislocations.

In the following section the important results are listed explicitly, followed by possible implications.

4.8.1 Selected results

Comparing the shuffle structure of the undissociated 60° dislocation with the glide structure, one finds the latter to be 630 meV/Å lower in line energy. However, since a shuffle → glide transition involves absorption or emission of vacancies or interstitials, one would expect shuffle segments to be present if formed by plastic deformation. These might then undergo the transition to glide structures in a thermal annealing process.

There is a barrier of around 180 meV/Å to the dissociation of the perfect 60° dislocation into a 30° and a 90° Shockley partial. Due to its width this is a considerable barrier which is not easily overcome. This explains the experimental observation of undissociated 60° dislocations even though an overall energy-gain of around 680 meV/Å strongly favours dissociation.

The situation is different for the $\frac{1}{2}[1\bar{1}0]$ screw dislocation, here the three proposed structures (glide, shuffle and mixed) can be transformed into each other by the mere breaking and forming of bonds. The shuffle screw is found to be metastable, but around 860 meV/Å higher in energy than the glide structure which resembles two 30° glide partials in the same Peierls valley. The dissociation of those two partials yields a further energy of approximately 1 eV/Å without any considerable barrier. Thus spontaneous dissociation into two 30° glide partials is strongly suggested. Indeed no undissociated $\frac{1}{2}[1\bar{1}0]$ screw has been observed experimentally.

The equilibrium distance between two dissociated Shockley partials, or in other words the equilibrium stacking fault width, can be determined easily by means of elasticity theory only. Using the elastic constants and stacking fault energies obtained within the DFTB method, the equilibrium distance was evaluated to around 30 and 35 Å for the screw and 60° dislocation respectively. This corresponds well with the experimentally observed widths between 25 and 42 Å. The large variation in observed widths might be explained by the energy minimum being very flat compared to the activation energies required to overcome the barrier between adjacent Peierls valleys.

A comparison of core energies ruled out many of the possible dissociation reactions, leaving only the 30° and 90° glide partials and the vacancy structure of the 90° shuffle partial likely to exist. Consistent with earlier calculations, in this work the double-period core reconstruction of the 90° glide partial is favoured over its single-period reconstruction.

The glide motion of the two glide partials has been considered as a process of kink formation and subsequent migration. Under these assumptions the 90° glide dislocation proves to be the by far more mobile species, with a thermal activation energy of 3.5 eV (the sum of the double-kink formation energy and the migration barrier) for short dislocation segments. At the 30° glide partial the migration barriers are on average 0.5 eV larger and the double-kink formation energy exceeds that of the 90° partial even by 2 eV. The resulting average thermal activation energy of the 30° glide partial is found to be $\overline{Q}_{30} \approx 6.1$ eV. If a dissociated dislocation moves as a whole, then its speed is controlled by the slower partial. Hence the thermal activation energy for both the dissociated screw and for the dissociated 60° glide dislocation will be given as that of the 30° glide partial.

Finally turning towards the electronic properties, only the perfect and partial shuffle dislocations appear to induce deep electronic states/bands in the band gap. The majority of these states arises from dangling bonds at the dislocation line and is located around mid-gap. As shown, the gap states should lead to a considerable increase in low-loss EELS absorption and might also well explain the supplementary core EELS absorption below the conduction band edge, which is experimentally observed in natural type-II diamond. This would then further suggest shuffle segments to be present in that material, which could well be responsible for the well-known band-*A* cathodoluminescence around 2.8–2.9 eV.

In polycrystalline CVD diamond, low-loss EEL spectra obtained near grain boundaries on presumably dissociated screw dislocations seem to correspond well with the calculated spectra of 30° glide partials. However, for a good comparison and to draw strong conclusions, further experimental EEL spectra of different partial and perfect dislocations would be required.

4.8.2 The decolouring of brown diamonds by HPHT treatment

The brown colouring, which is found in most natural diamonds, renders them useless as gem stones. Even though it is assumed to be related to plastic deformation and extended defects, its exact origin remains unknown. As mentioned in the introduction to this chapter, in recent years colour improvement via HPHT annealing became feasible, threatening the gem market. Hence a deeper understanding of the origin and annealing of the colouring is of particular interest and there is a strong interest to find definite methods of identifying HPHT treated stones.

Several indications point towards a possible scenario involving a shuffle $\rightarrow$ glide transition in the annealing and decolouring process:

1. The absence of deep electronic gap states for all perfect and partial glide dislocations excludes them from being responsible for the brown colouring if undecorated. All low energy shuffle structures however, give rise to deep bands in the upper two thirds of the electronic gap, which might account for the colour.

2. The shuffle dislocation is found to be higher in line energy, suggesting a shuffle $\rightarrow$ glide transition in the annealing process.

3. Calculated vibrational spectra show modes (around 1500 cm^{-1}) above the Raman frequency for the shuffle dislocation only [105]. These modes might account for a Raman line found in brown diamond, which anneals out above 2000°C [106], further supporting the transition mentioned above. However, the Raman signal appears to disappear *before* the actual decolouring.

4. In the case of dissociated shuffle dislocations, vacancy structures are more stable than those involving interstitials. Thus, a shuffle $\rightarrow$ glide transition at partials would always involve vacancy emission (or interstitial trapping). This is in agreement with the observed formation of vacancy–nitrogen complexes subsequent to annealing [72]. This process might well be fed by the vacancies released during a shuffle $\rightarrow$ glide transition.

Summarising all findings, it still remains unclear whether a shuffle $\rightarrow$ glide transition occurs *before* the actual annealing of the colour, or if the 1500 cm^{-1} Raman signal becomes just too small to be detected. It can also not be excluded that the observed line has a different origin and is not related to dislocations at all.

4.8.3   Outlook

Further (HR)TEM studies and EELS experiments on HPHT treated decoloured diamond as well as on untreated brown diamond are necessary. Then, in combination with the methods presented here, it might be possible to decide in favour or against an annealing mechanism involving dislocations undergoing a shuffle $\rightarrow$ glide transition.

Further, the electronic structure of kinks and jogs might be of interest, as well as dislocations with line directions different from $\langle 110 \rangle$.

Methodological: In the long run, the way of combining different methods to model the effects of dislocations on different length scales could be improved. A true embedding approach is imaginable, where the region treated atomistically is embedded in a larger region which is described by elasticity theory. The crucial point would then be the interface/coupling between the different regions, which is far from trivial.

Chapter 5

Dislocations in Silicon Carbide

In this chapter selected dislocation-related problems in silicon carbide (SiC) will be addressed, which are of particular interest and technologically relevant. Threading dislocations as mentioned in Section 3.3, which are formed during growth, are very common defects in SiC, but excluded from this work. Hence the only dislocations considered are the low energy dislocations in the basal plane[1].

After giving a short introduction into SiC as a material and explaining the different polytypes, the glide set of basal dislocations is analysed, discussing the main core reconstructions and energies. This is followed by an investigation of the glide motion of 30° and 90° partial dislocations, assuming a kink formation and migration process as already introduced for the case of diamond. Finally, the electronic structure of the discussed defects is examined.

Since all the basic concepts of the modelling of dislocations were discussed in detail with diamond as an example, this chapter will mainly present results and their implications. Whenever there is a considerable difference or problems specific to modelling dislocations in SiC, then those will be pointed out.

5.1   Introduction and background

Silicon carbide is a compound semiconductor consisting of the two components silicon and carbon. It possesses a wide electronic band gap of $2.39 - 3.33$ eV (depending on the polytype [107]) and has extreme properties similar to diamond. Its high thermal conductivity, extreme hardness and resistance to thermal shock, oxidation and corrosion, are useful for engineering applications in extreme environments [108]. As a semiconductor, the rather high breakdown electric field strength make SiC a promising material for high-power, high-temperature and high-frequency devices.

Single crystal SiC CVD-growth is technologically reasonably controlled and appreciable n-type conductivity can be achieved by doping with nitrogen. For p-type usually aluminium or boron are used as dopants [109, 110, 111].

---

[1]The term "basal plane" is used here for the hexagonal polytypes and the cubic polytype alike. In the latter it simply refers to an arbitrary {111} plane.
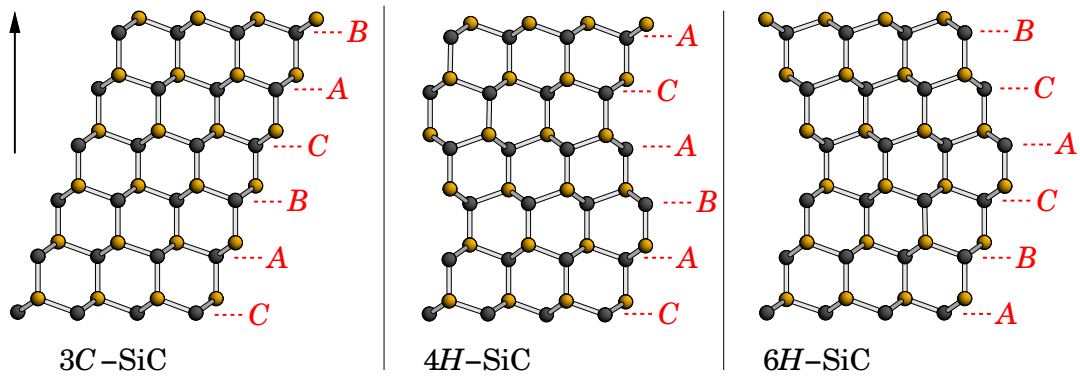
Figure 5.1: The Stacking sequences of 3*C*-, 4*H*- and 6*H*-SiC. Carbon atoms are represented in dark grey, Silicon atoms in yellow (light grey in black and white prints). Only the planes of one sub-lattice are labelled. The arrow indicates the ⟨111⟩ direction for the cubic polytype (3*C*) and ⟨0001⟩ for the hexagonal polytypes (4*H* and 6*H*).

In device development, a strong emphasis lies on high-power diodes: Bipolar 6*H*-SiC diodes have been reported with reverse breakdown voltages as high as 4.5 kV [111] and more recently 4*H*-SiC PiN diodes have shown reverse breakdown voltages > 5 kV with capacities to transmit very high powers > 100 MW. However, these devices show a considerable degradation under forward-biased operation [112, 113, 114, 115]. More details on this problem are given in Section 5.1.2.

Further, high current operation of devices is limited by the material quality. In particular micropipes, which thread the layer and are commonly associated with dislocations, have a disastrous effect: Even a single micropipe can lead to an early electrical breakdown of the device [116]. However as mentioned above, micropipes and threading dislocations will not be discussed further in this work.

5.1.1 The different polytypes of SiC

Unlike most other semiconductors, which only occur in one or two different crystal structures each, silicon carbide is known to exist in over 170 different polytypes [117] with cubic, hexagonal and rhombohedral symmetry. Three very common polytypes are 3*C* (cubic) with the stacking sequence $\cdots ABC|ABC\cdots$ and 4*H* and 6*H* (hexagonal) with the respective stacking sequences $\cdots ABAC|ABAC\cdots$ and $\cdots ABCACB|ABCACB\cdots$ as shown in Fig. 5.1. More details on the concept of crystal stacking can be found in Section 3.2.

Table 5.1 gives the experimental lattice constants and the electronic band gap for different polytypes. In all cases the average bond lengths resulting from the lattice constants are almost identical. The band gaps differ however. Especially that of cubic SiC is considerably smaller than those of the hexagonal polytypes.

The energy differences between the different polytypes are rather small, and consequently stacking faults are low energy defects. Just to give an example, Fig. 5.2 shows an extended stacking fault node in a plastically deformed 4*H*-SiC single crystal. The stacking fault and three equivalent partial dislocations with Burgers vectors $\boldsymbol{b} = 1/3[1\bar{1}00]$, $\boldsymbol{b} = 1/3[01\bar{1}0]$ and $\boldsymbol{b} = 1/3[10\bar{1}0]$ can be identified using
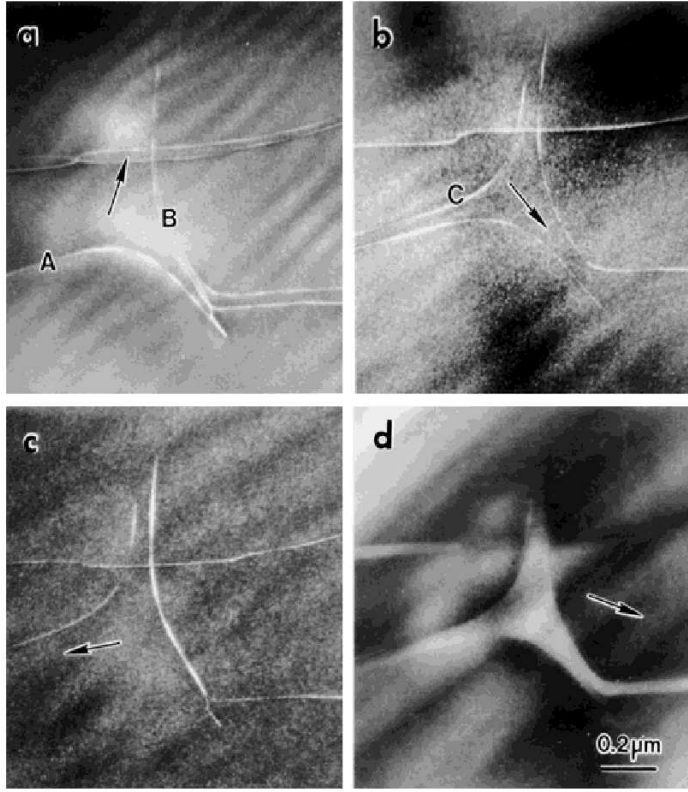
Figure 5.2: TEM weak-beam contrast experiments carried out on an extended dislocation node in a 4*H*-SiC single crystal. The crystal was plastically deformed at 1300°C. Arrows indicate the diffraction vector: (a) $g = [\bar{1}2\bar{1}0]$, (b) $g = [2\bar{1}\bar{1}0]$, (c) $g = [\bar{1}\bar{1}20]$, (d) $g = [\bar{1}01\bar{1}]$. The beam direction for (a) – (c) was $B = [0001]$ and $B = [\bar{1}012]$ for (d). Using the $g \cdot b = 0$ invisibility criterion in (a) – (c) allows us to identify the dislocations labelled A, B and C as partials with a respective Burgers vector of $1/3[1\bar{1}00]$, $1/3[01\bar{1}0]$ and $1/3[10\bar{1}0]$. Similarly, in (d) the stacking fault node between the partials can be seen. (Reproduced from Ref. [118] with kind permission from the authors.)

the $g \cdot b = 0$ invisibility criterion [118] (see also Appendix B). Experimental results agree on stacking fault energies of around 15 and 3 mJ/m^2 for 4*H*- and 6*H*-SiC respectively [119, 120]. These numbers accord well with recent theoretical calculations by Miao et al. [121]. Further, the calculated energy of an intrinsic fault in 3*C*-SiC was found to be negative (-3.4 mJ/m^2 [122] and -0.14 mJ/m^2 [121]). Since a stacking fault in cubic material is nothing else but a hexagonal inclusion, this negative stacking fault energy reflects 3*C* being a less stable polytype than for example 4*H* or 6*H*. Compared to diamond with a stacking fault energy around 300 mJ/m^2 (Table 4.2), all values reported for SiC are absolutely miniscule. Hence extended stacking faults are relatively easy to form. This results in devastating effects on SiC devices, which will be discussed in the next section.

Table 5.1: The experimental lattice constants (a_0, c_0/p) and the excitonic band gap E_g of different SiC polytypes. The unit cells for 3*C* and 2*H* are given in Fig. 3.1, those of 4*H* and 6*H* are similar to that of 2*H*, but include not two, but four and six bilayers along [0001] respectively. c_0 is divided by the number p of bilayers to facilitate comparison. All values are taken from Ref. [107]. Those of Ref. [123] are almost identical.

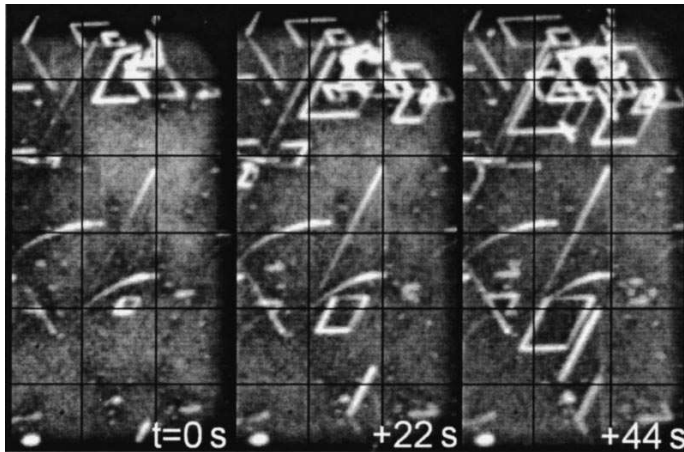|		3*C*	2*H*	4*H*	6*H*
a_0	(Å)	4.360	3.076	3.073	3.081
c_0/p	(Å)	—	2.524	2.513	2.520
E_g ($T = 4$ K)	(eV)	2.39	3.33	3.26	3.02

Figure 5.3: A time sequence of plan-view electroluminescence images of partial dislocation motion in 4*H*-SiC at 30°C and under a current density of 7 A/cm^2. The reference grid size is 100 × 100 μm. The image shows only the spectral range of 700 ± 20 nm, revealing the partial dislocations as bright lines and threading dislocations as bright spots. (Reproduced from Ref. [115] with kind permission from the authors.)

5.1.2 The degradation of SiC PiN diodes under forward-bias

Despite all afore mentioned advantages of SiC for high-power applications, a major problem in SiC power-device technology remains unsolved: Recent experiments have shown, that SiC PiN diodes[2] degrade during forward-biased operation of the device even at moderate current densities [112, 113, 114, 115], sometimes below 1 A/cm^2 [124]. This degradation means a considerable drop in the forward voltage, rendering the device useless after few days of constant operation.

The voltage drop is accompanied by the formation, propagation and growth of stacking faults of triangular and sometimes rhombic shape [114]. Very recently the stacking fault edges, which lie along $\langle 11\bar{2}0 \rangle$ directions, were identified by means of transmission electron microscopy and X-ray topography to be Shockley partials with Burgers vectors of $\frac{1}{3} \langle 1\bar{1}00 \rangle$-type [124, 125]. Fig. 5.4 shows high-resolution TEM images of presumably 30° Shockley partials bordering a stacking fault in 4*H*- and 6*H*-SiC diodes.

Since neither the mechanical stresses nor the temperature are high enough to overcome the barrier to dislocation glide motion, it is believed that a *recombination-enhanced dislocation glide* mechanism (REDG) is responsible for the observed effect[3]. The REDG mechanism requires non-radiative electron-hole recombination sites along the dislocation line. The released recombination energy then has to be redirected to assist the formation and migration of kinks at the dislocation and thus substantially lower the thermal glide activation energy [128].

Galeckas et al. [115] used *optical emission microscopy* (OEM) in spectrally selective

---

[2] A PiN diode is a diode without a direct *p-n*-junction, but *p*- and *n*-doped material are separated by an undoped region (intrinsic). When unbiased, it has a wide forbidden zone. The *p*- and *n*-type regions are usually heavily doped. Under forward-bias this allows enough carriers from the doped zones to diffuse into the forbidden zone to enable sufficient current flow.

[3] Weeks et al. [126] first proposed a theory of recombination-enhanced defect reactions. If a defect, which is involved in a defect reaction (e.g., annealing reactions, defect diffusion), induces a deep level in the electronic band gap, then the reaction might be enhanced by an electron–hole recombination associated with that deep level. The theory was later refined by Sumi [127] and first applied to dislocation glide motion by Maeda and Takeuchi [128].
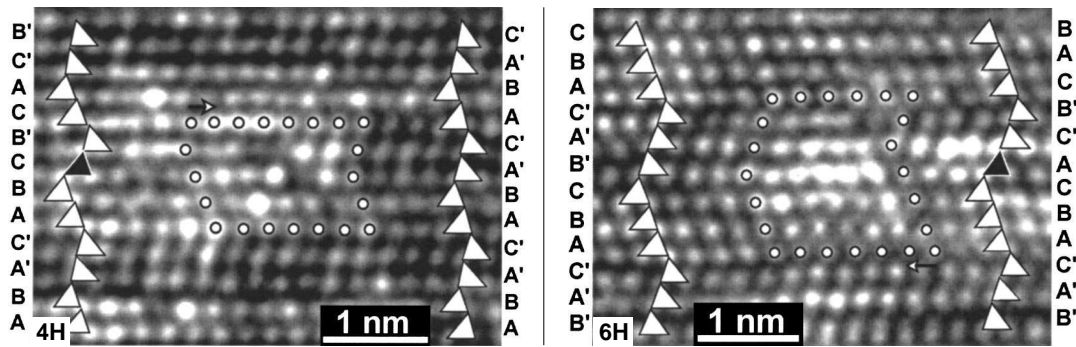
Figure 5.4: HRTEM images of partial dislocations in 4*H*- and 6*H*-SiC diodes. The basal plane lies horizontally and the stacking sequences are given at the left and right edges. Bulk stacking is indicated by white triangles and the fault by a black triangle. Both images are centred on the partial dislocation core, which is surrounded by a Burgers circuit (indicated by white dots). *Left:* 30° partial dislocation in 4*H*-SiC. *Right:* 30° partial dislocation in 6*H*-SiC. (Reproduced from Ref. [124] with kind permission from the authors.)

plan-view to demonstrate stacking fault expansion under low forward current. As can be seen in Fig. 5.3, the moving Shockley partials appear as bright lines, indicating the presence of radiative recombination centres on the dislocation line. Based on spectral analysis, a radiative recombination around 2.8 eV at the stacking fault and an also radiative recombination around 1.8 eV on the moving Shockley partials is identified. The findings for the stacking fault are in good agreement with recent theoretical work [121, 129]. Further the activation energy for partial glide motion under forward-bias is found to be 0.27 eV [115], substantially lower than the estimated value of 2.5 eV obtained from the temperature-dependence of the yield stress [130] or the brittle-to-ductile transition temperature [115, 131]. This supports a REDG mechanism as mentioned above and preliminary suggests a non-radiative recombination of approximately 2.2 eV at the Shockley partials [115].

If no solution is found to prevent the formation and propagation of stacking faults, then this might prove fatal for the development of a SiC-based high-power switching technology. In fact the Swiss-Swedish company ABB (Asea Brown Boweri) recently retreated from the SiC semiconductor market for exactly this reason.

5.1.3 Earlier theoretical work

As mentioned in the previous section, stacking faults play an important role in device degradation in SiC. Their energy and electronic structure have been investigated by means of DFT-based calculations recently [121, 122, 129]. However, only very little theoretical research has been carried out on the basal plane dislocations, which border the intrinsic faults and whose motion leads to stacking fault growth. In the literature the only DFT-based work on these dislocations is that of Sitch et al. [132], who modelled thermally activated dislocation motion in SiC. *However, their work was restricted to the 90° glide partial in 3C-SiC only.* Hence for a better understanding of the recombination-enhanced glide motion in SiC devices, further investigations are necessary.

