THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Proton conductivity of lanthanum and barium zirconate

Microscale aspects on first-principles basis

JOAKIM NYMAN

Department of Applied Physics
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2012

Proton conductivity of lanthanum and barium zirconate
Microscale aspects on first-principles basis
JOAKIM NYMAN
ISBN 978-91-7385-629-4

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 3310
ISSN 0346-718X

Department of Applied Physics
Chalmers University of Technology
SE-412 96 Göteborg, Sweden
Telephone: +46 (0)31–772 1000

The cover picture illustrates concentrations of oxygen vacancies and protons together with corresponding electrostatic potential across a grain boundary of barium zirconate, representing the work in Papers III–V.

Typeset in LaTeX; figures created using Matlab, Xfig, Vesta and Powerpoint.

Chalmers reproservice
Göteborg, Sweden 2012

Proton conductivity of lanthanum and barium zirconate
Microscale aspects on first-principles basis

JOAKIM NYMAN
Department of Applied Physics
Chalmers University of Technology

ABSTRACT

Fuel cells are devices which convert chemical energy into electrical energy cleanly and efficiently. Development of fuel cells compatible with hydrocarbon fuels would make more efficient use of present fossil and renewable fuels, and also enable progress towards a future hydrogen economy. One of the major hindrances to commercially viable fuel cell technologies is the lack of materials with properties appropriate for devices operating at temperatures high enough to allow carbon-containing fuels, while low enough to suppress the negative side effects of high temperatures. Part in a direct route towards intermediate-temperature fuel cell technologies is the optimization of the proton conductivity of solid, ceramic, oxide materials for use as electrolytes in so-called solid oxide fuel cells.

This thesis puts forward theoretical investigations into atomic and microscopic mechanisms which directly influence the proton conductivity of solid oxide materials posing as candidates for proton-conducting electrolyte materials. The foundation of this work is the description of the atomic and electronic structure of materials offered by methods based on density-functional theory. Combined with thermodynamic and electrostatic theory, the pressing issue of grain boundary resistivity in the otherwise promising proton-conducting solid oxide material barium zirconate ($BaZrO_3$), is addressed. Furthermore, fundamental aspects related to the optimization of proton conductivity by means of acceptor-doping are examined in the not as frequently studied material lanthanum zirconate ($La_2Zr_2O_7$). Acceptor-doping is intended to increase proton concentration by causing vacant oxygen positions, which, by incorporation of water molecules, can be filled with hydroxide ions.

The most important work and results can be summarized as follows: (i) By examining several different dopant species in $La_2Zr_2O_7$ it is shown that a poor choice of dopant can not only lead to inefficient concentration improvement but also to trapping of both protons and oxygen vacancies. In consistency with experimental observations, Ca and Sr are pointed to as the most promising dopants out of the twelve investigated species. (ii) By calculating the energy of oxygen vacancies in the vicinity of different grain boundary structures of $BaZrO_3$ it is demonstrated that accumulation of oxygen vacancies at the core of the grain boundary interfaces can significantly hamper the effective proton conductivity in the material. This accumulation leads to charged grain boundary cores and gives rise to a depletion of protons in the surrounding region. The magnitude of the effect corresponds well with experimental conductivity data.

Keywords: solid oxide fuel cell, electrolyte, conductivity, point defect, acceptor, oxygen vacancy, proton, trapping, grain boundary, segregation, space charge, depletion, pyrochlore, perovskite, La2Zr2O7,BaZrO3, first principles, DFT

LIST OF PUBLICATIONS

This thesis consists of an introductory text and the following papers:

I. Protonic defects in pure and doped $La_2Zr_2O_7$ pyrochlore oxide
M. E. Björketun, C. S. Knee, B. J. Nyman and G. Wahnström
Solid State Ionics **178**, 1642 (2008)

II. Substitutional doping and oxygen vacancies in $La_2Zr_2O_7$ pyrochlore oxide
B. J. Nyman, M. E. Björketun and G. Wahnström
Solid State Ionics **189**, 19 (2011)

III. Oxygen vacancy segregation and space-charge effects in grain boundaries of dry and hydrated $BaZrO_3$
B. J. Nyman, E. E. Helgee and G. Wahnström
Applied Physics Letters, doi:10.1063/1.3681169 (2012)

IV. Grain boundaries in BaZrO3: Oxygen vacancy segregation and space charge effects from first principles and atomistic simulations
E. E. Helgee, B. J. Nyman, A. Lindman and G. Wahnström
(in manuscript)

V. Equilibrium defect structure of yttrium-doped barium zirconate with oxygen deficient grain boundary cores
B. J. Nyman, P. Erhart and G. Wahnström
(in manuscript)

Specification of my contribution to the publications:

I. I did the proton barrier calculations and some more computations.

II. I did the majority of the calculations and wrote the paper.

III. I did the calculations and the modeling, and was the main author of the paper.

IV. I constructed the space charge model and wrote corresponding part of the paper.

V. I did a large part of the calculations and the modeling, and wrote the paper.

Contents

Chapter 1

Introduction

In todays fossil fuel economy, the majority of the energy consumed in our society continually comes from the combustion of oil, coal and natural gas. Meanwhile, the general awareness of negative effects in terms of pollution and global warming is steadily increasing, together with the understanding that renewable sources of energy are essential if the energy demands of our growing world population are to be sustainable. While efforts to reduce our dependency on fossil fuels are made in a range of areas, the notion of a future hydrogen economy is a tantalizing prospect.

In a successful hydrogen economy, energy from the Sun is converted into chemical energy via the splitting of water into hydrogen and oxygen. Effectively being stored in a fuel in the form of hydrogen molecules, this energy is then extracted as electricity in the reverse reaction by allowing hydrogen and oxygen to recombine into water in a controlled fashion. Depending on the application, solar energy is harvested either directly on-site where the power is required, or in centralized power plants for further distribution of hydrogen or electricity. The electric power can be used for most applications capable of running on direct-current electricity: from mobile electronic devices to transportation and housing.

In a quintessential incarnation of the hydrogen economy, solar energy is thus transformed into electricity via a closed cycle of hydrogen, oxygen and water. Pollution and production of greenhouse gases is eliminated and the Sun is the direct source of energy. The realization of a perfect hydrogen economy is, however, associated with many technological challenges: from efficient catalysis of the splitting of water, via safe storage, transportation and distribution of hydrogen, to cost-effective and reliable extraction of electric power from the hydrogen. Substantial political and financial commitments are also required in terms of for instance infrastructure in order to facilitate the transition from our current fossil-fuel driven energy economy.

Although implementation of an ideal hydrogen economy is a long-term ambition, development of the fuel cell devices capable of converting chemical energy directly into electricity is currently receiving significant attention (compare figure 1.1). Benefits of thus progressing towards a clean hydrogen economy along with more short-term gains from applying the attractive efficiency of the fuel cell principle on the already established hydrocarbon infrastructure are repeatedly being pointed out [1–7]. The present supply of hydrocarbon fuels and derivatives thereof, renew-

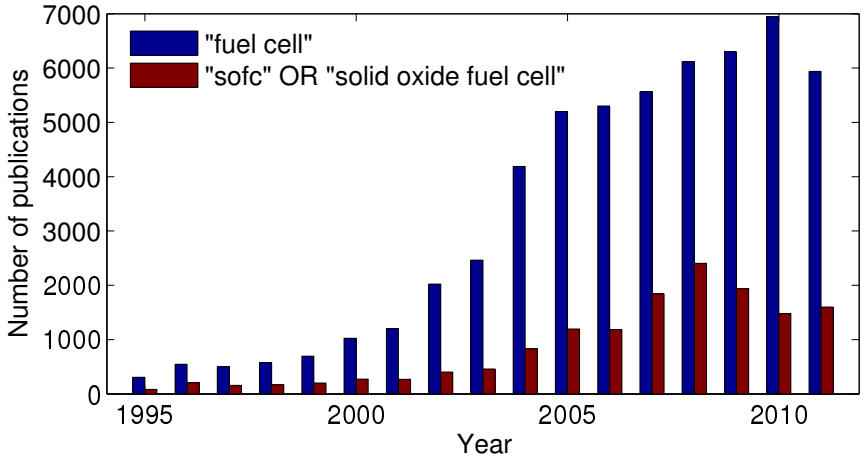


Figure 1.1: Annual publication numbers of scientific papers containing the indicated keywords, as indexed by *www.scopus.com* in October 2011.

able as well as fossil, is in other words believed to be put into more efficient use powering fuel cells instead of wasteful combustion engines. All the while developing clean, sustainable and effective means of producing and storing hydrogen, optimizing fuel cell systems tolerant to carbon-containing fuels in the meantime is thus considered highly desirable.

One class of fuel cells is particularly insensitive to impurities in the fuel: the solid oxide fuel cell. It takes its name from the ceramic oxide material used as the critical electrolyte component. So far the development of commercially viable solid oxide fuel cells have been hindered by issues related to the stability and performance of available materials [2, 8]. This thesis deals with aspects related to some of these issues. Taking advantage of the ability of density functional theory to realistically describe materials on the atomic scale, theoretical and computational methods are employed to study fundamental mechanisms related to the conductivity of two candidate materials for solid oxide fuel cell electrolytes. In doing so the performance with regards to the operating temperature of solid oxide fuel cells is adressed.

The content of the thesis is organized as follows. With the research presented mainly in the appended papers, the summarizing chapters 1-10 are intended to define and introduce the subject and put the results in a wider perspective. This is chapter 1. Chapters 2 and 3 serve as motivation and background to the work, introducing fuel cells and outlining the challenges of materials for solid oxide fuel cell electrolytes. Chapters 4 and 5 introduce the theoretical framework of the investigations, followed by a description of method and models in chapters 6–8. A summary of the papers is then provided in chapter 9, and finally a few routes for related work in the future are proposed in chapter 10.

Chapter 2

Fuel cell systems

Even though the principle of the fuel cell has been known for almost two centuries, it was not put into practical use until the US space program in the sixties and it took nearly three more decades for wide-spread development and commercialization to gather speed. [3] Over the last decade, the interest in the technology has been ever increasing, but still difficulties remain in finding suitable materials and materials processes which strike a viable balance between practicality, cost, efficiency and reliability. All of the content in the present chapter can be found in reviews on fuel cell technology and its status and challenges. [8–16] The intention is to briefly survey the area to provide a background for the research presented in this thesis. Of particular relevance are the effects of operating temperature, solid oxide fuel cells and the conductivity of electrolyte materials.

2.1 Basic principle

In electrolysis an electric current is driven through water, resulting in the splitting of water molecules into hydrogen and oxygen molecules. A fuel cell does the reverse; it combines hydrogen and oxygen to produce water and electricity. Splitting the water molecule stores free energy by rearranging the low-energy configuration of atoms in the water molecules into the higher-energy combination of hydrogen molecules and oxygen molecules. Similarly, the combination of atoms in other hydrogen-containing compounds like hydrocarbons correspond to a relatively speaking high-energy state. The fuel cell regains this energy by chemically facilitating the rearrangement into lower-energy configurations. It is this direct way of converting chemical energy into electricity that gives the fuel cell its advantage in efficiency over conventional heat engines such as petrol engines or steam turbines.

The basic fuel cell principle is outlined in figure 2.1 (a)–(b). Keeping in mind that in general the fuel may derive from hydrocarbons, the principle is illustrated with pure hydrogen for simplicity.[1] In the fuel cell, hydrogen (oxygen) is adsorbed onto the surface of the anode (cathode). Acting as catalysts, the electrodes then

[1] No particular loss in generality is implied since hydrocarbons for use with fuel cells are typically reformed into hydrogen, either separately in external reformers or internally at the fuel cell anode.

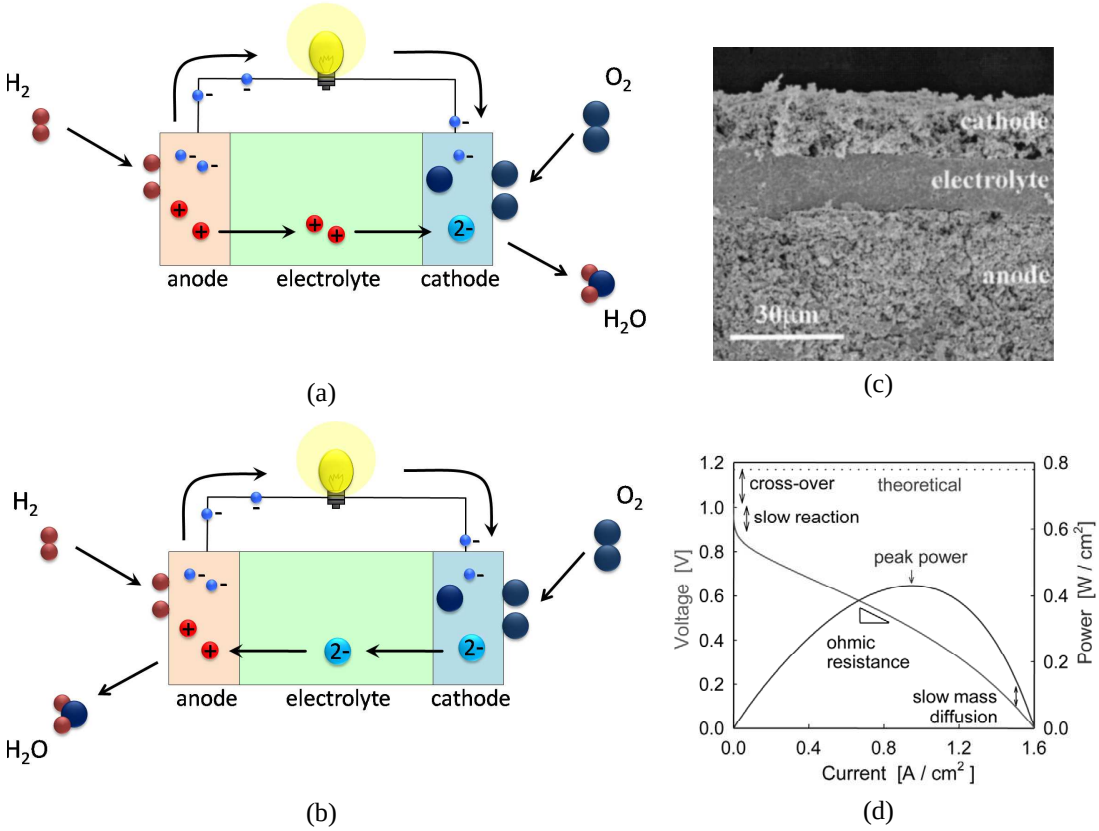


Figure 2.1: The fuel cell. (a)–(b): basic principle, (c): cross-sectional SEM micrograph of a solid oxide fuel cell [17], and (d): schematic performance characteristics (polarization curve) [18].

split respective molecule and incorporate the atoms into its structure where they are ionized. Electrons from the anode flow through the electrical load in the external loop to the cathode, while protons and/or oxygen ions are transported in the electrolyte. Depending on whether the electrolyte conducts oxide ions and/or protons, the components are then assembled into water at the anode and/or the cathode.

The reactions of the hydrogen fuel cell are thus:

$$\text{Full cell:} \quad \mathrm{H_2} + 1/2\,\mathrm{O_2} \longrightarrow \mathrm{H_2O} \tag{2.1}$$

$$\text{Anode:} \quad \mathrm{H_2} \longrightarrow 2\mathrm{H^+} + 2e^- \tag{2.2}$$

$$\text{Cathode:} \quad 1/2\,\mathrm{O_2} + 2e^- \longrightarrow \mathrm{O^{2-}} \tag{2.3}$$

$$\text{Water production:} \quad 2\mathrm{H^+} + \mathrm{O^{2-}} \longrightarrow \mathrm{H_2O}, \tag{2.4}$$

where the water production reaction takes place at the cathode for the case of proton-conducting electrolytes and at the anode for the case of oxide-conducting electrolytes.

2.2 Performance factors

From the principle of the fuel cell, several basic performance requirements are evident: The rate at which adsorption and dissociation of the fuel and oxygen molecules takes place at the **electrodes** must be high, as must the assembly and discharge of by-products (water molecules). The electrodes must also have both high electronic and ionic conductivity. An **electrolyte** on the other hand, should have high ionic conductivity but to prevent short-circuiting the cell it can not be permeable to electrons and molecules or atoms of the oxygen gas and the fuel.

Figure 2.1(d) schematically illustrates the effects of limitations in the basic fuel cell requirements [18]. Power density (power per fuel cell area) is given by voltage times current density: $P = U\,i$ and should be as high as possible to optimize the performance of the fuel cell with respect to size and weight and thus also cost. The dotted line indicates the electrostatic potential U_{theory} which in the ideal case is generated across the electrodes as a result of the available chemical energy (compare Box 1.1). In reality a number of losses are present and the potential can be written:

$$U = U_{\text{theory}} - U_{\text{co}} - U_{\text{sr}} - U_{\text{or}} - U_{\text{smd}}, \tag{2.5}$$

where the subscripts co, sr, or and smd refer to the losses indicated in the figure: cross-over, slow reaction, ohmic resistance and slow mass diffusion. **Cross-over** losses derive from imperfections in the insulating and gas-separating properties of the electrolyte: finite electronic conductivity and gas permeability or leaks due to for instance cracks or pores. **Slow reaction** rates at the electrodes generically becomes a problem at low temperatures or with ineffective catalyst materials. The **ohmic resistance** of the fuel cell gives rise to a linear drop in the voltage as function of current (Box 2.1 shows the effect on efficiency): $U_{\text{or}} = R\,i$, where R is the area-specific resistance (ASR) measured in Ωcm^2. **Slow mass diffusion**, finally, refers to limitations in the flow of fuel and oxygen to the electrodes and the flow of by-products away from them.

Many practical considerations also enter the picture, such as the compatibility of the components, their individual durability and cost, and the practicality and cost of the finished fuel cell device. All things considered, the operating temperature of a fuel cell device has great effects on its performance. A highly **beneficial** re-

Box 2.1: Theoretical voltage and efficiency of a hydrogen fuel cell.

If all the free energy ΔG^{f} released by reaction 2.1 is converted into work done by the two electrons produced in the anode, $W = -2eU_{\text{theory}}$, the voltage is:*

$$U_{\text{theory}} = -\Delta G^{\text{f}}/2e \sim 1.25\,\text{V} - 0.5\,T\,\text{mVK}^{-1}.$$

Assuming for illustration only ohmic losses and letting $P_{\text{theory}} = U_{\text{theory}} i$ and $P_{\text{or}} = U_{\text{or}} i = i^2 R$, the electrical efficiency of a fuel cell is:

$$\eta_{\text{el}} = \frac{P_{\text{theory}} - P_{\text{or}}}{P_{\text{theory}}} = 1 - \frac{iR}{U_{\text{theory}}}.$$

*Values represent standard pressures of H_2, O_2 and H_2O [19]. Implied is a maximum thermodynamic efficiency $\eta_G \sim 1 - 0.2T\ (1000\text{K})^{-1}$. Total efficiency would be $\eta = \eta_{\text{el}}\eta_G$.

sult of **high operating temperatures** is an increase of chemical reaction rates at the electrode surfaces, which reduces slow reaction losses. It also makes the fuel cell less sensitive to impurities in the fuel and more importantly enables the use of carbon-containing fuels. In devices operating at low temperatures, expensive electrode materials like platinum become necessary for efficient use even with pure hydrogen fuel. Further demands on the purity of the fuel are set by the susceptibility of platinum to CO poisoning. Last but not least, the conduction mechanisms of ions in solid oxides are thermally activated, which means electrolytes and electrodes made from such materials suffer less ohmic resistance losses at high temperatures.

There are however **disadvantages of excessive temperatures**. In addition to the decrease in thermodynamic efficiency (Box 2.1) it makes thermal insulation critical and forces the use of expensive ceramic or stainless steel materials in the construction of the fuel cell device. It also increases the degradation rate of the various components and may cause mechanical problems related to thermal expansion, particularly in the compatibility between electrolyte and electrode materials. Cracks and other imperfections may form, causing cross-over losses. Furthermore, the time it takes for the fuel cell to reach its operating temperature increases which leads to inconvenient start-up times and precludes use in applications with intermittent energy demands. Finally it can be remarked that devices with oxide-ion conducting electrolytes risk hampered efficiency at high temperatures due to formation of steam at the anodes, which dilutes the fuel.

