universität wien

DISSERTATION

Titel der Dissertation

Ab initio methods applied to oxides,
and atomic orbitals in the PAW framework

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr. rer.nat.)

Verfasser:	Dipl.-Ing. Roman Wahl
Matrikel-Nummer:	0156463
Dissertationsgebiet:	Physik
Betreuer:	Univ.-Prof. Dipl.-Ing. Dr. Georg Kresse

Wien, am 11. November 2010

Im Andenken an meine Mutter

Acknowledgements

First of all, I would like to thank Univ.-Prof. Dipl.-Ing. Dr. Jürgen Hafner and Univ.-Prof. Dipl.-Ing. Dr. Georg Kresse who gave me the unique opportunity to write this thesis at the Computational Materials Physics group. As my supervisor Prof. Kresse has always offered me his advice and support and I am especially grateful for his thoughtful behaviour in times of hardship.

Furthermore I would like to thank all my colleagues and friends who have supported me in many different ways throughout the years.

I am deeply indebted to Doris Vogtenhuber, Michael Pöltl, Laurids Schimka and Martijn Marsman who have always been prepared to listen to my professional and private concerns and who have thereby become real friends.

Last but not least I would like to thank my wife Marlene who has always supported me in every possible way, my sister Martina who has always given me her advice when needed and my parents whose love, confidence and patience have made me the person I am now.

Abstract

The accuracy of ab initio density functional theory calculations depends crucially on the applied exchange-correlation functional. In recent years, a true zoo of functionals has emerged in particular in the quantum chemistry community. Either the functional depends on the density alone (LDA), or additionally on the gradient of the density (GGA) and the kinetic energy density (Meta-GGA). Equally common are so called hybrid functionals, that add a small amount of exact Hartree-Fock like exchange. All these developments were aimed on an improved description of different properties of molecules, solids and surfaces. But very little is known for complex solids.
In the first part of this thesis we apply four different exchange-correlation functionals, the local density approximation (LDA), gradient corrected functionals (PBE, PBEsol) and hybrid functionals (HSE03) to two model perovskites, namely $SrTiO_3$ and $BaTiO_3$. We evaluate structural, electronic and phonon properties in the high temperature cubic as well as the tetragonal distorted phase. We focus our investigations specifically on the phonon frequencies and their dependence on the volume and used exchange-correlation functional. Due to the large variation of theoretical predictions for the frequency of the Γ_{15} (TO1) zone-centred phonon mode special attention is devoted to this particular mode. Our results show that the ferroelectric instability strongly depends on the volume and to a lesser extent on the applied functional. Specifically, a transition from a hard ferroelectric mode at small volumes to a soft mode at larger volumes is predicted, with the transition occurring roughly at the equilibrium volume of $SrTiO_3$. This makes it so particularly hard to predict whether $SrTiO_3$ is unstable towards a ferroelectric distortion or not.
The second part of the thesis concentrates on the description of polar surfaces, in particular ZnO. When wurtzite ZnO is sliced perpendicular to the (0001) axis, two different polar surface, the (0001)-Zn and $(000\bar{1})$-O terminated surfaces, are formed. In the simple ionic picture both surfaces are electrostatically unstable due to a diverging electrostatic energy. The surfaces have to adopt a modified surface structure to compensate for the polarity. In close collaboration with experiment, a novel reconstruction is suggested. The remarkable observation is that the previously investigated $(000\bar{1})$-O surface behaves very different than the (0001)-Zn surface. The

difference is explained by the different bond preferences of Zn and O atoms: as a *d*-element Zn atoms are more flexible than O atoms.
The final part of the thesis concentrates on the implementation of a local basis set code in VASP. We derive all required formulas for the projector augmented wave method and an atomic orbital basis set expanded in spherical Bessel functions. The atomic orbital expansion was implemented in the VASP code. Particular emphasis was placed on an implementation that maintains the translational invariance of the Hamiltonian to a very high accuracy, e.g. when all atoms are rigidly moved in the unit cell, the energy should not change significantly. The basis set was optimized to achieve this goal. This is demonstrated for few simple test cases.

Zusammenfassung

Die Genauigkeit von ab-initio Dichtefunktional Rechnungen hängt entscheidend von den verwendeten Austausch-Korrelations Funktionalen ab. In letzter Zeit hat sich hierbei, vor allem in der Quantenchemie, ein regelrechter Funktional-Zoo herausgebildet. Manche der Funktionale hängen nur von der Dichte ab (LDA), andere auch vom Gradienten der Dichte (GGA) oder sogar von der Dichte der kinetischen Energie (Meta-GGA). Ebenso verbreitet sind die sogenannten Hybridfunktionale, die einen gewissen Anteil von exaktem Hartree-Fock Austausch beimischen. All diese Entwicklungen haben zum Ziel, die Beschreibung der physikalischen Eigenschaften von Molekülen, Festkörpern und Oberflächen zu verbessern.
Im ersten Teil dieser Dissertation werden vier verschiedene Austausch-Korrelations Funktionale, und zwar die lokale Dichte Näherung (LDA), Gradientenkorrigierte Funktionale (PBE, PBEsol) und ein Hybridfunktional (HSE03) an zwei Perovskiten, $SrTiO_3$ und $BaTiO_3$, getestet. Dabei werden strukturelle, elektronische und phononische Eigenschaften in der kubischen und tetragonal verzerrten Phase berechnet. Im Zentrum der Aufmerksamkeit stehen dabei insbesondere die Phononenfrequenzen und ihre Abhängigkeit vom Volumen der Einheitszelle und dem jeweiligen Funktional. Da sich die bislang vorhandenen theoretischen Vorhersagen der Frequenz des zonen-zentrierten Phononenmodes Γ_{15} (TO1) stark unterscheiden, wird diese besonders intensiv untersucht. Unsere Ergebnisse zeigen, dass die ferroelektrische Instabilität stark vom Volumen, jedoch kaum vom verwendeten Funktional abhängt. Die vorliegenden Berechnungen sagen für die ferroelektrische Mode einen Übergang von hart auf weich bei Vergrößerung des Volumens voraus. Das Volumen, bei dem der Übergang stattfindet, liegt in der Nähe des Gleichgewichtsvolumens, weshalb es so schwierig ist, theoretisch vorherzusagen, ob sich $SrTiO_3$ gegenüber einer ferroelektrischen Verzerrung stabil verhält oder nicht.
Der zweite Teil dieser Arbeit ist der Beschreibung polarer Oberflächen, insbesondere jener von ZnO, gewidmet. Wenn man ZnO in der Wurtzitstruktur senkrecht zur (0001)-Achse spaltet, entstehen zwei nicht äquivalente Oberflächen, die (0001)-Zink und die $(000\bar{1})$-Sauerstoff begrenzte Oberfläche. Im Rahmen einer einfachen ionischen Beschreibung erweisen sich beide Oberflächen als instabil, aufgrund der divergierenden elektrostatischen Energie.

Die Oberflächen müssen daher eine modifizierte Oberflächenstruktur annehmen, um die Polarität zu kompensieren. Wir schlagen in enger Anlehnung an experimentelle Befunde ein neues Modell vor. Dabei stellt sich heraus, dass sich die Sauerstoff begrenzte Oberfläche anders verhält als die Zink begrenzte Oberfläche. Dieser Unterschied kann durch das verschiedene Bindungsverhalten der Zink- und der Sauerstoffatome erklärt werden: die d-Elektronen im Zink sind flexibler als die p-Elektronen im Sauerstoff.
Der letzte Teil dieser Dissertation beschäftigt sich mit der Implementierung eines lokalen Basissatzes in VASP. Alle für die PAW-Methode notwendigen Formeln werden abgeleitet. Der lokale Basissatz wird nach sphärischen Besselfunktionen entwickelt. Besonderes Augenmerk wird auf den Erhalt der Translationsinvarianz des Hamiltonoperators gelegt. Der Basissatz wird dahingehend optimiert, dass dieses Ziel erreicht wird. Dies wird anhand einiger Testrechnungen überprüft.

Contents

List of Figures

List of Tables

Part I

Theoretical background

Chapter 1

Many-body problem and density functional theory

1.1 Many-body Schrödinger equation

The time and spin independent Hamiltonian H for an interacting system of n electrons and N nuclei is given by

$$H = -\frac{\hbar^2}{2m_e}\sum_{i=1}^{n}\nabla_i^2 - \sum_{I=1}^{N}\frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_{i=1}^{n}\sum_{I=1}^{N}\frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I=1}^{N}\sum_{J>I}^{N}\frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}. \quad (1.1)$$

The electrons with mass m_e are denoted by lower case indices and the nuclei, with masses M_I and charges Z_I are denoted with upper case indices. Electronic and ionic spatial degrees of freedom are abbreviated in the following with

$$\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n \equiv \mathbf{r} \qquad \mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_N \equiv \mathbf{R}. \quad (1.2)$$

In Eq. (1.1) the first two terms are the kinetic energy operators for the electrons (T_e) and nuclei (T_N), the third term represents the interaction between the electrons and the nuclei (V_{Ne}). The last two terms describe the repulsion between electrons (V_{ee}) and the nuclei (V_{NN}) respectively.
It is common practice in the density functional community to use Hartree atomic units ($\hbar = m_e = e = 4\pi/\epsilon_0 = 1$) and with the former abbreviations we can simplify the notation of Eq. (1.1) to

$$H = T_e + T_N + V_{Ne}(\mathbf{r}, \mathbf{R}) + V_{ee}(\mathbf{r}) + V_{NN}(\mathbf{R}) \quad (1.3)$$

and finally get to the Schrödinger equation

$$\left[T_e + V_{Ne}(\mathbf{r}, \mathbf{R}) + V_{ee}(\mathbf{r}) + T_N + V_{NN}(\mathbf{R})\right]\Phi(\mathbf{r}, \mathbf{R}) = E\ \Phi(\mathbf{r}, \mathbf{R}). \quad (1.4)$$

1.1.1 Born-Oppenheimer approximation

Within the Born-Oppenheimer or adiabatic approximation it is assumed that the motion of the electrons and the ionic motion are on different time scales due to the large mass difference between the electrons m_e and the ions M_I. (A proton is approximately 2000 times heavier than an electron.) It is further assumed that the electrons instantaneously follow the motion of the nuclei and the wave function $\Phi(\mathbf{r}, \mathbf{R})$ can be separated in an electronic wave function $\Psi(\mathbf{r}; \mathbf{R})$, parametrically depending on the ionic coordinates $\mathbf{R}$ which is depicted by a semicolon, and the ionic wave function $\chi(\mathbf{R})$:

$$\begin{aligned}\left[T_e + V_{Ne}(\mathbf{r}; \mathbf{R}) + V_{ee}(\mathbf{r}) + T_N + V_{NN}(\mathbf{R})\right] \Psi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}) \\ = E_{\text{tot}}\ \Psi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}).\end{aligned} \tag{1.5}$$

Another consequence of the mass difference between electrons and nuclei is that the nuclear components of the wave function are spatially more localized than the electronic components. As a consequence it follows that the nuclear wave function rises more steeply than the electronic wave function, which means that $\nabla_I \chi(\mathbf{R}) \gg \nabla_I \Psi(\mathbf{r}; \mathbf{R})$, from which we can approximate

$$\begin{aligned}T_N \Psi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}) &= -\frac{1}{2M_I} \sum_{I=1}^{N} \nabla_I^2 \Psi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}) \\ &= -\frac{1}{2M_I} \sum_{I=1}^{N} \big[\Psi(\mathbf{r}; \mathbf{R})\nabla_I^2 \chi(\mathbf{R}) + 2 \cdot \nabla_I \Psi(\mathbf{r}; \mathbf{R})\nabla_I \chi(\mathbf{R}) \\ &\qquad + \chi(\mathbf{R})\nabla_I^2 \Psi(\mathbf{r}; \mathbf{R})\big] \\ &\approx -\frac{1}{2M_I} \sum_{I=1}^{N} \Psi(\mathbf{r}; \mathbf{R})\nabla_I^2 \chi(\mathbf{R}) \\ &\approx \Psi(\mathbf{r}; \mathbf{R}) T_N \chi(\mathbf{R}).\end{aligned} \tag{1.6}$$

Since the electronic kinetic operator T_e does not act on the ionic wave function, and the contributions of the ionic kinetic operator T_N acting on the electronic wave function can be neglected, as shown in Eq. (1.6), the Hamiltonian can be separated to

$$\frac{\left[T_e + V_{Ne}(\mathbf{r}; \mathbf{R}) + V_{ee}(\mathbf{r})\right] \Psi(\mathbf{r}; \mathbf{R})}{\Psi(\mathbf{r}; \mathbf{R})} = \frac{\left[E_{\text{tot}} - T_N - V_{NN}(\mathbf{R})\right] \chi(\mathbf{R})}{\chi(\mathbf{R})}. \tag{1.7}$$

The right hand side is depending only on the nuclear position and abbreviated with $E_e(\mathbf{R})$. Since the dependence of the nuclear positions are only of parametric type for the electrons, Eq. (1.7) can be decomposed into two equations, which solely depend on the ionic and electronic wave function.

$$\left[T_e + V_{Ne}(\mathbf{r}; \mathbf{R}) + V_{ee}(\mathbf{r})\right] \Psi(\mathbf{r}; \mathbf{R}) = E_e(\mathbf{R})\Psi(\mathbf{r}; \mathbf{R}) \tag{1.8}$$

$$\left[T_N + V_{NN}(\mathbf{R}) + E_e(\mathbf{R})\right] \chi(\mathbf{R}) = E_{\text{tot}}\chi(\mathbf{R}) \tag{1.9}$$

The electronic equation (1.8) is the fundamental Schrödinger equation for the theory of the electronic structure. The electronic energy $E_e(\mathbf{R})$ acts as a potential energy in the nuclear equation (1.9) and the nuclei move on this potential energy surface.
Since we will only use the electronic Schrödinger equation throughout this thesis we will not continue to denote the parametric dependence of the electronic wave functions on the nuclear coordinates for the ease of reading.

1.2 Hohenberg-Kohn theorems - the roots of DFT

The solution of the electronic Schrödinger equation (1.8), the many-body wave function $\Psi(\mathbf{r})$, depends on all electronic spatial coordinates. Within the current computational resources it is and will be not possible in the near future to solve for the many-body wave function for realistic systems and molecules.

As a result approximations have to be applied to the many-body problem to make calculations feasible. In 1964, P. Hohenberg and W. Kohn [1] published the theoretical basis for Density Functional Theory (DFT). They presented a formalism which was not based on the many-body wave function, but used only the electron density as key quantity. The two basic theorems and the main concept of DFT are presented and proven in the following.

Starting from the electronic many-body Schrödinger equation in the Born-Oppenheimer approximation (1.8), where we substituted V_{Ne} with the more general expression V_{ext},

$$(T_e + V_{\text{ext}} + V_{ee}) \left|\Psi\right\rangle = E_e \left|\Psi\right\rangle \tag{1.10}$$

one recognizes that for a fixed number of electrons, the external potential V_{ext} solely determines the wave functions $|\Psi\rangle$ (assuming non degenerated eigenstates)

$$V_{\text{ext}}(\mathbf{r}) \longrightarrow \Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n). \tag{1.11}$$

Additionally from each wave function Ψ the electron density $n(\mathbf{r})$ can be calculated via

$$n(\mathbf{r}) = \int d\mathbf{r}_2 \ldots d\mathbf{r}_n \Psi^*(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_n)\Psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_n) \tag{1.12}$$

providing a direct mapping from each eigenstate of Eq. (1.10) to the ground state density. In conjunction with the former considerations we obtain the following mapping between the external potential, eigenfunction of the corresponding Hamiltonian and the electron density:

$$V_{\text{ext}}(\mathbf{r}) \longrightarrow \Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n) \longrightarrow n(\mathbf{r}). \tag{1.13}$$

The first Hohenberg-Kohm theorem proves that both mappings, which are surjective by construction, are also injective and therefore invertible. Thus the external potential is connected via a one to one relationship to the electron density. As a consequence the electron density can be used as key quantity which is shown in the second theorem.

First Theorem: *For any system of interacting particles the external potential is uniquely defined, except for a constant, by the ground state particle density* $n_0(\mathbf{r})$.

Proof

First we have to proof the injectivity of $V_{\text{ext}}(\mathbf{r}) \longrightarrow \Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n)$. Thus two potentials V_{ext}, V'_{ext}, which differ by more than a constant, always lead to two different ground state wave functions $|\Psi\rangle$, $|\Psi'\rangle$. Let us suppose the wave functions are identical $|\Psi\rangle = |\Psi'\rangle$. Subtraction of the corresponding Schrödinger equations from each other yields

$$(V_{\text{ext}} - V'_{\text{ext}})\,|\Psi\rangle = (E - E')\,|\Psi\rangle = c\,|\Psi\rangle \tag{1.14}$$

which is obviously valid only if $(V_{\text{ext}} - V'_{\text{ext}}) = c$. The injectivity of the mapping between the wave function $|\Psi\rangle$ and the electron density n can be proven in the following way.
Suppose there are two different external potentials V^1_{ext} and V^2_{ext}, differing by more than a constant, leading to the same ground state density $n(\mathbf{r})$. The two external potentials lead to two different Hamiltonians H^1 and H^2, which have two different ground state wave functions $|\Psi^1\rangle$ and $|\Psi^2\rangle$ leading to the same ground state density. Since $|\Psi^2\rangle$ is not the ground state wave function of H^1 it follows

$$E^1 = \left\langle \Psi^1 \right| H^1 \left| \Psi^1 \right\rangle < \left\langle \Psi^2 \right| H^1 \left| \Psi^2 \right\rangle. \tag{1.15}$$

Reformulation of the inequality yields

$$\begin{aligned} \left\langle \Psi^1 \right| H^1 \left| \Psi^1 \right\rangle &< \left\langle \Psi^2 \right| H^2 \left| \Psi^2 \right\rangle + \left\langle \Psi^2 \right| H^1 - H^2 \left| \Psi^2 \right\rangle \\ E^1 &< E^2 + \int d^3r \left[V^1_{\text{ext}}(\mathbf{r}) - V^2_{\text{ext}}(\mathbf{r}) \right] n(\mathbf{r}). \end{aligned} \tag{1.16}$$

Making the same with index 1 and 2 interchanged leads to

$$E^2 < E^1 + \int d^3r \left[V^2_{\text{ext}}(\mathbf{r}) - V^1_{\text{ext}}(\mathbf{r}) \right] n(\mathbf{r}). \tag{1.17}$$

Adding both equations together we get

$$E^1 + E^2 < E^1 + E^2 \tag{1.18}$$

which is obviously a contradiction. Hence we have proven, that there can not be two different external potentials which give the same non-degenerated ground state charge density. Thus all properties of the system under consideration are completely determined, when the ground state density $n_0(\mathbf{r})$ is known.

Second theorem: *A universal functional for the energy E can be defined in terms of the density $n(\mathbf{r})$ valid for any external potential $V_{\text{ext}}(\mathbf{r})$. The exact ground state energy of the system is the global minimum of this functional and the density $n(\mathbf{r})$, which minimizes this functional is the exact ground state density $n_0(\mathbf{r})$.*

Proof

Using the first theorem, we known that all properties of a system are uniquely determined by the density $n(\mathbf{r})$, and therefore we can express them as functionals of $n(\mathbf{r})$. The total energy functional then reads

$$\begin{aligned} E_{\text{HK}}[n] &= T[n] + E_{\text{ee}}[n] + \int d^3r V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \\ &= F_{\text{HK}}[n] + \int d^3r V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \end{aligned} \tag{1.19}$$

where the universal functional $F_{\text{HK}}[n]$ includes all contributions of the interacting system to the Hamiltonian, such as kinetic energy and the interaction energy between the electrons.
Let us determine the Hamiltonian by fixing the ground state density to $n^1(\mathbf{r})$, which leads to an external potential $V^1_{\text{ext}}(\mathbf{r})$ and to the ground state wave functions Ψ^1.

$$E^1 = E_{\text{HK}}[n^1] = \left\langle \Psi^1 \right| H^1 \left| \Psi^1 \right\rangle \tag{1.20}$$

When we consider a different density $n^2(\mathbf{r})$, corresponding to a different wave function $\Psi^2(\mathbf{r})$, it follows immediately that the energy of this state E^2 is greater than E^1.

$$E^1 = \left\langle \Psi^1 \right| H^1 \left| \Psi^1 \right\rangle < \left\langle \Psi^2 \right| H^1 \left| \Psi^2 \right\rangle = E^2 \tag{1.21}$$

Evaluating Eq. (1.19) at the ground state density $n_0(\mathbf{r})$ gives always energies lower than for any other density $n(\mathbf{r})$. If we would know the functional $F_{HK}[n]$, we could find the ground state energy of the system, by minimizing the total energy functional with respect to variations in the electron density.

1.3 Replacing one problem with another – the Kohn-Sham approach

With the Hohenberg-Kohn theorems the high dimensionality of the many-body wave function, and the problems inherent to it (storage, calculation time), can be replaced by the electron density $n(\mathbf{r})$, which only depends on three spatial coordinates. The drawback of this switch to the electron density as fundamental quantity is the introduction of a universal functional F_{HK} in the total energy functional E_{HK}, which exact form is not known explicitly.

The ingenious idea of W. Kohn and L. J. Sham [2] was a mapping between the desired interacting system and an auxiliary non-interacting reference system. It is assumed that the ground state density of the original interacting system is equal to that of the non-interacting system. This leads to independent particle equations for the non-interacting system, similar to the solutions of the Hartree equations, which are in principle exactly solvable numerically.
Using independent particle orbitals $|\psi_i\rangle$ we can introduce the independent particle kinetic energy operator

$$T_s[n] = -\frac{1}{2}\sum_i^n \langle\psi_i|\,\nabla^2\,|\psi_i\rangle \tag{1.22}$$

and the Coulomb interaction energy of the electron density $n(\mathbf{r})$ interacting with itself

$$E_{\mathrm{H}}[n] = \frac{1}{2}\int d\mathbf{r}d\mathbf{r}'\;\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{1.23}$$

and reformulate the Hohenberg-Kohn expression for the ground state energy functional

$$\begin{aligned} E_{\mathrm{KS}}[n] &= F_{\mathrm{HK}}[n] + \int d\mathbf{r}\; V_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r}) \\ &= T_s[n] + E_{\mathrm{H}}[n] + \int d\mathbf{r}\; V_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r}) + T[n] - T_s[n] + E_{\mathrm{ee}}[n] - E_{\mathrm{H}}[n] \\ &= T_s[n] + E_{\mathrm{H}}[n] + \int d\mathbf{r}\; V_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r}) + E_{\mathrm{xc}}[n]. \end{aligned} \tag{1.24}$$

Within the Kohn-Sham approach the universal functional $F_{\mathrm{HK}}[n]$, which explicit form is unknown within the interacting particle picture, is replaced by its non-interacting counterpart. All many-body effects of correlation and exchange, which are not present in the non interacting picture, are lumped together in the exchange and correlation energy functional $E_{\mathrm{xc}}[n]$.

1.3.1 The Kohn-Sham equations

The solution of the Kohn-Sham energy functional can be found by minimization of the total energy functional E_{KS} with respect to the electron density $n(\mathbf{r})$. Since the non interacting kinetic energy operator $T_s[n]$ is formulated in terms of the non interacting orbitals, a minimization with respect to the Kohn-Sham orbitals is more practicable:

$$\frac{\delta E_{\mathrm{KS}}}{\delta \psi_i^*(\mathbf{r})} = \frac{\delta T_s[n]}{\delta \psi_i^*(\mathbf{r})} + \left[\frac{\delta E_{\mathrm{H}}[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{\mathrm{ext}}[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})}\right] \frac{\delta n(\mathbf{r})}{\delta \psi_i^*(\mathbf{r})} = 0. \quad (1.25)$$

Including the usual orthonormalization constraint via the Lagrange multiplier method yields the Schrödinger like Kohn-Sham equations

$$\begin{aligned} &\left[-\frac{1}{2}\nabla^2 + \frac{\delta E_{\mathrm{H}}[n]}{\delta n(\mathbf{r})} + V_{\mathrm{ext}} + \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})}\right] \psi_i = \varepsilon_i \psi_i \\ &\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r})\right] \psi_i = \varepsilon_i \psi_i. \end{aligned} \quad (1.26)$$

Since no approximations have been applied to the exchange-correlation potential $V_{\mathrm{xc}}(\mathbf{r})$ so far, the ground state orbitals of the Kohn-Sham system give the exact ground state density and energy of the many-body interacting system.
Finding accurate approximations to the unknown exchange and correlation potential is an ongoing field of research (see Sec. 1.4).

1.4 Approximations for the XC-functional

As stated in Sec. 1.3.1, KS-DFT is in principle an exact reformulation of the initial many-electron problem, if and only if the exact form of the exchange and correlation energy functional is known. Since the exact form of the energy functional is unknown, finding accurate and "transferable" approximations to the exchange-correlation potential is an ongoing field of research. Today there are many different flavours of functionals available which are more or less suitable for different systems.
Within this section we will briefly outline approximations used in this thesis. For further reading we refer to the cited papers.

1.4.1 Local density approximation

The first approximation introduced by Kohn and Sham for the exchange-correlation energy functional was the local density approximation (LDA) [2], which is based on the homogeneous electron gas (HEG).
Within the LDA approximation the exchange-correlation energy functional is decomposed into the exchange and correlation energy

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int d\mathbf{r}\; e_{\mathrm{xc}}^{\mathrm{HEG}}[n(\mathbf{r})] = \int d\mathbf{r}\; n(\mathbf{r}) \epsilon_{\mathrm{xc}}^{\mathrm{HEG}}[n(\mathbf{r})] \tag{1.27}$$

$$= E_{\mathrm{x}}^{\mathrm{LDA}}[n] + E_{\mathrm{c}}^{\mathrm{LDA}}[n], \tag{1.28}$$

where $\epsilon_{\mathrm{xc}}^{\mathrm{HEG}}$ is the energy density per electron of the homogeneous electron gas. Both terms in Eq. (1.28) can be treated independently and it is assumed, that the exchange and correlation energy, depending on a non homogeneous electron density, can be approximated locally by the exchange and correlation energy functionals of the homogeneous electron gas of this density.
The exchange energy of the HEG can be calculated analytically, yielding

$$E_{\mathrm{x}}^{\mathrm{LDA}}[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int d\mathbf{r}\; n^{4/3} \tag{1.29}$$

whereas for the correlation energy functional $E_{\mathrm{c}}^{\mathrm{LDA}}[n]$ analytic expressions are only known for the low and high density limit. For intermediate values accurate Quantum Monte Carlo simulations exist [3]. Different analytic interpolation functions were proposed in the literature for the correlation energy, which reproduce the low and high density limit. Usually the analytic form provides enough flexibility to reproduce the intermediate Monte Carlo data accurately [4, 5, 6]. Therefore all of them yield very similar results and are often generically referred to as LDA functionals.
Well known shortcomings of the LDA functionals are the chronic overbinding, which leads to an underestimation of equilibrium lattice constants by $\approx 1\%$. But the most troublesome point is that binding energies of molecules and solids are overestimated in the order of ≈ 1 eV/atom.

1.4.2 Generalized gradient approximations

To overcome the shortcomings of the local density approximation, a natural way is the inclusion of the density gradient in the calculation of the exchange-correlation energy. These functionals are referred to as the generalized gradient approximation (GGA) functionals and exhibit in general the form

$$E_{xc}^{\mathrm{GGA}}[n(\mathbf{r})] = \int n(\mathbf{r}) \left\{ \epsilon_x^{\mathrm{HEG}}[n(\mathbf{r})]\, F_x[n,s] + \epsilon_c^{\mathrm{HEG}}[n(\mathbf{r})]\, F_c[n,t] \right\} d\mathbf{r}, \tag{1.30}$$

where $\epsilon_x^{\mathrm{HEG}}$ and $\epsilon_c^{\mathrm{HEG}}$ are the exchange and correlation energy densities per particle for the uniform electron gas (as in the LDA approximation), F_x and F_c are enhancement factors and s and t are dimensionless and dimensionless spin scaled density gradient respectively.

Today the most widely used GGA-functional is the parametrization of Perdew-Burke-Ernzerhof [7] referred to as PBE. The exchange enhancement factor is a functional of the dimensionless density gradient and two non empirical parameters

$$F_x[n,s] = 1 + \frac{\mu s^2}{1+\mu s^2/\kappa} \tag{1.31}$$

where $\kappa = 0.804$ and $\mu = 0.21951$. The correlation enhancement factor can be written as

$$F_c[n,t] = \left[1 + \frac{H(r_s,\zeta,t,\beta)}{\epsilon_c^{\mathrm{HEG}}[n(\mathbf{r})]}\right] \tag{1.32}$$

with t being a dimensionless spin scaled density gradient, r_s being the local Seitz radius, ζ the relative spin polarization $(n_\uparrow - n_\downarrow)/n$ and $\beta = 0.066725$ being an empirical parameter. The general form of the functional H is complex and therefore we refer to [7] for further reading.
In general, the parametrization according to Perdew-Burke-Ernzerhof constitutes an improvement over LDA, but it also has its problems. Although PBE reduces the overbinding of LDA when applied to solids, it still tends to overbind small molecules by 20-30 kJ/mol. Furthermore surface energies are even lower in PBE than in LDA.

One of the foremost problems of the commonly used density functionals is that none of them predict an equilibrium volume close to experiment. To resolve this issue, Wu and Cohen (WC) [8] suggested to attenuate the exchange enhancement factor in the PBE functional. This approach allows to obtain equilibrium volumes that are much closer to experiment than either LDA or PBE.
A similar route was followed by Perdew *et al.* [9]. They pointed out that a GGA with accurate atomic exchange and correlation energies will violate the gradient expansion for slowly-varying densities, which is valid for solids and their surface. Therefore the construction of a GGA with improved atomization and total energies will worsen bond lengths in solids, and vice versa. Perdew *et al.* proposed a functional specially aimed at an improved description of the geometric degrees of freedom for bulk systems and surfaces, via concise modification of the exchange and correlation enhancement factors, while accepting worsening of atomization energies [10]. With the modified functional (PBEsol) excellent agreement with experiment was found for bulk lattice constants for a wide class of materials.
These approaches were partly inspired by the work of A. Mattsson [11], who used a subsystem approach for simultaneously treating solids and surfaces. The final functional also takes on the form of a generalized gradient functional with similar enhancement factors as the PBEsol functional and yields almost exactly the same lattice constants as the PBEsol functional [11, 12].

1.4.3 Combination with exact exchange - hybrid functionals

All local and gradient corrected functionals severely underestimate the band gap. Although it is often argued that density functional theory is in principle not suitable to predict band gaps, the underestimation leads to a significant overestimation of the static screening properties (a profound ground state property), and a one to one correspondence between the band gap error and the static screening properties is observed [13].
Hybrid functionals such as B3LYP, PBEh, and HSE06 address this point by partly including the exact non-local exchange of Hartree-Fock theory [14, 15, 16, 17]. These functionals yield generally much improved band gaps and dielectric properties for extended systems [17, 18, 19], and simultaneously the PBEh and HSE06 functional improve the equilibrium volume [20].
Only the B3LYP functional increases the equilibrium volume, a result of the incorrect behaviour of the LYP correlation energy for homogeneous systems, and it is therefore not well suited for the description of extended systems [21].

The HSE03 and HSE06 functionals are based on the "parameter free" PBEh hybrid functional (sometimes also termed PBE0 or PBE1PBE functional) [15, 22], which itself is constructed by mixing 25% non-local Fock exchange E_x with 75% PBE exchange E_x^{PBE} and using the full electron correlation part of the PBE functional [7]:

$$E_{xc}^{\mathrm{PBEh}} = \frac{1}{4}E_x + \frac{3}{4}E_x^{\mathrm{PBE}} + E_c^{\mathrm{PBE}}. \tag{1.33}$$

A drawback of this functional is the slow convergence of the exact non-local exchange interaction with distance, necessitating dense k-point grids for metals. To overcome this problem, Heyd *et al.* [23] proposed a decomposition of the exact non-local exchange into a long and a short range part in real space:

$$E_{xc}^{\mathrm{HSE}} = \frac{1}{4}E_x^{\mathrm{sr},\mu} + \frac{3}{4}E_x^{\mathrm{PBE,sr},\mu} + E_x^{\mathrm{PBE,lr},\mu} + E_c^{\mathrm{PBE}}. \tag{1.34}$$

The range separation parameter μ controls the characteristic distance at which the long range non-local interaction becomes negligible. Based on a slightly erroneous implementation, Heyd originally proposed a value of $\mu = 0.15$ a.u.$^{-1} \approx 0.3$ Å^{-1} for the range separation parameter, whereas recently Krukau *et al.* [17] suggested to use $\mu = 0.11$ a.u.$^{-1} = 0.207$ Å^{-1}. The implementation using the latter range separation parameter is usually termed HSE06, whereas for the first one the abbreviation HSE03 is common. We use a variant of the HSE06 functional that observes the homogeneous electron gas limit and all important sum rules. Contrary to the conventional HSE06 functional, the screening parameter is set to $\mu = 0.300$ Å^{-1} in both the semi-local as well as non-local part of the exchange functional, whereas

the recommended choice is $\mu = 0.207$ Å^{-1}. As has been shown previously, the specific choice of the screening parameter μ has very little influence on the total energies, but slightly affects band gaps [17, 20]. From now on we will refer to this functional as HSE. A detailed discussion of the implementation of the HSE functional in VASP is given in Ref. [20].

1.5 Plane waves as basis set

Due to the periodicity of bulk materials and the independent electron approximation, inherent to the Kohn-Sham equations, the Bloch theorem can be applied to the KS orbitals. Each orbital can be written as the product of a Bloch function times a Bloch phase factor $e^{i\mathbf{kr}}$

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}} u_{n\mathbf{k}}(\mathbf{r}), \tag{1.35}$$

where the Bloch function has the same periodicity as the Bravais lattice of the solid under consideration. An expansion of $u_{n\mathbf{k}}$ in plane waves is an obvious choice, since only vectors of the reciprocal lattice $\mathbf{G}$ contribute to the expansion and plane waves provide a convenient basis for the approximation of the orbitals, at least in the interstitial regions of a bulk material:

$$u_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{Gr}}. \tag{1.36}$$

In general electron orbitals exhibit rapid oscillations due to the orthonormality condition to the other states close to the ionic core. An expansion in plane waves would require the inclusion of a high number of reciprocal lattice points $\mathbf{G}$, which will increase the computational cost of a calculation drastically. Several approaches to circumvent this problem were proposed in the past, with the projector augmented wave (PAW) method [24, 25] being one of the most accurate and efficient ones.

1.6 Projector augmented wave method

Prior to the discussion of the PAW formalism, we will introduce the main concepts of pseudopotentials (PP).

1.6.1 Pseudopotentials

The key idea of all pseudopotentials is the separation of the electrons of an atom in electrons attributed to the ionic core and the valence electrons. The core electrons are tightly bound and more or less inactive, whereas the valence electrons play the dominant part in the bonding between different

atoms. Therefore the core electrons can be attributed to the core leading to a "screened" and less attractive core potential.
With the removal of the core electrons, the high number of nodes in the valence electron wave function caused by the orthogonality to the core electrons, can be reduced by introducing "pseudo" valence wave functions, which have to be orthogonal only among each other.
For the construction of a pseudopotential one usually solves for all electronic radial orbitals $|\phi_i\rangle$ of the free atom within the Kohn-Sham approach.
In a second step the all electron orbitals of the valence electrons are replaced within a specified core radius r_c by smooth and nodeless pseudo wave functions $|\tilde{\phi}_i\rangle$ and a screened pseudopotential is introduced, consisting of the bare attractive Coulomb long range part and the repulsive short range potential attributed to the core electrons surrounding the ionic core. Different methods for the construction of pseudopotentials were proposed in literature, e.g. Bachelet, Hamann Schlüter PP (BHS) [26], Vanderbilt PP [27] and Troullier, Martins (TM) PP [28]. Within the pseudopotentials supplied with the VASP code the Rappe, Rabe, Kaxiras, Joannopoulos (RRKJ) PP [29] is most often used.
The introduced valence pseudo wave functions should fulfil the following properties to ensure a reasonable behaviour of the pseudopotential in environments different to the bare atom (this is known under the term transferability) [30].

- The KS eigenvalues of the pseudo and all electron wave functions have to be the same.
- The shape of the pseudized wave function and the all electron wave function has to coincide exactly outside the core radius r_c.
- The logarithmic derivatives of the real all electron and pseudo wave function, evaluated at the core radius r_c, as function of the KS-eigenvalue of the orbital under consideration matches the all electron counterpart as closely as possible. It can be shown that this can be reached by constructing the pseudopotential in such a manner that it has the same norm as the all electron orbital.

Especially the last property is essential for a high degree of transferability of a pseudopotential.

1.6.2 The PAW methodology

It is evident, that a pseudopotential approach with nodeless pseudo wave functions is the first choice within a plane wave basis set, since it keeps the number of expansion coefficients and therefore the computational cost low in comparison to an all electron calculation. Obviously the downside is loss

of information, since the nodal structure of the valence electrons near the core is replaced by the pseudized orbital.
The strength of the PAW approach is that it combines the advantages of plane waves with an all electron approach. It retains the nodal structure of the valence electrons, but at the same time allows the treatment of the valence electrons using plane waves.
For a detailed derivation of the PAW formalism we refer to the original paper by Blöchl [24] and the detailed derivation following [25] given in Appendix A. We will only sketch the main concepts here.
In a conventional DFT implementation using pseudopotentials the solutions are pseudo orbitals $|\tilde{\Psi}_n\rangle$, usually expanded in plane waves.
In the PAW approach, the space Ω is divided into two regions. The augmentation region Ω_A is a sphere centred around each of the atomic sites with the radius r_c, whereas the remaining part of the system is assigned to the interstitial region Ω_I.
The PAW spheres are allowed to overlap, but the contributions from these overlaps are not calculated, since the PAW method only includes and allows for one centre terms.
The ultimate goal of this method is to map the all electron orbitals on computationally convenient pseudo orbitals. To this end a transformation T is introduced between the all electron picture and the pseudized description of the system under consideration

$$|\Psi_n\rangle = T|\tilde{\Psi}_n\rangle. \tag{1.37}$$

Since, in the construction of the pseudopotential, it is demanded that the pseudo and all electron orbital coincide outside the augmentation spheres, the transformation has no effect on the interstitial region, and only local contributions in the augmentation spheres are present. The transformation can be written as

$$T = 1 + \sum_R T_R \tag{1.38}$$

where the contributions T_R are only acting at the atomic site R within the PAW-sphere Ω_A.
A set of pseudo $|\tilde{\phi}_i\rangle$ and all electron orbitals $|\phi_i\rangle$ at different energies is usually determined in a similar manner as when a pseudopotential in constructed and via the invertible transformation T one can calculate one out of the other

$$|\phi_i\rangle = (1 + T_R)|\tilde{\phi}_i\rangle \quad \text{at atomic position } R. \tag{1.39}$$

Lets assume that the valence orbitals $|\tilde{\Psi}_n\rangle$ of our system, represented by plane waves, can be expanded using the pseudo partial waves $|\tilde{\phi}_i\rangle$ in the augmentation region Ω_A

$$|\tilde{\Psi}_n\rangle = \sum_i c_{i,n}|\tilde{\phi}_i\rangle \quad \text{in} \quad \Omega_A. \tag{1.40}$$

For continuity reasons the coefficients $c_{i,n}$ are the same in the expansion of the all electron orbital in AE partial waves (see Eq. (A.9), $|\phi_i\rangle$ and $|\tilde{\phi}_i\rangle$ match smoothly onto each other at the sphere boundary),

$$|\Psi_n\rangle = \sum_i c_{i,n} |\phi_i\rangle \quad \text{in} \quad \Omega_A. \tag{1.41}$$

The pseudo and all electron orbitals are related by the simple transformation

$$|\Psi_n\rangle = |\tilde{\Psi}_n\rangle - |\tilde{\Psi}_n\rangle_{\Omega_A} + |\Psi_n\rangle_{\Omega_A} \tag{1.42}$$

$$= |\tilde{\Psi}_n\rangle - \underbrace{\sum_i c_{i,n}|\tilde{\phi}_i\rangle + \sum_i c_{i,n} |\phi_i\rangle}_{\text{contributions only in } \Omega_A} \tag{1.43}$$

$$= |\tilde{\Psi}_n\rangle + T_R|\tilde{\Psi}_n\rangle. \tag{1.44}$$

By comparison with the structure of the desired transformation operator T_R we obtain

$$c_{i,n} = \langle \tilde{p}_i|\tilde{\Psi}_n\rangle, \tag{1.45}$$

where the projector function $|\tilde{p}_i\rangle$ is introduced, which determines the amount or strength of the pseudo partial wave in a given pseudo orbital. As shown in Appendix A, the following relations have to hold for the projector functions

$$\langle \tilde{p}_i|\tilde{\phi}_j\rangle = \delta_{ij} \tag{1.46}$$

$$\sum_i |\tilde{\phi}_i\rangle\langle \tilde{p}_i| = 1. \tag{1.47}$$

The projectors have to be dual to the partial waves they were constructed for

$$|\Psi_n\rangle = |\tilde{\Psi}_n\rangle - \sum_i \langle \tilde{p}_i|\tilde{\Psi}_n\rangle|\tilde{\phi}_i\rangle + \sum_i \langle \tilde{p}_i|\tilde{\Psi}_n\rangle|\phi_i\rangle \tag{1.48}$$

$$= |\tilde{\Psi}_n\rangle - |\tilde{\Psi}^1_n\rangle + |\Psi^1_n\rangle. \tag{1.49}$$

In the last line we adopted the usual notation within the PAW formalism, where a tilde sign $\tilde{}$ indicates a pseudo property and the ^{1} indicates a on-centre term, which is present within an augmentation sphere only.
In principle all local operators decompose in a similar way into the three contributions given in Eq. (1.49) (see App. A.1.1), like the density operator (see App. A.1.2) or the total energy (see App. A.1.6)

$$E = \tilde{E} - \tilde{E}^1 + E^1 \tag{1.50}$$

$$n = \tilde{n} - \tilde{n}^1 + n^1. \tag{1.51}$$

Minimizing the total energy expression with respect to the pseudo density operator (for a detailed derivation see App. A.1.8) one finally obtains the Schrödinger like DFT equation within the PAW formalism

$$\left[T + \tilde{V}_{\text{eff}} + \sum_{i,j} |\tilde{p}_i\rangle \left(\hat{D}_{ij} + D^1_{ij} - \tilde{D}^1_{ij}\right) \langle \tilde{p}_j|\right] |\tilde{\Psi}_n\rangle$$
$$= \left[\mathbf{1} + \sum_{i,j} |\tilde{p}_i\rangle Q_{ij} \langle \tilde{p}_j|\right] |\tilde{\Psi}_n\rangle. \quad (1.52)$$

Part II

$SrTiO_3$ and $BaTiO_3$ revisited using the projector augmented wave method: Performance of hybrid and semilocal functionals

Chapter 2

Introduction

Perovskite materials exhibit many interesting properties like colossal magnetoresistance, ferroelectricity, superconductivity, charge ordering, spin dependent transport and high thermopower. These compounds are used as sensors and catalyst electrodes in certain types of fuel cells and are candidates for memory devices and spintronics applications.
In general materials are called perovskites when their crystal structure is of the same type as calcium titanium oxide ($CaTiO_3$) which was first discovered in the Ural mountains of Russia, and named after the Russian mineralogist, L. A. Perovski (1792-1856).
The general chemical composition of a perovskite is of ABX_3 type, where the A and B are the cations and X is an anion which bonds to A and B. The high temperature structure (shown in Fig. 4.1) is of cubic symmetry with the A cations (green) at the 12-fold coordinated cube corners, the B cations (blue) in the 6-fold coordinated centre of the cell (bcc like), surrounded by an octahedron of X anions (red) which occupy the face centres of the cube. By lowering the temperature perovskites tend to undergo a variety of phase transitions, which made them attractive materials for fundamental studies. In the present work, we decided to focus our investigations on two "classical" model perovskite crystals, strontium titanate ($SrTiO_3$, STO) and barium titanate ($BaTiO_3$, BTO).

2.1 $SrTiO_3$

STO is probably the experimentally best investigated ABO_3 perovskite type oxide due to its extraordinary low temperature behaviour. Above the critical temperature (T_c) of about $105.5 \pm 0.5\,\mathrm{K}$ [31, 32] it has the usual cubic perovskite structure (O_h). Decreasing the temperature below T_c, STO undergoes an antiferrodistortive (AFD) structural phase transition to a tetragonal phase lowering the symmetry from O_h to D_{4h} induced by the condensation

of a soft phonon mode at the R-point [33]. This phase transition involves a rotation of the TiO_6 octahedron around one of the unit cell axes and consequently, due to the decreasing distance between the rotated oxygen and the strontium in the cube corners, an increase in the c lattice constant. This rotation takes place in opposite directions in adjacent unit cells.
In the last ten years renewed interest in the low temperature phase of $SrTiO_3$ has emerged because of experiments on the static dielectric constant by Müller and Burkard [34] and Rupprecht and Bell [35]. They showed that the dielectric constant in the cubic phase closely follows the Curie-Weiss law of the form $\epsilon \approx (T - T_{\text{ferro}})^{-1}$ with a critical temperature of 35.5 K and 37 K respectively.
In addition Viana *et al.* [36] showed that, applied to the low temperature tetragonal phase, the Curie-Weiss law yields a critical temperature of about 20 K. However, the predicted transition to the ferroelectric (FE) phase at T_{ferro}, which is caused by the condensation of a Γ-point phonon forming Ti-O dipoles, was never observed in experiment. Instead of the predicted divergence according to the Curie-Weiss law, the static dielectric constant saturates at an enormous value of approximately $2.4 \cdot 10^4$ when the temperature approaches zero [34].
Therefore $SrTiO_3$ is called an *incipient ferroelectric* material. Additional evidence that $SrTiO_3$ is close to a ferroelectric state is provided by the experiments done by Uwe and Sakudo [37], which induced the ferroelectric state by applying uniaxial stress, and Itoh *et al.* [38], who found a transition temperature of 23 K by isotopical substitution of ^{16}O with ^{18}O.
It was first suggested by Müller and Burkard [34] that zero-point quantum fluctuations of the titanium atoms suppress the condensation of the FE-phonon mode, leading to a so called "quantum paraelectric state". This assumption was corroborated later by Jauch and Palmer [39] using γ-ray diffraction and theoretically anticipated by Zhong and Vanderbilt [40]. From experiment, one can thus conclude that, neglecting quantum effects, STO should show a ferroelectric instability at the experimental volume. However, literature gives ambiguous results for the theoretically predicted frequency of the FE phonon mode ranging from soft to hard ($94i$ - 64 cm^{-1}), strongly depending on the chosen volume of the unit cell and the applied functional. We will address this point by carefully evaluating the volume dependence of the phonon frequencies using different functionals. The LDA, PBE, PBEsol and HSE06 functional will be covered in the present study.

2.2 $BaTiO_3$

$BaTiO_3$, on the other hand, is a prototypical ferroelectric perovskite, which has the typical perovskite cubic structure at high temperatures. At 393 K it undergoes a structural phase transition to a tetragonal ferroelectric phase. Between 278 K and 183 K it has an orthorhombic structure, and below 183 K it is rhombohedral [41]. In this work we will focus our investigations only on the paraelectric cubic and ferroelectric tetragonal phase.

Within our investigations we examine the structural properties of the cubic and antiferrodistortive phases of STO and the ferroelectric tetragonal phase of BTO and compare our results to experimental as well as recently published theoretical results. The electronic structures are shortly summarized, followed by the discussion of the phonon modes of STO and BTO. In particular, we focus on the variation of the TO phonon modes upon varying cell volumes and exchange-correlation functionals.

Chapter 3

Computational details

3.1 Calculation details

We used the *Vienna ab initio simulation package* [42, 43] for the present calculations. The PAW atomic reference configurations are listed in Tab. 3.1. Since perovskite properties are strongly dependent on the precise electronic structure, an accurate sampling of the Brillouin zone is of outermost importance. Additionally the accuracy of the calculations is also determined by the cutoff energy for the expansion of the wave functions in plane waves and by the discretization of the charge density and total potential in real space. Therefore, we performed for all important quantities extensive convergence tests using different Monkhorst-Pack [44] k-point meshes, different cutoff energies for all functionals and also different accuracy of the real space integration (controlled via the `PREC` flag in).
It turned out that for calculations using the simple cubic perovskite unit cell a $6 \times 6 \times 6$ Monkhorst-Pack k-mesh for the Brillouin zone integration, 500 eV cutoff energy for the expansion of the wave functions in plane waves and setting the precision flag `PREC` to `normal` yield converged results for the lattice parameters and the bulk moduli.

Due to the large volume dependence of the phonon frequencies of $SrTiO_3$

	valence e^-	r_s [Å]	r_p [Å]	r_d [Å]
Ti	$3s^2 3p^6 3d^2 4s^2$	0.953	1.217	1.217
O	$2s^2 2p^4$	0.794	1.429	
Sr	$4s^2 4p^6 5s^2$	1.323	1.323	
Ba	$5s^2 5p^6 6s^2$	1.482	1.429	

Table 3.1: PAW atomic valence configurations and core radii for $SrTiO_3$ and $BaTiO_3$.

(see Chap. 5), the determination of the theoretical equilibrium volume of the high temperature cubic phase is of outermost importance. We performed a damped molecular dynamics relaxation of the internal parameters at fixed volume and retained symmetry for all four exchange correlation functionals and fitted the energy volume data using

$$E(V) = E_0 + \frac{B_0 V}{B_0'}\left(\frac{(V_0/V)^{B_0'}}{B_0' - 1} + 1\right) - \frac{B_0 V_0}{B_0' - 1}. \tag{3.1}$$

Eq. (3.1) is the well known Murnaghan equation of state from which we extracted the equilibrium lattice constant a_0 and the bulk modulus B_0.