5.2   Modelling bulk SiC — the elastic constants

The dislocation structures investigated in this section will be very similar to those discussed in the case of diamond in the preceding chapter. For this reason the details of modelling and convergence tests will *not* be presented again at this point. It is important however, to check the accuracy of the applied method with respect to the elastic properties of SiC, since this can and will vary with material and method, basis sets or parameters[4]. Therefore in the following the anisotropic and isotropic elastic constants of SiC will be modelled within the DFTB method.

As discussed in preceding sections, SiC exists in many polytypes. However, their difference in stacking does not change the elastic properties much. Especially among the hexagonal types we can expect the elastic constants to be almost identical. Hence, for reasons of simplicity, the calculations here are restricted to 3*C* and 2*H*, the latter being a representative of the hexagonal polytypes. Even though experimentally 2*H* plays almost no role, it is chosen here since it has the smallest unit cell of all hexagonal polytypes, allowing a rather inexpensive calculation.

Just as in the case of diamond, the independent elastic constants of 3*C*- and 2*H*-SiC were obtained by suitably deforming a conventional unit cell and calculating the respective total energies. The integration over the Brillouin zone was accomplished using a Monkhorst-Pack-optimised set of $3 \times 3 \times 3$ *k*-points [84]. Tables 5.2 and 5.3 show the results for the cubic and the hexagonal polytype alongside with an incomplete selection of experimental and previous theoretical data.

3*C*-SiC: Unfortunately, not much experimental data is available on 3*C*-SiC. In fact there seems to be no confirmed direct measurement of the full set of elastic constants. As already noticed by Lambrecht et al. [135], Lee and Joannopoulos [134] refer to "experimental data" but do not give any reference or mention how it was obtained. In later publications — including the reference work of Landolt and Börnstein [107] — these values were often cited as experimental values, even though the work of Lee and Joannopoulos [134] is purely theoretical and the experimental work cannot be traced back from that reference. Hence their origin remains a mystery and they are not quoted here[5].

Instead in Table 5.2 the empirical tight-binding results are given alongside with calculations using Tersoff potentials (T), full-potential linear-muffin-tin-orbital (LMTO) calculations and valence force fields (VFF). The Tersoff results are remarkably close to the much less approximate LMTO calculations. Overall however, the VFF approach shows the best agreement with experiments — especially if one compares μ and ν. The empirical tight-binding calculations (ETB) as well as the density-functional based tight-binding approach of this work (DFTB) deviate especially in the off-diagonal constant c_{12} — a trend already observed in the case of diamond (Table 4.1). The experimental values given for c_{11}, c_{12} and c_{44} have been derived by Lambrecht et al. [135] from sound velocities measured by Feldman et al. [138].

---

[4] Within the DFTB method especially the choice of the DFT reference systems can have a major influence on quality of the calculations in this respect (see Section 1.1.3 and in particular page 10).

[5] Even the original authors — who quote those results without a reference — could not help in that matter [142].

Table 5.2: Elastic properties of 3C-SiC: A comparison between calculated and experimental data. The first three rows give the independent elastic constants c_{ij} of 3C-SiC. Those values for the shear modulus μ and the Poisson's ratio ν marked with an asterisk are calculated as Voigt averages following Eq. (1.49) and (1.46). All values are given in GPa, except for ν, which is dimensionless.

	DFTB[a] (3C)	T[b] (3C)	ETB[c] (3C)	LMTO[d] (3C)	VFF[e] (3C)	Exp.[f] (3C)	Exp.[g] (3C)
c_{11}	**487**	420	363	420	428		390
c_{12}	**218**	120	209	126	165		142
c_{44}	**232**	260	149	287	246		256
μ	**193**[*]	216[*]	120[*]	231[*]	200	192	203[*]
ν	**0.241**[*]	0.130[*]	0.300[*]	0.116[*]	0.187	0.168	0.153[*]

[a]Values obtained using the DFTB method.

[b]Theoretical results of Tersoff [133] using semiempirical interatomic potentials.

[c]Theoretical results of Lee and Joannopoulos [134] using an empirical tight-binding scheme.

[d]Theoretical results of Lambrecht et al. [135] using a full-potential LMTO approach.

[e]Theoretical work of Mirgorodsky et al. [136] in a valence force field approach.

[f]Experimental data by Carnahan [137].

[g]Data derived from sound velocities measured by Feldman et al. [138]. For details see Ref. [135].

Table 5.3: Elastic properties of hexagonal SiC: A comparison between calculated and experimental data. The first five rows give the independent elastic constants c_{ij} of the respective polytype. Those values for the shear modulus μ and the Poisson's ratio ν marked with an asterisk are calculated as Voigt averages following Eq. (1.50) and (1.46). All values are given in GPa, except for ν, which is dimensionless.

	DFTB[a] (2H)	VFF[b] (2H)	Exp.[c] (6H)	Exp.[d] (4H)	Exp.[d] (6H)	Exp.[e] (4H/6H)
c_{11}	**563**	520	502			501
c_{12}	**193**	145	95			111
c_{13}	**140**	89				52
c_{33}	**672**	585	565	605	565	553
c_{44}	**162**	170	169			163
μ	**190**[*]	201				194[*]
ν	**0.242**[*]	0.174	0.207	0.212	0.212	0.161[*]

[a]Values obtained using the DFTB method.

[b]Theoretical work of Mirgorodsky et al. [136] in a valence force field approach.

[c]Experimental data by Arlt and Schodder [139], average of two measurements.

[d]Experimental data by Karmann et al. [140] by measuring piezoelectric properties.

[e]Experimental data by Kamitani et al. [141] (Brillouin scattering).

Hexagonal SiC: Unlike the case of the cubic polytype, for 4*H*- and 6*H*-SiC plenty
of experimental data is available. However, the experimental determination of c_{13}
seems to be a problem, and the only value given here is derived from measurements
on only one particular sample, as explained in Ref. [141]. In the same work, no ma-
jor differences between the two hexagonal polytypes 4*H* and 6*H* could be found.
The difference between 4*H* and 6*H* in Ref. [140] might be explained by the different
temperature of the sample: The value for 6*H* was determined at room temperature,
whereas that for 4*H* was measured at 20 K. If this reflects a general trend, then the
elastic constants of SiC depend considerably on the temperature, and lower temper-
atures lead to larger values. This can then explain, why many calculations — usu-
ally zero-temperature calculations — give larger elastic constants than determined
in experiments.

As mentioned in Section 4.2, the relatively large deviations in ν do not influence the
accuracy of dislocation modelling much, since the important quantities are μ, $1 \pm \nu$
and $1 - 2\nu$. As can be seen in Tables 5.2 and 5.3 the deviation of the DFTB result
from the experimental value is negligible in the case of μ. However, the error in ν
cannot be quantified precisely since the experimental results are not coherent: The
experimental values themselves differ from each other by as much as 40 %.

5.3   Straight Shockley partials in the basal plane

As pointed out in Section 5.1.2, the technologically relevant dislocations in the basal
plane of SiC are the 30° and 90° Shockley (glide) partial dislocations bordering stack-
ing faults. Hence the investigations presented here will be entirely restricted to those
partials.

5.3.1   Core structures

Since the basal plane in all hexagonal polytypes and the {111} planes in 3*C*-SiC are
identical except for the surrounding crystal stacking, the core structures of disloca-
tions whose line direction and Burgers vectors lie in these planes can be expected to
be identical within the accuracy of the methods employed in this work. So at least in
terms of structures and line energies it would be sufficient to limit all investigations
to one polytype only. However, to be on the safe side, all core structures will be mod-
elled in both 3*C*-SiC and 2*H*-SiC. The latter polytype is chosen as a representative for
hexagonal stacking for reasons of simplicity, even though experimentally 2*H* plays
no significant role. These two polytypes allow us, to look at the (basal) glide plane
embedded in either a locally cubic or hexagonal stacking sequence respectively[6].

The structural optimisation was performed with the DFTB method in supercell-
cluster hybrid models of similar dimensions to those described in Section 4.2.2. All
structures and also energies presented in the following were obtained *without* the

---

[6] Possible glide planes in 4*H* and 6*H* are locally embedded in cubic stacking of the nearest and
next-nearest planes on one side and hexagonal stacking on the other. Further in 6*H*, glide planes with
locally cubic stacking on both sides are possible (Fig. 5.1).

self-consistent-charge extension of DFTB[7]. Further all calculations were performed at the Γ-point only, which — as demonstrated in Section 4.2.2 — is sufficient. Details on the structural optimisation process can be found in Section 4.3.1.

As already mentioned in Section 3.1.2, within compound semiconductors each dislocation core with an edge component can exist in two variants. Depending on the atom type of the terminating line of atoms, we will speak of a silicon core or a carbon core. This combined with the three low energy glide partial cores — 30° and single- and double-period reconstruction of the 90° partial — allows six low energy core structures. Fig. 5.5 shows the optimised structures in the 3*C* polytype. The following paragraphs give a brief description:

The 30° glide partial (Fig. 5.5 (a),(d)): Very similar to the case in diamond, the 30°glide partial reconstructs forming a line of bonded atom pairs. As given in Table 5.4, for the silicon core the Si–Si reconstruction bond length is almost exactly like in bulk silicon. The equivalent C–C reconstruction bonds in the the carbon core however, are stretched by 17 % compared to bulk diamond.

The 90° glide partial (SP) (Fig. 5.5 (b),(e)): In the single-period structure of the 90° glide partial, reconstruction bonds are formed which connect the faulted with the unfaulted region of the glide plane. In the silicon-terminated core these bonds are of comparable length to bulk silicon. Just as at the 30° partial, at the carbon core the reconstruction bonds are very stretched (Table 5.4).

The 90° glide partial (DP) (Fig. 5.5 (c),(f)): This double-period structure can be obtained from the single-period structure by introducing alternating kinks. For both the silicon and the carbon core this results in a shortening of the reconstruction bonds, leaving the Si–Si bonds slightly compressed compared to bulk silicon. The C–C bonds, however, still appear about 10 % stretched compared to bulk diamond.

The undissociated dislocations are not shown. Their structures are similar to the equivalent structures in diamond. Although as for the partials presented here, the carbon- and silicon-terminated dislocation cores differ from each other in reconstruction bond lengths and angles.

---

[7] Even though in contrast to diamond, silicon carbide is a rather polar compound semiconductor, the self-consistent-charge variant of the DFTB method (SCC-DFTB, see Section 1.1.3) gave no substantial difference in structure or relative energies compared to the standard variant. This might be due to the rather bulk-like bonding situation for all glide dislocation cores.

Table 5.4: Relative reconstruction bond lengths of the Shockley partials in 3*C*-SiC. All bond lengths are given relative to the respective bulk bond length in silicon or diamond, as obtained within the DFTB method.

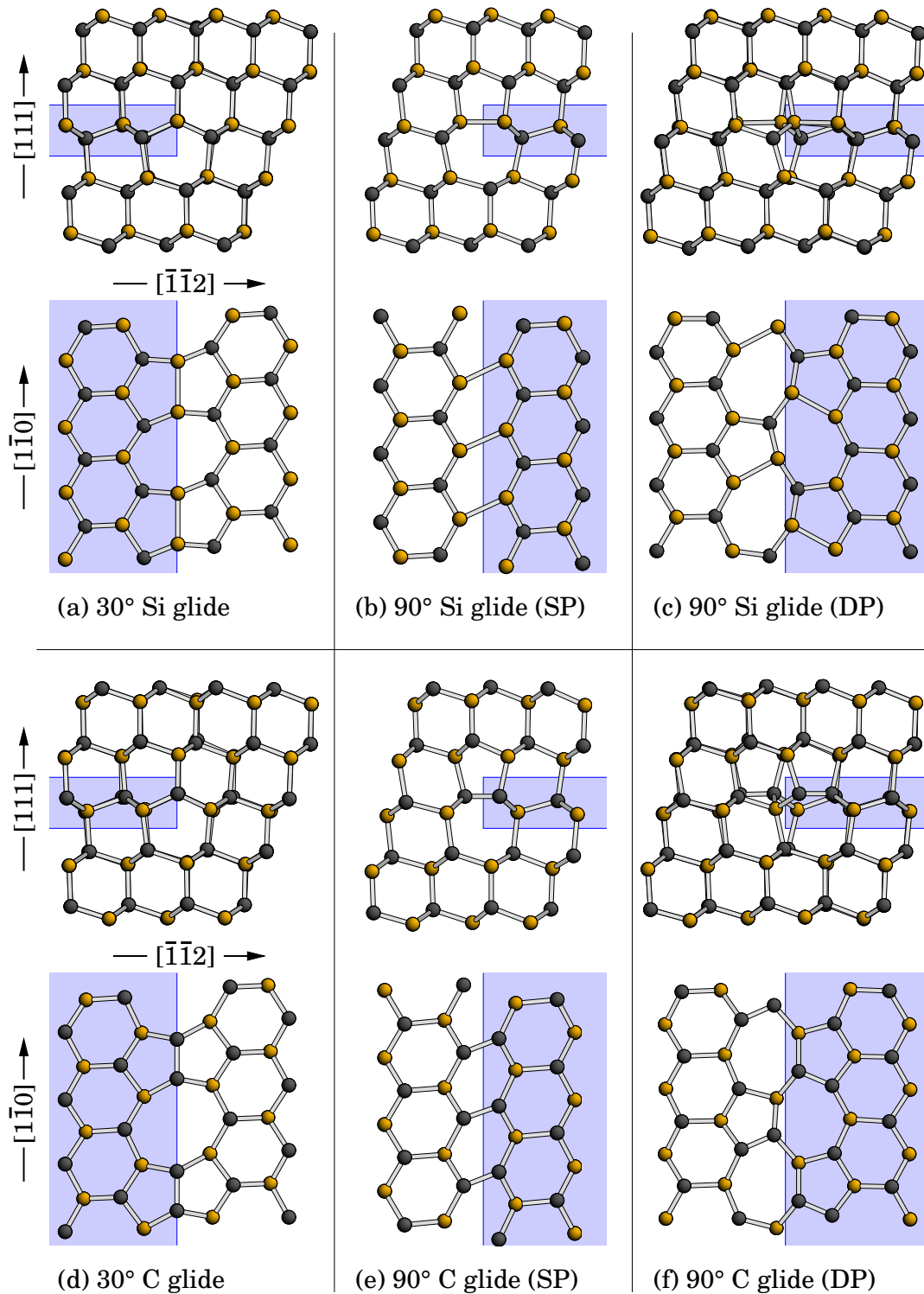|		30° glide	90° (SP) glide	90° (DP) glide
Si–Si	(bulk Si)	+0.7 %	−0.4 %	−1.4 %
C–C	(bulk diamond)	+17.0 %	+14.6 %	+9.7 %

Figure 5.5: The relaxed core structures of the Shockley partials in the {111} plane of 3*C*-SiC. (a) – (c) show the silicon-terminated partials. (d) – (f) show the carbon-terminated partials. For each structure, the upper figure shows the view along the dislocation line projected into the (1$\bar{1}$0) plane, and the lower figure shows the (111) glide plane. The region of the intrinsic stacking fault accompanying the partials is shaded. In case of the 90°partial SP and DP denote the single-period and double-period core reconstruction.

Figure 5.6: The relaxed core structure of a 30° carbon Shockley partial in 2*H*-SiC. *Left:* The view along the dislocation line projected into the (11$\bar{2}$0) plane. *Right:* The view projected into the (0001) basal plane. The region of the intrinsic stacking fault accompanying the partial is shaded.
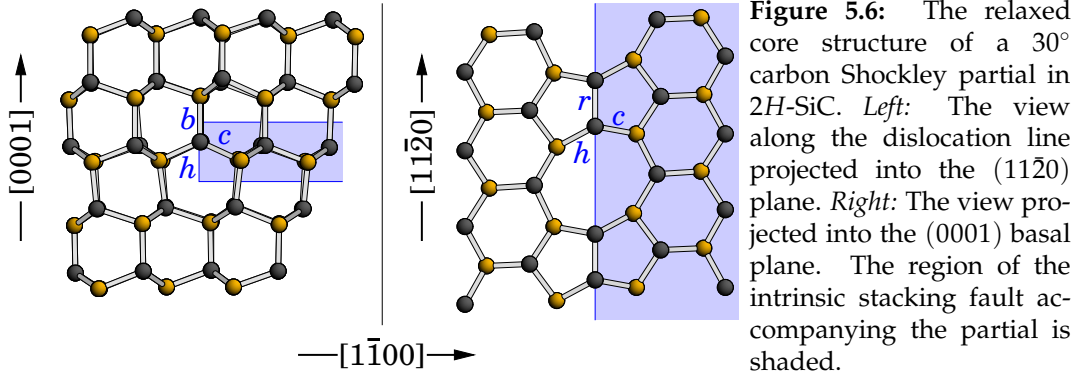
Fig. 5.5 shows the partial dislocations in the cubic stacking sequence only. All structures however, have also been optimised in the hexagonal 2*H* stacking. Fig. 5.6 gives the core structure of the carbon-terminated 30° partial in 2*H*-SiC as an example. Compared to its equivalent structure in 3*C*-SiC (Fig. 5.5 (d)), where the stacking fault was given as the hexagonal region of the glide plane (left half), here the situation is reversed and the cubic region (right half) defines the stacking fault. The actual bond angles and lengths within the glide plane appear to be identical within the accuracy of the methods applied. Table 5.5 gives the differences in selected bond lengths for all investigated partial core structures. The largest deviation is found in the case of the carbon-terminated 90° double-period partial. But even here the difference of approximately 0.4 % is negligible, if one keeps in mind that the respective bond is 10 % stretched compared to bulk diamond.

As a consequence for all partials, the core structures seem to be almost entirely independent of the surrounding crystal stacking[8]. The same can be expected for the core energies and will be tested in the next section.

---

[8] 3*C* and 2*H* give an either purely cubic or purely hexagonal stacking environment. Since the differences between the corresponding core structures in both stackings are miniscule, one can expect them to be just as small in other stacking sequences, which are (at least locally) of mixed cubic and hexagonal nature, like 4*H* and 6*H*.

Table 5.5: Differences in core bond lengths of the Shockley partial dislocations in 2*H*-SiC compared to 3*C*-SiC. As shown in Fig. 5.6, *r* denotes the reconstruction bond. *c* and *h* are the bonds connecting the central line of atoms with the cubic and hexagonal region of the glide plane respectively. *b* connects the central line of atoms with the bulk region in the adjacent plane. $\bar{r}$, which is given for the double-period 90°partials, gives the average of the two slightly different reconstruction bonds in each structure. The numbers given represent the change in bond length from 3*C* to 2*H* for Si-core and C-core structures.

	30° glide				90° (SP) glide	90° (DP) glide
	r	*c*	*h*	*b*	*r*	$\bar{r}$
Si	+0.08 %	+0.06 %	+0.09 %	+0.10 %	−0.17 %	−0.10 %
C	+0.08 %	+0.08 %	+0.08 %	+0.17 %	−0.25 %	−0.39 %

Table 5.6: The calculated energy factors $k(\beta)$ for 30° and 90° partials in cubic (C) and hexagonal SiC (H). Both isotropic and anisotropic elasticity theory are applied.

		30°		90°	
		C	H	C	H
$k(\beta)$ isotropic[a]	(GPa)	208 (212)	205 (203)	254 (240)	251 (231)
$k(\beta)$ anisotropic[b]	(GPa)	195 (191)	191 (191)	249 (231)	246 (227)

[a]Evaluated using Eq. (2.5). The first number is given by the elastic constants obtained in the DFTB method, the number in brackets by the experimental constants based on Feldman et al. [138] (3C) and Kamitani et al. [141] (4H/6H) (for both see Table 5.2 and 5.3).

[b]Evaluated using linear *anisotropic* elasticity theory as suggested in [93] and [41].

5.3.2   Core energies

Before discussing the core energies it is useful to investigate the influence of the polytype on the energy factors for the two partials, as these play an important role in obtaining the core energies following Eq. (2.4). With the elastic constants calculated in Section 5.2, the energy factors can be obtained easily using Eq. (2.5) or — a bit more elaborately — by means of anisotropic theory as described in [41]. Table 5.6 gives the energy factors obtained for the 30° and 90° partial dislocations in both 3C- and 2H-SiC, based on the DFTB elastic constants. For comparison the values based on the experimental constants for 3C and 4H/6H are shown in brackets. The difference between the energy factors in cubic and hexagonal stacking appears to be negligible in the case of the DFTB energy factors. Only the isotropic values based on experiments differ by up to 5 %. One has to keep in mind though, that the observed elastic constants for cubic and hexagonal stacking were obtained in very different experiments[9]. Overall the absolute values obtained in the anisotropic calculation are slightly lower than the isotropic values. The same effect was already observed for dislocations in diamond (see Table 4.3).

In Section 4.2.2 the spatial distribution of the formation energy was used to obtain the core radius R_c and the core energy E_c of a dislocation in comparison with the elastic energy as given in continuum theory (Eq. (2.4)). Now the same procedure will be applied to the Shockley partials in 3C- and 2H-SiC. However, since SiC is a compound semiconductor, one has to be careful when speaking of formation energies: In this section the radial formation energies and also the resulting core energies are all given relative to a stoichiometric crystal[10].

---

[9] Taking μ and ν of Carnahan [137] as the cubic isotropic constants instead of Feldman et al. [138], reduces the difference between the cubic and hexagonal case to almost zero.

[10] As mentioned in Section 4.2.2, within the DFTB method for each atom in a modelled structure an atomic total energy can be defined easily. The "formation energy" of a particular atom in that structure is then obtained by subtracting the total energy of the same type of atom in a perfect — and hence *stoichiometric* — SiC crystal (Eq. (4.2)). In terms of chemical potentials, here those chemical potentials, which correspond to a stoichiometric crystal growth for silicon and carbon, are assumed.
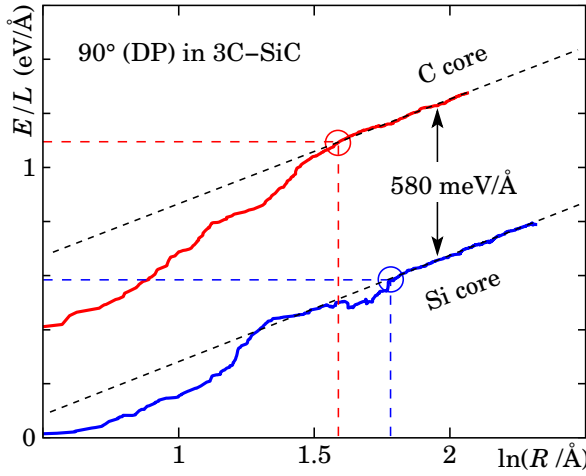
Figure 5.7: The radial formation energy of the double-period reconstruction of the 90° glide partial in 3C-SiC. $E_f(R)/L$ is given by Eq. (4.3). The gradient of the dashed line is obtained as a fit to the linear part of the energy, and corresponds to the elastic energy in Eq. (2.4). The fit yields $k(\beta) = 244$ GPa for the C-terminated core and $k(\beta) = 250$ GPa for the Si-terminated core. A circle marks the approximate point for each core, where for $R \rightarrow 0$ the energy deviates considerably from the straight line. This point defines the respective core radius R_c and the core energy E_c.