2.3 Technologies

A number of different ways of realizing the relatively simple fuel cell principle have been explored over the years. Implementations are to a great extent contrasted in the choice of electrolyte and are in general named correspondingly. While this thesis deals solely with issues related to the solid oxide fuel cell (SOFC), a brief orientation on the alternatives may be in order.[2] A summary of conventional fuel cell types with some of their characteristic traits is given in table 2.1. Quite striking is the division that can be made into low-temperature and high-temperature technologies. In the temperature range between 200 and 800°C only the molten-carbonate fuel cell (MCFC) can be found, at 650°C. This noticeable lack of devices capable of operating in the intermediate temperature range spurs much of the current research regarding fuel cells, with the outlook of finding a middle ground between the advantages and disadvantages of high temperatures.

At the low-temperature end the **polymer electrolyte membrane fuel cell (PEMFC)** is found. It is arguably the one garnering the greatest amount of attention due to its high power density along with the quick start-up times affiliated with its low operating temperature. These attributes make it suitable for a wide range of applications, including mobile installations such as cars. Unfortunately the proton-transporting water molecules inherent in the sulfonated polymer electrolyte make temperatures above the boiling point of water, 100°C, difficult to achieve.

[2] For further information on different fuel cell technologies see references 10, 11, 15.

Table 2.1: Summary of fuel cell types and characteristics. [15]

Parameters	Fuel cell types					
	PEMFC	AFC	PAFC	MCFC	SOFC	DMFC
Electrolyte	Solid polymer membrane (Nafion)	Liquid solution of KOH	Phosphoric acid (H_3PO_4)	Lithium and potassium carbonate (LiAlO2)	Stabilized solid oxide electrolyte (Y_2O_3, ZrO_2)	Solid polymer membrane
Operating temperature (°C)	50–100	50–200	~200	~650	800–1000	60–200
Anode reaction	$H_2 \rightarrow 2H^+ + 2e^-$	$H_2 + 2(OH^-) \rightarrow 2H_2O + 2e^-$	$H_2 \rightarrow 2H^+ + 2e^-$	$H_2O + CO_3^{2-} \rightarrow H_2O + CO_2 + 2e^-$	$H_2 + O^2 \rightarrow H_2O + 2e^-$	$CH_3OH + H_2O \rightarrow CO_2 + 6H^+ + 6H^-$
Cathode reaction	$1/2O_2 + 2H^+ + 2e^- \rightarrow H_2O$	$1/2O_2 + H_2O + 2e^- \rightarrow 2(OH)^-$	$1/2O_2 + 2H^+ + 2e^- \rightarrow H_2O$	$1/2O_2 + CO_2 + 2e^- \rightarrow CO_3^{2-}$	$1/2O_2 + 2e^- \rightarrow O^{2-}$	$3O_2 + 12H^+ + 12H^- \rightarrow 6H_2O$
Charge carrier	H^+	OH^-	H^+	$CO_3^=$	$O^=$	H^+
Fuel	Pure H_2	Pure H_2	Pure H_2	H_2, CO, CH_4, other hydrocarbons	H_2, CO, CH_4, other hydrocarbons	CH_3OH
Oxidant	O_2 in air	O_2 in air	O_2 in air	O_2 in air	O_2 in air	O_2 in air
Efficiency	40–50%	~50%	40%	>50%	>50%	40%
Cogeneration	–	–	Yes	Yes	Yes	No
Reformer is required	Yes	Yes	Yes	–	–	–
Cell Voltage	1.1	1.0	1.1	0.7–1.0	0.8–1.0	0.2–0.4
Power density (kW/m^3)	3.8–6.5	~1	0.8–1.9	1.5–2.6	0.1–1.5	~0.6
Installation Cost (US $/kW)	<1500	~1800	2100	~2000–3000	3000	–
Capacity	30 W, 1 kW, 2 kW, 5 kW, 7 kW, 250 kW	10–100 kW	100 kW, 200 kW, 1.3 MW	155 kW, 200 kW, 250 kW 1 MW, 2 MW	1 kW, 25 kW, 5 kW, 100 kW, 250 kW, 1.7 MW	1 W to 1 kW, 100 kW to 1 MW (Research)
Applications	Residential; UPS; emergency services such as hospitals and banking; industry; transportation; commercial	Transportation; space shuttles; portable power	Transportation; commercial cogeneration; portable power	Transportations (e.g. marine-ships; naval vessels; rail); industries; utility power plants	Residential; utility power plants; commercial cogeneration; portable power.	It is used to replace batteries in mobiles; computers and other portable devices
Advantages	High power density; quick start up; solid non-corrosive electrolyte	High power density; quick start up	Produce high grade waste heat; stable electrolyte characteristics	High efficiency; no metal catalysts needed	Solid electrolyte; high efficiency; generate high grade waste heat	Reduced cost due to absence of fuel reformer
Drawbacks	Expensive platinum catalyst; sensitive to fuel impurities (CO, H_2S)	Expensive platinum catalyst; sensitive to fuel impurities (CO, CO_2, CH_4, H_2S)	Corrosive liquid electrolyte; sensitive to fuel impurities (CO, H_2S)	High cost; corrosive liquid electrolyte; slow start up; intolerance to sulfur	High cost; slow start up; intolerance to sulfur	Lower efficiency and power density

Despite 40 years of success in the space craft industry, terrestrial use of the **alkaline fuel cell (AFC)** is limited due to the need of preventing contact with CO_2, which otherwise reacts with, and rapidly degrades the KOH electrolyte liquid. The **phosforic acid fuel cell (PAFC)** was the most popular technology in the early 1990s and several systems are in operation in Europe, USA and Japan. Although operating at twice as high temperatures as the PEMFC, expensive catalysts and pure fuel are still required.

From the perspective of using hydrocarbon derivatives as fuel, the **direct methanol fuel cell (DMFC)** and the **molten carbonate fuel cell (MCFC)** might seem interesting competitors to the SOFC. DMFCs like the PEMFCs have a polymer electrolyte and inherit many of their properties. The methanol fuel however causes cross-over and slow reaction losses which severely limit the power density and restrict the use to low-power applications. In the temperature-wise relatively pleasant MCFC, the molten salt (typically a Li/Na/KCO_3 mixture) used as electrolyte instead causes inconveniences. First of all its corrosive nature places high demands on container materials. Secondly, CO_2 forms at the anode and must be recirculated to the cathode by pumps. Third, the electrolyte can only withstand a limited number of thermal cycles and must therefore be kept above its melting temperature at all time.

The ceramic electrolyte material of **solid oxide fuel cells (SOFCs)** finally, brings many of the advantages of such materials. Solid oxides are often insulators or only slightly semiconducting, with suitable electronic properties. Their durability and mechanical robustness are also pleasing qualities. Drawbacks are generally speaking limited to the high-temperature disadvantages mentioned in section 2.2 and narrowing the fuel cell temperature gap from above by developing **intermediate-temperature solid oxide fuel cells (IT-SOFCs)** therefore indeed stands out as an attractive ambition.

Chapter 3

Materials for solid oxide electrolytes

What turns out to be critical in the development of intermediate-temperature solid oxide fuel cells is that the dense crystal structure of the ceramic electrolyte materials not only has the benefit of making them impermeable to fuel and oxygen molecules, but also brings the disadvantage of restricting the mobility of ions incorporated in the structure. The conduction mechanism is thermally activated, and as a consequence the resistivity rapidly increases with decreasing temperatures.

In terms of an acceptable resistance, the need for an overall area specific resistance not exceeding 0.5 Ωcm^2 of a finished fuel cell has been pointed out [9]. Assuming an electrolyte thickness of 15 μm and an allowed contribution of 0.15 Ωcm^2 to the overall resistance, this corresponds to a specific electrolyte conductivity of 10^{-2} Scm^{-1} [9], a number frequently encountered as the minimum target for solid oxide electrolyte materials. The material is then suitable for use at or above the temperature at which the number is achieved.

In this chapter the materials which are studied in Papers I–V are introduced, along with a condensed overview of other materials and considerations presently relevant to the field. More detailed reviews and examples of conventional as well as novel electrolyte materials can be found in for instance references 8, 13, 20.

3.1 Oxide-ion conductors

Conventional solid oxide fuel cells today employ oxide-ion conducting electrolytes based on zirconium oxide, ZrO_2. To enhance the ionic conductivity in the material it is doped, which means a certain amount of a foreign metal ion is introduced into it. This stabilizes the structure in its higher-conducting cubic fluorite phase and cause oxygen vacancies to form which enables the oxide ions to migrate in the material.[1] Typically yttrium (Y) or scandium (Sc) is used, giving rise to the names yttria- and scandia-stabilized zirconia respectively (YSZ and ScSZ). Characteristic for zirconia electrolytes is good structural and mechanical stability and sufficient ionic conductivity at the elevated temperatures solid oxide fuel cells operate today (800–1000°C).

[1] The mechanisms of doping and vacancy formation will be returned to in chapters 4 and 5.

Zirconia as an electrolyte has been studied for many years and its properties are relatively well explored. [13, 20]

Attempts have also been made to produce oxide-ion conducting electrolytes based on ceria, doped CeO_2. Compared to its fluorite relative zirconia, these materials display better ionic conductivity but during certain conditions a non-negligible electronic conductivity is also seen, with corresponding cross-over losses. Furthermore they struggle with issues related to stability and processing of the material. Ceria is considered promising for temperatures around 500–650°C and research is ongoing. [13, 20]

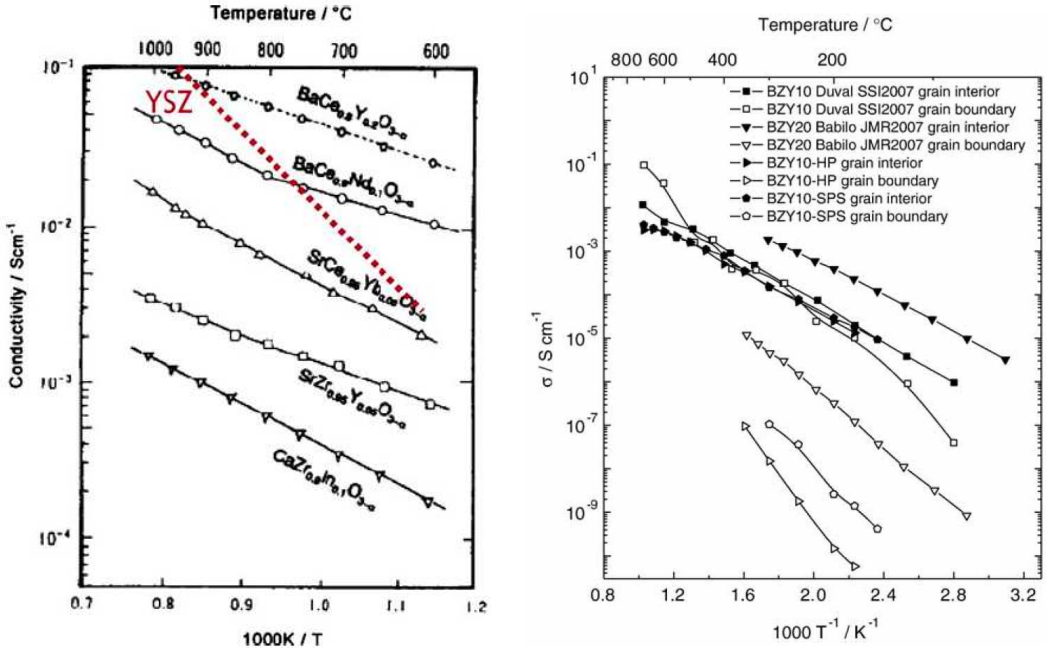
Other classes of oxide-ion conducting materials are also being investigated, including apatites and LAMOX as well as perovskites and perovskite-related structures. Ionic conductivities greater than that of zirconia has encouragingly been reported for all of these classes and might indicate that new insights or even future electrolyte candidates can be found among them. Apatites (*eg.* $La_9Sr(SiO_4)_6O_{2.5}$) are difficult to synthesize but are believed to have the feature of conducting oxide ions via a qualitatively different, interstitial, mechanism than for instance zirconia. LAMOX ($La_2Mo_2O_9$) undergoes a phase transition to a highly conducting phase at 580°C, but suffers from chemical and thermal incompatibility with electrode materials. In certain perovskite-related structures, interesting order-disorder transitions with noticeable effects on ionic oxygen conductivity are seen but need more investigation along with chemical instabilities reported in the materials. [13, 20]

3.2 Proton conductors

Leaving the oxide ion conductors, perovskite-structured materials are of even higher interest and connections to zirconia and ceria persist. The two materials investigated in Papers I–V of this thesis for instance, are both *zirconates*. In a wider perspective it should be mentioned that solid oxides may exhibit conductivity of both protons and oxygen ions, as protons and oxygen vacancies may well coexist in the material. Typical proton conductors exhibit dominating proton conductivity at low temperatures but transition to oxide ion conduction at high temperatures. The relevant mechanism of proton incorporation will be explained in chapter 4.

Generally speaking, proton conducting electrolytes bring the already pointed out basic benefit of water being produced at the oxygen-side of the fuel cell, allowing full use of the hydrogen fuel. However, the greatest reason solid oxide proton conductors attract much attention as electrolyte candidates for intermediate-temperature fuel cells is the generally lower conduction activation energy they exhibit. This allows for high ionic conductivities at fundamentally lower temperatures than conventional oxide ion conductors (*cf.* the difference in slopes seen in figure 3.1 (a)). [8]

Pure and high protonic conductivity in solid oxide materials was first reported for the cerate perovskites $SrCeO_3$ [25, 26] and $BaCeO_3$ [27] by Iwahara and coworkers in the eighties when they found increased conductivities in doped samples exposed to humid air. Later, spurred by the troublesome reactivity of cerates with CO_2 and the known better stability of zirconates, Iwahara et al investigated the conductivity of



(a) (b)

Figure 3.1: Comparison of (a): the temperature-dependence of conductivities in typical oxide-ion conductors (YSZ) and proton conductors. [14, 21], and (b): the conductivity of interior and boundary of $BaZrO_3$ grains. [22–24]

$CaZrO_3$, $SrZrO_3$ and the material studied in Papers III–V of this thesis: $BaZrO_3$ [28]. The cubic perovskite structure of $BaZrO_3$ is shown in figure 3.2 (a).

Since Iwaharas ground-breaking work, proton conductivity in barium cerate ($BaCeO_3$, BCO) and barium zirconate ($BaZrO_3$, BZO) has been extensively investigated and the materials are today two of the top contenders for proton-conducting electrolytes. They exhibit in some sense mutually opposite benefits and drawbacks, seemingly inherited from their oxide ion conducting fluorite cousins: BCO displays some of the highest proton conductivities but suffers from poor chemical stability and contributions of electronic conductivity, while BZO on the other hand has excellent chemical stability but in general show total conductivities about an order of magnitude lower than BCO. Additionally, BZO is troublesome to synthesize and requires high sintering temperatures to obtain dense, homogenous samples. Given the antagonistic assets of the two materials, ongoing investigations into mixed solutions can be mentioned, along with optimization of doping schemes. The outlook is hopeful enough that chemically compatible electrodes for use with cerates and zirconates in fuel cells are being researched and the performance of assembled fuel cell structures are being evaluated. [8, 14, 20, 30]

In the light of recent findings, the prospect of BZO is perhaps particularly promising. Encouraging results on the sintering issue has been reported by the inclusion of sintering aids such as ZnO and different materials processing techniques. Moreover, it has become apparent that the low measured conductivities are due to blocking ef-

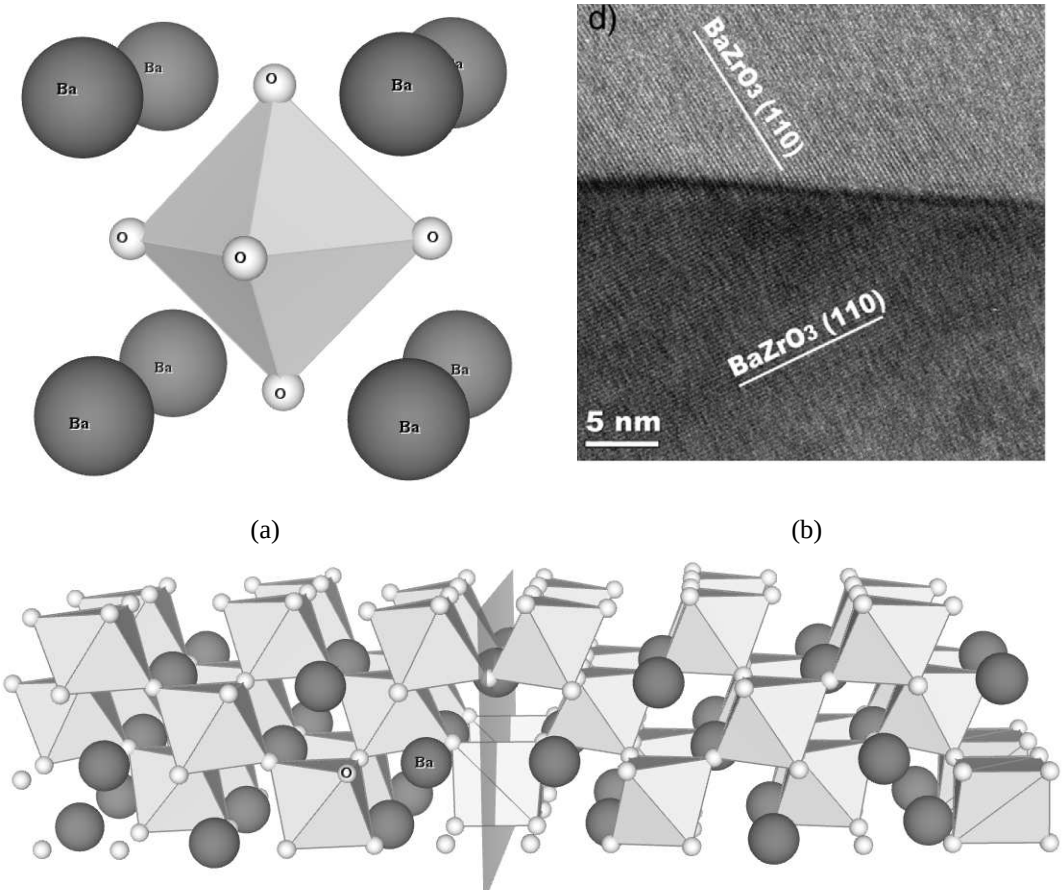


Figure 3.2: $BaZrO_3$ (barium zirconate). (a): the cubic perovskite structure. The zirconium ion is in the center of the oxygen octahedron. (b): HRTEM micrograph of the boundary between two grains of $BaZrO_3$ [29]. (c): one of the grain boundary structures studied in Papers III–V.

fects caused by grain boundaries (compare figure 3.1 (b)). The resistivity of the grain boundaries thus overshadows the high conductivity of ideal bulk material, shown in figure 3.3, which was predicted to be even higher than BCO by Kreuer [31] a decade ago. [8, 20, and references therein]

The grain-boundary blocking effect in BZO is clearly in need of elucidation. Based on experience from zirconia and ceria [33], it has been speculated that an aggregation of positive charge at the boundaries is likely a contributing cause [22,34] and recent experimental evidence seems to support the hypothesis [35, 36]. Such an accumulation of positive charge gives rise to an electrostatic potential barrier and depletion of the positive oxygen vacancies and protons in the grain boundary region. As of yet the cause of the positive grain boundaries has not been unraveled and shedding light on this issue opens for devising schemes to counteract the problem. This is what motivates the research of Papers III-V, where it is investigated whether