Due to the doubling of the unit cell in the tetragonal phase of STO as described later, a reduced mesh of $4 \times 4 \times 4$ points is sufficiently dense to yield reliable results for this phase. As BTO does not undergo an AFD distortion, a $6 \times 6 \times 6$ k-mesh was used for all considered phases.

For phonon calculations, a denser mesh is necessary: $8 \times 8 \times 8$ k-points are required for the cubic phase of STO and BTO. Due to the required high accuracy of the results, the electronic and ionic convergence criteria had to be chosen much more stringent than for standard bulk calculations. The geometries were relaxed until the maximum remaining force was smaller than 0.005 eV/Å.

3.2 Applied functionals

In the last 20 years, STO and BTO and perovskites in general have been thoroughly theoretically investigated mainly by using *ab initio* methods, in particular density functional theory (DFT) methods [1, 2]. In most cases, the local density approximation (LDA) has been used for the exchange and correlation part of the density functional, and to a lesser extent the generalized gradient approximation (GGA) and, in very few cases, hybrid functionals have been applied [45, 46]. Both, the LDA and PBE approximations, are subject to well known limitations: LDA tends to "overbind" solids and molecules, which leads to too small equilibrium volumes. Although the PBE functional constitutes an improvement over the LDA for most material properties, it tends to "overcorrect" the unit cell volume. Because of the strong volume dependence of the (anti) ferroelectric instabilities, both approaches cannot be expected to give a quantitative description of ABO_3 perovskite properties.
For instance, we will show in this work that in $SrTiO_3$, the ferroelectric (FE) instability depends very sensitively on the applied functional *and* the volume. At their respective theoretical equilibrium volume, LDA and GGA

give qualitatively different results (GGA is unstable against ferroelectric distortions, LDA is stable), and the situation is not improved by evaluating the properties at the experimental volume (GGA is now stable, LDA unstable against ferroelectric distortion). A quantitative description of the FE instability is therefore clearly out of reach, and worse, present theoretical methods even fail to predict concisely whether $SrTiO_3$ is at all unstable against a ferroelectric distortion.

Obviously, the foremost problem is that none of the commonly used density functionals predict an equilibrium volume close to experiment. Therefore Perdew *et al.* recently proposed the PBEsol functional (see Sec. 1.4.2). With this modified form of the PBE functional excellent agreement with experiment was found for bulk lattice constants for a wide class of materials. One can expect that this functional is more suitable for describing the instability of ferroelectric materials, in particular, the dilemma whether to use the experimental or theoretical volume is reduced.

However, a second issue remains unaddressed by this modified GGA functional. As all other GGA functionals it severely underestimates the band gap. Since lattice instabilities are strongly linked to the electronic states at the valence band maximum and conduction band minimum, and because dielectric properties are of fundamental importance for ferroelectric materials, an improved description of the band gap and screening properties might be equally important as an improved description of the equilibrium volume.

Hybrid functional try to correct the underestimation of the band gap by partially including the exact non-local Hartree-Fock exchange. We used the HSE06 hybrid functional throughout this study as described in Sec. 1.4.3.

Chapter 4

Structural and electronical properties of $SrTiO_3$ and $BaTiO_3$

4.1 Cubic phase

The high temperature perovskite structure is of cubic symmetry (shown in Fig. 4.1), with the (Sr, Ba) cations (green) at the 12-fold coordinated cube corners, the Ti cations (blue) in the 6-fold coordinated centre of the cell (bcc like), surrounded by an octahedron of O anions (red) which occupy the face centres of the cube.

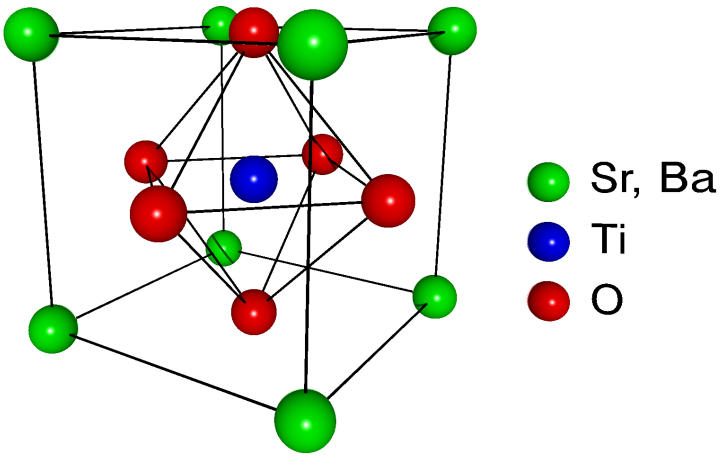


Figure 4.1: Cubic perovskite structure.

	$SrTiO_3$							
	LDA	PBE	PBEsol	HSE	B3LYP	B3PW	B1-WC	Exp.
a_0 [Å]								
present	3.863	3.943	3.898	3.904				3.890 [47]
Ref. [48]	3.86	3.93			3.94			3.900 [32]
Ref. [49]	3.86	3.94			3.94			
Ref. [50]	3.878	3.94						
Ref. [45]							3.880	
Ref. [46]						3.912		
B_0 [GPa]								
present	198	168	185	192				179 [47]
Ref. [48]	215	195			187			
Ref. [49]	214	169			177			
Ref. [50]	191	169						

Table 4.1: Comparison of calculated and experimental lattice constants a_0 [Å] and bulk moduli B_0 [GPa] of STO in the cubic phase for several GGA and hybrid functionals.

In Tab. 4.1 and Tab. 4.2 we present the lattice constants and bulk moduli of STO and BTO respectively for the four used exchange-correlation functionals together with the most recent theoretical and experimental data.

Using the standard LDA and PBE functionals we find values almost in perfect agreement with previous calculations performed by Heifets *et al.* [48] and Piskunov *et al.* [49], with the CRYSTAL and by Mete *et al.* [50] with the ABINIT program package.

As expected LDA underestimates the lattice constants by approximately 1 % whereas the PBE functional overestimates it by the same amount. The GGA functional PBEsol performs significantly better in determining the equilibrium lattice constant for both perovskites. The deviation from the experimentally observed values is only ≈ 0.05 %. The performance of the hybrid functional HSE is roughly similar, although a trend to slightly larger lattice constants and harder bulk moduli is clearly visible.

Other hybrid functionals, which were already applied on these two materials are the B3LYP, B3PW and B1-WC functional. The B3LYP functional yields almost or exceeds the PBE equilibrium lattice constant, which is a consequence of the incorrect behaviour of the LYP correlation energy for homogeneous systems [21].

In the B3PW functional, 20 % of exact non-local Fock exchange is mixed with Becke's GGA functional [52] for the exchange, whereas for the correlation functional the parametrization by Perdew-Wang [53, 54] is used. Due to the fact that B3PW posses a similar correlation as HSE, it also predicts

	$BaTiO_3$							
	LDA	PBE	PBEsol	WC	HSE	B3LYP	B1-WC	Exp.
a_0 [Å]								
present	3.953	4.035	3.986		3.995			4.00 [47]
Ref. [49]	3.96	4.03				4.04		
Ref. [45]	3.958	4.035		3.990		4.036	3.971	
B_0 [GPa]								
present	189	162	177		181			162 [47]
Ref. [49]	196	175				176		
Ref. [51]	195	160						

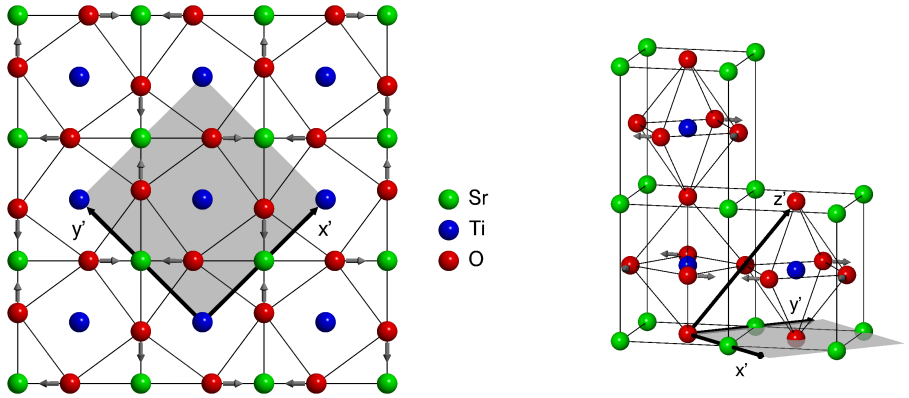
Table 4.2: Comparison of calculated and experimental lattice constants a_0 [Å] and bulk moduli B_0 [GPa] of BTO in the cubic phase for several GGA and hybrid functionals.

similar lattice constants as HSE. The recently proposed B1-WC hybrid functional [45] combines the gradient corrected Wu-Cohen functional [8], which behaves generally very similar to PBEsol, with 16 % non-local exact exchange. Since the WC exchange functional shrinks the lattice constants compared to PBE, and inclusion of exact exchange reduces the lattice constants even further, this functional predicts lattice constants that are smaller than for PBEsol and HSE.

4.2 Tetragonal phases

4.2.1 Antiferrodistortive phase (AFD) of STO

In the 1950s and 1960s a number of experiments, using different techniques such as microwave loss [35], neutron spectroscopy [55] or elastic constant [56] measurements, gave evidence for a structural phase transition in STO at a critical temperature of $105.5 \pm 0.5\,\mathrm{K}$ [32, 31].
In 1964 Cowley [55] investigated the phonon dispersion curves via neutron spectroscopy close to the Γ-point. He used his experimental results to obtain parameters for different theoretical models. These models suggested an instability of the STO crystal against two phonon modes at the M and R-point.
In 1967 Unoki and Sakudo [57] resolved the question of the low temperature phase by their measurements using the electron spin resonance of Fe^{3+} in doped STO single crystals. They showed that below the critical temperature the oxygen octahedron rotates around one of the unit cell axis (see Fig. 4.2(a)). The angle of rotation gradually increases when the temperature is lowered. However they were not able to determine the periodicity along



(a) Slice through the STO crystal along the x-y plane through the cube centred titanium atom, visualizing the rotation scheme of the oxygen octahedra.

(b) Three adjacent unit cells of the tetragonal phase of STO, showing the opposite rotation direction and the redefined unit cell.

Figure 4.2: Structural distortions of the STO crystal in its low temperature tetragonal phase.

the c tetragonal axis, but they assumed an alternating direction of rotation along the z-direction with a periodicity of $2c$.

In 1969 Shirane and Yamada [33] confirmed the assumption of Unoki and Sakudo that adjacent unit cells in x, y and z direction have an opposite direction of rotation by inelastic neutron scattering (see Fig. 4.2(b)). Additionally they showed conclusively that the condensation of the Γ_{25}-phonon mode at the R point causes the antiferrodistortive phase transition (AFD).

Müller and Burkard [34] and Viana *et al.* [36] measured the dielectric constant for STO in both, the cubic and tetragonal, phases. The critical temperature according to the Curie-Weiss law was determined to 35.5 K and 20 K respectively. However Müller and Burkard [34] observed a saturation of the dielectric constant below ≈ 10 K instead of the predicted divergence according to the high temperature dielectric constant measurements, and therefore the predicted transition to a ferroelectric state is not observed in experiment. It can be easily induced however for example by isotopical substitution of ^{16}O with ^{18}O [38] or applying uniaxial stress to the crystal [37]. Therefore STO is called "*incipient ferroelectric*". It was first suggested by Müller and Burkard [34] that quantum fluctuations suppress the formation of the ferroelectric phase. Twenty years later Jauch and Palmer [39] measured the zero point motion of the Ti centre atom of STO at ≈ 50 K with γ-ray diffraction and confirmed the assumption of the so called "quantum paraelectric state". The reason for the influence of the quantum zero point motion of the atoms on the suppression of the ferroelectric state in STO is

	LDA	PBE	PBEsol	HSE	B3PW	Experiment
a [Å]						
present	3.847	3.933	3.886	3.900		3.898 [58] (78 K)
Ref. [46]					3.910	3.898 [32] (65 K)
c/a [Å]						
present	1.008	1.004	1.006	1.001		1.00056 [32] (65 K)
Ref. [46]					1.0006	1.0009 [59] (10 K)
θ_z [°]						
present	6.05	4.74	5.31	2.63		1.4 [57] (77 K)
Ref. [46]					1.95	2.1 [57] (4.2 K)
ΔE [eV]						
present	0.019	0.007	0.011	0.002		
Ref. [46]					0.0002	

Table 4.3: Structural properties of the antiferrodistortive phase of $SrTiO_3$ below the critical temperature T_c. θ_z is the rotational angle of the octahedra around the z-axis, and $\Delta E = E_{\text{cub}} - E_{\text{AFD}}$ describes the energy gain for the antiferrodistortive phase.

not an unusually large absolute magnitude of the vibrational amplitude of the titanium atom, but the smaller displacement of it in the ferroelectric state and therefore the formation of static Ti-O dipoles and ferroelectricity is prevented.

To be able to describe the AFD structure of $SrTiO_3$ in terms of the smallest possible unit cell, one has to double the cubic unit cell to $\sqrt{2} \times \sqrt{2}$ in the x-y plane, due to the clockwise and neighbouring anti-clockwise rotation of the TiO_6 octahedra around the z-axis (see Fig. 4.2(b)). Furthermore, in addition to the in-plane rotation of the TiO_6 octahedra, the AFD structure involves a rotation of octahedra against each other in adjacent layers of STO in z-direction and therefore the z-axis has to be tilted such, that it points, in the next layer, to a cell with the same sense of rotation for the TiO_6 octahedron.
In the case of $SrTiO_3$, the relaxation into the AFD equilibrium structure was performed by rotating the oxygen ions out of the cubic equilibrium position and performing a damped molecular dynamics simulation in which the lattice shape and size was allowed to change.

Lixin Cao *et al.* [32] measured the temperature dependence of the lattice parameters of bare and epitaxially stressed STO between 65 K and 125 K

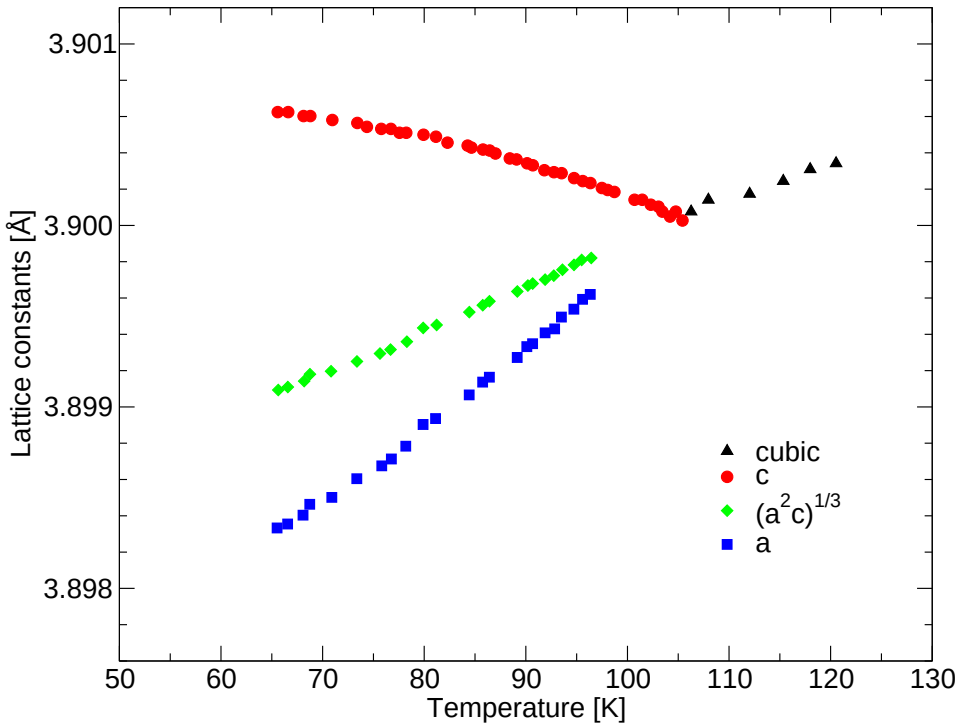


Figure 4.3: Fig. 4(a) from Lixin Cao *et al.* [32] showing the tetragonal lattice parameters a and c dependent on temperature and measured via X-ray backscattering.

using X-ray backscattering. From the observed temperature dependence of the lattice parameters a and c, as shown in Fig. 4.3 one can estimate, that at $T = 65\,\mathrm{K}$, c is almost identical to its zero temperature value (3.9006 Å), whereas some further decrease of a is to be expected (3.8982 Å), thereby increasing c/a as $T = 0\,\mathrm{K}$ is approached. In addition to the c/a ratio, the rotation angle of the oxygen octahedron is a characteristic criterion to estimate the quality of the used functional.

As expected, one obtains a similar behaviour for the equilibrium lattice constants of the AFD phase as for the cubic phase, *i.e.* LDA and PBE underestimate and overestimate the volume, respectively (see Tab. 4.3). Furthermore, LDA and PBE both overestimate the rotation angle and the c/a ratio significantly, LDA overestimating it more than PBE. PBEsol yields properties in between the two extremes, LDA and GGA. For the volume this is desirable, and in fact the equilibrium volume is very well described by PBEsol. However, the c/a ratio and the rotation angle are much too large as it is the case for LDA and PBE. The hybrid functional HSE performs exceptionally well,

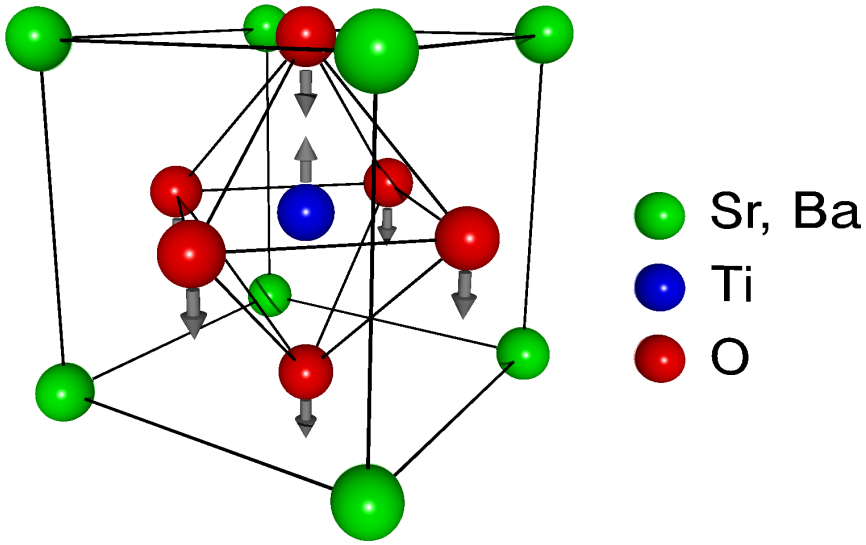


Figure 4.4: Schematic ball model of the ferroelectric distortions in $BaTiO_3$.

much better than any of the gradient corrected functionals. In particular, the c/a ratio and the rotation angle are very close to experiment, whereas the volume is slightly overestimated compared to the measurements of Heidemann [59] at 10 K. Previous studies utilizing atomic basis sets within the CRYSTAL code (Ref. [46, 48, 49]), give results almost identical to our study for both semilocal (LDA, PBE) and hybrid functionals (B3PW). One can thus conclude that inclusion of Fock exchange weakens the tendency of STO to undergo an antiferrodistortive phase transition, whereas modifications of the density functional part alone have little influence on the magnitude of the displacement in the antiferrodistortive phase transition.

4.2.2 Ferroelectric phase (FE) of BTO

For the relaxation of BTO into the ferroelectric tetragonal structure a similar approach was adopted. For $BaTiO_3$, the FE phase transition involves a distortion of Ti and O in z-direction only (see Fig. 4.4) and hence a supercell is not required to model the distortion. A $6 \times 6 \times 6$ k-mesh was used for all considered phases.

In the tetragonal phase of BTO, the c/a ratio and the distortions of the titanium and oxygen ions are the relevant properties for the description of the tetragonal ferroelectric phase. Our results as well as experimental reference data are presented in Tab. 4.4.

Although LDA underestimates, as expected, the lattice parameters and the

	LDA	PBE	PBEsol	WC	HSE	B3LYP	B1-WC	Exp.
a [Å]								
present	3.945	4.005	3.971		3.969			3.997[a]
Ref. [45]	3.954	4.013		3.982		3.996	3.962	3.986[b]
c [Å]								
present	3.978	4.210	4.054		4.135			4.031[a]
								4.026[b]
c/a [Å]								
present	1.009	1.050	1.021		1.042			1.0086[a]
Ref. [45]	1.006	1.035		1.012		1.066	1.015	1.0100[b]
V [Å^3]								
present	61.91	67.23	63.94		65.13			64.41[a]
Ref. [45]	62.2	66.9		63.9		68.0	63.2	63.97[b]
dz_{Ti}								
present	0.011	0.018	0.015		0.019			0.020[a]
Ref. [45]	0.011	0.018		0.013		0.019	0.015	0.015[b]
$dz_{\mathrm{O}_\parallel}$								
present	-0.017	-0.047	-0.028		-0.042			-0.026[a]
Ref. [45]	-0.014	-0.039		-0.022		-0.057	-0.024	-0.023[b]
$dz_{\mathrm{O}_\perp}$								
present	-0.011	-0.027	-0.017		-0.022			-0.012[a]
Ref. [45]	-0.009	-0.022		-0.013		-0.031	-0.014	-0.014[b]
P [C/m^2]								
present	0.216	0.435	0.314		0.407			0.27[c]
Ref. [45]	0.20	0.39		0.26		0.48	0.28	

[a] Ref. [41] measured at 280 K.
[b] Ref. [60]
[c] Experimental data from Ref. [61]

Table 4.4: Structural properties and macroscopic polarization P of the ferroelectric tetragonal phase of $BaTiO_3$, where a and c are the fully relaxed lattice parameters, c/a is the tetragonal distortion, V the unit cell volume, and $dz_{\mathrm{O}_\perp}$ and $dz_{\mathrm{O}_\parallel}$ stands for the atomic displacements orthogonal or parallel to the Ti-O bond in fractions of the c lattice constant.

unit cell volume, it gives very good results for the c/a ratio and reasonable, but to small results for the displacements of the titanium and oxygen ions. The PBE functional strongly overestimates the c/a ratio and the oxygen and titanium displacements. PBEsol again yields values that are in between LDA and PBE and therefore agreement with experiment is overall very good. We found that the magnitude of the ferroelectric displacement depends strongly on the volume; it is fairly small for the LDA volume, and almost

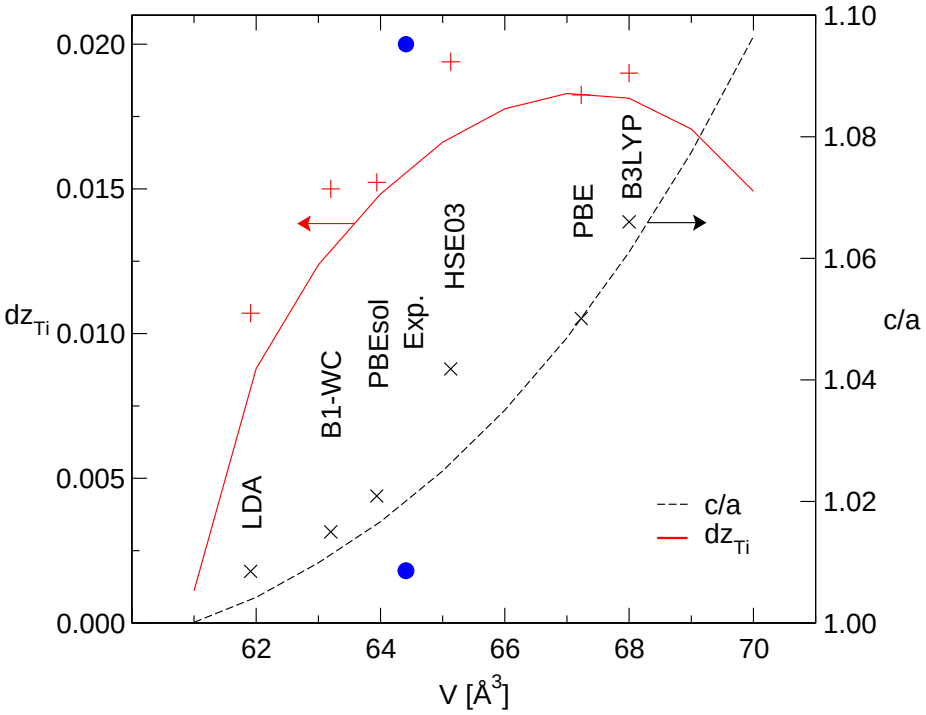


Figure 4.5: Volume dependence of the c/a ratio (dashed line) and of the relative displacement of Ti ion (dz_{Ti}, full red line) in the tetragonal phase of $BaTiO_3$ using the PBE functional. Symbols represent results using various functionals at their theoretical equilibrium volume. The experimental values are marked with blue filled circles.

doubles for the PBE volume. The functional itself has little direct influence on the displacement pattern, e.g. if the c/a ratio and displacement pattern is predicted using the PBE *at the LDA volume*, results are very close to the LDA values.

In Fig. 4.5 we show the displacement of the Ti atom and the c/a ratio for various volumes evaluated using the PBE functional and using various functionals at their respective equilibrium volume. Our data calculated using different functionals and the data for B1-WC and B3LYP available from literature [45, 49] fit very well on this curve, if we use the volume as control parameter. In this specific case, the predicted volume is therefore the single most important parameter determining the magnitude of the ferroelectric distortion.

This behaviour agrees well with the fact that BTO is a B-site driven ABO_3 perovskite, where the titanium ion is too small for its site due to the enlarged cell volume caused by the barium ions at the cube corners.

	LDA	PBE	PBEsol	HSE	B3LYP	B3PW	B1-WC	Exp.
				cubic				
$\Gamma - \Gamma$								
present	2.15	2.18	2.18	3.47				3.75[a]
Ref. [49]	2.36	2.35			3.96	3.89		
Ref. [45]							3.91	
$R - \Gamma$								
present	1.81	1.80	1.82	3.07				3.25[a]
Ref. [49]	2.04	1.99			3.63	3.57		
Ref. [45]							3.57	
				tetragonal				
$\Gamma - \Gamma$								
present	2.21	2.12	2.21	3.48				
Ref. [46]						4.03		
$R - \Gamma$								
present	1.97	1.79	1.93	3.11				
Ref. [46]						3.72		

[a] Ref. [64]

Table 4.5: Comparison of calculated and experimental direct $\Gamma - \Gamma$ (and indirect $R - \Gamma$ band gap) in eV for cubic and tetragonal distorted $SrTiO_3$.

In this light, it is not astonishing that the GGA functionals PBEsol and WC as well as the hybrid functional B1-WC give overall the best description of the distortion because they yield the best volumes, followed by the HSE functional as can be seen in Tab. 4.4. The macroscopic polarization P, which was calculated using the Berry phase approach [62, 63], can also be seen as a quality criterion for the performance of the different functionals and it is a result of the ferroelectric distortion.

Due to the underestimation of the ionic displacements using the LDA functional the polarization is too small in comparison to the experimentally found value, whereas for PBE the overestimation of the displacement of the titanium and oxygen ions leads to a too large value. As stated above the PBEsol, WC and B1-WC functionals describe the ferroelectric phase best and therefore yield values closest to the experimental data. The HSE hybrid functional overestimates the displacement of the oxygen atoms by a factor of two and yields a too large polarization.

	LDA	PBE	PBEsol	HSE	B3LYP	WC	B1-WC	Exp.
				cubic				
$\Gamma - \Gamma$								
present	1.84	1.87	1.91	3.07				
Ref. [45]	2.2	2.2			3.8	2.2	3.45	
$R - \Gamma$								
present	1.73	1.70	1.77	2.92				3.2[a]
Ref. [45]	2.1	2.1			3.7	2.1	3.39	
				tetragonal				
$\Gamma - \Gamma$								
present	2.09	2.31	2.16	3.55				
Ref. [45]							3.73	
$R - \Gamma$								
present	1.84	1.76	1.76	3.02				3.4[a]
Ref. [45]	2.10				3.80		3.44	

[a] Ref. [65]

Table 4.6: Comparison of calculated and experimental direct $\Gamma - \Gamma$ (and indirect $R - \Gamma$ band gap) in eV for cubic and tetragonal distorted $BaTiO_3$.

4.3 Electronic properties

For the converged lattice constants of STO and BTO, we have calculated the density of states by using the tetrahedron method for the Brillouin zone sampling. Tab. 4.5 and Tab. 4.6 show the direct ($\Gamma - \Gamma$) and indirect ($R - \Gamma$) band gaps for both, the cubic and tetragonal distorted structure, and the results from former calculations and experiments.
We observe for both perovskites the well known behaviour for LDA, PBE and PBEsol: LDA and PBE quite significantly underestimate the band gap, and PBEsol essentially gives results equivalent to LDA and PBE. Due to the fact that the WC-functional is a GGA functional it suffers from the same drawbacks concerning electronic properties as all other GGAs. As expected, the hybrid functionals performs best in comparison to experiment. The B1-WC functional overestimates the band gaps approximately by the same amount as the HSE functional underestimates it. Only in the case of the ferroelectrically distorted BTO the B1-WC functional almost exactly reproduces the experimentally observed band gap, whereas the HSE result is 0.4 eV off.

Chapter 5

Phonon properties

With five atoms in the unit cell and 12 degrees of freedom (disregarding the translational mode), the perfect cubic perovskites exhibit four distinct phonon modes: three triply degenerate phonon modes with the irreducible representation Γ_{15}, and one triply degenerate phonon mode with the irreducible representation Γ_{25}. In the present case, we have not calculated the longitudinal optical-transverse optical (LO-TO) splitting, and we will therefore restrict our discussion to TO modes.

In Tab. 5.1 and Tab. 5.2 we present our results and the theoretical and experimental data from other publications. Concerning the phonon frequencies, we see from experiment [55, 69] that an optical absorption feature is observed for $SrTiO_3$ at room temperature around 90 cm^{-1}, and that this absorption feature drops in frequency when the temperature is lowered. However, this mode Γ_{15} (TO1) is certainly not related to a harmonic phonon mode: as already discussed in the introduction, $SrTiO_3$ shows a divergence of the dielectric constant at low temperatures, which indicates that the Γ_{15} (TO1) mode should be soft in a *harmonic* description.

It is immediately obvious that the theoretically predicted phonon frequencies show a huge scatter for $SrTiO_3$, in particular, for the low frequency soft Γ_{15} mode. Sai [67] and Zhong [68] report phonon frequencies for the FE phonon mode Γ_{15} (TO1) that differ by over 80 cm^{-1}, although their applied lattice constants and exchange correlation functionals are identical. Moreover, Sai [67] found the mode to be hard, whereas Zhong [68] predicted it to be soft, indicating a ferroelectric instability. Additionally Tinte [51] and Sai [67] also used the experimental lattice constant for their calculations, which induces an external, fictitious negative pressure on the unit cell.

Such a situation is certainly discomforting and it reflects the difficulties in the construction of suitable pseudopotentials for Ti, and the rather slow convergence of the low frequency branch with respect to the number of k-points. The present calculations were performed using a full potential method, and treating the entire Ti $3s$, $3p$ and $3d$ shell as valence in order to

work	[Å]		Γ_{15} (TO1)	Γ_{15} (TO2)	Γ_{25}	Γ_{15} (TO3)
present	3.863	LDA	80	177	226	563
present	3.943	PBE	115i	147	234	512
present	3.898	PBEsol	29i	160	229	536
present	3.904	HSE	74i	162	250	533
Ref. [66]	3.904	LDA	87i	149	223	519
Ref. [67]	3.865	LDA	42	168		549
	3.897	LDA	94i	151		521
Ref. [68]	3.865	LDA	41i	165		546
Ref. [51]	3.894	LDA	57i	157		532
	3.894	GGA	64	166		551
Ref. [55]	296 K	Exp	91.1			
Ref. [55]	90 K	Exp	42.3	170	265	
Ref. [33]	120 K	Exp	52.5			
Ref. [69]	300 K	Exp	87.7	178		544
Ref. [70]	297 K	Exp	91.7	169±3	265±5	547±3
Ref. [71]	295 K	Exp	88±1	175±2	266±3	545±1

Table 5.1: Zone centre phonon frequencies for the cubic structure of $SrTiO_3$. In the second column we give the used lattice constant (in Å) for the calculations and the temperature (in K) for the experimental data (Tinte [51] and Ghosez [66] used the experimental lattice constant in their calculations, Sai [67] used it also for comparison to the theoretical lattice constant). Phonon frequencies are given in units of [cm^{-1}].

avoid any errors related to the frozen core approximation applied in our PAW calculations. Furthermore, it is nowadays certainly possible to fully converge the results with respect to the k-point set for local and gradient corrected functionals, however, for the hybrid functionals computational costs increase as the square of the number of k-points, and therefore certain compromises were required. To remain concise, the same set of k-points was used for all calculations ($6 \times 6 \times 6$), and hence the error bar for the low frequency imaginary branch might approach 10 cm^{-1}.

Before discussing the behaviour of the low frequency mode in more detail, we first want to elaborate briefly on the other modes. It is obvious that LDA performs quite well in predicting the vibrational frequencies. The Γ_{15} (TO2) and the Γ_{15} (TO3) modes match almost exactly the experimental data for both STO, as well as BTO, despite — or rather because — the calculations were done at the LDA volume, which is significantly too small. The PBE

work	[Å]		Γ_{15} (TO1)	Γ_{15} (TO2)	Γ_{25}	Γ_{15} (TO3)
present	3.953	LDA	$139i$	186	291	479
present	4.035	PBE	$239i$	169	286	453
present	3.986	PBEsol	$188i$	178	287	465
present	3.995	HSE	$241i$	185	310	480
Ref. [72]	3.94	LDA	$113i$	184	288	481
Ref. [51]	4.03	PBE	$203i$	168		463
Ref. [45]	3.990	WC	$128i$	186	282	469
Ref. [45]	3.971	B1-WC	$145i$	195	299	482
Ref. [73]	395 K	Exp	soft	182	308	482

Table 5.2: Zone centre phonon frequencies for the cubic structure of $BaTiO_3$. In the second column we give the used lattice constant (in Å) for the calculations and the temperature (in K) for the experimental data. Phonon frequencies are given in units of [cm^{-1}].

functional generally yields much too soft modes at the theoretical PBE volume. We note that most phonon modes are fairly independent of the applied functional, and the difference between LDA and GGA is mainly related to the different volume at which the calculations were performed [74]. As a result of the intermediate volume, the PBEsol functional yields values pretty much in between the LDA and PBE. If the results were entirely independent of the functional, one would expect similar values for the HSE functional. But this is not generally the case. On the contrary, inclusion of exact exchange stiffens the bonds somewhat and increases the phonon frequencies by typically 5 % compared to a local or gradient corrected functional. This behaviour is well known for molecules [75], and is also observed for extended bulk systems [74]. Here we also observe that the HSE functional yields consistently larger phonon frequencies than the PBEsol functional, although the equilibrium volume is slightly larger for HSE than for PBEsol. For BTO, the agreement is exceptionally good, whereas some discrepancies remain for the STO case. But even for STO the HSE functional yields overall the best agreement with experiment. We emphasize that the frequency of the fairly volume independent Γ_{25} mode is only accurately described by the HSE functional, whereas all non-hybrid functionals underestimate its frequency. A similar behaviour was observed by K. Hummer *et al.* [74] for different group-IV elemental semiconductors and insulators, where the optical phonon mode at the zone centre was very sensitive to the inclusion of exact exchange.

We now return to the low frequency "soft" mode. Even our own data show a large scatter among the four applied functionals, and to clarify whether

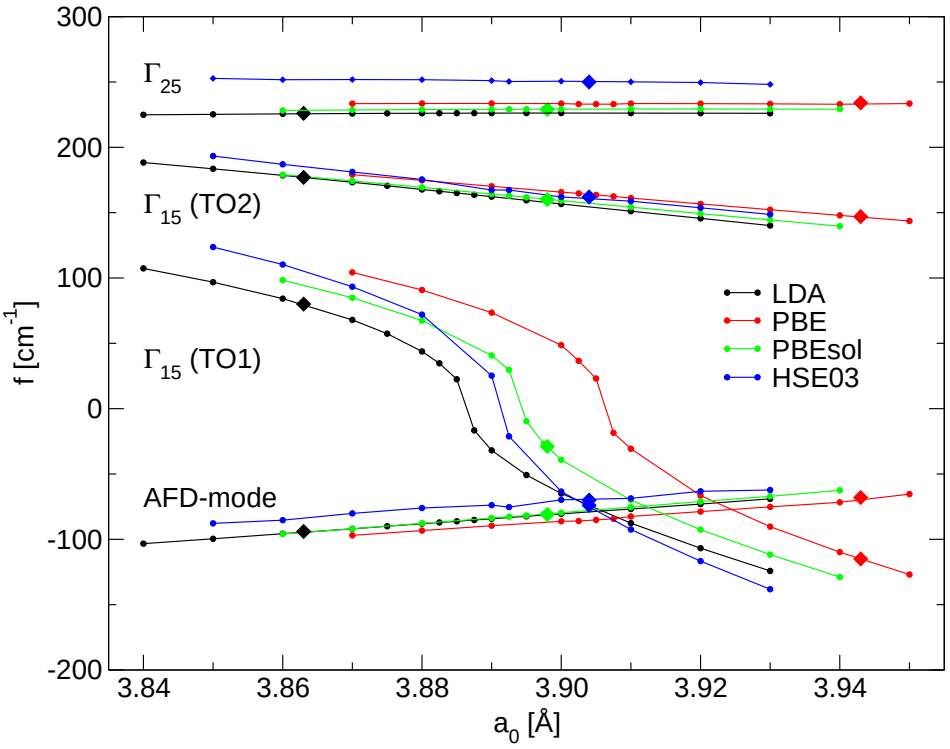


Figure 5.1: Volume dependence of the AFD and FE phonon mode of STO. Imaginary frequencies are represented by negative numbers. Phonon frequencies at the equilibrium volume are highlighted by filled diamonds.

these differences are related to different equilibrium volumes or to the applied exchange correlation functional, we investigated the volume dependence of the phonon modes of STO and BTO for all applied functionals. The AFD mode, corresponding to the zone boundary phonon at the R-point, was included in this investigation by doubling the unit cell, as already described in Sec. 4.

The results confirm our previous discussion. Most of the phonon modes are quite insensitive to the applied functional, although as already discussed the hybrid HSE functional shows a tendency to stiffen the modes, in particular the Γ_{25} mode. Even the imaginary frequency of the AFD mode only slightly depends on the functional, but we recall that the final displacement pattern, which also depends on quadratic terms, showed a strong dependence on the functional (see Tab. 4.3). The frequency of the ferroelectric mode, however, clearly changes from one functional to the other. This applies to both STO, as well as BTO. We see that the mode becomes soft upon increasing the volume. For STO the transition occurs very close to the experimental (and PBEsol and HSE) volume. A very simple model that is capable

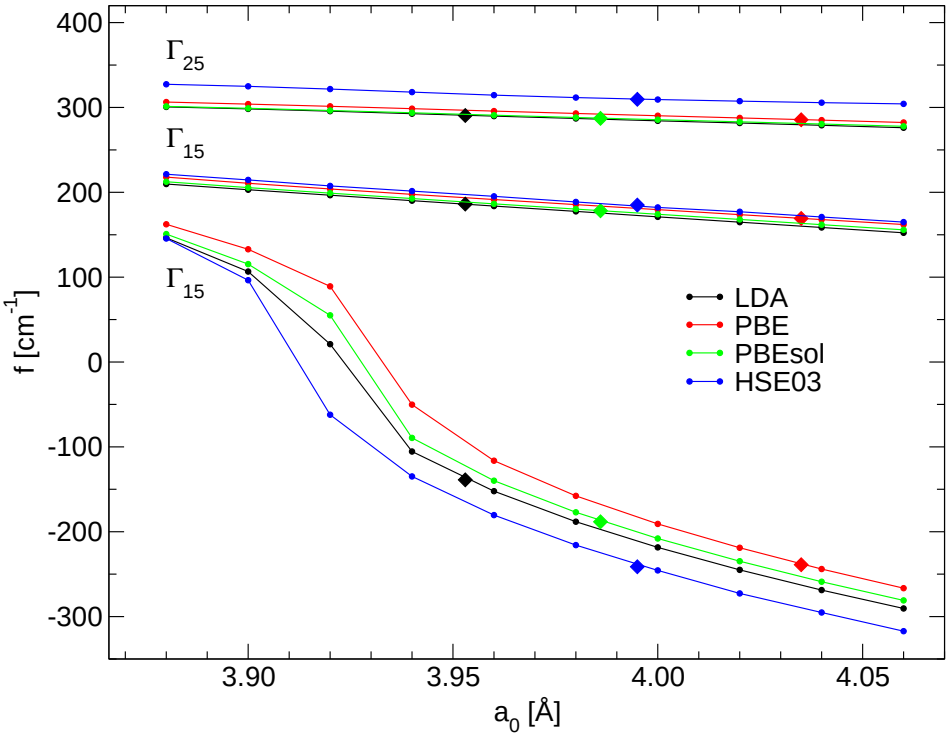


Figure 5.2: Volume dependence of FE phonon modes for cubic $BaTiO_3$. Imaginary frequencies are presented by negative numbers. Phonon frequencies at the equilibrium volume are denoted by filled diamonds.

to describe the observed behaviour is now discussed. Whether (S,B)TO is unstable against a ferroelectric instability seems to be determined by the space available to the Ti atom in the O-octahedron cage and by the preferred Ti-O distance. The space available to the Ti atom is proportional to half of the lattice constant. To determine the preferred Ti-O bond length, we optimized the TiO_2 rutile structure for each of the considered functionals and determined the respective theoretical ideal Ti-O bond length in this structure. The relevant values are given in the first line of Tab. 5.3.

Indeed, the table suggests that at the theoretical equilibrium volume, the FE mode of STO is stable in the LDA ($a_{\mathrm{STO}}/2 = 1.931 < 1.932$), strongly unstable for PBE ($a_{\mathrm{STO}}/2 = 1.972 > 1.962$), and slightly unstable for PBEsol ($a_{\mathrm{STO}}/2 = 1.949 > 1.945$) and HSE ($a_{\mathrm{STO}}/2 = 1.952 > 1.944$). This agrees with the predictions in Fig. 5.1.

Additionally we plotted the lattice constants, where the transition from a soft to hard phonon occurs, over the doubled lattice constant of the optimized rutile TiO_2 structure. We see in Fig. 5.3 an almost linear relation between the two lattice constants, suggesting that whether STO is unstable

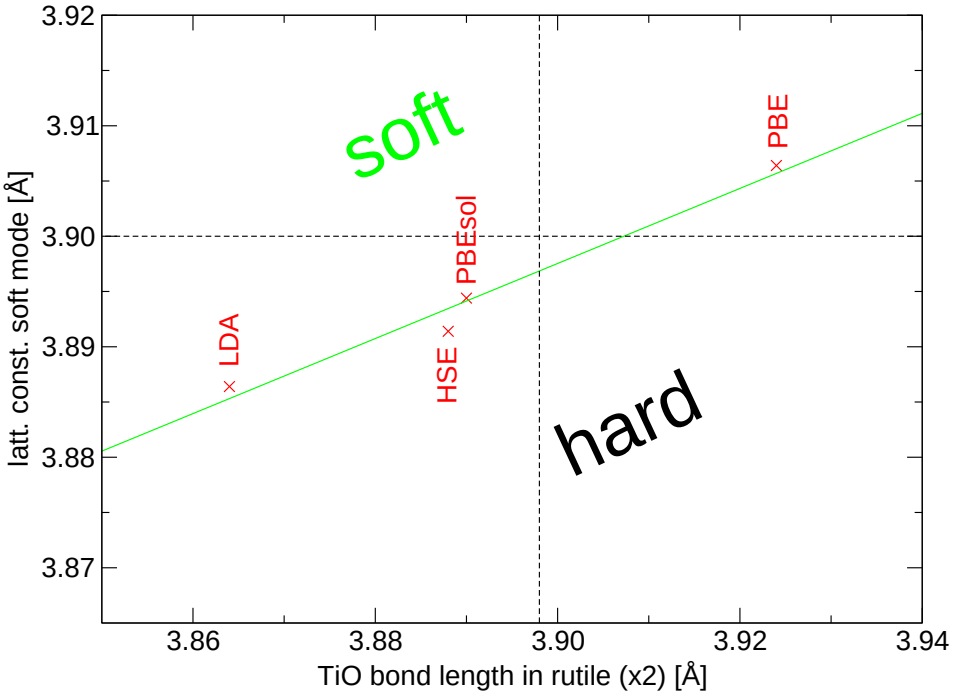


Figure 5.3: Comparison between the Ti-O bond length in rutile TiO_2 and the Ti-O bond length in STO at the transition point from a hard to a soft phonon mode.

against an FE instability seems to be determined by the space available for the Ti atom in the O octahedron and the preferred Ti-O bond length.

Comparison with experiment suggests that both HSE and PBEsol predict the $SrTiO_3$ volume and the preferred Ti-O bond length in TiO_2 reasonably well, the PBEsol functional more so than HSE. LDA more strongly underestimates the $SrTiO_3$ volume, confining the Ti-atom inside the oxygen octahedron and therefore predicting $SrTiO_3$ to be stable against a FE distortion of the lattice at the theoretical volume. On the other hand PBE predicts a much too large volume for $SrTiO_3$, expanding the oxygen cage and providing room for the Ti-atom to go off symmetry. Therefore PBE overestimates the instability at the theoretical volume.

For BTO the transition from hard to soft is observed far below the equilibrium volume, at lattice constants only slightly larger than those of STO, as can be seen from Fig. 5.2. The $a_{\mathrm{BTO}}/2$ distances at the transition are 1.962, 1.963, 1.964 and 1.966 Å for HSE, LDA, PBEsol and PBE, all well below the respective equilibrium Ti-O distances in BTO but only slightly larger than Ti-O in TiO_2. Thus BTO is without question, and for all applied functionals,

	LDA	PBE	PBEsol	HSE	Exp.
			Ti–O bond		
TiO_2	1.932 (−0.8)	1.962 (+0.7)	1.945 (−0.2)	1.944 (−0.3)	1.949
$SrTiO_3$	1.931 (−1.0)	1.972 (+1.1)	1.949 (−0.1)	1.952 (+0.1)	1.950
$BaTiO_3$	1.976 (−1.2)	2.018 (+0.9)	1.993 (−0.3)	1.997 (−0.1)	2.000
			Sr–O bond		
SrO	2.538 (−1.6)	2.602 (+0.9)	2.567 (−0.5)	2.583 (+0.1)	2.580
$SrTiO_3$	2.731 (−1.0)	2.788 (+1.1)	2.756 (−0.1)	2.760 (+0.1)	2.758
			Ba–O bond		
BaO	2.739 (−1.1)	2.807 (+1.4)	2.767 (−0.1)	2.787 (+0.6)	2.770
$BaTiO_3$	2.795 (−1.2)	2.853 (+0.9)	2.819 (−0.3)	2.825 (−0.1)	2.828

Table 5.3: Comparison of equilibrium M(=Ti,Sr,Ba)–O bond lengths (in Å). As reference for the stable metal-oxide bulk structures, the rutile (TiO_2) and the NaCl structure (SrO and BaO) were used. For STO and BTO the cubic structure was used. The numbers in parentheses give the deviations of the equilibrium distances from experiment (in percent).

a ferroelectric material, whereas STO is just on the fringe of a ferroelectric instability, making it difficult to place a firm number on the frequency of the Γ_{15} (TO1) mode. Fortunately, HSE and PBEsol, both functionals that predict reasonable equilibrium volumes and Ti-O bond lengths, show that $SrTiO_3$ is indeed unstable against a ferroelectric distortion, in agreement with the conclusion one must draw from the available experimental data.

Chapter 6

Summary and conclusion

In the present study, we have shown that the use of the novel PBEsol and HSE functionals gives a significant improvement of the description of the ABO_3 perovskites $SrTiO_3$ and $BaTiO_3$ in their cubic and tetragonal phases. Whereas the equilibrium geometries are sufficiently well captured by PBEsol (being a "classical" GGA potential with just the scaling coefficients adapted to improve the structural properties of solids and surfaces), it is inevitable to admix exact non-local Fock exchange to the standard density functional, as done in HSE, in order to obtain accurate electronic structure results. Only the latter approach gives satisfactory agreement of the calculated band gaps with experiment.