As an example, Fig. 5.7 shows the resulting radial formation energy per unit length for the double-period 90° partial dislocations in 3C-SiC (Fig. 5.5 (c),(f)). The C core structure is found about 580 meV/Å higher in line energy than the Si core.

The fitted energy factors, core radii and core energies for all investigated Shockley partials are given in Table 5.7. For all partials the energy factors obtained in the fitting procedure agree with those found in elasticity theory (given in Table 5.6). On average the agreement is better with anisotropic elasticity theory. Compared to diamond, all core energies are very small.

Some particular results will be discussed in the following paragraphs.

Table 5.7: The calculated energy factors $k(\beta)$, core radii R_c and core energies E_c of the basal plane Shockley partials in SiC. The corresponding structures can be found in Fig. 5.5 for 3C-SiC in the same sequence. To facilitate comparison between different dislocations with different core radii, the core energy E_c' corresponding to a radius of 6 Å is introduced.

		30° Si		90° Si (SP)		90° Si (DP)	
		3C	2H	3C	2H	3C	2H
$k(\beta)^a$	(GPa)	203	194	249	251	250	242
R_c	(Å)	5.3	5.5	3.9	3.4	6.0	5.2
E_c/L	(eV/Å)	0.44	0.45	0.51	0.46	0.59	0.54
$E_c'(R = 6 \text{ Å})/L$	(eV/Å)	**0.48**	**0.47**	**0.68**	**0.69**	**0.59**	**0.60**

		30° C		90° C (SP)		90° C (DP)	
		3C	2H	3C	2H	3C	2H
$k(\beta)^a$	(GPa)	194	197	242	237	244	244
R_c	(Å)	4.5	4.7	5.0	4.1	4.7	4.1
E_c/L	(eV/Å)	0.77	0.78	1.20	1.13	1.08	1.02
$E_c'(R = 6 \text{ Å})/L$	(eV/Å)	**0.86**	**0.85**	**1.27**	**1.27**	**1.17**	**1.17**

[a]Fit to $E_f(R)$ vs. $\ln(R/R_c)$ plot for $R \geq R_c$ following Eq. (2.4).

Silicon- and carbon-terminated partials: Compared with the Si core structures, the corresponding C core structures have a higher line energy: 380 meV/Å in case of the 30° partials and around 580 meV/Å in case of the 90° partials. To some extent this higher line energy might be attributed to the reconstruction bonds: In the Si core the reconstruction bonds are of about the same length as they would be in perfect bulk silicon, giving a rather low energy configuration. In the C core however, the bonds are stretched by 10 % or more compared to diamond (Table 5.4), which is energetically less favourable.

Single- and double-period reconstructions: In both polytypes the Si core double-period reconstruction of the 90° partial appears around 90 meV/Å lower in energy than the single-period structure. Similarly, the difference for the carbon-terminated structures is 100 meV/Å. With the difference that small, both reconstructions can be expected to co-exist.

Effects of the polytype: Since the polytype seems not to have any notable effect on the structure of the partials in the glide plane, one would expect the same for the energies. And in fact there appears to be no considerable difference in the line energies of equivalent partials — the maximum differences in $E'_\mathrm{c}(R = 6\ \text{Å})/L$ are of the order of 10 meV/Å, which is negligible. Larger deviations occur only for the core radii and the resulting energies. But as mentioned in Section 4.2.2, the definition of these core radii is rather vague anyway, and no strong conclusions can be drawn from it.

With the Shockley partials showing the same energies in the 3*C* and the 2*H* polytype, the same can be expected for 4*H* and 6*H* as well. Hence in terms of structure and energy the stacking sequence seems to have no impact on the partial dislocations in the basal plane of SiC.

5.4 Dislocation glide motion

Recombination-enhanced glide motion of partial dislocations presumably causes the degradation of bipolar SiC devices as described in Section 5.1.2. To gain more insight into this mechanism, it is an important prerequisite to understand the underlying atomistic processes of dislocation glide. Therefore, in this section the glide motion of the involved partial dislocations is modelled as a thermally activated process involving the formation and migration of kinks — details on this process, which assumes the absence of strong obstacles to dislocation motion, were given in Section 4.5.1. The resulting kink formation energies and migration barriers will then allow us to estimate the thermal activation energies for glide motion in SiC.

The cluster models and methods used in this section are *exactly* the same as described in Section 4.5 for dislocation glide in diamond. For this reason most of the details will not be repeated here. The calculations, which are all performed using the DFTB method, are restricted to the cubic polytype only. But as we have learnt in the preceding sections, changing the polytype has almost no impact on the dislocation structures and resulting energies in the basal plane. Hence all calculations described below will be representative for all polytypes alike.

5.4.1 The glide motion of 90° partial dislocations

Fig. 5.8 (far left panels) shows relaxed kink structures at the 90° silicon and the 90° carbon single-period partial projected into the basal plane. All atoms are four-fold coordinated. The difference in the reconstruction bond lengths of the two different partials (Si–Si or C–C bonds respectively) leads to slightly different bond angles, but qualitatively both structures are very similar and close to that found for diamond (Fig. 4.15).

The kink formation energy E_f is obtained by comparing a cluster containing a straight dislocation segment with one of the same stoichiometry containing a double-kink. The elastic kink–kink interaction energy and the stacking fault energy are subtracted following Eq. (4.8) and (4.9). Table 5.8 gives the resulting single-kink formation energies — for the silicon and the carbon partial both around 0.5 eV.

The elementary kink migration step of the right kink is depicted in Fig. 5.8, where the kink **K**, the saddle point **S** and the migrated kink **K′** are shown for both partials. In the process only two atoms move considerably and break and form new bonds. The motion of these two atoms — one silicon and one carbon atom — is used to parameterise the migration step (see Fig. 4.17). Fig. 5.9 (left) gives the corresponding two-dimensional energy surfaces[11]. The energy surfaces differ considerably for the two 90° partials. For the carbon partial the saddle point is well pronounced and the barrier can only be overcome by varying both parameters. The silicon and the carbon atom seem to be equally involved in the process. At the silicon par-

---

[11] As in Section 4.5 the energy surface is obtained at 10×10 points in the two-dimensional parameter space by relaxing the whole structure subject to constraining the two primary atoms to lie in a plane perpendicular to the connecting line between the initial and final position. In the vicinity of the saddle point **S**, the parameter mesh is refined by a factor of 10. The model used is a dislocated cluster containing a single kink.
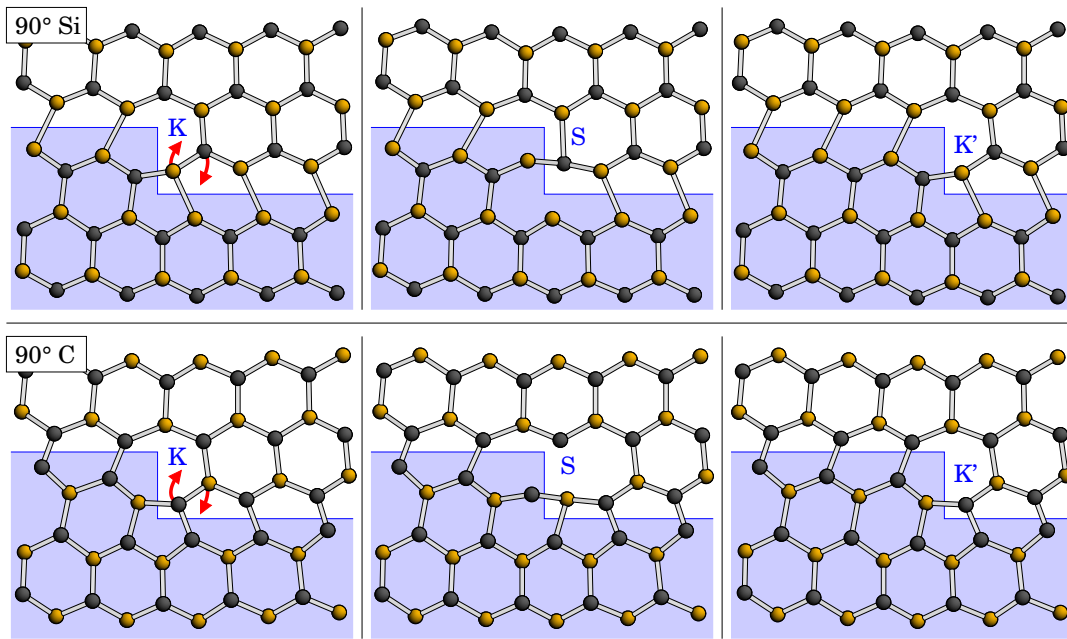
Figure 5.8: Kink migration at the 90° (SP) glide partial. The relaxed structures of the starting kink **K**, the saddle point **S** and the migrated kink **K′** are shown projected into the glide plane. The faulted region is shaded and arrows indicate the motion of the two involved atoms. *Upper panel:* The silicon-terminated core. *Lower panel:* The carbon-terminated core.

tial, however, the motion of the carbon atom clearly dominates the process[12]. The whole procedure was also applied to the two left kinks. But since the corresponding structures are approximately symmetric in the glide plane, no difference in the migration barriers could be found. As can be seen in Table 5.8, the barrier at the silicon-terminated partial is about 1 eV larger compared to the carbon-terminated partial. With diamond having a larger cohesive energy than silicon, this might seem contra-intuitive at first glance. However one has to keep in mind, that the C–C reconstruction bonds at the partial are considerably stretched compared to diamond, which lowers the energy required to break the bond.

From the kink formation energies and the migration barriers, the activation energy of the kink migration process can be evaluated following Eq. (4.7). Table 5.8 gives the results. With the kink formation energies being almost the same at both partials, the migration barrier is the deciding quantity: In a glide process controlled by the formation and migration of kinks, when no strong obstacles to dislocation motion are present, the carbon partial is clearly the more mobile 90° partial dislocation. The energy corresponding to the whole glide process of the 90° partial is schematically shown in Fig. 5.9 (right) for the first four kink migration steps, assuming short dislocation segments. There the dashed line connecting the minima represents the formation energy of the kink pair and the solid line the energy of the minimum en-

---

[12] To some extent this might be explained by the Si atom only moving by roughly 0.6 Å with respect to the surrounding crystal, whereas the C atom moves by 1.2 Å. This of course scales the two parameters accordingly. Also the different bonding situation for the two atoms will play a role: The movement of the Si atom involves the breaking and reforming of Si–Si bonds, which are rather "soft", whereas the movement of the C atom breaks and reforms "hard" Si–C bonds.
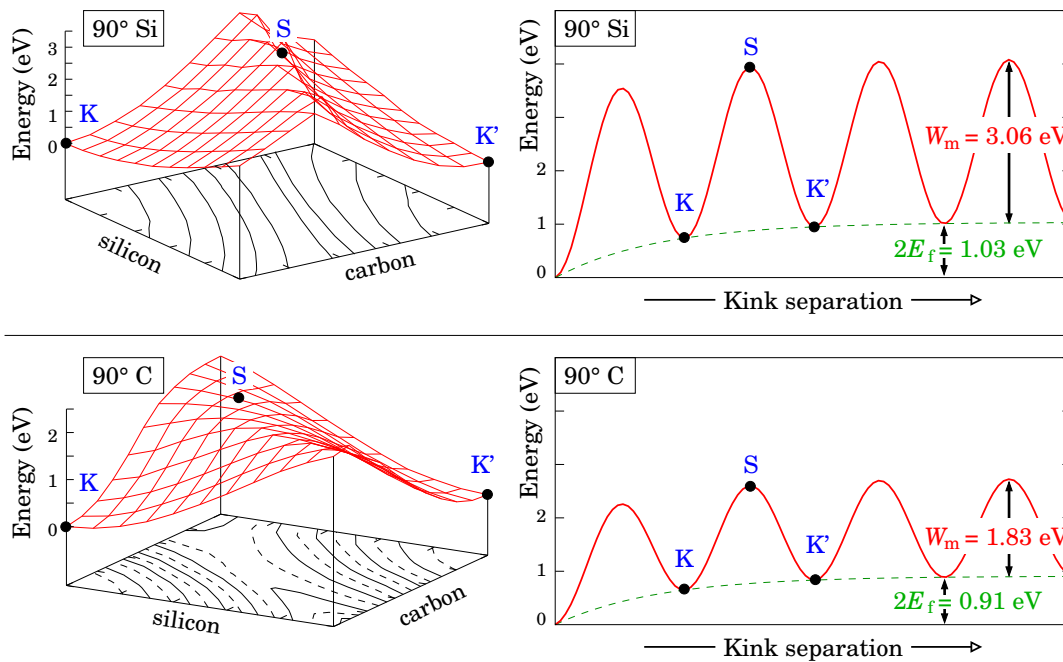
Figure 5.9: The energies and barriers of kink formation and migration at the 90° (SP) glide partial. *Left diagrams:* The energy surface of the process leading from the starting kink **K** via the saddle point **S** to the migrated kink **K'**. The two parameters are defined via the two involved atoms (see text and Fig. 4.17). *Right diagrams:* Schematic representation of the energy of the glide process. As shown in Fig. 4.14, a kink pair is formed and subsequent migration of the two kinks enlarges their separation. **K**, **S** and **K'** are labelled for an arbitrary migration step. The dashed line connecting the minima represents the formation energy of the kink pair. The energy contribution of the expanding stacking fault is not included in the graph. *Upper panel:* The silicon-terminated core. *Lower panel:* The carbon-terminated core.

ergy path for glide motion. The first few minima are considerably lower due to the attractive kink–kink interaction.

For reasons of simplicity dislocation glide was once again modelled for the single-period reconstruction of the 90° partial only. However since the elementary processes are very similar for the double-period structure, similar kink formation energies and migration barriers can be expected.

Table 5.8: Kink formation energies E_f and migration barriers W_m for the 90° Shockley partials. The resulting thermal glide activation energy $Q_{90} = 2E_f + W_m$ assumes short dislocation segments. The number in brackets gives the respective value for long segments ($E_f + W_m$).

		90° (SP) glide	(all energies in eV)
	E_f	W_m	Q_{90}
Si	0.515	3.06	**4.09 (3.58)**
C	0.455	1.83	**2.74 (2.29)**
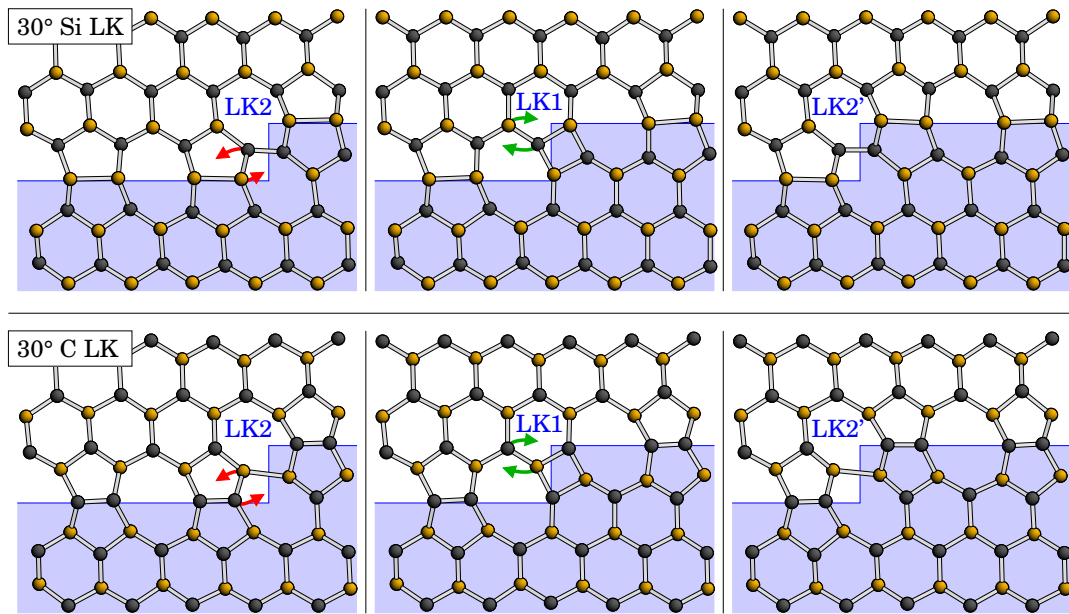
Figure 5.10: The elementary kink migration steps of left kinks at the 30° glide partial: LK2 → LK1 → LK2′. The relaxed structures of the high-energy kink **LK2**, the low energy kink **LK1** and the migrated high-energy kink **LK2′** are shown projected into the glide plane. The faulted region is shaded and arrows indicate the motion of the two involved atoms. *Upper panel:* The silicon-terminated core. *Lower panel:* The carbon-terminated core.
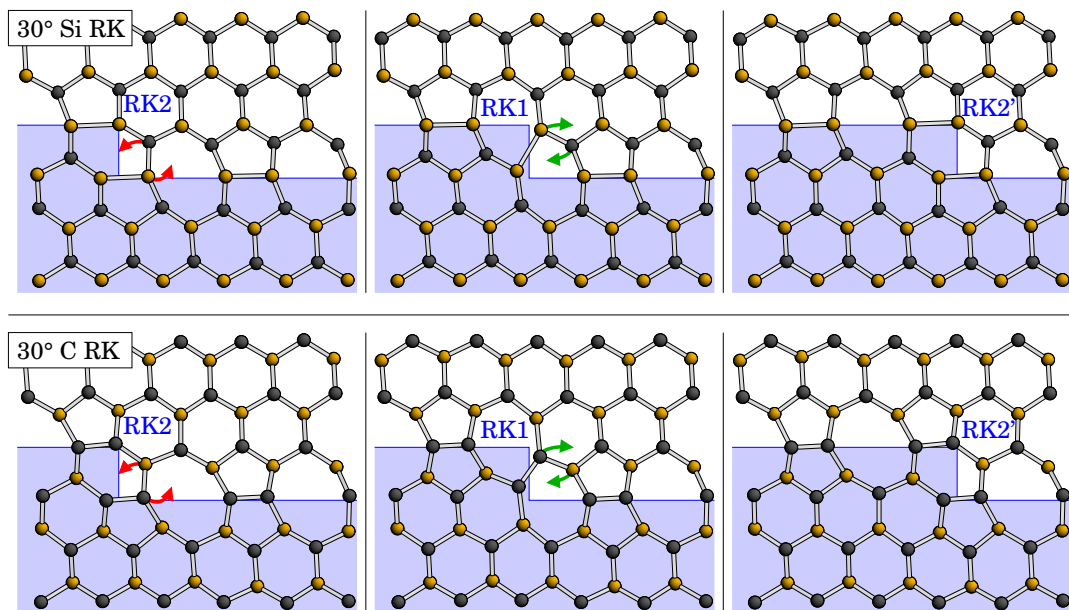
Figure 5.11: The elementary kink migration steps of right kinks at the 30° glide partial: RK2 → RK1 → RK2′. The relaxed structures of the high-energy kink **RK2**, the low energy kink **RK1** and the migrated high-energy kink **RK2′** are shown projected into the glide plane. The faulted region is shaded and arrows indicate the motion of the two involved atoms. *Upper panel:* The silicon-terminated core. *Lower panel:* The carbon-terminated core.

5.4.2   The glide motion of 30° partial dislocations

Since the growing stacking faults observed in degrading bipolar SiC devices are bordered by dislocations on four sides (Fig. 5.3), there will be both 90° and 30° partials present. Further, as already demonstrated in Section 4.5.3 for diamond, the kink structures and the mechanisms for kink migration at the 30° partial are different from those at the 90° partial, leading to substantially different glide activation barriers. Therefore in this section kink formation and migration at both 30° partials are examined in detail.

Due to the double-periodicity of the reconstructed 30° dislocation shown in Fig. 5.5 (a) and (d), two variants of left kinks (LK1 and LK2) and two variants of right kinks (RK1 and RK2) are possible. The left and middle panels of Fig. 5.10 depicts the different relaxed and reconstructed kink structures for the left kinks, distinguishing between the Si-terminated and the C-terminated partial. Fig. 5.11 shows the same for the right kinks. As can be seen, in the case of LK1 and RK1 the kink is located on a reconstruction bond, whereas LK2 and RK2 are located between two reconstruction bonds. As in the case of diamond, all atoms in all structures are fourfold coordinated. In principle the structures at the silicon core and the carbon core are very similar, with only differences due to the different character of the Si–Si and C–C bonds. The bond lengths and as a consequence also the bond angles differ between the two partials. The high-energy left kinks (LK2) appear with an interesting feature: The kink forms a reconstruction bond, which is of the opposite type to the reconstruction bonds of the respective partial — at the silicon-terminated partial with Si–Si reconstruction bonds the kink reconstructs with an "alien" C–C bond, and vice versa (Fig. 5.10 (left)).

The kink formation energies are obtained by introducing kink pairs in dislocated cluster models. The energy of the generated faulted area and the kink–kink interaction energy have to be subtracted according to Eq. (4.10) and (4.11). Since the structures of left and right kinks are very different, one cannot expect similar formation energies. Hence the modelling of kink pairs does not allow a calculation of explicit single-kink energies, but only the sums of two kink formation energies. For

Table 5.9: Kink formation energies E_f and migration barriers W_m for the 30° Shockley partials. One has to differentiate between left and right kink (LK and RK). $\Delta E_f(\mathrm{LK})$ gives the energy difference between the low energy (LK1) and the high-energy left kinks (LK2) and $\Delta E_f(\mathrm{RK})$ accordingly for the right kinks. To facilitate comparison with the 90° partials, for the 30° partials an average activation energy $\overline{Q}_{30} = (2E_f(\mathrm{LK1}) + W_m(\mathrm{LK}) + 2E_f(\mathrm{RK1}) + W_m(\mathrm{RK}))/2$ is introduced. This glide activation energy assumes short dislocation segments. The number in brackets gives the respective value for long segments.

