FIRST-PRINCIPLES CALCULATIONS OF NATIVE DEFECTS IN TETRAHEDRALLY-COORDINATED ISOVALENT SEMICONDUCTORS

Juha Markus Lento

Laboratory of Physics
Helsinki University of Technology
Espoo, Finland

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Department of Engineering Physics and Mathematics, Helsinki University of Technology for public examination and debate in Auditorium E at Helsinki University of Technology (Espoo, Finland) on the 7th of November, 2003, at 12 o'clock.

Dissertations of Laboratory of Physics, Helsinki University of Technology
ISSN 1455-1802

Dissertation 126 (2003):
Juha Markus Lento: First-principles calculations of native defects in tetrahedrally-coordinated isovalent semiconductors
ISBN 951-22-6769-5 (print)
ISBN 951-22-6770-5 (electronic)

OTAMEDIA OY
ESPOO 2003

Abstract

This thesis describes the development, the application and the analysis of the accuracy of state-of-the-art *ab initio* calculations in the description of intrinsic point defects in technologically important tetrahedrally coordinated isovalent semiconductors.

The calculations presented in this thesis are based on the density-functional theory. The effective single-particle equations derived from the density-functional theory in the Kohn-Sham scheme are solved numerically using the plane-wave basis representation of the valence electrons and the pseudopotential description of the core electrons.

The use of the plane-wave basis enforces periodic boundary conditions. The calculation of the properties of isolated defects within periodic boundary conditions is customarily referred to as the supercell approximation. The supercell method is analyzed in detail in the thesis, with a special emphasis on the calculation of charged point defects.

The developments in the numerical methods presented in this thesis include the implementation of a non-local screened-exchange operator for the improved description of the exchange and correlation energy and a non-uniform charge-compensation scheme for charged point defects in a massively-parallel plane-wave pseudopotential software package.

The included papers present the most accurate numerical electronic structure calculations to date for vacancies in silicon and silicon-germanium, and for interstitials in silicon carbide.

Preface

The thesis has been prepared in the Laboratory of Physics at the Helsinki University of Technology during the years 1998-2003. The calculations have mainly been performed using the massive parallel supercomputers at CSC, the Finnish IT center for science.

I wish to thank my supervisor Academy Professor Risto Nieminen for guidance and support in all areas of the thesis work during the whole period. The hands-on introduction to the computational physics and electronic structure calculations in particular was given by Professor Juhani von Bœhm, Dr. Tomi Mattila, Dr. Sami Pöykkö and Dr. Marko Pesola, whose previous work most of this thesis work is built on. The importance of analyzing the results of calculations using intuitive models, as well as the "let's just try how it works (now!)" attitude was taught by (late) Dr. Jose-Luis Mozos. Professor Martti Puska has been the person whom I have always been able to turn to with my questions concerning the general solid-state physics and specialized point-defect related theory. I also want to express my thanks to Sami Siljamäki and Dr. Tuomas Torsti, with whom I've been sharing the office for a long time, for their help and great company. It is my pleasure to thank the whole staff of the Laboratory for such a relaxed and enjoyable working atmosphere.

I express my special gratitude to my wife Sini, daughter Kaisla, my parents, brother and the rest of the family for the constant support and encouragement throughout the years.

Otaniemi, 2003

Juha Lento

Contents

List of publications

This thesis work consists of an overview and the following publications:

I J. Lento, M. Pesola, J.-L. Mozos and R. M. Nieminen, *Vacancies in SiGe: Jahn-Teller distortion and spin effects*, Applied Physics Letters **77**, 232 (2000).

II J. Lento, J.-L. Mozos and R. M. Nieminen, *Charged point defects in semiconductors and the supercell approximation*, Journal of Physics: Condensed Matter **14**, 2637 (2002).

III S.-L. Sihto, J. Slotte, J. Lento, K. Saarinen, E. V. Monakhov, A. Yu. Kuznetsov and B. G. Svensson, *Vacancy-phosphorus complexes in strained $Si_{1-x}Ge_x$: Structure and stability*, Physical Review B **68**, 115307 (2003).

IV J. Lento and R. M. Nieminen, *Non-local screened-exchange calculations for defects in semiconductors: vacancy in silicon*, Journal of Physics: Condensed Matter **15**, 4387 (2003).

V J. Lento, L. Torpo, T. E. M. Staab and R. M. Nieminen, *Self-Interstitials in 3C-SiC*, Journal of Physics: Condensed Matter (submitted 6.8.2003).

VI M. Kaukonen, C. J. Fall and J. Lento, *Interstitial H and H_2 in SiC*, Applied Physics Letters **83**, 923 (2003).

The author's contribution

The author has had an active role in all stages of the work reported in this thesis. The author has written the body of the text in Publications I, II, IV and V, and has had the main responsibility for the work reported in them. He has performed all electronic structure calculations reported in Publications I-V, excluding the positron calculations in Publication III, in which he has contributed the theoretical analysis of the microscopic structure of the defect. In Publication VI the author has contributed in the geometry optimizations using the plane-wave pseudopotential calculations and the analysis of the supercell effects for charged supercells.

1 Introduction

The macroscopic material properties of semiconductors are largely determined by intentionally introduced microscopic impurity (dopant) atoms and unintentionally created point defects. Intentionally introduced dopant atoms induce electronic levels within the energy gap of the semiconductor near the valence-band or the conduction-band edges. These shallow levels may release electrons to the conduction band (donors) or introduce electron holes to the valence band (acceptors), leading to what is called n-type or p-type conductivity, respectively.

The crystal growth and the later device manufacturing processes necessarily result in the creation of some amount of point (and extended) defects. The point defects may interfere with the intended doping by forming defect complexes with the dopant atoms thus passivating them, or by introducing electron levels in the energy band gap region compensating the intended doping. The unintentionally introduced defect levels may be located near the band edges (shallow levels) or in the mid-gap region (deep levels).

The identification of the microscopic structure of the point defects leading to specific experimentally measured levels is a difficult task, as the experimental techniques are usually rather indirect. A fruitful approach to provide microscopic interpretation of the measurements is given by *ab initio* methods. In *ab initio* methods one constructs a numerically solvable model system, and calculates defect properties in that model system from the first principles, *i.e.* from the Schrödinger equation. *Ab initio* methods, accompanied by experimental measurements and analytic model systems, can provide thorough understanding of the microscopic structure of the point defects.

It is generally impossible to directly solve the many-particle Schrödinger equation for the whole macroscopic crystal with roughly 10^{23} cm^{-3} atoms and 10^{24} cm^{-3} interacting electrons numerically. Basically two approximations are necessary. First, the many-particle interactions are to be somehow simplified. Second, the whole crystal has to be described using a model system with fewer particles. This leads to two general underlying topics of the thesis work, the density-functional theory (DFT) and the supercell method.

2 Density-functional theory

The physics of point defects and their electronic structure in the Ångström length-scale is described by quantum mechanics. Although the exact quantum-mechanical description is in principle known in the form of the Schrödinger equation, its direct numerical solution for more than just a few mutually interacting electrons is generally impossible. An alternative formulation of the quantum mechanics for the ground state (GS) of the many-electron system is given by the DFT [1]. The DFT states that the GS energy of the system of interacting electrons is a functional of the electron density, and that the GS energy can be found by minimizing the functional with respect to the density.

The most successful and widely used practical applications of the DFT are based on the so-called Kohn-Sham (KS) scheme, the local-density approximation (LDA) [2] and the generalized gradient approximation (GGA) [3]. The KS scheme maps the GS density, and thus the GS energy, of the original system of the interacting electrons to a system of non-interacting particles in an effective potential, and leads to a set of numerically solvable single-particle equations. In this section we describe the standard DFT only to the extent that is necessary for the discussion of the generalized Kohn-Sham (GKS) theory and the screened-exchange local-density functional [4]. An extended review of the DFT is presented *e.g.* in reference [5]. The Hartree atomic units are used throughout the text.

2.1 Hohenberg-Kohn theorem

Hohenberg and Kohn [1] derived the basic theorems of the DFT for the total energy of N electrons in a local external potential V_{ext}, described by the Hamiltonian

$$H = T + V_{ee} + \sum_{i=1}^{N} V_{ext}(\mathbf{r}_i), \tag{1}$$

where T and V_{ee} are the kinetic and electron-electron interaction operators, respectively. The basic theorem is that

1. the total energy is a functional of the electron density ρ,

$$E[\rho] = \int d\mathbf{r}\, V_{ext}(\mathbf{r})\rho(\mathbf{r}) + F[\rho], \tag{2}$$

2. the variational principle with respect to density applies, *i.e.*

$$E[\rho] \geq E_{GS} \text{ and } E[\rho_{GS}] = E_{GS}, \tag{3}$$

where E_{GS} and ρ_{GS} are the GS total energy and the electron density, respectively. In the extended derivation of the theorem, Levy [6] defines the Hohenberg-Kohn functional $F[\rho]$ using constrained minimization as

$$F[\rho] = \min_{\Psi \to \rho} < \Psi | T + V_{ee} | \Psi >, \tag{4}$$

where the minimum is taken over all N-electron wave functions Ψ that give the density ρ, *i.e.* ρ must be N-representable. The explicit functional dependence of the Hohenberg-Kohn functional $F[\rho]$ on the density is not known.