Furthermore, our results show that the ferroelectric instability strongly depends on the volume and to a lesser extent on the applied functional. Specifically, a transition from a hard ferroelectric mode at small volumes to a soft mode at larger volumes is predicted, with the transition occurring roughly at the equilibrium volume of $SrTiO_3$. This makes it so particularly hard to predict whether $SrTiO_3$ is unstable towards a ferroelectric distortion or not. Maybe the most crucial result of our study is that the lattice constant at which the transition is predicted to occur is roughly twice the theoretical equilibrium Ti-O bond length in TiO_2. This relation is clearly observed for $SrTiO_3$, but it also approximately applies to $BaTiO_3$. Since the Ti-O bond length in TiO_2 varies greatly between the functionals, quite different transition points are observed in $SrTiO_3$. Because LDA generally predicts too short bonds, and GGA generally predicts too large bonds, the best strategy seems to judge the stability of the mode by calculating its frequency at the theoretical equilibrium volume. Such a strategy should take care of the fact that the LDA and GGA "world" are too small and large, respectively, but unfortunately the strategy fails. The local density approximation more severely underestimates the Sr-O bond length than the Ti-O bond length (see Tab. 5.3). Therefore the $SrTiO_3$ volume, which is a compromise between Ti-O and Sr-O bonds, is much too small and the Ti atoms are locally

confined without the ability to undergo an off centre movement. The usual gradient corrected functional (PBE) shows the opposite trend, with a more severe overestimation of the Sr-O bond length than the Ti-O bond length. In this case, the ferroelectric instability is most likely overestimated at the theoretical $SrTiO_3$ equilibrium volume. The HSE functional predicts 1% smaller lattice constants than PBE, but the relative bond length errors are very similar to PBE. For this particular problem, the PBEsol functional seems to offer the best compromise with accurate predictions for all bonds (Sr-O, Ti-O and Ba-O).
The semi-local exchange functional suggested by Wu and Cohen (WC) [8] works very similar as PBEsol in most respects. This is related to the fact that the PBEsol functional modifies the semi-local exchange in very much the same manner as Wu and Cohen suggested for their WC functional. The hybrid B1-WC functional [45], which combines the WC functional with 16 % non local exchange, yields slightly too small lattice constants, since the inclusion of non-local exchange always shrinks the lattice constants compared to the corresponding semi-local functional (typical errors for $BaTiO_3$ and $SrTiO_3$ are -1 %, see Ref. [45]). On the other hand, the HSE functional tends to overestimate the volumes slightly, a behaviour that carries over from its parent PBE functional (errors for $BaTiO_3$ and $SrTiO_3$ are 1 %). Whether B1-WC or HSE are ultimately better suited for a specific problem might be largely system dependent, but we note that for $BaTiO_3$ and $PbTiO_3$, B1-WC predicts exceptionally good ferroelectric displacements and polarizations (at slightly too small equilibrium volumes).
Our important conclusion is that a design goal of functionals is the proper prediction of the *bond length balance* in very diverse materials. Although obvious in hindsight, this statement is hardly ever borne out as clearly as in the present study. It is noted that this observation applies to many other problems, for instance, strain and stress relief at interfaces.
Since the PBEsol and WC functionals work so exceptionally well, the question arises, whether there is any need for hybrid functionals. After all, they are at least one order of magnitude more expensive than semi-local functionals, and the PBEsol and WC functionals work exceedingly well for $BaTiO_3$ and $SrTiO_3$. So, why bother with a huge increase in computational costs? In fact, apart from the band gaps, which are arguably not ground state properties, there are a few other properties that are predicted clearly better by the hybrid functionals. For $SrTiO_3$ the ground state structure of the distorted antiferroelectric phase is in better agreement with experiment for the hybrid functional HSE, whereas all local functionals predict a somewhat too strong distortion. Second, the phonon modes are much more accurate using the hybrid functionals. In fact, the only other functional that allows for a reasonable prediction of the phonon frequencies is the LDA, if the frequencies are evaluated at the theoretically predicted equilibrium volume. LDA predicts so good frequencies because of a sizeable error compensation:

LDA frequencies at the experimental volume are too soft, but the too small volume and the strong increase of the phonons with volume compensates for this. GGA drastically underestimates the phonon frequencies as a result of a too large equilibrium volume, and as a result of a general underestimation of phonon frequencies using semi-local functionals. The hybrid functionals (HSE and B1-WC) work exceptionally well, and they are in fact the only functionals that allow to make accurate predictions for the Γ_{25} mode.
The verdict which functional to prefer for ferroelectric materials is certainly still open and requires further theoretical investigations. At this point it is only clear that theory is moving in the right direction. Both, the revised PBEsol functional and the WC functional as well as the hybrid HSE and B1-WC functionals are capable to make reliable predictions for $SrTiO_3$ and $BaTiO_3$, avoiding many problems inherent to the PBE and LDA functionals.

Part III

Stabilization mechanism for the polar ZnO(000$\bar{1}$)-O surface

Chapter 7

Introduction

The first use of zinc oxide (ZnO) is impossible to be traced back in time, since zinc oxide was already used by early humans. Until the 20^{th} century the main usage of ZnO was as ingredient in medical ointments and as pigment in paints, especially in zinc white. Today ZnO is usually processed in its white powder form and it is used as additive in many different materials like plastic, glass and rubber.

The crystal form of ZnO (zincite) is very rare on earth and almost all commercially and scientifically used crystals are produced synthetically. It is a semiconducting transparent metal oxide crystal with a direct band gap of ≈ 3.3 eV [76] and an interesting combination of semiconducting and optical properties. It is used in a wide range of applications like optoelectronics, light emitting devices and catalysis [77]. Although many chemicals are produced using heterogeneous catalysts based on zinc oxide or zinc oxide based compounds, the fundamental principles concerning the structure and reconstruction mechanisms of plain ZnO surfaces are still not explained adequately.

Bulk ZnO is built by stacking consecutively hexagonal layers of oxygen and zinc along the crystallographic c or (0001)-axis with alternating distances of $R_1 = 0.61$ Å and $R_2 = 1.99$ Å. When cleaving the ZnO crystal perpendicular to this axis, two of the most interesting and principally different surfaces are created. The oxygen terminated $(000\bar{1})$ surface, on which we will focus our investigations, and the zinc terminated (0001) surface, which was already investigated recently [78, 79]. Both of them are polar, belong to the so-called "Tasker-type(3)" surfaces [80] and are attributed to be unstable in their bulk terminated ideal form due to a diverging electrostatic surface energy. Polar surfaces can be stabilized by modifying either the charge or the stoichiometry in the surface layers with respect to the bulk properties. Different mechanisms are conceivable to compensate for the instability of the bulk terminated surface. Surface atoms can desorb leading to reconstructions of

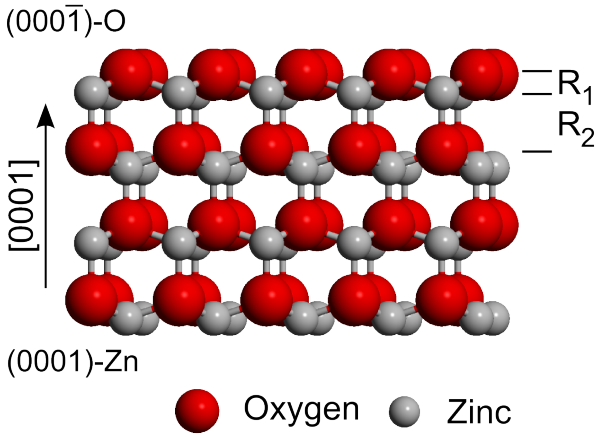


Figure 7.1: Ball model of the side view of the ZnO structure. ZnO adopts alternating hexagonal layers of four-fold coordinated Zn and O atoms with alternating distances of $R_1 = 0.61$ Å and $R_2 = 1.99$ Å.

the surface. Atoms from the residual atmosphere in the experimental setup can adsorb on the surface, providing the necessary charge compensation. Finally charge transfer within the material under consideration may happen as response to the polar electrostatic field [81].
The latter mechanism, also referred to as "metallization of the polar surfaces" was not observed in a recent scanning tunnelling spectroscopy (STS) experiment on ZnO [82]. Consequently we investigate in detail possible ad- and desorption mechanisms on the (000$\bar{1}$)-O terminated surface within the DFT-framework as implemented in the VASP program. The results of these calculations on the different surface terminations are then used for the construction of a phase diagram and the investigation of the thermodynamic stability of the surface.
In the subsequent sections of this introduction we will briefly explain our computational setup, as well as our way from the bulk unit cell of ZnO to the slab calculations as necessary within VASP. In Chapter 8 we will introduce the formalism used for the thermodynamical treatment of the DFT calculations and the calculation of the phase diagram, and finally the results of all different stabilization mechanisms are presented in Chapter 9.

7.1 Bulk ZnO

The first step towards an accurate simulation of the (000$\bar{1}$)-O terminated surface within the periodic boundary conditions implemented in VASP is the calculation of the theoretical lattice parameters of bulk ZnO. As already mentioned in the introduction, ZnO crystallizes in a wurtzite type structure.

	a	c	c/a	u
PW91	3.283 (+1.0)	5.303 (+1.8)	1.615 (+0.8)	0.3788 (-1.0)
PBE	3.287 (+1.1)	5.306 (+1.9)	1.614 (+0.8)	0.3790 (-0.9)
PBEsol	3.237 (−0.4)	5.228 (+0.4)	1.615 (+0.8)	0.3787 (-1.0)
HSE06	3.259 (+0.2)	5.239 (+0.6)	1.607 (+0.3)	0.3803 (-0.6)
Exp.	3.250	5.207	1.602	0.3825

Table 7.1: Lattice parameters a and c and the c/a ratio for bulk ZnO for different DFT exchange-correlation functionals are presented. Additionally the parameter $u = R_2/c$ (see Fig. 7.1) is shown, which is an internal coordinate of the wurtzite structure, determining the relative position of the oxygen and zinc sub-lattices along the (0001) axis. The numbers in parenthesis are the percental deviations from the corresponding experimental value.

Each of the two individual atom species form hexagonal sublattices, which are stacked on top of each other with alternating distances of $R_1 = 0.61$ Å and $R_2 = 1.99$ Å along the (0001) axis (see Fig. 7.1). In a perfect bulk environment, each oxygen and zinc atom has a fourfold (tetrahedral) coordination.
Similar to Part II we used different exchange-correlation (XC) functionals to calculate the equilibrium lattice parameters of ZnO. Beside the already used PBE [7], PBEsol [9] and HSE06 [23] hybrid functional, we also included the PW91 XC functional [83, 84], since it was already used in previous studies. The calculational details for the bulk calculation were as follows. For all applied XC functionals we used a dense $16 \times 16 \times 16$ Γ-centred k-point mesh. The cutoff energy for the plane wave basis set was chosen to be 600 eV.
Due to the hexagonal lattice with its additional internal degrees of freedom a fitting procedure to the Murnaghan equation of state as described in Sec. 3.1 is not suitable and therefore a quasi-Newton algorithm as implemented in VASP was used to relax the ions into their equilibrium ground state positions and to find the correct lattice parameters. For relaxation of the ions only the z-coordinate of the two oxygen ions was allowed to relax. In this case, we found that it is absolutely essential to use dense k-point grids and a high energy cutoff, since otherwise the energy minimization can give an incorrect c/a ratio and inaccurate lattice parameters.
When the behaviour of the calculated equilibrium lattice parameters for the used XC-functionals is compared to experiment, as given in Tab. 7.1, a similar trend as for the the structural properties of STO and BTO can be observed: PBEsol improves the lattice parameters a and c compared to the standard PW91 and PBE results, but it has no effect on the c/a ratio and the internal lattice parameter u. As for the antiferrodistortive phase of STO (see Tab. 4.3) only the HSE functional improves all important struc-

	valence e^-	r_s [Å]	r_p [Å]	r_d [Å]
Zn	$3d^{10}4s^2$	1.217		1.217
Zn	$4s^2$	1.217		
Zn	$4s^{1.5}$	1.217		
O	$2s^2 2p^4$	0.794	1.005	
H	$1s^1$	0.582		

Table 7.2: PAW atomic valence configurations and core radii used for the calculations.

tural parameters simultaneously. Although the HSE functional reproduces the structural properties of bulk ZnO more accurately than any of the other functionals, it is not suitable for this type of calculations, since large supercells, as used within this work, are computational too demanding.

The PW91 and PBE parametrizations of the XC-functional give almost identical results on the lattice parameters, although they have different analytic forms and are derived in different ways. Indeed it is commonly believed in the DFT community that both PW91 and PBE are interchangeable functionals and even our former results corroborate this assumption. However, Mattson *et al.* [85] showed recently that for surface calculations the differences between the PW91 and PBE functionals can give small but noticeable differences in calculated properties like vacancy formation energies.
The PBEsol functional would have been the best choice within this set of functionals, since it is designed for accurate surface energies and an improved description of the lattice parameters of solids and surfaces. As already mentioned before the use of the PBEsol functional does not improve the values of the c/a ratio and the internal parameter u in comparison to the PW91 and PBE values.
Ultimately we decided to use the PW91 XC-functional throughout our calculations to maintain compatibility to the recent studies on the (0001)-Zn terminated surface side [78, 79]. The PAW atomic reference configurations are listed in Tab. 7.2.

7.2 From the bulk to the surface

Since periodic boundary conditions are inherent to the plane wave code VASP, slabs are used to model surfaces. For the $(000\bar{1})$-O terminated surface, the slab is constructed by stacking bulk unit cells on top of each other in z-direction. After a sufficient number of unit cells, a vacuum layer of several ångström thickness is added on top of the last unit cell to interrupt the pe-

riodicity in the z-direction. If the vacuum width is sufficient, the interaction between the repeating slabs in z-direction is small.

Since large supercells are needed to model different surface reconstructions, a compromise between calculational accuracy and computational parameters like slab size or vacuum has to be made.
As we will show later, the quantity of interest for the construction of a phase diagram is the change in the surface energy $\Delta\gamma$ (see Eq. 8.9). To judge the reliability of a chosen set of simulation parameters we calculated the change in the surface energy of two different surface reconstructions for different numbers of ZnO double layers, different vacuum widths and cutoff energies.

The chosen reference systems were based on a (1×2) supercell with one (coverage $c_v = 0.5$) and two oxygen surface atoms ($c_v = 1.0$) removed from the topmost surface layer. For all slabs under consideration we started with a clean unreconstructed surface and allowed the three outermost double layers on both sides of the slab to relax [78]. Only for the slabs with only 6 double layers thickness 2 double layer on each side were relaxed. For the calculations, where one or two oxygen surface atoms were removed, we used the clean slab calculation as starting point, removed the corresponding number of oxygen atoms from the surface layer and held the Zn terminated (0001) surface layer fixed during the structural optimization (as required by Eq. 8.3).
Additionally we introduce the two following modifications to the Zn atoms on the Zn terminated side to accelerate the calculation on the O-terminated $(000\bar{1})$ surface: First, the Zn surface atoms were replaced by artificial atoms with the valence of 1.5 e^- similar to [78]. This makes the surface stoichiometric, i.e. the surface posses no additional holes in the valence and electrons in the conduction band. Second, in the adjacent layers up to the last four layers underneath the oxygen surface layer, the Zn atoms with 12 electrons in the valence $(s + d)$ were replaced by Zn atoms with just two s valence electrons. Since a relaxation of the ions in this cell is not feasible due to the pseudo atom, the relaxed supercell of the 280 eV calculation was used and accordingly adopted.
We calculated the change of the surface energy for the two structures mentioned above in dependence on the number of double layers of zinc and oxygen, at two different cut off energies, 280 eV and 400 eV, and with different amount of vacuum (8, 10, 14 Å). For all calculations a $(12 \times 6 \times 1)$ Γ-centred k-point mesh was used.
Both surface terminations show almost identical convergence behaviour with respect to the slab thickness and cutoff energy, where the cutoff energy has only a minor influence on the convergence. Similar to the observed behaviour in [78] a faster convergence with respect to the slab thickness is achieved by reducing the electrons on the outermost Zn atoms on the (0001)-Zn side

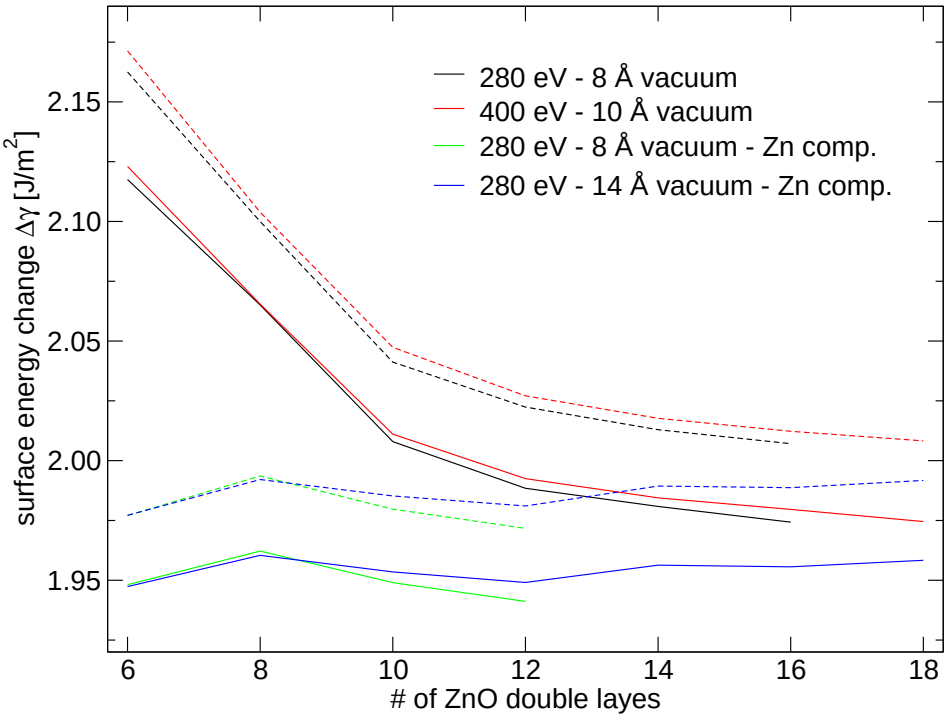


Figure 7.2: The changes in the surface free energy of two structures based on a (1×2) supercell with an oxygen vacancy concentration of $c_v = 0.5$ (solid line) and $c_v = 1.0$ (dashed line) for different vacuum widths and cutoff energies are presented. All dashed lines were shifted downwards by 2.59 J/m^2 to allow for presentation on a single plot.

via the artificial Zn pseudo atom, as shown in Fig. 7.2. While using the compensated slab, 8 double layer of ZnO yield almost the same change in the surface energy as 18 double layers, for the uncompensated slab, the energy difference is almost 0.1 J/m^2.
Therefore, the thickness of the slab was fixed to 8 double layers of ZnO throughout this work. The modification on the Zn atoms were applied and 8 Å of vacuum was added on top of the $(000\bar{1})$-O surface layer, since the difference to 14 Å of vacuum is negligible. This choice represents a well balanced compromise between calculational accuracy and computational efficiency and we expect our results to be converged within 0.02 J/m^2.

Chapter 8

Thermodynamics of the $(000\bar{1})$-O surface

Density functional theory (DFT), as used throughout this thesis, is naively applied as a zero-temperature and zero-pressure theory. Although experiments can be performed close to $T = 0$ K, the majority of processes relevant to catalysis and surface physics explicitly depend on temperature and pressure.
Therefore comparison between experimental quantities observed at finite temperatures and the results of DFT calculations can be rather challenging (pressure gap), especially, if the observed physical quantities are temperature and pressure dependent. The determination of the most stable structure and the composition of a surface exposed to certain environmental conditions, including different temperatures and partial pressures needs to be done with the some care.
It is a matter of fact that the structure of this particular surface evolves by taking into account adsorption-, desorption- and reconstruction processes. If we are interested in equilibrium properties and consider the kinetics to be sufficiently fast to reach their equilibrium, we can use standard equilibrium thermodynamics. The formalism we present here has already been successfully applied in a variety of surface studies before [78, 86, 87]. In the following we will briefly introduce the main concepts, for further reading we refer to Ref. [87].

8.1 Surface free energy

Lets consider a surface of area A per supercell, exposed to a surrounding atmosphere of particles consisting of different elements x, characterized by pressure p and temperature T. The atmosphere acts as a reservoir, where the particles under consideration are able to adsorb on and desorb from the surface without changing the temperature or the pressure. The thermody-

namic quantity of interest to judge the stability is the surface free energy per surface area

$$\gamma(T, \{p_x\}) = \frac{1}{A}\left[G(T, p, \{n_x\}) - \sum_x n_x \mu_x(T, p_x)\right]. \tag{8.1}$$

In Eq. (8.1), G is the Gibbs free energy of the solid exposing a surface to the environment, n_x are the number of particles of type x in the solid, and μ_x and p_x represent the chemical potential and the partial pressure of particles x in the reservoir. The most stable surface reconstruction will minimize this quantity at a given temperature T and partial pressures p_x.
As already discussed in Sec. 7.2, VASP is based on 3D-periodic boundary conditions, and hence finite sized slabs are used to model any surface of interest. This implies that both surface terminations are always present in a slab calculation and both sides are exposed to the vacuum and the reservoir respectively. Therefore it is not possible to discriminate between the different surface energies of both slab sides within a single slab calculation. Eq. (8.1) implies that the surface energy obtained from a free slab calculation is the sum of the surface energies of both sides of the slab

$$\gamma = \gamma_{0001} + \gamma_{000\bar{1}}. \tag{8.2}$$

Since we are only interested in the $(000\bar{1})$ side of the ZnO-surface, we keep the (0001) surface fixed during all our calculations, and we relate all energies relative to the ideal bulk truncated surface. The difference in the surface energy between the ideal (id) and the — in some way — reconstructed configuration (rec) is then simply the difference in the surface energy of the $(000\bar{1})$ side, since the (0001) surface is identical ($\gamma^{\text{id}}_{0001} = \gamma^{\text{rec}}_{0001}$) for both configurations

$$\begin{aligned}\Delta\gamma &= \gamma^{\text{rec}}_{0001} + \gamma^{\text{rec}}_{000\bar{1}} - \gamma^{\text{id}}_{0001} + \gamma^{\text{id}}_{000\bar{1}} \\ \Delta\gamma &= \gamma^{\text{rec}}_{000\bar{1}} - \gamma^{\text{id}}_{000\bar{1}}.\end{aligned} \tag{8.3}$$

As seen from Eq. (8.1), the whole formalism is based on the Gibbs free energy of the system (which represents the thermodynamic potential as a function of the variables which are most easily accessible to experiment, T and p), whereas we obtain total energies within the density functional framework from our calculations. In [78, 87, 88] it was shown that the Gibbs free energy can be approximated by the total energy, if zero-point vibrations, vibrational entropy contributions and enthalpy changes are neglected. Nevertheless care has to be taken since in a recent work [79] it has been shown that the vibrational entropy of the (0001)-Zn surface can play a decisive role at elevated temperatures, when different surface terminations with similar slope in the phase diagram are close to each other within a narrow energy range.

Within this study we concentrate on the three major atomic species involved in the experimental setup, namely the constituents of the bulk (Zn, O) and the molecules in the gaseous environment in the UHV chamber $\frac{1}{2}O_2$ and $\frac{1}{2}H_2$. Then Eq. (8.1) can be explicitly written as

$$\gamma(\{\mu_x\}) = \frac{1}{A}\left[E(n_{\mathrm{Zn}}, n_{\mathrm{O}}, n_{\mathrm{H}}^{\mathrm{add}}) - n_{\mathrm{Zn}}\mu_{\mathrm{Zn}} - n_{\mathrm{O}}\mu_{\mathrm{O}} - n_{\mathrm{H}}^{\mathrm{add}}\mu_{\mathrm{H}}\right], \tag{8.4}$$

where temperature and partial pressure dependencies have been abandoned for the ease of reading. It is possible to reduce the complexity of Eq. (8.4) by assuming that bulk ZnO is in thermal equilibrium with the Zn and O reservoirs. If this would not be the case, bulk ZnO would adsorb or desorb from the surface, which is undesirable in the experiment. This implies that the chemical potentials of both constituents are no longer independent, and have to satisfy

$$\mu_{\mathrm{Zn}} + \mu_{\mathrm{O}} = E_{\mathrm{ZnO}}^{\mathrm{DFT}}. \tag{8.5}$$

When inserting Eq. (8.5) in Eq. (8.4) it is possible to eliminate the chemical potential of zinc and obtain an oxygen and hydrogen chemical potential dependence of the surface energy only

$$\gamma(\mu_{\mathrm{O}}, \mu_{\mathrm{H}}) = \frac{1}{A}\Big[E(n_{\mathrm{Zn}}, n_{\mathrm{O}}, n_{\mathrm{H}}^{\mathrm{add}}) - n_{\mathrm{H}}^{\mathrm{add}}\mu_{\mathrm{H}} - n_{\mathrm{Zn}}E_{\mathrm{ZnO}}^{\mathrm{DFT}} - (n_{\mathrm{O}} - n_{\mathrm{Zn}})\mu_{\mathrm{O}}\Big]. \tag{8.6}$$

For the calculation of the change in the surface energy $\Delta\gamma$ as given in Eq. (8.3) we can explicitly write both constituents

$$\begin{aligned}
\gamma_{000\bar{1}}^{\mathrm{id}} &= \frac{1}{A}\left[E^{\mathrm{id}}(n_{\mathrm{Zn}}, n_{\mathrm{O}}) - n_{\mathrm{Zn}}E_{\mathrm{ZnO}}^{\mathrm{DFT}} - (n_{\mathrm{O}} - n_{\mathrm{Zn}})\mu_{\mathrm{O}}\right] \\
\gamma_{000\bar{1}}^{\mathrm{rec}} &= \frac{1}{A}\Big[E^{\mathrm{rec}}(n'_{\mathrm{Zn}}, n'_{\mathrm{O}}, n_{\mathrm{H}}^{\mathrm{add}}) - n_{\mathrm{H}}^{\mathrm{add}}\mu_{\mathrm{H}} - n'_{\mathrm{Zn}}E_{\mathrm{ZnO}}^{\mathrm{DFT}} - (n'_{\mathrm{O}} - n'_{\mathrm{Zn}})\mu_{\mathrm{O}}\Big]
\end{aligned} \tag{8.7}$$

and by introducing the Zn and O surplus of the reconstructed surface, as compared to the ideal surface,

$$n_{\mathrm{Zn}}^{\mathrm{vac}} = n_{\mathrm{Zn}} - n'_{\mathrm{Zn}} \qquad n_{\mathrm{O}}^{\mathrm{vac}} = n_{\mathrm{O}} - n'_{\mathrm{O}} \tag{8.8}$$

the difference of the reconstructed and ideal cleaved surface termination can be written as

$$\Delta\gamma(\mu_{\mathrm{O}}, \mu_{\mathrm{H}}) = \frac{1}{A}\Big[E^{\mathrm{rec}} - E^{\mathrm{id}} - n_{\mathrm{H}}^{\mathrm{add}}\mu_{\mathrm{H}} + n_{\mathrm{Zn}}^{\mathrm{vac}}E_{\mathrm{ZnO}}^{\mathrm{DFT}} + (n_{\mathrm{O}}^{\mathrm{vac}} - n_{\mathrm{Zn}}^{\mathrm{vac}})\mu_{\mathrm{O}}\Big]. \tag{8.9}$$

8.2 Reference energies

As we will see in the next section, binding and total energies of the hydrogen and oxygen dimer as well as of bulk ZnO and hexagonal Zn are needed for the calculation of the change in the surface free energy and boundaries of the phase diagrams in dependence of the oxygen and hydrogen partial pressures. Therefore we briefly summarize here the required reference energies.

All structures under consideration, except oxygen, were fully optimized using a cutoff energy of 600 eV. For the molecules a large unit cell was chosen to reduce the effect of the periodic boundary conditions. For the bulk systems a $16 \times 16 \times 16$ Γ-centred k-point mesh was used, for the free atoms and molecules, the Γ-point only was used. After the full optimization the cutoff was reduced to 280 eV to match the cutoff used throughout the slab calculations and the total energies $E_{\mathrm{tot}}^{\mathrm{sys}}$ were calculated for a single run without geometry optimization.
For the calculation of the binding energies E_{bind} also the atomic total energies $E_{\mathrm{tot}}^{\mathrm{atom}}$ are needed (Tab. 8.1)

$$E_{\mathrm{bind}} = \sum_{\mathrm{atom}} E_{\mathrm{tot}}^{\mathrm{atom}} - E_{\mathrm{tot}}^{\mathrm{sys}}. \tag{8.10}$$

For the hydrogen and the oxygen atom as well as the oxygen dimer a spin polarized calculation has to be used, to account for their natural ground states, following Hund's rule and the paramagnetic state for O_2, respectively. The oxygen pseudopotential as described in Tab. 7.2 and denoted with "soft" from now on, is well suited for slab calculations but when applied to the oxygen dimer the bond length is too large by 8.5 % in comparison to the experimental value of 1.21 Å. Therefore a much "harder" pseudopotential is used for the reference calculations on the oxygen atom and dimer with core radii of 0.582 Å for the s and p orbitals. With this "hard" pseudopotential the deviation from the experimental bond length can be reduced to 0.8 %. Furthermore the non-spherical contributions to the gradient corrections inside the PAW spheres are included (`LASPH=.TRUE.`) and the interpolation formula according to Vosko, Wilk and Nusair [6] is used for the correlation part of the exchange correlation functional (`VOSKOWN=1`).

Tab. 8.2 gives the data obtained within this study and the theoretical values and experimental heat of formations given in Ref. [89] for comparison. For consistency the total energy of the oxygen dimer has to be redefined, since we are using a soft and hard pseudopotential for the slab and the reference energy calculations, respectively:

$$E_{\mathrm{tot}}^{\mathrm{O_2}} = E_{\mathrm{O_2}}^{\mathrm{hard}} - 2E_{\mathrm{O}}^{\mathrm{hard}} + 2E_{\mathrm{O}}^{\mathrm{soft}} = -9.879 \text{ eV}. \tag{8.11}$$

[eV]	H	O^{soft}	O^{hard}	Zn
E_{tot}^{atom}	-1.11	-1.787	-1.846	0.00

Table 8.1: Total energies for the single atom calculations of H, O and Zn. For oxygen a "soft" and "hard" pseudopotential was used.

[eV]	H_2	O_2	bulk Zn	bulk ZnO
E_{tot}^{sys}	-6.74	-9.88	-1.11	-9.02
E_{bind}	4.52	6.31	1.11	2.97
Ref. [89]	4.50	6.61	1.12	2.84
H_{hf}^{exp}	4.52	5.17	1.35	3.50

Table 8.2: Calculated total E_{tot}^{sys} and binding E_{bind} energies for the isolated H_2 and O_2 molecule as well as the for bulk hexagonal Zn and ZnO. In addition, theoretical values and experimental heat of formations (H_{hf}^{exp}) are given from Ref. [89] for comparison. Energies are given per formula unit.

As one can see, our results are in overall good agreement with the former calculations done by Meyer [89], but we find stronger binding for bulk ZnO and weaker binding for O_2, which we attribute to more accurate potentials.

8.3 Simulation of the phase diagram

In principle, the most stable surface reconstruction at a given temperature and partial pressure of components x in the gas phase can be found by performing an unconstrained minimization of Eq. (8.9). Due to the lack of an algorithm suitable for this task, we proposed and investigated different types of surface reconstructions, involving oxygen and zinc vacancies or hydrogen adatoms.
For each type of surface reconstruction we calculated the change in the surface free energy $\Delta\gamma$ at the conditions imposed by the chemical potentials μ_O and μ_H. Using these values we can build up a two dimensional surface phase diagram, where each axis corresponds to a chemical potential and for each type of surface a different colour is used in the phase diagram. The surface with the lowest surface free energy at a given set of chemical potentials is then the most stable one at these environmental conditions.
It is important to keep in mind that this procedure is not equivalent to an unconstrained minimization of Eq. (8.9), since different types of surface reconstructions enter the calculation as proposed input structures, but not via an automatized structure fit. Therefore, it is in principle possible

that a particular surface reconstruction lower in energy than the considered surface models, is missing in the present study. To minimize this risk, we considered a large number of different types of surface terminations and reconstructions and, furthermore, we used the available experimentally expertise for the $(000\bar{1})$-O surface.

It is common practice, to reference all energies and chemical potentials to half of the binding energy of the molecules under consideration, i.e.:

$$\Delta\mu_{\mathrm{O}} = \mu_{\mathrm{O}} - 1/2E_{\mathrm{O}_2}, \quad \Delta\mu_{\mathrm{H}} = \mu_{\mathrm{H}} - 1/2E_{\mathrm{H}_2}. \tag{8.12}$$

This implies that the hydrogen and oxygen molecules have zero chemical potential $\Delta\mu_{\mathrm{O}}$, $\Delta\mu_{\mathrm{H}}$ at $T = 0$ K. Using Eq. (8.12) we can rewrite Eq. (8.9) to:

$$\begin{aligned}\Delta\gamma(\Delta\mu_{\mathrm{O}}, \Delta\mu_{\mathrm{H}}) =& \frac{1}{A}\Big[E^{\mathrm{rec}} - E^{\mathrm{id}} - n_{\mathrm{H}}^{\mathrm{add}}(\Delta\mu_{\mathrm{H}} + 1/2E_{\mathrm{H}_2}) \\ &+ n_{\mathrm{Zn}}^{\mathrm{vac}} E_{\mathrm{ZnO}}^{\mathrm{DFT}} + (n_{\mathrm{O}}^{\mathrm{vac}} - n_{\mathrm{Zn}}^{\mathrm{vac}})(\Delta\mu_{\mathrm{O}} + 1/2E_{\mathrm{O}_2})\Big]. \end{aligned} \tag{8.13}$$

The bounds of the chemical potentials for Zn, O and H are given by restrictions one has to impose on the system. If the chemical potential of hydrogen $\Delta\mu_{\mathrm{H}}$ or oxygen $\Delta\mu_{\mathrm{O}}$ is larger than 0 (which equals the binding energy of the molecule), molecular hydrogen and oxygen would start to condensate on the surface. The chemical potential of Zn has to be smaller than the energy of bulk Zn, because otherwise oxygen would desorb from the surface into the oxygen reservoir and the remaining Zn near the surface would reconstruct to hexagonal bulk Zn. In summary, the constraints are

$$\Delta\mu_{\mathrm{O}} < 0, \quad \Delta\mu_{\mathrm{H}} < 0, \quad \mu_{\mathrm{Zn}} < E_{\mathrm{Zn}}^{\mathrm{bulk}}. \tag{8.14}$$

The lower bound for the oxygen chemical potential $\Delta\mu_{\mathrm{O}}$ can be derived by inserting Eq. (8.5) in the upper bound for the chemical potential of zinc and using Eq. (8.12):

$$\begin{aligned} E_{\mathrm{ZnO}}^{\mathrm{DFT}} - E_{\mathrm{Zn}}^{\mathrm{bulk}} &< \mu_{\mathrm{O}} \\ E_{\mathrm{ZnO}}^{\mathrm{DFT}} - E_{\mathrm{Zn}}^{\mathrm{bulk}} - \frac{1}{2}E_{\mathrm{O}_2} &< \Delta\mu_{\mathrm{O}}. \end{aligned} \tag{8.15}$$

This results in the heat of formation of ZnO for the lower bound, which in our computational setup, neglecting zero point vibrations, computes to ≈ -3.0 eV.
Furthermore we restricted the lower bound of the chemical potential of hydrogen also to -3.0 eV. As we will see later when comparing to experiment, this is a sensible choice since experiments usually are performed at larger hydrogen chemical potentials. An additional process that has to be taken care of is the condensation of water on the surface. Avoidance of this is

H_2				O_2			
T [K]	$H(T)$	$S(T)$	$\mu^0_{H_2}(T)$	T [K]	$H(T)$	$S(T)$	$\mu^0_{O_2}(T)$
0	-8.467	0.000		0	-8.683	0.000	
298	0.000	130.680	-0.158	298	0.000	205.147	-0.272
700	11.749	155.606	-0.460	700	12.499	231.466	-0.730
800	14.702	159.548	-0.541	800	15.835	235.921	-0.851
1000	20.680	166.216	-0.710	1000	22.703	243.578	-1.010
1100	23.719	169.112	-0.797	1100	26.212	246.922	-1.227

Table 8.3: Values of enthalpy and entropy from Ref. [90]. The enthalpy H is given in units of $[\frac{\text{kJ}}{\text{mol}}]$ whereas the entropy T is given in units of $[\frac{\text{J}}{\text{K}\cdot\text{mol}}]$. The final chemical potentials at standard pressure μ^0_x are given in units of [eV], which is used throughout this work.

equivalent to the restriction that the sum of the chemical potential of oxygen and two times the chemical potential of hydrogen has to be smaller than the formation energy of water

$$\Delta\mu_O + 2\Delta\mu_H < E_{H_2O} - \frac{1}{2}E_{O_2} - E_{H_2} \approx -2.61 \text{ eV}. \tag{8.16}$$

This region will be marked in the phase diagrams by a white coloured area.

8.4 Pressure and temperature dependence of the chemical potentials

For comparison of the phase diagram calculated from ab initio with the experimental findings it is essential to know the temperature and pressure dependence of the chemical potentials used in the phase diagram. We use the ideal gas equation [87, II.E]

$$\mu_x(T, p_x) = \mu^0_x(T, p_0) + \frac{1}{2}k_B T \ln\left(\frac{p_x}{p_0}\right), \tag{8.17}$$

with

$$\mu^0_x(T, p_0) = \frac{1}{2}\left[H(T, p_0, x) - H(0\,\text{K}, p_0, x)\right] - \frac{T}{2}\left[S(T, p_0, x) - S(0\,\text{K}, p_0, x)\right]. \tag{8.18}$$

This yields the temperature dependence of the chemical potential μ^0_x from the differences of the enthalpy H and entropy S with respect to the zero temperature (0 K) limit.

T [K]	μ_x	p_H [mbar]	p_O [mbar]
298	0	$2.209 \cdot 10^{8}$	$1.585 \cdot 10^{12}$
	-3	$7.452 \cdot 10^{-94}$	$5.348 \cdot 10^{-90}$
723	0	$4.613 \cdot 10^{9}$	$3.694 \cdot 10^{13}$
	-3	$6.920 \cdot 10^{-33}$	$5.542 \cdot 10^{-29}$
1023	0	$1.558 \cdot 10^{10}$	$1.330 \cdot 10^{14}$
	-3	$4.304 \cdot 10^{-20}$	$3.674 \cdot 10^{-16}$

Table 8.4: Conversion table from chemical potentials at given temperature to corresponding oxygen and hydrogen partial pressures.

As will be seen later in this work, the temperatures which are relevant to this study are the room temperature 298 K (25°C), 723 K (450°C), which is the temperature were the STM images were recorded, and 1023 K (750°C), which is the preparation temperature of the ZnO samples in [91]. The values for the enthalpy and entropy are tabulated in dependence of the temperature for the standard pressure of $p_0 = 1$ bar in thermochemical tables such as Ref. [90]. We need to calculate the chemical potentials at standard pressure at temperatures above and below the target temperatures. The values for the element specific enthalpy and entropy as well as $\mu_x^0(T, p_0)$ are given in Tab. 8.3 for hydrogen and oxygen.
Since the values are given at intervals of 100 K, besides 298 K, we have to interpolate between the given temperatures. A linear interpolation turned out to be sufficient for getting accurate results for the phase diagrams. Experiments were performed exclusively at few fixed temperatures hence we are mainly interested in the conversion of the chemical potential into partial pressures at these given temperatures. Rewriting Eq. (8.17) yields the partial pressure p_x in dependence of the temperature T and the chemical potential μ_x:

$$p_x(T, \mu_x) = p_0 \cdot e^{\frac{2(\mu_x - \mu_x^0(T, p_0))}{k_B T}}. \tag{8.19}$$

In Tab. 8.4 we summarize the partial pressures for oxygen and hydrogen at the phase diagram boundaries, as determined in Sec. 8.3, in dependence of the temperatures used in experiment.

Chapter 9

Results

9.1 Compensation mechanisms

Oxygen has a high electron affinity in comparison to the electron affinity of zinc. Therefore, it is to be expected that in bulk ZnO the two $4s$ valence electrons of zinc are transferred to the oxygen atoms creating a closed oxygen $2p$ shell. In the pure bulk material each Zn ion is surrounded by 4 O, so in average half an electron is transferred along each bond to an oxygen ion, and no excess charge is present in the bulk.
The situation changes at the (0001) and $(000\bar{1})$ surface. Regardless of the slab surface side, each ion in the surface layer has only three next nearest neighbours. Therefore at the (0001) terminated surface the ionicity of the zinc surface ions is reduced to $Zn^{+3/2}$, whereas at the $(000\bar{1})$ terminated side oxygen has the ionicity of $O^{-3/2}$ and a closed $2p$ shell can not be reached.
This partial ionicity of $\pm 3/2$ forces electrons at the (0001) zinc terminated side into the conduction band, whereas the valence band is not entirely filled at the ideal $(000\bar{1})$ oxygen terminated side. Although these conclusions are drawn from an oversimplified atomistic picture, this model was confirmed to be applicable to ZnO [78] and also to be the main reason for the lower stability of polar surfaces in comparison to their non polar counterpart (Taskers rule does not really apply).
Stabilization of the surface layer can be achieved preferably by changing the surface stoichiometry by means of removing a certain amount of oxygen and zinc from the surface or by adsorbing hydrogen at the oxygen surface atoms.

As example let us consider a slab of a (2×2) supercell. In the surface layer 4 oxygen atoms are present receiving 6 electrons from their neighbouring Zn atoms; 2 electrons are missing to fill the entire $2p$ shell of all 4 oxygen atoms In other words, two electrons lack to fully populate the valence band at the oxygen terminated surface side. As stated above, the removal of one out of the four oxygen surface atoms is one of the possible mechanisms, because

then only 6 additional electrons are needed to fully occupy the oxygen $2p$ valence band. Therefore, the removal of a quarter of the oxygen surface atoms is likely to be the concentration of oxygen vacancies which is most stable.
The other possible mechanism is the adsorption of two hydrogen atoms on top of surface oxygen atoms and the formation of OH^- hydroxyl groups. Two electrons are then transferred to the oxygen surface atoms and the valence band is again fully populated. Here a coverage of half of the oxygen surface atoms by hydrogen should be the most stable configuration. Additionally the combination of both compensation mechanisms, as well as the simultaneous removal of oxygen and zinc atoms can lead to stable surface terminations.

In the following we will investigate the stability of the $(000\bar{1})$-O terminated surface with respect to more or less randomly distributed oxygen vacancies. In a second step we introduce hydrogen on the clean unreconstructed surface. Then we will focus on the simultaneous creation of oxygen and zinc vacancies, which is essential for more complex surface reconstructions like triangles and the hexagonal holes. The last part is dedicated to the comparison with recent experimental findings.

9.2 Oxygen vacancies

As mentioned in the preceding introduction the polar $(000\bar{1})$-O terminated ZnO surface is supposed to be stabilized by depleting oxygen from the surface. In a recent study, Meyer [89] investigated the behaviour of the surface free energy in dependence of the oxygen chemical potential and the concentration of oxygen vacancies in the surface layer.
The quantity of interest is the change of the surface free energy with respect to the surface energy of the bulk terminated slab. Using Eq. (8.13) we obtain

$$\begin{aligned} \Delta\gamma(\Delta\mu_{\mathrm{O}}) &= \gamma_{\mathrm{slab}}^{\text{O-vac}} - \gamma_{\mathrm{slab}} \\ &= \frac{1}{A}\left[E_{\mathrm{slab}}^{\text{O-vac}} - E_{\mathrm{slab}} + n_{\mathrm{O}}^{\mathrm{vac}}(\Delta\mu_{\mathrm{O}} + 1/2E_{\mathrm{O_2}})\right]. \end{aligned} \tag{9.1}$$

Meyer [89] used the smallest possible supercells to model the different vacancy concentrations c_v of 0, 1/4, 1/3, 1/2 and 1, namely a (1×1), (2×2), (1×3) and a (1×2) supercell. Initially, we adopted the supercells so that a comparison of our calculated data and the data presented in [89] is possible. As described before the chemical potentials are referenced to half of the binding energy of the oxygen and hydrogen molecule. In contrast to our study, where the calculated binding energy of the oxygen molecule is used throughout the work, Meyer [89] used an alternative approach. Meyer mixed experimental and theoretical values to obtain the reference energy for

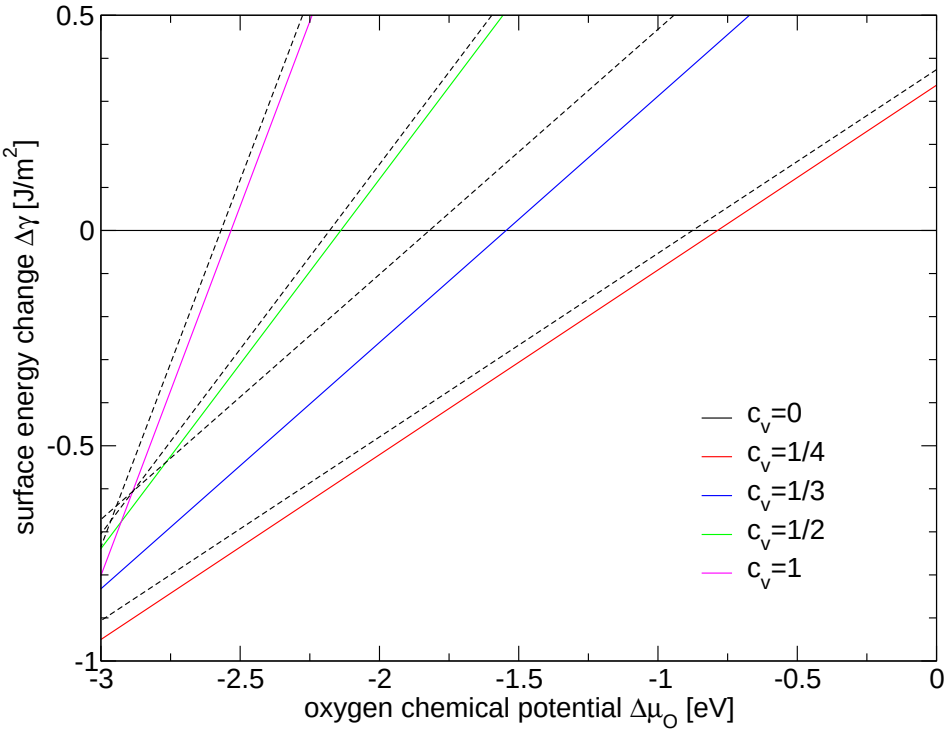


Figure 9.1: Comparison between our calculations and the results of Meyer [89]. We have shifted the oxygen chemical potential as given in the work of Meyer by 0.66 eV towards the positive direction for compatibility reasons, as noted in the text.

the oxygen dimer. To this end, he used the experimental formation energy of ZnO

$$E_{\mathrm{ZnO}}^{\mathrm{form}} = \left| E_{\mathrm{ZnO}} - E_{\mathrm{Zn}} - \frac{1}{2} E_{\mathrm{O_2}} \right| \tag{9.2}$$

to determine $E_{\mathrm{O_2}}$. In this way errors introduced by a poor description of the oxygen dimer within DFT, can be avoided. The point is that DFT is expected to be accurate for E_{ZnO} and E_{Zn}, $E_{\mathrm{ZnO}}^{\mathrm{form}}$ is known experimentally, thus $E_{\mathrm{O_2}}$can be deduced via Eq. (9.2).

The experimental value for the formation energy of ZnO is 3.50 eV, whereas the ab initio value was found to be 2.84 eV in the study of Meyer (see Tab. 8.2). Hence this redefinition of the oxygen chemical potential using experimental values corresponds to a shift of the oxygen reference energy $1/2E_{\mathrm{O_2}}$ by 0.66 eV. To compare with our first principles data (where such a correction is not applied), we need to shift the oxygen chemical potential $\Delta\mu_{\mathrm{O}}$ as give in Fig. 2 in [89] by 0.66 eV to the positive direction ($\Delta\mu_{\mathrm{O}} = \Delta\mu_{\mathrm{O}} + 0.66$ eV).

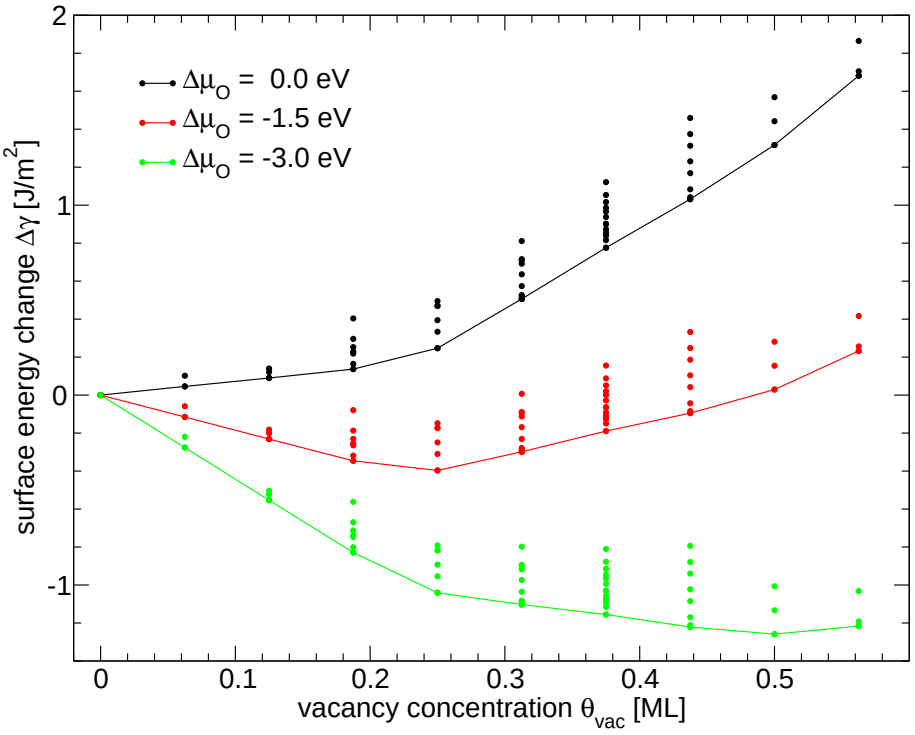


Figure 9.2: Change in the surface energy $\Delta\gamma$ in dependence of the oxygen vacancy concentration θ_{vac} for three different oxygen chemical potentials. Each of the 73 considered surface terminations is denoted by a filled circle and the lowest ones are interlinked by solid lines.