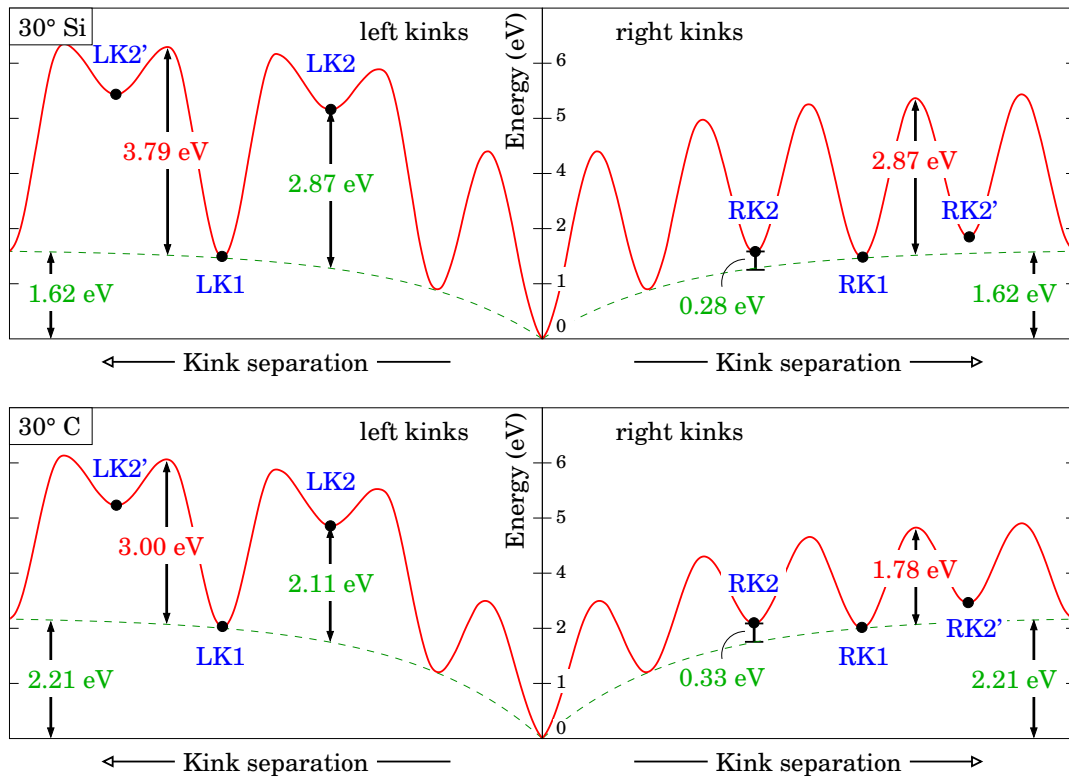|	30° glide	(all energies in eV)			
	$E_f(\mathrm{LK1}) + E_f(\mathrm{RK1})$	$\Delta E_f(\mathrm{LK})$	$W_m(\mathrm{LK})$	$\Delta E_f(\mathrm{RK})$	$W_m(\mathrm{RK})$	$\overline{Q}_{30}$
Si	1.62	2.87	3.79	0.28	2.87	**4.95 (4.14)**
C	2.21	2.11	3.00	0.33	1.78	**4.60 (3.50)**

Figure 5.12: The energies and barriers of kink formation and migration at the 30° glide partial. The graphs schematically represent the energy of the glide process: As shown in Fig. 4.14, a kink pair is formed and subsequent migration of the two kinks enlarges their separation. Since for the 30° partial the migration barriers differ for left and right kinks, the corresponding processes are shown separately. The dashed line connecting the minima represents the formation energy of the kink pair. The energy contribution of the expanding stacking fault is not included in the graph. *Upper panel:* The silicon-terminated core. *Lower panel:* The carbon-terminated core.

details see Section 4.5.3. Table 5.9 gives the resulting formation energies of the low energy kink pair $E_f(LK1) + E_f(RK1)$, as well as the extra energy ΔE_f needed if either the low energy left kink LK1, or the low energy right kink RK1, is replaced by the respective high-energy kink LK2 or RK2. With around 0.3 eV, this extra energy is particularly low for the right kinks RK2 at both partials and larger than 2 eV for the left kink LK2. The two right kinks appear with similar bond angle distortions, whereas among the two left kinks it is clearly LK2 which shows the strongest distortions ($> 20°$ at the two atoms closest to the kink). This might explain the large energy difference between LK1 and LK2.

As in the case of the 90° partial, the elementary migration steps are modelled in dislocated clusters containing a single kink. Fig. 5.10 and 5.11 show the kink migration steps LK2 → LK1 → LK2′ and RK2 → RK1 → RK2′ at the two different partials. The resulting energy surfaces for each single step look very similar to those at the 90° partial shown in Fig. 5.9 and again at the silicon partial it is the motion of the carbon atom, which dominates the process. Due to symmetry in the glide plane LK2 → LK1 and LK1 → LK2′ (or the equivalent right kink processes) are approximately twin processes and yield the same migration barriers within 1 %. Table 5.9

gives the resulting migration barriers W_m for both left and right kinks and Fig. 5.12 schematically shows the resulting energy of the migration processes of left and right kink at each partial. Overall the right kinks are more mobile.

As explained in Section 4.5.3, Eq. (4.12) describes the velocity of thermally activated glide motion with two distinct migration barriers for left and right kinks. However to allow an easy comparison, as a crude approximation an average activation barrier $\overline{Q}_{30} = (2E_f(LK1) + W_m(LK) + 2E_f(RK1) + W_m(RK))/2$ can be defined for short segments. For long segments the same expression holds, however with half the formation energies. Both values are given in Table 5.9. Among the 30° partials the carbon-core is slightly more mobile than the silicon-core partial, but the difference is less pronounced than for the 90° partial.

5.4.3 Dislocation glide motion — summary

The structures and basic processes involved in the formation and migration of kinks at basal plane partial dislocations in SiC are very similar to those found in the {111} plane of diamond. Also the general trends with respect to barrier heights and formation energies are the same. Overall, the 90° partial is the thermally more mobile species. Further, the purely thermal mobility of carbon partials in the absence of strong obstacles appears to be higher than that of silicon partials, and the velocities of undecorated neutral partials are largely controlled by their kink migration energy.

Comparison with experiments: Experimental results in 6*H*-SiC give activation energies in the range of $Q = 2.1$ to 4.8 eV [143, 144]. This is reasonably close to the activation energies found in this work, which range from 2.3 eV for the carbon-terminated 90°partial to 4.1 eV for the silicon-terminated 30°partial — assuming the glide process to involve the motion of long dislocation segments.

The carbon partials being the more mobile species is in contradiction to the conclusions of Pirouz and Yang [145]. However, LACBED (**l**arge **a**ngle **c**onvergent **b**eam **e**lectron **d**iffraction) experiments seem to indicate, that the Si partials in 3C-SiC are smooth whereas the C partials are zig-zagged [146], suggesting they are pinned by obstacles. In this case the glide process would not mainly be controlled by the formation and migration of kinks, but by the energy needed for the pinning and unpinning at the obstacles.

Comparison with independent theoretical work: In comparison with earlier DFT-based results for the 90° (SP) partials in SiC by Sitch et al. [132], the kink migration barrier found here is about 0.4 eV higher at both the silicon and the carbon partial. This might result from the use of much larger cluster models in this work, which give a more accurate representation of the stress response of the surrounding bulk material. In addition, in Ref. [132] the migration barrier was deduced from the barrier to double-kink formation, rather than kink expansion as modelled here. The formation energy of the smallest double-kink at the C partial is 0.4 eV lower than found in the present work. A more striking difference is found for the Si partial. In this work the kink formation energy appears to be similar to that at the C core, whereas in Ref. [132] it is negligible. Nevertheless, both calculations conclude that the thermal activation barrier for the carbon partial is lower than that of the silicon partial.

5.5   Electronic structure calculations

In this section the electronic structure of straight and kinked Shockley partial dislocations in the basal plane of SiC will be investigated.

The band structures of straight Shockley partials: In Section 5.3 it was shown that the effects of the stacking sequence of SiC on the atomic structure and line energy of the Shockley partials is negligible. Hence one could jump to the conclusion that the same holds as well for the electronic structure of the partials. However, as can be seen in Table 5.1, the electronic band gap varies from 2.39 to 3.33 eV, depending on the polytype [107, 123]. Also, the band structures themselves vary with the polytype. If the partial dislocations induce defect states in the band gap, then a wider band gap might allow states to appear in the gap, which in case of a smaller band gap would disappear into either the valence or the conduction bands. As a consequence, the electronic structure of the partials has to be investigated at least for the two extremes, which are once again 3C-SiC with a band gap of 2.39 eV and 2H-SiC with a band gap of 3.33 eV. Therefore, in this section the projected band structures of the discussed Shockley partials in 3C- and 2H-SiC are modelled within the DFT-pseudopotential approach (AIMPRO, see Section 1.1.2). As already described in Section 4.7, the band structure calculations are performed in smaller hybrid models than those for the calculation of the core energies — for SiC the model size is on average 110 atoms. The structures are relaxed with the AIMPRO code, using a set of 2 k-points along the the axis in k-space, which corresponds to the dislocation line direction. To then obtain the projected electronic band structure of the relaxed structure, the Kohn-Sham eigenstates are calculated at 21 different k-points along the same axis.

Fig. 5.13 shows the projected band structures of bulk 3C- and 2H-SiC as well as 2H-SiC containing an intrinsic stacking fault. The calculated width of the gap of bulk
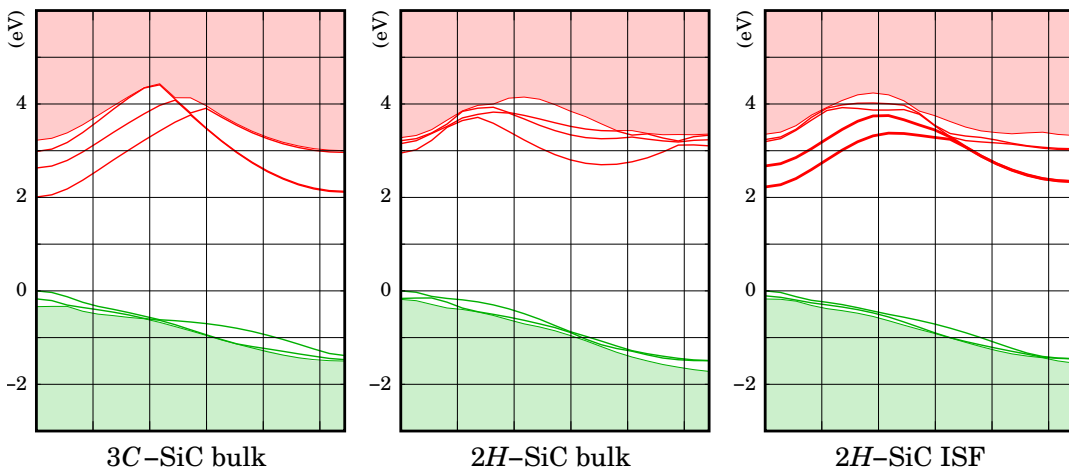
Figure 5.13: The band structures of bulk and faulted SiC projected onto the k-space axis corresponding to the basal plane dislocation line direction ([1$\bar{1}$0] in 3C and [11$\bar{2}$0] in 2H with respective periodicity lengths of $a_0/\sqrt{2}$ and a_0). Only those bands which are close to the gap are drawn, the rest of the valence and conduction band regions are shaded. The origin of the Brillouin zone is at the far left of each band structure.
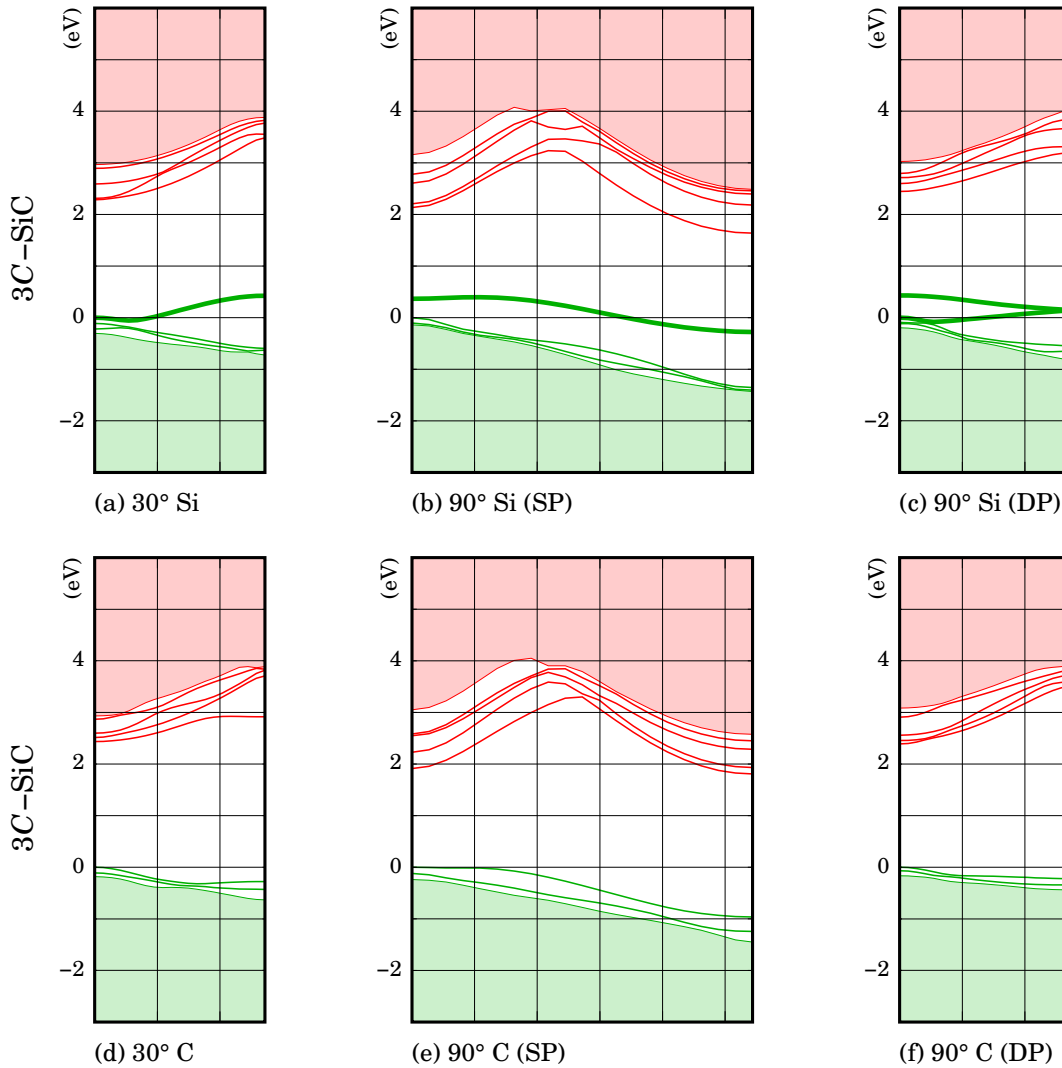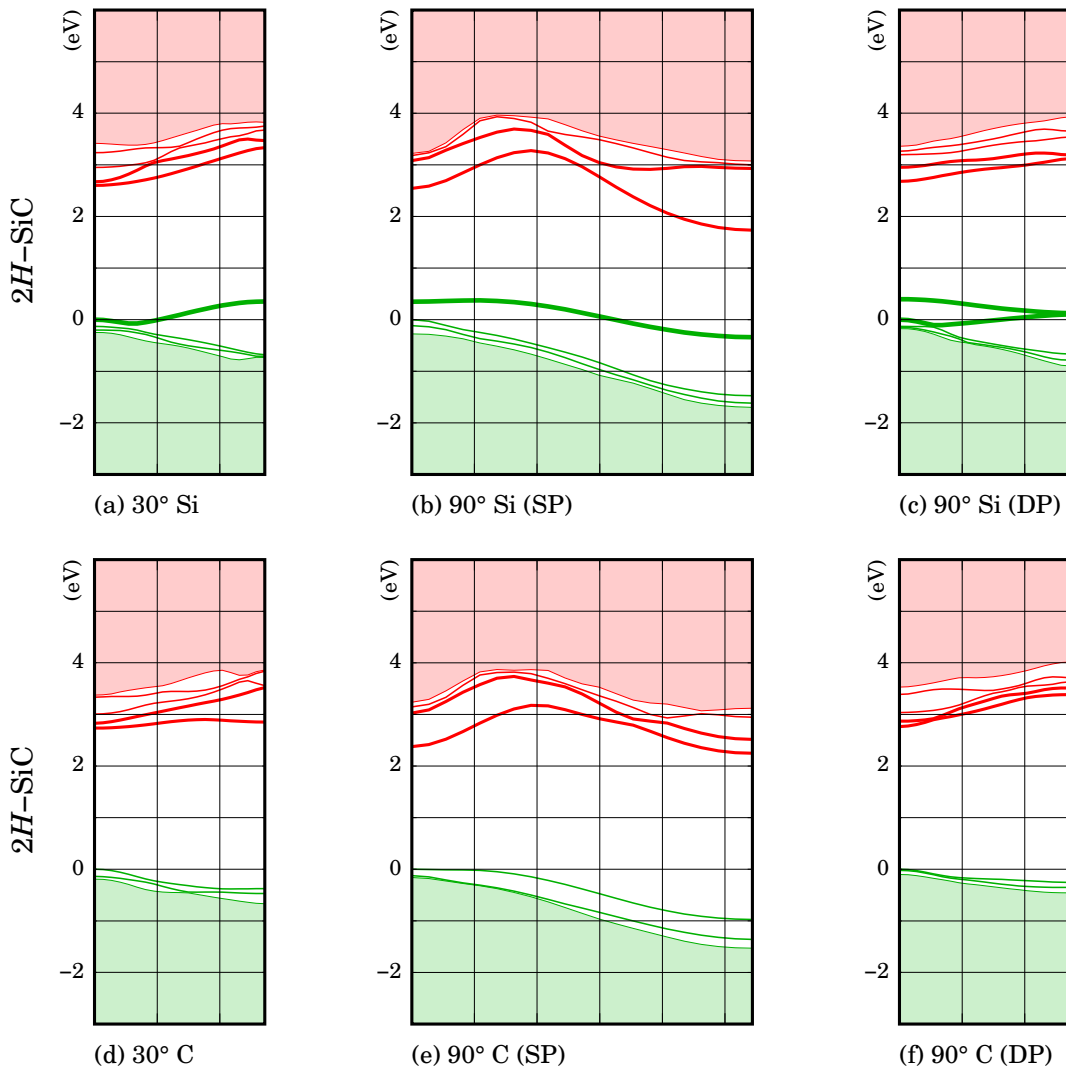
Figure 5.14: The band structures of glide partials in 3*C*-SiC projected onto the *k*-space axis corresponding to the dislocation line direction $[1\bar{1}0]$ (with a periodicity length of $a_0/\sqrt{2}$ for (b) and (e) and $\sqrt{2}\,a_0$ for all others). (a) – (c) show the band structures of the silicon-terminated partials and (d) – (f) show those of the carbon-terminated partials. Fig. 5.5 gives the corresponding core structures in the same sequence. Only those bands which are close to the gap are drawn, the rest of the valence and conduction band regions are shaded. The origin of the Brillouin zone is at the far left of each band structure. Being double-period reconstructions, (a), (c), (d) and (f) appear with a Brillouin zone of half width.

SiC deviates from experimental results. Two competing effects have to be considered: Commonly in LDA the band gap is *under-estimated*. On the other hand in the chosen supercell-cluster hybrid models the wavefunction is confined with respect to two dimensions, leading to a widening of the gap. Both effects compensate to some extent, yielding a gap just 18 % too small for both the 3*C* and 2*H* polytype. As predicted by Käckell et al. [122], the intrinsic stacking fault in 3*C*-SiC does not introduce any deep electronic gap states and therefore its band structure is not displayed. In 2*H*-SiC however, an intrinsic fault represents a layer of cubic inclusion in the lattice, so one would expect empty and localised gap states near the conduction

Figure 5.15: The band structures of glide partials in 2*H*-SiC projected onto the ***k***-space axis corresponding to the dislocation line direction [11$\bar{2}$0] (with a periodicity length of a_0 for (b) and (e) and 2 a_0 for all others). (a) – (c) show the band structures of the silicon-terminated partials and (d) – (f) show those of the carbon-terminated partials. Only those bands which are close to the gap are drawn, the rest of the valence and conduction band regions are shaded. The origin of the Brillouin zone is at the far left of each band structure. Being double-period reconstructions, (a), (c), (d) and (f) appear with a Brillouin zone of half width.

band minimum (CBM). Indeed localised bands are observed, which reach as far as approximately 0.5–0.6 eV below the CBM (see Fig. 5.14, right panel). In terms of energetic position and dispersion, these bands remarkably resemble the lowest conduction bands of the 3*C* polytype. Similar bands have been predicted by Miao et al. [121] for intrinsic faults in 4*H* ($\sim$ 0.3 eV below the CBM) and by Iwata et al. [129] in 4*H* and 6*H* ($\sim$ 0.2 eV below the CBM). Of course the supercell-cluster hybrids used here, which are optimised to model dislocations, are not ideal to describe bulk or faulted SiC. These models do not allow the calculation of a "full" band structure along selected high-symmetry directions in the Brillouin zone. Hence a direct comparison with Ref. [121] and [129] is impossible.

The projected band structures of the Shockley partials in 3*C*-SiC are displayed in Fig. 5.14. As all investigated dislocations are glide partials, no dangling bonds are present. For elementary semiconductors like diamond this is often equivalent to the absence of deep defect states in the gap. In a compound semiconductor like SiC, however, the reconstruction bonds of the glide partials are of a different nature (Si–Si or C–C) compared to those in bulk material (Si–C). Therefore the electronic states associated with those reconstruction bonds might have an energetic position very different from states in bulk material. As can be seen in Fig. 5.14, only the silicon-terminated partials (a) – (c) appear to introduce deep defect bands into the gap — reaching as far as 0.4 eV above the valence band maximum (VBM). Analysis shows that the corresponding wavefunctions are strongly localised at the Si–Si reconstruction bonds. At the 30° carbon partial an empty band splits off from the conduction bands at the edge of the Brillouin zone. Apart from that no further considerable changes in the band structure can be observed[13].

The equivalent band structures in 2*H*-SiC are shown in Fig. 5.15. Near the valence band they are all almost identical with the corresponding ones in 3*C*. The only considerable difference between the two polytypes is that in the case of 3*C* the lowest unoccupied bands are of conduction band character, whereas in the case of 2*H* they are localised bands of the intrinsic stacking fault.

The electronic structure of kinks: After having investigated the electronic structure of the straight Shockley partials, the question now arises if any additional electronic levels are introduced in the band gap due to the presence of kinks. To avoid kink–kink interaction, pure cluster models containing a single kink only are used[14]. The structures are relaxed using AIMPRO and their resulting electronic Kohn-Sham eigenvalue spectra are given in Fig. 5.16 and 5.17. The obtained gap widths vary due to the differences in cluster size and geometry for the different types of partials and kinks. As a general observation for both 30° partials, the high energy kinks (LK2 and RK2) give rise to a slightly more pronounced local gap narrowing than their low energy counterparts (LK1 and RK1). This can probably be interpreted as a consequence of the larger strain field associated with LK2 and RK2.

In the case of the silicon-terminated 30° and 90° partials the highest occupied levels are localised gap levels of the Si–Si reconstruction bonds. These characteristic levels correspond to the Si–Si related band above the VBM predicted in the band structures of Si-terminated partials (Fig. 5.14 and 5.15). Also most of the kink structures in carbon-terminated partials do not alter the original electronic structure of the partials much. Overall, only three kink structures show more outstanding electronic characteristics:

30° Si LK2 :   Here the highest occupied level is pushed about 0.4 eV into the band gap. This effect is probably related to the local strain at the kink structure.

---

[13] The slight variations in the width of the gap for the different partials result from differences in model geometry and size.