The main benefit of the DFT is that it provides a formal justification for working with the density instead of the many-particle wave function. If an approximation for the Hohenberg-Kohn functional can be found, one can find the corresponding approximations for the GS density and the GS energy by minimizing the total-energy functional with respect to the density.

2.2 Kohn-Sham scheme

The most common and successful applications of the DFT are based on the KS scheme [2]. The KS scheme is based on the partitioning of the total energy functional as

$$E[\rho] = T_0[\rho] + \int d\mathbf{r}\, V_{ext}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}'\, \frac{\rho(\mathbf{r}')\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[\rho], \tag{5}$$

where $T_0[\rho]$ is the kinetic energy of a model system of N non-interacting electrons and $\rho(\mathbf{r})$ is the electron density. The electron density $\rho(\mathbf{r})$ is constructed from the spin-orbitals ψ_i of the N non-interacting electron system

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2. \tag{6}$$

The central assertion of the KS scheme is that for any interacting system one can find a non-interacting model system, with some local effective potential $v_{eff}(\mathbf{r})$, that has the GS density that exactly equals the GS density of the interacting system (v-representability). The difference in the kinetic energy between the interacting electron system and the model non-interacting electron system, as well as the many-particle effects, the exchange and correlation energy contribution, are taken into account in the remaining $E_{xc}[\rho]$ term.

The exact form of the exchange and correlation functional is not known. However, a simple LDA

$$E_{xc}[\rho] \approx E_{xc}^{LDA}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r}) \epsilon_{xc}^{HEG}(\rho(\mathbf{r})), \tag{7}$$

where $\epsilon_{xc}^{HEG}(\rho(\mathbf{r}))$ is the accurately numerically evaluated exchange and correlation energy density of a homogeneous electron gas (HEG) [7] with an electron density $\rho(\mathbf{r})$, has proved to be sufficient for a wide range of systems [5].

Minimization of the total energy with the constraint of fixed number of electrons N leads to the single-particle equations

$$[-\frac{1}{2}\nabla^2 + v_{eff}^{LDA}(\mathbf{r})]\psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \qquad i = 1, \ldots, N, \tag{8}$$

where the local effective potential $v_{eff}^{LDA}(\mathbf{r})$ is

$$v_{eff}^{LDA}(\mathbf{r}) = V_{ext}(\mathbf{r}) + \int d\mathbf{r}' \, \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \left. \frac{d(\rho \epsilon_{xc}^{HEG}(\rho))}{d\rho} \right|_{\rho=\rho(\mathbf{r})}. \tag{9}$$

The GS total energy and density are obtained numerically by the self-consistent solution of KS equations (6), (8) and (9). In principle, if the exact effective potential were known the calculated GS energies and electron densities for the interacting and the non-interacting model systems would be exactly equal.

It is by no means obvious that the functional dependence of the exchange and correlation energy on the density is simple. In fact, there are many known deficiencies of the LDA [5], of which the most apparent ones related

to the point defects in semiconductors are the overestimation of cohesive energies and the underestimation of the energy band gap. The overestimation of cohesive energies and the resulting over-binding of ions may lead to qualitatively incorrect description of the local geometries of the defects especially in cases where a relatively flat Born-Oppenheimer surface possesses several competing minima and the minimum total energy configuration depends on several competing energy contributions, such as the exchange-interaction and the Jahn-Teller effect. The underestimation of the band gap introduces uncertainty in the determination of the absolute positions of the ionization levels. In addition, it enhances the spurious effects induced by the defect level dispersion in the supercell method. Also, as the DFT is a GS theory, it does not directly provide information on the excited states of the system. The physical interpretation of the KS eigenvalues and the KS single-particle orbitals is discussed exceedingly in the literature, for example in references [6] and in the references therein.

2.3 Screened-exchange local-density approximation

One of the promising extensions of the KS scheme and the LDA is the GKS scheme and in particular the screened-exchange local-density approximation (sX-LDA) therein [4, 8, 9]. The GKS scheme provides the theoretical framework in which one can develop new approximations for the exact Hohenberg-Kohn functional. As in the evaluation of the kinetic energy in the standard KS scheme, the auxiliary single-particle spin-orbitals can be used to improve the description of the exchange interaction in a similar manner as in the Hartree-Fock (HF) approximation. The unscreened exchange-interaction in the HF approximation is known to lead to some fundamental problems, such as the vanishing density of states at the Fermi surface for metals and the overestimation of the band gaps in semiconductors and insulators. The over-emphasis of the insulating character in the HF approximation and the over-emphasis of the metallic character in the LDA are tractable to the fundamental approximations within them. The LDA is based on the exchange and correlation description of the homogeneous electron gas, which possesses no gap and in that sense is a metallic system. The missing correlation effects and thus the absence of the screening of the exchange interaction in the HF approximation are responsible for the poor description of metallic systems. It is intuitively clear, that a working approach could be reached by combining the two somewhat complementary

approaches, the LDA and the HF approximations.

In the sX-LDA the exchange and correlation part of the total-energy functional is further partitioned to explicitly include the non-local screened-exchange interaction of the single-particle orbitals

$$\begin{aligned} E_{xc}[\rho] &\approx E_{xc}^{sX-LDA}[\rho] = E_x^{LDA}[\rho] - E_{sx}^{LDA}[\rho] + E_c^{LDA}[\rho] \\ &\quad - \sum_{i<j}^{N} \int d\mathbf{r} \int d\mathbf{r}' \frac{\psi_i^*(\mathbf{r})\psi_j^*(\mathbf{r}')e^{-k_{TF}|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{r})\psi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\delta_{s_i s_j}, \end{aligned} \tag{10}$$

where indexes s_i and s_j in the Kronecker delta $\delta_{s_i s_j}$ refer to the spins of the orbitals. The local contributions are defined as

$$E_c^{LDA}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r})\epsilon_c^{HEG}(\rho(\mathbf{r})), \tag{11}$$

$$E_x^{LDA}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r})\epsilon_x^{HEG}(\rho(\mathbf{r})), \tag{12}$$

$$E_{sx}^{LDA}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r})\epsilon_{sx}^{HEG}(\rho(\mathbf{r})). \tag{13}$$

Above, the correlation energy density $\rho\epsilon_c^{HEG}$ for the HEG is calculated numerically [7, 10] and the exchange and the screened-exchange contributions for the HEG

$$\epsilon_x^{HEG}(\rho) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\rho^{1/3} \tag{14}$$

and

$$\epsilon_{sx}^{HEG}(\rho) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\rho^{1/3}F(z), \tag{15}$$

respectively, are evaluated analytically. For the Thomas-Fermi screening used in this work the analytic expression for the function $F(z)$ in equation (15) is [11]

$$F(z) = 1 - \frac{4}{3}z\arctan\frac{2}{z} - \frac{z^2}{6}\left[1 - \left(\frac{z^2}{4}+3\right)\ln\left(1+\frac{4}{z^2}\right)\right]. \tag{16}$$

The parameter $z = k_{TF}/\bar{k}_F$ is defined as the ratio of the Thomas-Fermi and Fermi wave vectors corresponding to the average valence-electron density.

Similarly as in the standard KS scheme, the minimization of the total energy leads to a set of single-particle equations,

$$\begin{aligned}
&[-\frac{1}{2}\nabla^2 + v_{eff}^{sX-LDA}(\mathbf{r})]\psi_i(\mathbf{r}) \\
&\quad - \int d\mathbf{r}' \sum_{j=1}^{N} \frac{\psi_j^*(\mathbf{r}')e^{-k_{TF}|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\ \psi_i(\mathbf{r}')\delta_{s_i s_j} \\
&= \epsilon_i\psi_i(\mathbf{r}), \quad i = 1, \ldots, N,
\end{aligned} \tag{17}$$

with the effective local potential

$$\begin{aligned}
v_{eff}^{sX-LDA}(\mathbf{r}) &= V_{ext}(\mathbf{r}) + \int d\mathbf{r} \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\
&\quad + \frac{d(\rho\epsilon_x^{HEG}(\rho) - \rho\epsilon_{sx}^{HEG}(\rho) + \rho\epsilon_c^{HEG}(\rho))}{d\rho}.
\end{aligned} \tag{18}$$

In contrast to the standard scheme, we note that in the sX-LDA approximation the single-particle equations (i) contain a non-local contribution and (ii) are explicitly coupled. These two factors will induce considerable computational complexity compared to the standard LDA.