As can be seen in Fig. 9.1, our theoretical findings coincide well with the predicted behaviour by Meyer [89]. Meyer already pointed out in his study, that Fig. 9.1 can only be used to determine surface reconstructions within the set of considered structures lowest in energy at a certain oxygen or hydrogen partial pressure . It is evident that a restriction to a (2×2) supercell leads to a very limited number of surface structures, which are, in addition, highly periodic and perfectly ordered. For the (2×2) case, all possible reconstructions are already shown in Fig. 9.1. It is very likely that surface reconstructions exist which are lower in energy than those shown in Fig. 9.1, since pits and islands are found to be stable on the Zn terminated (0001) surface [78]. To extend the set of possible surface vacancy structures we expanded our calculations to a (4×4) supercell, which is 4 times larger than the previously used supercell. This approach has two major benefits: first of all, we are able to investigate the development of oxygen vacancies in dependence of the oxygen partial pressure more precisely. Second, for a given

oxygen vacancy concentration the number of possible vacancy arrangements increases drastically. For example there is only 1 possible vacancy structure for the (2×2) supercell to model the vacancy concentration $c_v = \frac{1}{4}$ whereas we investigated 7 symmetry inequivalent structures in the (4×4) supercell, exhibiting eg. clustering and different 2D ordering of the vacancies.

In total we investigated 73 different configurations of oxygen vacancies for different concentrations. To determine the most stable structures we plot the change in the surface free energy with respect to the oxygen vacancy concentration as given in Fig. 9.2 for three different oxygen chemical potentials. The variable transformation from the oxygen chemical potential $\Delta\gamma(\Delta\mu_{\mathrm{O}})$ as given in Eq. (9.1) to the oxygen vacancy concentration $\Delta\gamma(\theta_{\mathrm{vac}})$ can be performed by a Legendre transformation:

$$\begin{aligned} \frac{d\Delta\gamma(\Delta\mu_{\mathrm{O}})}{d\Delta\mu_{\mathrm{O}}} &= \frac{n_{\mathrm{O}}^{\mathrm{vac}}}{A} = \theta_{\mathrm{vac}} \\ \Delta\gamma^{*}(\theta_{\mathrm{vac}}) &= \Delta\gamma(\Delta\mu_{\mathrm{O}}) - \Delta\mu_{\mathrm{O}}\theta_{\mathrm{vac}} \\ &= \frac{1}{A}\left(E_{slab}^{\text{O-vac}} - E_{\mathrm{slab}}\right) + \frac{1}{2}\theta_{\mathrm{vac}}\ \overline{E}_{\mathrm{O}_2}. \end{aligned} \tag{9.3}$$

The chemical potential is absent from this equation but we may vary the oxygen reference energy $\overline{E}_{\mathrm{O}_2}$.
For oxygen rich conditions ($\overline{E}_{\mathrm{O}_2} = E_{\mathrm{O}_2}$) no energy is gained by removing oxygen from the surface, since oxygen is abundant available in the gaseous environment. When lowering the reference energy (which corresponds to a reduction of the oxygen partial pressure) the formation of oxygen vacancies become energetically favourable. For $\overline{E}_{\mathrm{O}_2} = E_{\mathrm{O}_2} - 1.5$ eV the change in the surface energy drops until the vacancy concentration of 0.25 monolayer (ML) is reached. When depleting the surface layer further the surface energy increases again until, at $\theta_{\mathrm{vac}} = 0.5$ ML, the formation of oxygen vacancies becomes energetically unfavourable. When almost no gaseous oxygen is present ($\overline{E}_{\mathrm{O}_2} = E_{\mathrm{O}_2} - 3.0$ eV), the energy change drops until half of the surface layer is depleted.
Additionally we see that the change of the slope of the enveloping function is largest at a quarter vacancy concentration. Therefore it is to be expected that this surface reconstruction is particularly stable in a wide range of the chemical potentials $\Delta\mu_{\mathrm{O}}$. Another feature we can extract from Fig. 9.2 is the fact whether a reconstruction appears in the phase diagram. If the enveloping function has a convex shape, the vacancy reconstructions will appear in the phase diagram. As one can see in Fig. 9.2, at $\theta_{\mathrm{vac}} = 6/16$ the data point is located slightly above the line between the neighbouring points and as a consequence the reconstruction at this vacancy concentration will not show up in the phase diagram (see Fig. 9.3).

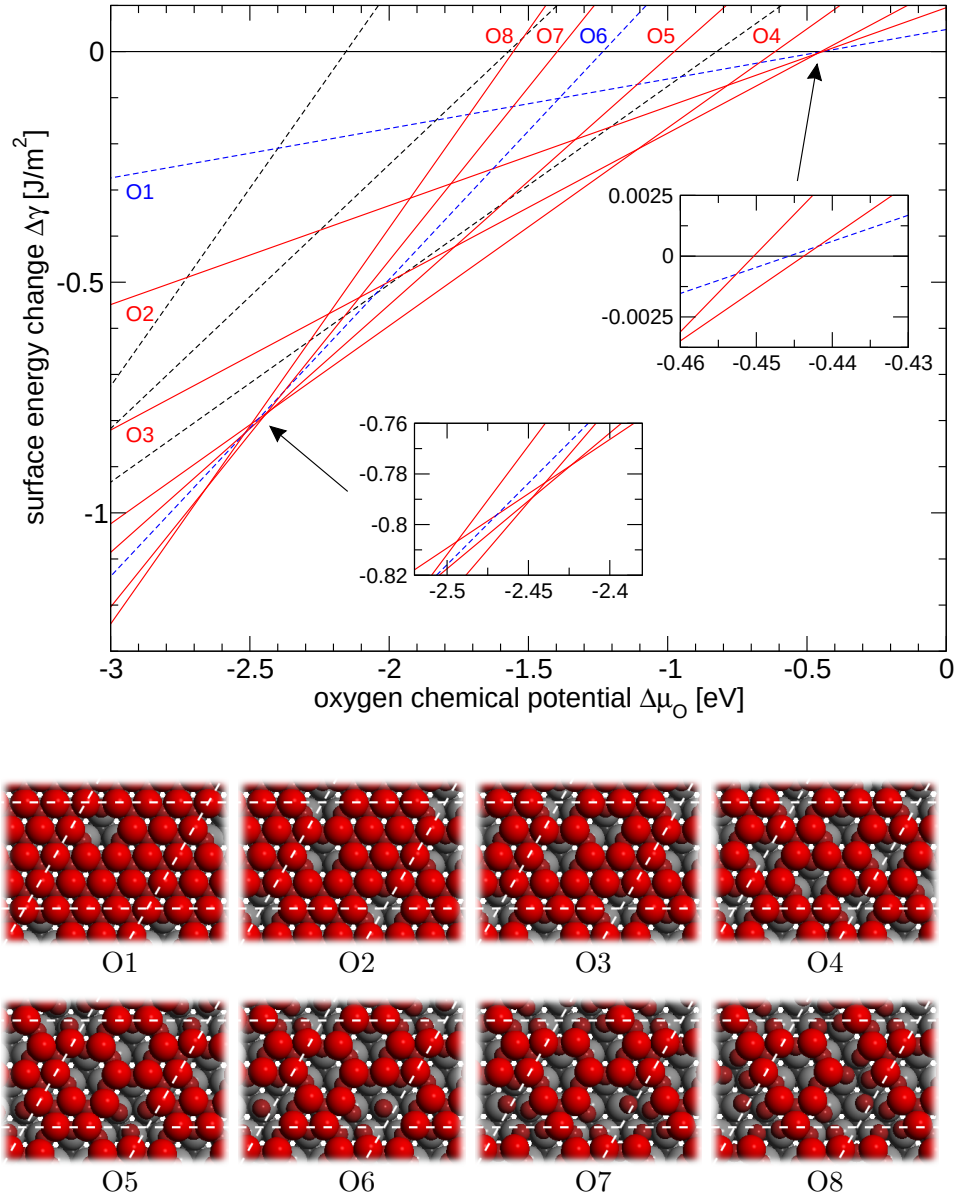


Figure 9.3: Phase diagram for the most stable oxygen vacancy configurations for different oxygen vacancy concentrations in dependence of the oxygen chemical potential. The black dotted lines are the phases from the small supercell (see Fig. 9.1). In the insets crucial points are magnified, where the blue dotted lines mark the phases with concentrations of 1/16 and 6/16 which do not show up in the phase diagram. Below the phase diagram we present the corresponding ball models of the vacancy configurations for concentrations of $c_v = n/16$.

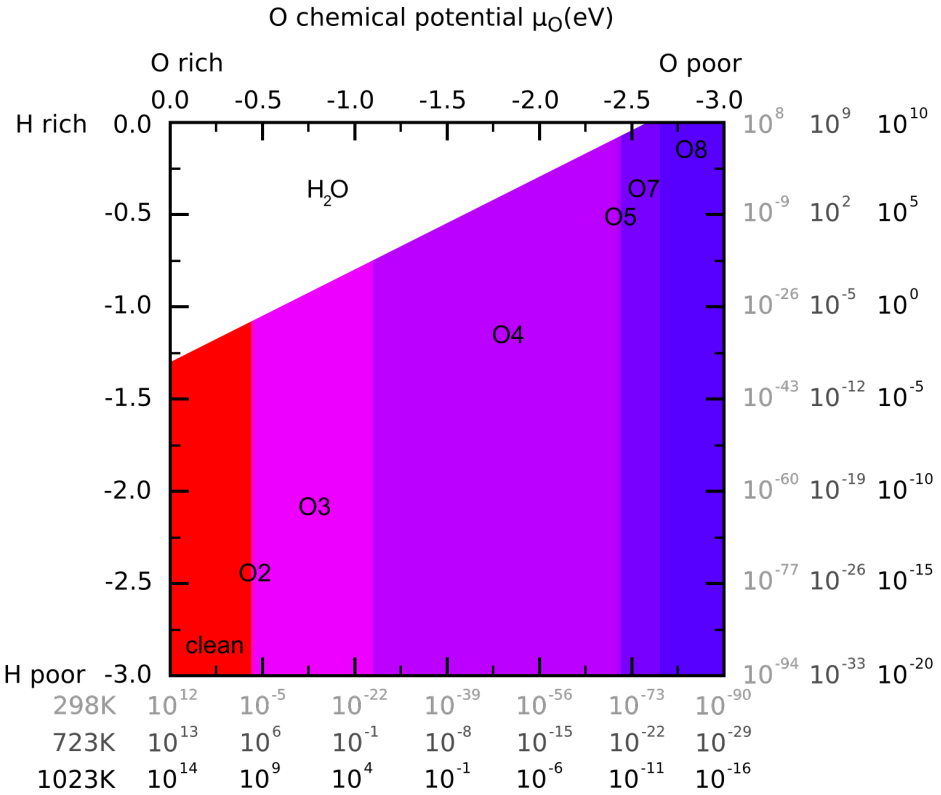


Figure 9.4: Colour coded 2 dimensional phase diagram of the O terminated ZnO-$(000\bar{1})$ surface as a function of O and H chemical potentials. Oxygen vacancies as described in the text were taken into account. Note that the scale of the oxygen chemical potential is reversed in comparison to Fig. 9.3. The corresponding ball models to the oxygen vacancy structures are given in Fig. 9.3

In Fig. 9.3 we present the most stable configurations together with the previously reported structures of the smaller supercells. In comparison to the smaller supercells ((1×2), (1×3) and (2×2)) we find energetically more favourable vacancy configurations. In the case of a vacancy concentration of 1/4 the (4×4) supercell structure is 0.091 J/m^2 lower in energy, and for the vacancy concentration of 1/2 the energy difference in favour to the larger supercell configuration is 0.516 J/m^2.

In addition, we find an almost smooth increase of the oxygen vacancy concentration with decreasing oxygen chemical potential. The decrease in the chemical potential is equivalent to a transition from an oxygen rich ($\Delta\mu_O = 0$ eV) to an oxygen poor ($\Delta\mu_O = -3$ eV) environment. The most

stable configuration in the investigated range of the oxygen chemical potential is the expected one quarter vacancy concentration, which is stable within -2.42 eV $< \Delta\mu_O < -1.10$ eV, so over a range of 1.32 eV in terms of the oxygen chemical potential.
Although the increase looks smooth at first glance, the insets in Fig. 9.3 reveal the absence of the 1/16 and 6/16 phases. For the former the energy difference to show up in the phase diagram is 0.2 mJ/m^2. This difference is negligible, within the estimated error of the present study and as a consequence we are not able to determine whether this phase is indeed unstable. For the 6/16 concentration the energy difference is larger 8 mJ/m^2, but still the same argument holds, since the energy to show up in the phase diagram is still smaller than the estimated error of the calculation.
In Fig. 9.4 we show the 2 dimensional surface phase diagram as a function of both the oxygen and the hydrogen chemical potential. As a matter of fact, no variation along the hydrogen chemical axis is present, since no hydrogen is taken into account so far. It is important to note that the oxygen chemical potential axis has been reversed in these plots, therefore zero concentration is depicted using red colour and when going to high vacancy concentrations the colour changes gradually from violet to blue. The phases with 2/16 and 5/16 concentration are small but still visible in the phase diagram. Additionally we added at the bottom and to the right of the phase diagram for all three temperatures the corresponding partial pressures to the oxygen and hydrogen chemical potential using the values given in Tab. 8.4.
The phases that show up in the phase diagram are shown schematically as ball models in Fig. 9.3. It turned out during the calculation of the different oxygen vacancy patterns that it is preferable to remove oxygen from the surface in a way that hexagonal rings are formed with the remaining surface oxygen atoms. By increasing the vacancy concentration these rings are broken up by removing more and more oxygen from the surface.

Low Energy Electron Diffraction (LEED) and Helium Atom Scattering (HAS) experiments performed by Kunat *et al.* [92] indicate the presence of a (1×3) structure for the O-terminated surface in an almost hydrogen free environment. This proposed structure is identical to the (1×3) supercell. However, as already pointed out by Meyer [89] and as can be seen in Fig. 9.1 and 9.4, our calculations do not corroborate the stability of this single missing row structure. Moreover a recent LEED experiment relates the (1×3) surface termination to the presence of impurities in the surface [82].

	c_H	on top		hcp hollow	fcc hollow	
clean surface	0.0	+3.009	(+3.16)	0.00	+0.044	(+0.05)
half covered	0.5	+2.826		0.00	+0.008	
full covered	1.0	+1.784	(+1.78)	0.00	−0.006	(−0.02)

Table 9.1: Total energy differences of alternative adsorption sites to the ideal $(000\bar{1})$ bulk truncated surface (hcp hollow site) in eV. As alternative adsorption sites we considered the position on-top of the underlying Zn atoms and in the fcc-hollow position. The pure oxygen surface layer, as well as half coverage and a full monolayer coverage where considered. The number in parenthesis are theoretical values from Ref. [89] for comparison.

9.3 Hydrogen adsorption

The adsorption of hydrogen at the $(000\bar{1})$-O surface may lead to a stabilization of the surface and the presence of hydrogen in the experiment, even under ultra high vacuum (UHV) conditions, makes the investigation of the influence on the surface worthwhile.
Hydrogen molecules H_2 can dissociate and the hydrogen atoms may adsorb on the $(000\bar{1})$-O surface forming OH^- groups. In principle, there are three different possible sites for the formed OH^- group. The conventional bulk continuation of ZnO (hcp hollow), the fcc continuation with no Zn underneath, and on-top of a Zn atom of the layer underneath. We calculated the total energy of all surface models at zero, half and full hydrogen coverage, and related the energies for the fcc hollow and on-top structure to the energy of the bulk continuation (hcp hollow) as presented in Tab. 9.1.
For zero hydrogen concentration c_H the hcp hollow positions for the oxygen surface atoms are the most stable ones. Covering the surface with an (2×1) overlay of hydrogen (Fig. 9.6 H8), the hcp hollow site is still the preferable position compared to the two other positions under consideration. One has to note that the difference between the hcp and fcc hollow position is now almost zero. When fully covering the surface with hydrogen atoms, the preferable surface site changes from hcp to the fcc hollow site for oxygen. Again the energy difference is almost zero and it is not relevant on the scale of Fig. 9.6. Therefore we performed the majority of the calculations at the hcp site in a (4×4) supercell.
The change of the surface free energy with respect to the bulk terminated slab can be deduced from Eq. (8.13):

$$\Delta\gamma(\Delta\mu_H) = \frac{1}{A}\left[E_{slab}^{H\text{-}add} - E_{slab} - n_H^{add}(\Delta\mu_H + 1/2E_{H_2})\right]. \tag{9.4}$$

In total we investigated 76 surface terminations and we applied similarly to Sec. 9.2 the Legendre transformation for the variable transformation from

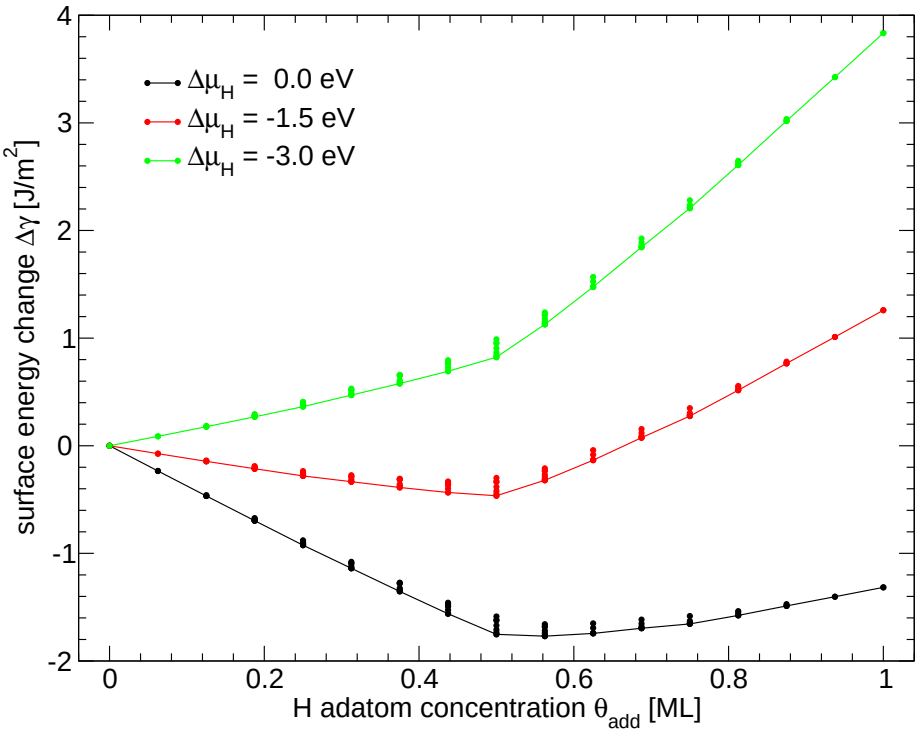
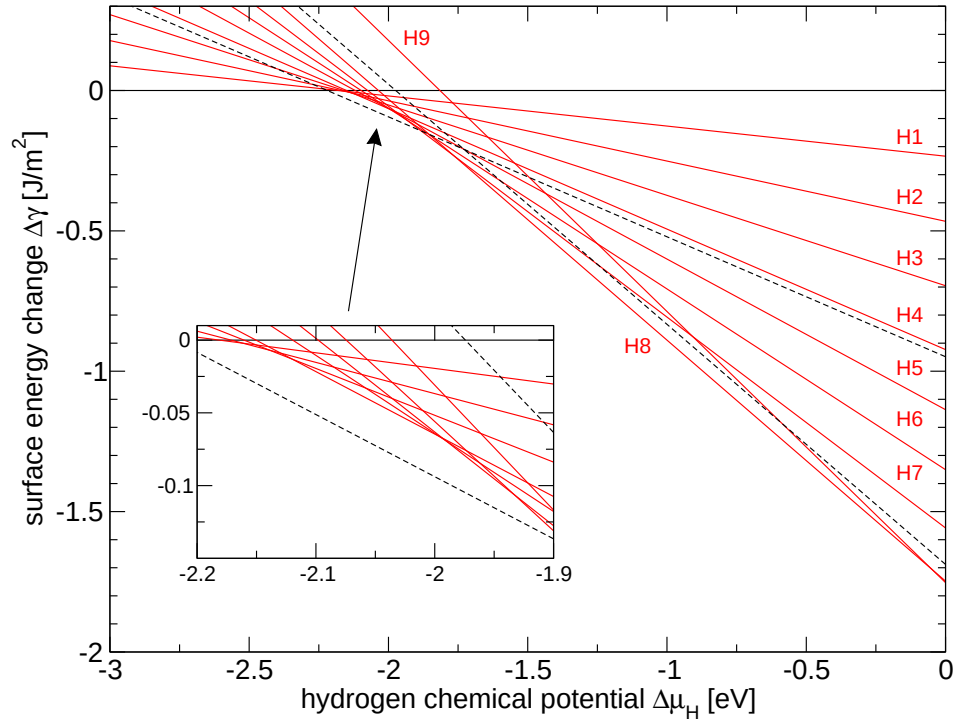


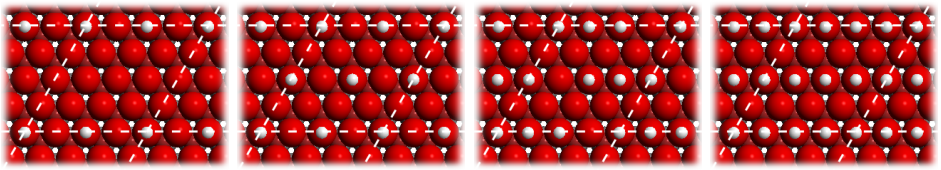
Figure 9.5: Change in the surface energy $\Delta\gamma$ in dependence of the hydrogen adatom concentration θ_{add} for three different hydrogen chemical potentials. Each of the 76 considered surface terminations is denoted by a filled circle and the lowest ones are interlinked by solid lines.

the hydrogen chemical potential $\Delta\gamma(\Delta\mu_{\text{H}})$ to the adatom concentration $\Delta\gamma(\theta_{\text{add}})$

$$\begin{aligned}\frac{d\Delta\gamma(\Delta\mu_{\text{H}})}{d\Delta\mu_{\text{H}}} &= -\frac{n_{\text{H}}^{\text{add}}}{A} = -\theta_{\text{add}} \\ \Delta\gamma^{*}(\theta_{\text{add}}) &= \Delta\gamma(\Delta\mu_{\text{H}}) + \Delta\mu_{\text{H}}\theta_{\text{add}} \\ &= \frac{1}{A}\left(E_{slab}^{\text{H-add}} - E_{\text{slab}}\right) - \frac{1}{2}\theta_{\text{add}}\ \overline{E}_{\text{H}_2}. \end{aligned} \tag{9.5}$$

In a hydrogen rich gaseous environment ($\overline{E}_{\text{H}_2} = E_{\text{H}_2}$) hydrogen will adsorb on the surface. The maximum in the change of the surface energy is reached at an adatom concentration of 9/16. When the hydrogen chemical potential is reduced ($\overline{E}_{\text{H}_2} = E_{\text{H}_2} - 1.5$ eV), which is equivalent to a reduction of the amount of hydrogen in the gas phase, an adatom concentration of 1/2 gives the largest change in the surface energy. Reducing the chemical potential of hydrogen even further ($\overline{E}_{\text{H}_2} = E_{\text{H}_2} - 3.0$ eV) adsorption of hydrogen





H2 H4 H6 H8

Figure 9.6: Change in the surface energy of different hydrogen concentrations on the $(000\bar{1})$-O terminated ZnO surface as function of the hydrogen chemical potential $\Delta\mu_{\mathrm{H}}$. The black dashed lines correspond to the data extracted from Fig. 3 of Ref. [89] for $c_{\mathrm{H}} = 1/4$ and $1/2$.

becomes energetically unfavourable. From the enveloping function we can deduce that the surface covered with half a monolayer of hydrogen will be a particularly stable surface termination in the chosen range of the hydrogen chemical potential $\Delta\mu_{\mathrm{H}}$.

Arranging the hydrogen atoms at the surface of a (4×4) supercell turned out to be less complex than arranging the oxygen vacancies in Sec. 9.2. As intuitively expected, maximizing the distance between the individual hydrogen adatoms is the key property to minimize the total energy of the

ZnO slabs. Surfaces with hydrogen clusters are always found to be higher in energy than configurations with an empty oxygen in between the OH^- complexes. Such an adsorption pattern can only be constructed until a hydrogen concentration of $c_H = 1/4$ is reached. At $c_H = 1/4$ the most stable configuration can be formed by placing 1 hydrogen in a (2×2) periodicity (see Fig. 9.6 H4). By further increasing the concentration of hydrogen, H has to be adsorbed on "interstitial" oxygen surface atoms, where a tendency to form parallel linear chains on the surface is observed. At half coverage the most stable configuration is formed by placing hydrogen atoms in a (1×2) periodicity.
Fig. 9.6 shows the change of the surface free energy in dependence of the hydrogen chemical potential relative to the clean $(000\bar{1})$-O surface. The respective hydrogen concentrations are shown above the corresponding red full lines. The black dashed lines correspond to the data extracted from Fig. 3 of Ref. [89] for $c_H = 1/4$ and $1/2$, which coincide quite well with our results. In the inset we zoomed in the area, where most of the transitions occur.
At low hydrogen pressure (-3.0 eV $< \Delta\mu_H < -2.181$ eV in the phase diagram) the clean surface (without hydrogen adsorbates) is the most stable structure. By increasing the hydrogen partial pressure (-2.181 eV $< \Delta\mu_H < -1.766$ eV) a continuous increase of the adsorbed hydrogen on the surface can be observed from $c_H = 1/16$ to $c_H = 7/16$. The most stable hydroxylated surface found in the phase diagram is the configuration that is half covered with hydrogen. It is the favoured surface termination over a range of 1.716 eV in the phase diagram. This agrees with the stabilization mechanism discussed in the introduction of this section. Only at very high hydrogen partial pressures, the hydrogen coverage can be increased beyond half a monolayer ($c_H = 1/2$). A hydrogen concentration beyond $c_H = 9/16$ is not realized for the considered allowed conditions.
In Fig. 9.7, we see the effect hydrogen adsorption has on an otherwise clean and unreconstructed O-terminated surface phase. Phases with adsorbed hydrogen atoms are coloured from red (0 hydrogen adatoms) over orange to yellow (9 hydrogen adatoms) depending on the number of hydrogen adatoms. At hydrogen rich conditions, half hydrogen coverage is the most stable surface termination, independent of the oxygen partial pressure.

In general, our ab initio results, extrapolated to room temperatures and UHV pressures coincide well with former theoretical findings [89] and also with available experimental diffraction data [92, 93]. Experiments in the past were, without exception, carried out at room temperate and under UHV conditions. Under these conditions both our present results and other previous studies show that the O-terminated ZnO surface quickly saturates with OH groups even in UHV. The surface then adopts a (1×1) terminated surface with half a monolayer coverage of OH (exemplified in our studies by

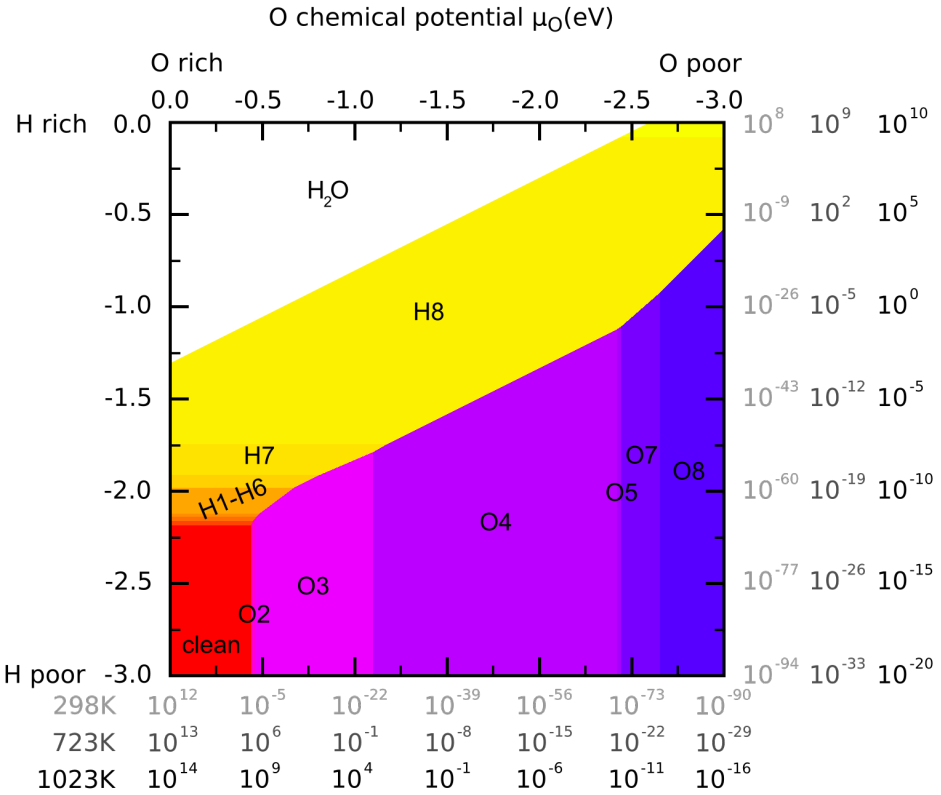


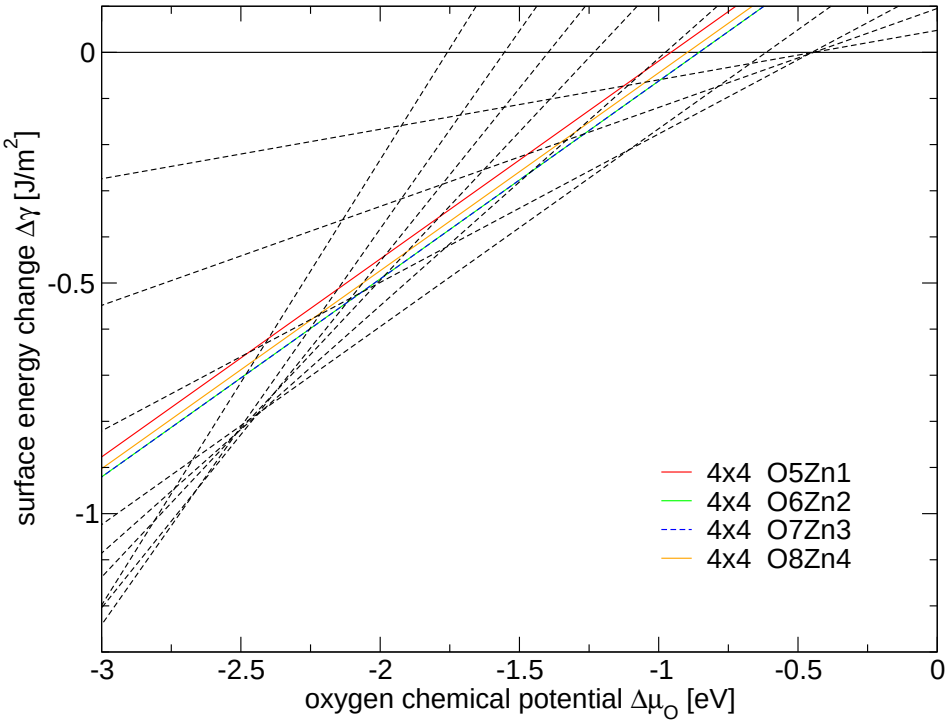
Figure 9.7: Color coded 2 dimensional phase diagram of the O terminated ZnO-$(000\bar{1})$ surface as a function of O and H chemical potentials. Oxygen vacancies and hydrogen add atoms as described in the text were taken into acount.

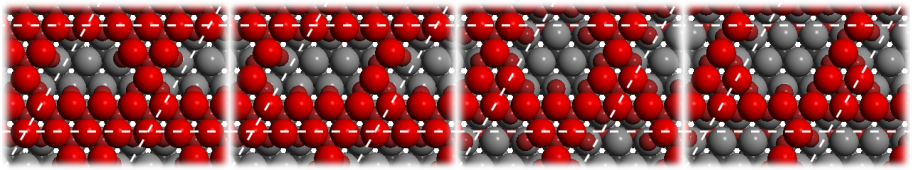
a (1×2)H overlayer structure as given by H8 in Fig. 9.6), which appears as a regular (1×1) structure in diffraction experiments.

9.4 Mixture of oxygen and zinc vacancies

In addition to the sole desorption of the surface oxygen atoms, as seen in Sec. 9.2, also a combination of oxygen and zinc desorption can stabilize the polar $(000\bar{1})$-O surface. By removing the hydrogen dependence in Eq. (8.13) we obtain the change in the surface free energy:

$$\Delta\gamma(\Delta\mu_{\mathrm{O}}) = \frac{1}{A}\Big[E_{\mathrm{slab}}^{\mathrm{Zn,O\text{-}vac}} - E_{\mathrm{slab}} + n_{\mathrm{Zn}}^{\mathrm{vac}} E_{\mathrm{ZnO}}^{\mathrm{DFT}} + (n_{\mathrm{O}}^{\mathrm{vac}} - n_{\mathrm{Zn}}^{\mathrm{vac}})(\Delta\mu_{\mathrm{O}} + 1/2E_{\mathrm{O_2}})\Big]. \quad (9.6)$$





O5Zn1	O6Zn2	O7Zn3	O8Zn4
$n_{\mathrm{O}}^{\mathrm{vac}} = 5$	$n_{\mathrm{O}}^{\mathrm{vac}} = 6$	$n_{\mathrm{O}}^{\mathrm{vac}} = 7$	$n_{\mathrm{O}}^{\mathrm{vac}} = 8$
$n_{\mathrm{Zn}}^{\mathrm{vac}} = 1$	$n_{\mathrm{Zn}}^{\mathrm{vac}} = 2$	$n_{\mathrm{Zn}}^{\mathrm{vac}} = 3$	$n_{\mathrm{Zn}}^{\mathrm{vac}} = 4$

Figure 9.8: Phase diagram similar to Fig. 9.3 with the inclusion of the 4 surface termination involving oxygen and zinc desorption from the first surface double layer with an effective oxygen vacancy concentration of $c_v = 0.25$. Underneath the phase diagram the ball models, the corresponding abbreviations and the number of oxygen $n_{\mathrm{O}}^{\mathrm{vac}}$ and zinc $n_{\mathrm{Zn}}^{\mathrm{vac}}$ vacancies in the (4×4) periodicity are given.

The additional removal of zinc atoms in the layer below the oxygen surface layer goes hand in hand with a larger number of oxygen surface vacancies and a strong clustering of surface vacancies. The first can easily be understood in the ionic picture: each removed zinc atom with ionicity +2 counterbalances the removal of one oxygen surface layer atom with ionicity −2. In other words, to maintain a stoichiometric surface Zn^{+2} and O^{-2} ions

always need to be removed in pairs. The clustering is a necessity for the removal of Zn atoms in the layer underneath the surface, since already for the removal of a single Zn atom the three oxygen surface atoms above have to be removed. Initially we adopted simple clustered vacancy structures as already used in Sec. 9.2, and removed as much zinc within the clustered vacancies as necessary to counterbalance the removal of a total of 4 oxygen atoms in the (4×4) periodicity. At the end we investigated 4 different oxygen and zinc vacancy arrangements, including trapezoidal, rhomboidal and hexagonal pits. All of them have an oxygen vacancy concentration of $c_v = 0.25$. Ball models and abbreviations of these surface terminations are given in Fig. 9.8.
All of the four structures are located close to each other in the phase diagram (Fig. 9.8). The structures O6Zn2 and O7Zn3 are energetically almost identical. The energy differences between the O5Zn1 and O6Zn2 structure to the O4 configuration are 0.147 J/m^2 and 0.103 J/m^2, respectively. Therefore non of the investigated structures is energetically favourable and the O4 structure, investigated in Sec. 9.2, remains the most stable surface termination for all reconstructions with an oxygen vacancy concentration of $c_v = 0.25$.

It has been shown in the past that triangular reconstructions represent an energetically favourable surface termination for the (0001)-Zn terminated side. Therefore we will focus our investigations in the following section on this reconstruction adopted to the $(000\bar{1})$-O surface side.

9.4.1 Triangular pits

Recent studies on the Zn terminated (0001) side of ZnO, involving STM experiments and extensive large scale DFT calculations [78, 94], showed that the stabilization of the (0001)-Zn terminated surface is accomplished by removal of zinc and oxygen atoms of the first and second layer – resulting in the formation of triangular shaped reconstructions. It is evident that this reconstruction mechanism can be adopted to the $(000\bar{1})$-O surface, by interchanging the Zn and O atoms. The formation of triangular reconstructions of length n involves the removal of $n \times (n+1)/2$ oxygen and $n \times (n-1)/2$ zinc atoms and corresponds to

$$n \times (n+1)/2 - n \times (n-1)/2 = n \tag{9.7}$$

effective oxygen vacancies. Removing a n-sized triangular vacancy from the oxygen surface layer requires at least a supercell of the size of:

$$n \text{ even} \quad \ldots \quad \left(\sqrt{\frac{3n^2+6n+4}{4}}\right) \times \left(\sqrt{\frac{3n^2+6n+4}{4}}\right)$$

n	2	3	4	5	6	7
$c_{\mathrm{Ovac}}^{\mathrm{triangle}}$	0.286	0.250	0.211	0.185	0.162	0.146

Table 9.2: Maximum oxygen vacancy concentration $c_{\mathrm{Ovac}}^{\mathrm{triangle}}$ for single triangular pits with side length n.

$$n \text{ odd} \quad \ldots \quad \left(\sqrt{\frac{3}{4}(n+1)^2}\right) \times \left(\sqrt{\frac{3}{4}(n+1)^2}\right). \tag{9.8}$$

The vacancy concentration is then given by:

$$\begin{aligned} n \text{ even} \quad \ldots \quad c_{\mathrm{Ovac}}^{\mathrm{triangle}} &= \frac{4n}{3n^2+6n+4} \quad \text{for} \quad n>1 \\ n \text{ odd} \quad \ldots \quad c_{\mathrm{Ovac}}^{\mathrm{triangle}} &= \frac{4n}{3(n+1)^2} \quad \text{for} \quad n>2. \end{aligned} \tag{9.9}$$

It can be seen from Eq. (9.9) that the oxygen vacancy concentration decreases when the side length of the triangular reconstruction is increased. Furthermore we see from Tab. 9.2 that the preferred vacancy concentration of a quarter of the oxygen surface atoms is only reached for triangular pits with a side length of 3 oxygen atoms.
Nevertheless, a way to increase the overall oxygen vacancy concentration for larger triangular pits is the removal of additional, smaller triangular shaped reconstructions in the second double layer within the pit of the first double layer.
We adopted the sizes of the plain and nested triangles from Ref. [78] with oxygen and zinc atoms interchanged using different supercells (from (4×4) to $(\sqrt{48} \times \sqrt{48})$) and all possible sizes of triangles and nested triangles. As can be seen from Fig. 9.9 the triangular reconstructions are approximately 0.032 J/m^2 more stable than the simple vacancy configurations (black dashed lines) at similar vacancy concentrations. This at least applies above an oxygen chemical potential of -2.48 eV.

Although the triangular reconstruction with a side length of 3 and a supercell size of $(\sqrt{12} \times \sqrt{12})$ has the desired one quarter oxygen vacancy concentration, it is less stable than all other triangular reconstructions with approximately the same vacancy concentration. This indicates that larger triangles are generally preferred over small triangles, which is in agreement with the observations made on the Zn-terminated (0001) surface side [78]. The (5×5) supercell with two nested 2 and 4 sized triangles in the surface layer has an oxygen vacancy concentration of 0.24 and comes closest to the desired 25 percent of vacancy concentration. It is by far the most stable triangular reconstruction (-2.48 eV $< \Delta\mu_{\mathrm{O}} < -1.01$ eV). At more oxygen

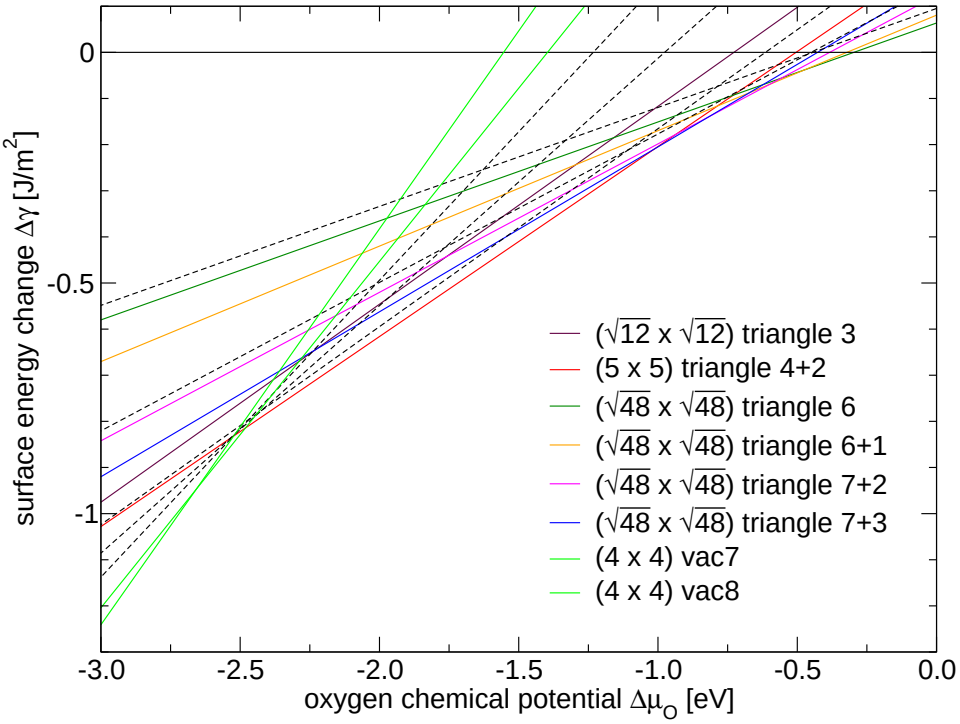


Figure 9.9: Phase diagram showing the most stable configurations including triangular shaped reconstructions in dependence of the oxygen chemical potential. The black dotted lines are the most stable phases found in Sec. 9.2. Triangular reconstructions are more stable than the previously found vacancy configurations for oxygen chemical potential greater than -2.48 eV. For smaller chemical potentials, 7 and 8 oxygen vacancy structures remain the most stable structures.

rich conditions triangular reconstructions with a smaller oxygen vacancy concentration become more stable.

In a $(\sqrt{48} \times \sqrt{48})$ supercell, 7+3 ($c_{\mathrm{Ovac}}^{\mathrm{triangle}} = 0.208$) and 7+2 ($c_{\mathrm{Ovac}}^{\mathrm{triangle}} = 0.1875$) triangles can be realized. Additionally we also considered a simple 6 sized triangle ($c_{\mathrm{Ovac}}^{\mathrm{triangle}} = 0.125$), with and without an additional single oxygen vacancy in the triangular pit ($c_{\mathrm{Ovac}}^{\mathrm{triangle}} = 0.146$) in the $(\sqrt{48} \times \sqrt{48})$ supercell. Under O poor conditions the previously found surface reconstructions involving 7 and 8 vacancies are the most stable ones, since it is not possible to model triangular reconstructions with such a high vacancy concentration.

In Fig. 9.10, we added the triangular reconstructions to the phase diagram. As can be seen, all previously found reconstruction structures involving oxygen vacancies, except the structures involving 7 and 8 oxygen vacancies, are

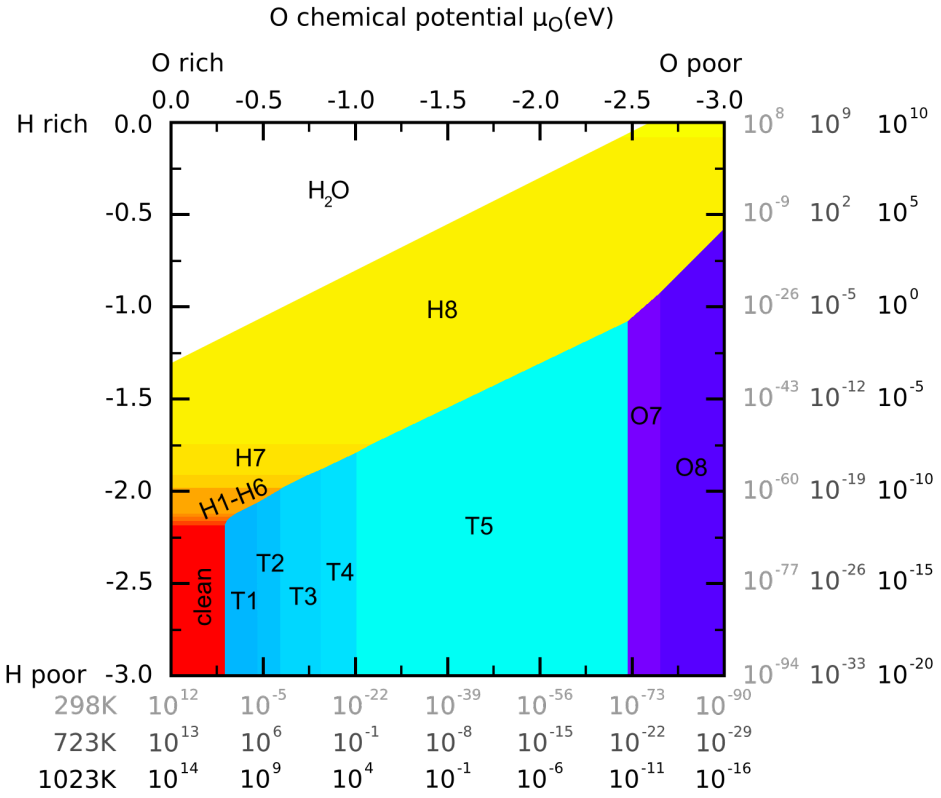
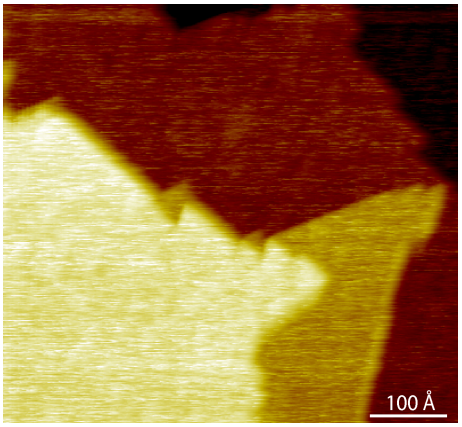


Figure 9.10: Phase diagram including plain oxygen vacancies, hydrogen adatoms, oxygen and zinc vacancies and triangular reconstructions. T1=($\sqrt{48}\times\sqrt{48}$) triangle 6, T2=($\sqrt{48}\times\sqrt{48}$) triangle 6+1, T3=($\sqrt{48}\times\sqrt{48}$) triangle 7+2, T4=($\sqrt{48}\times\sqrt{48}$) triangle 7+3, T5=(5×5) triangle 4+2.

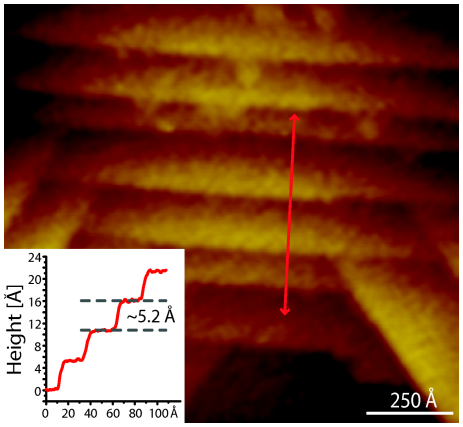
higher in energy than the triangular reconstructions. As a preliminary result, we can conclude that the triangular pits are the most stable reconstructions get considered. This is similar to the Zn-terminated surface, where the triangular pits are the most stable reconstruction type [78].

9.4.2 Hexagonal holes

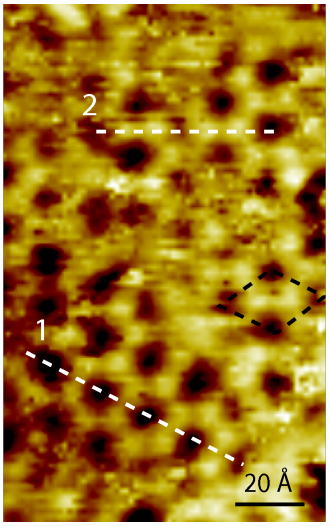
In the past, scanning tunneling microscopy (STM) was unable to reveal the atomic scale structure of the O-terminated $(000\bar{1})$ side of ZnO due to the weak conductivity of the surface at room temperature. Using non-contact atomic force microscopy (NC-AFM) in combination with integrated variable temperature STM, Lauritsen and co-workers [91] were able to investigate the surface structure with atomic resolution.
NC-AFM is able to record nano scale images of any flat surface independent



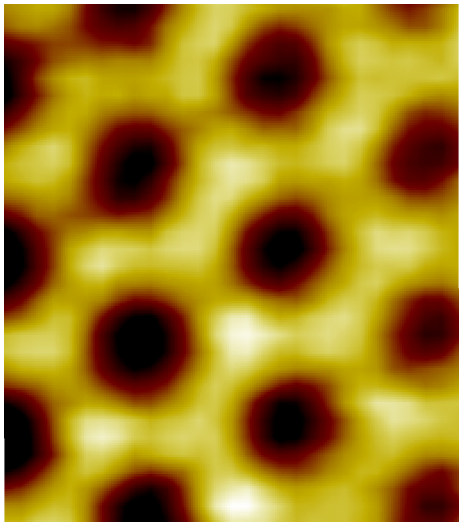
(a) STM image (600×600 Å^2) of the O-terminated ZnO $(000\bar{1})$ surface recorded at a substrate temperature of ≈ 723 K.