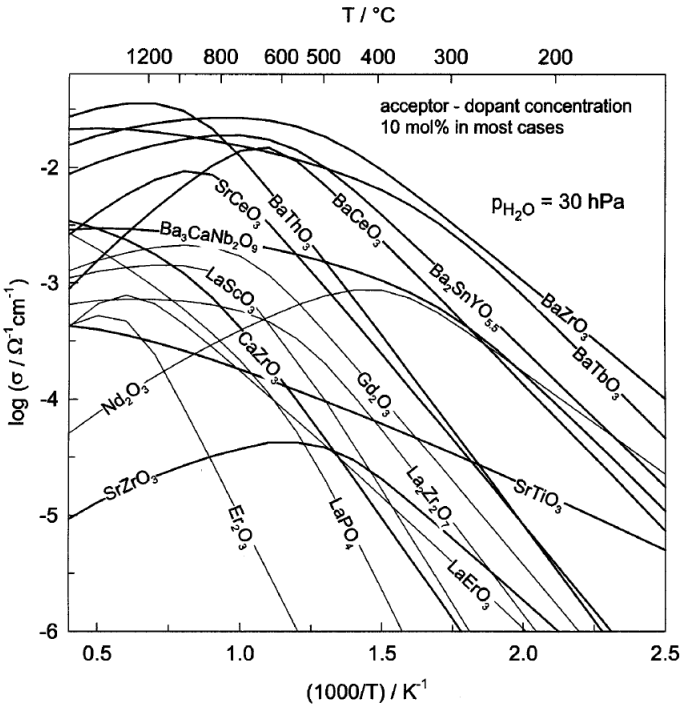
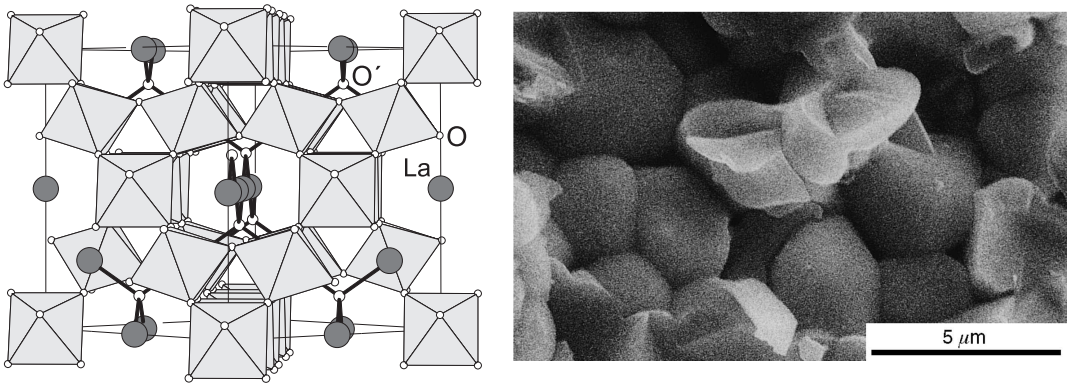


Figure 3.3: Proton conductivities of various oxides [31] as calculated from data on proton concentrations and mobilities [32]. Noteworthy is the high conductivity of $BaZrO_3$, but in the midst of the perovskites (composition ABO_3) are also other classes of less studied oxides, for instance the pyrochlore-structured $La_2Zr_2O_7$.

an aggregation of positive oxygen vacancies in the core of grain boundaries in BZO is a factor for the blocking effect. An HRTEM micrograph of a grain boundary and one of the grain boundaries in which oxygen vacancy segregation was explored are shown in figure 3.2 (b)–(c).

Although perovskites like BCO and BZO are the most mature proton conductors for use as solid oxide electrolytes, the search for suitable materials clearly does not end with them. Investigations of other classes of materials enable discovery of common denominators or qualitatively different mechanisms of proton incorporation and transport, which make possible a deeper understanding of properties relevant for proton-conducting electrolytes. Certain perovskite-related compounds (BCN18) for instance display intrinsic metal ion non-stoichiometry which gives rise to increased proton concentrations. Other perovskite relatives (elpasolites, cryolites, $Ba_2In_2O_5$) have intrinsically oxygen-deficient structures. Phosphates like $LaPO_4$ have received interest due to its unusual mechanisms of proton incorporation and transport. Hope for diminishing the fuel cell temperature gap from below has been raised by solid acids like $CsHSO_4$ and CsH_2PO_4, which show high proton conductivities at 120–300°C. Niobates and tantalates have enjoyed quite some attention due to excellent chemical stability.

One class of oxides which in the context of the present thesis is of extra interest is the rare-earth pyrochlore oxides, to which lanthanum zirconate ($La_2Zr_2O_7$, LZO),



(a) (b)

Figure 3.4: $La_2Zr_2O_7$ (lanthanum zirconate). In (a) the pyrochlore structure, as given by first-principles optimization of the ionic configuration (Paper I) is shown. Two inequivalent oxygen sites exist, O and O′. The Zr ions are in the center of the O octahedra. Figure (b) displays a SEM micrograph, showing grains of $La_2Zr_2O_7$ [38].

the material studied in Papers I–II belong. The structure, seen in figure 3.4, has similarities with the proton-conducting perovskites in that it contains interlinked oxygen octahedra with the rare-earth cation in the center. In a sense it is also related to zirconia as it can be seen as a fluorite super-structure, with the interesting difference of an intrinsically vacant oxygen site. Possibly motivated by these properties and the success of the perovskites, Iwahara and coworkers in 1996 investigated the proton conductivity of a number of pyrochlore zirconates, including LZO. [37] Although doping was still necessary to obtain ionic conductivity, indicating that the intrinsic vacancy position did not contribute, the proton conductivity was not insignificant (compare figure 3.3).

It can be added that the ion-conducting properties of LZO are relevant not only from a fundamental point of view but also from the perspective of fuel cell design, as restrictive layers of LZO have been seen to form in the interface between YSZ electrolytes and popular electrode materials containing La [39]. While such issues were not seen in recent compatibility tests between BZO and electrode materials [30], interdiffusion between Zr-based electrolytes and La-containing electrodes is not unlikely to cause segregation of LZO.

Despite of this, further investigations into proton conductivity in pyrochlore-structured oxides have been limited compared to the perovskites. Especially lacking are theoretical studies of the microscopic mechanisms related to proton conductivity. In Papers I–II a few of the most fundamental mechanisms of proton transport and doping optimization are examined in LZO.

Chapter 4

Ion conduction in solid oxides

Optimizing the ionic conductivity of materials for use as electrolytes in solid oxide fuel cells is a direct route towards improving fuel cell efficiency and narrowing the temperature gap in present fuel cell technologies. Understanding the mechanisms of first and foremost proton conductivity in solid oxide on a fundamental level is also the consistent theme of the present thesis. In this chapter an introduction to electrical conductivity in general is therefore provided in section 4.1. This is followed by an account of the effect of microstructure in terms of grains and grain boundaries on the effective, macroscopic conductivity in section 4.2. The remaining sections of the chapter then outlines the specific atomistic mechanisms which are explicitly or implicitly relevant for the work presented in the appended papers.

4.1 Definitions and introduction

The electrical resistance R of a piece of material is a measure of its inability to conduct direct electric current and is defined by the familiar ohm's law:

$$I = U\frac{1}{R}. \tag{4.1}$$

That is, applying a voltage U over a piece of material with the resistance R gives rise to the current I. Given a certain voltage, the total current depends on the geometry of the piece of material under consideration. For a uniform piece the current increases with its cross-sectional area and decreases with its length:

$$I = U\frac{1}{\rho}\frac{A}{L}, \tag{4.2}$$

which introduces the resistivity ρ of the material: the measure of the inability of the material itself to conduct direct electric current. Resistivity is thus the intensive material property corresponding to the extensive resistance. It is however often convenient to consider not the inability of a material to conduct current but instead its ability, which is the conductivity σ:

$$\sigma = \frac{1}{\rho}. \tag{4.3}$$

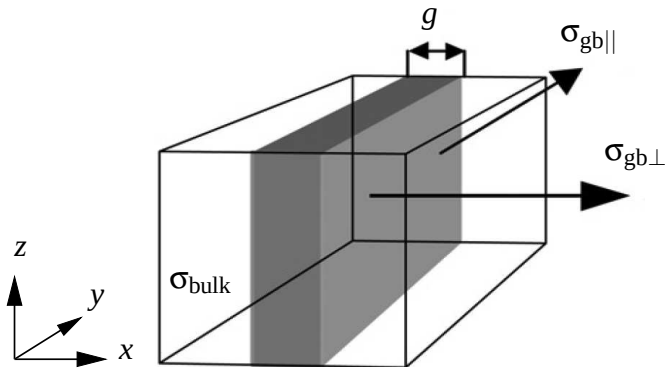


Figure 4.1: Illustration of grain boundary conductivity differing from bulk (grain interior). Inside the shaded grain boundary region, a different conductivity can be expected both in the parallel ($\sigma_{gb||}$) and perpendicular ($\sigma_{gb\perp}$) directions.

Since electric current is the transport or flux of electric charge, different charge carriers can generally speaking contribute to the conductivity of a material and the total conductivity is the sum of the conductivities related to each and every one of them. In solid oxide electrolytes the relevant charge carriers are in general oxide ions O^{2-} and/or protons H^+ and, hopefully to a negligible extent, electrons e^- and/or electron holes h^+:

$$\sigma = \sigma_{O^{2-}} + \sigma_{H^+} + \sigma_{e^-} + \sigma_{h^+}. \tag{4.4}$$

Sometimes the transference number t is used to signify the contributions of different charge carriers to the total conductivity:

$$\sigma = (t_{O^{2-}} + t_{H^+} + t_{e^-} + t_{h^+})\sigma. \tag{4.5}$$

In electrolytes it is desirable to obtain $t_{O^{2-}} + t_{H^+} >> t_{e^-} + t_{h^+}$. In oxide-ion conductors $t_{O^{2-}}$ dominates while in proton conductors t_{H^+} shows the largest contribution.

Although electronic conductivity is of implicit interest in the present work in that it should be low in electrolyte materials, ionic conductivity is the primary topic and will henceforth be the focus of attention.

4.2 Microstructure: grains and grain boundaries

In a material like conventionally synthesized $BaZrO_3$ with its grains and troublesome grain boundaries it is unlikely that the conductivity is homogenous and equal throughout the material. In the interfaces between the grains, the grain boundaries, the structure is per definition different than in the bulk interior of the grains and a different conductivity should be expected, compare figure 4.1.

Considering the structural arrangement of grains seen in for instance figure 3.4(b), a simplified model is needed to evaluate the effective conductivity of the material. A common approach is illustrated in figure 4.2. The grains are approximated as cubes with the side length G assigned the average grain size or diameter of grains in the

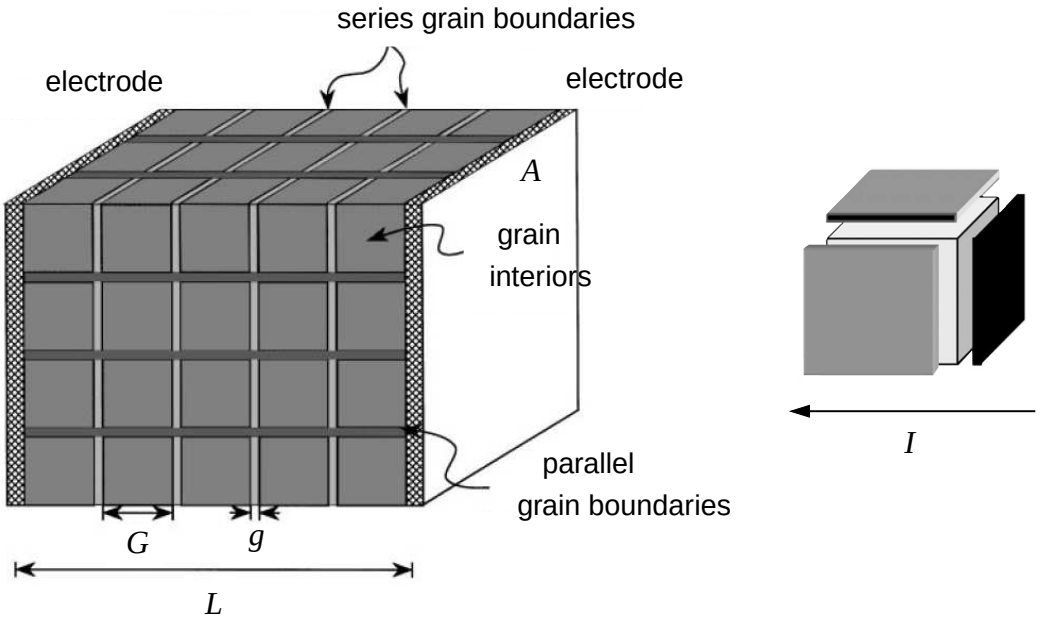


Figure 4.2: The brick layer model, approximating grains with an average size G separated by grain boundaries of thickness g. On the right side, one building block of the brick layer structure is shown, with the arrow indicating the imagined direction of the current through the sample.

sample. Between the cubes, flat layers model grain boundaries having some characteristic thickness g. One building block of this so-called brick layer model consists of a cubic bulk piece, one grain boundary layer perpendicular to the direction of the current and two parallel with it. If large grains and thin boundary layers are assumed, $G >> g$, the explicit expression for the effective conductivity σ_{eff} of this brick layer geometry is (details in Box 4.1):

$$\frac{1}{\sigma_{\text{eff}}} = \frac{1}{\sigma_{\text{bulk}} + 2\frac{g}{G}\sigma_{\text{gb}||}} + \frac{g}{G}\frac{1}{\sigma_{\text{gb}\perp}}, \tag{4.6}$$

which translates to an effective resistance $R_{\text{eff}} = L/A\sigma_{\text{eff}}$ of the sample. For grain boundaries with diminished conductivity compared to bulk, like is seen in BZO, equation 4.6 shows the relevance of grain size and the perpendicular component of the grain boundary conductivity: If $\frac{2g}{G}\sigma_{\text{gb}||} << \sigma_{\text{bulk}}$, the effective conductivity

Box 4.1: Effective conductivity from brick layer model.

Referring to figure 4.2 and assuming $G >> g$, the "electrical curcuit" of one unit in the brick layer geometry can be seen as one boundary layer with area G^2 and length g connected in series with a parallel connection of the bulk piece, area G^2 and length G, and two boundary layers, total area $2gG$ and length G. From the parallel connection, $\frac{G^2}{G}\sigma_1 = \frac{G^2}{G}\sigma_{\text{bulk}} + \frac{2gG}{G}\sigma_{\text{gb}||}$ is obtained. This is to be connected serially to $\frac{G^2}{G}\sigma_2 = \frac{G^2}{g}\sigma_{\text{gb}\perp}$. Thus writing $1/\sigma_{\text{eff}} = 1/\sigma_1 + 1/\sigma_2$ gives the expression in equation 4.6.

obeys

$$\frac{1}{\sigma_{\text{eff}}} = \frac{1}{\sigma_{\text{bulk}}} + \frac{g}{G}\frac{1}{\sigma_{\text{gb}\perp}}. \tag{4.7}$$

Here the second term on the right hand side is suppressed by the factor g/G, which mathematically confirms what is intuitively understood: that large grains and therefore a smaller ratio of grain boundary to bulk is preferable in the case of resistive grain boundaries. However, the exponential temperature dependence of the conductivities dominates over the size effect as temperature is elevated (*cf.* the diverging bulk and grain boundary graphs of figure 3.1(b)), making the impact of the grain boundaries sensitive to the activation energy of the grain boundary conductivity (see section 4.3).

In papers III–V, it is shown that concentration gradients are expected in the grain boundary regions of $BaZrO_3$ (compare figure 4.7). In such cases, the grain boundary conductivity can be found by averaging over the interface domain. To a first approximation it is reasonable to assume that the local conductivity in the grain boundary is homogenous parallel to the interface but varies perpendicularly. Referring to the coordinates indicated in figure 4.1 the grain boundary region can then be divided into slices with width dx. Parallel to the interface, the average conductivity is found by integrating the conductivity of each slice:

$$\sigma_{\text{gb}||} = \frac{1}{g}\int_{-g/2}^{g/2} \sigma(x)\mathrm{d}x, \tag{4.8}$$

where $x = 0$ has been placed at the center of the interface. In the perpendicular direction, the slices are instead serially connected and the average conductivity is given by:

$$\frac{1}{\sigma_{\text{gb}\perp}} = \frac{1}{g}\int_{-g/2}^{g/2} \frac{1}{\sigma(x)}\mathrm{d}x. \tag{4.9}$$

As demonstrated, the brick layer model provides a means to estimating the effective conductivity of a certain polycrystalline sample given the conductivity in different regions within the crystal, more specifically the bulk and grain boundary regions. It is also routinely used to separate the measured conductivity of a sample into bulk and grain boundary contributions, as seen in figure 3.1(b). A review of the impedance spectroscopy then employed is for instance given in reference 40, where further discussion on and references to the brick layer model can also be found.

4.3 Activation energy vs mobility and concentration

Measurements of ionic conductivity in solid oxides generically produce linear graphs when $\log(\sigma T)$ or for small temperatures T, $\log(\sigma)$ is plotted against $1/T$ (compare figures 3.1 and 3.3). In other words the conductivity typically takes the familiar arrhenius form:

$$\sigma = \frac{\sigma_0}{T}\exp(-E^{\text{a}}/k_{\text{B}}T), \tag{4.10}$$

which is usually taken as the definition of the activation energy E^{a} of the conductivity for a given material sample during specific measurement conditions.

To illuminate the origin of the expression in equation 4.10 it is necessary to consider the conductivity of ions migrating in a crystal lattice from an atomistic viewpoint. For each specific charge carrier, its contribution to the total conductivity depends on two factors: a) how many such carriers exist per unit of material and b) how mobile they are. For a certain species i that carries q_i number of fundamental charges e and has concentration c_i and mobility u_i, its conductivity can thus be written:

$$\sigma_i = q_i e u_i c_i. \tag{4.11}$$

There are in other words essentially two possible sources of a given activation energy: mobility or concentration of the charge carrier in question. It also follows that two ways exist of increasing the conductivity: increasing the concentration of the carrier or increasing its mobility.

Ions migrating in a dense crystal lattice are generally restricted to jumping between specific lattice sites. Consequently there has to exist an uninterrupted pathway of such sites in the lattice. Furthermore the sites may not be saturated by other copies of the migrating species. Considering further that a lattice site in a crystal to a lesser or greater extent represents a stable ion position corresponding to a local minimum in the free energy landscape, it follows that in essence the problem of finding a material with high ionic conductivity boils down to finding a material (i) capable of hosting a large concentration of the ion in question, that has (ii) partially occupied, interlinked ion sites, (iii) separated by an as low as possible migration barrier [41].

In the appended papers, migration barriers are only considered in Paper I and strong emphasis of the research is instead on the aspect of free carrier concentration. The mobility aspect is discussed in the next section, while the remaining content of the summarizing part of the thesis is devoted to concentration.