The original main motivation in the development of the GKS formalism and the sX-LDA appears to be in the improved description of the excited states and in the correction of the LDA band-gap underestimation [4]. The increased computational cost compared to the standard KS scheme and the LDA has limited the applications mainly to the calculation of the bulk properties [4, 8, 12]. Publication IV presents, to our knowledge, the first reported calculation of a point defect in a semiconductor within the sX-LDA implemented in the plane-wave pseudopotential (PWPP) method.

2.4 Implementation of the non-local sX-LDA functional in the plane-wave basis

The present implementation of the sX-LDA in a PWPP calculation requires modifications in two parts of the PWPP program, basically. The local ef-

fective potential $v_{eff}(\mathbf{r})$ needs to be modified and the non-local screened-exchange operator needs to be constructed. The modification of the local potential can be carried out simply according to the analytic expressions (14)–(18).

As the computational machinery of the fast Fourier transformation (FFT) is already implemented in the standard plane-wave (PW) programs, its utilization in the calculation of the non-local part of the effective potential is straightforward. A similar approach has been used to evaluate the exact-exchange operator in the PW basis [13]. We also note the similarity of the approach to the evaluation of the standard Hartree term.

In the PW basis-set expansion the single-particle spin-orbitals $\psi_{i\mathbf{k}}$ are stored as PW coefficients $\psi_{i\mathbf{k}}(\mathbf{G})$

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} \psi_{i\mathbf{k}}(\mathbf{G}) \exp[-i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}], \tag{19}$$

where the first index i labels the single-particle state (band and spin) and the second index $\mathbf{k}$ the Bloch state ($\mathbf{k}$ is a point in the first Brillouin zone). The transformation from the representation of the single-particle orbitals $\psi_{i\mathbf{k}}(\mathbf{G})$ in the reciprocal space to the uniform grid in the functions $\psi_{i\mathbf{k}}(\mathbf{r})$ in the real space is achieved efficiently using the inverse FFT, $\mathcal{F}^{-1}$. The transformation to the opposite direction is achieved using the forward FFT, $\mathcal{F}$. The electron density is defined as a suitable weighted sum over $\mathbf{k}$-points in the first Brillouin zone,

$$\rho(\mathbf{r}) = \sum_{i\mathbf{k}} w_{\mathbf{k}} |\psi_{i\mathbf{k}}(\mathbf{r})|^2. \tag{20}$$

Let us next take a closer look at the non-local screened-exchange term in the single-particle equations (17) for the state $\psi_{i\mathbf{k}}(\mathbf{r})$,

$$\hat{h}^{NL}_{sx}\psi_{i\mathbf{k}}(\mathbf{r}) = -\sum_{j\mathbf{q}} \int d\mathbf{r}' \frac{\psi^*_{j\mathbf{q}}(\mathbf{r}') e^{-k_{TF}|\mathbf{r}-\mathbf{r}'|} \psi_{j\mathbf{q}}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \psi_{i\mathbf{k}}(\mathbf{r}') \delta_{s_{i\mathbf{k}} s_{j\mathbf{q}}}, \tag{21}$$

where in comparison to Equation (17) we have now explicitly written out the Bloch state indexes $\mathbf{k}$ and $\mathbf{q}$. The procedure to calculate the convolution type integral in equation (21) using the FFT machinery starting from the reciprocal-space representation is as follows:

1. $\psi_{j\mathbf{q}}(\mathbf{r}) = \mathcal{F}^{-1}\{\psi_{j\mathbf{q}}(\mathbf{G})\}$,
2. $\tilde{\rho}_{ij\mathbf{kq}}(\mathbf{r}) = w_{\mathbf{q}}\psi^*_{j\mathbf{q}}(\mathbf{r})\psi_{i\mathbf{k}}(\mathbf{r})\delta_{s_{i\mathbf{k}}s_{j\mathbf{q}}}$,
3. $\tilde{\rho}_{ij\mathbf{kq}}(\mathbf{G}) = \mathcal{F}\{\tilde{\rho}_{ij\mathbf{kq}}(\mathbf{r})\}$,
4. $\int d\mathbf{r}' \frac{\psi^*_{j\mathbf{q}}(\mathbf{r}')e^{-k_{TF}|\mathbf{r}-\mathbf{r}'|}\psi_{j\mathbf{q}}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\ \psi_{i\mathbf{k}}(\mathbf{r}') = \psi_{j\mathbf{q}}(\mathbf{r})\mathcal{F}^{-1}\left\{\frac{\tilde{\rho}_{ij\mathbf{kq}}(\mathbf{G})}{|\mathbf{k}-\mathbf{q}+\mathbf{G}|^2+k^2_{TF}}\right\}$,

where the standard Fourier transform of the screened Coulomb interaction is

$$\frac{1}{|\mathbf{k}-\mathbf{q}+\mathbf{G}|^2+k^2_{TF}} = \mathcal{F}\left\{\frac{e^{-k_{TF}|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\right\}. \tag{22}$$

As the summation $\sum_{j\mathbf{q}}$ cannot be performed efficiently inside the convolution integral, as in the evaluation of the standard Hartree term, the calculation of the non-local term necessarily becomes computationally by far the most demanding part of the method. Additional care in the implementation compared to the standard PW program with local effective potentials has to be taken when the inversion and the possible point-group symmetries of the first Brillouin zone are to be utilized. In the parallel implementations utilizing parallelization over **k**-vectors, the evaluation of the non-local term necessitates also the communication of the single-particle orbitals at different **k**-points between the computer nodes. In the standard methodology with local effective potential only the communication of the sum of the wave-function squares, the partial density at different **k**-vectors, is required.

Additional and possibly necessary modifications to the current computational scheme as described here and used in Publication IV are the inclusion of the sX-LDA functional in the description of the core electrons and the use of improved screening models.

3 Supercell methods

Periodic boundary conditions (PBC) are the natural choice for the theoretical study of the bulk properties of crystalline solids. The crystal structure of a solid is defined using a periodically repeated unit cell containing a basis, *i.e.* the ion positions within the unit cell. As a computational model, the infinitely repeated unit cell does not have any surfaces, and the calculated properties correspond exactly to bulk properties of the macroscopic crystal, which by definition are independent of the surfaces of the crystalline solid.

The supercell method is used to model isolated imperfections in crystal structures. Examples of such deviations from the perfect crystal structure are surfaces, dislocations, or as in the scope of this thesis, point defects. The supercell is a unit cell, the size of which is made large enough to accommodate the defect and enough bulk crystal to separate the defects from their periodic images. As the supercell is made larger the unphysical interactions between periodic defect images in the superlattice become negligible or at least controllable, and eventually vanish to yield the desired limit corresponding to the isolated defect. The supercell sizes reachable with the current *ab initio* methods require careful analysis of the convergence with respect to the supercell size when the comparison of the calculated values to the experimental measurements of isolated point defects is to be made.

An alternative approach to the supercell method is the cluster method, in which the crystal is modeled by a large, properly terminated isolated cluster of ions cut from the original crystal structure. The defect is embedded into the cluster and the cluster size is increased until the desired properties have converged. In practice the advantage with respect to supercell method gained from the use of non-periodic boundary conditions is canceled by the necessity of considering the surface effects. The choice between the supercell and cluster methods also depends on the used computational scheme. The methods based on the expansion of the single-particle orbitals in localized atomic-orbital or Gaussian basis originating from the chemistry applications are more suitable to cluster boundary conditions, whereas the use of the traditional solid-state approach with the PW expansion of the single-particle orbitals necessitates the use of the PBC. The new approaches based on the numerical representation of the single-particle orbitals in real-space grids are flexible in the choice of the boundary conditions [14], allowing even mixed boundary conditions, *i.e.* PBC in some directions and non-periodic in others. Nevertheless, the effects of the chosen boundary conditions on

the calculated defect properties are to be estimated.

The spurious defect-defect interactions in the supercell method between periodic defect images are customarily treated in three parts, the elastic interactions mediated by ion relaxations around the defect, the dispersion of the defect levels in the periodic computational superlattice, and the electrostatic Madelung energy of the array of charged defects. As an important addition to this list, we also discuss briefly the dependency of the defect description on the particular choice of the supercell symmetry. In practice, it is not necessarily simple to identify and estimate different error contributions separately.