(b) Room temperature NC-AFM image (1200×1200 Å^2) of a region with a high step density.



(c) High resolution STM image (110×180 Å^2) of the hexagonal (5×5) reconstruction structure present on the terrace planes recorded at ≈ 723 K.



(d) Correlation averaged STM image (46×57 Å^2) of the (5×5) reconstruction.

Figure 9.11: STM images taken from Ref. [91].

of the surfaces conductivity. In Fig. 9.11(a) the STM image recorded at a substrate temperature of ≈ 720 K, shows a flat surface with rather large terraces. However, these regions are interrupted by regions with a high density of steps, as the NC-AFM image shown in Fig. 9.11(b) unravels. The majority of the step edges have a height of approximately 5.2 Å (see inset in Fig. 9.11(b)), which is almost identical to the height of two layers of ZnO. These two fold step edges are terminated by a stacking of O-terminated

$(\bar{1}010)$ edges and Zn-terminated $(10\bar{1}0)$ edges, and are therefore stoichiometric with respect to the amount of Zn and O atoms. This indicates that the surface roughness as seen in Fig. 9.11(b) does not contribute to the stabilization of the O-terminated surface. Moreover, this is in contrast to the stabilization principle observed on the opposite Zn-terminated surface, where the stabilization occurs through single-layer deep triangular pits and O-terminated step edges [78, 94, 95]. Therefore a different mechanism on the atomic scale has to be responsible for the stabilization of the polar $(000\bar{1})$-O surface.
High temperature STM images recorded at approximately 720 K reveal the presence of a reconstruction at the surface on an atomic scale (Fig. 9.11(c)). The reconstruction is characterized by a sixfold symmetry with depressions (black colour coded areas in Fig. 9.11(d)) defining the centre and the corners of the hexagon and six protrusions surrounding the centre of the depression. The lattice parameter of the unit cell is measured to be 16.3 ± 0.02 Å, which is almost exactly 5 times the length of the in plane unit cell of bulk ZnO ($a = 3.25$ Å). Therefore the observed hexagonal honeycomb like structure can be recognized as a (5×5) reconstruction of the surface. For a more detailed characterization of the reconstruction two cross-section scans, one along the in plane unit cell axis (1,0,0), the other along the [1,1,0] direction (Fig. 9.11(c)), in the STM images have been extracted (see Fig. 9.16). They reveal two characteristic properties for this particular termination of the $(000\bar{1})$ surface.

- The average depth of the depression relative to the bright regions is approximately 2.1 Å (Fig. 9.16), which coincide well with the height of a single ZnO layer ($R_1 + R_2 = 2.6$ Å in Fig. 7.1).
- The six protrusions exhibit a average height of 0.4 ± 0.1 Å relative to the region in between them.

Taken into account the experimental findings, we performed extensive DFT calculations on the (5×5) supercell for a number of different surface models, which have to be compatible with the six fold symmetry, a partial removal of the ZnO double layer, and the removal of approximately one quarter of the surface oxygen atoms for the electrostatic stabilization of the otherwise polar surface.
The most stable structure found is presented in Fig. 9.12. This surface reconstruction is modelled by removing 11 oxygen and 7 zinc atoms from the topmost layer of the $(000\bar{1})$ surface. With the remaining 14 oxygen atoms, 2 hexagons are formed, which are connected via fourfold coordinated zinc (4f) atoms. One of these hexagons remains on the regular wurtzite ZnO stacking sequence (red atoms), whereas the other hexagon terminates the surface at the atypical zinc-blende stacking sequence (orange atoms). These two hexagon are interlinked by threefold coordinated zinc (3f) atoms

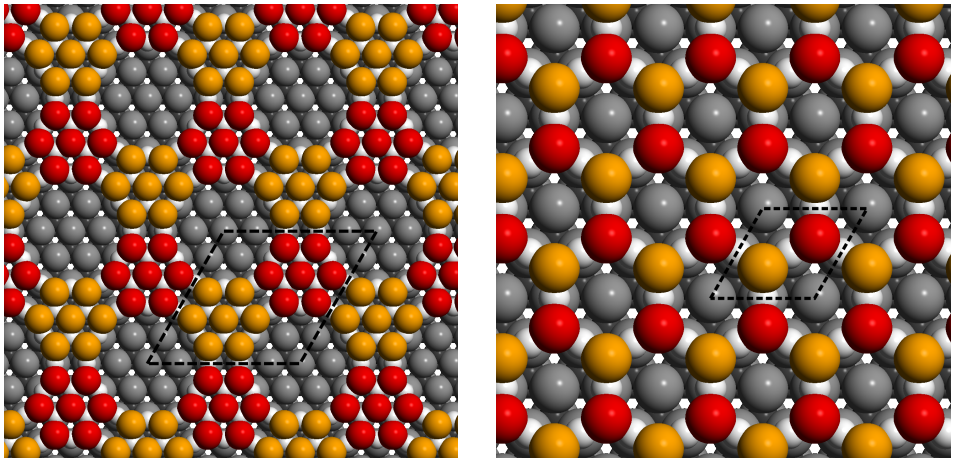


Figure 9.12: Schematic ball model of the most stable structure for the experimentally found hexagonal (5×5) supercell (left). Stoichiometric structure with a (2×2) periodicity (right), which is the most stable one at poor oxygen and hydrogen pressures. Oxygen atoms in the regular wurtzite stacking sequence are red colour coded, whereas atoms at the atypical zinc-blende stacking sequence have orange colour.

to form a honeycomb-like structure with a one layer hexagonal hole in the centre.

As already mentioned before, it is necessary to remove one quarter of the oxygen surface atoms (in the case of a (5×5) supercell this corresponds to $25/4 = 6.25 \approx 6$ atoms) to fully remove the polarity of O-terminated ZnO surface. In this reconstruction model, only 4 more oxygen than zinc atoms are removed from the surface, and thus the surface polarity is not fully compensated. The remaining partial polarity of the (5×5) structure can be compensated by the formation of hydroxy-groups (OH^-) on the surface. Up to 5 hydrogen atoms were adsorbed on the surface for compensation of the remaining 4.5 excess electrons (see Fig. 9.15).

Within our DFT calculations we also identified a stoichiometric structure with a (2×2) periodicity to be more stable than previously discussed oxygen vacancies and triangular shaped reconstructions at oxygen poor and hydrogen poor conditions. This structure is constructed along the same principles as the (5×5) structure. Two oxygen and one zinc atom are removed from the surface layer and one of the remaining oxygen atoms is moved from the wurtzite to the zinc blende stacking sequence. In contrast to the (5×5) structure the (2×2) reconstruction fully compensates for the surface polarity without the need of additional hydrogen adsorbed at the surface. All oxygen atoms are interlinked by zinc atoms, which all possess an unusual threefold coordination to the oxygen atoms.

In Fig. 9.13 we finally present the phase diagram, which is based on all

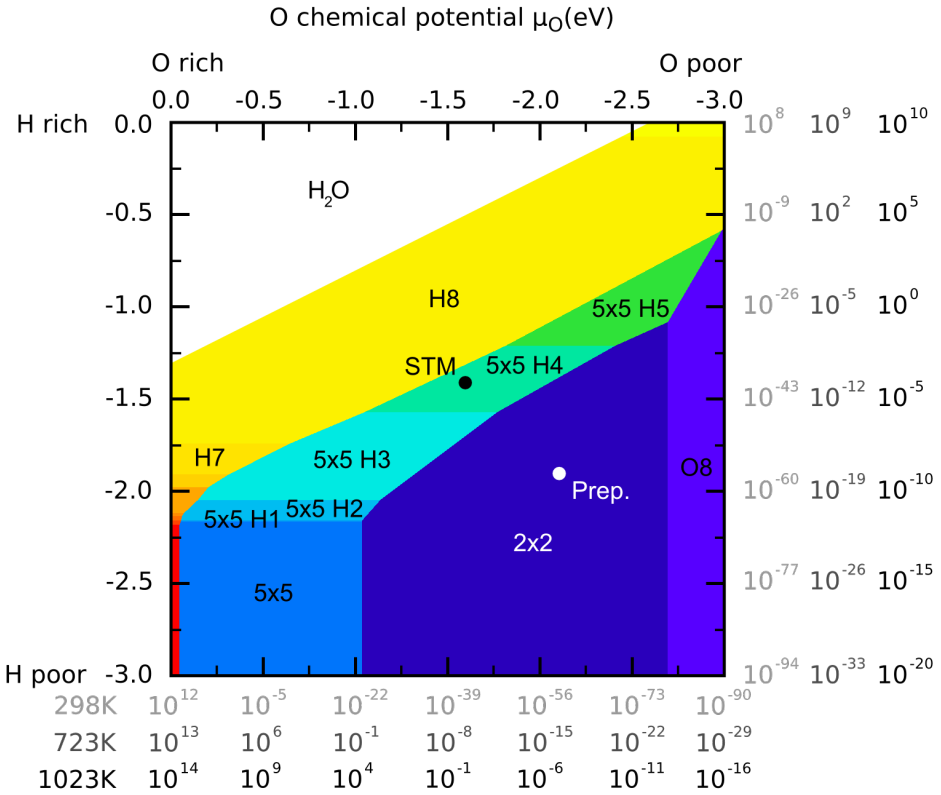


Figure 9.13: Phase diagram of all considered reconstruction types of the ZnO-$(000\bar{1})$ surface. Additionally we indicate the preparation conditions (1023 K) and the conditions where the STM image (Fig. 9.11(d)) was recorded (723 K) with black dotes.

surface terminations taken under consideration so far. In comparison to the former phase diagram, we recognize the absence of the triangular shaped vacancy structures, which are all lower in energy than the (2×2) (the large blue area) and (5×5) (light blue $c_H = 0$ to green $c_H = 5/25$ areas) surface terminations.

The essential finding of this study is that the predominant surface structures present on the two related polar (0001)-Zn and $(000\bar{1})$-O surfaces are fundamentally different. To corroborate this point we recalculated the phase diagram of the (0001)-Zn terminated surface side as given in Ref. [78] and included the new surface terminations of the $(000\bar{1})$-O side with zinc and oxygen atoms interchanged. Since we did not found any difference to the previously published phase diagram the fundamental difference between the ZnO surfaces has to be related to the specific bonding preferences of Zn and

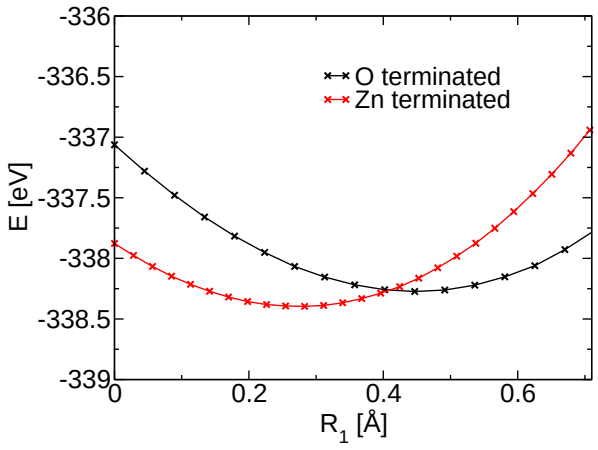


Figure 9.14: Dependence of the total energy on the height R_1 (see Fig. 7.1) of surface Zn and O atoms of a non-polar (compensated) (0001)-Zn and $(000\bar{1})$-O terminated surface, where every fourth Zn or O atom is removed from the surface, respectively. The Zn^{2+} ions are located in a more planar configuration and protrude only 0.3 Å above the second layer O^{2-} ions, whereas O^{2-} ions are located 0.5 Å above the second layer Zn^{2+} ions.

O atoms on the surfaces. Oxygen prefers a tetrahedral or pyramidal bonding configuration, in which all three O $2p$ orbitals can interact to form bonding and anti-bonding linear combinations, whereas as a $3d$ element, Zn is more flexible with respect to bond formation. In the pyramidal configuration, the out of plane d_{xz} and d_{yz} orbitals dominate, whereas for the in plane bonding, the d_{xy} and d_{x2-y2} orbitals contribute, *i.e.* the bond order varies only little between the pyramidal and planar configurations. The strength of this effect can be demonstrated by DFT calculations (see Fig. 9.14) on the total energy of (0001)-Zn and $(000\bar{1})$-O terminated surfaces as a function of the displacement of the terminating surface layer for a non-polar reconstruction. The polarity in the used (2×2) supercell was compensated by the removal of a quarter of the respective surface atoms. The clear observation is that the Zn^{2+} ions sink deeper into the surface than the O^{2-} ions. A thorough investigation of all considered structural models show a similar trend for all threefold Zn^{2+} and O^{2-} surface ions and it is summarized in Tab. 9.3.

Inspecting Tab. 9.3 we conclude that the O ions on the $(000\bar{1})$-O surface clearly prefer a pyramidal coordination, protruding typically 0.5 Å above the three neighbouring Zn ions, whereas the equivalent Zn ions on the (0001)-Zn surface are much more free to relax and typically protrude only 0.3 Å. The flexibility of Zn coordination is essential for the observed stability of the (5×5) and (2×2) structures, since 2/3 or all of the surface Zn atoms, respectively are placed in a flat configuration with three neighbouring O

	(0001) Zn-terminated			(000$\bar{1}$) O-terminated		
	wurtzite	zinc	2nd layer	wurtzite	zinc	2nd layer
(5×5) hexhol	0.33	0.35	0.28	0.53	0.56	0.48
(2×2) hexhol	0.33	0.34	0.34	0.54	0.57	0.67
(2×2) vac1	0.25			0.54		

Table 9.3: Average heights of the terminating Zn and O surface atoms over the three neighbouring O or Zn ions, respectively. All values are given in Å. "wurtzite" and "zinc" denotes the stacking sequence of the first layer for the the (2×2) and (5×5) hexagonal hole reconstructions. Average heights from the first layer inside the hexagonal holes are denoted with "2nd layer".

atoms being co-planar with the Zn atom. Hypothetical (5×5) and (2×2) structures on the Zn-terminated surface would force O into a coplanar configuration with three neighbouring Zn atoms, which is energetically very unfavourable.

9.5 Comparison between theory and experiment

9.5.1 Scanning tunnelling microscopy (STM)

For comparison with experimentally measured STM images [91], constant current STM images were simulated using the model of Tersoff and Hamann by plotting contours of the local density of states projected onto the tip axes [96]. In Fig. 9.15 we present the bare (5×5) surface termination and surfaces with up to four adsorbed hydrogen atoms. The adsorption position of the hydrogen atoms (blue balls) is indicated in the ball models underneath each STM image. The last image is the experimentally measured STM image at 450°C.

Although comparison does not allow for a unique assignment of the theoretical structures good agreement is seemingly observed for the hydroxylated surface terminations involving three or four hydrogen atoms. As already stated in Sec. 9.4.2, an amount of four hydrogen atoms in a (5×5) unit cell would almost fully compensate for the polarity of the surface. For a more detailed comparison, and to support the assignment of the (5×5)H4 termination, cross section scans, indicated by the solid and dashed line in the experimental STM image in Fig. 9.15, where performed experimentally and also extracted from the theoretical STM images.

From line scan 2 (presented in Fig. 9.16), which is along one of the unit cell axes, no further information on the structural details of the surface can be extracted. The periodicity of experiment and simulation is identical, but any information inside the hexagonal pits remains undetermined, since the

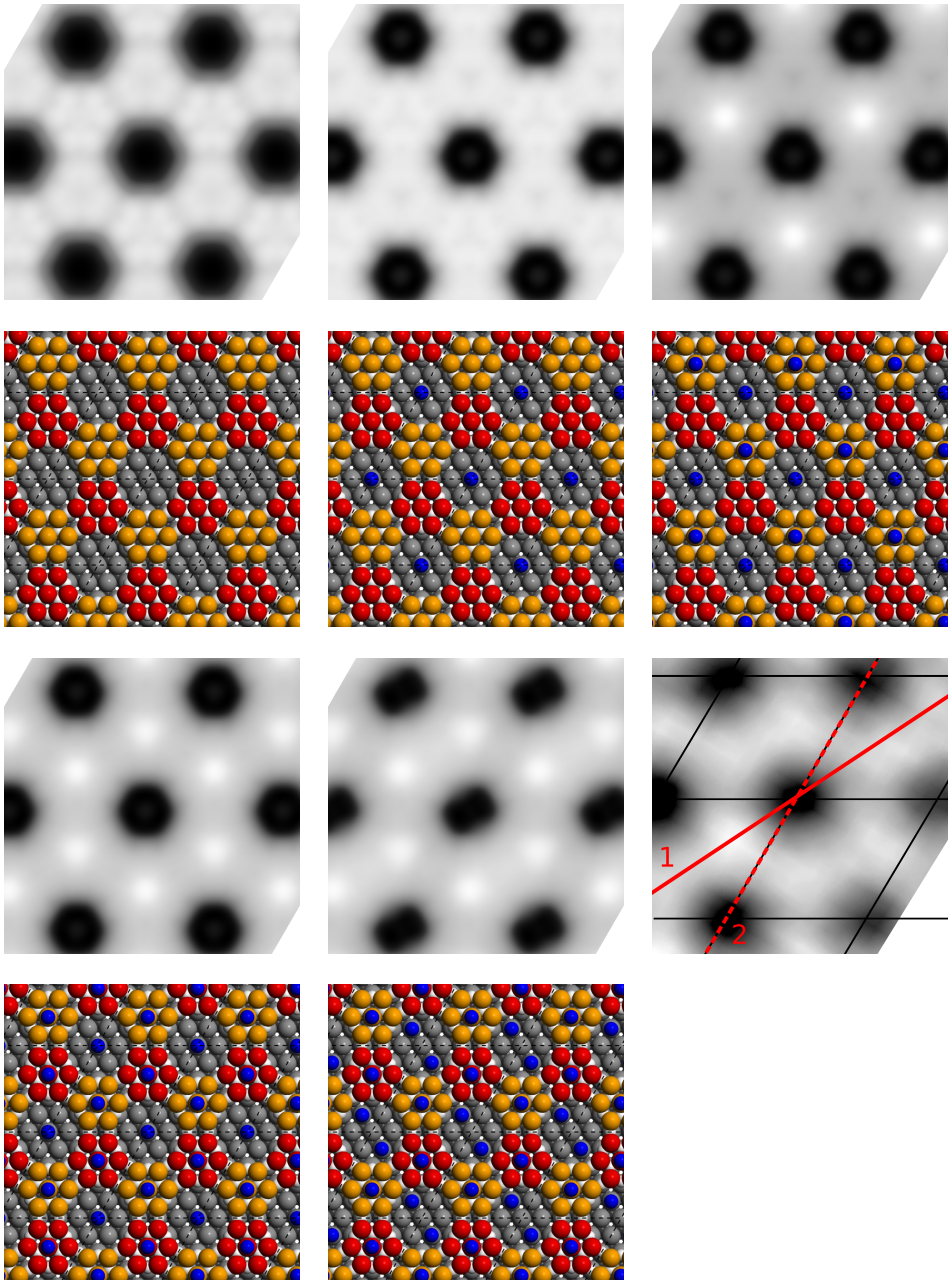


Figure 9.15: Theoretical STM images of the (5×5) surface reconstruction with different number of hydrogen atoms adsorbed at the surface. Under each STM image a ball model is given, where the red balls denote oxygen atoms at the regular wurtzite stacking sequence, orange balls highlight oxygen in the atypical zinc blende stacking sequence and blue balls are hydrogen atoms adsorbed on top of oxygen atoms. The last image with the indicated line scans is the experimental STM image given for comparison.

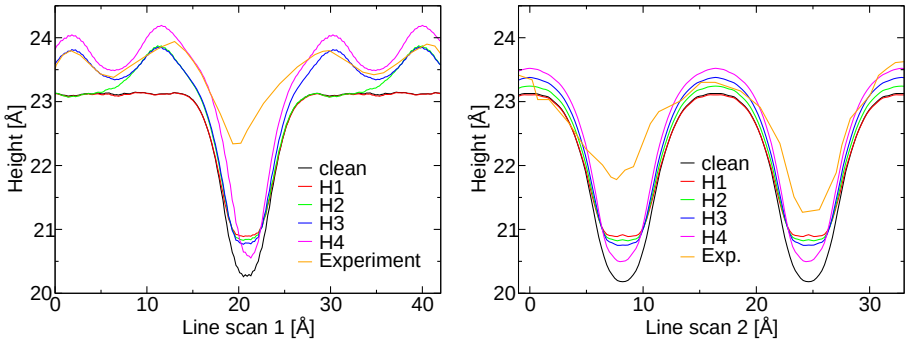


Figure 9.16: Comparison of experimental and theoretical line scans of the (5×5) structure with different hydrogen concentrations. The position of the line scans are indicated by red lines in the last picture of Fig. 9.15. The left graph is depicted as solid red line, the right graph as dashed red line.

real tip in the experiment will not be able to go as deep into the hexagonal pits as the Tersoff Hamann STM simulation suggests.
In contrast, line scan 1 offer more insight in the atomic details of the surface. The experimental line scan features four distinctive elevations to the left and right of the observed pit.
The simulated clean structure as well as the structure with one hydrogen atom placed inside the pit does not show any such feature. When placing hydrogen atoms at the centre of the hexagons forming the surface layer, protrusions show up in the simulated scan lines. Best agreement can be found with structures with a hydrogen atom in the middle of each hexagon, i.e. the H3-H5 structures respectively. Although the amount of hydrogen in the pit can not be determined by this study, it is energetically favourable that hydrogen adsorbs on the entire surface (pits and terraces) as we have seen at the beginning in Sec. 9.3.
In total, all the STM investigations corroborate the existence of the novel (5×5) surface termination at environmental conditions reported before. Although this is already a strong proof for the accurateness of our assumptions a second method will be discussed in the next section to provide further evidence.

9.5.2 Near edge X-ray adsorption fine structure (NEXAFS) spectroscopy

The fundamental phenomenon underlying near edge X-ray adsorption fine structure (NEXAFS) spectroscopy is the adsorption of X-ray radiation by atomic core electrons in solids and the resulting emission of a photo electron. The resulting core hole is then filled either via an Auger process or by capture of an electron from another shell followed by the emission of a fluorescent

photon. By measuring the emitted Auger electrons at a given energy and varying the X-ray photon energy, a NEXAFS spectrum can be obtained. As for all X-ray spectroscopy methods this spectrum is element-specific, and due to the limited range of the Auger electrons in solids the obtained signal can be assigned to the first few layers underneath the surface. Additionally NEXAFS experiments have been proven to be very sensitive to different properties like bond angles, bond lengths and coordination numbers.
The key feature to distinguish the three surface terminations within a NEXAFS experiment is the amount of threefold zinc in the surface layer. In a hydrogen rich environment, where the surface is unreconstructed but covered with half a monolayer of hydrogen (1×2)H no threefold zinc atoms are present at the surface. In the (5×5) reconstructed surface, 12 out of the 18 surface zinc atoms have a threefold coordination, whereas in the stoichiometric (2×2) surface termination all zinc atoms in the surface layer exhibit a threefold coordination.
To obtain additional evidence that the proposed (2×2) and (5×5) surface terminations are identical to the experimentally observed surface reconstruction, Lauritsen and coworkers [91] performed NEXAFS spectroscopy measurements on the L-edge of zinc at those three temperature and environmental conditions that lead to the (1×2)H, (2×2) and (5×5) surface termination.
The simplest approximation for the calculation of a NEXAFS spectrum by properties available from DFT calculations is the use of the ground state partial density of the Zn s and d states [97]. For most transition metals this approximation leads to bad results when the L-edge is used in experiment. This is caused by the so-called multiplet effect, where the p core hole can overlap with d valence holes. This effect can be rather large for open shell transition metals (TM). The zinc L-edge is special since the $3d$ band is fully occupied and no $2p$ core hole and $3d$ valence hole overlap is present [98]. Therefore multiplet effects are not important, and the NEXAFS spectrum should correspond well to the partial density of the Zn s and d states.
To acquire data from our DFT simulations we used the following procedure. For the three topmost layers, the local partial density of s and d states were determined. The layers are counted from the first layer exposed to the vacuum, *i.e.* Zn atoms inside a hexagonal hole and exposed to the vacuum belong to the first layer. To account for the final penetration depth of the experimental probe, the layers are weighted from top to bottom using a weighting coefficient of 1.0, 0.6 and 0.3 and the weighted contributions from each layer are summed up. To account for experimental broadening, the final data are folded using a Lorentzian function ($\gamma = 1.0$ eV) [98].
To determine which features in the NEXAFS spectra are fingerprints for a certain local environment, we also included the calculated spectra of the first zinc layer of the (2×2) surface termination, which exclusively contains threefold zinc, and of the bulk termination, corresponding to fourfold

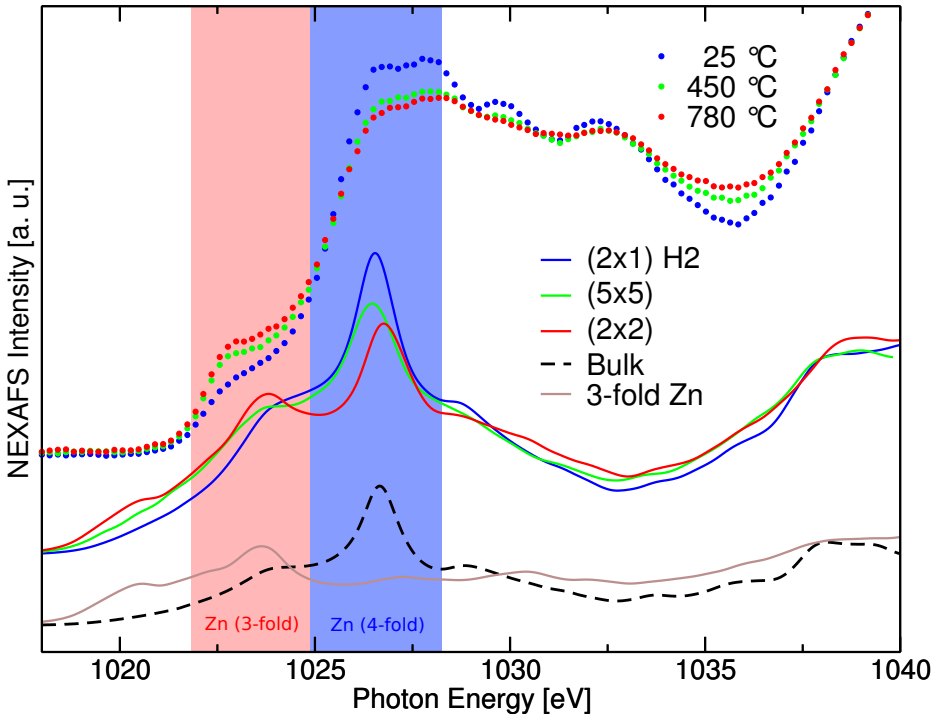


Figure 9.17: Experimental NEXAFS spectra of the ZnO $(000\bar{1})$ surface at three different temperatures (full circles) [91]. Lines indicate simulated NEXAFS spectra corresponding to the three major reconstruction types, (1×2)H, (5×5) and (2×2). For reference also NEXAFS-simulations of a hypothetical isolated layer of threefold (3f) Zn (brown solid line) and bulk (4f) ZnO (black dashed line) are included.

coordinated zinc atoms only. As can be seen in Fig. 9.17 the reference spectrum for threefold zinc (brown solid line) has the main peak at ≈ 1023 eV, whereas the 4 folded zinc spectrum (black dashed line) has a strong peak at ≈ 1027 eV. The NEXAFS spectroscopy curve at room temperature (blue dotted line in Fig. 9.17) is dominated by fourfold coordinated zinc. At 450°C a significant Zn(3f) feature develops around ≈ 1022 eV and simultaneously the main Zn(4f) peak drops considerably in intensity. This indicates the formation of the (5×5) surface termination as seen in the theoretical phase diagram at this temperature and environmental conditions. By increasing the temperature to 780°C the observed trend of the transformation of fourfold to threefold zinc is further amplified, supporting the development of the theoretically observed (2×2) surface termination. In principle, the same behaviour is observed in the theoretical calculated NEXAFS spectra, strongly corroborating the existence of the novel honeycomb like structures.

Chapter 10

Summary

Within this study we calculated the phase diagram of the $(000\bar{1})$-O terminated surface of ZnO in dependence of the partial pressures of the two main constituents of the gaseous environment in a UHV chamber, namely oxygen and hydrogen by investigating 205 different surface terminations, including hydrogen ad- and oxygen and zinc desorption. Novel honeycomb like surface terminations in (2×2) and (5×5) supercells were found to be energetically more favourable than triangular pit vacancy reconstructions, which are the predominant surface terminations at the (0001)-Zn surface [78, 79] at low temperatures. Additional STM-images in the Tersoff-Hamann framework [96] and NEXAFS-spectra were simulated and gave overall good agreement with recent experiments [91]. The stabilization of the (2×2) and (5×5) surface terminations can clearly be attributed to the more flexible bonding behaviour of zinc atoms in comparison to oxygen, making these terminations energetically unfavourable on the (0001)-Zn terminated side.

We showed in the present study in combination with recent experimental insights that the surface terminations on the two polar (0001) and $(000\bar{1})$ surfaces of ZnO are of fundamentally different type. Furthermore the existence of the honeycomb like terminations on the ZnO $(000\bar{1})$ surface provides new essential insight into the mechanism governing the cancellation of surface polarity on ZnO(0001) surfaces, and the results demonstrate that more elaborate models which include large atomic relaxations must be added to the conventional and widely applied models based on cancellation of the dipole by the partial removal of atoms from their bulk positions. Such a model should be applicable to the analysis of the atomic-scale structures and stability of other polar surfaces from the large and important group of ionic compounds (e.g. oxides, sulphides), which consist of transition metals, with a bonding flexibility similar to that of Zn, in combination with lighter anions with a strict preferential p-type bonding, such as O. The new specific atomic-scale insight into the structures of the polar ZnO surfaces has

widespread implications for the fundamental understanding of their physical and chemical properties. The design of functional ZnO nano-structures often relies on a detailed knowledge of the controlled growth on selected ZnO facets and an exact atomic-scale understanding of, e.g. adhesion on metals or other metal oxides. Additionally, the present findings demonstrate how the ZnO surface structures may dynamically respond to changes in the surrounding ambient atmosphere.

Part IV

Implementation of a minimal basis set in

Chapter 11

Introduction

The ongoing improvements in computer hardware have enabled the scientific community to push the limits of computational simulations on solids and surfaces, both in accuracy and simulation size, further and further.
With VASP it is today possible to calculate accurately the adsorption of molecules on surfaces with up to 50 atoms in the supercell (e.g. random phase approximation (RPA) within many electron methods such as the adiabatic connection fluctuation dissipation framework (ACFDT) [99] or MP2 (Möller-Plesset 2)). On the other hand it is possible to simulate systems with up to 1000 atoms within the DFT framework and the generalized gradient approximation as implemented in the VASP package.
An even larger number of atoms can only be simulated by the use of so-called order-n-methods $\mathcal{O}(n)$, where the computational time and memory usage scales linearly with the system size.
In the last decade order-n methods were implemented within the DFT framework [100], and instead of the usually used plane wave basis in combination with periodic boundary conditions, the use of an atomic centred local basis set as applied in most quantum chemical programs[1,2] has turned out to be a better choice.
We developed and implemented a self-consistent density functional method using the projector augmented wave method based on a minimal basis set of atomic orbitals.
The long term purpose is two fold. First we aim at $\mathcal{O}(n)$ calculations of the electronic properties. Second, but probably more importantly, VASP could tremendously benefit from a mixed basis set implementation, where plane waves (PW) are complemented by LCAO like functions at the atomic sites. The present work will form the basis for these future developments.

[1] SIESTA: `http://www.icmab.es/siesta/`
[2] GAUSSIAN: `http://www.gaussian.com`

Chapter 12

Basis set - the PAW atom

The first step towards a localized atomic orbital basis set is naturally the calculation of the atomic orbitals of all elements present in the system under consideration. At first glance this task looks like an easy one, since during the construction of the PAW pseudopotential, the all electron and pseudo orbitals are already calculated. But the pseudo valence orbitals are presently not stored in the final `POTCAR` file, since they are not required in the original PAW formalism.

The integration of the atomic orbitals into the `POTCAR` file seems to be a simple solution, but has two major drawbacks. First the inclusion in the `POTCAR` file can cause inconsistencies with older versions of VASP, which would be unable to read the new file format.

Secondly, and this is the main disadvantage, it is desirable to tune the atomic orbitals for example by adding a confinement potential to the original atomic potential [101] or restricting the atomic orbital to a given radius R_{cut} around the atomic centre [102]. The inclusion of this procedures in the pseudopotential generation code is not of advantage since then this program would be required to perform a minimal basis set calculation. More flexibility can be achieved by integrating these calculations directly in the VASP package.

Out of this considerations, we decided to implement the calculation of the atomic basis functions within the VASP program package. The flags `LCAO=.TRUE.` and `LREAL=A` have to be set in the `INCAR` file to initiate a minimal basis set calculation.

12.1 The radial Schrödinger equation in the PAW formalism

According to the derivations given in Appendix A.1.8 the Schrödinger equation for a single atom within the PAW formalism has the following form

$$\left[T + \tilde{V}_{\text{eff}}[n] + \sum_{i,j} |\tilde{p}_i\rangle \underbrace{\left(\hat{D}_{ij} + D^1_{ij} - \tilde{D}^1_{ij}\right)}_{D_{ij}} \langle\tilde{p}_j|\right] |\phi_{nl}\rangle = \varepsilon_n \left(\mathbf{1} + \sum_{i,j} |\tilde{p}_i\rangle Q_{ij} \langle\tilde{p}_j|\right) |\phi_{nl}\rangle \,. \quad (12.1)$$

Using the fact that the atomic orbitals as well as the projector functions are products of spherical harmonics and radial functions we can adopt the same procedures to obtain a radial Schrödinger equation as for the simple hydrogen atom.
The effective potential $\tilde{V}_{\text{eff}}$ is of spherical symmetry, the kinetic energy operator can be rewritten in terms of the angular momentum operator, and the projector operators and the orbitals are products of a spherical harmonic $Y_{lm}(\theta, \varphi)$ and a radial function depending only on r:

$$\begin{aligned} \tilde{V}_{\text{eff}}[n](\mathbf{r}) &= \tilde{V}_{\text{eff}}[n](r) \\ T &= -\frac{\hbar^2}{2mr}\frac{d^2}{dr^2}r + \frac{\mathbf{L}^2}{2mr^2} \\ \tilde{p}_i(\mathbf{r}) &= Y_{lm}(\theta, \varphi)\tilde{P}_i(r) \\ \phi_{nl}(\mathbf{r}) &= Y_{lm}(\theta, \varphi)R_{nl}(r). \end{aligned}$$

Using atomic units, used throughout this work and the fact that the spherical harmonics are eigenfunctions of the angular momentum operator, and the orthonormality relation for spherical harmonics, as given in Eq. (B.4) for the integrals between the projector functions $\langle\tilde{p}_j|$ and the atomic orbitals $|\phi_{nl}\rangle$, we can simplify Eq. (12.1) to

$$\left[-\frac{1}{r}\frac{d^2}{dr^2}r + \frac{l(l+1)}{r^2} + \tilde{V}_{\text{eff}}[n](r)\right] |R_{nl}\rangle + \sum_{i,j} |\tilde{P}_i\rangle D_{ij} \langle\tilde{P}_j|R_{nl}\rangle = \varepsilon_n \left(|R_{nl}\rangle + \sum_{i,j} |\tilde{P}_i\rangle Q_{ij} \langle\tilde{P}_j|R_{nl}\rangle\right). \quad (12.2)$$

Substituting $R_{nl}(r)$ with $\phi_{nl}(r)/r$ and $\tilde{P}_i(r)$ with $\tilde{p}_i(r)/r$ and multiplying with r we finally obtain the radial Schrödinger equation for an atom in the PAW formalism:

$$\left[-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}+\tilde{V}_{\text{eff}}[n](r)\right]|\phi_{nl}\rangle+\sum_{i,j}|\tilde{p}_i\rangle D_{ij}\langle\tilde{p}_j|\phi_{nl}\rangle = \varepsilon_n\left(|\phi_{nl}\rangle+\sum_{i,j}|\tilde{p}_i\rangle Q_{ij}\langle\tilde{p}_j|\phi_{nl}\rangle\right). \quad (12.3)$$

We introduce the following abbreviations for the non local part and the overlap operator:

$$V_{\text{NL}}|\phi_{nl}\rangle=\sum_{i,j}|\tilde{p}_i\rangle D_{ij}\langle\tilde{p}_j|\phi_{nl}\rangle \quad (12.4)$$

$$S|\phi_{nl}\rangle=|\phi_{nl}\rangle+\sum_{i,j}|\tilde{p}_i\rangle Q_{ij}\langle\tilde{p}_j|\phi_{nl}\rangle \quad (12.5)$$

and finally obtain

$$\left[-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}+\tilde{V}_{\text{eff}}[n](r)+V_{\text{NL}}\right]|\phi_{nl}\rangle=\varepsilon_n\ S|\phi_{nl}\rangle\,. \quad (12.6)$$

12.2 Solving the radial PAW Schrödinger equation

The ground state orbitals for a given potential can be found by separating Eq. (12.6) into two regions, outside and inside the PAW sphere.
Outside the PAW sphere the terms involving the projector operators are not present due to their confinement inside the sphere, and the Schrödinger equation simplifies to

$$\left[-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}+\tilde{V}_{\text{eff}}[n](r)\right]|\phi_{nl}\rangle=\varepsilon_n|\phi_{nl}\rangle. \quad (12.7)$$

This equation is solved by inwards integration on a logarithmic grid from ≈ 200 Å to the border of the PAW sphere. Inside the sphere the following scheme has to be applied due to the non local contributions.
First the homogeneous and inhomogeneous differential equations are solved by outward integration:

$$\left(T+\tilde{V}_{\text{eff}}-\epsilon\right)|\alpha_{\text{loc}}\rangle = 0 \quad (12.8)$$

$$\left(T+\tilde{V}_{\text{eff}}-\epsilon\right)|\alpha_i\rangle = |\tilde{p}_i\rangle\,. \quad (12.9)$$

The solution of the eigenvalue equation $\left(T + \tilde{V}_{\text{eff}} + V_{NL} - S\varepsilon\right)|\phi_{nl}\rangle = 0$ is then given by the Ansatz:

$$|\phi_{nl}\rangle = \sum_i A_i |\alpha_i\rangle + |\alpha_{\text{loc}}\rangle \,. \tag{12.10}$$

Inserting this in the original eigenvalue equation together with the definition of V_{NL} and S one obtains:

$$\begin{aligned}
0 &= \left(T + \tilde{V}_{\text{eff}} + V_{NL} - S\varepsilon\right)|\phi_{nl}\rangle \\
0 &= \left[T + \tilde{V}_{\text{eff}} - \varepsilon + \sum_{i,j}\left(D_{ij} - \varepsilon Q_{ij}\right)|\tilde{p}_j\rangle\langle\tilde{p}_i|\right]\left(\sum_k A_k |\alpha_k\rangle + |\alpha_{\text{loc}}\rangle\right) \\
0 &= \underbrace{\left(T + \tilde{V}_{\text{eff}} - \varepsilon\right)|\alpha_{\text{loc}}\rangle}_{=0} + \sum_{i,j}\left(D_{ij} - \varepsilon Q_{ij}\right)|\tilde{p}_j\rangle\,\langle\tilde{p}_i|\,\alpha_{\text{loc}}\rangle \\
&\quad + \sum_k A_k \underbrace{\left(T + \tilde{V}_{\text{eff}} - \varepsilon\right)|\alpha_k\rangle}_{=|\tilde{p}_k\rangle} + \sum_k \sum_{i,j}\left(D_{ij} - \varepsilon Q_{ij}\right)|\tilde{p}_j\rangle\,\langle\tilde{p}_i|\,A_k\,|\alpha_k\rangle \\
0 &= \sum_{i,j}\left(D_{ij} - \varepsilon Q_{ij}\right)|\tilde{p}_j\rangle\,\langle\tilde{p}_i|\,\alpha_{\text{loc}}\rangle \\
&\quad + \sum_k A_k \left[|\tilde{p}_k\rangle + \sum_{i,j}\left(D_{ij} - \varepsilon Q_{ij}\right)|\tilde{p}_j\rangle\,\langle\tilde{p}_i|\,\alpha_k\rangle\right].
\end{aligned} \tag{12.11}$$

Multiplying from the left with the pseudo partial wave $\langle\tilde{\phi}_j|$ and using relation (A.18) one finally obtains the following system of linear equations which can be solved easily numerically:

$$\begin{aligned}
\sum_k \underbrace{\left[\delta_{jk} + \sum_i \left(D_{ij} - \varepsilon Q_{ij}\right)\langle\tilde{p}_i|\,\alpha_k\rangle\right]}_{B_{jk}} A_k &= \underbrace{-\sum_i \left(D_{ij} - \varepsilon Q_{ij}\right)\langle\tilde{p}_i|\,\alpha_{\text{loc}}\rangle}_{C_j} \\
\sum_k B_{jk} A_k &= C_j.
\end{aligned} \tag{12.12}$$

The solution of the original problem Eq. (12.6) can be found by varying the eigenenergy so that the first derivative of the calculated wave function at the matching point (which is the radius of the PAW-sphere) is identical for inward and outward integration and the number of nodes is identical to the desired number of nodes (usually zero for the pseudized orbital).

12.3 The atomic PAW wave functions and their representation in VASP

As mentioned in the preceding section the atomic orbitals are determined on a logarithmic grid which extends up to ≈ 200 Å from the atomic centre. This has the advantage of a large number of grid points in the vicinity of the atomic centre, where the variation of the orbitals are largest, and a decreasing number of grid points when moving away from the atomic centre. As we will see later, setting up the Hamiltonian with the minimal basis as defined before on a real space 3 dimensional grid (atomic orbitals are separated in spherical harmonics times a radial part), involves the evaluation of integrals on this real space grid and naturally the introduction of a cutoff radius R_{cut} to make calculations feasible.
In VASP, the evaluation of the projector operators in real space is already implemented (`LREAL=A` flag). The routines involved in this calculations are in principle performing exactly what we need for the determination of the Hamiltonian within our minimal basis set.
In order to use adapted versions of this routines the type `nonlr_struct` found in `nonlr.F` has to be used for the representation of the atomic orbitals $|\phi_{nl}\rangle$ instead of the projector functions $|\tilde{p}_i\rangle$. Internally the `nonlr_struct` type uses an equidistant one dimensional grid to perform a spline interpolation of the values of the projectors onto any points in the three dimensional real space grid.

We implemented a revised version of the `nonlr_struct` which handles all necessary information needed, including the atomic orbitals and the kinetic energy operator multiplied with the orbitals. It turned out that the main numerical problem is the transformation of the atomic orbitals $|\phi_{nl}\rangle$ and $T|\phi_{nl}\rangle$, represented on a logarithmic grid, to the equidistant grid used in the `nonlr_struct` construct.

12.3.1 Spline interpolation

Bringing the atomic orbitals from the logarithmic to the equidistant grid can be done in different ways. The simplest one is to use a spline interpolation routine. The kinetic contribution $T|\phi_{nl}\rangle$ can be evaluated in two ways:

(i) Inherent to the spline interpolation procedure is the possibility of evaluating the second derivative and consequently $T|\phi_{nl}\rangle$ at any point.

(ii) Another option is to rewrite the Schrödinger equation so that $T|\phi_{nl}\rangle$ is on one side of the equation alone:

$$T|\phi_{nl}\rangle = \left(S\varepsilon - \tilde{V}_{\text{eff}} - V_{NL}\right)|\phi_{nl}\rangle\,. \tag{12.13}$$

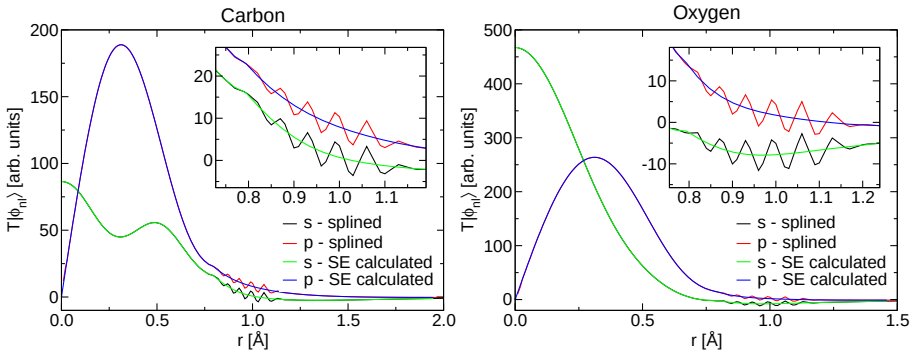


Figure 12.1: $T\,|\phi_{nl}\rangle$ of carbon and oxygen s and p orbital calculated by using the second derivative from the spline interpolation routine and by rewriting the Schrödinger equation.

A comparison between both methods for obtaining $T\,|\phi_{nl}\rangle$ for the carbon and oxygen s and p orbitals is given in Fig. 12.1.
Unfortunately, as one can see in Fig. 12.1 and in the insets, it turned out that using the second derivative results in undesired noise due to discontinuities in higher derivatives of the orbitals at the matching point, where the projector functions become zero.
The alternative way is to rewriting the Schrödinger equation as given in Eq. (12.13). To this end we have to spline all required quantities, the potential, the projectors and the orbitals to the linear grid. The overlap integrals between the projectors and wave functions can still be calculated on the logarithmic grid. Summing all contributions up properly leads to the smoother curves in Fig. 12.1. This is clearly the numerically desired procedure.

Although the second method gives us a good description of $T\,|\phi_{nl}\rangle$ on the linear grid , we still have to face few problems. The subroutine implemented already in VASP for the conversion from the one dimensional to the three dimensional grid uses internally a spline routine, too. Hence the oscillation visible in the second derivative $T\,|\phi_{nl}\rangle$ will inadvertently affect the calculation of $|\phi_{nl}\rangle$ on the grid (e.g. by introducing undesirable noise).

Another problem is that the grid spacing on the real space grid is exceedingly coarse and smooth orbitals would definitely be desirable. Additionally the orbitals $|\phi_{nl}\rangle$ and $T\,|\phi_{nl}\rangle$ we introduced here, inside the atomic sphere with radius R_{cut}, have non zero value at the sphere boundary as well as a non vanishing first derivative.
This obviously is a disadvantage. Consider for example a dimer: when the distance between the two atoms is decreased so that the spheres begin to overlap, one will see an abrupt decrease in energy and not a smooth variation

as desired. Therefore orbitals $|\phi_{nl}\rangle$ and $T\,|\phi_{nl}\rangle$ should smoothly become zero at some present radius R_{cut}.
To overcome all these problems we constructed a method, which approximates the numerically found orbital by one which has all the desired properties: zero value at R_{cut}, infinitely many continuous derivatives, etc..

12.3.2 Expansion in Bessel functions

We are free to choose "well suited" basis functions for the LCAO-type calculations. In analogy to plane waves we expanded the atomic pseudoorbitals in linear combinations of spherical Bessel functions (LCSBF):

$$\chi_n^l(r) = \sum_{i=1}^{n+1} A_i^n \; j_l(q_i r), \quad \text{with } n \leq N, \tag{12.14}$$

where N is the maximum number of basis functions. N is determined by requiring that the kinetic energy of the maximum q_n is smaller than the cutoff energy E_{cut}

$$q_i^2/2 < E_{\text{cut}}, \quad \text{for } i = 1, ..., N. \tag{12.15}$$

The careful choice of an appropriate value for E_{cut} results in very smooth orbitals as we will see later in Sec. 12.3.4.
Bessel functions have the advantage to be eigenfunctions of the kinetic energy operator T (again in analogy to PW's)

$$\left(-\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr} + \frac{l(l+1)}{r^2}\right) j_l(q_i r) = q_i^2 \; j_l(q_i r), \tag{12.16}$$

so that $T|\chi_m^l\rangle$ can be easily calculated as:

$$T|\chi_n^l\rangle = \sum_{i=1}^{n+1} A_i^n \; q_i^2 \; j_l(q_i r). \tag{12.17}$$

The q_i are defined such that $j_l(q_i R_{\text{cut}}) = 0$. Therefore our basis functions $\chi_n^l(r)$ have the value zero at the cutoff radius.
The coefficients A_i^n are determined by the constraint we have to impose for getting a advantageous basis. The constraints are as follows:

- The first derivative at R_{cut} has to be zero: ${\chi_n^l}'(R_{\text{cut}}) = 0$
- Normalization and orthogonality: $\int dr \chi_m^l(r)\chi_n^l(r) = \delta_{mn}$.

To realize these particular constraints, the construction of the n^{th} basis function start with the following Ansatz:

$$\chi_n^l = A_n^n \; j_l^n + A_{n+1}^n \; j_l^{n+1} \qquad n = 1, ..., N. \tag{12.18}$$

The radial dependence of the Bessel and basis functions are omitted for improving readability $j_l(q_n r) \to j_l^n$. A_n^n is set to 1 and A_{n+1}^n is set so that the first derivative is zero at R_{cut}:

$$A_{n+1}^n = -\frac{j_l{}'(q_n R_{\text{cut}})}{j_l{}'(q_{n+1} R_{\text{cut}})}. \tag{12.19}$$

To ensure orthogonality between the n^{th} basis function to all other $n-1$ basis functions a Gram-Schmidt orthogonalization is performed.

$$\left|\chi_n^l\right\rangle \leftarrow \left|\chi_n^l\right\rangle - \sum_{m=1}^{n-1} \underbrace{\left\langle \chi_m^l | \chi_n^l \right\rangle}_{C_{m,n}} \left|\chi_m^l\right\rangle \tag{12.20}$$

Finally the basis function is numerically normalized

$$\left|\chi_n^l\right\rangle \leftarrow \frac{\left|\chi_n^l\right\rangle}{\left\langle \chi_n^l | \chi_n^l \right\rangle} \tag{12.21}$$

and it can be written in the form of Eq. (12.14):

$$\chi_n^l(r) = \sum_{i=1}^{n+1} A_i^n \; j_l^i, \quad \text{with } n \le N, \tag{12.22}$$

The coefficients A_i^n can explicitly be calculated:

$$\begin{aligned} A_{n+1}^n &= A_{n+1}^n \\ A_n^n &= \begin{cases} A_1^1 & n = 1 \\ -C_{n-1,n} A_n^{n-1} + A_n^n & n > 1 \end{cases} \\ A_i^n &= \sum_{j=1}^{n-1} -C_{j,n} A_i^j \qquad\qquad i < n. \end{aligned} \tag{12.23}$$

Using these relations for the coefficients of the Bessel functions the numerical implementation and calculation of the basis functions is straight forward (see Fig. 12.2).