[14] The cluster stoichiometries are: $Si_{94}C_{89}H_{128}$ (straight 30° Si), $Si_{94}C_{89}H_{124}$ (left kinks in 30° Si), $Si_{110}C_{103}H_{142}$ (right kinks in 30° Si) and $Si_{125}C_{125}H_{150}$ (straight and kinked 90° partials). The 30° C clusters are constructed by simply swapping Si and C. The chosen stacking sequence is 3*C*.
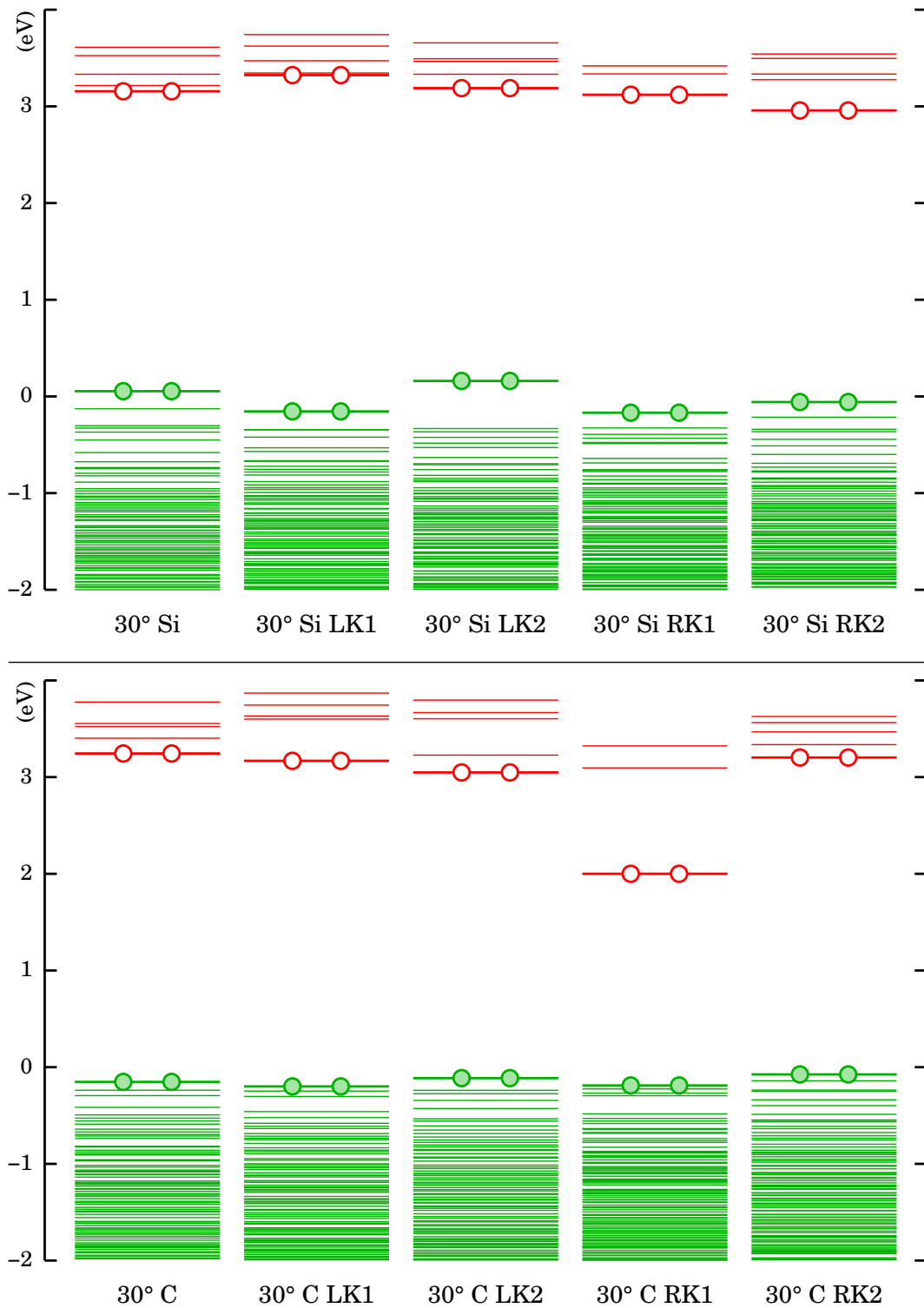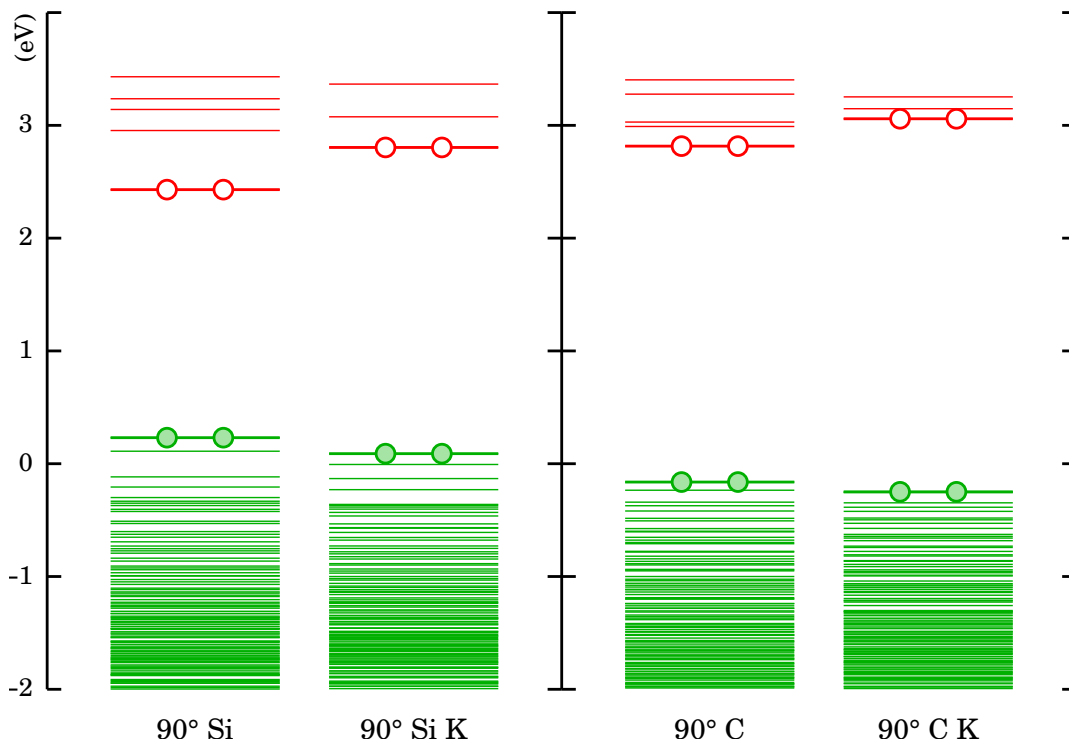
Figure 5.16: The Kohn-Sham eigenvalue spectra of kinked 30° partials in SiC. The highest occupied and the lowest unoccupied states are indicated by two filled and two empty circles respectively. As some of these states are localised gap levels (see text), they do not necessarily represent the valence band maximum or the conduction band minimum. The spectra are unscaled and energy-zero is chosen arbitrarily through a shift constant for all structures. *Upper panel:* The Si-terminated partial. *Lower panel:* The C-terminated partial. The far left spectra give the respective unkinked partial.

Figure 5.17: The Kohn-Sham eigenvalue spectra of kinked 90° partials in SiC. The highest occupied and the lowest unoccupied states are indicated by two filled and two empty circles respectively. As some of these states are localised gap levels (see text), they do not necessarily represent the valence band maximum or the conduction band minimum. The spectra are unscaled and energy-zero is chosen arbitrarily through a shift constant for all structures. *Left pair:* The Si-terminated partial. *Right pair:* The C-terminated partial. The left spectrum in each pair gives the respective unkinked partial.

30° C LK2 : A feature common to all silicon-terminated partials can also be found in the carbon-terminated LK2 structure: As shown in Fig. 5.10, this is the only kink in a C-terminated partial which appears with an "alien" Si–Si reconstruction bond. It is this bond which in the neutral charge state gives rise to the occupied gap level close to the VBM — very similar to the Si partials.

30° C RK1 : As can be seen in Fig. 5.11, the kink reconstructs with a considerable bond angle distortion. This distortion might give rise to the empty level observed in the upper half of the gap.

5.6   Summary and conclusions (SiC)

In this chapter the approach and methods, which were presented in detail in Chapter 4, and tested with diamond as an example, have been applied to model 30° and 90° Shockley partials in the basal plane of silicon carbide. Their low energy core reconstructions, energies and electronic structures were investigated in both the cubic (3*C*) and one hexagonal polytype (2*H*). Furthermore, dislocation glide was modelled as a mechanism of kink formation and migration. The obtained thermal activation energies were in reasonable agreement with experimental results and show the same trends as independent theoretical work.

Recombination-enhanced dislocation glide might cause the experimentally observed degradation of bipolar devices. Therefore, the electronic structures of all involved kinks have been calculated as they could play a role in this process.

In the following section the important results are listed explicitly, followed by a discussion of possible implications.

5.6.1   Selected results

Generally speaking, the core reconstructions of partial dislocations in the basal plane of SiC are very similar to those found in the preceding chapter for diamond. Unlike in diamond though, in SiC one has to distinguish between two types of core reconstructions for each dislocation: Silicon-terminated and carbon-terminated cores with respectively Si–Si and C–C reconstruction bonds. For all investigated structures, the Si–Si core bonds are of comparable length to bonds in bulk Si. The C–C bonds of the carbon-terminated cores, however, are all stretched by approximately 15 % compared to the diamond equilibrium bond length. As a consequence the carbon-terminated cores possess higher core energies.

Furthermore, the results strongly indicate that for all partials both energy and local structure are almost entirely independent of the surrounding crystal stacking. The same can be expected for 4*H*, 6*H* and all further hexagonal stacking sequences.

Comparing the single-period and the double-period reconstructions of the 90° partials, in SiC the latter is only about 90–100 meV/Å lower in energy. Thus both will probably co-exist at room temperature, and most certainly at device operating temperatures.

The glide motion of the four different Shockley partials — 30° and 90°, both Si- and C-terminated — has been considered as a process of kink formation and subsequent migration. Overall, the 90° partial was found to be the thermally more mobile species. Further, at least in the absence of strong obstacles, the thermal mobility of carbon-terminated partials appears to be higher than that of silicon partials. The glide activation energies found — ranging from 2.3 eV to 4.1 eV for long dislocation segments — agree well with experimental results, but the carbon partials being the more mobile species is in contradiction to the conclusions of Pirouz and Yang [145]. However, LACBED (**l**arge **a**ngle **c**onvergent **b**eam **e**lectron **d**iffraction) experiments seem to indicate, that the Si partials in 3C-SiC are smooth whereas the C partials are zig-zagged [146], suggesting they are pinned by obstacles. Then the glide process

would not be mainly controlled by the formation and migration of kinks, but by the energy needed for the pinning and un-pinning at the obstacles.

Since the band gap varies with the polytype, the electronic band structures of the partials were calculated both in 3*C*- and 2*H*-SiC. In both polytypes only the Si partials appear with a localised band, which is as deep as 0.4 eV above the VBM. These bands originate from the Si–Si reconstruction bonds. In contrast to this, the C partials lack any deep states in the gap.

Independent of the type of partial, for the 2*H* polytype additional empty bands below the CBM are found. These bands are associated with the intrinsic stacking fault accompanying the partial dislocation and reach approximately 0.5–0.6 eV into the gap. Due to the smaller gap, for 4*H*- and 6*H*-SiC one would expect the stacking fault bands to be closer to the conduction band. And indeed theoretical investigations by Miao et al. [121] for 4*H* and Iwata et al. [129] for 6*H* predict similar bands at 0.3 and 0.2 eV below the CBM respectively. Except for the presence and positions of the stacking fault related bands, the polytype seems to have no further influence on the electronic structure of the basal plane partials.

Out of the 10 elementary kinks, only three locally introduce additional states in the gap: At the 30° Si partial the high-energy left kink (LK2) appears with an occupied level 0.4 eV above the VBM. At the C partial the same kink (LK2) reconstructs with a Si–Si bond. Similar to the Si–Si core bonds in all Si partials, this leads to a localised level above the VBM. Finally, the low energy right kink (RK1) at the 30° C partial gives rise to an empty level in the upper half of the band gap. In contrast to the 30° partials, all kinks at the 90° partials were found to be electrically inert.

5.6.2 Recombination-enhanced dislocation glide

As presented in the introduction of this chapter in detail, rapid degradation under forward-biased operation turned out to be the major problem of bipolar devices based on hexagonal SiC. In this irreversible process stacking faults, bounded by Shockley partials, propagate and expand, rendering the device useless. Neither the temperature nor the stress in these devices are sufficient to overcome the barriers to dislocation glide. Therefore it is commonly believed that the bounding Shockley partials move in a process of recombination-enhanced dislocation glide (REDG), where the glide activation energy is substantially lowered by energy from non-radiative recombination at the dislocation core. So far the origin of this non-radiative recombination could not be resolved experimentally. However, there are several important experimental findings, which might help to understand the process:

1. In 4*H*-SiC, electroluminescence experiments by Galeckas et al. [115] reveal a 2.8 eV radiative recombination localised at the stacking faults involved in the degradation process. This recombination has been attributed to electronic bands induced by the fault at 0.2–0.6 eV below the conduction band minimum, which were predicted by Miao et al. [121], Iwata et al. [129] and also in this work.

2. Under forward-bias a further ~1.8 eV radiative recombination is found at the Shockley partials bordering the growing stacking faults [115].

3. Skowronski et al. [147] observed 30° partials in 4*H*-SiC moving under forward-bias. In their experiments, light emission of considerable intensity as well as glide motion appeared to be restricted to one species of 30° partial only. The authors of Ref. [147] propose it is the silicon partial which is mobile under forward-bias and emits light.

4. The activation energy to partial glide motion under forward-bias is found to be 0.27 eV [115]. Comparing this with similar experimental estimates without bias, the authors of Ref. [115] suggest a non-radiative recombination of 2.2 eV at the Shockley partials to be responsible for the REDG mechanism. However, no further experimental indication of such recombination exists.

As found in this work, the main difference between Si and C partials in terms of electronic structure appears to be the presence of states 0.4 eV above the VBM in the case of the Si partials. This, in combination with the afore mentioned experimental findings, points towards a non-radiative recombination process between the gap states of the Si-terminated partials and those induced by the stacking fault. Assuming the latter states to be around 0.3 eV below the CBM [121] together with a gap of 3.3 eV, and the states of the Si partials to be 0.4 eV above the VBM (this work), gives an estimate of 2.6 eV for the non-radiative recombination. Considering the band gap error in DFT, this is in reasonable agreement with the 2.2 eV suggested in Ref. [115]. It further explains why only one type of partial — the Si partial — is highly mobile under forward-bias.

All conclusions drawn here are still rather speculative. Further experiments might yield results to support or contradict the REDG model presented above.

Also, the origin of the 1.8 eV emission remains unclear. It cannot be explained by the electronic gap states found for some of the kink structures either, as those do not appear on every type of partial.

5.6.3   Outlook — an alternative model

First test calculations for the 30° Si partial indicate, that positively charging the dislocation line empties the Si–Si bonding states, leaving the reconstruction bonds broken. Since under forward-bias the active region of the device contains many carriers (electrons and holes), holes might be trapped at the Si–Si bonding states, breaking the bond, and thus lowering the glide activation barrier. As the C–C bonds do not induce localised states above the VBM, this mechanism would only work for the Si-terminated partials. Hence it explains part of the experimental observations. However, the effect of the broken bonds on the glide activation energies has to be tested and quantified for all involved partials.

Chapter 6

Summary and Outlook

In this work the atomic structures, energies, electronic properties and the mobility of dislocations in diamond and silicon carbide were studied. The approach presented combines DFT-based atomistic calculations with linear isotropic (and to some extent anisotropic) elasticity theory. The latter has proven invaluable to describe the long range elastic strain effects associated with dislocations and further served as a good convergence criterion with respect to the size of the atomistic models used. A convenient determination of the various energy offsets — unaccessible in elasticity theory — became possible through atomistic modelling within the DFT-based tight-binding method (DFTB). More accurate electronic structure calculations were also performed in a localised basis pseudopotential approach (AIMPRO). Wherever there was overlap between the elasticity description of dislocations and their atomistic modelling, the agreement was very good and the elastic limit of the atomistic calculations well fulfilled.

The first three chapters of this work gave a short introduction of the methods used and dislocation basics in general. The relevant results of elasticity theory were presented and explained.

In Chapter 4 the low energy dislocations of the $\{111\} \langle 110 \rangle$ slip system in diamond were examined. The predicted low energy core structures and dissociation distances match well with those observed experimentally in (HR)TEM. In particular to compare the calculated structures with high-resolution micrographs, the application of image simulations based on the calculated coordinates was crucial. For the 60° dislocation a wide barrier to dissociation into Shockley partials was found, which can explain the experimental observation of undissociated 60° dislocations in HRTEM even though an overall energy-gain strongly favours dissociation.

A comparison of core energies ruled out many of the possible dissociation reactions for the $\frac{1}{2}[1\bar{1}0]$ screw and the 60° dislocation, leaving only the 30° and 90° glide partials and the vacancy structure of the 90° shuffle partial likely to exist. Consistent with earlier calculations, in this work the double-period core reconstruction of the 90° glide partial is favoured over its single-period reconstruction by around 170 meV.

Based on the predicted core structures it was then possible to calculate the corresponding electronic structures. Only the perfect and partial shuffle dislocations ap-

pear to induce deep levels in the band gap — predominantly caused by dangling bonds. As shown in simulations, these gap states should lead to an increase in low-loss EELS absorption and supplementary core EELS absorption below the conduction band edge, which is observed experimentally in type-II diamond. This would then suggest shuffle segments are present in that material, which could well be responsible for the observed band-*A* cathodoluminescence. Furthermore experimental EEL spectra taken on a 30° partial agree with those predicted for the respective glide partial.

For the two probably predominant partial dislocations — the 30° and the 90° glide partials — the kinetics of thermally activated glide motion have been modelled assuming a kink formation and migration mechanism. The 90° glide dislocation was found to be the by far more mobile species, with a thermal activation energy of 3.5 eV for short dislocation segments as opposed to $\sim$ 6.1 eV for the 30° partial.

Comparing the shuffle structures of the 60° dislocation with the glide structures, the latter appeared $\sim$ 0.6 eV/Å lower in line energy. However, since a shuffle $\rightarrow$ glide transition involves absorption or emission of vacancies or interstitials, one would expect shuffle segments to be present if formed by plastic deformation. These might then undergo the transition to glide structures in a thermal annealing process. Furthermore among the shuffle partials it is the vacancy structures which were found to be energetically most stable. This, combined with results from Raman and photoluminescence experiments, supports a scenario for the observed high-pressure, high-temperature decolouring of brown diamond, which involves a shuffle $\rightarrow$ glide transition via the emission of vacancies. Further experiments are necessary to draw final conclusions in this matter.

In Chapter 5 the 30° and 90° Shockley partials in silicon carbide were investigated. These partials border stacking faults in the basal plane whose growth is believed to be the cause of the degradation of bipolar devices under forward bias. Core structures and energies were found to be independent of the crystal stacking. Since the band gap varies with the polytype however, the electronic band structures of the partials were calculated both in 3*C*- and 2*H*-SiC. Only the Si partials appear with a localised band as deep as 0.4 eV above the valence band maximum. The C partials lack any deep states in the gap. For the 2*H* polytype, additional empty bands below the conduction band minimum are found independent of the partial. These bands are associated with the intrinsic stacking fault accompanying the partial dislocation and reach approximately 0.5–0.6 eV into the gap.

The structures and basic processes involved in the formation and migration of kinks in basal plane partial dislocations in SiC were found to be very similar to those in the {111} plane of diamond. Also the general trends for thermal activation energies to glide motion are the same. Overall, the 90° partial appeared to be the more mobile species and the purely thermal mobility of carbon partials in the absence of strong obstacles appears to be higher than that of silicon partials. However, some experiments indicate that pinning by obstacles might play a role as well for some partials. Nevertheless the activation energies of between 2.3 and 4.1 eV found for long dislocation segments agree well with experimental observations in 6*H*-SiC.

The glide activation energies, in combination with the electronic structure calculations, can be related to a recombination-enhanced dislocation glide mechanism,

where non-radiative electron-hole recombination occurs at sites along the dislocation line. The released recombination energy then has to be redirected to assist the formation and migration of kinks at the dislocation and thus substantially lowers the thermal glide activation energy. The results of this work — combined with previous experimental evidence — strongly point towards a non-radiative recombination between the gap levels induced by the stacking fault and those of the bordering silicon-terminated Shockley partials. This recombination would enhance the partial mobility considerably, and thus expedite stacking fault growth in forward-biased bipolar SiC devices — leading to their irreversible degradation.

Methodological outlook

In the long run, the way of combining different methods to model the effects of dislocations on different length scales could be improved. A true embedding approach is imaginable, where the region treated atomistically is embedded in a larger region which is described by elasticity theory. The crucial point would then be the interface/coupling between the different regions, which is far from trivial. As a successful implementation would increase the accuracy of the core energies and electronic structures, and at the same time it would reduce the computational effort as the volume treated atomistically could be minimised.

Appendix A

Straight Dislocations in Elasticity Theory

Throughout this work many fundamental expressions for dislocation energies and forces are used, which are based on linear elasticity theory. In Chapter 2 these expressions are introduced without giving any derivation, as deducing them would clearly lead beyond the scope of this thesis and has been presented elsewhere. Still, to give the reader a feeling for the basic approach, this appendix will give some simple examples. The presentation is brief, but chosen to be close to that of Hirth and Lothe [41].

A.1 Screw dislocations in isotropic media

In this section the elastic strain energy of a straight screw dislocation in an infinite and otherwise defect-free isotropic material will be deduced. To construct the dislocation let us start with a cylinder of unstrained material orientated along the z or x_3 axis. On this cylinder one then applies a shear displacement $u(r)$ along the x_3 direction across the (x_1, x_3) plane as demonstrated in Fig. A.1. The cut/shear surface is defined by $x_1 > 0$, $x_2 = 0$. There is no displacement in the x_1 and x_2 direction $(u_1(r) = u_2(r) = 0)$ and the displacement at the cut surface is discontinuous:

$$\lim_{\eta \to 0} u_3(x_1, -\eta, x_3) - u_3(x_1, +\eta, x_3) = |b|, \qquad \eta > 0, x_1 > 0 \tag{A.1}$$

A Burgers circuit varying ϑ from 0 to 2π gives the Burgers vector b. In an isotropic medium by symmetry the displacement must vary uniformly with ϑ and one obtains:

$$u_3(x_1, x_2, x_3) = \frac{|b|}{2\pi} \arctan\left(\frac{x_2}{x_1}\right) \quad = \quad u_z(r, \vartheta, z) = |b| \frac{\vartheta}{2\pi} \tag{A.2}$$

The right hand side gives the expression in cylindrical coordinates (r, ϑ, z) as indicated in Fig. A.1. Following Eq. (1.33) the resulting non-zero strains are:

$$\varepsilon_{13} = \varepsilon_{31} = -\frac{|b|}{4\pi} \frac{x_2}{x_1^2 + x_2^2} \quad , \quad \varepsilon_{23} = \varepsilon_{32} = \frac{|b|}{4\pi} \frac{x_1}{x_1^2 + x_2^2} \tag{A.3}$$

115
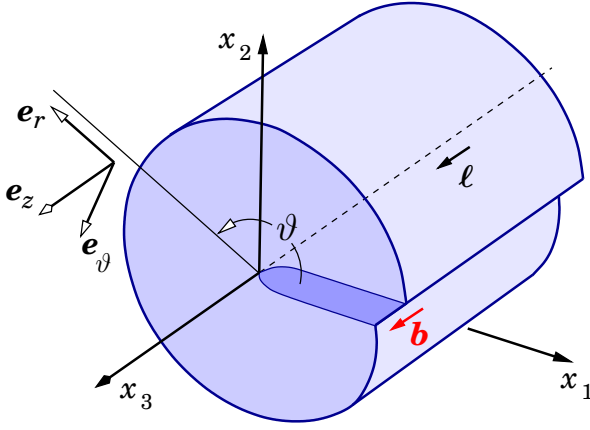
Figure A.1: A screw dislocation oriented along the z-axis with a cylinder of the surrounding material drawn. Both the Cartesian coordinate system (x_1, x_2, x_3) and a local set of unit vectors (e_r, e_ϑ, e_z) of the cylindrical coordinate system are given. ℓ gives the line direction and a Burgers circuit varying ϑ from 0 to 2π gives the Burgers vector b.