Clear answers to the question of the convergence of the calculated properties as a function of the supercell size can be obtained by repeating calculations for several supercell sizes and symmetries. In practice the calculations are often done using the largest supercell computationally feasible at the time. Meticulous and laborious convergence analysis is then disregarded with a false thought that "This is already the best that can be done at the moment." In fact, a careful convergence analysis using several different supercells does not only provide error estimates for the calculated results, but in many cases provides an important method to extrapolate the supercell results to the experimentally relevant limit of isolated defects.

3.1 Elastic interactions

Point defects induce elastic stress in the host lattice which is relieved by ion displacements, *i.e.* lattice relaxation. Although the ionic relaxations in the calculations are not limited to first neighbour shells nowadays, the lattice relaxation pattern is still restrained by the supercell geometry. The argument used often in the supercell calculations that the ion displacements vanish near the borders of the supercell does not necessarily guarantee that the long-range ionic relaxations are correctly described, as the supercell symmetry itself often fixes the positions of the border ions. According to the elastic continuum theory [15], the strain field at large distances from the point defect should fall off as $|\mathbf{r}|^{-3}$ and the ionic displacements should fall off as $|\mathbf{r}|^{-2}$.

A related question is also whether to use the theoretical lattice constant of the ideal crystal structure in the defect calculation or to let also the supercell size and shape to relax. The main reason for the usual choice of

using the perfect crystal structure lattice constant is practical, because the relaxation of the supercell size and shape would require considerable extra computational effort. Calculations using silicon vacancy as a test case indicate that for a supercell of 64 atomic sites the calculated defect properties are not significantly affected by the changes in the lattice parameter of up to 1% [16]. The relaxation of the supercell size and shape may also induce unphysical changes in the description of the bulk states.

In the tetrahedrally coordinated covalent materials the distances between the periodic defect images along the rigid [110] zigzag-chain directions are probably more important to the convergence of the lattice relaxations than the volume relaxations. In the case of the silicon vacancy this is discussed in references [16, 17]. It was also shown that the Born-Oppenheimer energy surface for the silicon vacancy in particular is extremely flat, with the total-energy differences between different ionic configurations a few tenths of electron volt.

3.2 Defect level dispersion

Ideally, the defect levels induced by isolated point defects corresponding to localized deep-level states are flat, *i.e.* they show no dispersion as a function of the wave vector **k**. In the supercell method the defects form a superlattice, which introduces spurious dispersion in the defect levels. For the silicon vacancy the dispersion of the deep-level states was estimated from the free electron dispersion to be 0.7 eV for the supercell of 64 atomic sites and still 0.2 eV for the supercell of 216 atomic sites [16]. These numbers should be compared to a typical LDA energy band gap of 0.5 eV in silicon. The typical LDA underestimation of the band gap of roughly 50 % in semiconductors and insulators further enhances the errors caused by the defect level dispersion. A careful analysis of the dispersion characteristics of the deep levels can be used to explain the wide spectrum of results for the silicon vacancy calculated using different **k**-sets for first Brillouin zone sampling [16].

The dispersion of the defect levels may lead to the mixing of the valence or conduction band wave functions with the defect induced deep-level wave functions leading to the enhancement of the delocalized character of the defect orbitals and even to an incorrect occupation of the single-particle states in some regions of the Brillouin zone. The choice of generally optimal **k**-point sampling is discussed in references [18, 19, 20, 21, 22], with the work

presented in the last reference [22] more or less completing the discussion. In fact the discussion of the generally optimal Brillouin-zone sampling is analogous to the optimal sampling of a general periodic function, as needed for example in the context of the discrete Fast-Fourier Transform (FFT) [22, 20].

3.3 Electrostatic Madelung energy

Due to the long range of the Coulomb interaction, the Madelung energy of a charged infinite (super)lattice per unit cell is not well defined. A simple standard solution, which is especially easy to implement in the PW expansion, is to introduce a homogeneous, neutralizing background charge density (Figure 3.1). In the limit of the infinitely large supercell the induced error vanishes as the defect-defect distance increases and the density of the introduced compensation charge vanishes. However, there is no guarantee that the convergence with respect to the supercell size would be particularly fast.

The origin and the necessity of the Madelung-type correction for the electrostatic energy of a charged point defect in the supercell method can be demonstrated by a simple example. Let us consider the Coulomb energy of an isolated Gaussian charge distribution and compare that to the value calculated with the supercell approximation, *i.e.* , an infinite array of corresponding charge distributions immersed in a uniform, neutralizing background. The size of the supercell is given in terms of its linear dimension L, defined as the cube root of the supercell volume, $L = \sqrt[3]{V}$.

The Coulomb energy of an isolated Gaussian charge distribution

$$\rho_G(\mathbf{r}) = \left(\sqrt{\frac{\kappa}{\pi}}\right)^3 \exp(-\kappa r^2), \tag{23}$$

with the corresponding electrostatic potential

$$\phi_G(\mathbf{r}) = \frac{\mathrm{erf}\,(\sqrt{\kappa} r)}{r}, \tag{24}$$

$$\mathrm{erf}\,(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt, \tag{25}$$

is easily evaluated from

$$E_{isolated} = \frac{1}{2} \int \phi_G(\mathbf{r}) \rho_G(\mathbf{r}) \mathbf{dr} = \frac{1}{2} 4\pi \int r^2 \phi_G(r) \rho_G(r) dr = \sqrt{\frac{\kappa}{2\pi}}. \tag{26}$$

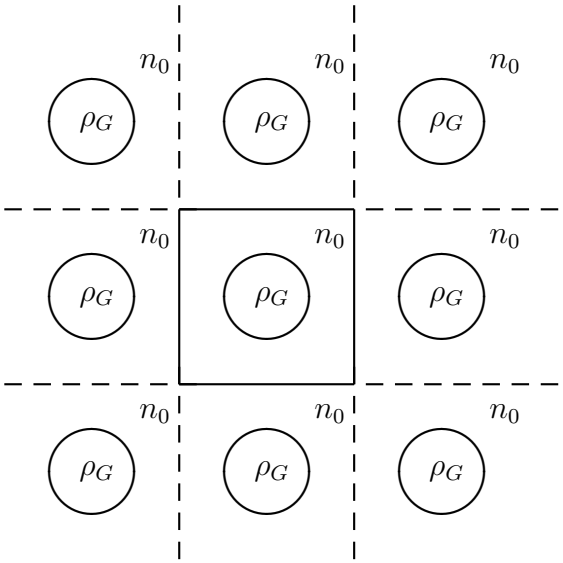

Figure 3.1: Periodic array of defects in the supercell method. The basic method to render the supercell charge-neutral in the calculations of charged defects is to use a uniform, homogeneous background charge density n_0. In the large supercell limit the spurious defect-defect interactions between the defect-charge distributions ρ_G and the uniform compensating background vanish.

The energy per one unit cell for an array of charge distributions with the uniform neutralizing background is evaluated using the Fourier series representation for the corresponding potential

$$\phi_G^P(\mathbf{r}) = \sum_{\mathbf{k}\neq 0} \frac{4\pi}{\Omega k^2} \exp(\frac{k^2}{4\kappa}) \exp(i\mathbf{k}\cdot\mathbf{r}), \tag{27}$$

where $\mathbf{k}$ runs over the reciprocal lattice, leading to

$$\begin{aligned} E_{periodic} &= \frac{1}{2}\int \phi_G^P(\mathbf{r})\rho_G(\mathbf{r})\mathbf{dr} \\ &= \frac{1}{2}\frac{4\pi}{\Omega}\left(\frac{\kappa}{\pi}\right)^{3/2} \sum_{\mathbf{k}\neq 0} \frac{1}{k^2} \exp(\frac{k^2}{4\kappa}) \underbrace{\int \exp(-\kappa r^2)\exp(i\mathbf{k}\cdot\mathbf{r})\mathbf{dr}}_{\left(\frac{\kappa}{\pi}\right)^{3/2}\exp(\frac{k^2}{4\kappa})} \\ &= \frac{2\pi}{\Omega}\sum_{\mathbf{k}\neq 0}\frac{1}{k^2}\exp(\frac{k^2}{2\kappa}). \end{aligned} \tag{28}$$

In the spherical unit cell approximation and at the continuum limit we can integrate analytically $\left(k^2 \to \frac{2\pi}{L}, \sum_{\mathbf{k}\neq 0} \to 4\pi\int_{r_0}^{\infty} r^2 dr, r_0 = \sqrt[3]{\frac{3}{4\pi}}\right)$:

$$\begin{aligned} E_{periodic} &= \sqrt{\frac{\kappa}{2\pi}}\left[1 - \mathrm{erf}\left(\sqrt[3]{\frac{3\pi}{4}}\sqrt{\frac{2\pi}{\kappa}}\frac{1}{L}\right)\right] \\ &= \sqrt{\frac{\kappa}{2\pi}} - \sqrt[3]{\frac{6}{\pi}}\frac{1}{L} + \frac{\pi}{\kappa}\left(\frac{1}{L}\right)^3 + \mathcal{O}\left[\left(\frac{1}{L}\right)^5\right], \end{aligned} \tag{29}$$

and the series expansion clearly brings up the L dependent difference in the Coulomb energy between the isolated charge distribution and the corresponding supercell approximation.