The calculated χ_n^l are the basis functions for our LCSBF approach. A linear combination of the χ_n^l's are used to obtain an "atomic" orbital for the later LCAO-type calculation within VASP.

$$\begin{aligned} |\bar{\phi}_{nl}\rangle &= \sum_{i=1}^{N} B_i |\chi_i^l\rangle \\ &= \sum_{i=1}^{N} B_i \sum_{j=1}^{i+1} A_j^i \; j_l(q_j r) \end{aligned} \tag{12.24}$$

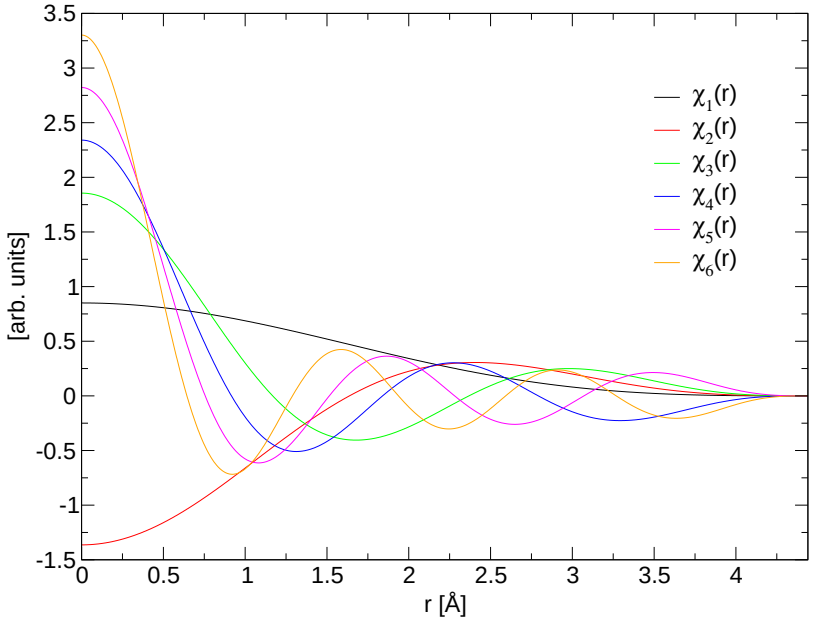


Figure 12.2: First six LCSBF for s orbital in carbon. The cutoff radius is 4.44 Å.

The calculation is performed on a equidistant linear grid for the B_is, which has the same number of grid points as required by VASP. This has the advantage that we do not have to spline the orbitals $|\bar{\phi}_{nl}\rangle$ and $T|\bar{\phi}_{nl}\rangle$ to the VASP support grid again, but can directly copy them.
To obtain the B_i we performed a single shot evaluation of the Schrödinger equation in the LCSBF basis using the ground state potential, which was splined to the linear grid. This leads to a matrix equation of the form

$$\begin{pmatrix} H_{11} & \cdots & H_{1N} \\ \vdots & \ddots & \vdots \\ H_{N1} & \cdots & H_{NN} \end{pmatrix} \cdot \begin{pmatrix} B_1 \\ \vdots \\ B_N \end{pmatrix} = E \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \cdot \begin{pmatrix} B_1 \\ \vdots \\ B_N \end{pmatrix}, \quad (12.25)$$

with $H_{ij} = \langle \chi_i^l | H | \chi_j^l \rangle$ and $S_{ij} = \delta_{ij}$, which can be solved easily by LAPACK routines.
Once we have determined the coefficients B_i for the orthogonal basis we can perform a simple transformation to the conventional basis of Bessel functions using

$$D_j = \sum_{i=k}^{N} B_i A_j^i \qquad \left| \begin{array}{ll} k = 1 & \text{if } j = 1 \\ k = j-1 & \text{if } j \geq 2 \end{array} \right. \quad (12.26)$$

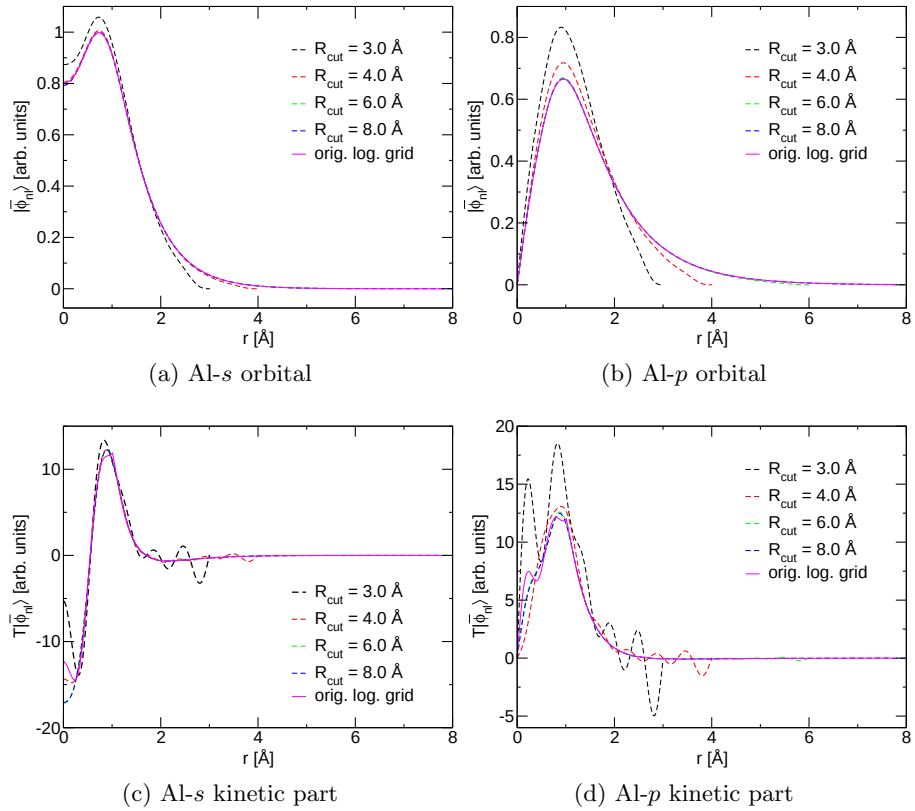
and obtain for the “atomic” orbital:

$$|\bar{\phi}_{nl}\rangle = \sum_{j=1}^{N+1} D_j \; j_l(q_j r). \tag{12.27}$$

The results of the expansion in Bessel functions depend only on two parameters, namely the radius for the sphere R_{cut} in which the atomic orbital is non zero and the energy E_{cut} which restricts the number of Bessel functions.

12.3.3 Impact of the cutoff radius R_{cut}

With the parameter R_{cut} one can restrict the spatial extension of the orbital under consideration to a given sphere with the radius R_{cut}. In Fig. 12.3 we can see the effect of the cutoff radius on the orbitals for aluminium calculated with an energy cutoff E_{cut} of 240.3 eV, which is equal to the energy cutoff supplied with the pseudopotential. A small cutoff radius, here for aluminium



(a) Al-s orbital (b) Al-p orbital

(c) Al-s kinetic part (d) Al-p kinetic part

Figure 12.3: Atomic orbitals $|\bar{\phi}_{nl}\rangle$ and kinetic part $T\,|\bar{\phi}_{nl}\rangle$ using Bessel basis functions for the s and p orbitals of aluminium with different cutoff radii R_{cut}. The original, on the logarithmic grid calculated counterparts, are also given for comparison.

around 3 to 4 Å, leads to more localized orbitals (see [102]), as can be seen in Fig. 12.3(a) and 12.3(b). A larger cutoff leads to orbitals that are almost identical to the original orbitals. The impact of the cutoff radius can be easier seen in the kinetic term, presented in Fig. 12.3(c) and 12.3(d). The spatial confinement leads to strong oscillations in the kinetic part, especially near the cutoff radius. This is not desirable and we will focus on this issue in the next section.

12.3.4 Impact of the cutoff energy E_{cut} and penalty for the kinetic energy operator

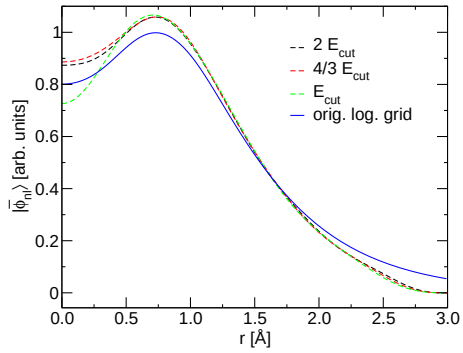
The cutoff energy E_{cut}^{PW} within an ordinary plane wave VASP calculation can either be set via the `ENCUT` tag in the `INCAR` file, or as default behaviour the energy supplied with the pseudopotential is used. Since this cutoff energy E_{cut}^{PW} is optimized for a plane wave calculation we have to investigate and relate the value of the introduced E_{cut} to the default E_{cut}^{PW}.
We recall that the cutoff energy E_{cut} determines the number of Bessel functions used for approximating the atomic orbitals $|\bar{\phi}_{nl}\rangle$ within the framework presented so far. In Fig. 12.4 we used three different energies for E_{cut}, namely E_{cut}^{PW}, 4/3 E_{cut}^{PW} and 2 E_{cut}^{PW}, for the construction of the atomic orbitals. For aluminium E_{cut}^{PW} was chosen to be the default value of 240 eV supplied within the `POTCAR` file. As one can see the difference in the orbital $|\bar{\phi}_{nl}\rangle$ and in the kinetic energy operator applied to the orbital $T|\bar{\phi}_{nl}\rangle$ between 2 E_{cut}^{PW} and 4/3 E_{cut}^{PW} is fairly small when compared to the single pseudopotential cutoff energy E_{cut}^{PW}. Therefore the cutoff energy is set to

$$E_{\text{cut}} = \frac{4}{3} E_{\text{cut}}^{PW} . \tag{12.28}$$

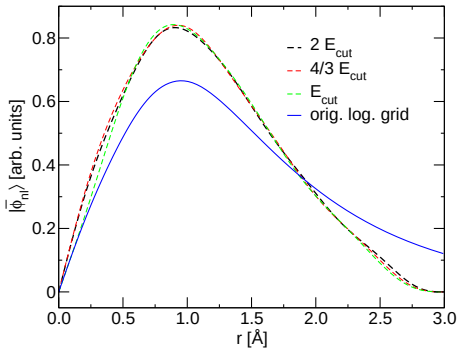
The oscillations in the numerical kinetic energy term can be reduced by introducing a new kinetic energy operator T' where a penalty function is added. The penalty function "punishes" Bessel functions with a larger q value (high Fourier components) and the contributions of these basis functions are lowered so that the influence of their oscillatory behaviour is decreased.

$$T' = \begin{cases} q_n^2 & \text{if } q_n^2 \leq E_{\text{cut}} \\ q_n^2 \cdot \left(1 + \tan\left(\frac{q_n^2 - E_{\text{cut}}}{E_{\text{cut}}} \frac{\pi}{2}\right)\right) & \text{if } E_{\text{cut}} < q_n^2 < 2\, E_{\text{cut}} \end{cases} \tag{12.29}$$

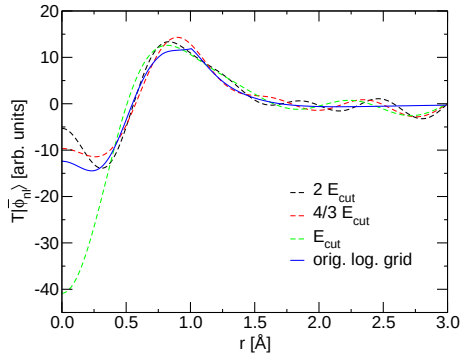
In Fig. 12.4(e) we see a comparison between the calculated kinetic energy with and without the penalty function. In Fig. 12.4(f) we see a comparison of the calculated kinetic term with and without the penalty function. It is clearly observed that the oscillating behaviour is decreased and numerically advantageous orbitals are obtained in this manner.



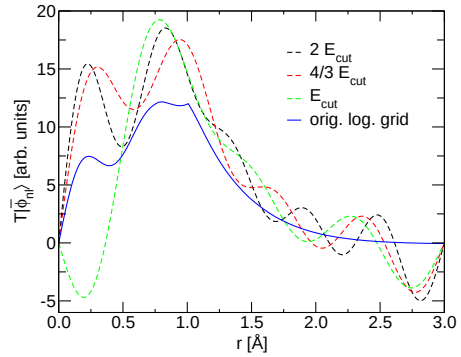
(a) Al-s orbital



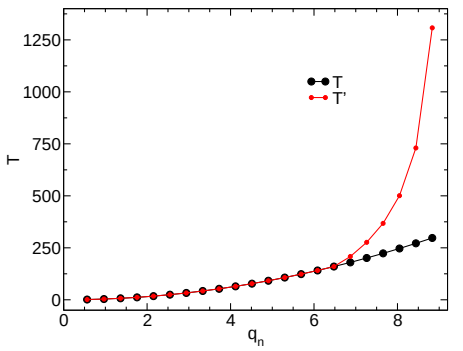
(b) Al-p orbital



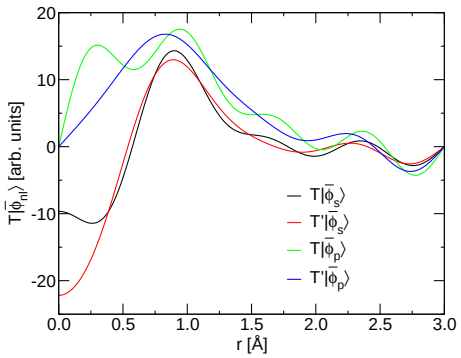
(c) Al-s kinetic part



(d) Al-p kinetic part



(e) Kinetic energy operator as defined in Eq. (12.29) for the p-atomic orbital of Silicon with a cutoff energy of $E_{\text{cut}} \approx 327$ eV.



(f) Comparison between the original and the kinetic operator including the penalty function acting on the s and p atomic orbitals using Bessel basis functions of aluminium.

Figure 12.4: (a)-(d): Atomic orbitals $\left|\bar{\phi}_{nl}\right\rangle$ and kinetic part $T\left|\bar{\phi}_{nl}\right\rangle$ for the s and p orbitals of aluminium with different energy cutoff E_{cut} using Bessel basis functions. The original, on the logarithmic grid calculated counterparts, are also given for comparison. (e)-(f): Penalty kinetic energy operator T' and its impact on the kinetic part of the atomic orbital within the expansion in Bessel functions.

Chapter 13

Linear combination of atomic orbitals

13.1 Minimal basis set

The linear combination of atomic orbitals method (LCAO), in solid state physic also sometimes termed Bloch or tight binding (TB) method, is one of the oldest methods for solving the periodic potential problem in condensed matter physics. It was originally proposed by Bloch [103] in 1928.

In the TB approach, the single particle eigenstates $|\psi_n\rangle$ are expanded into a suitable set of atomic orbitals $\phi_{lm,i}(\mathbf{r}-\mathbf{R_i})$ localized at point $\mathbf{R}_i$ with expansion coefficients $a^n_{lm,i}$. The atomic orbitals are given by a radial function $\bar{\phi}_{l,i}(r)$ and a real-valued spherical harmonic $Y_{lm}(\Omega_\mathbf{r})$:

$$\phi_{lm,i}(\mathbf{r}) = Y_{lm}(\Omega_\mathbf{r})\bar{\phi}_{l,i}(r). \tag{13.1}$$

For a finite system the expansion in atomic orbitals is given by a simple linear combination of all orbitals. The index i runs over all atoms in the system under consideration:

$$\psi_n(\mathbf{r}) = \sum_i \sum_{lm} a^n_{lm,i}\ \phi_{lm,i}(\mathbf{r}-\mathbf{R}_i). \tag{13.2}$$

Using the fact that we carry out our calculations in an infinitely periodic crystal, where the crystal can be built up from a repeating unit cell, we can decompose the position of the ions $\mathbf{R}_i$ in vector $\mathbf{R}$ which indexes all unit cells and a vector $\mathbf{R}'_i$ which determines the position of the atoms in the unit cell.

$$\psi_n(\mathbf{r}) = \sum_i \sum_{lm} a^n_{lm,i} \sum_\mathbf{R} \phi_{lm,i}(\mathbf{r}-(\mathbf{R}+\mathbf{R}'_i)) \tag{13.3}$$

We will drop the index $'$ from here on, and i and $\mathbf{R}_i$ will simply index the atoms in the unit cell. A drawback of this form is that it fulfils the Bloch Theorem only for $\mathbf{k} = \mathbf{0}$.

$$\begin{aligned}\psi_n(\mathbf{r}+\mathbf{R}_0) &= \sum_i \sum_{lm} a^n_{lm,i} \sum_{\mathbf{R}} \phi_{lm,i}(\mathbf{r}+\mathbf{R}_0-(\mathbf{R}+\mathbf{R}_i) \\ &= \sum_i \sum_{\nu} a^n_{lm,i} \sum_{\mathbf{R}'} \phi_{lm,i}(\mathbf{r}-(\mathbf{R}'+\mathbf{R}_i) \\ &= e^{i0\cdot\mathbf{R}_0}\ \psi_n(\mathbf{r}) \end{aligned} \tag{13.4}$$

Therefore it is more convenient to introduce a Bloch phase factor and apply the following ansatz for the one electron orbitals

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_i \sum_{lm} a^{n\mathbf{k}}_{lm,i} \underbrace{\frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}(\mathbf{R}+\mathbf{R}_i)}\ \phi_{lm,i}(\mathbf{r}-(\mathbf{R}+\mathbf{R}_i))}_{\beta^{\mathbf{k}}_{lm,i}(\mathbf{r}-\mathbf{R}_i)}, \tag{13.5}$$

where N is the number of unit cells summed up by $\sum_{\mathbf{R}}$. The atomic orbitals are multiplied by a phase factor $e^{i\mathbf{k}(\mathbf{R}+\mathbf{R}_i)}$ and consequently the coefficients $a^n_{lm,i}$ become $\mathbf{k}$-point depended, which is denoted in the superscript of $a^{n\mathbf{k}}_{lm,i}$. This form fulfils the Bloch Theorem:

$$\begin{aligned}\psi_{n\mathbf{k}}(\mathbf{r}+\mathbf{R}_0) &= \sum_{\mathbf{R}} \sum_i e^{i\mathbf{k}(\mathbf{R}+\mathbf{R}_i)} \frac{1}{\sqrt{N}} \sum_{lm} a^{n\mathbf{k}}_{lm,i}\ \phi_{lm,i}(\mathbf{r}+\mathbf{R}_0-\mathbf{R}-\mathbf{R}_i) \\ &= \sum_{\mathbf{R}} \sum_i e^{i\mathbf{k}(\mathbf{R}+\mathbf{R}_i)} \frac{1}{\sqrt{N}} \sum_{lm} a^{n\mathbf{k}}_{lm,i}\ \phi_{lm,i}(\mathbf{r}-\mathbf{R}_i-\underbrace{(\mathbf{R}-\mathbf{R}_0)}_{\mathbf{R}'}) \\ &= \sum_{\mathbf{R}'} \sum_i e^{i\mathbf{k}(\mathbf{R}'+\mathbf{R}_0+\mathbf{R}_i)} \frac{1}{\sqrt{N}} \sum_{lm} a^{n\mathbf{k}}_{lm,i}\ \phi_{lm,i}(\mathbf{r}-\mathbf{R}_i-\mathbf{R}') \\ &= e^{i\mathbf{k}\mathbf{R}_0} \sum_{\mathbf{R}'} \sum_i e^{i\mathbf{k}(\mathbf{R}'+\mathbf{R}_i)} \frac{1}{\sqrt{N}} \sum_{lm} a^{n\mathbf{k}}_{lm,i}\ \phi_{lm,i}(\mathbf{r}-\mathbf{R}_i-\mathbf{R}') \\ &= e^{i\mathbf{k}\mathbf{R}_0}\ \psi_{n\mathbf{k}}(\mathbf{r}). \end{aligned} \tag{13.6}$$

An alternative form of the Bloch theorem states, that the eigenstates of a one-electron Hamiltonian with a potential periodic with the Bravais lattice, can be written as the Bloch phase factor times a function with the periodicity of the potential.

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \tag{13.7}$$

In VASP this form is used for the orbital, where $u_{n\mathbf{k}}(\mathbf{r})$ has the same periodicity as the Bravais lattice. Expressing Eq. (13.5) in the form of Eq. (13.7) yields for the cell periodic part

$$\begin{aligned} u_{n\mathbf{k}}(\mathbf{r}) &= e^{-i\mathbf{k}\mathbf{r}} \sum_i \sum_{lm} a_{lm,i}^{n\mathbf{k}} \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}(\mathbf{R}+\mathbf{R}_i))} \phi_{lm,i}(\mathbf{r} - (\mathbf{R}+\mathbf{R}_i)) \\ &= \sum_i \sum_{lm} a_{lm,i}^{n\mathbf{k}} \underbrace{\frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{-i\mathbf{k}(\mathbf{r}-(\mathbf{R}+\mathbf{R}_i))} \phi_{lm,i}(\mathbf{r} - (\mathbf{R}+\mathbf{R}_i))}_{\bar{\beta}_{lm,i}^{\mathbf{k}}(\mathbf{r}-\mathbf{R}_i)} \\ &= \sum_i \sum_{lm} a_{lm,i}^{n\mathbf{k}} \; \bar{\beta}_{lm,i}^{\mathbf{k}}(\mathbf{r}-\mathbf{R}_i) \end{aligned} \tag{13.8}$$

where $\bar{\beta}_{lm,i}^{\mathbf{k}}(\mathbf{r})$ is the Bloch periodic part of the basis functions (Bloch basis function).

13.2 Generalized eigenvalue equation

The Hamilton operator within the PAW formalism is calculated in the Appendix (see Sec. A.1.8, Eq. A.69). The one particle Schrödinger equation has the following form

$$\begin{aligned} H\,|\psi_{n\mathbf{k}}\rangle &= \epsilon_{n\mathbf{k}} S\,|\psi_{n\mathbf{k}}\rangle \\ \left[T + \tilde{V}_{\text{eff}} + \sum_{i,j} |\tilde{p}_i\rangle D_{ij} \langle \tilde{p}_j|\right] |\psi_{n\mathbf{k}}\rangle &= \epsilon_{n\mathbf{k}} \left(\mathbf{1} + \sum_{i,j} |\tilde{p}_i\rangle Q_{ij} \langle \tilde{p}_j|\right) |\psi_{n\mathbf{k}}\rangle \end{aligned} \tag{13.9}$$

where T is the kinetic energy. $\tilde{V}_{\text{eff}}$ is the local potential, where Hartree and exchange and correlation potential are summed up, and the last term is the non local part of the potential.

To obtained a closed matrix equation, Eq. 13.9 is multiplied by $|\beta_{\nu,i}^{\mathbf{k}}\rangle$ from the left. It is convenient to describe the resulting matrix elements in terms of the Bloch periodic functions $|\bar{\beta}_{\nu,i}^{\mathbf{k}}\rangle$, because then the calculations can be restricted to the first primitive cell. ν and μ are abbreviations for the combination of the l and m quantum numbers.

$$\begin{aligned}
\left\langle \beta^{\mathbf{k}}_{\mu,j} \middle| H \middle| \psi_{n\mathbf{k}} \right\rangle &= \sum_i \sum_\nu a^{n\mathbf{k}}_{\nu,i} \left\langle \beta^{\mathbf{k}}_{\mu,j} \middle| H \middle| \beta^{\mathbf{k}}_{\nu,i} \right\rangle \\
&= \sum_i \sum_\nu a^{n\mathbf{k}}_{\nu,i} \left\langle \beta^{\mathbf{k}}_{\mu,j} \middle| H \middle| e^{i\mathbf{kr}} e^{-i\mathbf{kr}} \beta^{\mathbf{k}}_{\nu,i} \right\rangle \\
&= \sum_i \sum_\nu a^{n\mathbf{k}}_{\nu,i} \left\langle \beta^{\mathbf{k}}_{\mu,j} e^{i\mathbf{kr}} \middle| \bar{H} \middle| e^{-i\mathbf{kr}} \beta^{\mathbf{k}}_{\nu,i} \right\rangle \\
&= \sum_i \sum_\nu a^{n\mathbf{k}}_{\nu,i} \underbrace{\left\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} \middle| \bar{H} \middle| \bar{\beta}^{\mathbf{k}}_{\nu,i} \right\rangle}_{\bar{H}^{\mathbf{k}}_{\mu\nu,ij}}
\end{aligned} \tag{13.10}$$

From line 2 to 3 in Eq. (13.10) the kinetic energy operator T changes:

$$\begin{aligned}
T|\beta^{\mathbf{k}}_{\nu,i}\rangle &= -\frac{\hbar^2}{2m}\nabla^2 e^{i\mathbf{kr}} \bar{\beta}^{\mathbf{k}}_{\nu,i}(\mathbf{r}) \\
&= e^{i\mathbf{kr}} \frac{\hbar^2}{2m} \left(-\nabla^2 - 2ik + k^2\right) \bar{\beta}^{\mathbf{k}}_{\nu,i}(\mathbf{r}) \\
&= e^{i\mathbf{kr}} \underbrace{\frac{\hbar^2}{2m} \left(\frac{1}{i}\nabla + \mathbf{k}\right)^2}_{\bar{T}} \bar{\beta}^{\mathbf{k}}_{\nu,i}(\mathbf{r}).
\end{aligned} \tag{13.11}$$

which we denote by $\bar{T}$ and $\bar{H}$. The basis function, which we multiplied from the left can be combined with the phase factor to form the cell periodic function $|\bar{\beta}^{\mathbf{k}}_{\mu,j}\rangle$. The same operation can be performed for the overlap operator S, which does not change S ($[S, \mathbf{k}] = 0$). We finally obtain the generalized eigenvalue equation:

$$\left[\bar{H}^{\mathbf{k}}_{\mu\nu,ij} - \epsilon_{n\mathbf{k}} S^{\mathbf{k}}_{\mu\nu,ij}\right] a^{n\mathbf{k}}_{\nu,i} = 0. \tag{13.12}$$

13.3 Calculation of the matrix elements

Using Eq. (13.12) and the Hamilton operator in the PAW formalism (see Eq. (13.9)) we can explicitly write the different contributions to the matrices in the Hamilton as well as the overlap operator:

$$\begin{aligned}
\bar{H}^{\mathbf{k}}_{\mu\nu,ij} = &\left\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} \middle| \bar{T} \middle| \bar{\beta}^{\mathbf{k}}_{\nu,i} \right\rangle + \left\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} \middle| \tilde{V}_{\text{eff}} \middle| \bar{\beta}^{\mathbf{k}}_{\nu,i} \right\rangle \\
&+ \sum_l \sum_{\mu',\nu'} \left\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} \middle| \tilde{p}^{\mathbf{k}}_{\mu',l} \right\rangle D_{\mu'\nu',l} \left\langle \tilde{p}^{\mathbf{k}}_{\nu',l} \middle| \bar{\beta}^{\mathbf{k}}_{\nu,i} \right\rangle
\end{aligned} \tag{13.13}$$

$$S^{\mathbf{k}}_{\mu\nu,ij} = \left\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} \middle| \left(1 + \sum_l \sum_{\mu',\nu'} \left| \tilde{p}^{\mathbf{k}}_{\mu',l} \right\rangle Q_{\mu'\nu',l} \left\langle \tilde{p}^{\mathbf{k}}_{\nu',l} \right| \right) \middle| \bar{\beta}^{\mathbf{k}}_{\nu,i} \right\rangle. \tag{13.14}$$

Note that

$$\tilde{p}^{\mathbf{k}}_{\nu,l}(\mathbf{r}) = \tilde{p}_{\nu,l}(\mathbf{r}) \cdot e^{-i\mathbf{k}\mathbf{r}}, \tag{13.15}$$

which arises from the additional phase factors in the cell periodic part of the wave function. The summations in the non local part are performed over all atoms in the unit cell and μ' and ν' are indexing all projectors used in the pseudopotential for the considered atoms.

13.4 Evaluation of the integral via a 3-d real space integration

The most simple implementation to calculate $\bar{H}^{\mathbf{k}}_{\mu\nu,ij}$ and $S^{\mathbf{k}}_{\mu\nu,ij}$ as given in Eq. (13.13) and Eq. (13.14) is to define a regular 3-dimensional real space grid in the primitive unit cell of the system under consideration and to calculate the different contributions to $\bar{H}^{\mathbf{k}}_{\mu\nu,ij}$ and $S^{\mathbf{k}}_{\mu\nu,ij}$ directly.
According to Eq. (13.14) the contributions to the matrix elements have the following form in real space

$$\begin{aligned} O^{\mathbf{k}}_{\mu\nu,ij} &= \langle \bar{\beta}^{\mathbf{k}}_{\mu,j} | \bar{O} | \bar{\beta}^{\mathbf{k}}_{\nu,i} \rangle \\ &= \int d^3r' \int d^3r \langle \bar{\beta}^{\mathbf{k}}_{\mu,j} | \mathbf{r}' \rangle \langle \mathbf{r}' | \bar{O} | \mathbf{r} \rangle \langle \mathbf{r} | \bar{\beta}^{\mathbf{k}}_{\nu,i} \rangle \\ &= \frac{1}{N} \int d^3r \sum_{\mathbf{R}'} e^{i\mathbf{k}(\mathbf{r}'-(\mathbf{R}'+\mathbf{R}_j))} \phi^*_{\mu,j}(\mathbf{r}' - (\mathbf{R}' + \mathbf{R}_j)) \; \bar{O}(\mathbf{r}', \mathbf{r}) \\ &\qquad \cdot \sum_{\mathbf{R}} e^{-i\mathbf{k}(\mathbf{r}-(\mathbf{R}+\mathbf{R}_i))} \phi_{\nu,i}(\mathbf{r} - (\mathbf{R} + \mathbf{R}_i)), \end{aligned} \tag{13.16}$$

where we use the operator $\bar{O}$ as placeholder for an arbitrary operator. Even those are non-local operators:

$$\bar{O}(\mathbf{r}', \mathbf{r}) = \langle \mathbf{r}' | \tilde{p}^{\mathbf{k}}_{\mu',l} \rangle Q_{\mu'\nu',l} \langle \tilde{p}^{\mathbf{k}}_{\nu',l} | \mathbf{r} \rangle. \tag{13.17}$$

As one can see the calculation of $|\bar{\beta}^{\mathbf{k}}_{\nu,i}\rangle$ involves a summation over the entire Bravais lattice which can be reduced to a summation by the following considerations: the atomic orbitals, which are calculated by an expansion in Bessel functions as given in Sec. 12.3.2, have a finite cutoff radius R_{cut} and can be imagined as spheres in the unit cell. If the sphere exceeds the border of the unit cell it will re-enter the unit cell on the opposite site. This corresponds to a contribution from the orbital translated by a lattice vector $\mathbf{R}$ (see Fig. 13.1).
Using this idea, the kinetic energy operator and the local contributions to the Hamilton matrix are easily calculated.

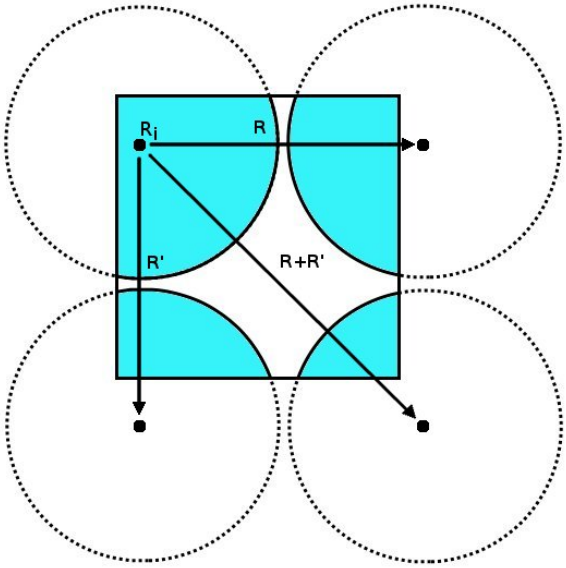


Figure 13.1: Systematic sketch of the construction of the wave functions in the primitive unit cell.

This method, which evaluates the matrix elements via a 3D integration has several drawbacks. It is the least efficient method and the memory requirement is highest because of the necessity to save the wave functions as well as the kinetic terms and the operators during execution of the program. Last but not least it turns out (see Sec. 14.1) by careful investigation of the different contributions to the Hamilton matrix elements, that for an accurate evaluation of the kinetic matrix elements a huge number of grid points is necessary which dramatically slows down the execution speed. Therefore we implemented a more sophisticated and precise method for evaluating the problematic matrix elements.

13.5 Reducing the dimensionality of integration by analytic evaluation

Due to the fact that the atomic orbitals $|\phi_{lm}\rangle$, kinetic energy operators acting on an orbital $T|\phi_{lm}\rangle$ and the projector functions $|\tilde{p}_{\nu,i}\rangle$ are products of spherical harmonics and a radial dependent function, the angular dependent integrations can be carried out analytically reducing the dimensionality of the integrations from three to one dimensional integrals. One can distinguish between two cases, namely the one centre term and the two centre terms, where the two functions building up the integrand are not located at the same atomic position.

13.5.1 Fourier transformation of LCAO basis function $\phi_{lm}(\mathbf{r})$

As discussed above the final LCAO basis functions consist of a radial and spherical part:

$$\phi_{lm}(\mathbf{r}) = Y_{lm}(\Omega_{\mathbf{r}})\ \bar{\phi}_l(r). \tag{13.18}$$

The spherical part is set up form the orthogonal Bessel basis discussed in Sec. 12.3.2.

$$\bar{\phi}_l(r) = \sum_{t=1}^{\mathrm{N+1}} D_t\ j_l(q_t r)\ \Theta(R_{\mathrm{cut}} - r) \tag{13.19}$$

Θ is the stepfunction and R_{cut} is the radius of the sphere for the specific atom under consideration. The Fourier Transformation of these functions is given by

$$\begin{aligned}
\phi_{lm}(\mathbf{q}) &= \frac{1}{\sqrt{(2\pi)^3}} \int d^3r\ \phi_{lm}(\mathbf{r})\ e^{-i\mathbf{rq}} \\
&= \frac{1}{\sqrt{(2\pi)^3}} \int d^3r\ Y_{lm}(\Omega_{\mathbf{r}}) \bar{\phi}_l(r) \\
&\qquad \cdot 4\pi \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} (-i)^{l'}\ Y_{l'm'}(\Omega_{\mathbf{r}}) Y_{l'm'}(\Omega_{\mathbf{q}}) j_{l'}(qr) \\
&= \frac{4\pi}{\sqrt{(2\pi)^3}} \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} \frac{1}{i^{l'}}\ Y_{l'm'}(\Omega_{\mathbf{q}}) \underbrace{\int d\Omega_{\mathbf{r}} Y_{lm}(\Omega_{\mathbf{r}}) Y_{l'm'}(\Omega_{\mathbf{r}})}_{\delta_{ll'}\delta_{mm'}} \\
&\qquad \cdot \int_0^{\infty} dr\ r^2 \bar{\phi}_l(r)\ j_{l'}(qr)\ \Theta(R_{\mathrm{cut}} - r) \\
&= Y_{lm}(\Omega_{\mathbf{q}})\ \underbrace{\sqrt{\frac{2}{\pi}} \frac{1}{i^l} \overbrace{\int_0^{R_{\mathrm{cut}}} dr\ r^2 \bar{\phi}_l(r)\ j_l(qr)}^{\hat{\phi}_l(q)}}_{\bar{\phi}_l(q)}.
\end{aligned} \tag{13.20}$$

13.6 Calculation of the overlap and kinetic energy operator

The matrix element of two Bloch basis functions is given by

$$\begin{aligned}
\left\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} \middle| \bar{\beta}^{\mathbf{k}}_{\nu,i} \right\rangle &= \int d^3r \, \langle \bar{\beta}^{\mathbf{k}}_{\mu,j} | \mathbf{r} \rangle \langle \mathbf{r} | \bar{\beta}^{\mathbf{k}}_{\nu,i} \rangle \\
&= \frac{1}{N} \int d^3r \sum_{\mathbf{R},\mathbf{R}'} e^{i\mathbf{k}(\mathbf{r}-(\mathbf{R}'+\mathbf{R}_j))} \phi^*_\mu(\mathbf{r}-(\mathbf{R}'+\mathbf{R}_j)) \\
&\qquad \cdot e^{-i\mathbf{k}(\mathbf{r}-(\mathbf{R}+\mathbf{R}_i))} \phi_\nu(\mathbf{r}-(\mathbf{R}+\mathbf{R}_i)) \\
&= \frac{1}{N} \sum_{\mathbf{R},\mathbf{R}'} e^{i\mathbf{k}(\mathbf{R}'-\mathbf{R}+\mathbf{R}_j-\mathbf{R}_i)} \\
&\qquad \cdot \int d^3r \, \phi^*_\mu(\mathbf{r}-(\mathbf{R}'+\mathbf{R}_j)) \phi_\nu(\mathbf{r}-(\mathbf{R}+\mathbf{R}_i)) \\
&= \sum_{\mathbf{R}} e^{i\mathbf{k}(\mathbf{R}+\mathbf{R}_{ij})} \underbrace{\int d^3r \, \phi^*_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}-(\mathbf{R}+\mathbf{R}_{ij}))}_{S_{ij}(\mathbf{R}+\mathbf{R}_{ij})} . \qquad (13.21)
\end{aligned}$$

From line 3 to 4 we used the fact that the double sum $\sum_{\mathbf{R},\mathbf{R}'}$ can be simplified to $N \cdot \sum_{\mathbf{R}'-\mathbf{R}}$. Therefore we redefined $\mathbf{R} = \mathbf{R}' - \mathbf{R}$ and additionally we introduce $\mathbf{R}_{ij} = \mathbf{R}_j - \mathbf{R}_i$.
For $i = j$, a one centre term is encountered, which can be calculated analytically using the orthogonality relation for Bessel functions (see Eq. C.6). If the distance is non zero ($i \neq j$), we have to calculate a two centre integral, which can be done analytically using the convolution theorem for Fourier transformations. First we calculate the equations for the overlap operator. The kinetic matrix element can easily be calculated from the overlap matrix element.

13.6.1 One centre term

$$\begin{aligned}
S_{ii}(0) &= \int d^3r \, \phi^*_{lm}(\mathbf{r}) \phi_{l'm'}(\mathbf{r}) \\
&= \int d^3r \, Y_{lm}(\Omega_\mathbf{r}) \bar{\phi}^*_{lm}(r) Y_{l'm'}(\Omega_\mathbf{r}) \bar{\phi}_{l'm'}(r) \\
&= \underbrace{\int d\Omega_\mathbf{r} \, Y_{lm}(\Omega_\mathbf{r}) Y_{l'm'}(\Omega_\mathbf{r})}_{\delta_{ll'}\delta_{mm'}} \\
&\qquad \cdot \int dr \, r^2 \sum_{t=1}^{N+1} D_t \, j_l(q_t r) \sum_{t'=1}^{N+1} D_{t'} \, j_{l'}(q_{t'} r) \, \Theta(R_{\mathrm{cut}} - r)
\end{aligned}$$

$$= \sum_{t=1}^{N+1}\sum_{t'=1}^{N+1} D_t D_{t'} \underbrace{\int_0^{R_{\text{cut}}} dr\ r^2\ j_l(q_t r) j_l(q_{t'} r)}_{\frac{R_{\text{cut}}^3}{2} j_{l+1}^2(q_t R_{\text{cut}})\bar{\delta}_{q_t q_{t'}}}$$

$$= \frac{R_{\text{cut}}^3}{2} \sum_{t=1}^{N+1} D_t^2\ j_{l+1}^2(q_t R_{\text{cut}}) \tag{13.22}$$

From line 4 to line 5 we used Eq. (C.6). The one centre kinetic matrix element is almost identical but includes an additional factor q_t^2 (see Eq. 12.16) and has the following form

$$K_{ii}(0) = \frac{R_{\text{cut}}^3}{4} \sum_{t=1}^{N+1} D_t^2\ q_t^2\ j_{l+1}^2(q_t R_{\text{cut}}). \tag{13.23}$$

13.6.2 Two centre term

The two centre term is calculated using the convolution theorem for Fourier transformations.

$$\begin{aligned}
S_{ij}(\mathbf{R}) &= \int d^3r\ \phi_{lm}^*(\mathbf{r})\phi_{l'm'}(\mathbf{r}-\mathbf{R}) \\
&= \int d^3q\ \phi_{lm}^*(\mathbf{q})\phi_{l'm'}(\mathbf{q}) e^{-i\mathbf{q}\mathbf{R}} \\
&= \int d^3q\ Y_{lm}(\Omega_{\mathbf{q}})\bar{\phi}_l^*(q)\ Y_{l'm'}(\Omega_{\mathbf{q}})\bar{\phi}_{l'}(q) \\
&\qquad \cdot 4\pi \sum_{L=0}^{\infty}\sum_{M=-L}^{L} (-i)^L j_L(qR) Y_{LM}(\Omega_{\mathbf{R}}) Y_{LM}(\Omega_{\mathbf{q}}) \\
&= 4\pi \sum_{L=0}^{\infty}\sum_{M=-L}^{L} (-i)^L Y_{LM}(\Omega_{\mathbf{R}}) \underbrace{\int d\Omega_{\mathbf{q}}\ Y_{lm}(\Omega_{\mathbf{q}}) Y_{l'm'}(\Omega_{\mathbf{q}}) Y_{LM}(\Omega_{\mathbf{q}})}_{C_{lml'm'LM}} \\
&\qquad \cdot \int_0^{Q_{\max}} dq\ q^2 \bar{\phi}_l^*(q)\bar{\phi}_{l'}(q) j_L(qR) \\
&= 8 \sum_{L=|l-l'|}^{l+l'}\sum_{M=-L}^{L} i^{l'-l-L} Y_{LM}(\Omega_{\mathbf{R}})\ C_{lml'm'LM} \\
&\qquad \cdot \underbrace{\int_0^{Q_{\max}} dq\ q^2 \hat{\phi}_l^*(q)\hat{\phi}_{l'}(q) j_L(qR)}_{\bar{S}_{ll'L}(R)} \\
&= 8 \sum_{L=|l-l'|}^{l+l'}\sum_{M=-L}^{L} i^{l'-l-L} Y_{LM}(\Omega_{\mathbf{R}})\ C_{lml'm'LM}\ \bar{S}_{ll'L}(R)
\end{aligned} \tag{13.24}$$

The kinetic energy matrix elements are again calculated similar to Eq. (13.23) but we have to include an additional q^2 term

$$K_{ij}(\mathbf{R}) = 4 \sum_{L=|l-l'|}^{l+l'} \sum_{M=-L}^{L} i^{l'-l-L} Y_{LM}(\Omega_{\mathbf{R}})\ C_{lml'm'LM}\ \bar{K}_{ll'L}(R) \qquad (13.25)$$

with

$$\bar{K}_{ll'L}(R) = \int_0^{Q_{\max}} dq\ q^4 \hat{\phi}_l^*(q)\hat{\phi}_{l'}(q) j_L(qR). \qquad (13.26)$$

The one dimensional integrals to calculate $\bar{S}_{ll'L}(R)$ and $\bar{K}_{ll'L}(R)$ are presently calculated numerically using a radial equally spaced grid and a simple Simpson integration.

13.7 Calculation of the projection of the wave function

In addition to the kinetic and overlap term we also have to calculate the projection of the atomic orbitals onto the projector functions supplied with the pseudopotential. Due to the fact that the projector functions are known on discrete sampling points in reciprocal space and are also a composition of a spherical harmonic and radial dependent part, the calculation can be carried out similar to the above given derivations.

13.7.1 One centre term

$$\begin{aligned}
P_{ii}(0) &= \int d^3q\ \tilde{p}^*_{l'm'}(\mathbf{q})\phi_{lm}(\mathbf{q}) \\
&= \int d^3q\ \frac{1}{4\pi}\sqrt{\frac{2}{\pi}} i^{l'} Y_{lm}(\Omega_{\mathbf{q}})\bar{p}^*_{l'}(q) Y_{lm}(\Omega_{\mathbf{q}})\bar{\phi}_l(q) \\
&= \frac{1}{4\pi}\sqrt{\frac{2}{\pi}} i^{l'} \underbrace{\int d\Omega_{\mathbf{q}}\ Y_{l'm'}(\Omega_{\mathbf{q}}) Y_{lm}(\Omega_{\mathbf{q}})}_{\delta_{ll'}\delta_{mm'}} \int dq\ q^2\ \bar{p}^*_{l'}(q)\sqrt{\frac{2}{\pi}} i^{-l}\hat{\phi}_l(q) \\
&= \frac{1}{2\pi^2}\int_0^{Q_{\max}} dq\ q^2\ \bar{p}^*_{l'}(q)\hat{\phi}_l(q) \qquad (13.27)
\end{aligned}$$

The factor $\frac{1}{4\pi}\sqrt{\frac{2}{\pi}}$ accounts for the fact, that the projector functions stored on the `POTCAR` are scaled. $\bar{p}_{l'}(q)$ stands for the radial part of the projector function as supplied in the `POTCAR` file.

13.7.2 Two centre term

Similar to the two centre terms for the overlap and kinetic matrix elements, the convolution theorem is used to simplify the two centre terms.

$$
\begin{aligned}
P_{ii}(\mathbf{R}) &= \int d^3r\ \tilde{p}^*_{l'm'}(\mathbf{r})\phi_{lm}(\mathbf{r}-\mathbf{R}) \\
&= \int d^3q\ \tilde{p}^*_{l'm'}(\mathbf{q})\phi_{lm}(\mathbf{q})e^{-i\mathbf{q}\mathbf{R}} \\
&= \frac{1}{4\pi}\sqrt{\frac{2}{\pi}}i^{l'}\int d^3q\ Y_{l'm'}(\Omega_{\mathbf{q}})\bar{p}^*_{l'}(q)\ Y_{lm}(\Omega_{\mathbf{q}})\bar{\phi}_l(q) \\
&\qquad \cdot 4\pi\sum_{L=0}^{\infty}\sum_{M=-L}^{L}(-i)^L j_L(qR)Y_{LM}(\Omega_{\mathbf{R}})Y_{LM}(\Omega_{\mathbf{q}}) \\
&= \sqrt{\frac{2}{\pi}}i^{l'}\sum_{L=0}^{\infty}\sum_{M=-L}^{L}(-i)^L Y_{LM}(\Omega_{\mathbf{R}}) \\
&\qquad \cdot \underbrace{\int d\Omega_{\mathbf{q}}\ Y_{lm}(\Omega_{\mathbf{q}})Y_{l'm'}(\Omega_{\mathbf{q}})Y_{LM}(\Omega_{\mathbf{q}})}_{C_{lml'm'LM}} \\
&\qquad \cdot \int_0^{Q_{\max}} dq\ q^2\bar{p}^*_{l'}(q)\bar{\phi}_l(q)j_L(qR) \\
&= \frac{2}{\pi}\sum_{L=|l-l'|}^{l+l'}\sum_{M=-L}^{L} i^{l'-l-L}Y_{LM}(\Omega_{\mathbf{R}})\ C_{lml'm'LM}\cdot \\
&\qquad \int_0^{Q_{\max}} dq\ q^2\bar{p}^*_{l'}(q)\hat{\phi}_l(q)j_L(qR)
\end{aligned}
\tag{13.28}
$$

13.8 Technical aspects

The grid spacing on the one dimensional q space grid is chosen equal for all atomic types included in the calculation, so that the evaluation of the integrals is kept as simple as possible. Due to the fact that the grid is differently spaced in general for the projector functions, the projectors are interpolated onto this grid using a spline interpolation. Additionally the coefficients for the construction of the LCSBF basis are stored, so that the one centre terms can be evaluated analytically as described above. The only term which has to be evaluated on the 3D real space grid is the expectation value of the local potential $\langle\bar{\beta}^{\mathbf{k}}_{\mu,j}|\tilde{V}_{\text{eff}}|\bar{\beta}^{\mathbf{k}}_{\nu,i}\rangle$.

Chapter 14

Numerical precision and simple applications

14.1 Translational invariance of the total energy

An important test for the precision of the numerical routines is the "smoothness" of the total energy. It is clear, that the total energy of a system remains constant when moving all atoms in a unit cell, since the crystal structure does not changes upon such a translation. Due to the discretization of the unit cell in real space, the numerical solution will not be exactly translational invariant but will show a variation in the total energy when moving all atoms of the basis from one real space grid point to a neighbouring one. The magnitude of the variation, is therefore a good criterion to judge the numerical accuracy of the integration methods used for the evaluation of the matrix elements.
From Eq. (13.13) and Eq. (13.14) we see that three different matrix elements contribute to the Hamilton and overlap matrices. The kinetic energy matrix $\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} | \bar{T} | \bar{\beta}^{\mathbf{k}}_{\nu,i} \rangle$, the matrix elements involving the potential $\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} | \tilde{V}_{\text{eff}} | \bar{\beta}^{\mathbf{k}}_{\nu,i} \rangle$ and the projector matrix elements $\langle \tilde{p}^{\mathbf{k}}_{\nu',l} | \bar{\beta}^{\mathbf{k}}_{\nu,i} \rangle$.