Or in cylindrical coordinates simply:

$$\varepsilon_{\vartheta z} = \varepsilon_{z\vartheta} = \frac{|b|}{4\pi} \frac{1}{r} \tag{A.4}$$

Eq. (1.35) and the generalised Hooke's law for isotropic media Eq. (1.45) finally yield the non-zero stress components in cylindrical coordinates:

$$\sigma_{\vartheta z} = \sigma_{z\vartheta} = \frac{\mu |b|}{2\pi} \frac{1}{r} \tag{A.5}$$

This stress field is also known as the self-stress of the screw dislocation. It trivially satisfies the equilibrium equations of classical elasticity Eq. (1.32). With the strains and stresses known one can now calculate the strain energy density. With Eq. (1.37) and (1.36) we obtain:

$$e_s(r) = \frac{1}{2} c_{ijkl} \, \varepsilon_{ij} \, \varepsilon_{kl} = \frac{1}{2} \sigma_{ij} \varepsilon_{kl} = \frac{1}{2} \left(\sigma_{\vartheta z} \, \varepsilon_{z\vartheta} + \sigma_{z\vartheta} \, \varepsilon_{\vartheta z} \right) = \frac{\mu |b|^2}{8\pi^2} \frac{1}{r^2} \tag{A.6}$$

The volume integral over the infinite medium diverges. However, one can calculate the strain energy per unit length contained between two co-axial cylinders centred on the dislocation line:

$$\frac{E(r_0, R)}{L} = \int_{r_0}^{R} e_s(r) \, 2\pi r \, \mathrm{d}r = \frac{\mu |b|^2}{4\pi} \ln \left(\frac{R}{r_0} \right) \tag{A.7}$$

Here L is the length of the integration volume along z and r_0 and R the radius of the inner and outer cylinder respectively. The expression diverges for both the limits $R \to \infty$ and $r_0 \to 0$. The former divergence is physical: The strain energy per unit length of a dislocation in an infinite volume is not defined and the absolute strain energy of a dislocation depends on the size of the crystal. In the case of a single dislocation in an otherwise defect free crystal one can approximately choose R as the average distance of the dislocation line from the surface. Due to the logarithmic dependence in Eq. (A.7) the choice of R is not critical.

The divergence for $r_0 \to 0$ shows a general problem of linear continuum elasticity theory as presented here: The underlying assumption that the displacements vary slowly over the dimensions of the volume investigated breaks down in the core region of a dislocation. Furthermore, for interatomic length scales continuum theory is unable to describe the discrete system of atoms correctly. Therefore we define the so called core radius R_c around the dislocation line. R_c is chosen as the smallest possible radius, where for $R \geq R_c$ elasticity theory is able to describe the strains and energies associated with the dislocation "correctly". Of course this definition is rather vague.

The so called core energy E_c stored inside the cylinder of radius R_c cannot be obtained in elasticity theory. Here atomistic calculations are essential and in many cases quantum mechanical effects contribute considerably. With this core energy introduced, Eq. (A.7) becomes:

$$\frac{E(R)}{L} = \frac{\mu |\boldsymbol{b}|^2}{4\pi} \ln \left(\frac{R}{R_c} \right) + \frac{E_c}{L}, \quad R \geq R_c \tag{A.8}$$

This expression gives the energy per unit length contained in a cylinder of radius R around the dislocation.

An alternative to introducing a core energy is to simply adjust r_0 so that Eq. (A.7) gives the right energy. In the literature this adjusted radius r_0 is often called the core radius as well. However, r_0 is not a physical radius but just a different representation of the core energy — it should not be confused with the core radius R_c. In this work r_0 does not play any major role.

The screw dislocation in a finite volume: In a finite volume one has to consider surface effects. The displacement field of a dislocated rod of finite length and with free ends appears with a twist — the so called *Eshelby twist* [148]. This twist can be observed experimentally in long thin whiskers containing a screw dislocation [149, 150, 151]. It also occurs when a dislocation with a screw component is modelled atomistically in a cluster where the surfaces are allowed to relax freely. Taking the free-surface terms into consideration, the energy of a dislocation in such a cluster of radius R is given as [41]:

$$\frac{E}{L} = \frac{\mu |\boldsymbol{b}|^2}{4\pi} \left\{ \ln \left(\frac{R}{R_c} \right) - 1 \right\} + \frac{E_c}{L} \tag{A.9}$$

A.2 General straight dislocations in isotropic media

Just as with the screw dislocation in the last section, the edge dislocation induces planar strain into the surrounding material — this time however defined by $u_3 = 0$ and $\partial u_i / \partial x_3 = 0$. Starting from this, one can derive the non-zero stress components via an *Airy stress function* as described in Ref. [41]:

$$\sigma_{rr} = \sigma_{\vartheta\vartheta} = -\frac{\mu |\boldsymbol{b}| \sin\vartheta}{2\pi(1-\nu)r}, \quad \sigma_{r\vartheta} = \sigma_{\vartheta r} = \frac{\mu |\boldsymbol{b}| \cos\vartheta}{2\pi(1-\nu)r}, \quad \sigma_{zz} = \nu(\sigma_{rr} + \sigma_{\vartheta\vartheta}) \tag{A.10}$$

Similarly to Eq. (A.6), the strain energy density can be determined from the self-stresses and integration between two co-axial cylinders of radius r_0 and R yields:

$$\frac{E(r_0, R)}{L} = \frac{\mu |\boldsymbol{b}|^2}{4\pi(1-\nu)} \ln\left(\frac{R}{r_0}\right) \tag{A.11}$$

This expression is very similar to Eq. (A.7) and hence all which has been said in the context of the screw dislocation applies here as well. An expression similar to Eq. (A.8) can be derived.

Since only *linear* theory is used, the results for the screw and the edge dislocation can be superimposed to describe a mixed dislocation. Let us assume a dislocation where the Burgers vector is inclined to the line direction at an angle β. Then the screw component of that vector is projected as $|\boldsymbol{b}| \cos\beta$ and the edge component as $|\boldsymbol{b}| \sin\beta$. Superposition then gives the energy of the general straight dislocation:

$$\frac{E(R)}{L} = \mu \left(\cos^2\beta + \frac{\sin^2\beta}{1-\nu} \right) \frac{|\boldsymbol{b}|^2}{4\pi} \ln\left(\frac{R}{R_{\mathrm{c}}}\right) + \frac{E_{\mathrm{c}}}{L}, \quad R \geq R_{\mathrm{c}} \tag{A.12}$$

The term which only depends on the elastic constants and the angle β between the line direction and the burgers vector is in this work called energy factor $k(\beta)$:

$$k(\beta) = \mu \left(\cos^2\beta + \frac{\sin^2\beta}{1-\nu} \right) \tag{A.13}$$

A.3   Straight dislocations in anisotropic media

For good reason the two preceding sections were restricted to isotropic media only. As the calculation of self-stresses in anisotropic media is by far more complicated. Very often it is sufficient to describe the anisotropic crystal as approximately isotropic, with the elastic constants given by the Voigt or the Reuss average (see Section 1.2.4). But this might not always be accurate enough. Fully anisotropic theory yields an expression for the self-stress energy similar to Eq (A.12):

$$\frac{E(R)}{L} = \frac{K|\boldsymbol{b}|^2}{4\pi} \ln\left(\frac{R}{R_{\mathrm{c}}}\right) + \frac{E_{\mathrm{c}}}{L}, \quad R \geq R_{\mathrm{c}} \tag{A.14}$$

Here K is the energy factor derived from the displacement field $\boldsymbol{u}(\boldsymbol{r})$. However the determination of the displacement field through Eq. (1.32) and (1.36) generally requires the solution of a sixth-order polynomial equation [41]. This can be done numerically and was demonstrated for the determination of dislocation energies by Teutonico [93, 152].

In this work the same method is used as implemented by Heggie and Nylén [87] and the anisotropic energy factors are calculated for all dislocations discussed.

Appendix B

Transmission Electron Microscopy

There are several experimental techniques which allow the visualisation of dislocations and other structural defects in semiconductors. Among these the most common are surface etching in combination with conventional microscopy, X-ray topography, and conventional or high-resolution electron microscopy. There exists a variety of further methods, like EBIC (electron beam induced current microscopy), STEM (scanning transmission electron microscopy), infrared microscopy and more. However, as this is not meant to be a complete guide to imaging techniques, they are excluded from this overview. Primarily this appendix presents the two transmission electron microscopy (TEM) methods, which in Chapters 4 and 5 provide valuable information and allow to verify some of the theoretical results.

When "shining" an electron beam onto a thin crystal specimen as used in TEM, there are two important interactions of the electrons with the crystal's atoms: Elastic scattering and inelastic scattering. Electrons which are not scattered are simply transmitted through the crystal. Both the elastically scattered and the transmitted electrons are used in TEM to visualise structural defects. Inelastic electron scattering is utilised in electron energy-loss spectroscopy as described in Appendix C. In the now following sections the basic concepts of TEM will be explained. A more detailed description can be found in the books of Hirsch et al. [153] and Williams and Carter [154].

B.1   Conventional transmission electron microscopy

In conventional transmission electron microscopy a monochromatic and parallel electron beam shines on a thin sample of usually $\leq$ 100 nm thickness. The basic setup is shown in Fig. B.1. When passing through the sample, some electrons are transmitted through without interacting, while others are scattered. For TEM one tries to minimise inelastic scattering/absorption as it does not contain much local structural information. However, it usually cannot be avoided completely and inelastically scattered electrons lead to features in TEM micrographs known as *Kakuchi lines*, whose width is inversely proportional to the atomic spacing in the crystal. The theory of inelastic scattering is rather complicated and shall not be discussed here. In the following only elastic scattering will be considered.
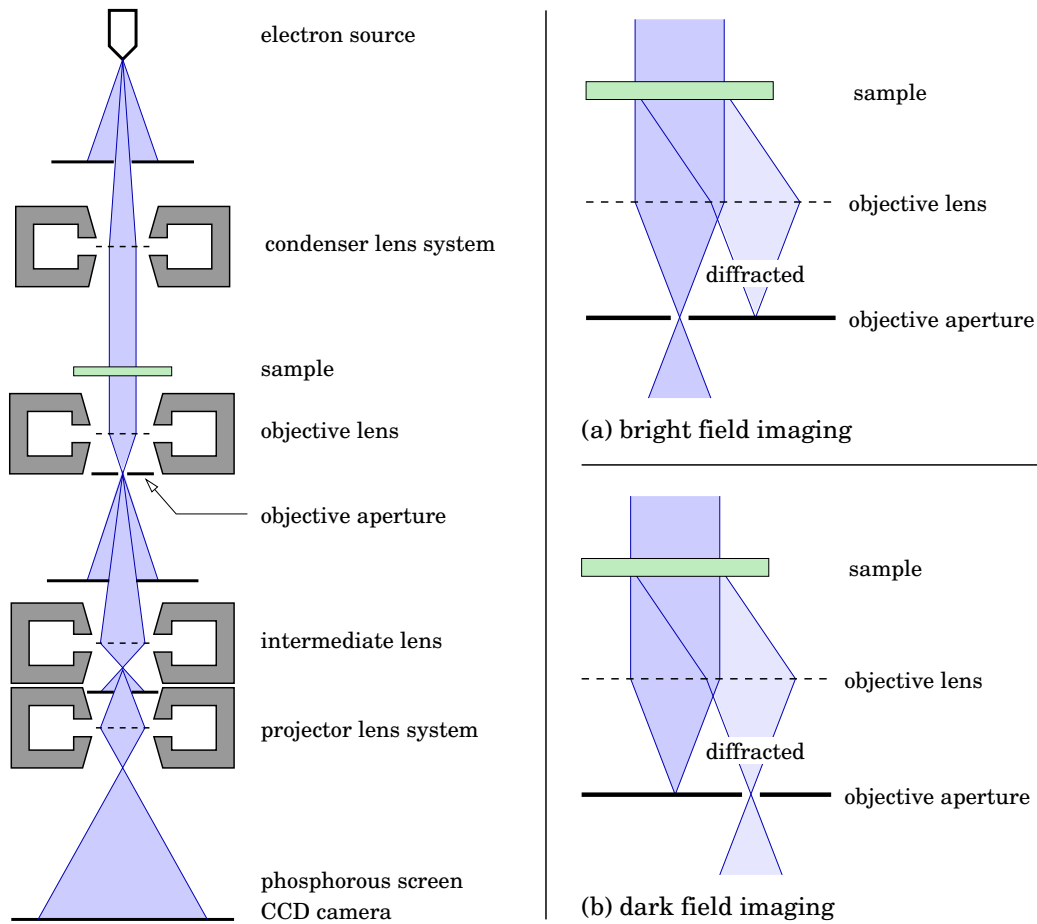
Figure B.1: Schematic sketch of a transmission electron microscope. *Left:* A schematic simplified representation of a TEM setup with condenser, objective, intermediate and projector lenses. Usually between the objective aperture and the intermediate lens a further so called diffraction lens is inserted. *Right:* The two basic imaging modes. In bright field imaging the diffracted beam is masked by the objective aperture and the transmitted beam is used. In the case of dark field imaging the situation is just the other way round.

The origin of the TEM image contrast: Elastic scattering in a crystal occurs with highest intensity whenever the *Bragg condition for diffraction* is met (see Fig. B.2):

$$n\,\lambda = 2\,d_{hkl}\,\sin\theta, \qquad n = 1, 2, 3, \ldots \tag{B.1}$$

Here λ is the wavelength of the electron beam, d_{hkl} the spacing of a set of (hkl) crystallographic planes and θ gives the angle between the beam and that set of planes. The diffraction intensity is strongest for planes with low indices hkl. Since in a crystal there are several such planes, generally several diffracted beams or reflexes are observed. It is common to label the different Bragg reflexes with the reciprocal vector g_{hkl} of their corresponding plane — in this context also called the diffraction vector.

In the vicinity of a structural defect, such as a dislocation, the lattice is strained and the lattice planes are bent locally. This results in local changes of the Bragg condition and thus in a change in intensity for the diffracted and transmitted beams.

If all diffracted beams are masked by the objective aperture and hence only the transmitted beam is used for the imaging process, then one speaks of a *bright field* image (see Fig. B.1 (a)). Bright field images are useful to give a first impression of the microstructure.

For *dark field* imaging one of the diffracted beams is used (Fig. B.1 (b)). To obtain the best contrast, the specimen is tilted in the beam in a way that the so called *two-beam condition* is met: The Bragg condition is nearly fulfilled (low excitation error) for only one specific set of crystallographic low-index planes, resulting in only two beams of considerable intensity — the transmitted and the diffracted beam of the chosen plane. Usually the incident beam is tilted such that the selected diffracted beam lies along the optical axis of the

Figure B.2: Bragg diffraction in the TEM sample. g denotes the corresponding diffraction vector.

microscope. This so called centred dark field method is essential in order to obtain good quality images. Overall the dark field method is lower in intensity and yields a better contrast than bright field microscopy.

When it comes to imaging very fine dislocation structures, then often the resolution of dark field imaging as described above is not sufficient. However, it can be further improved by so called *weak-beam* imaging, where the sample is tilted away from the ideal two-beam condition, such that the exact Bragg condition is only met in a small region close to the dislocation (large Bragg excitation error for bulk material). If this region is narrow enough, then for example dislocation dissociation into partials can be resolved in plan-view with resolutions around 1 nm. A typical weak-beam image is given in Fig. 4.1.

The invisibility criterion: Plan-view imaging can not only be used to visualise dislocations, but also to determine their Burgers vector b by making use of selected diffraction vectors under dark field two-beam or under weak-beam conditions.

Fig. B.3 schematically shows a displaced simple cubic lattice around an edge dislocation as obtained in (anisotropic) elasticity theory. As one can see, the set of lattice planes with the reciprocal vector g_{100} is not distorted / bent at all. Hence for these planes the Bragg diffraction condition in the vicinity of the dislocation core remains the same as in bulk material. Consequently the dislocation remains in-

Figure B.3: Illustration of the invisibility criterion for an edge dislocation: Planes with the reciprocal vector g_{100} are not bent.

visible in dark field imaging *if the corresponding Bragg reflex g_{100} is selected exclusively*. g_{100} is specified by being perpendicular to the Burgers vector b and further lying in

the glide plane, which is defined by b and the dislocation line direction ℓ. From this one obtains the *invisibility criterion* for an edge dislocation:

$$g \cdot b = 0 \quad \text{and} \quad g \cdot (b \times \ell) = 0 \tag{B.2}$$

As can be seen in Fig. B.3, the glide plane itself is usually only weakly bent (reciprocal vector g_{010}). Therefore a g_{010} Bragg reflex would give only very weak contrast. Hence $g \cdot b = 0$ alone is usually sufficient as a less rigorous invisibility criterion. For the pure screw dislocation the invisibility criterion is also given by Eq. (B.2), with the right hand term fulfilled automatically as $b_s \times \ell = 0$.

Imaging a dislocation with different diffraction vectors and comparing the contrast allows an exact determination of the direction of the Burgers vector via the invisibility criterion. Fig 5.2 gives a typical example.

B.2   High-resolution transmission electron microscopy

Conventional TEM in plan-view can be a useful tool to observe dislocations and to determine their line directions and Burgers vectors. Its resolution is too limited though, to provide information about the actual atomic structure at the dislocation core. This can be achieved by high-resolution electron microscopy (HRTEM). To obtain the required resolution of the order of interatomic distances (~ 0.2 nm), high-resolution microscopes need a better mechanical and electrical stability and higher accelerating voltages (> 300 keV) than a conventional TEM. In contrast to the latter, the HRTEM image is a result of a multitude of differently scattered beams and their phase difference (see next paragraph). In order to obtain a good quality image, the sample has to be very thin (~ 10 nm) and the optical axis has to lie precisely along a crystallographic low-index direction. For a good contrast, the atoms along the chosen low-index direction have to lie exactly on top of each other. Therefore line defects like dislocations can only be imaged in cross-section, with the dislocation line as parallel to the electron beam as possible.

The next paragraph will give the rough idea of HRTEM imaging, for more details see for example Ref. [155].

The origin of the HRTEM image contrast: Basically HRTEM works as a phase contrast microscope: After passing through the sample, the diffraction pattern of the electron beam is given as the Fourier transform of the periodic potential experienced. The following objective lens brings all diffracted beams and the transmitted beam back together again, resulting in an interference pattern which is described by the Fourier back-transformation. The created image thus gives a magnified representation of the periodic potential which the electrons experienced when passing through the sample. This image is magnified further by the subsequent electron-optical system. Overall, magnifications of the order of magnitude of 10^6 can be achieved. As there is no such thing as a perfect phase contrast electron microscope, the obtained image depends on several factors. The most important are the alignment of the beam, its coherence, the specimen thickness, the defocus of the objective lens and the chromatic and spherical aberration.

HRTEM simulation: A defect destroys the periodicity of the lattice. Since intuition fails when it comes to predict the effects of that broken periodicity on the Fourier transform, one cannot interpret the observed contrast patterns of defects easily. Interpretation is much easier and safer in comparison with simulated HRTEM images based on atomistic defect models.

Image simulation can be performed using Bloch waves, or in a multi-slice approach as described by Moodie [156]: The atomistic model is divided into slices perpendicular to the incident beam. For each slice the crystal potential is projected onto a plane. The propagation of the beam through the sample is then described as a succession of scattering events of the incident wavefront at the planes representing the slices with intermediate propagation through vacuum between the slices. The rather complicated scattering theory involved is described in detail in Ref. [157]. The subsequent objective lens is modelled by a Fourier transform which has to take the defocus into account as well as possibly several other specifications of the HRTEM setup. In this procedure, the defocus caused by the imperfect objective lens can be described by a so called phase contrast transfer function. This function is specific for each microscope and has to be determined experimentally by imaging the same area of a sample with varying defocus.

The so obtained images can then be compared with experimental images as demonstrated in Section 4.6. All simulated images shown have been generated with the multi-slice method using the commercial CRYSTAL KIT^{TM} and MACTEMPAS^{TM} software packages.

B.3 Alternative techniques

Although this appendix mainly covers transmission electron microscopy, two alternative or supplementary techniques — X-ray topography and surface etching — are now explained briefly.

X-ray topography: Unlike in the case of light or electron beams, there are no efficient lenses for X-rays. Hence the construction of an imaging apparatus similar
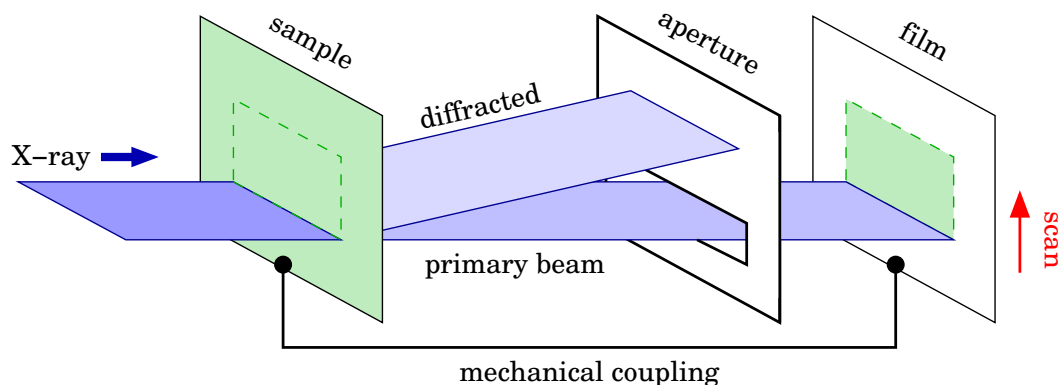
Figure B.4: Schematic X-ray topography setup. The diffracted beam is masked by an aperture. Sample and film are mechanically coupled during the scan.

to a conventional or transmission electron microscope is impossible. Nevertheless X-rays can be used in a scanning technique as illustrated in Fig. B.4: An X-ray beam with a quasi one-dimensional cross section shines on a thin sample. This sample is oriented that the Bragg condition for diffraction is (nearly) met for one set of low-index lattice planes. As described earlier in Appendix B.1 for electron beams, the strain field around the dislocation locally deforms the lattice and thus changes the Bragg condition in its vicinity. As a result, the intensity of the primary and of the diffracted beam vary according to the strain field. To image the strain field, the sample is scanned with one of the beams — usually the diffracted beam — masked by an aperture.