The spurious Madelung energy contribution in the supercell method was first realized by Leslie and Gillan [23]. Assuming that the unit cell is large enough, the difference between the Coulomb energy of an isolated defect and that of a superlattice of defects can be estimated using a macroscopic approximation. Leslie and Gillan considered the Madelung energy of the periodic array of point charges q with a neutralizing uniform background

immersed in structureless dielectric material described by the dielectric constant ϵ of the perfect crystal. This immediately leads to the first-order (monopole) correction

$$\Delta E_1 = \frac{q^2 \alpha}{2\epsilon L}, \tag{30}$$

where α is the Madelung constant corresponding to the Bravais lattice of the supercell. We see immediately that the leading order of the correction is $1/L$, *i.e.* the convergence with respect to supercell size is relatively slow, and that the absolute magnitude of the error increases for more ionic materials with smaller dielectric screening constants. We also note that the correction depends on the charge state of the defect as q^2, quickly leading to large corrections for the high charge states, ± 3, ± 4, etc.

Makov and Payne [24] took the idea of point charges further by considering charge distributions in cubic lattices. A general systematic correction was also proposed by Nozaki and Itoh [25], regrettably omitting the effects of screening in their method and using a test calculation which does not allow the polarization of the crystal. A slightly different approach, the local-moment counter-charge (LMCC) method, was proposed by Schultz [26, 27], which offers an alternative way to estimate the defect-defect Coulomb interactions. The LMCC method is also based on the classical electrostatic description and the linearity of the Poisson equation.

The main reason why the proposed corrections do not offer an immediate solution in insulators and semiconductors for intermediate-size supercells is that we cannot simply separate the external charge distribution from the induced screening charge as required by the classical treatment. For large enough supercells the changes in the local screening around the defect become smaller and the remaining long-range screening effects can be approximated by direct extrapolation from different supercell sizes according to the $1/L$ convergence. A rather clear picture of the convergence of the total energies can be achieved by using simultaneously two different charge compensation schemes (Figure 3.2), uniform and localized, as in Publication II of this thesis.

Finally we note that in the case of systems separated by vacuum, where the charge density vanishes at the borders of the supercell, all the proposed corrections based on the classical analysis can be successfully used up to high precision in practice. At the other extreme limit there are metallic

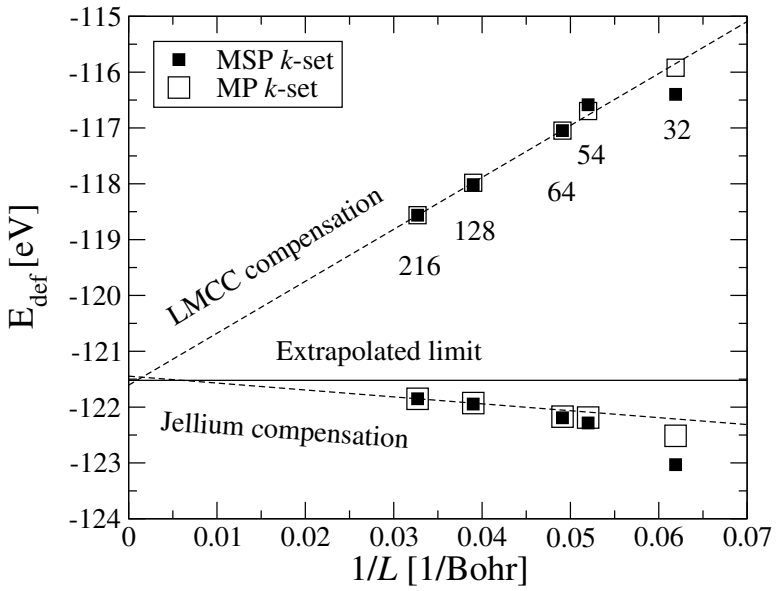

Figure 3.2: The total-energy convergence as a function of the linear dimension of the supercell L for a silicon interstitial at tetrahedral site with the charge $q = +2$. The error in the total-energy difference $E_{def}(L) = E_{N+1}(L) - E_N(L)$, where E_{N+1} and E_N are the total-energies of the supercells with and without the interstitial, respectively, directly transfers to the formation energy estimates. The total-energies are calculated using both the jellium and the LMCC charge-compensation schemes. The filled (MSP) and open squares (MP) refer to two different first Brillouin-zone sampling schemes. The horizontal line is drawn at the average value of the jellium and LMCC extrapolations at 121.5 eV. The clear superiority of the homogeneous compensation over the LMCC method in this particular case, *i.e.* the smaller slope, is explained by the rather large dielectric screening constant of silicon, $\epsilon = 12$. The situation would be reversed for more ionic materials with ϵ closer to unity [Publication II].

systems, in which the metallic screening completely shields the defect charge already over short length scales, and which thus require no Madelung-type electrostatic corrections.

3.4 Supercell symmetry

An interesting effect besides the elastically mediated defect-defect interaction was reported in a thorough study of the neutral silicon vacancy by Probert and Payne [17]. They found that the formation energy of the unrelaxed vacancy depends strongly on the supercell symmetry, the body-centered cubic (BCC) cell yielding a well-converged value already with a 32 atomic-site supercell, superior to the 64 atomic-site simple cubic (SC) or the 250 atomic-site face-centered cubic (FCC) supercells. The reason for this lies in the directionality of the defect-defect interaction and becomes apparent when plotting the charge density difference between the unrelaxed vacancy and the perfect crystal unit cell. The superlattice gives rise to a spurious long-range component in the aperiodic defect charge aligned along the nearest neighbour defect-defect directions. For the BCC supercells this spurious long-range component is in the [111] direction, and thus is commensurate with the underlying host-lattice bond directions. On the other hand, the long-range component of the aperiodic charge for the SC and the FCC supercells with [100] and [110] nearest neighbour defect directions, respectively, is not commensurate with the host-lattice bond directions. In fact, for the FCC superlattice the long-range component of the aperiodic charge coincides with the elastically rigid [110] direction.

4 Single-electron models for vacancies

The structure of tetrahedrally coordinated semiconductors is characterized by covalent, directional sp^3 bonding. When a neighbouring ion is missing, as for example in the case of a vacancy or a surface, the broken-bond orbitals left are called dangling bonds (DB). Due to the electronic and ionic relaxations DB is a qualitative concept.

One of the most studied defects is the mono-vacancy in silicon, which is often taken as the prototype of point defects in semiconductors. Curiously, it has also proved to be one of the most demanding test systems for the state-of-the-art calculations.

Vacancies in silicon are usually produced by using high-energy electron collisions at low temperatures. The interstitials in the collision-induced interstitial-vacancy Frenkel pairs are extremely mobile even at 4.2 K, and thus the Frenkel pair is either immediately annihilated or the interstitial escapes and gets trapped at some other impurity. As the migration energy of the vacancy is around 150 K, the isolated vacancies can be observed experimentally.

The experimental data for the silicon vacancy is provided mostly by the electron paramagnetic resonance (EPR) spectroscopy, the electron-nuclear double resonance (ENDOR) spectroscopy and the deep-level transient spectroscopy (DLTS) techniques and are summarized in the review article by Watkins [28]. These methods yield information on the symmetry and the spatial distribution of the highest unpaired localized electron state, and the ionization levels defined as the values of the electron chemical potential at which the defect changes its charge state. The positron lifetime (PLT) spectroscopy directly measures the changes in the open volumes in the lattice [29]. The open volumes associated with the vacancy-phosphorous pair have also been estimated using DLTS under hydrostatic pressure [30].