4.4 Mobility

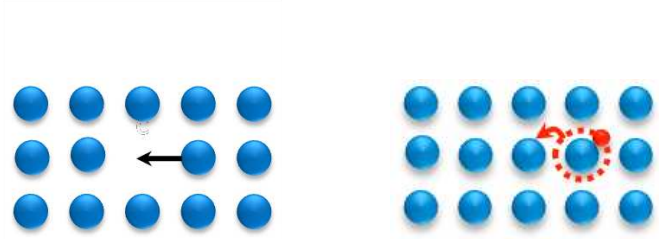
Starting from equation 4.11, the mobility of an ion is linked to its self-diffusivity D via the Nernst-Einstein relation [41]:

$$u = \frac{qeD}{k_{\mathrm{B}}T}. \tag{4.12}$$

An explicit expression for the self-diffusion coefficient can in turn be found from random walk theory [41]:

$$D = \tilde{\gamma} z (1-n) \nu_0 \exp\left(-\frac{\Delta G^{\mathrm{m}}}{k_{\mathrm{B}}T}\right), \tag{4.13}$$

where $\tilde{\gamma}$ contains geometrical details of the lattice and possible corrections for non-randomness in the diffusion, z is the number of neighboring equivalent sites, n



(a) (b)

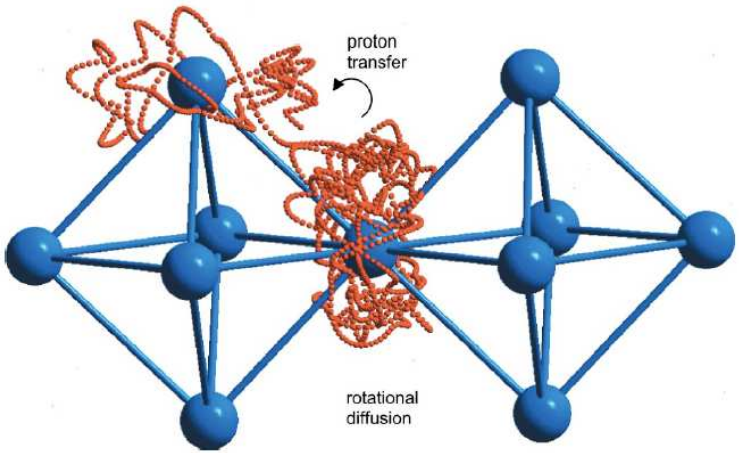
Figure 4.3: Schematic illustration of the vacancy mechanism of oxide ion migration (a) and the grotthuss mechanism of proton migration (b). For simplicity only the oxygen sublattice is shown.

is the occupied fraction of those sites, and ν_0 is a characteristic vibrational frequency of the lattice, expressing the frequency of "jump attempts" of the migrating ion. $\Delta G^{\mathrm{m}} = \Delta H^{\mathrm{m}} - T\Delta S^{\mathrm{m}}$ finally, is the barrier in free energy between two equivalent sites. Here, H^{m} is enthalpy and S^{m} entropy. Noteworthy is that the temperature dependency of the diffusivity is determined by the migration enthalpy $\Delta H^{\mathrm{m}} = \Delta E^{\mathrm{m}} + p\Delta V^{\mathrm{m}}$. If for convenience a new parameter $\gamma = \tilde{\gamma} z \nu_0 \exp(\Delta S^{\mathrm{m}}/k_{\mathrm{B}})$ is defined, the conductivity of an ionic species i in a crystal lattice can now be written by combining equations 4.11–4.13:

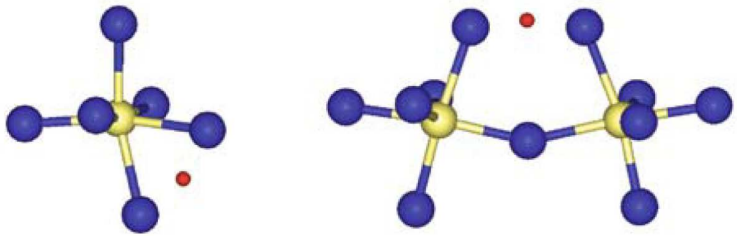
$$\sigma_i = \frac{c_i(1 - c_i/N_i)q^2e^2\gamma}{k_{\mathrm{B}}T} \exp\left(-\frac{\Delta H_i^{\mathrm{m}}}{k_{\mathrm{B}}T}\right). \tag{4.14}$$

Here the density of lattice sites N_i offered the migrating species was also introduced. It can be added that usually the concentration of charge carriers is relatively small so that the factor $(1 - c_i/N_i)$ is well approximated by unity. The case of oxide ion conduction is slightly different in this respect, as ideal oxides have $c_{\mathrm{O}^{2-}}/N_{\mathrm{O}} = 1$ (all oxygen sites are occupied by oxygen ions). After noting that $c_{\mathrm{O}^{2-}}(1 - c_{\mathrm{O}^{2-}}/N_{\mathrm{O}}) = (N_{\mathrm{O}} - c_{\mathrm{V_O}})c_{\mathrm{V_O}}/N_{\mathrm{O}}$, where $\mathrm{V_O}$ denotes a vacant oxygen site, it is therefore more common to think of oxygen migration as the migration of oxygen vacancies, and consider the conductivity $\sigma_{\mathrm{V_O}} \propto c_{\mathrm{V_O}}(1 - c_{\mathrm{V_O}}/N_{\mathrm{O}}) \approx c_{\mathrm{V_O}}$.

Figure 4.3 (a) illustrates the vacancy-mediated diffusion process of oxide ion conduction. Regarding proton migration, protons naturally associate with oxygen ions to form hydroxide ions, $(\mathrm{OH})^-$. In solid oxides this corresponds to the proton being located in a position between lattice sites, covalently bonded to an oxygen ion. To migrate, the proton has to reorient itself around the oxygen ion and then reassociate to a neighboring one. This two-step rotation-jump operation, called the *Grotthuss mechanism*, is believed to be the main channel for proton migration in perovskites and is schematically illustrated in figure 4.3 (b). Computer simulations have shown that the highest migration barrier is associated with the jump step. It has further become clear that the migration is enabled by vibrations of the oxygen lattice, where distortions allow the protons to be handed over from one oxygen ion to the next, thus reducing the migration barrier. It is believed that the softness



(a)



(b)

Figure 4.4: More detailed illustration of proton migration in perovskites. Quantum molecular dynamics simulations showing (a): the proton trajectory, with the proton associated to an oxygen ion most of the time and only occasionally migrating to a neighboring ion [31]. (b): lattice distortion during the proton transfer step [42, 44, 45].

of the lattice in perovskites is indeed what makes them particularly good proton conductors. [31, 42, 43] The proton trajectory and lattice distortions, from quantum molecular dynamics simulations, are shown in figure 4.4. In Paper I it is shown that a similar proton pathway as seen in perovskites exists also in pyrochlore oxides, and the migration barrier is evaluated.

4.5 Concentration

The concentration of a free charge carrier i such as presently discussed is for a dilute, non-interacting solution at thermodynamic equilibrium given by the Boltzmann distribution:

$$c_i = N_i e^{-\Delta G_i^{\mathrm{f}}/k_B T}, \tag{4.15}$$

where ΔG_i^{f} is the change in Gibbs free energy associated with the introduction of one such carrier into the lattice. The underlying thermodynamics will be discussed

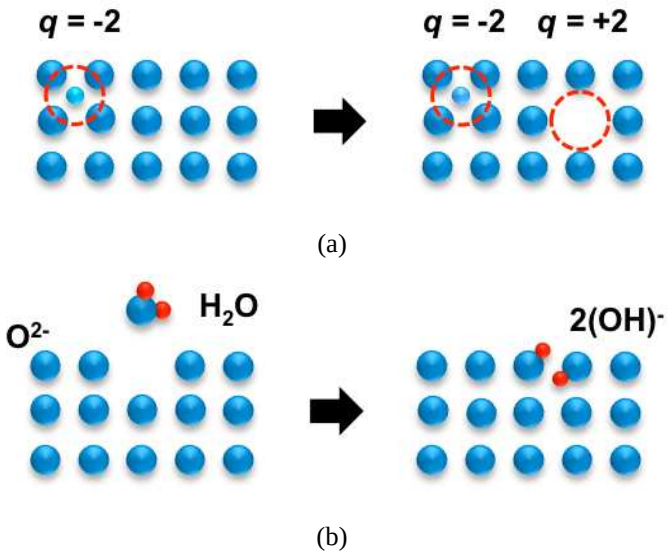


Figure 4.5: Schematic illustration of acceptor doping giving rise to a charge-compensating concentration of oxygen vacancies, (a), subsequently replaced by hydroxide ions in a humid environment, (b).

in chapter 5 and methods and considerations for calculating realistic free formation energies will be reviewed in some detail in chapter 7. For the purposes of the present chapter it however suffices to ascertain that as a consequence the conductivity of carrier i is now written:

$$\sigma_i = \frac{N_i q^2 e^2 \gamma}{k_B T} \exp\left(-\frac{\Delta H_i^m + \Delta H_i^f}{k_B T}\right), \tag{4.16}$$

where it has been assumed that $1 - c_i/N_i \approx 1$ and the exponential factor comprising the entropy and volume of formation has similar to the migration counterparts been consumed by the factor γ. It is seen that the formation enthalpy of the charge carrier also contributes to the activation energy. The sum $\Delta H_i^m + \Delta H_i^f$ expresses the fact that a low concentration and high mobility (high ΔH_i^f, low ΔH_i^m) can give rise to the same conductivity as a high concentration and low mobility (low ΔH_i^f, high ΔH_i^m).

Regarding the concentration of protons in most solid oxide proton conductors, it has already been touched upon in chapter 3 that significant protonic conductivity only appears after acceptor doping, an intentional incorporation of negatively charged point defects in the material, followed by exposure of the sample to a humid atmosphere. The reason is that the negative charges introduced in the material have to be compensated for, either by the creation of positive charges or the removal of other negative charges. The former can for instance occur by the creation of electron holes or protonic defects, while the latter can be achieved by the removal of oxygen ions. While protons as a result in principle can be incorporated directly in connection to the doping process, oxygen vacancies are more stable at the high temperatures in question (both oxygen and hydrogen tend towards vaporization as the temperature is increased). As the sample cools, water molecules from a humid

Figure 4.6: Schematic illustration of acceptor dopants trapping oxygen vacancies (left) and protons (right).

atmosphere are instead incorporated in the vacant oxygen positions, effectively creating two hydroxide ions per vacancy, *cf.* figure 4.5. The effect of this *hydration reaction* is that protons are introduced in the material and the proton concentration is increased.

In Paper II the charge-compensating efficiency of a number of different dopant species incorporated in LZO at typical synthesis conditions is evaluated.

4.6 Trapping

While introducing negative acceptor dopants in the material increases the concentration of oxygen vacancies and/or protons, it also has a potential downside. The positively charged protons and oxygen vacancies can be expected to be electrostatically attracted to the dopants, thus introducing local low-energy positions for the migrating species, compare figure 4.6. In order for the migrating species to dissociate from this local bound state and enter the normal, conductive, state it needs to obtain an enthalpy ΔH^{t}. In consequence, the activation energy of the species in question—in its trapped state—is split further to include also the dissociation or trapping enthalpy ΔH^{t} [46]:

$$E^{\mathrm{a}} = \Delta H^{\mathrm{m}} + \Delta H^{\mathrm{f}} + \Delta H^{\mathrm{t}}. \tag{4.17}$$

Note that here the formation enthalpy refers to carriers in the trapped state. During conditions such that the trapped states are occupied (most prominently low temperatures), this formation enthalpy is small and the trapped carriers experience a migration barrier $\Delta H^{\mathrm{m}} + \Delta H^{\mathrm{t}}$. In the case of proton conductors, this trapping effect may have not only direct bearings to the ionic conductivity but trapping of oxygen vacancies may also have indirect effects in terms of proton concentration in that trapped oxygen vacancies can not migrate to the surface of the sample and take part in the hydration reaction. Trapping of protons by acceptor dopants in LZO is investigated in Paper I, and trapping of oxygen vacancies is studied in Paper II.

4.7 Depletion

The present chapter is concluded with figure 4.7, illustrating the effect of an accumulation of positive charge at the core of grain boundaries. The positive charge

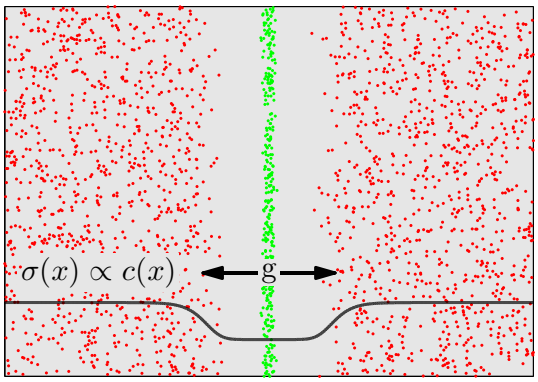


Figure 4.7: Schematic illustration of an accumulation of positive charge (green) at the core of a grain boundary causing depletion of positive charge carriers (red), for instance protons, in the grain boundary region.

electrostatically repulses positive charges in the region, causing a space-charge region depleted of positive charge carriers such as protons. Due to the suppressed concentration, the conductivity of these charge carriers will be severely hampered, with a harmful effect on the effective conductivity of the material, as discussed in section 4.2. In Papers III–IV the proton depletion associated with the accumulation of positive oxygen vacancies at the core of grain boundaries in BZO is examined.

Chapter 5

Defects in semiconductors and insulators

Ideal, crystalline, solid state materials are constructed by an infinite periodic repetiton in space of some basic building block – a unit cell consisting of one or more atoms. No finite, real-world, pieces of material are however perfect crystals. Somewhere the crystal has to end in a surface, or an interface to another crystal with a different orientation or a different composition or arrangement of atoms. Such interfaces are also commonly found inside materials, between grains of bulk material or to local precipitates of different phases. There may for instance also be dislocations between atomic planes, and imperfections due to missing, added or switched atoms. These various imperfections in a crystal can be classified by their spatial dimensionality as bulk-, plane-, line- and point defects.

Such defects can greatly affect the macroscopic properties of a material, like its ability to withstand thermal and mechanical stresses. More importantly for the subject of the present thesis, the electronic properties of the material is also affected. In a periodically symmetric material the probability of finding an electron at one point in a unit cell must be the same in the equivalent point of every unit cell in the crystal. Imperfections however break this symmetry and may introduce new electronic states associated with the imperfection. Analogously, point defects may find positions of higher or lower stability in the vicinity of other defects, be it other point defects (*cf.* trapping) or higher-dimensional ones like grain boundaries (*cf.* segregation). Accumulation of charged point defects at grain boundaries causes local electric fields and possible depletion of charge carriers. Understanding the physics of defects in semiconductors is essentially what has enabled the engineering and development of the plethora of electronic semiconductor devices which are an integrated part of our society. As seen in previous chapters, point defects are also a necessary ingredient in the ionic conductivity of solid oxides and much like the properties of electronic semiconductor devices are dictated by the electronic structure, the properties of solid oxide fuel cell electrolytes are dictated by the defect structure. By the discussion above the structure of defects and electrons are however intrically connected, and treatments of the two are in a wider perspective difficult to separate.

This chapter offers a brief introduction to point defects and grain boundaries,

with the purpose of clarifying definitions and notation (section 5.1), as well as providing an introductory model for understanding the electrical properties of point defects (section 5.2) and grain boundaries (section 5.5) in semiconductors. Furthermore the thermodynamics and chemistry of point defects is introduced in sections 5.3–5.4.

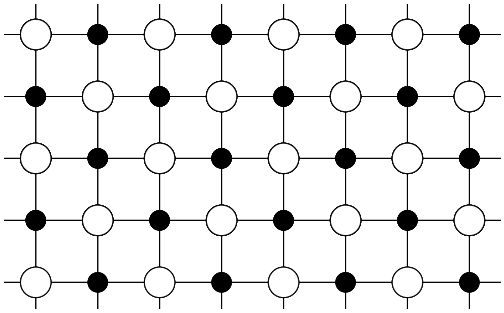
5.1 Definitions and notation

Point defects can be defined as zero-dimensional imperfections in a crystal, meaning a single atom is removed, added or replaced in a way that is not consistent with the ideal crystal lattice. In the present thesis it is practical to divide point defects into three categories: *vacancies*, *substitutional* defects and *interstitial* defects. These are illustrated in figure 5.1. Vacancies arise when a native atom is removed from its position in the crystal lattice (figure 5.1(b)). Substitutional point defects are the result of the replacement of an atom with another, either native or foreign to the material (figure 5.1(c)). Interstitial point defects, finally, are atoms occupying positions between points belonging to the crystal lattice (figure 5.1(d)).

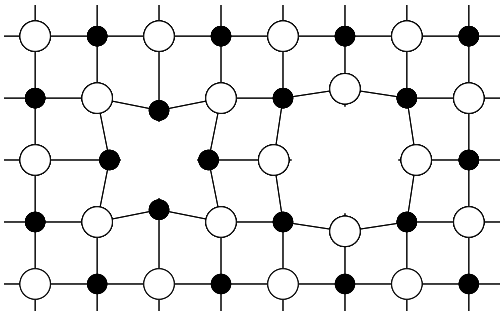
By convention the *Kröger-Vink notation* is used to denote point defects. The species M and lattice site s of the defect is expressed as M_s. For example, a substitutional point defect constructed by the incorporation of titanium on a site which in the pure material is occupied by lanthanum, is denoted $\mathrm{Ti}_{\mathrm{La}}$. For interstitial point defects the position of the defect is described with the subscript i. For instance Ti_i denotes a titanium atom in an interstitial position. Vacancies are most commonly denoted V, so that for example a vacancy on a lanthanum site is denoted V_{La}. Sometimes a lower-case v is used instead, particularly if there is a chance of confusion with Vanadium. In addition to species and position, point defects are characterized by their *effective charge*. The creation of a point defect in a semiconductor or insulator is typically accompanied by a more or less local change in charge, the magnitude of which can vary depending on the occupation of electronic states associated with the defect. Often one needs to specify to which charge state of the defect one refers, and this is done by adding a superscript to the notation. Dots, primes and crosses ($M^{\bullet}$, M' and $M^{\times}$) respectively denote a positive, negative and neutral effective charge, measured in elementary units. For example, removing an O^{2-} ion from its position in an ionic oxide nominally results in a vacancy where the charge has increased by two units, why the point defect would be denoted $V_{\mathrm{O}}^{\bullet\bullet}$. Placing the oxygen ion in an interstitial position would similarly result in an O_i'' defect, while replacing say a La^{3+} with a Nd^{3+} corresponds to a $\mathrm{Nd}_{\mathrm{La}}^{\times}$ defect.

5.2 Donors and acceptors

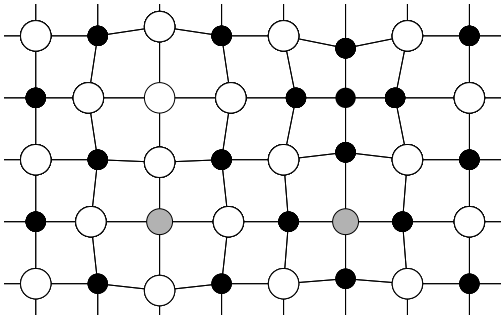
As previously mentioned, point defects in insulators and semiconductors can carry an effective charge and give rise to electronic states with energies in the band gap. Depending on the occupation of these states the charge of the defect changes. Posi-



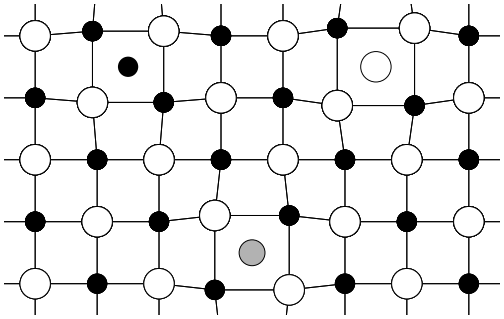
(a) Perfect crystal lattice without point defects.



(b) *Vacancies*: one on the B-site (V_B), one on the W-site (V_W).



(c) *Substitutional* point defects: two constructed from native atomic species occupying the wrong lattice site (W_B and B_W) and two from a foreign species (G_B and G_W).



(d) *Interstitial* point defects: two constructed from native atomic species (B_i and W_i) and one from a foreign species (G_i).

Figure 5.1: Point defects in a crystal constructed from two atomic species denoted B, for black, and W, for white. Foreign atomic species have been colored gray and are denoted G.

tively charged point defects have released, or donated, one or more electrons to the host material and are consequently called donors. Negatively charged point defects have instead trapped, or accepted, electrons and are called acceptors. It can be stated that the foundation for understanding these properties was laid in 1940 when Mott and Gurney described what came to be known as the effective mass or hydrogenic theory for point defects [47]. A frequently cited review on the topic was written by Kohn in 1957 [48]. The theory is outlined here in a conceptual form as an introduction to effective charge, localized states, and acceptors and donors.