X-ray topography usually gives a resolution of around 5 μm and allows a limited Burgers vector analysis via the Bragg condition.

Chemical surface etching: By means of conventional optical microscopy the limited resolution usually prohibits to directly image dislocations intersecting with an untreated surface of the crystal. The basic idea is to visualise the location of a dislocation by marking it with surface features of sufficient size, which can then be detected in conventional microscopy. This can be done by surface etching: Some solutions preferentially dissolve the strained material around structural defects intersecting the surface. Exposing a polished surface to such a solution for

Figure B.5: Sketch of an etched surface. Dislocations are drawn as lines.

an appropriate amount of time results in the formation of so called *etch pits* where the defect intersects the surface. This is schematically shown in Fig. B.5. It is possible to determine the type of the defect by the shape of the etch pit, revealing dislocations, stacking faults and other defects.

The advantages of chemical etching are its high sensitivity and low cost. On the other hand however, since the etching mechanisms are usually not understood in detail, the interpretation of the obtained images is rather prone to errors. Also the detection limit remains generally unclear and Burgers vectors or core structures cannot be identified[1].

---

[1] Only in some cases, when screw and edge dislocations result in etch pits of distinct shape or depth, then the type of a dislocation may be identified as screw, edge or mixed.

Appendix C

Electron Energy-Loss Spectroscopy

Some of the dislocation core structures discussed in this work appear with under-coordinated atoms or with reconstruction bonds of very different nature compared to bulk material. In many cases this results in electronic states deep in the band gap. This change in electronic structure causes various effects — for example optical activity or recombination-enhanced dislocation glide as described in Chapter 5. All these effects can be measured and quantified experimentally, however the most *direct* spatially resolved observation and "measurement" of electronic gap states can be achieved by *electron energy-loss spectroscopy* (EELS).

As explained in Appendix B, electron microscopy makes use of transmitted and elastically scattered (diffracted) electrons to generate an image of the investigated structure. In contrast EELS utilises *inelastic* scattering of electrons to probe the electronic structure. Implemented in an electron microscope, the fine electron beam allows high spatial resolutions of the order of atomic distances[1]. Combining EELS with TEM has the further advantage that obtained spectra can be directly correlated with defect structures observed in TEM. Microanalytical EELS experiments were first proposed and demonstrated by Hillier and Baker [158] more than half a century ago. However, only several decades later did experimental techniques become advanced and stable enough to allow an easy application.

The following sections will give a brief introduction into the basics of EELS and EELS simulation. More details can be found in the book of Egerton [159] and in Refs. [160, 161].

C.1 The basic principles

As primary high-energy electrons pass through a thin sample, some are scattered inelastically and lose a fraction of their energy. Fig. C.1 (left) illustrates how this energy-loss is measured in EELS. The energy is lost as electronic transitions between the valence band and conduction band states of the solid are induced, resulting in so called *low-loss* spectra. The interaction can be that of single valence electron exci-

---

[1] Typically EEL spectrometer are combined with scanning electron microscopes (STEM) and the technique is often referred to as high-resolution EELS (HREELS).
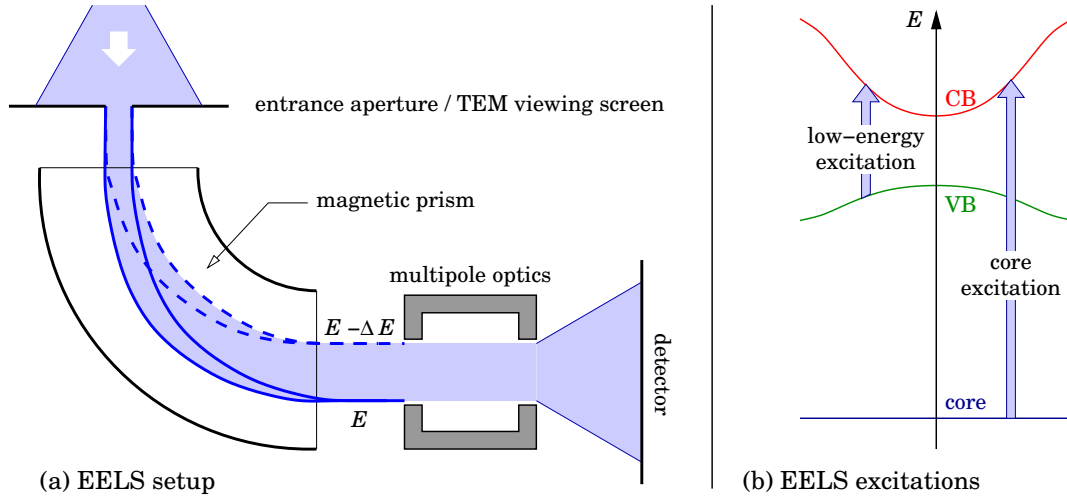
Figure C.1: A simplified EELS setup and the principle EELS excitations. *Left:* A simplified sketch of a spectrometer as it can be installed underneath an electron microscope as shown in Fig. B.1. *Right:* Representation of the principle EELS excitations / transitions. **core** denotes a deep atomic core state and **VB** and **CB** a valence and conduction band states respectively.

tation events or can also encompass collective modes of the solid, called plasmons. If the transitions occur between tightly bound atomic core states and the conduction band one speaks of *core-excitation* EELS or *energy-loss near edge structure* (ELNES) as shown in Fig. C.1 (right). Typical spectral energy ranges are 0–50 eV for low-loss and 100-300 eV for core-excitation EELS. Very small losses of the order of 1–100 meV, which arise from the excitation of phonons, often cannot be resolved by the energy filter of a (S)TEM EEL spectrometer and then simply contribute to the zero-loss peak of (approximately) elastically scattered electrons.

Inelastic scattering and the dielectric function: The interaction of high-energy electrons passing through a solid can be described as an interaction with the field of polarisation which is caused by the electric field of the solid's electrons [162, 163]. The quantity actually observed in EELS experiments is the fraction of electrons which are scattered into a solid angle $d\Omega$ having lost an energy between ΔE and $\Delta E + dE$. To link this quantity with the interaction with the field of polarisation, we introduce the differential cross section $\frac{d\sigma^2}{d\Omega dE}$. Assuming an isotropic medium, this cross section can be related to the macroscopic dielectric function $\varepsilon(\boldsymbol{q}, \omega)$ [164, 165]:

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

Here $\boldsymbol{q}$ gives the momentum transfer and $\omega = E/\hbar$. The last term, the imaginary part of the negative reciprocal dielectric function, is also known as the *loss function*. In the single particle approximation the imaginary part of the dielectric function

$\varepsilon(q, \omega) = \varepsilon_1(q, \omega) + i\varepsilon_2(q, \omega)$ can be written as [166, 167]:

$$\varepsilon_2(q, \omega) = \frac{4\pi e^2}{\Omega q^2} \sum_{n,j} |\langle \Psi^n | e^{iq \cdot r} | \Psi^j \rangle|^2 \, \delta(E^n - E^j - \hbar\omega) \qquad \text{(C.2)}$$

Where Ω is the unit cell volume, and $|\Psi^n\rangle$ are single particle states, with respective energies E^n. The real part of the dielectric function can be obtained easily via a Kramers-Kronig transformation.

C.2 The simulation of EELS

Eq. (C.2) basically represents a joint density of states, weighted by matrix elements $\langle \Psi^n | e^{iq \cdot r} | \Psi^j \rangle$. Electronic structure calculations can be used to calculate these matrix elements. In this work EEL spectra are shown which were obtained based on pseudopotential calculations (AIMPRO). The next sections will give some details on how the necessary matrix elements can be approximated in an approach for low-loss and core-excitation EELS respectively.

Low-loss EELS: As mentioned above, low-loss EELS involves excitations from occupied valence band states to empty conduction band states. In the long-wavelength dipole approximation ($q \to 0$), the matrix elements in Eq. (C.2) can be simplified further using $e^{iq \cdot r} \approx 1 + iq \cdot r$:

$$\varepsilon_2^\ell(E) = \frac{4\pi e^2}{\Omega} \sum_{c,v} \int_{\text{BZ}} |\langle \Psi_k^c | r^\ell | \Psi_k^v \rangle|^2 \, \delta(E_k^c - E_k^v - E) \, \text{d}^3 k \qquad \text{(C.3)}$$

Here c and v denote conduction and valence bands respectively and k gives the Brillouin zone vector — excitations with $\Delta k \neq 0$ are excluded in this simplified approach. As the dielectric function depends on the polarisation of the exciting wave through q, it has to be calculated for different orientations ℓ of r of the electron beam. The integral over k covers the whole Brillouin zone which is sampled with 1000–5300 Monkhorst-Pack-optimised [84] k-points for bulk cells and 100–150 k-points for larger models containing dislocations. As LDA underestimates the band gap, a scissors operator is applied to the band structure, shifting the conduction bands upwards to match the experimental gap. Defect states in the gap are shifted proportionally to their distance from the valence band maximum. A Lorentzian broadening scheme for the δ-function in Eq. (C.3) allows an analytical Kramers-Kronig transformation to obtain the real part of the dielectric function.

Core-excitation EELS: Core-excitation spectra result from transitions from atomic core states to empty conduction band states. However, in pseudopotential calculations as presented in this work, the core states are not treated explicitly but are only included through the pseudopotential. Therefore here Eq. (C.3) is not applicable to compute core-excitation spectra. Still, core-excitations can be modelled at a very simple level: Final state effects (i.e. electron-hole interactions) are completely
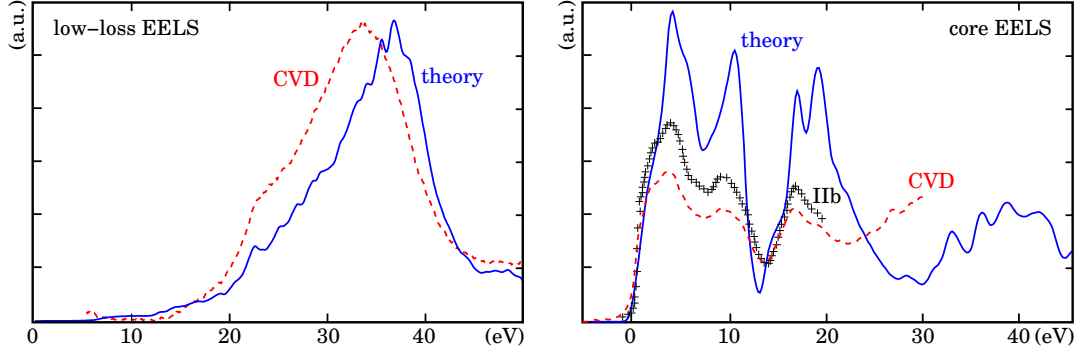
Figure C.2: Theoretical and experimental EELS on diamond. *Left:* Low-loss EEL spectrum (solid line) compared with experimental data on CVD diamond (dashed line) [6]. *Right:* Theoretical *K*-edge EEL spectrum of bulk diamond (solid line), compared with experimental data on CVD diamond (dashed line) [6] and on natural type-IIb diamond (plus signs) from Bruley and Batson [104]. The experimental spectra have been arbitrarily scaled and the energy zero is set at the conduction band minimum, where the spectra have been aligned.

neglected, the core bands are assumed to have no dispersion and the initial state is assumed to be localised on a single atom. Under these assumptions, core-excitation spectra are related to an angular momentum projected local density of states on that atom. For an *s*-like initial core state — as in the diamond *K*-edge spectrum — only transitions into *p*-like states contribute. The spectrum at one atom is then proportional to the *p*-projected density of conduction band states localised at $\boldsymbol{R}$.

$$I_{\boldsymbol{R}}(E) = \sum_c \int_{\mathrm{BZ}} |\langle \Psi_{\boldsymbol{k}}^c | p_{\boldsymbol{R}} \rangle|^2 \delta(E_{\boldsymbol{k}}^c - E_{i,\boldsymbol{R}} - E) \; \mathrm{d}^3 \boldsymbol{k} \qquad (C.4)$$

The angular part of the projecting function $p_{\boldsymbol{R}}$ is the standard Legendre polynomial while the radial part is chosen to be constantly 1 (in the respective units) for radii smaller 1 a.u., and zero otherwise. Varying the radius has no considerable influence on the computed EEL spectra [6]. As discussed in Ref. [13], the energy $E_{i,\boldsymbol{R}}$ of the core level at $\boldsymbol{R}$ can be reconstructed from the differences in the potential at the atomic core between bulk material and the atom of interest. Consequently Eq. (C.4) allows us to approximately obtain the spectral contributions from one atom, without the need to reconstruct the core-wavefunction of the final state. Summation over all atoms within a chosen volume finally reproduces the situation in an electron microscope, where the electron beam probes a small region containing several atoms.

Fig. C.2 shows the resulting calculated low-loss and core-excitation spectra of bulk diamond in comparison with experimental data. Qualitatively the theoretical and experimental spectra agree reasonably well and the peak positions are reproduced. Above 20 eV the differences between theory and experiment become significant. These may be due to surface effects in the performed experiments or the crude approximations made in the theoretical approach.

Bibliography

Page backreferences: To facilitate locating the context of citations, numbers at the very end of an entry indicate the page or pages on which the entry is cited.

The author's publication list: The first 16 entries in this bibliography represent the author's personal publication list sorted by material (Si, diamond, GaN and mixed).

[1] A. T. Blumenau, R. Jones, S. Öberg, T. Frauenheim, and P. R. Briddon, *Optical bands related to dislocations in Si*, J. Phys.: Condens. Matter **12**, 10123 (2000).

[2] A. T. Blumenau, R. Jones, S. Öberg, P. R. Briddon, and T. Frauenheim, *Dislocation related photoluminescence in silicon*, Phys. Rev. Lett. **87**, 187404 (2001). 39, 64

[3] R. Jones and A. T. Blumenau, *Interaction of dislocations in Si with intrinsic defects*, Scr. Mater. **45**, 1253 (2001). 64

[4] J. P. Goss, T. A. G. Eberlein, R. Jones, N. Pinho, A. T. Blumenau, T. Frauenheim, P. R. Briddon, and S. Öberg, *Planar interstitial aggregates in Si*, J. Phys.: Condens. Matter **14**, 12843 (2002).

[5] A. T. Blumenau, M. I. Heggie, C. J. Fall, R. Jones, and T. Frauenheim, *Dislocations in diamond: Core structures and energies*, Phys. Rev. B **65**, 205205 (2002).

[6] C. J. Fall, A. T. Blumenau, R. Jones, P. R. Briddon, T. Frauenheim, A. Gutiérrez-Sosa, U. Bangert, A. E. Mora, J. W. Steeds, and J. E. Butler, *Dislocations in diamond: Electron energy-loss spectroscopy*, Phys. Rev. B **65**, 205206 (2002). 74, 128, 139

[7] A. T. Blumenau, R. Jones, T. Frauenheim, B. Willems, O. I. Lebedev, G. Van Tendeloo, D. Fisher, and P. M. Martineau, *Dislocations in diamond: Dissociation into partials and their glide motion*, Phys. Rev. B (submitted 2002). 139

[8] J. Elsner, A. T. Blumenau, T. Frauenheim, R. Jones, and M. I. Heggie, *Structural and electronic properties of line defects in GaN*, MRS Internet J. Nitride Semicond. Res. **5S1**, W9.3 (2000).

[9] A. T. Blumenau, J. Elsner, R. Jones, M. I. Heggie, S. Öberg, T. Frauenheim, and P. R. Briddon, *Dislocations in hexagonal and cubic GaN*, J. Phys.: Condens. Matter **12**, 10223 (2000). 40

[10] T. Remmele, M. Albrecht, H. P. Strunk, A. T. Blumenau, M. I. Heggie, J. Elsner, T. Frauenheim, H. P. D. Schenk, and P. Gibart, *Core structure of dislocations in GaN revealed by transmission electron microscopy*, in *Microscopy of Semiconducting Materials 2001* (IOP, Bristol, UK, 2001), vol. 169 of *Inst. Phys. Conf. Ser.*, p. 323.

[11] C. J. Fall, R. Jones, P. R. Briddon, A. T. Blumenau, T. Frauenheim, and M. I. Heggie, *Theory of electron energy loss spectroscopy and its application to threading edge dislocations in GaN*, Mat. Res. Soc. Symp. Proc. **693**, I10.1 (2002). 71

[12] C. J. Fall, R. Jones, P. R. Briddon, A. T. Blumenau, T. Frauenheim, and M. I. Heggie, *Influence of dislocations on electron energy-loss spectra in gallium nitride*, Phys. Rev. B **65**, 245304 (2002).

[13] C. J. Fall, J. P. Goss, R. Jones, P. R. Briddon, A. T. Blumenau, and T. Frauenheim, *Modeling electron energy-loss spectra of dislocations in silicon and diamond*, Physica B **308**, 577 (2001). 128

[14] U. Gerstmann, A. T. Blumenau, and H. Overhof, *Transition metal defects in group-III nitrides: An ab initio calculation of hyperfine interactions and optical transitions*, Phys. Rev. B **63**, 075204 (2001).

[15] A. T. Blumenau, C. J. Fall, R. Jones, M. I. Heggie, P. R. Briddon, T. Frauenheim, and S. Öberg, *Straight and kinked 90° partial dislocations in diamond and 3C-SiC*, J. Phys.: Condens. Matter **14**, 12741 (2002).

[16] R. Jones, C. J. Fall, A. Gutiérrez-Sosa, U. Bangert, M. I. Heggie, A. T. Blumenau, T. Frauenheim, and P. R. Briddon, *Calculated and experimental low-loss electron energy loss spectra of dislocations in diamond and GaN*, J. Phys.: Condens. Matter **14**, 12793 (2002).

[17] R. Jones and P. R. Briddon, *The* ab initio *cluster method and the dynamics of defects in semiconductors*, in *Identification of Defects in Semiconductors*, edited by M. Stavola (Academic Press, Bristol, Boston, MA, 1998), vol. 51A of *Semiconductors and Semimetals*, chap. 6. 1, 6

[18] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, and R. Kaschner, *Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon*, Phys. Rev. B **51**, 12947 (1995). 2, 8

[19] T. Frauenheim, G. Seifert, M. Elstner, Z. Hajnal, G. Jungnickel, D. Porezag, S. Suhai, and R. Scholz, *A self-consistent charge density-functional based tight-binding method for predictive materials simulations in physics, chemistry and biology*, phys. stat. sol. (b) **217**, 41 (2000). 2, 8, 11

[20] M. J. P. Musgrave and J. A. Pople, *A general valence force field for diamond*, Proc. R. Soc. Lond. A **268**, 474 (1962). 3

[21] P. N. Keating, *Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure*, Phys. Rev. **145**, 637 (1966). 3

[22] F. H. Stillinger and T. A. Weber, *Computer simulation of local order in condensed phases of silicon*, Phys. Rev. B **31**, 5262 (1985). 3

[23] J. Tersoff, *New empirical model for the structural properties of silicon*, Phys. Rev. Lett. **56**, 632 (1986). 3

[24] K. Ohno, K. Esfarjani, and Y. Kawazoe, *Computational Materials Science* (Springer Verlag, Berlin, 1999). 3

[25] P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Phys. Rev. **136**, B864 (1964). 4

[26] W. Kohn and L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev. **140**, A1133 (1965). 4, 5

[27] E. H. Lieb, *Density functionals for Coulomb systems*, International Journal of Quantum Chemistry **24**, 243 (1983). 4

[28] T. Koopmans, *Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den einzelnen Elektronen eines Atoms*, Physica **1**, 104 (1934). 5

[29] H. Hellmann, *Einführung in die Quantenchemie* (Franz Deutike, Leipzig, 1937). 6

[30] R. P. Feynman, *Forces in molecules*, Phys. Rev. **56**, 340 (1939). 6

[31] P. Pulay, *Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. I. theory*, Mol. Phys. **17**, 197 (1969). 6

[32] P. Briddon and R. Jones, *LDA calculations using a basis of Gaussian orbitals*, phys. stat. sol. (b) **217**, 131 (2000). 6

[33] W. A. Harrison, *Pseudopotentials in the theory of metals* (W. A. Benjamin, New York, 1966). 6

[34] A. M. Stoneham, *Theory of Defects in Solids* (Oxford University Press, London, 1975). 6

[35] G. B. Bachelet, D. R. Hamann, and M. Schlüter, *Pseudopotentials that work: From H to Pu*, Phys. Rev. B **26**, 4199 (1982). 6, 7

[36] G. Seifert, H. Eschrig, and W. Bieger, *Eine approximative Variante des LCAO-X_α Verfahrens*, Z. Phys. Chem. **267**, 529 (1986). 8

[37] D. Porezag, *Development of ab-initio and approximate density functional methods and their application to complex fullerene systems*, Dissertation (PhD thesis), Fakultät für Naturwissenschaften, Technische Universität Chemnitz-Zwickau (1997). 8

[38] M. Sternberg, *The atomic structure of diamond surfaces and interfaces*, Dissertation (PhD thesis), Fachbereich Physik, Universität Paderborn (2001). 8

[39] J. F. Janak, *Proof that $\partial E/\partial n_i = \epsilon_i$ in density-functional theory*, Phys. Rev. B **18**, 7165 (1978). 8

[40] W. M. C. Foulkes and R. Haydock, *Tight-binding models and density-functional theory*, Phys. Rev. B **39**, 12520 (1989). 9

[41] J. P. Hirth and J. Lothe, *Theory of Dislocations* (Wiley, New York, 1982), 2nd ed. 12, 15, 17, 22, 23, 24, 25, 28, 31, 36, 43, 49, 54, 58, 92, 115, 117, 118

[42] A. E. H. Love, *The Mathematical Theory of Elasticity* (Cambridge University Press, Cambridge, 1927). 12

[43] I. S. Sokolnikoff, *Mathematical Theory of Elasticity* (McGraw-Hill, New York, 1956). 12

[44] J. F. Nye, *Physical Properties of Crystals* (Oxford University Press, London, 1957). 12, 14

[45] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, New York, 1959). 12

[46] W. Voigt, *Lehrbuch der Kristallphysik* (Teubner, Leipzig, 1928). 15

[47] A. Reuss, *Berechnung der Fliessgrenze von Mischkristallen*, Z. ang. Math. Mech. **9**, 55 (1929). 15

[48] C. Kittel, *Introduction to Solid State Physics* (John Wiley & Sons, New York, 1986), 6th ed. 18, 27, 69

[49] F. R. N. Nabarro, *The mathematical theory of stationary dislocations*, Adv. Phys. **1**, 269 (1952). 20

[50] J. Frenkel, *Zur Theorie der Elastizitätsgrenze und der Festigkeit kristallinischer Festkörper*, Z. Phys. **37**, 572 (1926). 23