4.1 Watkins and Schlüter models

The symmetry relaxation patterns for the vacancy in different charge states observed in experiments are usually discussed using the Watkins model [28], based on the linear combination of atomic orbitals (LCAO). In the doubly positive charge state the fully symmetric linear combination of the sp^3 type

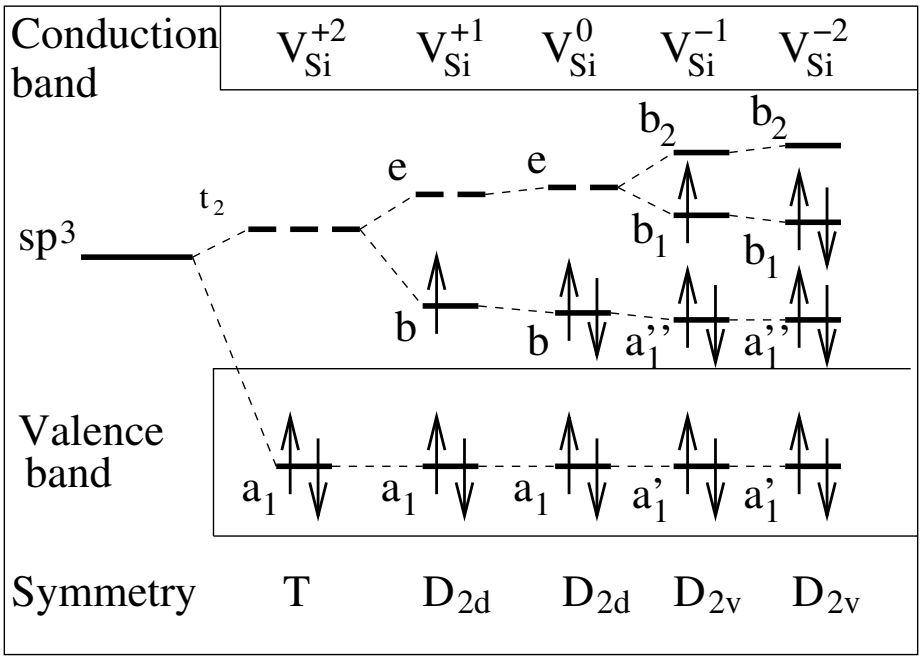

Figure 4.1: Schematic presentation of the filling of the defect levels in the Watkins model for the vacancy.

DB orbitals leads to the defect level a_1 which lies energetically deep inside the valence band. The unoccupied triply degenerate defect level t_2 lies in the energy gap region and the symmetry of the defect remains compatible with the perfect lattice point group T_d. As one electron is added to the defect, changing the charge state of the defect from +2 to +1, the symmetry is lowered to D_{2d}, and the degeneracy of the t_2 level is lifted. The splitting of the degenerate level associated with the symmetry lowering Jahn-Teller (JT) distortion leads to a lower total energy for the system. In the negative charge states the symmetry of the defect is further reduced to C_{2v}. The filling of the single-particle levels in different charge states according to the Watkins model is shown in Figure 4.1.

The competing total-energy contributions in the JT effect and the symmetry-breaking pairing-mode distortion (E) can be analyzed using the simple Schlüter model Hamiltonian [31],

$$E_E = -V_E Q_E + \frac{1}{2} k_E Q_E^2, \tag{31}$$

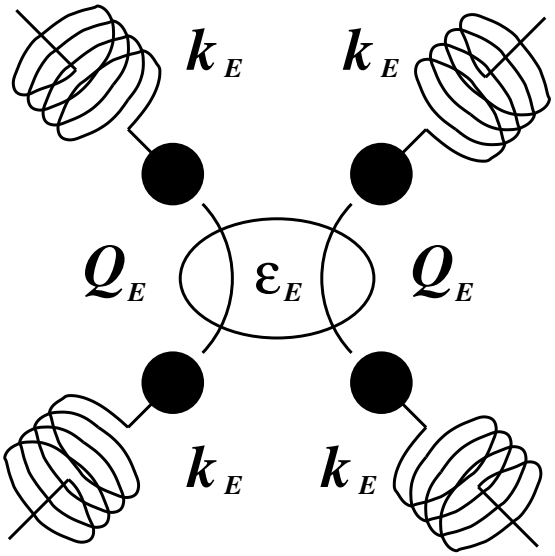

Figure 4.2: Schematic presentation of the Schlüter model for the vacancy.

schematically presented in Figure 4.2. The linear Jahn-Teller coupling coefficient $V_E = -\partial\epsilon_E/\partial Q_E$ describes how the orbital energy changes as a function of the generalized coordinate Q_E associated with the pairing mode. The corresponding force constant k_E can be estimated from phonon frequency calculations [Publication I]. The magnitude of the energy gain $E_{\rm JT}$ from the JT distortion can be estimated by minimizing the energy with respect to Q_E, leading to

$$E_{\rm JT} = -V_E^2/2k_E. \tag{32}$$

An interesting effect arises when the energy gain from the JT distortion exceeds the Coulomb repulsion U between two localized electrons at the defect. This circumstance of the negative-effective-U effect for the silicon vacancy is realized in the direct transition from the doubly positive to the neutral charge state $(++/0)$. The electron capture in the doubly positive charge state is immediately followed by the capture of a second electron and the singly positive charge state is not stable for any value of the electron chemical potential [28].

In addition to the pairing mode distortion leading to the Jahn-Teller effect, vacancies also undergo breathing mode relaxations. The breathing mode

relaxation, which can be included in the Schlüter model analysis, does not induce a change in the defect symmetry.

An alternative occupation scheme for the defect levels is realized in the cases where the exchange interaction dominates over JT effect. The strong exchange interaction leads to the occupation of the triply degenerate t_2 level according to the Hund's rule. The resulting high-spin ground states for the electronic configuration preserve the T_d point group symmetry. Examples of such systems are vacancies in diamond [32], silicon vacancy in SiGe [Publication I] and in SiC [33, 34].

Despite its apparent simplicity, the calculation of the electronic structure and the ionic relaxations of the vacancy in silicon has proved to be a highly non-trivial task. The wide spread of the theoretical results is comprehensively discussed in References [16, 17] and in Publication II and in Publication IVof this thesis. The main reasons for the spread in the theoretical results are the delicate coupling between the ionic and electronic degrees of freedom and the relatively flat Born-Oppenheimer energy surface with multiple competing minima, which both make the system particularly sensitive to various computational parameters. To summarize, the calculations (within the LDA) seem to have converged with respect to the symmetry lowering lattice distortions, and reproduce the experimentally observed symmetry-relaxation patterns, at least for the positive and neutral charge states. Interestingly, the DFT calculations employing the standard LDA and the generalized-gradient approximation (GGA) predict a strong *inward* relaxation toward the center of the vacancy, in contrast to earlier Green's function calculations [35] and the intuition based on the relaxation of the surface atoms toward the bulk of the crystal [31]. *Ab initio* calculations using the screened-exchange local-density approximation (sX-LDA), described in Section 2.3, seem to indicate an *outward* relaxation of the vacancy [Publication IV]. Unfortunately, experimental studies [29, 30] do not offer a definite answer, and some ambiguity in even the sign of the breathing-mode relaxation remains.

5 Ground-state properties of point defects

The experimentally measured physical properties of the point defects discussed in this work which are accessible with the total-energy minimization based *ab initio* DFT methods are the local symmetries, the relaxation volumes, the equilibrium concentrations and the ionization levels of point defects.

5.1 Defect geometry

The experimental determination of the local symmetry and the relaxation volume of the vacancy in silicon was already discussed in Section 4. In principle the comparison of the experimentally measured symmetry and the volume relaxations to the calculated ones is straightforward, as the ionic coordinates are directly accessible from the calculations. In practice, however, the Born-Oppenheimer surface may be shallow, possessing several competing local minima, resulting in a delicate dependence of the theoretical results on the computational parameters used, see Section 3.

5.2 Formation energy

The equilibrium concentrations and the ionization levels of point defects are estimated using formation energies E_F calculated from the supercell total energies. At low temperatures the phonon entropy effects are disregarded and the free energy of formation is largely determined by the formation energy [36]. In practice the relative equilibrium concentrations c of point defects are thus estimated using the Boltzmann factor

$$c \sim \exp(-E_F/k_BT). \tag{33}$$

The standard formalism to calculate the formation energies in supercell methods was introduced by Zhang and Northrup [37]. The formation energy of a defect as a function of the electron chemical potential μ_e in a charge state q is calculated from

$$E_F = E_{tot}(q) - \sum_s n_s\mu_s + q(E_V + \mu_e), \tag{34}$$

where $E_{tot}(q)$ is the total energy of the defect-containing supercell. The electron chemical potential μ_e is measured relative to the valence-band maximum E_V of the corresponding bulk-crystal reservoir [Publication IV]. The number of the atom species s in the supercell is n_s, and μ_s is the corresponding atomic chemical potential.

For native defects in elemental semiconductors, such as mono-vacancies or interstitials in Si and Ge [Publication II, Publication IV], the atomic reference energy ($\sum_s n_s\mu_s$) is calculated from the total energy of the supercell of the perfect crystal structure, using $\frac{N\pm1}{N}E_N$, where $N \pm 1$ is the number of atoms in the defect cell and N the number of atoms in the defect free supercell with a total energy of E_N.