The effective mass concept concerns the behaviour of electrons close to the bottom of the conduction band, and electron holes close to the top of the valence band. Close to the edges, electronic energy bands can often be estimated with a parabolic function, as depicted in figure 5.2. Recalling the energy of a free electron,

$$E_{\mathrm{e}}^{\mathrm{free}}(k) = \hbar^2 k^2 / 2m_{\mathrm{e}},$$

where m_{e} is the free electron mass, it is realized that the energy of electrons near the minimum of the conduction band, $E_{\mathrm{e}}^{\mathrm{CBM}}$, and holes near the maximum of the

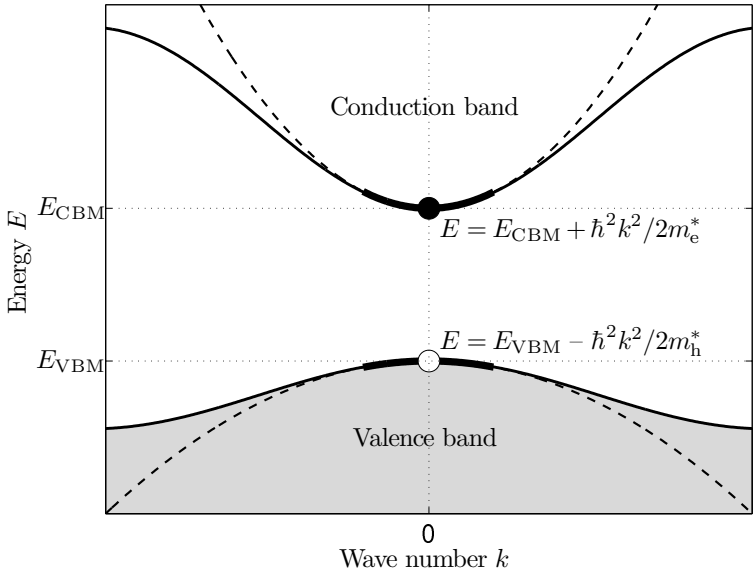


Figure 5.2: Depiction of effective masses of electrons in the conduction band and electron holes in the valence band. Close to the edges, simple bands are approximately parabolic, *ie.* $E \propto k^2$. Comparing to the energy of a free electron, $E_{\mathrm{e}}^{\mathrm{free}} = \hbar^2 k^2 / 2m_{\mathrm{e}}$, electrons and holes close to the band edges can be seen as free-like electrons with some effective masses m_{e}^* and m_{h}^* respectively.

valence band, $E_{\mathrm{h}}^{\mathrm{VBM}}$, can be written[1]:

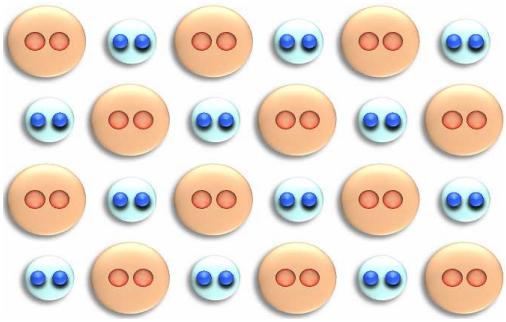
$$E_{\mathrm{e}}^{\mathrm{CBM}}(k) = E_{\mathrm{CBM}} + \frac{\hbar^2 k^2}{2m_{\mathrm{e}}^*} \tag{5.1}$$

$$E_{\mathrm{h}}^{\mathrm{VBM}}(k) = E_{\mathrm{VBM}} + \frac{\hbar^2 k^2}{2m_{\mathrm{h}}^*}, \tag{5.2}$$

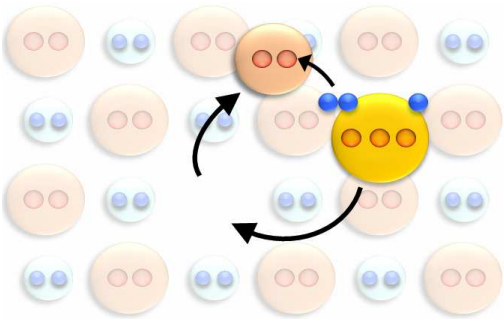
where E_{CBM} and E_{VBM} are the energies at respective band edge. Having the same dispersion relation as free electrons, electrons and holes close to the edges of the conduction- and valence bands can thus be expected to behave like such, only with some effective masses m_{e}^* and m_{h}^* respectively.

It is now time to introduce a point defect. For illustrative purposes an ionic compound will form the basis of the discussion, but the arguments can be applied to covalent crystals as well. Figure 5.3 illustrates the effective changes imposed on the system when a *donor* defect is created. For instance, replacing a divalent cation with a trivalent, a point defect with an effective charge $+1$ is created. One excess electron is also introduced to the system. This electron can be placed in the bottom of the conduction band, where it will behave like a free electron with effective mass m_{e}^* and energy given by equation 5.1. It will however be attracted by the positive charge of the defect, indicating there has to exist a bound state with energy lower than the free-like state. The situation is very similar to a free electron and a proton, an ionized

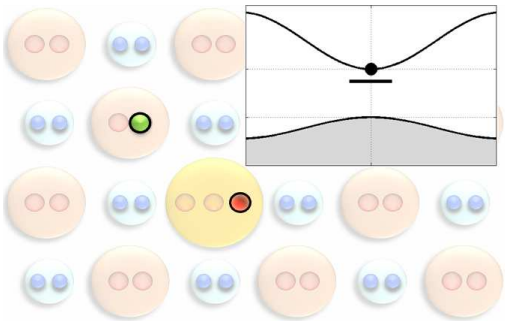
[1]Notice opposite sign in the hole energy compared to the VBM electron energy in figure 5.2; in an electron energy diagram a decrease in electron energy is an increase in hole energy.



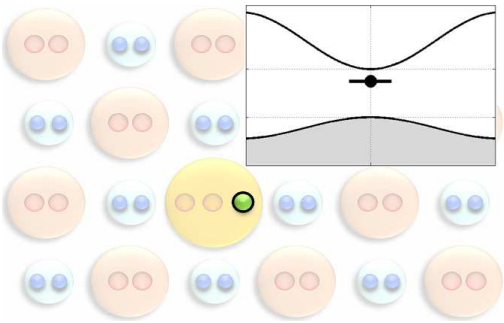
(a) Before introducing a point defect. Perfect ionic crystal constructed from two ions. The anions (small, cyan-colored) are negatively charged due to electrons occupying two valence states which are empty in the neutral atom. The cations (large, orange) are positive due to two missing valence electrons.



(b) Introducing a substitutional point defect. A trivalent atom replaces one of the divalent cations. Two of the three new valence electrons are accounted for by making the replaced cation neutral. One extra valence electron is however available to the system.



(c) The effective change to the system is the extra valence state and electron. If the electron is *not* occupying the localized state, the point defect has effectively a positive charge and the electron is in the free-like state in the conduction band. The situation is similar to a proton and free electron, *ie.* an ionized hydrogen atom.



(d) If the electron *is* occupying the localized state, the point defect is effectively neutral and the electron is bound to the point defect. The situation is similar to a hydrogen atom.

Figure 5.3: Illustration of the formation of a donor defect, in this example by the substitution of a divalent cation with a trivalent. The bound, localized, state that appears must be lower in energy than the free-like state in the conduction band. Comparing with a hydrogen atom the energy of the localized state can be estimated, as can the orbital extension of the electron (see the text). Note that an equivalent change in effective local charge and available electrons as described here occur for different point defects, for instance creating an anion vacancy. In the cases of vacancies, the localized states can be expected to be characterized by the valence states of neighboring ions instead of valence states of the substituent.

hydrogen atom. If the crystal hosting the point defect and electron is estimated with a screening medium with some dielectric constant ε, the energy of the localized defect state can be evaluated. The effective positive charge of the point defect establishes a Coulomb field, additional to all the other crystal fields that existed before the point defect was introduced. Due to the screening of the host medium, the Coulomb field is not given by e/r but instead

$$U_{\text{donor}}(r) = \frac{e/\varepsilon}{r},$$

where e is the elementary charge. The electron in the hydrogen atom has the ground state energy

$$E_{\text{e}}^{\text{hydrogen}} = -e^4 m_{\text{e}}/2\hbar^2,$$

which is in reference to the free vacuum state. Here the relevant reference energy is instead that of the free-like state at the minimum of the conduction band, and it is immediately concluded that the bound state of the donor defect has an energy

$$E_{\text{donor}} = E_{\text{CBM}} - \frac{e^4 m_{\text{e}}^*}{2\hbar^2\varepsilon^2},$$

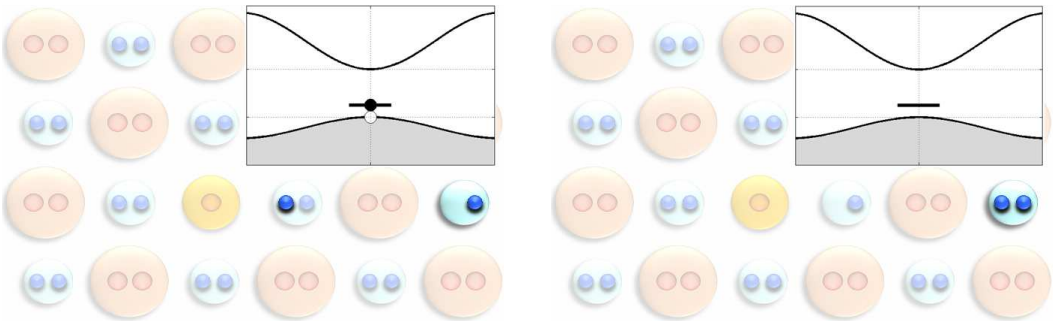
or

$$E_{\text{donor}} = E_{\text{CBM}} - E_{\text{e}}^{\text{hydrogen}} \frac{m_{\text{e}}^*}{m_{\text{e}}\varepsilon^2}. \tag{5.3}$$

The dielectric constant ε typically has a value of order 10, and the values of m_{e}^* can range in order between $0.01m_{\text{e}} - 1m_{\text{e}}$. This corresponds to a value of E_{donor} between 13.6×10^{-4} and 13.6×10^{-2} eV, or from about 1 to about 140 meV, below the conduction band. Compared to band gaps of order 1 eV or more, the archetypal donor level is thus close to the conduction band. The weak bonding of the electron also implies a large spatial extension which is far from point-like. Using the expression for the bohr radius, $\text{a}_0 = \hbar^2/m_{\text{e}}e^2 \sim 0.5$ Å, the typical extension of an electron in a donor state is between 5 and 500 Å.

The qualitative description of *acceptor* defects is analogous to that of donors. As an example, figure 5.4 illustrates the effects of replacing a divalent cation in the lattice with a monovalent. In removing a divalent atom from the system and replacing it with a monovalent, one electron has to effectively be removed from the system. This electron can be taken from bulk valence states, *ie.* the valence band, which results in an effectively negative point defect, *cf.* figure 5.4(a). It can also be taken from valence states localized at the point defect as seen in figure 5.4(b), which corresponds to a neutral point defect. The local valence electrons have lost one partnering empty valence state in the cation and thus represent a higher energy than the bulk valence states in the valence band. In other words, figure 5.4(a) represents a situation with an electron trapped in a localized defect state with energy higher than the valence band, while figure 5.4(b) represents a situation with no electron trapped or conversely a hole trapped in a bound state at the negatively charged point defect. To estimate the energy of the localized state, comparison with a hydrogen-like situtation is again possible, only the Coulomb field is set up by a negative charge,

$$U_{\text{acceptor}}(r) = -\frac{e/\varepsilon}{r},$$



(a) The electron removed from the system has been taken from the valence states of the bulk. This is illustrated with the missing electron on the anion to the far right and the hole in the valence band. Locally at the point defect, the charge has decreased with one unit, why the effective defect charge is -1. This corresponds to a local excess of one electron, which is attributed to one valence electron in one of the anions closest to the replaced cation. This electron has lost its corresponding empty valence state in the cation and is more weakly bound, meaning higher in energy, than valence electrons in the bulk (the valence band).

(b) The electron removed from the system has here been taken from the local valence states at the defect, illustrated with the missing electron on one of the anions closest to the replaced cation. Thus the point defect is neutral, with no local excess electron, and there is no electron missing in the bulk valence states. In the band picture this corresponds to the localized state not being occupied. Compared to the situation in (a), the localized electron has filled the hole in the valence band. Conversely the situation can be thought of as the positive valence band hole being trapped in a bound state at the negatively charged point defect.

Figure 5.4: Schematic illustration of how an acceptor defect can form, in this example by substituting a divalent cation with a monovalent. In order to make the removed divalent cation into a neutral atom, two electrons are needed. Since the substituent only has one valence electron, one electron must effectively be removed from the system. That electron can either be taken from bulk valence states (a), or from valence states localized at the point defect (b).

and attracts the positive hole instead of a negative electron. The bound state is by an electron seen as higher in energy than the states in the valence band and in the band picture the energy of the bound state of the hole must thus be

$$E_{\text{acceptor}} = E_{\text{VBM}} + E_{\text{e}}^{\text{hydrogen}} \frac{m_{\text{h}}^*}{m_{\text{e}}\varepsilon^2}. \tag{5.4}$$

In full analogy with donor states, values for $E_{\text{e}}^{\text{hydrogen}} m_{\text{h}}^*/m_{\text{e}}\varepsilon^2$ are small compared to the band gap and it is concluded that typical, simple, acceptor states are located in the band gap close to the valence band.

5.3 Concentration at thermodynamic equilibrium

Common for all concentration considerations in the present thesis is the assumption of thermodynamic equilibrium. In this section the thermodynamics of point defect

formation is therefore sketched, leading to expressions for the concentration of a given point defect in a crystalline material.[2]

Considering some system of which the crystalline material of interest is part, the formation of a point defect is described by a chemical equilibrium relation:

$$\text{system including perfect crystal} \rightleftharpoons \text{system including defective crystal.}$$

One may equivalently write

$$\text{defect species in source} \rightleftharpoons \text{defect species in crystal,}$$

where by source it is here meant for instance the oxygen gas surrounding the crystal in the case of formation of oxygen vacancies in ambient conditions. Equivalently the gas acts as a sink for the oxygen atom removed from the crystal. The source or sink varies depending on the defect and defect process in question and can for instance also be a different region within the crystal itself. For the purpose of the present discussion the point defect formation equilibrium is therefore written:

$$\text{defect species in } \mathrm{X}_1 \rightleftharpoons \text{defect species in } \mathrm{X}_2.$$

where $\mathrm{X}_{1,2}$ is used to denote different thermochemical "environments" or "states" of the defect species.

Equilibrium corresponds to a minimum in Gibb's free energy:

$$G = H - TS = E + pV - TS, \tag{5.5}$$

where H is the enthalpy, S the entropy, E the energy and V the volume of the system. The formation of N_i defects of type i in the crystal, X_2, imposes a change in the free energy:

$$\begin{aligned} G = G(\mathrm{X}_1) + G(\mathrm{X}_2) \longrightarrow G = & G(\mathrm{X}_1) + G(\mathrm{X}_2) \\ & + \Delta G_i(\mathrm{X}_1) + N_i \Delta G_i^{\circ}(\mathrm{X}_2) - k_\mathrm{B} T \ln \Omega_i(\mathrm{X}_2), \end{aligned} \tag{5.6}$$

where $\Delta G_i(\mathrm{X}_1)$ denotes all changes in free energy of X_1, while for X_2 the change was explicitly divided into two parts: $\Delta G_i^{\circ}(\mathrm{X}_2)$ comprises all changes in free energy related to the creation of one defect, except for the change in configurational entropy $k_\mathrm{B} \ln \Omega_i(\mathrm{X}_2)$. Here $\Omega_i(\mathrm{X}_2)$ is the number of microstates available to the point defects. If N_max denotes the number of positions available to formation of the N_i defects, the configurational entropy is:

$$\begin{aligned} k_\mathrm{B} \ln \Omega_i(\mathrm{X}_2) = k_\mathrm{B} \ln \binom{N_\mathrm{max}}{N_i} &= k_\mathrm{B} \ln \frac{N_\mathrm{max}!}{N_i!(N_\mathrm{max} - N_i)!} \\ &\approx k_\mathrm{B} N_\mathrm{max} \ln \frac{N_\mathrm{max}}{N_\mathrm{max} - N_i} - k_\mathrm{B} N_i \ln \frac{N_i}{N_\mathrm{max} - N_i}. \end{aligned} \tag{5.7}$$

where Stirling's formula was used. To find the equilibrium number of defects in the crystal the free energy is minimized by taking the derivative of G with respect to N_i

[2] A detailed treatment is found in for instance Maier [49].

and equating with zero. After noting that a change $\mathrm{d}N_i$ in X_2 is a change $-\mathrm{d}N_i$ in X_1, the result is:

$$\frac{\partial}{\partial N_i}\Delta G_i(\mathrm{X}_1) = \Delta G_i^{\circ}(\mathrm{X}_2) + k_{\mathrm{B}}T\ln\frac{N_i}{N_{\max}-N_i}. \tag{5.8}$$

At this point it is noted that a change in Gibb's free energy with respect to the number of particles is defined as the chemical potential μ of that particle: $\frac{\partial}{\partial N_i}\Delta G_i(\mathrm{X}_1) = \mu_i(\mathrm{X}_1)$. If the change in X_2 is for the moment treated the same way, the equilibrium relation equation 5.8 can be phrased:

$$\mu_i(\mathrm{X}_1) = \mu_i(\mathrm{X}_2), \tag{5.9}$$

which states that the chemical potential of species i is equal in X_1 and X_2, given thermodynamic equilibrium. Furthermore it is now possible to replace the numbers N_i and $N_{\max}$ with concentrations c_i and $c_{\max}$ in units of choice, for instance defects and sites per volume or per formula unit of the host crystal (mole). Consequently, equation 5.8 can be written:

$$\mu_i(\mathrm{X}_1) = \mu_i^{\circ}(\mathrm{X}_2) + k_{\mathrm{B}}T\ln\frac{c_i}{c_{\max}-c_i}, \tag{5.10}$$

where $\mu_i^{\circ}(\mathrm{X}_2) = \Delta G_i^{\circ}(\mathrm{X}_2)$. Exponentiation of equation 5.10 gives

$$c_i = (c_{\max}-c_i)\exp\left(-\frac{\Delta G_i^f}{k_{\mathrm{B}}T}\right), \tag{5.11}$$

where

$$\Delta G_i^f = \mu_i^{\circ}(\mathrm{X}_2) - \mu_i(\mathrm{X}_1). \tag{5.12}$$

Rewriting equation 5.11 now yields an explicit expression for the concentration of point defects of type i in the crystal X_2:

$$c_i = c_{\max}\frac{1}{1+\exp(\Delta G_i^f/k_BT)}. \tag{5.13}$$

It is seen that the occupation of the available defect sites $c_{\max}$ is given by the Fermi-Dirac distribution. For low defect concentrations, $c_i << c_{\max}$, the Boltzmann approximation can instead be used:

$$c_i = c_{\max}\exp\left(-\frac{\Delta G_i^f}{k_{\mathrm{B}}T}\right), \tag{5.14}$$

in which case the statistical decrease in occupation probability as sites are being used up by species i is lost.

The maximum concentration, $c_{\max}$, presently used refers to the number of sites available for formation of the point defect under consideration. In general, several different species can compete for the same sites (*cf.* for instance hydroxide ions and

oxygen vacancies, which together with oxygen ions compete for oxygen lattice sites in solid oxides during hydration) and

$$c_{\max} = c_{\text{tot}} - \sum_j c_j, \tag{5.15}$$

with c_{tot} the total density of lattice sites in the crystal and j running over the competing species.

Apart from the explicit statistical approximations, the derivation above silently assumed a concentration-independent change $\mu_i^\circ(\mathrm{X}_2)$. While a reasonable approximation in many cases, it may not be true for extremely high defect concentrations or highly interacting point defects. Formally, this discrepancy is rectified by defining an effective number or concentration of defects, a so-called activity a_i which is conveniently defined in the Boltzmann form so that it subsumes also the statistical approximations:

$$\mu_i = \mu_i^\circ + k_\mathrm{B} T \ln \mathrm{a}_i. \tag{5.16}$$

The activity is in turn sometimes written in terms of an activity coefficient f_i: $\mathrm{a}_i = f_i N_i / N_{\max}$. In practice, defect interaction is more readily evaluated in an approximate manner by explicitly considering pairs of defects (*cf.* trapping) or more complicated clusters. Such clusters or complexes can be included in a complete defect treatment by defining them as independent defect species.