[51] Y. I. Frenkel and T. A. Kontorova, *On the theory of plastic deformation and twinning I*, Zh. Eksp. Teor. Fiz (JETP) **8**, 89 (1938). 23

[52] J. Frenkel and T. Kontorova, *On the theory of plastic deformation and twinning*, Phys. Z. Sovj. **13**, 1 (1938). 23

[53] U. Dehlinger and A. Kochendörfer, *Eigenbewegungen in Kristallgittern*, Z. Phys **116**, 576 (1940). 23

[54] R. E. Peierls, *The size of a dislocation*, Proc. Phys. Soc. Lond. **52**, 34 (1940). 23

[55] F. R. N. Nabarro, *Dislocations in a simple cubic lattice*, Proc. Phys. Soc. Lond. **59**, 256 (1947). 23

[56] X. J. Ning, F. R. Chien, and P. Pirouz, *Growth defects in GaN films on sapphire: The probable origin of threading dislocations*, J. Mater. Res. **11**, 580 (1996). 33

[57] A. R. Lang, *Internal structure*, in *The properties of diamond*, edited by J. E. Field (Academic Press, London, New York, San Francisco, 1982), p. 425. 35, 36

[58] J. Isberg, J. Hammersberg, E. Johansson, T. Wikström, D. J. Twitchen, A. J. Whitehead, S. E. Coe, and G. A. Scarsbrook, *High carrier mobility in single-crystal plasma-deposited diamond*, Science **297**, 1670 (2002). 35, 36

[59] K. Haenen and M. Nesladek, *Optical properties of n-type diamond and p-n junctions* (2002), talk at the De Beers Diamond Conference, Oxford. 36

[60] D. J. Twitchen (De Beers Industrial Diamonds (UK)), *High voltage switching with single crystal CVD diamond* (2002), talk at the De Beers Diamond Conference, Oxford. 36

[61] H. B. Dyer, F. A. Raal, L. Du Preez, and J. H. N. Loubser, *Optical absorption features associated with paramagnetic nitrogen in diamond*, Philos. Mag. **11**, 763 (1965). 36

[62] I. Kiflawi and A. R. Lang, *Linearly polarized luminescence from linear defects in natural and synthetic diamond*, Philos. Mag. **30**, 219 (1974). 36

[63] P. L. Hanley, I. Kiflawi, and A. R. Lang, *Topographically identifiable sources of cathodoluminescence in natural diamonds*, Phil. Trans. R. Soc. Lond. A **284**, 330 (1977). 36

[64] N. Sumida and A. R. Lang, *Cathodoluminescence evidence of dislocation interactions in diamond*, Philos. Mag. A **43**, 1277 (1981). 36

[65] J. E. Graebner, M. E. Reiss, L. Seibles, T. M. Hartnett, R. P. Miller, and C. J. Robinson, *Phonon scattering in chemical-vapor-deposited diamond*, Phys. Rev. B **50**, 3702 (1994). 36

[66] J. W. Steeds, , A. E. Mora, J. E. Butler, and K. M. Bussmann, *Transmission electron microscopy investigation of boron-doped polycrystalline chemically vapour-deposited diamond*, Philos. Mag. A **82**, 1741 (2002). 36, 75

[67] P. Humble and R. H. J. Hannink, *Plastic-deformation of diamond at room-temperature*, Nature, Lond. **273**, 37 (1978). 36

[68] N. Sumida and A. R. Lang, *Cathodoluminescence and TEM studies of dislocation-rich natural diamonds*, in *Microscopy of semiconducting materials 1981*, edited by A. G. Cullis and D. C. Joy (Institute of Physics, Bristol, London, 1981), no. 60 in conference series, p. 319. 36

[69] P. Pirouz, D. J. H. Cockayne, N. Sumida, Sir P. Hirsch, and A. R. Lang, *Dissociation of dislocations in diamond*, Proc. R. Soc. Lond. A **386**, 241 (1983). 36, 37, 41, 56, 57, 65

[70] J. Ruan, K. Kobashi, and W. J. Choyke, *On the "band-A" emission and boron related luminescence in diamond*, Appl. Phys. Lett. **60**, 3138 (1992). 36

[71] D. Takeuchi, H. Watanabe, S. Yamanaka, and H. Okushi, *Origin of band-A emission in diamond thin films*, Phys. Rev. B **63**, 245328 (2001). 36, 74

[72] A. T. Collins, H. Kanda, and H. Kitawaki, *Colour changes produced in natural brown diamonds by high-pressure, high-temperature treatment*, Diamond Relat. Mater. **9**, 113 (2000). 36, 79

[73] T. Moses (Gemological Institute of America) (2002), talk at the De Beers Diamond Conference, Oxford. 37

[74] P. M. Martineau (DTC Research Centre (De Beers), Maidenhead (UK)) (2002), private communication. 37

[75] D. Fisher (DTC Research Centre (De Beers), Maidenhead (UK)) (2002), private communication. 37

[76] A. S. Nandedkar and J. Narayan, *Atomic structure of dislocations in silicon, germanium and diamond*, Philos. Mag. A **61**, 873 (1990). 37, 51

[77] P. K. Sitch, R. Jones, S. Öberg, and M. I. Heggie, *An ab initio study of the 90 degree partial dislocation core in diamond*, Journal de Physique III **7**, 1381 (1997). 37, 64

[78] R. W. Nunes, J. Bennetto, and D. Vanderbilt, *Core reconstruction of the 90° partial dislocation in nonpolar semiconductors*, Phys. Rev. B **58**, 12563 (1998). 37, 50

[79] X. Blase, K. Lin, A. Canning, S. G. Louie, and D. C. Chrzan, *Structure and energy of the 90° partial dislocation in diamond: A combined ab initio and elasticity theory analysis*, Phys. Rev. Lett. **84**, 5780 (2000). 37, 40, 50, 51

[80] M. I. Heggie, S. Jenkins, C. P. Ewels, P. Jemmer, R. Jones, and P. R. Briddon, *Theory of dislocations in diamond and silicon and their interaction with hydrogen*, J. Phys.: Condens. Matter **12**, 10263 (2000). 37, 64

[81] S. Jenkins and M. I. Heggie, *Quantitative analysis of bonding in 90 degrees partial dislocation in diamond*, J. Phys.: Condens. Matter **12**, 10325 (2000). 37

[82] J. P. Goss, R. Jones, M. I. Heggie, C. P. Ewels, P. R. Briddon, and S. Öberg, *Theory of hydrogen in diamond*, Phys. Rev. B **65**, 115207 (2002). 37

[83] C. P. Ewels, N. T. Wilson, M. Heggie, R. Jones, and P. R. Briddon, *Graphitisation at diamond dislocation cores*, J. Phys.: Condens. Matter **13**, 8965 (2001). 37, 47

[84] H. J. Monkhorst and J. D. Pack, *Special points for Brillouin-zone integrations*, Phys. Rev. B **13**, 5188 (1976). 38, 70, 86, 127

[85] H. J. McSkimin and W. L. Bond, *Elastic moduli of diamond*, Phys. Rev. **105**, 116 (1957). 38

[86] M. H. Grimsditch and A. K. Ramdas, *Brillouin scattering in diamond*, Phys. Rev. B **11**, 3139 (1975). 38

[87] M. Heggie and M. Nylén, *Dislocation core structures in alpha-quartz derived from a valence force potential*, Philos. Mag. B **50**, 543 (1984). 39, 118

[88] T. A. Arias and J. D. Joannopoulos, *Ab initio theory of dislocation interactions: From close-range spontaneous annihilation to the long-range continuum limit*, Phys. Rev. Lett. **73**, 680 (1994). 40

[89] J. E. Northrup, *Screw dislocations in GaN: The Ga-filled core model*, Appl. Phys. Lett. **78**, 2288 (2001). 40

[90] K. Suzuki, M. Ichihara, S. Takeuchi, N. Ohtake, M. Yoshikawa, K. Hirabayashi, and N. Kurihara, *Electron-microscopy studies of dislocations in diamond synthesized by a CVD method*, Philos. Mag. A **65**, 657 (1992). 41

[91] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, San Diego, 1996). 42

[92] G. Voronoi, *Nouvelles applications des paramètres continus à la théorie des formes quadratiques, deuxième memoire, recherches sur les paralleloèdres primitifs*, Journal für die Reine und Angewandte Mathematik **134**, 198 (1908). 42

[93] L. J. Teutonico, *Uniformly moving dislocations of arbitrary orientation in anisotropic media*, Phys. Rev. **127**, 413 (1962). 43, 49, 92, 118

[94] J. Bennetto, R. W. Nunes, and D. Vanderbilt, *Period-doubled structure for the 90° partial dislocation in silicon*, Phys. Rev. Lett. **79**, 245 (1997). 47

[95] E. Weber and H. Alexander, *EPR of dislocations in silicon*, Journal de Physique (Paris) **40**, C6/101 (1979). 48

[96] N. Lehto and S. Öberg, *Interaction of vacancies with partial dislocations in silicon*, Phys. Rev. B **56**, R12706 (1997). 48

[97] J. F. Justo, M. de Koning, W. Cai, and V. V. Bulatov, *Vacancy interaction with dislocations in silicon: The shuffle-glide competition*, Phys. Rev. Lett. **84**, 2172 (2000). 48

[98] H. Koizumi, Y. Kammura, and T. Suzuki, *Core structure of a screw dislocation in a diamond-like structure*, Philos. Mag. A **80**, 609 (2000). 53

[99] V. Celli, M. Kabler, T. Ninomiya, and R. Thomson, *Theory of dislocation mobility in semiconductors*, Phys. Rev. **131**, 58 (1963). 58, 64

[100] V. V. Rybin and A. N. Orlov, *Theory of dislocation mobility in the low-velocity range*, Sov. Phys. Solid State **11**, 2635 (1970). 58, 64

[101] R. W. Nunes, J. Bennetto, and D. Vanderbilt, *Structure, barriers, and relaxation mechanisms of kinks in the 90° partial dislocation in silicon*, Phys. Rev. Lett. **77**, 1516 (1996). 59

[102] R. W. Nunes, J. Bennetto, and D. Vanderbilt, *Atomic structure of dislocation kinks in silicon*, Phys. Rev. B **57**, 10388 (1998). 59, 62, 64

[103] W. L. Ng, M. A. Lourenco, R. M. Gwilliam, S. Ledain, G. Shao, and K. P. Homewood, *An efficient room-temperature silicon-based light-emitting diode*, Nature, Lond. **410**, 192 (2001). 70

[104] J. Bruley and P. E. Batson, *Electron-energy-loss studies of dislocations in diamond*, Phys. Rev. B **40**, 9888 (1989). 74, 76, 128

[105] R. Jones (University of Exeter, UK) (2001), unpublished. 79

[106] P. Spear (DTC Research Centre (De Beers), Maidenhead (UK)) (2001), unpublished. 79

[107] W. v. Münch, *Physics of group IV elements and III-V compounds*, in *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung (Springer, Berlin, 1982), vol. 17a of *New Series, Group III*, chap. 1.5. 81, 83, 86, 102

[108] M. M. Dobson, *Silicon Carbide Alloys* (Parthenon, Lancashire, 1986). 81

[109] T. Kimoto, H. Nishino, W. S. Yoo, and H. Matsunami, *Growth mechanism of 6H-SiC in step-controlled epitaxy*, J. Appl. Phys. **73**, 726 (1993). 81

[110] J. A. Powell, D. J. Larkin, L. G. Matus, W. J. Choyke, J. L. Bradshaw, L. Henderson, M. Yoganathan, J. Yang, and P. Pirouz, *Growth of high quality 6H-SiC epitaxial films on vicinal (0001) 6H-SiC wafers*, Appl. Phys. Lett. **56**, 1442 (1990). 81

[111] O. Kordina, J. P. Bergman, A. Henry, E. Janzén, S. Savage, J. André, L. P. Ramberg, U. Lindefelt, W. Hermansson, and K. Bergman, *A 4.5 kV (6H) silicon carbide rectifier*, Appl. Phys. Lett. **67**, 1561 (1995). 81, 82

[112] H. Lendenmann, F. Dahlquist, N. Johansson, R. Söderholm, P. Å. Nilsson, J. P. Bergman, and P. Skytt, *Long term operation of 4.5kV PiN and 2.5kV JBS diodes*, Mater. Sci. Forum **353-356**, 727 (2000). 82, 84

[113] A. Galeckas, J. Linnros, B. Breitholtz, and H. Bleichner, *Investigation of electroluminescence across 4H-SiC p(+)/n(-)/n(+) structures using optical emission microscopy*, Mater. Sci. Forum **353-356**, 389 (2000). 82, 84

[114] J. P. Bergman, H. Lendenmann, P. Å. Nilsson, U. Lindefelt, and P. Skytt, *Crystal defects as source of anomalous forward voltage increase of 4H-SiC diodes*, Mater. Sci. Forum **353-356**, 299 (2000). 82, 84

[115] A. Galeckas, J. Linnros, and P. Pirouz, *Recombination-enhanced extension of stacking faults in 4H-SiC p-i-n diodes under forward bias*, Appl. Phys. Lett. **81**, 883 (2002). 82, 84, 85, 109, 110

[116] P. G. Neudeck and J. A. Powell, *Performance limiting micropipe defects in silicon-carbide wafers*, IEEE Electron. Dev. Lett. **15**, 63 (1994). 82

[117] W. J. Choyke and R. P. Devaty, *Fundamental aspects of SiC*, in *The Handbook of Semiconductor Technology*, edited by K. A. Jackson and W. Schröter (Wiley-VCH, Weinheim, Germany, 2000), chap. 11, p. 661. 82

[118] A. V. Samant, W. L. Zhou, and P. Pirouz, *The relationship between activation parameters and dislocation glide in 4H-SiC single crystals*, phys. stat. sol. (b) **222**, 75 (2000). 83

[119] K. Maeda, K. Suzuki, S. Fujita, M. Ishihara, and S. Hyodo, *Defects in plastically deformed 6H SiC single-crystals studied by transmission electron-microscopy*, Philos. Mag. A **57**, 573 (1988). 83

[120] M. H. Hong, A. V. Samant, and P. Pirouz, *Stacking fault energy of 6H-SiC and 4H-SiC single crystals*, Philos. Mag. A **80**, 919 (2000). 83

[121] M. S. Miao, S. Limpijumnong, and W. R. L. Lambrecht, *Stacking fault band structure in 4H-SiC and its impact on electronic devices*, Appl. Phys. Lett. **79**, 4360 (2001). 83, 85, 104, 109, 110

[122] P. Käckell, J. Furthmüller, and F. Bechstedt, *Stacking faults in group-IV crystals: An ab initio study*, Phys. Rev. B **58**, 1326 (1998). 83, 85, 103

[123] M. E. Levinshtein, S. L. Rumyantsev, and M. S. Shur, eds., *Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, and SiGe* (John Wiley and Sons, New York, 2001). 83, 102

[124] J. Q. Liu, M. Skowronski, C. Hallin, R. Söderholm, and H. Lendenmann, *Structure of recombination-induced stacking faults in high-voltage SiC p-n junctions*, Appl. Phys. Lett. **80**, 749 (2002). 84, 85

[125] P. O. Å. Persson, L. Hultman, H. Jacobson, J. P. Bergman, E. Janzén, J. M. Molina-Aldareguia, W. J. Clegg, and T. Tuomi, *Structural defects in electrically degraded 4H-SiC $p^+/n^-/n^+$ diodes*, Appl. Phys. Lett. **80**, 4852 (2002). 84

[126] J. D. Weeks, J. C. Tully, and L. C. Kimerling, *Theory of recombination-enhanced defect reactions in semiconductors*, Phys. Rev. B **12**, 3286 (1975). 84

[127] H. Sumi, *Dynamic defect reactions induced by multiphonon nonradiative recombination of injected carriers at deep levels in semiconductors*, Phys. Rev. B **29**, 4616 (1984). 84

[128] K. Maeda and S. Takeuchi, *Enhancement of dislocation mobility in semiconducting crystals by electronic excitation*, in *Dislocations in Solids*, edited by F. R. N. Nabarro and M. S. Duesbery (North-Holland, Amsterdam, 1996), vol. 10, p. 444. 84

[129] H. Iwata, U. Lindefelt, S. Öberg, and P. R. Briddon, *Localized electronic states around stacking faults in silicon carbide*, Phys. Rev. B **65**, 033203 (2002). 85, 104, 109

[130] J.-L. Demenet, M. H. Hong, and P. Pirouz, *Plastic behavior of 4H-SiC single crystals deformed at low strain rates*, Scr. Mater. **43**, 865 (2000). 85

[131] M. Zhang, H. M. Hobgood, J.-L. Demenet, and P. Pirouz, *The brittle to ductile transition in 4H-SiC*, Mater. Sci. Forum **389**, 767 (2002). 85

[132] P. K. Sitch, R. Jones, S. Öberg, and M. I. Heggie, *Ab initio investigation of the dislocation structure and activation energy for dislocation motion in silicon carbide*, Phys. Rev. B **52**, 4951 (1995). 85, 101

[133] J. Tersoff, *Modeling solid-state chemistry: Interatomic potentials for multicomponent systems*, Phys. Rev. B **39**, 5566 (1989). 87

[134] D. H. Lee and J. D. Joannopoulos, *Simple scheme for deriving atomic force constants: Application to SiC*, Phys. Rev. Lett. **48**, 1846 (1982). 86, 87

[135] W. R. L. Lambrecht, B. Segall, M. Methfessel, and M. van Schilfgaarde, *Calculated elastic constants and deformation potentials of cubic SiC*, Phys. Rev. B **44**, 3685 (1991). 86, 87

[136] A. P. Mirgorodsky, M. B. Smirnov, E. Abdelmounîm, T. Merle, and P. E. Quintard, *Molecular approach to the modeling of elasticity and piezoelectricity of SiC polytypes*, Phys. Rev. B **52**, 3993 (1995). 87

[137] R. D. Carnahan, *Elastic properties of silicon carbide*, J. Am. Ceram. Soc. **51**, 223 (1968). 87, 92

[138] D. W. Feldman, J. H. Parker, Jr., W. J. Choyke, and L. Patrick, *Phonon dispersion curves by raman scattering in SiC, polytypes 3C, 4H, 6H, 15R, and 21R*, Phys. Rev. **173**, 787 (1968). 86, 87, 92

[139] G. Arlt and G. R. Schodder, *Some elastic constants of silicon carbide*, J. Acoust. Soc. Am. **37**, 384 (1965). 87

[140] S. Karmann, R. Helbig, and R. A. Stein, *Piezoelectric properties and elastic constants of 4H and 6H SiC at temperatures 4–320 K*, J. Appl. Phys. **66**, 3922 (1989). 87, 88

[141] K. Kamitani, M. Grimsditch, J. C. Nipko, C.-K. Loong, M. Okada, and I. Kimura, *The elastic constants of silicon carbide: A Brillouin-scattering study of 4H and 6H SiC single crystals*, J. Appl. Phys. **82**, 3152 (1997). 87, 88, 92

[142] D. H. Lee (University of California, Berkeley) (2002), private communication. 86

[143] S. Fujita, K. Maeda, and S. Hyodo, *Dislocation glide motion in 6H SiC single crystals subjected to high-temperature deformation*, Philos. Mag. A **55**, 203 (1987). 101

[144] A. V. Samant, W. L. Zhou, and P. Pirouz, *Effect of test temperature and strain rate on the yield stress of monocrystalline 6H-SiC*, phys. stat. sol. (a) **166**, 155 (1998). 101

[145] P. Pirouz and J. W. Yang, *Polytypic transformations in SiC — the role of TEM*, Ultramicroscopy **51**, 189 (1993). 101, 108

[146] X. J. Ning and P. Pirouz, *Determination of the core nature of partial dislocations in indentation-deformed 3C-SiC*, Inst. Phys. Conf. Ser. **142**, 449 (1996). 101, 108

[147] M. Skowronski, J. Q. Liu, W. M. Vetter, M. Dudley, C. Hallin, and H. Lendenmann, *Recombination-enhanced defect motion in forward-biased 4H-SiC p-n diodes*, J. Appl. Phys. **92**, 4699 (2002). 110

[148] J. D. Eshelby, *Screw dislocations in thin rods*, J. Appl. Phys. **24**, 176 (1953). 117

[149] W. W. Webb, *Point defect trapping in crystal growth*, J. Appl. Phys. **33**, 1961 (1962). 117

[150] R. D. Dragsdorf and W. W. Webb, *Detection of screw dislocations in α-Al_2O_3 whiskers*, J. Appl. Phys. **29**, 817 (1958). 117

[151] C. M. Drum, *Axial imperfections in filamentary crystals of aluminum nitride.I*, J. Appl. Phys. **36**, 816 (1965). 117

[152] L. J. Teutonico, *Dynamical behavior of dislocations in anisotropic media*, Phys. Rev. **124**, 1039 (1961). 118

[153] P. B. Hirsch, A. Howie, R. B. Nicholson, D. W. Pashley, and M. Whelan, *Electron Microscopy of Thin Crystals* (Butterworths, London, 1965). 119

[154] D. B. Williams and C. B. Carter, *Transmission Electron Microscopy* (Plenum Press, New York, 1996). 119

[155] J. C. H. Spence, *Experimental High Resolution Electron Microscopy* (Oxford University Press, 1988), 2nd ed. 122

[156] A. F. Moodie, *Reciprocity and shape functions in multiple-scattering diagrams*, Z. Naturforsch. **27**, 437 (1972). 123

[157] K. Ishizuka and N. Uyeda, *A new theoretical and practical approach to the multislice method*, Acta Cryst. **A33**, 740 (1977). 123

[158] J. Hillier and R. F. Baker, *Microanalysis by means of electrons*, J. Appl. Phys. **15**, 663 (1944). 125

[159] R. F. Egerton, *EELS in the Electron Microscope* (Plenum Press, New York, 1996), 2nd ed. 125

[160] C. J. Pickard, *Ab initio electron energy loss spectroscopy*, PhD thesis, Christ's College, University of Cambridge (1997). 125

[161] U. Falke, *Anwendungen der Elektronen-Energieverlust-Spektroskopie in der Materialwissenschaft*, Dissertation (PhD thesis), Fakultät für Naturwissenschaften, Technische Universität Chemnitz-Zwickau (1997). 125

[162] J. Geiger, *Elektronen und Festkörper* (Vieweg, Braunschweig, 1968). 126

[163] P. Schattschneider, *Fundamentals in inelastic scattering* (Springer, Wien, New York, 1986). 126

[164] D. Nozières and D. Pines, *Electron interaction in solids. Characteristic energy loss spectrum*, Phys. Rev. **113**, 1254 (1959). 126

[165] P. M. Platzman and P. A. Wolff, *Waves and interactions in solid state plasmas*, vol. suppl. 13 of *Solid state physics* (Academic Press, New York, London, 1972). 126

[166] H. Ehrenreich and M. H. Cohen, *Self-consistent field approach to the many-electron problem*, Phys. Rev. **115**, 786 (1959). 127

[167] G. F. Bassani and G. P. Parravicini, *Electronic states and optical transitions in solids*, vol. 8 of *International series of monographs in the science of the solid state* (Pergamon Press, Oxford, 1975), edited by R. A. Ballinger. 127

Acknowledgements