In Sec. 13.6 we derived an evaluation procedure entirely in reciprocal space for the kinetic energy and projector matrix elements. For the potential matrix elements the evaluation in reciprocal space is not applicable since the potential is given on the 3D real space grid.

In the following sections we present the variations of the total energy for a single atom and bulk system using different combinations of the two integration methods for the kinetic and projector matrix elements. For comparison we included the variation of the total energy when using the standard plane wave basis set within VASP. The nomenclature for the different combinations of integration methods is given in Tab. 14.1.

	$\langle \bar{\beta}^{\mathbf{k}}_{\mu,j} \vert \bar{T} \vert \bar{\beta}^{\mathbf{k}}_{\nu,i} \rangle$	$\langle \tilde{p}^{\mathbf{k}}_{\nu',l} \vert \bar{\beta}^{\mathbf{k}}_{\nu,i} \rangle$
krgrid	3D real	reciprocal
prgrid	reciprocal	3D real
kprgrid	3D real	3D real
fourier	reciprocal	reciprocal

Table 14.1: Glossary of the different combinations of integration methods as used in Fig. 14.1 and Fig. 14.2. "3D real" denotes an integration on the 3D real space grid, whereas "reciprocal" denotes an integration using the convolution theorem for Fourier transformation in reciprocal space as described in Sec. 13.6.

14.1.1 Single atom in a box

As initial and simplest test case, we used a single free atom in a box with two different radial cutoffs. One was chosen to be sufficiently large to ensure that the calculated basis functions are almost identical to the orbitals of the free atom. In this way we avoided a spatial confinement of the orbitals. The second radial cutoff was set to approximately half of the first cutoff for a strong spatial confinement.

The simulation box is chosen large enough to enclose the entire atomic sphere to avoid periodicity effects. The cutoff energy for the expansion in Bessel functions was chosen to be the standard cutoff energy as found in the `POTCAR` file and the Γ-point only was used. In the used pseudopotential for aluminium the $3s^2 3p^1$ electrons were in the valence, the core radius was 1.005 Å and the PBE parametrization of the exchange-correlation potential was used.

In Fig. 14.1 the variation in the total energy with respect to the translation of the atomic basis from one real space grid point to the next neighbouring point is presented for both radial cutoffs. First let us focus on the right graph, where the basis functions of the aluminium atom are not spatially confined. We see that the "prgrid" and "kprgrid" termed combinations of integration methods yield larger variations upon translation than the other combinations. In both combinations, the projector matrix elements are calculated on the 3D real space grid. The kinetic energy matrix elements are evaluated in these two combinations by two different methods and are therefore not the origin of the variations. This behaviour can be explained with the construction of the projector functions, which are localized within the PAW sphere. The PAW sphere itself has a much smaller radius (see Tab. 3.1 and Tab. 7.2) than the radius R_{cut} enclosing the atomic orbitals. As a consequence variations of the projector functions are not represented adequately on the coarse 3D real space grid, resulting in the behaviour ob-

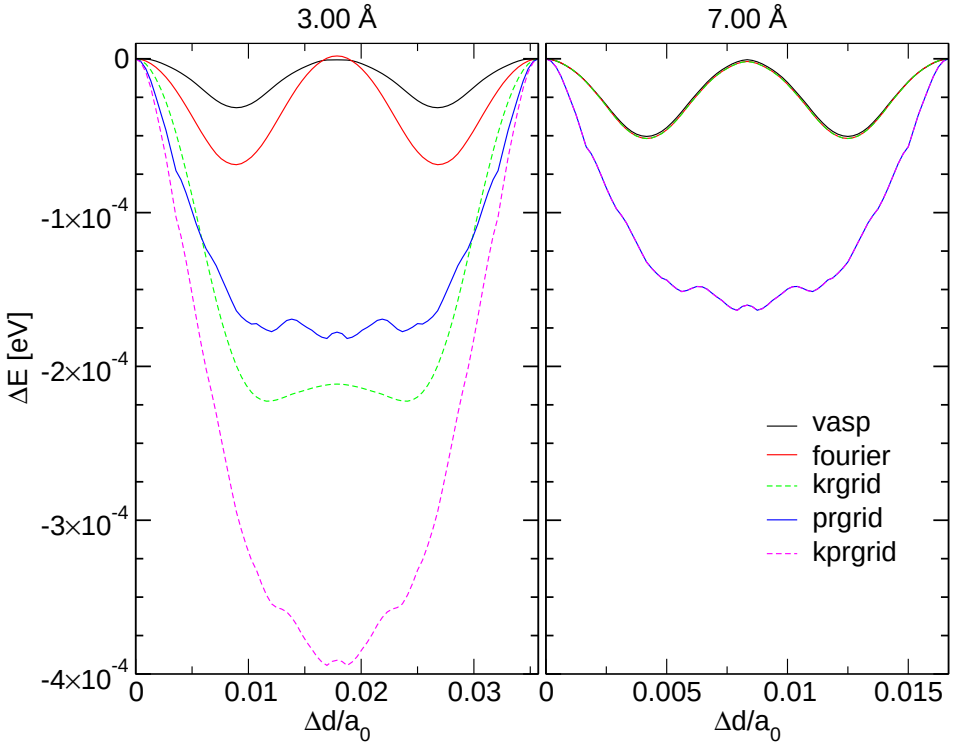


Figure 14.1: Comparison of the different integration methods using a single aluminium atom in a box with the corresponding plane wave calculation. All energy data points were referenced to their first energy with $\Delta d/a_0 = 0$. For the left graph the basis function radial cutoff R_{cut} was set to 3.00 Å and for the right graph to 7.00 Å.

served in Fig. 14.1. Both combinations where the projector matrix elements ("fourier" and "krgrid") are evaluated in reciprocal space are superior to the latter one. The chosen evaluation procedure of the kinetic energy matrix element does not effect the overall translational accuracy, which indicates that the resulting function when applying the kinetic energy operator on the basis functions is well represented on the 3D real space grid. Overall the variations are very similar to the plane wave VASP calculation for "fourier" and "krgrid". Therefore the variation in energy for these three calculations can be entirely attributed to the potential energy matrix, which is evaluated on the 3D real space grid. When confining the basis functions, as done on the left graph of Fig. 14.1, the translational variation of the kinetic energy evaluation on the 3D real space grid changes drastically. In principle, the same argument applies to this as for the projector functions. More precisely, when applying the kinetic energy operator on the confined basis functions oscillating functions are obtained (see Sec. 12.3.4) which can not be rep-

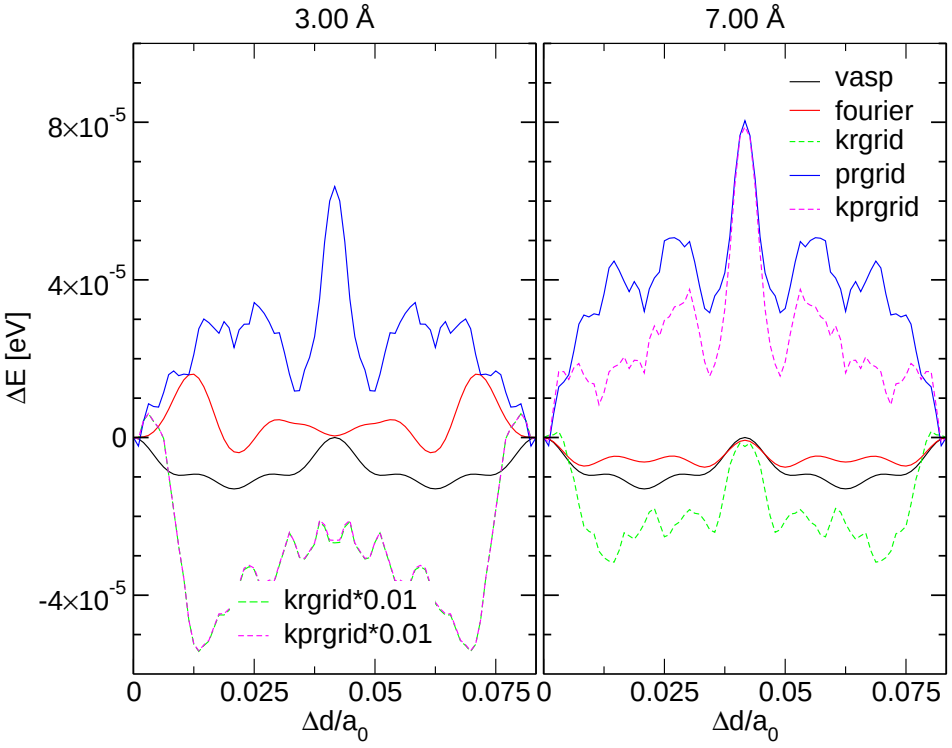


Figure 14.2: Comparison of the different integration methods using a aluminium in its fcc structure with the corresponding plane wave calculation. All energy data points were referenced to their first energy with $\Delta d/a_0 = 0$. For the left graph the basis function radial cutoff R_{cut} was set to 3.00 Å and for the right graph to 7.00 Å.

resented accurately on the 3D real space grid. Therefore, the integration methods that evaluate the kinetic matrix elements in reciprocal space perform better then other ones.
Again the entire evaluation of all matrix elements in reciprocal space is the most accurate one with a similar precision as using the plane wave basis set.

14.1.2 Crystal structure

As in the previous section we used aluminium as test element, but now we consider it in the fcc crystal structure. The calculations were performed for the equilibrium lattice constant of 4.041 Å obtained with the plane wave basis set of VASP. A $8\times8\times8$ Γ-centred k-point mesh was used for the Brillouin zone integration. All other settings were identical to the calculations in the previous section.
In the case of unconfined free atomic orbitals, as presented in the right graph of Fig. 14.2, the magnitudes of the variations of the total energies for

	valence e^-	E_{cut} [eV]	r_s [Å]	r_p [Å]
Si	$3s^2 3p^2$	245	1.005	1.005
Ge	$4s^2 4p^2$	174	1.217	1.217

Table 14.2: PAW atomic valence configurations and core radii for silicon and germanium.

combinations including integration on the 3D real space grid performs worst. Again the entire evaluation in reciprocal space is the favourable procedure with a similar precision to the plane wave basis set.
Identical to the previous section, a confined atomic orbital basis was used for the calculations on the translational variations in the left graph of Fig. 14.2. Note that, the "krgrid" and "kprgrid" curves in the left graph of Fig. 14.2 are scaled by 0.01 to be able to show them in a single graph for comparison. The same argumentation applies as for the calculation on a single atom.
All in all, we see that as expected, the evaluation of the kinetic energy matrix elements as well as the projector matrix elements in reciprocal space is the method of choice. Since the integration is not carried out on possibly too coarse 3D real space grid, a precision similar to the plane wave basis set as used in a standard VASP calculation is obtained.

14.2 Applications

Although it is well known that a minimal basis set is not capable to yield results similar to a high quality plane wave calculation we applied this minimal basis set to diamond silicon and germanium just to show that this procedure works and is a good starting point for further developments on a localized basis set capable for linear scaling ab initio calculations or the implementation of a mixed basis set. An outlook on how the accuracy of the basis set can be improved is given in the conclusion (see Chap. 15).
The following calculations were performed using the PBE parametrization of the exchange correlation functional as described in Sec. 1.4.2. For lattice parameter calculations, a $8 \times 8 \times 8$ Γ-centred k-point mesh is dense enough to yield converged results. The cutoff for expansion in plane waves or Bessel functions was chosen 50 percent higher than the standard cutoff energy E_{cut} as given in Tab. 14.2.

14.2.1 Lattice constants versus LCAO sphere radius $\mathbf{R_{cut}}$

With the introduction of a minimal, singe-ζ, basis set we also introduced a spatial cutoff radius R_{cut} to make the basis functions more convenient for calculation. Confining the atomic orbitals will decrease the inter atomic

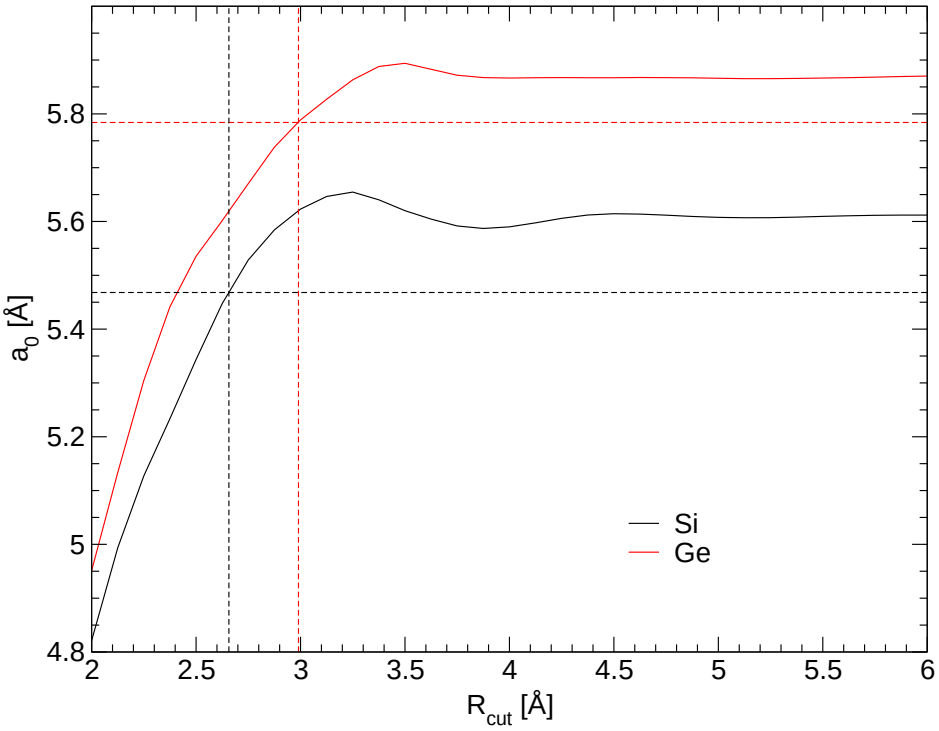


Figure 14.3: Lattice parameter in dependence of the radial cutoff radius R_{cut} for silicon and germanium in the diamond structure. The dashed horizontal lines are the equilibrium lattice constants obtained with a plane wave basis set.

bond lengths and the lattice parameter of bulk systems. Therefore the cutoff radius R_{cut} is the most important parameter controlled by the user.
For clarification we calculated for silicon and germanium in their diamond crystal structure for a set of cutoff radii the equilibrium lattice constants and plotted them in Fig. 14.3. Additionally we included the equilibrium lattice constants for both systems obtained with a plane wave basis set (horizontal dashed lines). From Fig. 14.3 we determine the cutoff radii where the minimal basis set calculation reproduces the plane wave calculation. The radii are 2.65 Å for silicon and 2.99 Å for germanium. Using not confined atomic orbitals for the minimal basis set one would yield too large lattice constants by 1-2 %.

14.2.2 Band structure and density of states of bulk silicon and germanium

For band structure and density of states calculations we used a $16 \times 16 \times 16$ Γ-centred k-point mesh to obtain a converged charge density. The opti-

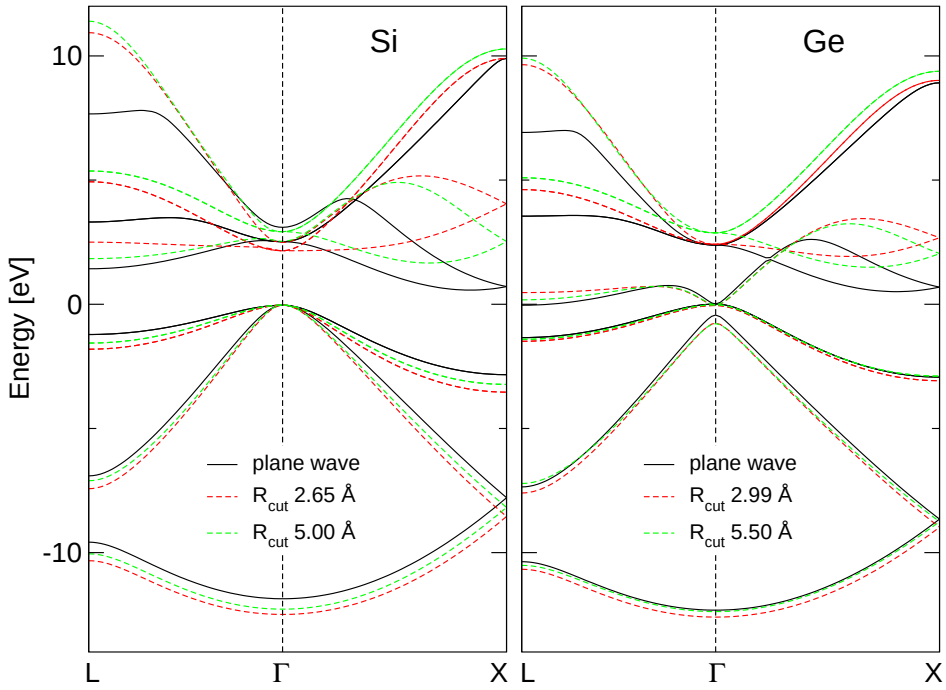


Figure 14.4: Band structure of bulk diamond silicon and germanium. The plane wave basis set calculation is shown with solid black lines, calculations using the minimal basis set are shown with dashed lines. For the red dashed line the radial cutoff was set to reproduce the plane wave lattice parameters, whereas for the green dashed line a cutoff was used to reproduce not confined atomic orbitals.

mized theoretical lattice constant of the plane wave calculation was used for all calculations. We used two different cutoff radii in the calculations for comparison. One radius was chosen to be the value where the minimal basis set reproduces the plane wave lattice parameter as reported before. The second radius was chosen large enough to minimize the effect of confinement on the atomic orbitals. Therefore we used 5.00 Å for silicon and 5.50 Å for germanium in agreement with Fig. 14.3. For plotting the band structure we used the usual high symmetry path in the Brillouin zone of the diamond lattice, namely $L - \Gamma - X$.

As can be seen in Fig 14.4 the agreement between the plane wave and minimal basis set for the valence band is acceptable. A larger cutoff radius seems to improve slightly the description of the valence bands compared to the plane wave calculation. Since we did not include any unoccupied orbitals within the minimal basis set, the conduction band is not represented ade-

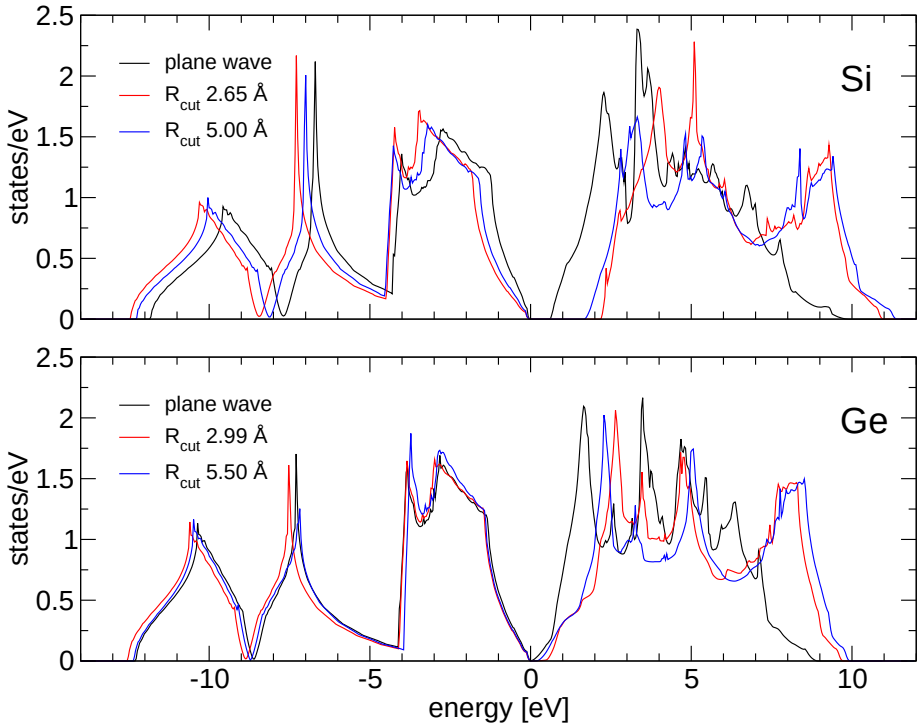


Figure 14.5: Density of states of bulk diamond silicon and germanium for different radial cutoffs and with the plane wave basis set.

quately. Similar to a tight binding approach the inclusion of the unoccupied d orbitals would improve the description of the conduction bands [104].
A similar observation is made for the density of states as shown in Fig. 14.5. Below the Fermi level the density of states is slightly shifted but comparable to the plane wave calculation. Again the conduction bands are not reproduced adequately and as a consequence the band gaps are also too large compared to the plane wave basis set calculation. Furthermore even the type of band gap changes from a direct ($R_{\text{cut}} = 2.65$ Å) to the correct indirect band gap ($R_{\text{cut}} = 5.00$ Å) in silicon when increasing the cutoff radius.

Chapter 15

Conclusion and outlook

We have implemented a minimal, single-ζ, basis set in VASP. Our implementation is based on the PAW formalism, as already implemented in the VASP code, and allows a seamless integration of the PAW method with local orbitals. The crucial step of this work is to expand the atomic orbitals in spherical Bessel functions that are spatially confined in a sphere of a chosen radius. The discontinuity of the orbitals at the sphere boundary is reduced by modifying the kinetic energy operator and penalizing high energy components in the kinetic energy operator. The final orbitals are smooth in real space and can be accurately presented on a fairly coarse 3D real space grid. In principle, this allows to calculate all required integrals directly using a simple 3D grid. For the kinetic energy and the non-local contributions to the PAW potential we, however, also explored a partially analytical integration scheme (requiring only a single numerical 1D integration). In fact, the analytical integration is significantly more accurate than the one using the 3D grid, and hence the preferred choice. We have extensively tested the present implementation, paying particular attention to the attainable numerical accuracy. To this end we have determined how the total energy changes upon rigidly shifting all atoms in the unit cell from one grid point to the next (egg-box effects or aliasing errors). If the kinetic energy term and the non-local term in the PAW formalism are evaluated analytically, the numerical accuracy is found to be at least on par with standard plane wave calculations. This is extremely encouraging, and suggests that the present construction of the basis functions allows for *numerically* highly accurate calculations.

Although, the focus of the present work was not on obtaining results that are converged with respect to the basis set, we have presented a few results for very simple test cases (C, Si, Ge). Of course, it is well known that a minimal basis set (as the one used here) is not sufficient to reproduce lattice constants or band gaps. But extension of the present code to larger basis sets is entirely straightforward and requires only minimal modifications of

the code. Matter of fact, all the developed algorithms will remain applicable for more extended basis sets.
A more interesting route, that will be pursue in the future is to merge the local basis set with a plane wave expansion (mixed basis set implementation). This route seems to be particularly promising, and will allow to reduce the basis sets for 2^{nd} row elements and $3d$ elements significantly. We envisage that all elements in the periodic table could be described using basis sets corresponding to a plane wave cutoff of around 150-200 eV. For the aforementioned, 2nd row elements and $3d$ elements, this will require one to supplement the plane waves with local orbitals for the spatially confined $2p$ and $3d$ orbitals. Other natural extensions are downfolding of the PW results to a minimal basis set to aid chemical interpretation of the results.

Appendix A

PAW-formalism

The derivation follows closely [24, 25]. The aim of this method is to map the physically relevant wave functions on computationally convenient pseudo wave functions. The transformation between these two Hilbert spaces is defined as

$$|\Psi\rangle = T|\tilde{\Psi}\rangle. \tag{A.1}$$

The expectation value of an operator can then be derived in both representation through

$$\langle\Psi|\, A\,|\Psi\rangle = \langle\tilde{\Psi}|T^{\dagger}AT|\tilde{\Psi}\rangle = \langle\tilde{\Psi}|\tilde{A}|\tilde{\Psi}\rangle \tag{A.2}$$

with $\tilde{A} = T^{\dagger}AT$. The transformation is defined in such a way as to account for the atomic configuration of the system under consideration. The transformation differs from identity only by a sum of local atom centred contributions T_R,

$$T = 1 + \sum_R T_R, \tag{A.3}$$

which act only within some augmentation region Ω_R enclosing the atom. Therefore both all electron (AE) and pseudo (PS) wave functions are identical outside the augmentation region. The local terms T_R are defined for each augmentation region individually by

$$|\phi_i\rangle = (1 + T_R)|\tilde{\phi}_i\rangle, \tag{A.4}$$

where $|\phi_i\rangle$ is an AE partial wave and $|\tilde{\phi}_i\rangle$ is a pseudo partial wave . The AE partial waves are solutions of a one-electron radial Schrödinger (Kohn-Sham) equation for an isolated atom. Therefore the index i is a combined index R, l, m, n. Here R stands for the atomic site, l, m are the angular momentum quantum numbers and n is an additional degree of freedom to label different partial waves for the same site and angular momentum. Let

us assume that there exists for each $|\tilde{\phi}_i\rangle$ an corresponding $|\phi_i\rangle$ and that they are identical if $r > \Omega_R$.
Each pseudo orbital can be expanded into terms of pseudo partial waves

$$|\tilde{\Psi}\rangle = \sum_i c_i |\tilde{\phi}_i\rangle \text{ within } \Omega_R. \tag{A.5}$$

One can show with Eq. (A.3), (A.4) and the fact the T_R acts only on the augmentation region at R that the coefficients c_i stay the same in the expansion of the all-electron wave function

$$|\Psi\rangle = T|\tilde{\Psi}\rangle = T\sum_i c_i|\tilde{\phi}_i\rangle \tag{A.6}$$

$$= \sum_i c_i \left(1 + \sum_R T_R\right)|\tilde{\phi}_i\rangle \tag{A.7}$$

$$= \sum_i c_i \left(1 + T_R\right)|\tilde{\phi}_i\rangle \tag{A.8}$$

$$= \sum_i c_i \left|\phi_i\right\rangle. \tag{A.9}$$

Thus, the all electron wave function can be expanded in the following way

$$|\Psi\rangle = |\tilde{\Psi}\rangle - |\tilde{\Psi}\rangle + |\Psi\rangle \tag{A.10}$$

$$= |\tilde{\Psi}\rangle - \sum_i c_i|\tilde{\phi}_i\rangle + \sum_i c_i \left|\phi_i\right\rangle. \tag{A.11}$$

The c_i are still not known. Comparing this expansion with the desired structure of Eq. (A.3)

$$|\Psi\rangle = (1 + \sum_R T_R)|\tilde{\Psi}\rangle \tag{A.12}$$

we find that

$$c_i = \langle\tilde{p}_i|\tilde{\Psi}\rangle. \tag{A.13}$$

The $|\tilde{p_i}\rangle$ are called projector functions and for each pseudo partial wave there is one projector function. Furthermore the projector functions have to fulfil

$$\sum_i |\tilde{\phi}_i\rangle\langle\tilde{p}_i| = 1 \tag{A.14}$$

because otherwise the one centre expansion of the pseudo wave function is not identical to itself

$$|\tilde{\Psi}\rangle = \sum_i c_i|\tilde{\phi}_i\rangle \tag{A.15}$$

$$= \sum_i \langle\tilde{p}_i|\tilde{\Psi}\rangle|\tilde{\phi}_i\rangle \tag{A.16}$$

$$= \underbrace{\sum_i |\tilde{\phi}_i\rangle\langle\tilde{p}_i|}_{=1}\,\tilde{\Psi}\rangle. \tag{A.17}$$

Additionally one can show that the following relation has to be fulfilled so that Eq. (A.14) holds

$$\langle \tilde{p}_i | \tilde{\phi}_j \rangle = \delta_{ij}. \tag{A.18}$$

Using Eq. (A.14)

$$1 \stackrel{!}{=} \sum_i |\tilde{\phi}_i\rangle\langle\tilde{p}_i| \sum_j |\tilde{\phi}_j\rangle\langle\tilde{p}_j| \tag{A.19}$$

$$\stackrel{!}{=} \sum_{i,j} |\tilde{\phi}_i\rangle \underbrace{\langle\tilde{p}_i|\tilde{\phi}_j\rangle}_{\delta_{ij}}\langle\tilde{p}_j| \tag{A.20}$$

$$\stackrel{!}{=} \sum_i |\tilde{\phi}_i\rangle\langle\tilde{p}_i|. \tag{A.21}$$

In summary, one obtains for the transformation the following form

$$T = 1 + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle \right) \langle\tilde{p}_i|. \tag{A.22}$$

We see that three quantities determine the transformation

- the all electron partial waves $|\phi_i\rangle$
- the pseudo partial wave $|\tilde{\phi}_i\rangle$, which is identical to $|\phi_i\rangle$ outside Ω_R
- the projector function $|\tilde{p}_i\rangle$ to each $|\tilde{\phi}_i\rangle$ within the augmentation region, which obeys $\langle\tilde{p}_i|\tilde{\phi}_j\rangle = \delta_{ij}$:

$$|\Psi\rangle = |\tilde{\Psi}\rangle + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle \right) \langle\tilde{p}_i|\tilde{\Psi}\rangle. \tag{A.23}$$

By means of Eq. (A.23) the pseudo wave functions $|\tilde{\Psi}\rangle$ become the variational parameters and $|\Psi\rangle$ is uniquely determined by $|\tilde{\Psi}\rangle$.

A.1 Operators within the PAW-method

Since the pseudo wave functions are the variational parameters we have to express the expectation value of an operator in terms of $|\tilde{\Psi}\rangle$.

$$\langle A \rangle = \sum_n f_n \langle\Psi| A |\Psi\rangle = \sum_n f_n \langle\tilde{\Psi}|\tilde{A}|\tilde{\Psi}\rangle \tag{A.24}$$

A.1.1 Local operator

With the exact expression of the transformation T (see Eq. (A.22)) we can calculate the exact form of the operator $\tilde{A}$.

$$
\begin{aligned}
\tilde{A} =& T^\dagger A T \\
=& \left[1 + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle\right) \langle\tilde{p}_i|\right]^\dagger A \left[1 + \sum_j \left(|\phi_j\rangle - |\tilde{\phi}_j\rangle\right) \langle\tilde{p}_j|\right] \\
=& A + \sum_i |\tilde{p}_i\rangle \left(\langle\phi_i| - \langle\tilde{\phi}_i|\right) A + \sum_j A \left(|\phi_j\rangle - |\tilde{\phi}_j\rangle\right) \langle\tilde{p}_j| \\
&+ \sum_{i,j} |\tilde{p}_i\rangle \left(\langle\phi_i| - \langle\tilde{\phi}_i|\right) A \left(|\phi_j\rangle - |\tilde{\phi}_j\rangle\right) \langle\tilde{p}_j| \\
=& A + \sum_i |\tilde{p}_i\rangle \langle\phi_i| A - \sum_i |\tilde{p}_i\rangle\langle\tilde{\phi}_i|A + \sum_j A |\phi_j\rangle \langle\tilde{p}_j| - \sum_j A|\tilde{\phi}_j\rangle\langle\tilde{p}_j| \\
&+ \sum_{i,j} \Bigg[|\tilde{p}_i\rangle \langle\phi_i| A |\phi_j\rangle \langle\tilde{p}_j| - \underbrace{|\tilde{p}_i\rangle\langle\tilde{\phi}_i|}_{=1} A |\phi_j\rangle \langle\tilde{p}_j| \\
&\qquad - |\tilde{p}_i\rangle \langle\phi_i| A \underbrace{|\tilde{\phi}_j\rangle\langle\tilde{p}_j|}_{=1} + |\tilde{p}_i\rangle\langle\tilde{\phi}_i|A|\tilde{\phi}_j\rangle\langle\tilde{p}_j| \Bigg] \\
=& A - \sum_i |\tilde{p}_i\rangle\langle\tilde{\phi}_i|A - \sum_j A|\tilde{\phi}_j\rangle\langle\tilde{p}_j| \\
&+ \sum_{i,j} \left[|\tilde{p}_i\rangle \langle\phi_i| A |\phi_j\rangle \langle\tilde{p}_j| + |\tilde{p}_i\rangle\langle\tilde{\phi}_i|A|\tilde{\phi}_j\rangle\langle\tilde{p}_j| \right] \\
=& A + \sum_{i,j} \left[|\tilde{p}_i\rangle \langle\phi_i| A |\phi_j\rangle \langle\tilde{p}_j| - |\tilde{p}_i\rangle\langle\tilde{\phi}_i|A|\tilde{\phi}_j\rangle\langle\tilde{p}_j| \right] \\
=& A + \sum_{i,j} |\tilde{p}_i\rangle \left[\langle\phi_i| A |\phi_j\rangle - \langle\tilde{\phi}_i|A|\tilde{\phi}_j\rangle \right] \langle\tilde{p}_j| \\
=& A + A^1 - \tilde{A}^1
\end{aligned}
\tag{A.25}
$$

Furthermore there is additional freedom to add an arbitrary operator inside the augmentation region to $\tilde{A}$. Due to the completeness relation Eq. (A.14)

$$
B - \sum_{i,j} |\tilde{p}_i\rangle\langle\tilde{\phi}_i|B|\tilde{\phi}_j\rangle\langle\tilde{p}_j| = 0. \tag{A.26}
$$

The expectation value equals zero for any pseudo wave function.

A.1.2 Derivation of the charge density

The charge density $n(\mathbf{r})$ at $\mathbf{r}$ is given by the expectation value of the real space projection operator $|r\rangle\langle r|$:

$$
\begin{aligned}
|r\rangle \langle r| &= \sum_n f_n \langle \Psi_n|\ r\rangle \langle r|\ \Psi_n\rangle \\
&= \sum_n f_n \langle \tilde{\Psi}_n | T^\dagger r\rangle \langle rT | \tilde{\Psi}_n\rangle. \qquad \text{(A.27)}
\end{aligned}
$$

Using Eq. (A.25) we get for the real space projection operator

$$
T^\dagger\, |r\rangle \langle r|\, T = |r\rangle \langle r| + \sum_{i,j} |\tilde{p}_i\rangle \left[\langle \phi_i|\ r\rangle \langle r|\ \phi_j\rangle - \langle \tilde{\phi}_i | r\rangle \langle r | \tilde{\phi}_j\rangle \right] \langle \tilde{p}_j| \qquad \text{(A.28)}
$$

and finally for the charge density in the PAW formalism

$$
\begin{aligned}
n(\mathbf{r}) = & \underbrace{\sum_n f_n \langle \tilde{\Psi}_n | r\rangle \langle r | \tilde{\Psi}_n\rangle}_{\tilde{n}(\mathbf{r})} + \underbrace{\sum_{n,i,j} f_n \langle \tilde{\Psi}_n | \tilde{p}_i\rangle \langle \phi_i|\ r\rangle \langle r|\ \phi_j\rangle \langle \tilde{p}_j | \tilde{\Psi}_n\rangle}_{n^1(\mathbf{r})} \\
& - \underbrace{\sum_{n,i,j} f_n \langle \tilde{\Psi}_n | \tilde{p}_i\rangle \langle \tilde{\phi}_i | r\rangle \langle r | \tilde{\phi}_j\rangle \langle \tilde{p}_j | \tilde{\Psi}_n\rangle}_{\tilde{n}^1(\mathbf{r})}
\end{aligned}
\qquad \text{(A.29)}
$$

using

$$
\rho_{ij} = \sum_n f_n \langle \tilde{\Psi}_n | \tilde{p}_i\rangle \langle \tilde{p}_j | \tilde{\Psi}_n\rangle \qquad \text{(A.30)}
$$

which are the occupancies of each augmentation channel (i,j) we get

$$
\begin{aligned}
n(\mathbf{r}) &= \underbrace{\sum_n f_n \langle \tilde{\Psi}_n | r\rangle \langle r | \tilde{\Psi}_n\rangle}_{\tilde{n}(\mathbf{r})} + \underbrace{\sum_{i,j} \rho_{ij} \langle \phi_i|\ r\rangle \langle r|\ \phi_j\rangle}_{n^1(\mathbf{r})} - \underbrace{\sum_{i,j} \rho_{ij} \langle \tilde{\phi}_i | r\rangle \langle r | \tilde{\phi}_j\rangle}_{\tilde{n}^1(\mathbf{r})} \\
&= \tilde{n}(\mathbf{r}) + n^1(\mathbf{r}) - \tilde{n}^1(\mathbf{r}). \qquad \text{(A.31)}
\end{aligned}
$$

From now on we will use the frozen core approximation within the calculations. Therefore we restrict all quantities like $\tilde{n}(\mathbf{r}), ...$ and the index n over the valence bands only. This leads to the introduction of four more quantities that describe the core charge densities.

n_c ... frozen core all-electron charge density of the reference atom

$\tilde{n}_\mathrm{c}$ ... partial electronic core density, equivalent to n_c outside a radius r_pc.

n_Zc ... point charge density n_Z plus n_c

$\tilde{n}_\mathrm{Zc}$ ... charge distribution equal to n_Zc outside some radius and shall have the same moment inside the core region

$$
\int_{\Omega_r} n_\mathrm{Zc}(\mathbf{r})\ d^3r = \int_{\Omega_r} \tilde{n}_\mathrm{Zc}(\mathbf{r})\ d^3r. \qquad \text{(A.32)}
$$

A.1.3 Hartree energy

The Hartree energy is given by

$$E_H[n_T] = \frac{1}{2}\int d^3r \int d^3r' \frac{n_T(\mathbf{r})n_T(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{A.33}$$

where

$$n_T(\mathbf{r}) = n(\mathbf{r}) + n_{Zc}(\mathbf{r}). \tag{A.34}$$

This leads to

$$\begin{aligned} E_H[n+n_{Zc}] &= \frac{1}{2}\int d^3r \int d^3r' \frac{(n(\mathbf{r})+n_{Zc}(\mathbf{r}))(n(\mathbf{r}')+n_{Zc}(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|} \\ &= \underbrace{\frac{1}{2}\int d^3r \int d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}}_{e-e} + \underbrace{\int d^3r \int d^3r' \frac{n(\mathbf{r})n_{Zc}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}}_{e-ion} \\ &\quad + \underbrace{\frac{1}{2}\int d^3r \int d^3r' \frac{n_{Zc}(\mathbf{r})n_{Zc}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}}_{ion-ion}. \end{aligned} \tag{A.35}$$

Because of the non-locality of the operator, the calculation of the Hartree energy is not entirely straightforward. The total charge density $n_T(\mathbf{r})$ needs to be decomposed in the following way:

$$\begin{aligned} n_T &= n + n_{Zc} \\ &= \tilde{n} + n^1 - \tilde{n}^1 + n_{Zc} \\ &= (\tilde{n}+\hat{n}) + (n^1 + n_{Zc}) - (\tilde{n}^1 + \hat{n}) \\ &= \underbrace{(\tilde{n}+\hat{n}+\tilde{n}_{Zc})}_{\tilde{n}_T} + \underbrace{(n^1+n_{Zc})}_{n_T^1} - \underbrace{(\tilde{n}^1+\hat{n}+\tilde{n}_{Zc})}_{\tilde{n}_T^1}. \end{aligned} \tag{A.36}$$

The crucial step here is the introduction of a compensation charge $\hat{n}$. Using the following abbreviation

$$(a)(b) = \int d^3r \int d^3r' \frac{a(\mathbf{r})b(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{A.37}$$

the Hartree energy is decomposed into the following terms

$$\begin{aligned} E_H[n_T] &= \frac{1}{2}(n_T)(n_T) \\ &= \frac{1}{2}(\tilde{n}_T + n_T^1 - \tilde{n}_T^1)(\tilde{n}_T + n_T^1 - \tilde{n}_T^1) \\ &= \frac{1}{2}(\tilde{n}_T)(\tilde{n}_T) + (n_T^1 - \tilde{n}_T^1)(\tilde{n}_T) + \frac{1}{2}(n_T^1 - \tilde{n}_T^1)(n_T^1 - \tilde{n}_T^1). \end{aligned} \tag{A.38}$$

Since the potentials created by n_T^1 and $\tilde{n}_T^1$ are exactly identical outside the PAW sphere the second and third terms in Eq (A.38) vanishes there. Inside the spheres the second term constitutes a problem, $\tilde{n}_T$ is calculated on a plane wave grid, whereas $(n_T^1 - \tilde{n}_T^1)$ is saved on a radial grid. Interaction terms between both must be avoided for reasons of efficiency, hence one uses the usual completeness assumption $\tilde{n}_T \approx \tilde{n}_T^1$. The approximation holds only if a complete set of projectors is used. If this is not the case an error of $\overline{(n_T^1 - \tilde{n}_T^1)(\tilde{n}_T - \tilde{n}_T^1)}$ is introduced. With this approximation no interaction terms between the two used grids are present. The overline means that the quantity is evaluated on a radial grid.

$$E_H[n_T] = \frac{1}{2}(\tilde{n}_T)(\tilde{n}_T) - \frac{1}{2}\overline{(\tilde{n}_T^1)(\tilde{n}_T^1)} + \frac{1}{2}\overline{(n_T^1)(n_T^1)} \tag{A.39}$$

In the next part we will insert the definitions of Eq. (A.36) into Eq. (A.39) and calculate each term separately

$$\begin{aligned}
\frac{1}{2}(\tilde{n}_T)(\tilde{n}_T) &= \frac{1}{2}(\tilde{n} + \hat{n} + \tilde{n}_{Zc})(\tilde{n} + \hat{n} + \tilde{n}_{Zc}) \\
&= \frac{1}{2}(\tilde{n} + \hat{n})(\tilde{n} + \hat{n}) + (\tilde{n}_{Zc})(\tilde{n} + \hat{n}) + \frac{1}{2}(\tilde{n}_{Zc})(\tilde{n}_{Zc}) \\
&= \frac{1}{2}(\tilde{n} + \hat{n})(\tilde{n} + \hat{n}) + (\tilde{n}_{Zc})(\tilde{n} + \hat{n}) + \frac{1}{2}\overline{(\tilde{n}_{Zc})(\tilde{n}_{Zc})} + U(\mathbf{r}, Z_{ion})
\end{aligned}$$

$$\begin{aligned}
\frac{1}{2}\overline{(\tilde{n}_T^1)(\tilde{n}_T^1)} &= \frac{1}{2}\overline{(\tilde{n}^1 + \hat{n} + \tilde{n}_{Zc})(\tilde{n}^1 + \hat{n} + \tilde{n}_{Zc})} \\
&= \frac{1}{2}\overline{(\tilde{n}^1 + \hat{n})(\tilde{n}^1 + \hat{n})} + \overline{(\tilde{n}_{Zc})(\tilde{n}^1 + \hat{n})} + \frac{1}{2}\overline{(\tilde{n}_{Zc})(\tilde{n}_{Zc})}
\end{aligned}$$

$$\begin{aligned}
\frac{1}{2}\overline{(n_T^1)(n_T^1)} &= \frac{1}{2}\overline{(n^1 + n_{Zc})(n^1 + n_{Zc})} \\
&= \frac{1}{2}\overline{(n^1)(n^1)} + \overline{(n^1)(n_{Zc})} + \frac{1}{2}\overline{(n_{Zc})(n_{Zc})}.
\end{aligned} \tag{A.40}$$

A.1.4 Exchange and correlation energy

For the exchange and correlation energy only the electronic charge density is important. Therefore the total charge density is different to the one used in the Hartree potential. The decomposition is similar to the one used in the Hartree energy:

$$n + n_c = (\tilde{n} + \hat{n} + \tilde{n}_c) + (n^1 + n_c) - (\tilde{n}^1 + \hat{n} + \tilde{n}_c). \tag{A.41}$$

For local or semi local energy expressions we obtain

$$E_{xc}[n+n_c] = E_{xc}[\tilde{n}+\hat{n}+\tilde{n}_c] + \overline{E_{xc}[n^1+n_c]} - \overline{E_{xc}[\tilde{n}^1+\hat{n}+\tilde{n}_c]}. \quad \text{(A.42)}$$

A.1.5 Kinetic energy

The kinetic energy $\langle T \rangle$ can be decomposed into the following terms using Eq. (A.30)

$$\begin{aligned}
\langle T \rangle &= \sum_n f_n \langle \tilde{\Psi}_n | \tilde{T} | \tilde{\Psi}_n \rangle \\
&= \sum_n f_n \langle \tilde{\Psi}_n | T | \tilde{\Psi}_n \rangle + \sum_{n,i,j} f_n \langle \tilde{\Psi}_n | \tilde{p}_i \rangle \left[\langle \phi_i | \, T \, | \phi_j \rangle - \langle \tilde{\phi}_i | T | \tilde{\phi}_j \rangle \right] \langle \tilde{p}_j | \tilde{\Psi}_n \rangle \\
&= \sum_n f_n \langle \tilde{\Psi}_n | T | \tilde{\Psi}_n \rangle + \sum_{i,j} \rho_{ij} \, \langle \phi_i | \, T \, | \phi_j \rangle - \sum_{i,j} \rho_{ij} \langle \tilde{\phi}_i | T | \tilde{\phi}_j \rangle. \qquad \text{(A.43)}
\end{aligned}$$

A.1.6 Total energy expression

The total energy is given by

$$\begin{aligned}
E &= \sum_n f_n \, \langle \Psi_n | \, H \, | \Psi_n \rangle \\
&= \sum_n f_n \, \langle \Psi_n | \, T \, | \Psi_n \rangle + E_H[n+n_{Zc}] + E_{xc}[n+n_c]. \qquad \text{(A.44)}
\end{aligned}$$

Using now the results of the previous section we can decompose the energy functional in three terms

$$E = \tilde{E} + E^1 - \tilde{E}^1 \qquad \text{(A.45)}$$

namely

$$\begin{aligned}
\tilde{E} &= \sum_n f_n \langle \tilde{\Psi}_n | T | \tilde{\Psi}_n \rangle + E_H[\tilde{n}+\hat{n}] + \int V_H[\tilde{n}_{Zc}][\tilde{n}(\mathbf{r}) + \hat{n}(\mathbf{r})] \; d^3r \\
&\qquad + E_{xc}[\tilde{n}+\hat{n}+\tilde{n}_c] + U(\mathbf{R}, Z_{ion}) \\
E^1 &= \sum_{i,j} \rho_{ij} \, \langle \phi_i | \, T \, | \phi_j \rangle + \overline{E_H[n^1]} + \int V_H[n_{Zc}] n^1(\mathbf{r}) \; d^3r + \overline{E_{xc}[n^1+n_c]} \\
\tilde{E}^1 &= \sum_{i,j} \rho_{ij} \langle \tilde{\phi}_i | T | \tilde{\phi}_j \rangle + \overline{E_H[\tilde{n}^1+\hat{n}]} + \int V_H[\tilde{n}_{Zc}][\tilde{n}^1(\mathbf{r}) + \hat{n}(\mathbf{r})] \; d^3r \\
&\qquad + \overline{E_{xc}[\tilde{n}^1+\hat{n}+\tilde{n}_c]}
\end{aligned} \qquad \text{(A.46)}$$

with

$$V_H[n] = \int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{A.47}$$

$$E_H[n] = \frac{1}{2}(n)(n) = \frac{1}{2}\int d^3r \int d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{2}\int d^3r V_H[n][n(\mathbf{r})]. \tag{A.48}$$

A.1.7 Overlap operator

In the all-electron calculation the overlap matrix is simply given by the unity operator **1** due to the orthonormalization of the wave functions

$$\begin{aligned}\tilde{S} &= S + \sum_{i,j} |\tilde{p}_i\rangle \left[\langle\phi_i| \, S \, |\phi_j\rangle - \langle\tilde{\phi}_i|S|\tilde{\phi}_j\rangle\right] \langle\tilde{p}_j| \\ &= \mathbf{1} + \sum_{i,j} |\tilde{p}_i\rangle \left[\langle\phi_i| \, \phi_j\rangle - \langle\tilde{\phi}_i|\tilde{\phi}_j\rangle\right] \langle\tilde{p}_j| \\ &= \mathbf{1} + \sum_{i,j} |\tilde{p}_i\rangle Q_{ij} \langle\tilde{p}_j|\end{aligned} \tag{A.49}$$

with

$$Q_{ij} = \langle\phi_i| \, \phi_j\rangle - \langle\tilde{\phi}_i|\tilde{\phi}_j\rangle. \tag{A.50}$$

A.1.8 Hamilton operator

Due to the fact that the pseudo wave functions $|\tilde{\Psi}_n\rangle$ are now the variational parameters, we have to find the minimum of the PAW-energy functional

$$\frac{dE}{d\tilde{\rho}} = H \tag{A.51}$$

with respect to the pseudo density operator

$$\tilde{\rho} = \sum_n f_n |\tilde{\Psi}_n\rangle\langle\tilde{\Psi}_n|. \tag{A.52}$$

The density operator enters the PAW-energy functional in three places:

- directly like in the kinetic energy

$$\langle\tilde{A}\rangle = \sum_n f_n \langle\tilde{\Psi}_n|\tilde{A}|\tilde{\Psi}_n\rangle = Tr[\rho\tilde{A}] \tag{A.53}$$

- via the pseudo charge density $\tilde{n}$

$$\tilde{n} = \sum_n f_n \langle\tilde{\Psi}_n|r\rangle\langle r|\tilde{\Psi}_n\rangle \tag{A.54}$$

- through the occupancies of the augmentation channels ρ_{ij}

$$\rho_{ij} = \sum_n f_n \langle \tilde{\Psi}_n | \tilde{p}_i \rangle \langle \tilde{p}_j | \tilde{\Psi}_n \rangle. \tag{A.55}$$