5.4 Defect chemistry and mass action

In general defect processes, more than one constituent is involved and it is useful to find relations between the concentrations or partial pressures of the constituents. Consider therefore the following chemical reaction model:

$$a\,\mathrm{A} + b\,\mathrm{B} \leftrightharpoons c\,\mathrm{C} + d\,\mathrm{D}, \tag{5.17}$$

where a, b, c, d denotes the number of species A,B,C,D taking part in the reaction. With the definition of chemical potential from section 5.3 in place, the free energy balance can be written

$$a\,\mu_\mathrm{A} + b\,\mu_\mathrm{B} = c\,\mu_\mathrm{C} + d\,\mu_\mathrm{D}, \tag{5.18}$$

which after rearranging and using equation 5.16 reads:

$$a\,\mu_\mathrm{A}^\circ + b\,\mu_\mathrm{B}^\circ - c\,\mu_\mathrm{C}^\circ - d\,\mu_\mathrm{D}^\circ = k_\mathrm{B} T \ln \frac{\mathrm{a}_\mathrm{C}^c \mathrm{a}_\mathrm{D}^d}{\mathrm{a}_\mathrm{A}^a \mathrm{a}_\mathrm{B}^b}.$$

Let $\Delta G^\circ = a\,\mu_\mathrm{A}^\circ + b\,\mu_\mathrm{B}^\circ - c\,\mu_\mathrm{C}^\circ - d\,\mu_\mathrm{D}^\circ$. Then:

$$\exp\left(-\frac{\Delta G^\circ}{k_\mathrm{B} T}\right) \equiv K = \frac{\mathrm{a}_\mathrm{C}^c \mathrm{a}_\mathrm{D}^d}{\mathrm{a}_\mathrm{A}^a \mathrm{a}_\mathrm{B}^b}, \tag{5.19}$$

where K is called the equilibrium constant of the reaction and ΔG° is the difference in standard potentials, or the total difference in free energy except for configurational entropy, between the right and left hand sides of the reaction. The relationship

between a chemical reaction (here equation 5.17) and an expression like that of equation 5.19 is called the mass action law.

Since equation 5.19 is written in the Boltzmann form, the site restriction in a crystal lattice, equation 5.15, must be remembered when replacing the activities with concentrations. Moreover, since implicitly the interest revolves around charged acceptor and/or donor defects, consistency with respect to charge neutrality of the material as a whole has also to be taken into account.

Taking the equilibrium formation of oxygen vacancies in an undoped oxide as an example, conditions such that the removal of an oxygen atom from the material to ambient oxygen gas results in a fully ionized defect are assumed. The vacancy will then have an effective charge +2:

$$\mathrm{O}_\mathrm{O}^\times \rightleftharpoons \mathrm{V}_\mathrm{O}^{\cdot\cdot} + \frac{1}{2}\mathrm{O}_2(g) + 2\,\mathrm{e}'.$$

For gas phase constituents the activity corresponds to the partial pressure of the gas, if ideal gases are assumed. The mass action principle gives:[3]

$$K_1 = \frac{c_\mathrm{V}\, n_\mathrm{e}^2\, p_{\mathrm{O}_2}^{1/2}}{c_\mathrm{O}},$$

where c_V denotes the concentration of vacancies, n_e the concentration of electrons and c_O the concentration of non-vacant oxygen sites. If low vacancy concentration is assumed and no other defects compete for the oxygen sites the Boltzmann approximation is valid and $c_\mathrm{O} = N_\mathrm{O}$, with N_O denoting the density of oxygen sites in the lattice. With no other defects, charge neutrality is also simple: $n_\mathrm{e} = c_\mathrm{V}/2$, and an expression for the vacancy concentration as function of oxygen partial pressure is obtained:

$$c_\mathrm{V} = (4N_\mathrm{O}K_1)^{1/3} p_{\mathrm{O}_2}^{-1/6}.$$

The vacancy concentration decreases with increasing oxygen pressure.

For hydration of an acceptor-doped solid oxide, the relations become more complicated. The hydration reaction is written:

$$\mathrm{H_2O}(g) + \mathrm{V}_\mathrm{O}^{\cdot\cdot} + \mathrm{O}_\mathrm{O}^\times \rightleftharpoons 2(\mathrm{OH}_\mathrm{O}^\cdot), \tag{5.20}$$

and the mass-action relationship is

$$K_2 = \frac{c_\mathrm{OH}^2}{c_\mathrm{V}\, c_\mathrm{O}\, p_{\mathrm{H_2O}}}. \tag{5.21}$$

Further, site restriction now gives:

$$N_\mathrm{O} = c_\mathrm{O} + c_\mathrm{V} + c_\mathrm{OH} \tag{5.22}$$

[3]Note that here the formation free energy is implicitly assumed to be given at a reference oxygen pressure of 1, in the units used for p_{O_2} (tabulated values are usually in bar).

and, assuming a single acceptor dopant A, for instance Y'_{Zr} (Y^{3+} substituted for Zr^{4+}), charge neutrality dictates (here it is assumed only vacancies and protons compensate for the dopants):

$$0 = 2c_V + c_{OH} - c_A. \tag{5.23}$$

Combining equations 5.21–5.23 gives the dependence of the proton concentration on dopant concentration and water partial pressure:

$$c_{OH} = N_O \frac{\kappa}{\kappa - 4}\left[1 - \sqrt{1 - \frac{\kappa - 4}{\kappa}\left(2\frac{c_A}{N_O} - \frac{c_A^2}{N_O^2}\right)}\right], \tag{5.24}$$

where $\kappa = p_{H_2O} K_2$.

The mass-action approach and the defect chemistry outlined here have been used in Papers III and IV to evaluate the effect of hydration on the equilibrium concentrations of oxygen vacancies and protons in grain boundaries of $BaZrO_3$. Equivalently, the concentrations of point defects in a material can be found by employing equation 5.11 on the defects in question. Relationships between the various concentrations are then manifested in the relation between the chemical potentials (in their full form, *cf.* equation 5.18) of the different constituents, which appear in the formation energy ΔG^f. By introducing the chemical potential of electrons, μ_e, and self-consistently applying charge neutrality, the intricate charge balance between acceptors, donors, electrons and electron holes can be adressed, resulting in the equilibrium concentrations of the various defects. This formalism is described in chapter 7 and has been explicitly used in Papers II and V. Implicitly it is used in all papers since it enters when comparing the energy (and thus concentration) of incorporation on different lattice sites, including segregation to grain boundaries and trapping energies.

5.5 Space-charge at interfaces

Figure 5.5 illustrates how an accumulation of excess charge in a spatially localized region in a material will give rise to charge-compensating areas of so-called space charge around the region. While the figure illustrates compensation by depletion of charge carriers with equal sign as the charged region (electron holes and oxygen vacancies in the inset, protons in the large illustration), accumulation of carriers with opposite sign is equally possible and it is the concentration and mobility of different charge carriers in the material that determine the outcome. It is noted that similarly the ”excess” charge discussed here may arise from a diminished concentration of opposite charge.

As figure 5.5 suggests, boundaries between grains are frequently seen as two opposing grain surfaces, possibly with a layer of non-crystalline, amorphous, material or even secondary phases in between. Furthermore the accumulated charge is often thought of as immobile, with a frozen concentration in the core of the interface. Given such circumstances, it is reasonable to take the charge in the core

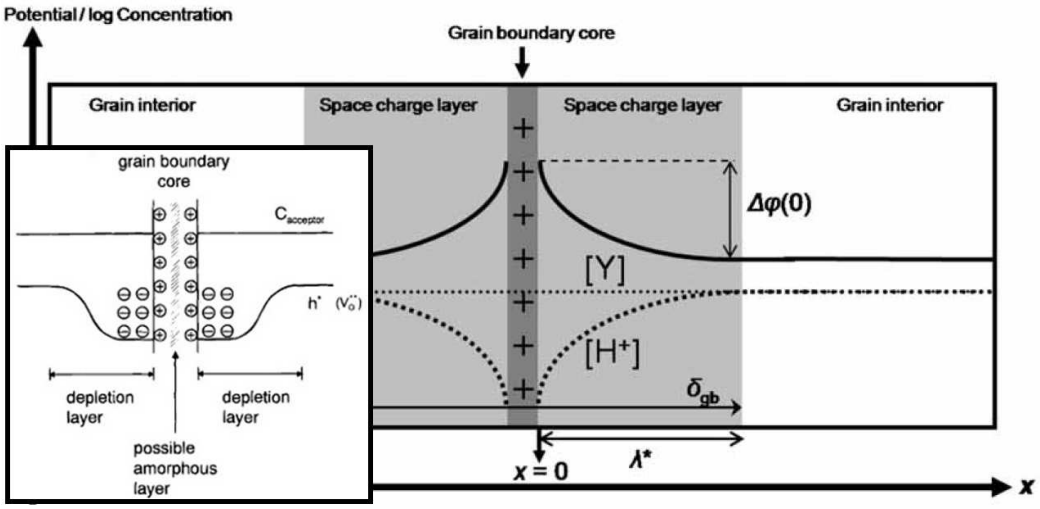


Figure 5.5: Two illustrations of excess positive charge at a grain boundary interface and resulting depletion of positive charge adjacent to the interface [22, 50]. See also figure 4.7.

as a boundary condition for finding the compensating concentration in the space-charge region, and treat the core charge as a parameter (*cf.* Gouy-Chapman and Mott-Schottky approximations [49, 51]). However, the electrostatic potential caused by the core charge will tend to deplete also the concentration of core charge itself and if the carriers of this charge are mobile, a steady-state concentration will be established, balancing the effect of the electrostatic potential and the driving force for accumulation.

Both cases however consider thermodynamic equilibrium of mobile charge carriers; in the first case only in the space charge region, while in the second case the consideration includes also the core. In a one-dimensional approximation which is usually employed, the electrochemical potential (electrochemical since charged species in an electrostatic potential is considered) of mobile charge carriers is then considered as function of distance x from the grain boundary interface and equilibrium with the grain interior ($x = \infty$) is assumed (for simplicity the Boltzmann approximation is used here):

$$\mu_i^\circ(x) + k_\mathrm{B}T \log \frac{c_i(x)}{c_i^{\max}} = \mu_i^\circ(\infty) + k_\mathrm{B}T \log \frac{c_i(\infty)}{c_i^{\max}}. \tag{5.25}$$

Only the difference in free energy between the grain interior and grain boundary region enters; $\Delta G_i^\circ(x) = \mu_i^\circ(x) - \mu_i^\circ(\infty)$. If no explicit driving force for accumulation exists, this difference is given by the potential energy obtained by the charge carrier due to the grain boundary electrostatic potential, $\Delta G_i^\circ(x) = q_i e \Delta\varphi(x)$. To find the electrostatic potential, Poisson's equation needs to be solved:

$$\frac{\mathrm{d}}{\mathrm{d}x}\Delta\varphi(x) = -\frac{1}{\varepsilon}\rho(x), \tag{5.26}$$

where $\rho(x)$ is the charge density, here determined by the concentrations of charge carriers:

$$\rho(x) = \sum_i q_i e c_i(x). \tag{5.27}$$

More generally, given the structural difference across a grain boundary even in the absence of amorphous layers or secondary phases, it is unlikely that a given charge carrier has a constant chemical potential across a grain boundary. Local differences give rise to positive or negative segregation energies, striving to enrich or deplete the grain boundary area of the carrier in question:

$$\Delta G_i^{\circ}(x) = \mu_i^{\circ}(x) - \mu_i^{\circ}(\infty) = \Delta G_i^{\mathrm{seg}}(x) + q_i e \Delta\varphi(x). \tag{5.28}$$

Equations 5.25 and 5.26 (with $\rho(x)$ given by equation 5.27 and the difference in chemical potentials by equation 5.28) constitute a system of coupled equations which must be solved simultaneously. In Papers III–V an iterative numerical scheme has been used to find the electrostatic potential and free carrier concentrations (proton, oxygen vacancies, electrons and electron holes) driven by energetic segregation ($E_{\mathrm{V}_{\ddot{\mathrm{O}}}}^{\mathrm{seg}} < 0$) of mobile oxygen vacancies to the core of grain boundaries in $BaZrO_3$.

Chapter 6

Computational method

In the research put forward in the present thesis, the energetics of point defect formation is of central importance. This chapter introduces the first principle computational atomistic method based on density-functional theory which has been the primary means to evaluating energies of point defect formation, trapping and segregation. Ab-initio or first-principle methods refer to the derivation of physical properties from fundamental quantum-mechanical principles without involving empirical models or severe simplifications of the underlying physical laws. Density-functional theory (DFT) is a means to such simplifications and has since its conception in the 1960's become a house-hold word among chemists and physicists. Numerous review articles (*eg.* [52–54]) and textbooks (*eg.* [55, 56]) cover theory and application in great detail. The aim of the present text is to provide an overview of the topic.

On the atomic scale, matter can be described as a collection of nuclei and electrons, interacting via electrostatic forces. In principle all properties of such a system can be derived from the quantum-mechanical wave function of the full system, $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{R}_1, \mathbf{R}_2, \ldots; t)$, where $\mathbf{r}_i$ and $\mathbf{R}_I$ denote the coordinates of the electrons and nuclei in the system, respectively. In theory it is possible to determine the wavefunction using the time-independent Schrödinger equation:

$$\mathcal{H}\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{R}_1, \mathbf{R}_2, \ldots) = E\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{R}_1, \mathbf{R}_2, \ldots). \tag{6.1}$$

In anticipation of doing so it is straight-forward to write the Hamiltonian $\mathcal{H}$ of the system:

$$\begin{aligned} \mathcal{H} = &-\frac{1}{2}\sum_I \frac{1}{M_I}\nabla_I^2 + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &-\frac{1}{2}\sum_i \nabla_i^2 \quad\quad + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{I,I}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|}, \end{aligned} \tag{6.2}$$

where atomic units, $\hbar = m_e = e = 1$ were used, and Z_I, M_I respectively denote the charge and mass of the nuclei. However, a finite-size piece of material contains a number of nuclei and electrons on the order of Avogadro's constant, $N_\mathrm{A} \sim 10^{23}$. Solving any equation with that number of degrees of freedom is clearly a formidable task, far beyond the capabilities of any theoretical or computational methods. It is clear that simplifications and approximations are required.

6.1 The Born-Oppenheimer approximation

A common starting point in simplifying many-atom problems is to separate the nuclear and electronic degrees of freedom. While the electrostatic forces acting on the atomic nuclei are comparable in magnitude to the forces on the electrons, the mass of the nuclei is much larger. This fact implies that the dynamics of the heavy nuclei is much slower than the dynamics of the light electrons. In the Born-Oppenheimer approximation [57], the electrons are assumed to adiabatically follow the nuclear motion. In other words, at every instance on the time scale of the nuclei, the electrons relax to a ground state given by the potential of the nuclei in the current configuration. This enables the separation of the many-atom problem into an electronic problem and a nuclear problem.

6.2 Electronic structure: density-functional theory

Having separated the electronic and nuclear degrees of freedom, the electronic problem is solved for a fixed nuclear configuration giving rise to an external potential V_{ext}. The Hamiltonian for the electronic system can then be written

$$\mathcal{H}_e = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + V_{\text{ext}}. \tag{6.3}$$

Still the complexity of this problem poses a tremendous computational challenge due to the electron-electron interaction expressed in the second term of equation 6.3, which prevents the use of efficient diagonalization techniques. It is in reducing the complexity by removing these non-diagonal terms that density functional theory enters.

The essence of DFT is to transform the problem of determining the *wave-function* of the system of *interacting* electrons into the problem of determining the electronic *density*, which can be found from a system of *independent* electrons.

In 1964 Hohenberg and Kohn [58] formulated a theorem in two parts regarding the ground-state electronic density. First, the external potential V_{ext} of an electronic system is up to a constant uniquely determined by the ground-state electronic density $n_0(\mathbf{r})$. Since formally the wave-function is determined by the external potential, the direct implication is that all properties of an electronic system is completely determined by the electronic density. In particular this should be true for the energy of the system, and the second part of the theorem concerns the relationship between the ground state energy and the electronic density: There exists an energy functional

$$E[n(\mathbf{r})] = F[n(\mathbf{r})] + \int n(\mathbf{r})V_{\text{ext}}[n(\mathbf{r})]d\mathbf{r}, \tag{6.4}$$

whose minimum is given by the ground state density $n_0(\mathbf{r})$ and corresponds to the ground state energy E_0 of the system:

$$E_0 = E[n_0(\mathbf{r})] = \min_{n(\mathbf{r})} E[n(\mathbf{r})]. \tag{6.5}$$

In principle it is thus possible to find the ground-state energy and electron density by minimizing the energy functional, which is a 3-dimensional problem in contrast to the original $3N$-dimensional problem. Unfortunately the complex many-body effects are inescapable and so far no approximations have been introduced. An explicit form of $F[n(\mathbf{r})]$, which contains all effects of electron-electron interaction, is therefore not known.

In 1965 Kohn and Sham [59] introduced a practical scheme for computing the ground state energy and electronic density. The key idea was that the Hohenberg-Kohn theorem only applies to the electronic density, not the specific properties of the electronic system. The density could for instance be derived from a fictitous system of *independent* electrons. Kohn and Sham proposed to consider such a system and isolate the many-body effects of the true system into one unknown energy term. Kohn and Sham showed that it is then possible to write an explicit expression of the energy functional in the following form:

$$E_{\mathrm{KS}}[n(\mathbf{r})] = T_s[n(\mathbf{r})] + E_{\mathrm{H}}[n(\mathbf{r})] + E_{\mathrm{xc}}[n(\mathbf{r})] + \int n(\mathbf{r}) V_{\mathrm{ext}}[n(\mathbf{r})] d\mathbf{r}, \tag{6.6}$$

where $T_s[n(\mathbf{r})]$ is the kinetic energy functional of the independent electrons and $E_{\mathrm{H}}[n(\mathbf{r})]$ is the Hartree energy, the electrostatic energy stored in the charge density $n(\mathbf{r})$. Finally $E_{\mathrm{xc}}[n(\mathbf{r})]$ is the energy due to exchange (the Pauli principle), correlation (Coulomb repulsion) and many-body contributions to kinetic energy in the real, interacting, system of electrons. This exchange-correlation functional is consequently what now contains all many-body effects of the original problem. Minimizing the Kohn-Sham functional of equation 6.6 leads to the Kohn-Sham equations:

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{KS}}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = E_i \phi_i(\mathbf{r}), \tag{6.7}$$

where $V_{\mathrm{KS}}(\mathbf{r})$ is given by

$$V_{\mathrm{KS}}(\mathbf{r}) = V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r}) + V_{\mathrm{ext}}(\mathbf{r}), \tag{6.8}$$

and in turn $V_{\mathrm{H}}(\mathbf{r}) = \int (n(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'|) d\mathbf{r}'$ and $V_{\mathrm{xc}}(\mathbf{r}) = \delta E_{\mathrm{xc}}[n(\mathbf{r})]/\delta n(\mathbf{r})$. Equation 6.7 is a system of Schrödinger equations which are coupled via the electronic density $n(\mathbf{r}) = \sum_i f_i |\phi_i|^2$, where f_i is the occupation number of the independent-electron orbital ϕ_i. The equations can be solved iteratively given an initial guess of $n(\mathbf{r})$ until a self-consistent solution is found. The solution is the ground-state electronic density $n_0(\mathbf{r})$, and it can be shown that the corresponding ground state energy is given by

$$E_0 = \sum_i f_i E_i - E_{\mathrm{H}}[n_0(\mathbf{r})] + E_{\mathrm{xc}}[n_0(\mathbf{r})] - \int n_0(\mathbf{r}) V_{\mathrm{xc}}(\mathbf{r}) d\mathbf{r}. \tag{6.9}$$

Still, no approximations have been introduced, and if an exact form of the exchange-correlation functional could be found, the Kohn-Sham approach would lead to the exact ground-state energy and electron density corresponding to the original electronic problem. Such an exact form is however not known, and approximations must be used.

To recapitulate, what is accomplished with DFT is to trade the problem of the system of interacting electrons (*cf.* equation 6.3) for a problem involving independent particles, which has to be solved iteratively (equation 6.7) and in practice involves an approximation of the electron-electron interaction. Computationally speaking the latter turns out to be favorable since fast diagonalization schemes can be employed. The complexity of the original problem however persists in the form of finding a proper exchange-correlation approximation. Fortunately, it turns out that relatively simple schemes usually yield surprisingly accurate results.