For native defects in compound semiconductors, such as SiC which we use here as an example [36], only the sum rule

$$\mu_{\mathrm{SiC}} = \mu_{\mathrm{Si}} + \mu_{\mathrm{C}} \tag{35}$$

and the limits

$$\mu_{\mathrm{Si}} \leq \mu_{\mathrm{Si}}^{bulk}, \tag{36}$$
$$\mu_{\mathrm{C}} \leq \mu_{\mathrm{C}}^{bulk} \tag{37}$$

are available, where μ_{SiC} is the energy per Si-C pair in the SiC crystal, and μ_{Si}^{bulk} and μ_{C}^{bulk} are the respective elemental bulk values. The effect of different growth conditions on the defect formation energy become apparent when we rewrite Equation (34) as

$$E_F = E_{tot}(q) - \frac{1}{2}(n_{\mathrm{Si}} + n_{\mathrm{C}})\mu_{\mathrm{SiC}}^{bulk} - \frac{1}{2}(n_{\mathrm{Si}} - n_{\mathrm{C}})(\mu_{\mathrm{Si}}^{bulk} - \mu_{\mathrm{C}}^{bulk}) \tag{38}$$
$$- \frac{1}{2}(n_{\mathrm{Si}} + n_{\mathrm{C}})\Delta\mu + q(E_V + \mu_e), \tag{39}$$

where the dependence on the growth conditions is in the chemical potential difference defined as

$$\Delta\mu = (\mu_{\mathrm{Si}} - \mu_{\mathrm{C}}) - (\mu_{\mathrm{Si}}^{bulk} - \mu_{\mathrm{C}}^{bulk}). \tag{40}$$

The interval for $\Delta\mu$ is thus from C-rich defect formation conditions ($\Delta\mu = -\Delta H$) to Si-rich conditions ($\Delta\mu = \Delta H$), where the heat of formation of

the SiC crystal is defined as usual, $\Delta H = \mu_{\mathrm{Si}}^{bulk} + \mu_{\mathrm{C}}^{bulk} - \mu_{\mathrm{SiC}}^{bulk}$. In addition to the C- and Si-rich conditions, the calculated formation energies are often reported in the stoichiometric conditions ($\Delta\mu = 0$). For non-native elements, such as dopant or impurity atoms, an estimate for the (reservoir) atomic chemical potential is calculated from a suitable reference system.

Often, defects are not created in thermodynamical equilibrium conditions, but in fact far from it, for example in electron or proton collisions. In such conditions the concentrations of different defects are largely determined by kinetic effects. Also due to the inaccuracies in the determination of the chemical potentials of the constituent ions, some caution is advisable when the absolute values of formation energies are compared to experiments. However, in many cases the calculated relative formation energies can be directly compared to experiments.

In the case of silicon-germanium, which is a random alloy, we can also consider different variants of the vacancy with a different number of germanium neighbours. The formation energy of different variants was shown to depend nearly linearly on the number of germanium neighbours [38], which was rationalized in Publication I and applied in the interpretation of the PLT measurements in Publication III.

5.3 Ionization levels

Ionization levels are defined as the values of the electron chemical potential μ_e at which the thermodynamically stable charge state changes. The calculated ionization levels thus include the effects of lattice relaxations and are related to experiments in which the levels are probed by varying the position of the electron chemical potential (Fermi level) in the bulk electron reservoir. We note that especially for the ionization levels involving negative charge states the absolute positions of the levels are sensitive to the various supercell method and to the LDA induced inaccuracies. A typical formation energy plot including the positions of the ionization levels is shown in Figure 5.1.

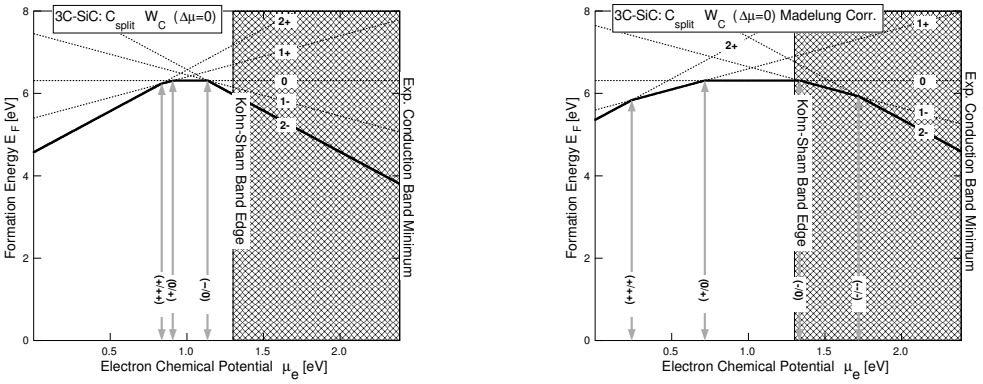

Figure 5.1: Formation energy of the carbon interstitial in the [110]-split configuration in 3C-SiC in different charge states ($\Delta\mu = 0$). The first panel omits the first order Madelung correction whereas the second one includes it [Publication V].

6 Summary of the Publications

Publications I, III, V and VI of this thesis focus on the application of the computational methods presented in this overview. They contribute in the development of the microscopic picture of intrinsic point defects in technologically important tetrahedrally coordinated semiconductors silicon-germanium and silicon-carbide. Publications II and IV focus on the analysis and improvement of the theoretical and computational methods used.

6.1 Publication I: Vacancies in SiGe: Jahn-Teller distortion and spin effects

A vacancy in silicon-germanium ($Si_{1-x}Ge_x$) was studied using a zinc-blende model structure (SiGe). The Schlüter model was used in the analysis of different formation energy contributions for the vacancy defect. The same model was used to characterize the vacancy dangling-bond orbitals in tetrahedrally bound covalent materials C, Si, Ge, SiGe and SiC, and explain the observed differences in the calculated ionic relaxations around the vacancy.

6.2 Publication II: Charged point defects in semiconductors and the supercell approximation

The calculation of charged point defects in semiconductors and insulators using the supercell approximation and several proposed corrections relating the calculated total energies to the corresponding values for isolated point defects were analysed using the silicon self-interstitial and the PWPP method as a relevant example. The shortcomings of the classical electrostatics description in the derivation of the corrections were discussed. A direct extrapolation using two complementary schemes was suggested to achieve reasonable confidence in the convergence of the calculated formation energies.

The standard uniform compensating-charge scheme was shown to correspond to metallic screening of the aperiodic defect charge in the sense that the average electrostatic potential outside the supercell is vanishing. The local-moment counter-charge (LMCC) scheme [26, 27], on the other

hand, was shown to correspond to the other extreme in which the potential outside the supercell equals the electrostatic potential of the isolated bare (unscreened) defect charge. The analytic post-corrections for both the compensation schemes derived from the classical electrostatics require the separation of the aperiodic defect charge from the periodic perfect crystal charge, which is easily achievable. What is not straightforward though, is the separation of the external charge distribution from the screening charge, required in the classical analysis using the static dielectric constant to describe the screening in the crystal.

6.3 Publication III: Vacancy-phosphorous complexes in strained $Si_{1-x}Ge_x$: Structure and stability

The experimental observation of the stabilizing effect of a germanium neighbour to a vacancy defect in silicon-germanium was analysed using a simple dangling-bond (DB) model and verified by direct *ab initio* calculations.

The positron annihilation spectroscopy measurements showed that the dominant defect in the proton irradiated $Si_{1-x}Ge_x$, the vacancy-phosphorous pair (V-P), transforms to a V-P-Ge complex at the annealing temperature of 150–175 °C. This complex anneals out at the temperature of 200 °C. The temperature interval of $\Delta T = 50\,^{\circ}\mathrm{C}$ in which the complex is stable was estimated to correspond to 0.1–0.2 eV larger binding energy for the V-P-Ge complex than for the V-P pair. This estimate is in good agreenment with the performed direct *ab initio* supercell-calculation estimate for the binding energy difference. Based on the previous analysis on Publication I, the result is explained with the different character of the DB orbitals of Si and Ge. The larger extent of the Ge DB orbital leads to a larger overlap of the DB orbitals in the vacancy and thus lower orbital energy for the binding in the linear combination of the DB orbitals.