We get

$$\frac{dE}{d\tilde{\rho}} = \frac{\partial E}{\partial \tilde{\rho}} + \int \frac{\partial E}{\partial \tilde{n}} \frac{\partial \tilde{n}}{\partial \tilde{\rho}} \, d^3r + \sum_{i,j} \frac{\partial E}{\partial \rho_{ij}} \frac{\partial \rho_{ij}}{\partial \tilde{\rho}}, \tag{A.56}$$

with

$$\frac{\partial \tilde{n}}{\partial \tilde{\rho}} = |r\rangle \langle r| \quad \text{and} \quad \frac{\partial \rho_{ij}}{\partial \tilde{\rho}} = |\tilde{p}_i\rangle \langle \tilde{p}_j| \,. \tag{A.57}$$

It follows

$$\frac{dE}{d\tilde{\rho}} = \frac{\partial E}{\partial \tilde{\rho}} + \frac{\partial E}{\partial \tilde{n}} + \sum_{i,j} \frac{\partial E}{\partial \rho_{ij}} |\tilde{p}_i\rangle \langle \tilde{p}_j| \,. \tag{A.58}$$

Now let us have a look at each term of the energy $E = \tilde{E} + E^1 - \tilde{E}^1$ separately.

$$\tilde{E} = \sum_n f_n \langle \tilde{\Psi}_n | T | \tilde{\Psi}_n \rangle + E_H[\tilde{n} + \hat{n}] + \int V_H[\tilde{n}_{Zc}][\tilde{n}(\mathbf{r}) + \hat{n}(\mathbf{r})] \, d^3r + E_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] + U(\mathbf{R}, Z_{ion}) \tag{A.59}$$

$$\begin{aligned}
\frac{\partial \tilde{E}}{\partial \tilde{\rho}} &= T \\
\frac{\partial \tilde{E}}{\partial \tilde{n}} &= \frac{\partial}{\partial \tilde{n}} \left\{ E_H[\tilde{n} + \hat{n}] + \int V_H[\tilde{n}_{Zc}][\tilde{n}(\mathbf{r}) + \hat{n}(\mathbf{r})] \, d^3r + E_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] \right\} \\
&= V_H[\tilde{n} + \hat{n}] + V_H[\tilde{n}_{Zc}] + V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] \\
&= V_H[\tilde{n} + \hat{n} + \tilde{n}_{Zc}] + V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] \\
&= \tilde{V}_{\text{eff}}
\end{aligned} \tag{A.60}$$

The ρ_{ij} enter only via the compensation charge $\hat{n}$ which is defined through

$$\hat{n} = \sum_{i,j,L} \rho_{ij} \hat{Q}_{ij}^L(\mathbf{r}) \tag{A.61}$$

$$\begin{aligned}
\frac{\partial \tilde{E}}{\partial \rho_{ij}} &= \int d^3r \; \underbrace{\frac{\partial \tilde{E}}{\partial \hat{n}}}_{\tilde{V}_{\text{eff}}} \frac{\partial \hat{n}}{\partial \rho_{ij}} = \int d^3r \; \tilde{V}_{\text{eff}}(\mathbf{r}) \frac{\partial}{\partial \rho_{ij}} \sum_{i,j,L} \rho_{ij} \hat{Q}_{ij}^L(\mathbf{r}) \\
&= \sum_L \int d^3r \; \tilde{V}_{\text{eff}}(\mathbf{r}) \hat{Q}_{ij}^L(\mathbf{r}) = \hat{D}_{ij}.
\end{aligned} \tag{A.62}$$

In E^1 and $\tilde{E}^1$ only terms containing ρ_{ij} are important. ρ_{ij} enters either directly or through $\hat{n}$, n^1 or $\tilde{n}^1$ (for details see Eq. A.31).

$$n^1 = \sum_{i,j} \rho_{ij} \langle \phi_i |\, r \rangle \langle r |\, \phi_j \rangle \tag{A.63}$$

$$\tilde{n}^1 = \sum_{i,j} \rho_{ij} \langle \tilde{\phi}_i | r \rangle \langle r | \tilde{\phi}_j \rangle \tag{A.64}$$

$$\begin{aligned} E^1 = & \sum_{i,j} \rho_{ij} \langle \phi_i | \, T \, | \phi_j \rangle + \overline{E_H[n^1]} \\ & + \int V_H[n_{Zc}] n^1(\mathbf{r}) \; d^3r + \overline{E_{xc}[n^1 + n_c]} \end{aligned} \tag{A.65}$$

$$\begin{aligned} \frac{\partial E^1}{\partial \rho_{ij}} = & \langle \phi_i | \, T \, | \phi_j \rangle + \int d^3r \; \frac{\partial}{\partial n^1} \left[\overline{E_H[n^1]} \right. \\ & \left. + \int V_H[n_{Zc}] n^1(\mathbf{r}) \; d^3r + \overline{E_{xc}[n^1 + n_c]} \right] \frac{\partial n^1}{\partial \rho_{ij}} \\ = & \langle \phi_i | \, T \, | \phi_j \rangle + \int d^3r \;\; \left[V_H[n^1](\mathbf{r}) \right. \\ = & \qquad\qquad \left. + V_H[n_{Zc}] + V_{xc}[n^1 + n_c] \right] \langle \phi_i |\, r \rangle \langle r |\, \phi_j \rangle \\ = & \langle \phi_i | \, T + V_H[n^1 + n_{Zc}] + V_{xc}[n^1 + n_c] \, | \phi_j \rangle \\ = & D^1_{ij} \end{aligned} \tag{A.66}$$

$$\begin{aligned} \tilde{E}^1 = & \sum_{i,j} \rho_{ij} \langle \tilde{\phi}_i | T | \tilde{\phi}_j \rangle + \overline{E_H[\tilde{n}^1 + \hat{n}]} \\ & + \int V_H[\tilde{n}_{Zc}] [\tilde{n}^1(\mathbf{r}) + \hat{n}(\mathbf{r})] \; d^3r + \overline{E_{xc}[\tilde{n}^1 + \hat{n} + \tilde{n}_c]} \end{aligned} \tag{A.67}$$

$$\begin{aligned} \frac{\partial \tilde{E}^1}{\partial \rho_{ij}} = & \langle \tilde{\phi}_i | T | \tilde{\phi}_j \rangle + \int d^3r \; \frac{\partial}{\partial \tilde{n}^1} \left\{ \overline{E_H[\tilde{n}^1 + \hat{n}]} \right. \\ & \left. + \int V_H[\tilde{n}_{Zc}] [\tilde{n}^1(\mathbf{r}) + \hat{n}(\mathbf{r})] \; d^3r + \overline{E_{xc}[\tilde{n}^1 + \hat{n} + \tilde{n}_c]} \right\} \frac{\partial \tilde{n}^1}{\partial \rho_{ij}} \\ & + \int d^3r \; \frac{\partial}{\partial \hat{n}} \left\{ \overline{E_H[\tilde{n}^1 + \hat{n}]} + \int V_H[\tilde{n}_{Zc}] [\tilde{n}^1(\mathbf{r}) + \hat{n}(\mathbf{r})] \; d^3r \right. \end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad + \overline{E_{xc}[\tilde{n}^1 + \hat{n} + \tilde{n}_c]} \Big\} \frac{\partial \hat{n}}{\partial \rho_{ij}} \\
&= \langle \tilde{\phi}_i | T | \tilde{\phi}_j \rangle \\
&\quad + \int d^3r \ \left\{ V_H[\tilde{n}^1 + \hat{n} + \tilde{n}_{Zc}] + V_{xc}[\tilde{n}^1 + \hat{n} + \tilde{n}_c] \right\} \langle \tilde{\phi}_i | r \rangle \langle r | \tilde{\phi}_j \rangle \\
&\qquad + \int d^3r \ \underbrace{\left\{ V_H[\tilde{n}^1 + \hat{n} + \tilde{n}_{Zc}] + V_{xc}[\tilde{n}^1 + \hat{n} + \tilde{n}_c] \right\}}_{\tilde{V}^1_{\text{eff}}} \sum_L \hat{Q}^L_{ij}(\mathbf{r}) \\
&= \langle \tilde{\phi}_i | T + \tilde{V}^1_{\text{eff}} | \tilde{\phi}_j \rangle + \sum_L \int d^3r \ \tilde{V}^1_{\text{eff}}(\mathbf{r}) \hat{Q}^L_{ij}(\mathbf{r}) \\
&= \tilde{D}^1_{ij}
\end{aligned}
\tag{A.68}
$$

Inserting all results in Eq. (A.58) finally yields

$$
H = T + \tilde{V}_{\text{eff}} + \sum_{i,j} |\tilde{p}_i\rangle \left(\hat{D}_{ij} + D^1_{ij} - \tilde{D}^1_{ij} \right) \langle \tilde{p}_j | . \tag{A.69}
$$

A.2 Double counting corrections

The total energy can also be expressed as the sum of the Kohn-Sham eigenvalues plus double counting corrections. In the following part, we will derive the energy expression resulting from summation of the Kohn-Sham eigenvalues. Comparison with Eq. (A.46) yields the corrections needed to obtain the exact total energy.

$$
\begin{aligned}
&\sum_n f_n \langle \tilde{\Psi}_n | H | \tilde{\Psi}_n \rangle = \\
&\quad = \sum_n f_n \langle \tilde{\Psi}_n | \left[T + \tilde{V}_{\text{eff}} + \sum_{i,j} |\tilde{p}_i\rangle \left(\hat{D}_{ij} + D^1_{ij} - \tilde{D}^1_{ij} \right) \langle \tilde{p}_j | \right] | \tilde{\Psi}_n \rangle \\
&\quad = \sum_n f_n \langle \tilde{\Psi}_n | T | \tilde{\Psi}_n \rangle + \sum_n f_n \langle \tilde{\Psi}_n | V_H[\tilde{n} + \hat{n} + \tilde{n}_{Zc}] + V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] | \tilde{\Psi}_n \rangle \\
&\qquad + \sum_n \sum_{i,j} f_n \langle \tilde{\Psi}_n | \tilde{p}_i \rangle \left(\hat{D}_{ij} + D^1_{ij} - \tilde{D}^1_{ij} \right) \langle \tilde{p}_j | \tilde{\Psi}_n \rangle \\
&\quad = \sum_n f_n \langle \tilde{\Psi}_n | T | \tilde{\Psi}_n \rangle + \sum_n f_n \langle \tilde{\Psi}_n | V_H[\tilde{n} + \hat{n} + \tilde{n}_{Zc}] + V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] | \tilde{\Psi}_n \rangle \\
&\qquad + \sum_{i,j} \rho_{ij} \sum_L \int \tilde{V}_{\text{eff}}(\mathbf{r}) \hat{Q}^L_{ij}(\mathbf{r}) d\mathbf{r} + \sum_{i,j} \rho_{ij} \left\langle \phi_i \right| T + V^1_{\text{eff}} \left| \phi_j \right\rangle \\
&\qquad - \sum_{i,j} \rho_{ij} \langle \tilde{\phi}_i | T + \tilde{V}^1_{\text{eff}} | \tilde{\phi}_j \rangle - \sum_{i,j} \rho_{ij} \sum_L \int \tilde{V}^1_{\text{eff}}(\mathbf{r}) \hat{Q}^L_{ij}(\mathbf{r}) d\mathbf{r}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_n f_n \langle \tilde{\Psi}_n | T | \tilde{\Psi}_n \rangle + \int \tilde{n} \left(V_H[\tilde{n} + \hat{n} + \tilde{n}_{Zc}] + V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] \right) d\mathbf{r} \\
&\quad + \sum_{i,j} \rho_{ij} \left\langle \phi_i \right| T \left| \phi_j \right\rangle + \int n^1 \left(V_H[n^1 + n_{Zc}] + V_{xc}[n^1 + n_c] \right) d\mathbf{r} \\
&\quad - \sum_{i,j} \rho_{ij} \langle \tilde{\phi}_i | T | \tilde{\phi}_j \rangle - \int \tilde{n}^1 \left(V_H[\tilde{n}^1 + \hat{n} + \tilde{n}_{Zc}] + V_{xc}[\tilde{n}^1 + \hat{n} + \tilde{n}_c] \right) d\mathbf{r} \\
&\quad + \sum_{i,j} \rho_{ij} \sum_L \int \left[\tilde{V}_{\text{eff}}(\mathbf{r}) - \tilde{V}^1_{\text{eff}}(\mathbf{r}) \right] \hat{Q}^L_{ij}(\mathbf{r}) d\mathbf{r}
\end{aligned} \tag{A.70}
$$

Comparing to Eq. (A.46) we immediately see that the last term is "counted twice" and that the kinetic energy contributions are correct. As in standard DFT we have to take a closer look at the terms involving the potential. Therefore we will discuss each of the three terms separately.

$$
\underbrace{\int \tilde{n}\ V_H[\tilde{n} + \hat{n} + \tilde{n}_{Zc}] d\mathbf{r}}_{\text{a}} + \underbrace{\int \tilde{n}\ V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] d\mathbf{r}}_{\text{b}} =
$$

$$
\begin{aligned}
a &= \int (\tilde{n} + \hat{n} - \hat{n}) \left(V_H[\tilde{n} + \hat{n}] + V_H[\tilde{n}_{Zc}] \right) d\mathbf{r} \\
&= \int (\tilde{n} + \hat{n}) V_H[\tilde{n} + \hat{n}] d\mathbf{r} - \int \hat{n}\ V_H[\tilde{n} + \hat{n} + \tilde{n}_{Zc}] d\mathbf{r} \\
&\qquad + \int (\tilde{n} + \hat{n}) V_H[\tilde{n}_{Zc}] d\mathbf{r} \\
&= 2E_H[\tilde{n} + \hat{n}] - \sum_{i,j,L} \rho_{ij} \int V_H[\tilde{n} + \hat{n} + \tilde{n}_{Zc}] \hat{Q}^L_{ij} d\mathbf{r} \\
&\qquad + \int (\tilde{n} + \hat{n}) V_H[\tilde{n}_{Zc}] d\mathbf{r}
\end{aligned} \tag{A.71}
$$

$$
\begin{aligned}
b &= \int \tilde{n}\ V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] d\mathbf{r} - \int (\tilde{n} + \hat{n}) V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] d\mathbf{r} \\
&= - \int \hat{n}\ V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] d\mathbf{r} \\
&= - \sum_{i,j,L} \rho_{ij} \int V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_{Zc}] \hat{Q}^L_{ij} d\mathbf{r}
\end{aligned} \tag{A.72}
$$

Both terms in a and b involving the compensation charge $\hat{n}$ can be added and cancel the corresponding $\sum_{i,j,L} \rho_{ij} \int \tilde{V}_{\text{eff}} \hat{Q}^L_{ij} d\mathbf{r}$. Therefore the double counting correction for the first term is given by

$$
\tilde{E}_{\text{dc}} = -E_H[\tilde{n} + \hat{n}] + E_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] - \int (\tilde{n} + \hat{n}) V_{xc}[\tilde{n} + \hat{n} + \tilde{n}_c] d\mathbf{r}. \tag{A.73}
$$

For the all electron contributions we obtain:

$$\begin{aligned}\int n^1 \left(V_H[n^1+n_{Zc}]+V_{xc}[n^1+n_c]\right) d\mathbf{r} = \\ = \int n^1\ V_H[n^1]d\mathbf{r} + \int n^1\ V_H[n_{Zc}]d\mathbf{r} + \int n^1\ V_{xc}[n^1+n_c]d\mathbf{r} \\ = 2E_H[n^1] + \int n^1\ V_H[n_{Zc}]d\mathbf{r} + \int n^1\ V_{xc}[n^1+n_c]d\mathbf{r}, \end{aligned} \tag{A.74}$$

and the correction is

$$\hat{E}_{dc} = -E_H[n^1] + E_{xc}[n^1+n_c] - \int n^1\ V_{xc}[n^1+n_c]d\mathbf{r}. \tag{A.75}$$

The third term is similar to the first term, just exchange $\tilde{n}$ with $\tilde{n}^1$, additionally it cancels the other term involving the compensation charge.

$$\tilde{E}^1_{\rm dc} = -E_H[\tilde{n}^1+\hat{n}]+E_{xc}[\tilde{n}^1+\hat{n}+\tilde{n}_c]-\int \left(\tilde{n}^1+\hat{n}\right) V_{xc}[\tilde{n}^1+\hat{n}+\tilde{n}_c]d\mathbf{r} \tag{A.76}$$

Finally the total energy expression can be cast in the following form

$$E = \sum_n f_n \langle\tilde{\Psi}_n|H|\tilde{\Psi}_n\rangle + \tilde{E}_{\rm dc} + \hat{E}_{dc} - \tilde{E}^1_{\rm dc}. \tag{A.77}$$

A.3 Reconstruction of the strength parameter for atomic calculations (VASP peculiarities)

In the Hamilton operator Eq. (A.69), three contributions to the total strength parameter are present. One part originates from the compensation part, one from the all electron partial and one from the pseudo partial contributions

$$D_{ij} = \hat{D}_{ij} + D^1_{ij} - \tilde{D}^1_{ij}, \tag{A.78}$$

where each part is given by:

$$\begin{aligned}\hat{D}_{ij} &= \sum_L \int d^3r\ \tilde{V}_{\rm eff}(\mathbf{r})\hat{Q}^L_{ij}(\mathbf{r}) \\ D^1_{ij} &= \langle\phi_i|\,T + V^1_{\rm eff}\,|\phi_j\rangle \\ \tilde{D}^1_{ij} &= \langle\tilde{\phi}_i|T + \tilde{V}^1_{\rm eff}|\tilde{\phi}_j\rangle + \sum_L \int d^3r\ \tilde{V}^1_{\rm eff}(\mathbf{r})\hat{Q}^L_{ij}(\mathbf{r}).\end{aligned}$$

On the `POTCAR` file an ionic D_{ij}^{ion} is stored, which is defined as

$$D_{ij}^{\text{ion}} = \langle\phi_i|\,T\,|\phi_j\rangle - \langle\tilde{\phi}_i|T|\tilde{\phi}_j\rangle + \langle\phi_i|\,V_{\text{atom}}^1\,|\phi_j\rangle - \langle\tilde{\phi}_i|\tilde{V}_{\text{atom},}^1|\tilde{\phi}_j\rangle - \sum_L \int d^3r\ \tilde{V}_{\text{atom}}^1(\mathbf{r})\hat{Q}_{ij}^L(\mathbf{r}). \tag{A.79}$$

The required D_{ij} (A.78) can easily be obtained as

$$D_{ij} = D_{ij}^{\text{ion}} + \sum_L \int d^3r\ \tilde{V}_{\text{eff}}(\mathbf{r})\hat{Q}_{ij}^L(\mathbf{r}) + \langle\phi_i|\,V_{\text{eff}}^1 - V_{\text{atom}}^1\,|\phi_j\rangle - \langle\tilde{\phi}_i|\tilde{V}_{\text{eff}}^1 - \tilde{V}_{\text{atom},}^1|\tilde{\phi}_j\rangle - \sum_L \int d^3r\ \left(\tilde{V}_{\text{eff}}^1 - \tilde{V}_{\text{atom}}^1\right)\hat{Q}_{ij}^L(\mathbf{r}). \tag{A.80}$$

In general, the effective potentials $\tilde{V}_{\text{eff}}^1$ and V_{eff}^1 might be different from $\tilde{V}_{\text{atom}}^1$ and V_{atom}^1. The terms including the effective and atomic potentials are simple one-centre terms, due to the fact, that their contributions are only contributing strictly within the PAW sphere on their atom

$$\rho_{ij} = \sum_n f_n\langle\tilde{\Psi}_n|\tilde{p}_i\rangle\langle\tilde{p}_j|\tilde{\Psi}_n\rangle \Rightarrow n^1 = \sum_{i,j}\rho_{ij}\,\langle\phi_i|\,r\rangle\,\langle r|\,\phi_j\rangle \Rightarrow V_{\text{eff}}^1[n^1]. \tag{A.81}$$

In the special case of atomic calculations, the terms involving the difference between effective and atomic potential should not contribute. If a confinement potential is introduced, these terms are non vanishing.

Appendix B

Spherical harmonics

B.1 Real-valued spherical harmonic basis functions

Spherical harmonic basis functions can be defined in various different ways depending on the derivation or normalisation requirements. Here we use the real spherical harmonics, which are defined in the following way

$$Y_{lm}(\phi,\theta) = \Theta_{l|m|}(\theta)\Phi_m(\phi) \tag{B.1}$$

with

$$\Theta_{lm}(\theta) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\theta) \tag{B.2}$$

$$\Phi_m(\phi) = \begin{cases} \sqrt{2}\cos(m\phi), & m > 0 \\ 1, & m = 0 \\ \sqrt{2}\sin(|m|\phi), & m < 0. \end{cases} \tag{B.3}$$

The (real) spherical harmonics are orthonormal when integrated over the unit sphere.

$$\int_0^{\pi} d\theta \ \sin\theta \int_0^{2\pi} d\phi \ Y_{lm}(\phi,\theta)Y_{l'm'}(\phi,\theta) = \delta_{ll'}\delta_{mm'} \tag{B.4}$$

B.2 Expansion of a plane wave in real spherical harmonics and Bessel functions

A plane wave can be conveniently expanded in a summation of spherical harmonics times Bessel functions. For a detailed derivation see Schwabl, QM 1, page 325. Only real spherical harmonics are required:

$$e^{i\mathbf{kr}} = 4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{l} i^l j_l(kr)Y_{lm}(\Omega_{\mathbf{k}})Y_{lm}(\Omega_{\mathbf{r}}), \tag{B.5}$$

with $\Omega_{\mathbf{r}}$ being the angular dependence in spherical coordinates and $r = |\mathbf{r}|$ the radial component of $\mathbf{r}$. The same holds for $\mathbf{k}$.

B.3 Spherical Bessel transformation

Forward transformation is given by

$$\bar{\beta}_l(q) = 4\pi \int_0^\infty j_l(qr)\ \beta_l(r)\ r^2 dr. \tag{B.6}$$

The backward transformation is given by

$$\beta_l(r) = \frac{1}{2\pi^2} \int_0^\infty j_l(qr)\ \bar{\beta}_l(q)\ q^2 dq, \tag{B.7}$$

and connected via the orthogonality relation of spherical Bessel functions

$$\int_0^\infty r^2 j_l(ur) j_l(vr) dr = \frac{\pi}{2u^2}\delta(u-v). \tag{B.8}$$

$$\begin{aligned}
\beta_l(r) &= \frac{1}{2\pi^2} \int_0^\infty j_l(qr)\ 4\pi \int_0^\infty j_l(qr')\ \beta_l(r')\ r'^2 dr' q^2 dq && \text{(B.9)}\\
&= \frac{4\pi}{2\pi^2} \int_0^\infty dr' \int_0^\infty dq\ q^2\ j_l(qr)\ j_l(qr')\ \beta_l(r')\ r'^2 && \text{(B.10)}\\
&= \frac{2}{\pi}\frac{\pi}{2r^2} \int_0^\infty dr'\ \delta(r-r')\beta_l(r')r'^2 && \text{(B.11)}\\
&= \frac{1}{r^2}\beta_l(r)\ r^2 && \text{(B.12)}\\
&= \beta_l(r) && \text{(B.13)}
\end{aligned}$$

B.3.1 Solution of the radial kinetic energy expression

When solving the Helmholtz equation in spherical coordinates by separation of variables, the radial equation has the form:

$$x^2 \frac{d^2}{dx^2} y(x) + 2x \frac{d}{dx} y(x) + [x^2 - n(n+1)] y(x) = 0. \tag{B.14}$$

This can easily be reformulated to the form of the kinetic energy known from the radial Schrödinger equation by using a variable substitution $x = q_n \cdot r$.

$$\left[-\frac{1}{r^2}\frac{d}{dr} r^2 \frac{d}{dr} + \frac{l(l+1)}{r^2} \right] j_l(q_n r) = q_n^2\ j_l(q_n r) \tag{B.15}$$

Hence the spherical Bessel functions are eigenfunctions of the kinetic energy operator in spherical coordinates.

B.4 Clebsch-Gordan coefficients

If the total angular momentum states is expanded in the uncoupled basis of $|lm\rangle$ and $|l'm'\rangle$ one get

$$|LM\rangle = \sum_{m=-l}^{l} \sum_{m'=-l'}^{l'} \left|lml'm'\right\rangle \left\langle lml'm'|LM\right\rangle . \tag{B.16}$$

The expansion coefficients $\langle lml'm'|LM\rangle$ are called Clebsch-Gordan coefficients.

$$\begin{aligned}
\left\langle lml'm'|LM\right\rangle &= \int d^3q \ \left\langle lml'm'|q\right\rangle \langle q|LM\rangle \\
&= \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi \ Y^*_{lm}(\Omega_{\mathbf{q}}) Y^*_{l'm'}(\Omega_{\mathbf{q}}) Y_{LM}(\Omega_{\mathbf{q}}) \\
&= \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi \ Y_{lm}(\Omega_{\mathbf{q}}) Y_{l'm'}(\Omega_{\mathbf{q}}) Y_{LM}(\Omega_{\mathbf{q}})
\end{aligned} \tag{B.17}$$

In the last line we used the fact that only real spherical harmonics are used. It can be derived that L must satisfy the triangular condition

$$|l - l'| \leq L \leq l + l'. \tag{B.18}$$

Appendix C

Spherical Bessel functions

C.1 Definition and properties

The spherical Bessel functions are special cases of the ordinary Bessel functions with half-integer order and defined as

$$j_n(z) = \sqrt{\frac{\pi}{2z}} J_{n+\frac{1}{2}}(z) \tag{C.1}$$

and fulfill the following orthogonality relation

$$\int_0^\infty dx\ x^2\ j_\alpha(ux)\ j_\alpha(vx) = \frac{\pi}{2u^2}\delta(u-v) \tag{C.2}$$

Using the above equation we can rewrite the orthogonality relation in terms of the general Bessel functions

$$\int_0^\infty dx\ x^2\ j_\alpha(ux)\ j_\alpha(vx) = \frac{\pi}{2} \int_0^\infty dx\ x\ J_{\alpha+\frac{1}{2}}(x)\ J_{\alpha+\frac{1}{2}}(x). \tag{C.3}$$

C.2 r-confined orthogonality relation

Within the analytic calculation of the kinetic energy operator we obtain terms which are similar to the orthogonality relation of the spherical Bessel functions, but where the integration constant is not infinity but a root of the Bessel function. Therefore it is important to calculate a relation for this expression. The following relation of the general Bessel functions is needed in the following where a and b are zeros of the Bessel function J_p:

$$\int_0^1 dx\ x\ J_p(ax)\ J_p(bx) = \begin{cases} 0 & a \neq b \\ \frac{1}{2}\left[J_p'(a)\right]^2 = \frac{1}{2}\left[J_{p-1}(a)\right]^2 = \frac{1}{2}\left[J_{p+1}(a)\right]^2 & a = b \end{cases} \tag{C.4}$$

Lets start, l is the angular momentum quantum number and q_t are chosen such that $j_l(q_t \cdot R_{\max})$ is zero

$$\begin{aligned}
\int_0^{R_{\max}} dx\ x^2 j_l(q_t x) j_l(q_{t'} x) &= \frac{\pi}{2q_t} \int_0^{R_{\max}} dx\ x\ J_{l+\frac{1}{2}}(q_t x)\ J_{l+\frac{1}{2}}(q_{t'} x) \\
&\overset{x'=\frac{x}{R_{\max}}}{=} \frac{\pi R_{\max}^2}{2q_t} \int_0^1 dx'\ x'\ J_{l+\frac{1}{2}}(\underbrace{q_t R_{\max}}_{a} x') \\
&\qquad \cdot J_{l+\frac{1}{2}}(\underbrace{q_{t'} R_{\max}}_{a'} x') \\
&= \frac{\pi R_{\max}^2}{2q_t} \int_0^1 dx\ x\ J_{l+\frac{1}{2}}(ax)\ J_{l+\frac{1}{2}}(a' x) \\
&\overset{a=a'}{=} \frac{\pi R_{\max}^2}{2q_t}\ \frac{1}{2} \left[J_{l+\frac{3}{2}}(q_t R_{\max}) \right]^2 \\
&= \frac{\pi R_{\max}^2}{4q_t}\ \frac{2q_t R_{\max}}{\pi} j_{l+1}^2(q_t R_{\max}) \\
&= \frac{R_{\max}^3}{2}\ j_{l+1}^2(q_t R_{\max}). \qquad \text{(C.5)}
\end{aligned}$$

In line four of the derivation we assumed that $a = a'$ because otherwise it follows from Eq. (C.4), that the term is zero. We finally get an orthogonality relation, where the $q_t \cdot R_{\max}$ have to be zeros of the spherical Bessel function.

$$\int_0^{R_{\max}} dx\ x^2 j_l(q_t x) j_l(q_{t'} x) = \begin{cases} 0 & q_t \neq q_{t'} \\ \frac{R_{\max}^3}{2}\ j_{l+1}^2(q_t R_{\max}) & q_t = q_{t'} \end{cases} \qquad \text{(C.6)}$$

To write this in a more convenient way we write it in the style of a Kronecker delta to symbolize that it is only non zero for equal q_t and $q_{t'}$. (For non integer q_t the Dirac delta function is usually a distribution and has only a concise meaning in an integral.)

$$\int_0^{R_{\max}} dx\ x^2 j_l(q_t x) j_l(q_{t'} x) = \frac{R_{\max}^3}{2}\ j_{l+1}^2(q_t R_{\max})\ \bar{\delta}_{q_t q_{t'}} \qquad \text{(C.7)}$$

Bibliography

[1] P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Phys. Rev. B **136**, 864 (1964).

[2] W. Kohn and L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev. A **140**, 1133 (1965).

[3] D. M. Ceperley and B. J. Alder, *Ground state of the electron gas by a stochastic method*, Phys. Rev. Lett. **45**, 566 (1980).

[4] U. von Barth and L. Hedin, *A local exchange-correlation potential for the spin polarized case: I*, J. Phys. C **5**, 1629 (1972).

[5] J. P. Perdew and A. Zunger, *Self-interaction correction to density-functional approximations for many-electron systems*, Phys. Rev. B **23**, 5048 (1981).

[6] S. H. Vosko, L. Wilk, and M. Nusair, *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, Can. J. Phys. **58**, 1200 (1980).

[7] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple*, Phys. Rev. Lett. **77**, 3865 (1996).

[8] Z. Wu and R. E. Cohen, *More accurate generalized gradient approximation for solids*, Phys. Rev. B **73**, 235116 (2006).

[9] J. P. Perdew, A. Ruzsinsky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, *Restoring the density-gradient expansion for exchange in solids and surfaces*, Phys. Rev. Lett. **100**, 136406 (2008).

[10] L. Schimka, J. Harl, and G. Kresse, *Improved hybrid functional for solids: The HSEsol functional*, To be published.

[11] R. Amiento and A. E. Mattsson, *Functional designed to include surface effects in self-consistent density functional theory*, Phys. Rev. B **72**, 085108 (2005).

[12] A. E. Mattsson, R. Armiento, J. Paier, G. Kresse, J. M. Wills, and T. R. Mattsson, *The AM05 density functional applied to solids*, J. Chem. Phys. **128**, 084714 (2008).

[13] M. Shishkin, M. Marsman, and G. Kresse, *Accurate quasiparticle spectra from self-consistent GW calculations with vertex corrections*, Phys. Rev. Lett. **99**, 246403 (2007).

[14] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields*, J. Phys. Chem. **98**, 11623 (1994).

[15] C. Adamo and V. Barone, *Toward reliable density functional methods without adjustable parameters: The PBE0 model*, J. Chem. Phys. **110**, 6158 (1999).

[16] M. Ernzerhof and G. E. Scuseria, *Assessment of the Perdew-Burke-Ernzerhof exchange-correlation functional*, J. Chem. Phys. **110**, 5029 (1999).

[17] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, *Influence of the exchange screening parameter on the performance of screened hybrid functionals*, J. Chem. Phys. **125**, 224106 (2006).

[18] J. Muscat, A. Wander, and N. M. Harrison, *On the prediction of band gaps from hybrid functional theory*, Chem. Phys. Lett. **342**, 397 (2001).

[19] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, *Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional*, J. Chem. Phys. **123**, 174101 (2005).

[20] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, *Screened hybrid density functionals applied to solids*, J. Chem. Phys. **124**, 154709 (2006) J. Chem. Phys. **125**, 249901(E) (2006).

[21] J. Paier, M. Marsman, and G. Kresse, *Why does the B3LYP hybrid functional fail for metals?*, J. Chem. Phys. **127**, 024103 (2007).

[22] J. P. Perdew, M. Ernzerhof, and K. Burke, *Rationale for mixing exact exchange with density functional approximations*, J. Chem. Phys. **105**, 9982 (1996).

[23] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *Hybrid functionals based on a screened Coulomb potential*, J. Chem. Phys **118**, 8207 (2003).

[24] P. E. Blöchl, *Projector augmented-wave method*, Phys. Rev. B **50**, 17953 (1994).

[25] G. Kresse and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method*, Phys. Rev. B **59**, 1758 (1999).

[26] G. B. Bachelet, D. R. Hamann, and M. Schlüter, *Pseudopotentials that work: From H to Pu*, Phys. Rev. B **26**, 4199 (1982).

[27] D. Vanderbilt, *Optimally smooth norm-conserving pseudopotentials*, Phys. Rev. B **32**, 8412 (1985).

[28] N. Troullier and J. L. Martins, *Efficient pseudopotentials for plane-wave calculations*, Phys. Rev. B **43**, 1993 (1991).

[29] A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, *Optimized pseudopotentials*, Phys. Rev. B **41**, 1227 (1990).

[30] D. R. Hamann, M. Schlüter, and C. Chiang, *Norm-conserving pseudopotentials*, Phys. Rev. Lett. **43**, 1494 (1979).

[31] M. Liu, T. R. Finlayson, and T. F. Smith, *High-resolution dilatometry measurements of $SrTiO_3$ along cubic and tetragonal axes*, Phys. Rev. B **55**, 3480 (1997).

[32] L. Cao, E. Sozontov, and J. Zegenhagen, *Cubic to tetragonal phase transition of $SrTiO_3$ under epitaxial stress: A X-ray backscattering study*, Phys. stat. sol (a) **181**, 387 (2000).

[33] G. Shirane and Y. Yamada, *Lattice-dynamical study of the 110 K phase transition of $SrTiO_3$*, Phys. Rev. **177**, 858 (1969).

[34] K. A. Müller and H. Burkhard, *$SrTiO_3$: An intrinsic quantum paraelectric below 4 K*, Phys. Rev. B **19**, 3593 (1979).

[35] G. Rupprecht and R. O. Bell, *Microwave losses in strontium titanate above the phase transition*, Phys. Rev. **125**, 1915 (1962).

[36] R. Viana, P. Lunkenheimer, J. Hemberger, R. Böhmer, and A. Loidl, *Dielectric spectroscopy in $SrTiO_3$*, Phys. Rev. B **50**, 601 (1994).

[37] H. Uwe and T. Sakudo, *Stress-induced ferroelectricity and soft phonon modes in $SrTiO_3$*, Phys. Rev. B **13**, 271 (1976).

[38] M. Itoh, R. Wang, Y. Inaguma, T. Yamaguchi, Y-J. Shan, and T. Nakamura, *Ferroelectricity induced by oxygen isotope exchange in strontium titanate perovskite*, Phys. Rev. Lett. **82**, 3540 (1999).

[39] W. Jauch and A. Palmer, *Anomalous zero-point motion in $SrTiO_3$: Results from γ-ray diffraction*, Phys. Rev. B **60**, 2961 (1999).

[40] W. Zhong and D. Vanderbilt, *Effect of quantum fluctuations on structural phase transitions in $SrTiO_3$ and $BaTiO_3$*, Phys. Rev. B **53**, 5074 (1996).

[41] G. H. Kwei, A. C. Lawson, and S. J. L. Billinge, *Structures of the ferroelectric phase of barium titanate*, J. Phys. Chem. **97**, 2368 (1993).

[42] G. Kresse and J. Furthmüller, *Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set*, Comp. Mat. Sci. **6**, 15 (1996).

[43] G. Kresse and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*, Phys. Rev. B **54**, 11169 (1996).

[44] H. J. Monkhorst and J. D. Pack, *Special points for Brillouin-zone integrations*, Phys. Rev. B **13**, 5188 (1976).

[45] D. I. Bilc, R. Orlando, R. Shaltaf, G.-M. Rignanese, J. Íniguez, and Ph. Ghosez, *Hybrid exchange-correlation functional for accurate prediction of the electronic and structural properties of ferroelectric oxides*, Phys. Rev. B **77**, 165107 (2008).

[46] E. Heifets, E. Kotomin, and V. A. Trepakov, *Calculations for antiferrodistortive phase of $SrTiO_3$ perovskite: hybrid density functional study*, J. Phys. Cond. Matt. **18**, 4845 (2006).

[47] Landolt-Börnstein, *Ferroelectrics and Related Substances, New Series, Vol. 3* Springer Verlag (1969).

[48] E. Heifets, R. I. Eglitis, E. A. Kotomin, J. Maier, and G. Borstel, *Ab initio modeling of surface structure for $SrTiO_3$ perovskite crystals*, Phys. Rev. B **64**, 235417 (2001).

[49] S. Piskunov, E. Heifets, R. I. Eglitis, and G. Borstel, *Bulk properties and electronic structure of $SrTiO_3$, $BaTiO_3$, $PbTiO_3$ perovskites: an ab initio HF/DFT study*, Comp. Mat. Sci. **29**, 165 (2004).

[50] E. Mete, R. Shaltaf, and Ş. Ellialtıoğlu, *Electronic and structural properties of a 4d perovskite: Cubic phase of $SrZrO_3$*, Phys. Rev. B **68**, 035119 (2003).

[51] S. Tinte, M. G. Stachiotti, C. O. Rodriguez, D. L. Novikov, and N. E. Christensen, *Applications of the generalized gradient approximation to ferroelectric perovskites*, Phys. Rev. B **58**, 11959 (1998).

[52] A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A **38**, 3098 (1988).

[53] J. Perdew and Y. Wang, *Accurate and simple analytic representation of the electron-gas correlation energy*, Phys. Rev. B **45**, 13244 (1992).

[54] J. P. Perdew and W. Yue, *Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation*, Phys. Rev. B **33**, 8800 (1986).

[55] R. A. Cowley, *Lattice dynamics and phase transitions of strontium titanate*, Phys. Rev. **134**, A981 (1964).

[56] R. O. Bell and G. Rupprecht, *Elastic constants of strontium titanate*, Phys. Rev. **129**, 90 (1963).

[57] H. Unoki and T. Sakudo, *Electron spin resonance of Fe^{3+} in $SrTiO_3$ with special reference to the 110 K phase transition*, J. Phys. Soc. Jap. **23**, 546 (1967).

[58] B. Alefeld, *Die Messung der Gitterparameteränderung von Strontiumtitanat am Phasenübergang bei 108 K*, Z. Physik A **222**, 155 (1969).

[59] A. Heidemann and H. Wettengel, *Die Messung der Gitterparameteränderung von $SrTiO_3$*, Z. Physik A **258**, 429 (1973).

[60] G. Shirane, H. Danner, and P. Pepinsky, *Neutron diffraction study of orthorhombic $BaTiO_3$*, Phys. Rev. **105**, 856 (1957).

[61] H. H. Wieder, *Electrical behaviour of barium titanate single crystals at low temperatures*, Phys. Rev. **99**, 1161 (1955).

[62] R. D. King-Smith and D. Vanderbilt, *Theory of polarization of crystalline solids*, Phys. Rev. B **47**, 1651 (1993).

[63] R. Resta, *Macroscopic polarization in crystalline dielectrics: the geometric phase approach*, Rev. Mod. Phys. **66**, 899 (1994).

[64] K. van Benthem, C. Elsässer, and R. H. French, *Bulk electronic structure of $SrTiO_3$: Experiment and theory*, J. Appl. Phys. **90**, 6156 (2001).

[65] S. H. Wemple, *Polarization fluctuations and the optical-absorption edge in $BaTiO_3$*, Phys. Rev. B **2**, 2679 (1970).

[66] Ph. Ghosez, D. Desquesnes, X. Gonze, and K. M. Rabe, *First-principle study of lattice instabilities in $Ba_xSr_{1-x}TiO_3$*, arXiv:cond-mat 0004464v1 (2000).

[67] N. Sai and D. Vanderbilt, *First-principles study of ferroelectric and antiferrodistortive instabilities in tetragonal $SrTiO_3$*, Phys. Rev B. **62**, 13942 (2000).

[68] W. Zhong, R. D. King-Smith, and D. Vanderbilt, *Giant LO-TO splittings in perovskite ferroelectrics*, Phys. Rev. Lett. **72**, 3618 (1994).

[69] W. G. Spitzer, Robert C. Miller, D. A. Kleinman, and L. E. Howarth, *Far infrared dielectric dispersion in $BaTiO_3$, $SrTiO_3$, and TiO_2*, Phys. Rev. **126**, 1710 (1962).

[70] W. G. Sterling, *Neutron inelastic scattering study of the lattice dynamics of strontium titanate: Harmonic models*, J. Phys. C **5**, 2711 (1972).

[71] H. Vogt and G. Neumann, *Observation of infrared active and silent modes in cubic crystals by hyper-Raman scattering*, Phys. Stat. Sol. B **92**, 57 (1979).

[72] Ph. Ghosez, X. Gonze, and J.-P. Michenaud, *Coulomb interaction and ferroelectric instability of $BaTiO_3$*, Europhys. Lett. **33**, 713 (1996).

[73] Y. Luspin, J. L. Servoin, and F. Gervais, *Soft mode spectroscopy in barium titanate*, J. Phys. C **13**, 3761 (1980).

[74] K. Hummer, J. Harl, and G. Kresse, *Heyd-Scuseria-Ernzerhof hybrid functional for calculating the lattice dynamics of semiconductors*, Phys. Rev. B **80**, 115205 (2009).

[75] A. P. Scott and L. Radom, *Harmonic vibrational frequencies: An evaluation of Hartree-Fock, Møller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors*, J. Phys. Chem. **100**, 16502 (1996).

[76] V. Srikant and D. R. Clarke, *On the optical band gap of zinc oxide*, J. Appl. Phys. **83**, 5447 (1998).

[77] K. Ellmer, A. Klein, and B. Rech, *Transparent conductive zinc oxide: basics and applications in thin film solar cells* Springer, Berlin (2008).

[78] G. Kresse, O. Dulub, and U. Diebold, *Competing stabilization mechanism for the polar ZnO(0001)-Zn surface*, Phys. Rev. B **68**, 245409 (2003).

[79] M. Valtiner, M. Todorova, G. Grundmeier, and J. Neugebauer, *Temperature stabilized surface reconstructions at polar ZnO(0001)*, Phys. Rev. Lett. **103**, 065502 (2009).

[80] P. W. Tasker, *The stability of ionic crystal surfaces*, J. Phys. C. **12**, 4977 (1979).

[81] C. Noguera, *Polar oxide surfaces*, J. Phys.: Condens. Matter **12**, R367 (2000).

[82] O. Dulub, L. A. Boatner, , and U. Diebold, *STM study of the geometric and electronic structure of ZnO (0001)-Zn, (000$\bar{1}$)-O, (10$\bar{1}$0), and (11$\bar{2}$0) surfaces*, Surf. Sci. **519**, 201 (2002).

[83] J. P. Perdew, J. A. Chevary, S. H. Vosko, Koblar A. Jackson, Mark R. Pederson, D. J. Singh, and C. Fiolhais, *Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation*, Phys. Rev. B **46**, 6671 (1992).

[84] J. P. Perdew, J. A. Chevary, S. H. Vosko, Koblar A. Jackson, Mark R. Pederson, D. J. Singh, and C. Fiolhais, *Erratum: Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation*, Phys. Rev. B **48**, 4978 (1993).

[85] A. E. Mattsson, R. Armiento, P. A. Schultz, and T. R. Mattsson, *Nonequivalence of the generalized gradient approximations PBE and PW91*, Phys. Rev. B **73**, 195123 (2006).

[86] I. G. Batyrev, A. Alavi, and M. W. Finnis, *Equilibrium and adhesion of Nb/sapphire: The effect of oxygen partial pressure*, Phys. Rev. B **62**, 4698 (2000).

[87] K. Reuter and M. Scheffler, *Composition, structure, and stability of RuO_2(110) as a function of oxygen pressure*, Phys. Rev. B **65**, 035406 (2001).

[88] Q. Sun, K. Reuter, and M. Scheffler, *Effect of a humid environment on the surface structure of RuO_2(110)*, Phys. Rev. B **67**, 205424 (2003).

[89] B. Meyer, *First-principles study of the polar O-terminated ZnO surface in thermodynamic equilibrium with oxygen and hydrogen*, Phys. Rev. B **69**, 045416 (2004).

[90] *NIST-JANAF Thermochemical Tables* Washington, DC : American Chemical Society ; New York : American Institute of Physics for the National Institute of Standards and Technology, 4th ed. (1998).

[91] J. V. Lauritsen, S. Porsgaard, M. K. Rasmussen, M. C. R. Jensen, R. Bechstein, K. Meinander, B. S. Clausen, S. Helveg, R. Wahl, G. Kresse, and F. Besenbacher, *Atomic-scale structure and stability of the polar O-terminated ZnO(000$\bar{1}$) surface*, submitted.

[92] M. Kunat, St. Gil Girol, Th. Becker, U. Burghaus, and Ch. Wöll, *Stability of the polar surfaces of ZnO: A reinvestigation using He-atom scattering*, Phys. Rev. B **66**, 081402–1 (2002).

[93] M. van Hove and R. Leysen, *LEED study of the polar surfaces of ZnO*, Phys. Stat. Sol. A **9**, 361 (1972).

[94] O. Dulub, U. Diebold, and G. Kresse, *Novel stabilization mechanism on polar surfaces: ZnO(000$\bar{1}$)-Zn*, Phys. Rev. Lett. **90**, 016102 (2003).

[95] S. Torbrügge, F. Ostendorf, and M. Reichling, *Stabilization of zinc-terminated ZnO(0001) by a modified surface stoichiometry*, J. Phys. Chem. C **113**, 4909 (2009).

[96] J. Tersoff and D. R. Hamann, *Theory of the scanning tunneling microscopy*, Phys. Rev. B **31**, 805 (1985).

[97] F. M. F. de Groot, *XANES spectra of transition metal compounds*, Journal of Physics: Conference Series **190**, 1 (2009).

[98] F. M. F. de Groot, *Multiplet effects in X-ray spectroscopy*, Coord. Chem. Rev. **249**, 31 (2005).

[99] L. Schimka, J. Harl, A. Stroppa, A. Grüneis, M. Marsman, F. Mittendorfer, and G. Kresse, *Accurate surface and adsorption energies from many-body perturbation theory*, Nat. Mat. **9**, 741 (2010).

[100] J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejon, and D. Sanchez-Portal, *The SIESTA method for ab initio order-N materials simulation*, J. Phys.: Condens. Matter **14**, 2745 (2002).

[101] D. Porezag, Th. Frauenheim, Th. 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).

[102] O. F. Sankey and D. J. Niklewski, *Ab initio tight-binding model for molecular dynamics simulations and other applications in covalent systems*, Phys. Rev. B **40**, 3979 (1989).

[103] F. Bloch, *Über die Quantenmechanik der Elektronen in Kristallgittern*, Z. Physik **52**, 555 (1928).

[104] N. Bernstein, M. J. Mehl, D. A. Papaconstantopoulos, N. I. Papanicolaou, Martin Z. Bazant, and Efthimios Kaxiras, *Energetic, vibrational, and electronic properties of silicon using a nonorthogonal tight-binding model*, Phys. Rev. B **62**, 4477 (2000).

List of publications

1 S. Janecek, M. S. Kaczmarski, R. Wahl, E. Hernandez and E. Krotscheck, *A new real-space algorithm for realistic density functional calculations*, Eur. Phys. J. D **43**, 173 (2007).

2 R. Wahl, D. Vogtenhuber and G. Kresse, *$SrTiO_3$ and $BaTiO_3$ revisited using the projector augmented wave method: performance of hybrid and semilocal functionals*, Phys. Rev. B **78**, 104116 (2008)

3 A. Nagoya, R. Asahi, R. Wahl and G. Kresse, *Defect formation and phase stability of Cu_2ZnSnS_4 photovoltaic material*, Phys. Rev. B **81**, 113202 (2010).

4 J. V. Lauritsen, M. C. R. Jensen, M. K. Rasmussen, S. Porsgaard, R. Bechstein, B. S. Clausen, S. Helveg, R. Wahl, G. Kresse and F. Besenbacher, *Atomic-scale structure and stability of the polar O-terminated $ZnO(000\bar{1})$ surface*, submitted.

Curriculum vitae

Personal information:

Name:	Dipl.-Ing. Roman Wahl
Date of birth:	May 11, 1982
Place of birth:	Linz
Nationality:	Austria
Family status :	Married

Education:

1988–1992	Primary school: Volksschule Perg
1992–1996	Secondary modern school: Hauptschule II Perg
1996–2000	Grammar school: BORG Perg
2000–2001	Military service
2001–2006	Studies of Technical Physics at the Johannes Keppler University Linz; Diploma Thesis at the Department of Theoretical Physics: *DFT calculations on small sodium and magnesium clusters* under the supervision of Univ.-Prof. Dr. Eckhard Krotscheck
	Graduation (Dipl. Ing.): 22.9.2006
2006–2010	PhD thesis at the Computational Materials Physics group at the University of Vienna: *Ab initio methods applied to oxides, and atomic orbitals in the PAW framework* under the supervision of Univ.-Prof. Dipl.-Ing. Dr. Georg Kresse