Exchange-correlation approximations

In the Kohn-Sham approach to DFT, the effects of electron-electron interaction are built into the exchange-correlation (xc) functional $E_{\mathrm{xc}}[n(\mathbf{r})]$, *cf.* equation 6.6. Here the basic properties of two of the simplest and most commonly employed approximations, namely LDA and GGA are briefly outlined.

The first approximation of the xc functional was proposed by Kohn and Sham in their original paper [59] and became known as the local density approximation (LDA). In LDA, a local xc energy density $\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}$ is defined, and in every point in space estimated with that of a homogeneous electron gas of density $n(\mathbf{r})$. The xc functional can then be written

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n(\mathbf{r})] = \int n(\mathbf{r})\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}\big(n(\mathbf{r})\big)d\mathbf{r}. \tag{6.10}$$

The essential benefits of LDA is arguably its simplicity and relative unambiguity. An exact expression for the exchange energy ε_{x} of a homogeneous electron gas is known, and the correlation energy ε_{c} can be calculated with Monte Carlo simulations [60]. The xc energy is given simply as the sum of the two: $\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}} = \varepsilon_{\mathrm{x}} + \varepsilon_{\mathrm{c}}$.

The electron density is rarely homogeneous. A natural improvement to LDA is to take local gradients of the density into account. This is exactly what the popular generalized gradient approximation (GGA) of the xc energy does. In other words, the xc energy is considered a function of the local density and its gradient; $\varepsilon_{\mathrm{xc}}^{\mathrm{GGA}} = \varepsilon_{\mathrm{xc}}^{\mathrm{GGA}}\big(n(\mathbf{r}), \nabla n(\mathbf{r})\big)$. GGA is however not one single approximation. It exists in many flavors, and expansions to different gradient orders have been experimented with. Functionals are constructed using both theoretical methods involving formal requirements posed by sum-rules, long-range decay etc., and by empirical fitting of parameters, aimed at reproducing experimental results on bond lengths, bond angles etc., found in molecular databases [56].

In practice, the choice of which xc functional to use often comes down to testing and empirical knowledge from comparison with experiments. In the present context of semiconductors and especially point defects in semiconductors, particular attention is warranted in regards to the size of the band gap. This however deserves a separate discussion, see section 8.2.

Practical implementation: plane waves, supercells and pseudopotentials

It has been established that the Kohn-Sham equations simplify the original electronic problem. Apart from the briefly mentioned iterative procedure, a solution scheme has however not yet been discussed.

A common computational approach is to transform the Schrödinger equations into a linear eigenvalue problem by expanding the Kohn-Sham wave-functions in some basis set. Different basis sets are clearly possible, but due to the inherent periodicity in crystalline system it is for condensed matter-investigations natural to use periodic boundary conditions and choose a basis set of plane waves. As a result of Bloch's theorem [61], a complete basis set is then given by plane waves with wave vectors $\mathbf{k}$ within the first Brilluin zone, according to:

$$\Psi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n,\mathbf{k}+\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \tag{6.11}$$

where n is the index of the eigensolutions and corresponds to different bands. The sum runs over all reciprocal lattice vectors $\mathbf{G}$, but is in practice truncated by choosing some cut-off energy E_{cut} and for each $\mathbf{k}$ only include lattice vectors such that $\frac{1}{2}|\mathbf{k}+\mathbf{G}| \leq E_{\text{cut}}$.

For perfect, crystalline, bulk systems this approach is ideal since calculations representing an infinite piece of material only need to be made for atoms of one unit cell. When studying non-periodic systems, large cells are however needed and the periodically repeated unit is called a supercell. For non-periodic systems special attention needs to be made in regards to the size of the supercell, which will be discussed in the context of point defects in section 8.1.

The downside of using plane-wave basis functions is the slow convergence of the sum for rapidly varying functions. Close to nuclei, electronic wave functions oscillate heavily, why plane wave basis sets are not ideal. Fortunately, while the electrons close to the nucleus can be polarized, they do not take active part in chemical bonding. It is therefore common to only include the valence electrons in DFT calculations. The potential of a nucleus is then replaced with a pseudopotential which includes the combined potential of the nucleus and the core electrons. Pseudopotentials are constructed to produce wave functions coinciding with the all-electron wavefunctions outside some cut-off radius, while inside the cut-off they are modified for a plane-wave friendly, smooth, behavior of the wave function.

6.3 Nuclear structure

In standard DFT-based computation schemes, the nuclear problem, or more accurately ionic problem in the case of pseudopotentials, is treated classically. The electronic problem, including the ground-state energy, depends parametrically on the ionic configuration. This means the total energy E^{tot} of the ion-electron system does

as well, and that the force $\mathbf{F}_I$ on ion I is given by

$$\mathbf{F}_I = -\frac{\partial E^{\text{tot}}}{\partial \mathbf{R}_I},$$

This force is readily obtained, using the Hellman-Feynman theorem [62], as the expectation value of the corresponding gradient of the Hamiltonian. For instance the ionic equilibrium configuration can then be determined by employing some minimization algorithm (*eg.* conjugate gradient, Newton schemes) and iterative solutions of the electronic and ionic problems.

Chapter 7

Ab initio modeling of point defect formation

During the last two decades, the formalism necessary for successfully using the computational methods based on density-functional theory to calculate the stability and electronic properties of point defects has been well established [63–73]. On the basis of thermodynamic equilibrium considerations, defect concentrations produced during different chemical environments, temperatures and pressures can be obtained from realistic formation energies calculated with respect to atomic chemical potentials. Similarly the equilibrium distribution of electrons over acceptor-, donor- and host states can be readily taken into account by self-consistent determination of the Fermi level of the system. The deliberations outlined here have been explicitly applied in Papers II and V but are of direct relevance for the work in all papers as the evaluation of trapping- and segregation energies involves comparing formation energies.

The present chapter takes its origin from the discussion on point defect concentration given in section 5.3 and aims to survey the use of ground-state energy output of DFT calculations in evaluating free energies of formation, $\Delta G^{\mathrm{f}}(T,p)$, of point defects. First, in section 7.1, an expression for the formation energy, $\Delta E_0^{\mathrm{f}} \equiv \Delta E^{\mathrm{f}}(T=0,p=0)$, of a point defect at zero temperature and pressure is derived. In section 7.2 methods for treating the atomic chemical potentials occurring in the formation energy are exemplified. The chemical potential of electrons is separately dealt with in section 7.3, and in section 7.4 the calculation and meaning of thermodynamic charge transition levels is commented on. To conclude the chapter, treatments of finite temperature and pressure are addressed in section 7.5.

7.1 Formulation at zero temperature and pressure

The energies before and after formation of a point defect are illustrated in figure 7.1 and comprise first and foremost the total internal energy of the crystal: $E^{\mathrm{tot}}_{\text{material w/o defect}}$ and $E^{\mathrm{tot}}_{\text{material w/ defect}}$ respectively. Furthermore the chemical potentials of the reference states with which atoms are exchanged need to be taken into

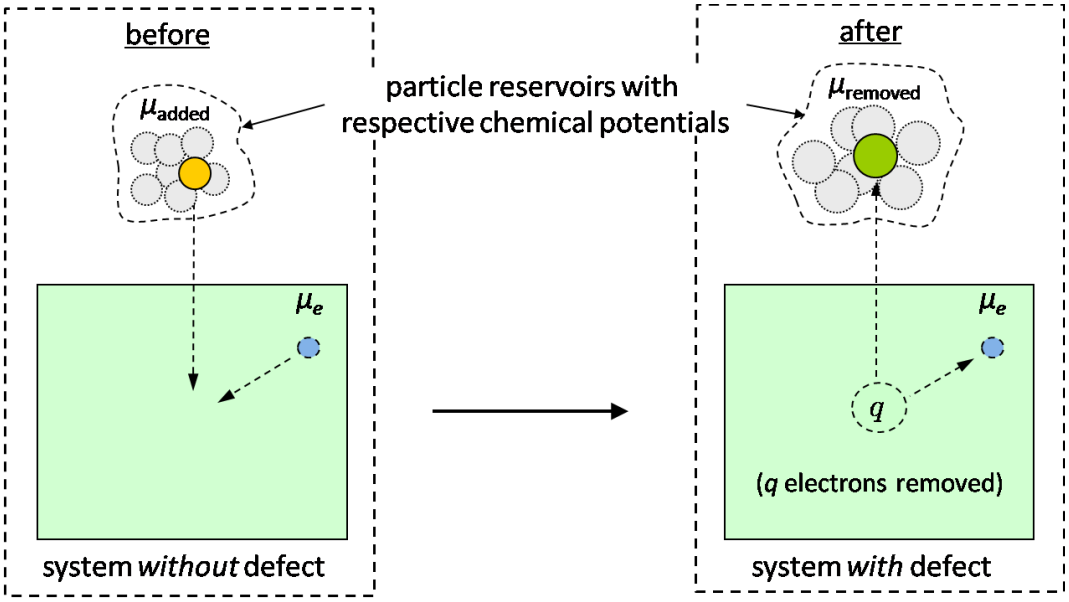


Figure 7.1: Formation of a point defect. Atoms added to or removed from the material are exchanged with corresponding reservoirs. Similarly the point defect may carry an effective charge, which means during formation one or more electrons are exchanged with the surrounding host material, acting as a reservoir for electrons.

account. In general the chemical potentials are treated as parameters except for well-defined situations where the values are pinned by thermodynamic equilibrium with specified material phases. Such pinning constraints are used to dictate limits on the chemical potentials, which will be illustrated in section 7.2. Presently the chemical potentials can be thought of as the energy cost of exchanging atoms with a reservoir of such atoms. If an atom is added to the material, the energy before formation of the defect comprises the chemical potential of the atom in its reservoir. Similarly, when defect formation involves removing an atom from the material, the energy after formation will have a contribution from the chemical potential of the removed atom in its corresponding reservoir. Vacancies, for instance, involve placing one atom in such a reservoir, while substitutional doping comprise both placing a host atom of the crystal in its reservoir and taking a dopant atom from its reservoir.

In the case of point defects carrying an effective charge q, electrons are during formation added to or removed from the energy levels associated with the defect. These electrons have been taken from, or placed in, the environment of the point defect, which usually only constitutes the host material. The energy of placing an electron in a material is, analogously to atom reservoirs, given by the chemical potential of electrons in the material, denoted μ_e. For a defect carrying an effective charge q measured in elementary charges, q electrons have been transferred between the defect and the surrounding host material. The energy related to electronic exchange can thus be expressed by adding $q\mu_\mathrm{e}$ to the energy after formation, in which case the situation with negative effective charges is included in the sign of q.

Concerning the energy of the crystal containing a charged defect, figure 7.2 il-

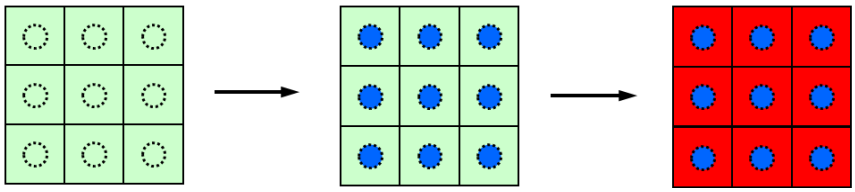


Figure 7.2: Modeling charged point defect in DFT. The number of electrons in the supercell is modified corresponding to the effective charge. To avoid diverging Coulomb energy due to interacting supercell images, a neutralizing homogeneous background charge is added.

lustrates the standard approach to model such charged point defects in DFT. First the number of electrons used in the calculation is modified according to the charge. For example a defect with charge $+2$ is modeled by using two electrons fewer than what would give a neutral system, while for negative charges electrons are instead added. The change in the number of electrons alter the occupation of the defect-associated states and the charge related to the defect. In DFT implementations employing periodic boundary conditions, an artificial homogeneous background charge of equal magnitude but opposite sign is added to neutralize the supercell.

To summarize, the energies before and after formation of a point defect can be written

$$\begin{aligned} E_{\text{before}} &= E^{\text{tot}}_{\text{material w/o defect}} + \mu_{0,\text{added atom}} \\ E_{\text{after}} &= E^{\text{tot}}_{\text{material w/ defect}} + \mu_{0,\text{removed atom}} + q\,\mu_{0,e}, \end{aligned}$$

where a zero in the subscripts was added as a reminder of zero temperature and pressure. After some rearrangment the final expression is:[1]

$$\begin{aligned} \Delta E^{\text{f}}_0 = & E^{\text{tot}}_{\text{material w/ defect}} - E^{\text{tot}}_{\text{material w/o defect}} \\ & + \mu_{0,\text{removed atom}} - \mu_{0,\text{added atom}} + q\,\mu_{0,e}. \end{aligned} \tag{7.1}$$

7.2 Chemical potential limits

It was noted in equation 5.18 and preceding text that thermodynamic equilibrium establishes relations between the chemical potentials of the components of the equilibrated system. Such relations can be used to pinpoint the values of the chemical potentials given known equilibrium conditions, and to find limits based on knowledge regarding phase stabilities [74].

To illustrate these deliberations, the pyrochlore composition $A_2B_2O_7$ is taken as an example and the three chemical potentials $\mu_{0,A}$, $\mu_{0,B}$ and $\mu_{0,O}$ are considered. The basic assumption is that the $A_2B_2O_7$-phase is stable (or at least metastable), which at equilibrium gives an equation involving the chemical potentials:

$$2\mu_{0,A} + 2\mu_{0,B} + 7\mu_{0,O} = \mu_{0,A_2B_2O_7} = E^{\text{tot}}[A_2B_2O_7], \tag{7.2}$$

[1] Several atoms are often added or removed during formation of the defect, with straight-forward generalization of equation 7.1.

where $E^{\mathrm{tot}}[\mathrm{A_2B_2O_7}]$ is the total energy of one formula unit $\mathrm{A_2B_2O_7}$, as given by a DFT calculation. Here the ratios of the constituents may vary, and in order to determine all three potentials, two more equilibrium equations are needed. A priori knowledge about which stable phases the constituents A, B and O can form can be used to consider different situations defined by simultaneous equilibrium with two of those phases at the time. Assume for simplicity that A has no stable bulk phase and only one stable binary oxide phase, $\mathrm{A_2O_3}$. Assume further that also B has only one stable oxide phase, $\mathrm{BO_2}$, but can also exist as bulk. Oxygen can exist in any of the oxide phases but also in the "bulk" gas phase: $\mathrm{O_2}(g)$. In principle six different points in a phase diagram can then be examined, each corresponding to equilibrium with $\mathrm{A_2B_2O_7}$ and two of the other phases, see table 7.1. Having used these equations to find the values of the chemical potentials, formation energies corresponding to each of the specific chemical environments may be calculated.

Usually it is not motivated to study all environmental conditions outlined in table 7.1. First, synthesis of a material for the most part takes place in equilibrium with an atmosphere containing oxygen. The value of $\mu_{0,\mathrm{O}}$ is then pinned, leaving only points 1–3. Second, the values of the metal atoms are in an oxygen-rich environment often limited by the formation of binary oxides, in the present example $\mathrm{A_2O_3}$ and $\mathrm{BO_2}$, and it is natural to consider only A- and B-rich environments, defined by chemical equilibrium with respective binary oxide. Only points 1 and 3 now remain. Each of these points correspond to a maximum value of respective chemical potential, $\mu_{0,\mathrm{A}}$ or $\mu_{0,\mathrm{B}}$. The A-rich limit of point 1 is defined by chemical equilibrium with $\mathrm{A_2O_3}$ and corresponds to an upper bound on $\mu_{0,\mathrm{A}}$:

$$\mu_{0,\mathrm{A}}^{\max} = \frac{1}{2}E^{\mathrm{tot}}[\mathrm{A_2O_3}] - \frac{3}{2}\mu_{0,\mathrm{O}}.$$

Table 7.1: Example environments as described in the text, with simultaneous equilibrium between two A-B-O phases. In conjunction with chemical equilibrium with $\mathrm{A_2B_2O_7}$ and equation 7.2, all three native chemical potentials $\mu_{0,\mathrm{A}}$, $\mu_{0,\mathrm{B}}$ and $\mu_{0,\mathrm{O}}$ can be computed at each of the six points, enabling the calculation of formation energies corresponding to each specific chemical environment. As described in the text it is however usually not necessary to consider all these combinations.

	phase 1	phase 2	equilibrium equation 1	equilibrium equation 2
1	O	$\mathrm{A_2O_3}$	$2\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{O_2}]$	$2\mu_{0,\mathrm{A}} + 3\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{A_2O_3}]$
2	O	B	$2\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{O_2}]$	$n\mu_{0,\mathrm{B}} = E^{\mathrm{tot}}[\mathrm{B}_n]$
3	O	$\mathrm{BO_2}$	$2\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{O_2}]$	$\mu_{0,\mathrm{B}} + 2\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{BO_2}]$
4	$\mathrm{A_2O_3}$	B	$2\mu_{0,\mathrm{A}} + 3\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{A_2O_3}]$	$n\mu_{0,\mathrm{B}} = E^{\mathrm{tot}}[\mathrm{B}_n]$
5	$\mathrm{A_2O_3}$	$\mathrm{BO_2}$	$2\mu_{0,\mathrm{A}} + 3\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{A_2O_3}]$	$\mu_{0,\mathrm{B}} + 2\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{BO_2}]$
6	B	$\mathrm{BO_2}$	$n\mu_{0,\mathrm{B}} = E^{\mathrm{tot}}[\mathrm{B}_n]$	$\mu_{0,\mathrm{B}} + 2\mu_{0,\mathrm{O}} = E^{\mathrm{tot}}[\mathrm{BO_2}]$

Similarly, point 3 corresponds to the B-rich limit and the maximum value of $\mu_{0,\mathrm{B}}$:

$$\mu_{0,\mathrm{B}}^{\max} = E^{\mathrm{tot}}[\mathrm{BO}_2] - 2\mu_{0,\mathrm{O}}.$$

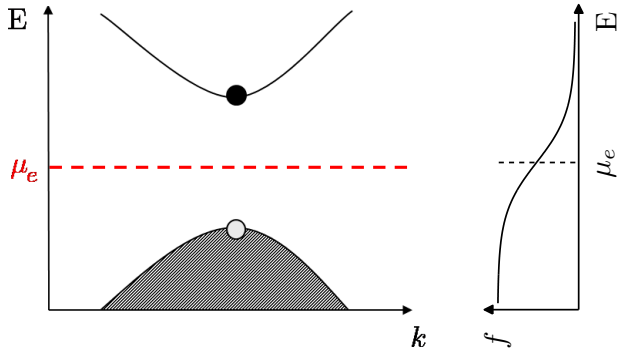
In the A-rich limit $\mu_{0,\mathrm{B}}$ is at its minimum value, which can be calculated with equation 7.2. Similarly for $\mu_{0,\mathrm{A}}$ in the B-rich limit. Thus only one degree of freedom in the chemical potentials of the native atoms remains, with bounds defined by the A- and B-rich limits. Notice here that at finite temperatures the values of all chemical potentials in this situation depends on the oxygen partial pressure (and temperature- and pressure variation in the enthalpy and vibrational entropy of oxygen). Compare section 7.5.

For point defects involving a foreign species M, for instance dopants or impurities, one more chemical potential comes into play. The value of this potential can be calculated in a similar manner as previously described, by considering different stable phases involving M. However, important conclusions can be drawn without doing so, specifically by studying properties given by differences in formation energies, for which the chemical potentials cancel. First, thermodynamic charge transition levels are calculated by comparing the formation energy of different charge states of a given defect. Secondly, by comparing the formation energy of a defect incorporated on different lattice sites, the distribution of the defect concentration over the different sites can be calculated. Similarly, in evaluation of trapping and segregation energies all reservoir chemical potentials cancel.