6.4 Publication IV: Non-local screened-exchange calculations for defects in semiconductors: vacancy in silicon

The screened-exchange local-density approximation was used to calculate the electronic structure of a vacancy in silicon, making this the first application of the method in a large-scale supercell calculation. The non-locality

of the functional introduces a considerable computational complexity compared to the standard LDA and GGA schemes, making the sX-LDA calculations comparable to the Hartree-Fock method in the computational effort. The efficient implementation in the plane-wave basis allowed the structural optimization in the 32-atomic-sites supercell and the self-consistent minimization with respect to the electronic degrees of freedom in the 256-atomic-sites supercell.

In contrast to the LDA and the GGA approximations, the sX-LDA was shown to lead to an outward relaxation of the vacancy nearest neighbours. This qualitative difference was explained with the removal of the overbinding generally associated with the LDA. The example demonstrates the significance of the development of new improved fuctionals, not only to estimate the excited-state properties but also to calculate the structural properties.

6.5 Publication V: Self-Interstitials in 3C-SiC

The formation energies and ionization level positions of the technologically important self-interstitial defects in silicon-carbide were calculated using the *ab initio* PWPP method. The work reported in Publication V is a direct continuation of the work reported in reference [36] and completes the theoretical analysis of simple intrinsic point defects, vacancies and antisites in SiC. The finite-size effects in the results of the supercell calculations were thoroughly analyzed in the light of the recent developements in Publication II and in reference [17].

The formation energies of the interstitial defects were found to lie between the formation energies of the antisite and vacancy type defects, thus indicating their relevance in the diffusion processes. The carbon interstitial was found to have several competing minimum total-energy configurations in the split-interstitial geometries, while the silicon interstitial had a single, clear minimum total-energy configuration at the tetrahedral site surrounded by carbon atoms. The carbon interstitial was found to introduce defect levels in the semiconductor energy gap. The defect levels introduced by the silicon interstitial at the tetrahedral site were shown to mix with the conduction band in the 128-atomic-sites supercell calculation.

6.6 Publication VI: Interstitial H and H_2 in SiC

The hydrogen impurity in SiC was studied using two complementary approaches, the Gaussian-type local basis and the PW basis, within the DFT.

Minimum total-energy configuration for H^+ (proton) was found at the site antibonding to the carbon atom (ab-C) and the minimum total-energy configuration for H^- was found to be a tetrahedral site surrounded by silicon atoms (T_{Si}). The ionization levels (+/-) were calculated to be 0.9 eV and 1.3 eV above the valence band maximum in the 3C and 4H polytypes, respectively. The neutral hydrogen atom (H^0) was not found to be stable in the thermodynamical equilibrium for any value of the electron chemical potential.

The calculated diffusion barrier for H^+ was found to be 0.5 eV, implying high mobility at the room temperature. The diffusion barrier for the singly negative H^- was estimated to be considerably higher, of the order of 2.0 eV. The high mobility of H^+ at room temperature, with the calculated binding energy for two neutral H atoms of 1.3 eV/molecule in 3C-SiC and 2.9 eV/molecule in 4H-SiC, compared to the binding free energy of an isolated H_2 of 4.52 eV, indicates a high likelihood of the H_2 formation in SiC. The minimum total-energy position in the SiC lattice for the H_2 was found at the T_{Si} site, with the axis of the molecule parallel to the lattice c-axis. The local vibrational modes of the H_2 molecule and those of the H impurity in different charge states were also calculated to facilitate comparison to Raman spectra measurements.

References

[1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864

[2] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133

[3] Perdew J P, Burke K and Ernzernhof M 1996 *Phys. Rev. Lett.* **77** 3865

Langreth D C and Mehl M J 1983 *Phys. Rev. B* **28** 1809

Becke A D 1988 *Phys. Rev. A* **38** 3098

[4] Seidl A, Görling A, Vogl P, Majewski J A and Levy M 1996 *Phys. Rev. B* **53** 3764

[5] Jones R O and Gunnarsson O 1989 *Rev. Mod. Phys.* **61** 689

[6] Levy M and Perdew J P 1985 in *Density functional methods in Physics*, vol 123 of *NATO Advanced Science Institutes Series* ed. Dreizler R M and da Providencia J (New York: Plenum Press) Chap. 2, p 11

Parr R G and Yang W 1989 *Density-Functional Theory of Atoms and Molecules*, International series of Monographs on Chemistry, No. 16 (New York: Oxford Unversity Press)

Dreizler R M and Gross E K U 1990 *Density-Functional Theory* (Berlin: Springer)

Levy M 1990 *Advances in Quantum Chemistry*, ed. Trickey S B (San Diego: Academic Press), vol 21, p 69

[7] Ceperley D M and Alder B J 1980 *Phys. Rev. Lett.* **45** 566

[8] Bylander B M and Kleinman L 1990 *Phys. Rev. B* **41** 7868

[9] Engel G E 1997 *Phys. Rev. Lett.* **78** 3515

[10] Perdew J P, Burke K and Wang Y 1996 *Phys. Rev.* B **54** 16533

Perdew J P, Jackson K A, Pederson M R, Singh D J and Fiolhais C 1992 *Phys. Rev.* B **46** 6671

Perdew J P, Jackson K A, Pederson M R, Singh D J and Fiolhais C 1993 *Phys. Rev.* B **48** 4978(E)

[11] Mohammed A R E and Sahni V 1984 *Phys. Rev. B* **29** 3687

[12] Asahi R, Mannstadt W and Freeman A J 1999 *Phys. Rev. B* **59** 7486

Picozzi S, Continenza A, Asahi R, Mannstadt W, Freeman A J, Wolf W, Wimmer E and Geller C B 2000 *Phys. Rev. B* **61** 4677

Stampfl C, Mannstadt W, Asahi R and Freeman A J 2001 *Phys. Rev. B* **63** 155106

[13] Chawla S and Voth G A 1998 *J. Chem. Phys.* **108** 4697

[14] Torsti T 2003 *Real-space electronic structure calculations for nanoscale systems*, CSC research reports R01/03 (Espoo: CSC - Scientific Computing Ltd.)

[15] Stoneham A M 1975 *Theory in Defects and Solids: Electronic Structure of Defects in Insulators and Semiconductors*, Monographs on the Physics and Chemistry of Materials (New York: Oxford University Press) p 181

[16] Puska M J, Pöykkö S, Pesola M, and Nieminen R M 1998 *Phys. Rev. B* **58** 1318

[17] Probert M I J and Payne M C 2003 *Phys. Rev. B* **67** 075204

[18] Chadi D J and Cohen M L 1973 *Phys. Rev. B* **8** 5747

[19] Monkhorst H J and Pack J D 1976 *Phys. Rev. B* **13** 5188

[20] Froyen S 1992 *Phys. Rev. B* **45** 3796

[21] Makov G, Shah R and Payne M C 1996 *Phys. Rev. B* **53** 15513

[22] Moreno J and Soler J M 1992 *Phys. Rev. B* **45** 13891

[23] Leslie M and Gillan M J 1985 *J. Phys. C: Solid State Phys.* **18** 973

[24] Makov G and Payne M C 1995 *Phys. Rev. B* **51** 4014

[25] Nozaki H and Itoh S 2000 *Phys. Rev. E* **62** 1390

[26] Schultz P A 1999 *Phys. Rev. B* **60** 1551

[27] Schultz P A 2000 *Phys. Rev. Lett.* **85** 1942

[28] Watkins G D 1986 in *Deep Centers in Semiconductors*, edited by S. T. Pantelides (New York: Gordon and Breach) p 147

[29] Mäkinen J, Hautojärvi P and Corbel C 1992 *J. Phys: Condens. Matter* **4** 5137

[30] Samara G A 1988 *Phys. Rev. B* **37** 8523

Samara G A 1989 *Phys. Rev. B* **39** 12764

[31] Baraff G A, Kane E O and Schlüter M 1980 *Phys. Rev. B* **21** 3563

Baraff G A, Kane E O and Schlüter M 1980 *Phys. Rev. B* **21** 5662

[32] Isoya J, Kanda H, Uchida Y, Lawson S C, Yamasaki S, Itoh H and Morita Y 1992 *Phys. Rev. B* **45** 1436

[33] Zywietz A, Furthmüller J and Bechstedt F 1999 *Phys. Rev. B* **59** 15167

[34] Torpo L, Nieminen R M, Laasonen K E, and Pöykkö S 1999 *Appl. Phys. Lett.* **74** 221

[35] Scheffler M, Vigneron J P and Bachelet G B 1985 *Phys. Rev. B* **31** 6541

[36] Torpo L, Marlo M, Staab T E M and Nieminen R M 2001 *J. Phys: Condens. Matter* **13** 6203

[37] Zhang S B and Northrup J E 1991 *Phys. Rev. Lett.* **67** 2339

[38] Boguslawski P and Bernholc J 1999 *Phys. Rev. B* **59** 1567