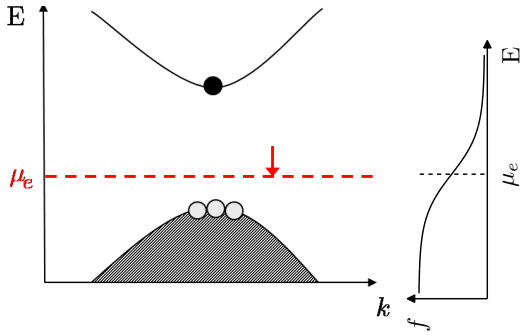
7.3 Fermi level and self-consistent concentrations

The last term appearing in equation 7.1 involves the electronic chemical potential, μ_{e}. Just like atomic chemical potentials this is a free energy related to the cost of adding one electron to a reservoir of electrons. Also in full analogy to atomic chemical potentials, the electronic chemical potential only takes a distinct value in specific equilibrium situations, determined by the thermodynamic conditions and electrochemical environment, as will be explained in the following. In such a specific environment, the pinned value of the electrochemical potential is suitably defined as the Fermi level of the material.

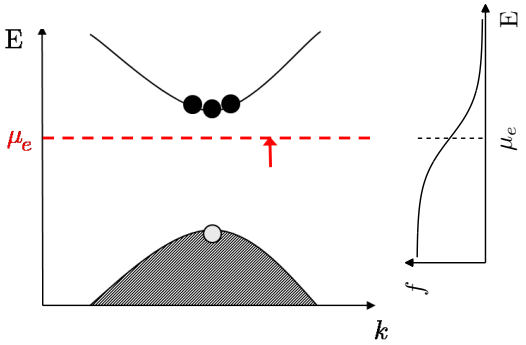
The appearance of $q\,\mu_{\mathrm{e}}$ in equation 7.1 measures the energetic cost associated with the exchange of q electrons between the defect and the electrochemical environment. As illustrated in figure 7.3 the value of μ_{e} will in a pure (non-defective) material be pinned to the center of the band gap (provided the density of states are equal in the valence and conduction bands), but in the presence of acceptor- and/or donor defects the value depends on the relative concentration and electronic properties of the different defects. Without detailed knowledge regarding the formation energy of all possible defects and their different charge states, μ_{e} can formally not be pinpointed and must be treated as a parameter. If that knowledge on the other hand can be gathered, it is possible to calculate the value of μ_{e} by considering charge neutrality. In a neutral material the concentration of all effective charge contributions must sum to zero. If $c_i(q_i)$ denotes the concentration of a certain point defect



(a) Pure (non-defective) semiconductors and insulators: The finite-temperature smearing of the Fermi function (right) is symmetric around μ_e. Assuming the density of states are equal near the edges of the valence- and conduction bands, the resulting hole-electron formation is symmetric around the center of the band gap. Hence the value of μ_e (the Fermi level) must be in the center of the band gap.



(b) With *acceptor* defects present: Electrons from the valence band can be trapped in the states associated with the defects, which reduces the concentration of electrons in the valence band. The Fermi function, and with it μ_e, must then shift downwards.



(c) With *donor* defects present: Electrons can be donated from the defect to the conduction band, leading to an increased concentration of electrons in the conduction band and a corresponding upwards shift of the Fermi function and μ_e.

Figure 7.3: Schematic illustration of the position of the Fermi level (the value of μ_e) in the band gap of insulators and semiconductors. Note that donors and acceptors regularly co-exist in a material, in which case the concentration of electrons donated to the conduction band competes with the concentration of holes created by acceptors in the valence band. The position of the Fermi level is then determined by the relative concentration of different defects in the material and the energy involved in moving electrons between them and the bands. Equation 7.3 expresses this relationship and is a means to calculate the Fermi level.

i in a specific state with charge q_i, and the concentration of electrons in the conduction band by n_{CB} and the concentration of holes in the valence band by p_{VB}, the charge-neutrality condition can be written:

$$\sum_i \sum_{q_i} q_i c_i(q_i) + p_{\mathrm{VB}} - n_{\mathrm{CB}} = 0, \tag{7.3}$$

where the outer sums runs over all species of point defects present in the system and the inner runs over all charge states of respective defect.

To solve equation 7.3, each of the relevant concentrations must be determined so that the only unknown is μ_{e}. In the Boltzmann approximation the dependence of the concentrations $c_i(q_i)$ on μ_{e} follows (*cf.* equations 5.14 and 7.1):

$$c_i(q_i) \propto e^{-q_i \mu_{\mathrm{e}}/k_{\mathrm{B}}T}. \tag{7.4}$$

The concentrations of conduction band electrons and valence band holes are found by integrating over the density of states $g(E)$ in the conduction band (CB) and valence band (VB) respectively:

$$\begin{aligned} n_{\mathrm{CB}} &= \int_{\mathrm{CB}} f(\mu_{\mathrm{e}}, E) g(E) \mathrm{d}E \\ p_{\mathrm{VB}} &= \int_{\mathrm{VB}} [1 - f(\mu_{\mathrm{e}}, E)] g(E) \mathrm{d}E, \end{aligned} \tag{7.5}$$

where $f(\mu_{\mathrm{e}}, E)$ is the Fermi function. Using the Boltzmann approximation of the Fermi function, equation 7.5 can be reduced to (*cf. eg.* reference 75):

$$\begin{aligned} n_{\mathrm{CB}} &= N_{\mathrm{CB}} e^{-(E_{\mathrm{CBM}} - \mu_{\mathrm{e}})/k_{\mathrm{B}}T} = \left(N_{\mathrm{CB}} e^{-E_{\mathrm{CBM}}/k_{\mathrm{B}}T} \right) e^{\mu_{\mathrm{e}}/k_{\mathrm{B}}T} \\ p_{\mathrm{VB}} &= N_{\mathrm{VB}} e^{-(\mu_{\mathrm{e}} - E_{\mathrm{VBM}})/k_{\mathrm{B}}T} = \left(N_{\mathrm{VB}} e^{E_{\mathrm{VBM}}/k_{\mathrm{B}}T} \right) e^{-\mu_{\mathrm{e}}/k_{\mathrm{B}}T}, \end{aligned} \tag{7.6}$$

where E_{CBM} and E_{VBM} are the energies at the conduction band minimum and valence band maximum, respectively. N_{CB} and N_{VB} in turn denote the effective density of states in the conduction- and valence bands, and are given by

$$\begin{aligned} N_{\mathrm{CB}} &= 2 \left(\frac{2\pi m_{\mathrm{h}}^* k_{\mathrm{B}} T}{h^2} \right)^{3/2} \\ N_{\mathrm{VB}} &= 2 \left(\frac{2\pi m_{\mathrm{e}}^* k_{\mathrm{B}} T}{h^2} \right)^{3/2}, \end{aligned} \tag{7.7}$$

where m_{h}^* and m_{e}^* are respectively the effective masses of holes in the conduction band and electrons in the valence band. Note that that for the Boltzmann approximation of the Fermi function to be valid, conditions such that $E_{\mathrm{CBM}} - \mu_{\mathrm{e}} \gg k_{\mathrm{B}}T$ and $\mu_{\mathrm{e}} - E_{\mathrm{VBM}} \gg k_{\mathrm{B}}T$ must hold.

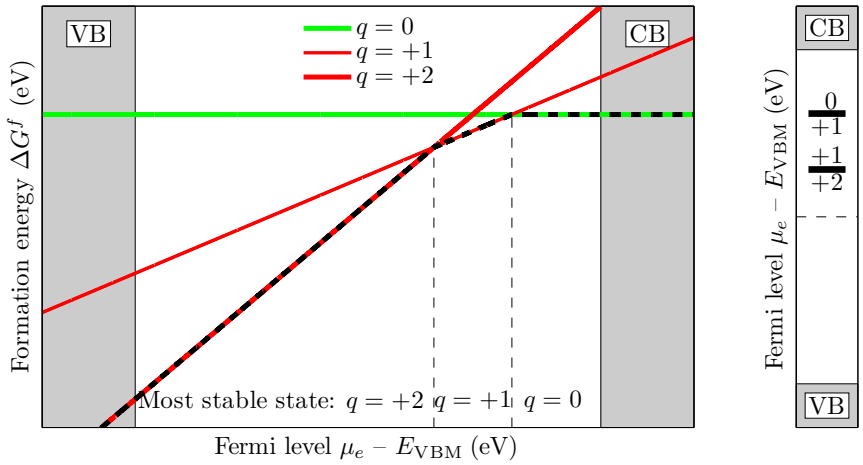


Figure 7.4: *In the figure, VB, VBM and CB stand for valence band, valence band maximum and conduction band.* Determination of thermodynamic charge transition levels from formation energies (example of oxygen vacancy from Paper II). Plotting the formation energy for three different charge states of the defect, it is found that for low values of μ_e the state with charge $+2$ is most stable. At $\mu_\mathrm{e} - E_\mathrm{VBM} \approx 2.5$ a transition occurs to the state with charge $+1$ having the lowest energy, defining a thermodynamic charge transition level as depicted in the right panel. Increasing μ_e further, another level is found for $\mu_\mathrm{e} - E_\mathrm{VBM}$ just over 3.

7.4 Thermodynamic charge transition levels

With the formalism for calculating formation energies in place it is straight-forward to calculate the thermodynamic charge transition levels induced in the band gap when introducing point defects in a semiconductor or insulator. At thermodynamic equilibrium, the total concentration of a particular defect will be distributed over its different charge states. As illustrated in figure 7.4 the value $\mu_\mathrm{e}^\mathrm{transition}$ of the electronic chemical potential where the defect changes from being thermodynamically most stable in charge state q, to charge state q', is found by comparing the formation energy of the different charge states. Below $\mu_\mathrm{e}^\mathrm{transition}$ the concentration of defects in charge state q will dominate and vice versa. It should be noticed that these levels will in general not coincide with experimentally determined optical transition levels. The ionic configuration in equilibrium may vary significantly between the states below and above a transition level, in which case these transition levels will differ from optical ones. Optical transition levels may be calculated by using the same defect geometry for the final charge state as for the initial. Thermodynamic transition levels can on the other hand be observed in experiments where the final charge state can fully relax to its equilibrium configuration, such as in deep-level transient spectroscopy (DLTS) [76–78].

7.5 Temperature, pressure and zero-point energy

Having so far formally restricted the treatment to zero temperature and pressure, it is now extended to finite temperature and pressure. Effects of zero-point motion are conveniently included in the discussion as well.

In the context of point defects in semiconductors, temperature and pressure effects related to solid phases are commonly neglected. Usually the contributions are small, and especially for solid reservoirs the similarities of crystal and reservoir lead to a great deal of cancellation. Gas-phase reservoirs on the other hand can contribute significantly to the free energy of formation and are more frequently taken into account. As a first estimation of finite temperature- and pressure effects it is thus reasonable to only take changes related to atomic exchange with gas phases into account.

Taking equilibrium with an oxygen-rich atmosphere as an example, the chemical potential of oxygen is defined as half Gibbs free energy of one O_2(g) molecule; $\mu_O(T,p) = 1/2 g_{O_2}(T,p)$. If the harmonic zero-point vibrational energy of the oxygen molecule, $\hbar\omega_{O_2}/2$, is explicitly included, and ideal-gas behavior is assumed, the chemical potential of oxygen is:

$$\mu_O(T,p) = \frac{1}{2}\left(E^{\text{tot}}[O_2] + \frac{\hbar\omega_{O_2}}{2} + h_{O_2}(T,p^\circ) - T s_{O_2}(T,p^\circ) + k_B T \ln\frac{p}{p^\circ}\right), \quad (7.8)$$

where p° is some reference pressure and h_{O_2}, s_{O_2} is respectively the enthalpy and entropy of one oxygen molecule. The energy scale is here chosen so that $h_{O_2}(0,p^\circ) = 0$. The harmonic zero-point energy is known for the oxygen molecule (about 0.1 eV) and values of enthalpy and entropy are readily available in thermodynamical tables [19].

While the approach taken in the present thesis is to neglect temperature and pressure effects of solid phases, a brief outline is nevertheless in order. It is first noted that a pressure of 1 atm translates to 10^{-3} meV/Å^3, which in comparison with formation energies on the order of eV makes contributions from the pV-term of Gibbs free energy safe to neglect. If the vibrations in a solid are modeled by approximating the N-atom lattice with $3N$ harmonic oscillators, each with a unique frequency ω_s, the explicit expression for the vibrational free energy $F^{\text{vib}}(T)$ of a crystal is [79]:

$$F^{\text{vib}}(T) = \sum_{s=1}^{3N}\left\{\frac{\hbar\omega_s}{2} + k_B T \ln\left[1 - \exp(-\hbar\omega_s/k_B T)\right]\right\}.$$

Notice that the zero-temperature vibration enters explicitly in the first term and that the classical result $3Nk_BT$ is recovered in the high-temperature limit $k_BT >> \hbar\omega$. The harmonic frequency modes ω_s can be calculated by diagonalizing the force-constant matrix corresponding to forces on the ions in the lattice. Using finite differences these forces can be calculated using DFT by applying small, subsequent perturbations of the ion coordinates. Another approach to calculating harmonic phonon frequencies is to combine DFT with linear response theory [80]. In the formation of

a point defect, phonon modes can disappear (vacancies) or appear (*eg.* interstitials), and the frequencies of existing modes can change due to changes in interatomic potentials. To a first approximation, a change of $3k_{\mathrm{B}}T$ can thus be expected during formation of a point defect in a solid crystal, and a practical improvement can be obtained by studying changes in the phonon spectrum related to the ion site where the defect is introduced, and to the nearest neighbors.

Chapter 8

Discussion of method

Methods based on density-functional theory currently represent the most popular way to perform theoretical research into the electronic and atomistic properties of crystalline materials. Despite this success and continued advances in computing power, the techniques are not flawless in terms of accuracy and performance. In relation to the calculation of formation energies of point defects, there are particularly two things which makes a combination of high accuracy and feasible performance less than straightforward to achieve. The first of the issues is an artifact of modeling non-periodic features in implementations of DFT employing periodic boundary conditions, and most notably concerns the formation energy of charged point defects. This item is attended to in section 8.1. The second subject is the inability of calculations utilizing standard LDA and GGA exchange-correlation functionals to accurately reproduce the band gap in insulators and semiconductors. An error in the band gap translates to an uncertainty in the relative formation energies of different defect charge states. Section 8.2 outlines the problem and routes to improved band gaps, including the use of hybrid exact exchange functionals which is employed in Paper V.

8.1 Supercell calculations: finite-size effects

The first term on the right hand side of equation 7.1 refers to the total energy of a piece of material containing one sample of the relevant point defect. This one point defect breaks the periodic symmetry of the host crystal. Modeling non-periodic structures in the commonly employed periodic boundary supercell implementations of DFT is inherently intricate due to the periodic repetition of supercells. This issue is always present when modeling systems not exhibiting the symmetry given by the supercell periodicity: atoms and molecules, surfaces and interfaces, line defects and point defects. If for instance one point defect is placed in a supercell, the result is a periodic array of point defects and not the single, isolated defect intended, compare figures 8.1 and 8.2. Interaction between periodic defect images gives rise to spurious contributions to the calculated energy. The generic solution to minimizing the energy error is to separate the non-periodic features by making the supercell large enough

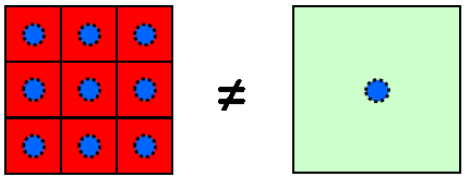


Figure 8.1: Electrostatic interaction between the periodic supercell images of charged point defects. Erroneous energy contributions arise from the defect images interacting with each other and with the neutralizing background.

so that they do not interact. The effect of increasing the size of the supercell can be checked by studying the convergence of properties of interest, *eg.* formation energies, with respect to supercell size. Furthermore, for quantities and conclusions based on comparison of the formation energy of similar point defects, much of the errors due to finite supercell sizes cancel. While the majority of the work in the present thesis is founded on such convergence tests and comparisons of formation energies, a brief discussion on these finite-size effects is nevertheless appropriate. With focus on potentially charged point defects, the spurious interactions may be divided into two categories: electrostatic and elastic.

Electrostatic interaction

For charged point defects, an erroneous electrostatic contribution to the formation energy comes about from the artificial array of charges and the homogenous background charge needed to make the supercell neutral, *cf.* figure 8.1. Several different approaches to improve supercell convergence by suppressing or correcting for the electrostatic error have been proposed and discussed [81–92], but it is one of the earliest and most straight-forward ones that has gained the most momentum and more or less become standard. In 1985 Leslie and Gillan pointed out that the coulombic error can be estimated by the Madelung energy of an array of point charges in a neutralizing background and a screening medium [81]. This energy is given by

$$E_{\mathrm{LG}} = -\frac{\alpha q^2}{2\varepsilon L},$$

where q denotes the defect charge, α the Madelung constant which depends on lattice type, and ε the macroscopic dielectric constant of the screening medium (the material hosting the defect). However, despite the descriptive name, the charge of point defects is on a microscopic level rarely point-like but generally has a finite extension and carries higher order electrostatic moments. In 1995 Makov and Payne proved that the Madelung energy caused by the monopole moment of the charges should be appended with a quadrupole term scaling as qQ/L^3 and that the total coulombic error can be written [82]:

$$E_{\mathrm{MP}} = -\frac{\alpha q^2}{2\varepsilon L} - \frac{2\pi qQ}{3\varepsilon L^3} + O(L^{-5}), \tag{8.1}$$

where Q denotes the quadropole moment of the charge. The derivation of this expression was based on ions or charged molecules, *ie.* on charged defects hosted by vacuum and not a condensed matter material such as a solid oxide. The screening effects of a host other than vacuum were included in a phenomenological manner and described by the dielectric constant. The dielectric constant is however a macroscopic quantity which not necessarily describes screening effects accurately on a microscopic level. Another implicit assumption is that the charge of the defect does not extend beyond the boundaries imposed by the supercell. Arguably, this may not necessarily be true considering the large bohr radius predicted by the hydrogen/effective mass model. The Makov-Payne correction can thus be expected to produce excellent results for ions and charged molecules but is uncertain for defects in condensed systems. Indeed, examples of failures of the Makov-Payne correction to improve supercell convergence for defects in condensed systems can be found in references 83–85.

Several authors have suggested a different approach to removing the spurious interaction, which involves truncating or compensating the long-range tail of the Coulomb potential induced by the point defect [86, 89, 93]. However, these methods do not take into account polarization of the material outside the supercell. This leads to an error estimated to be larger than in the standard approach for materials with $\varepsilon > 2.8$, which includes most solids of interest [91]. In a recent publication [91] a new, fully *ab initio*, approach to the coulombic interaction was presented which might hold some promise, but evaluation and tests of the method are still few.

For an estimation of the errors due to electrostatic interaction involved in the present work, the leading monopole term of the Makov-Payne/Leslie-Gillan correction was calculated to 0.08 eV for the formation energy of a defect in charge state ± 2 (*eg.* a fully ionized oxygen vacancy) in the 135-atom BZO supercells used in Paper V.

Elastic interaction

Elastic interaction arises from the fact that ions surrounding the defect site relax to a different stable configuration when the defect forms. As illustrated in figure 8.2, finite supercells prevents the relaxation to take place the way it would for an isolated defect. To avoid errors caused by elastic interaction, supercells therefore need to have a size that essentially encloses the elastic strain field caused by the defect, *ie.* so that the ions at the edges of each supercell do not experience any significant strain caused by the defect.

Models of point defects in isotropic elastic media [94] show that the energy caused by spurious elastic interaction falls off as $1/R^3$, where R is the distance from the defect. In supercell calculations one has a choice in boundary conditions, keeping the volume of the supercell constant or keeping the pressure in the system constant equal to zero. According to the models, the elastic energy is either overestimated or underestimated depending on these boundary conditions. If the volume

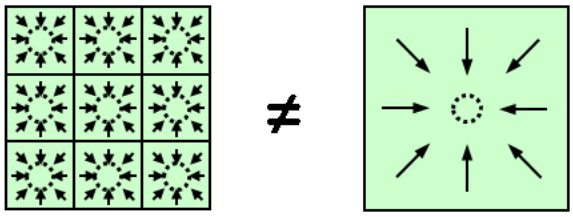


Figure 8.2: Elastic interaction between the periodic supercell images of point defects. If the supercell is small, the ions surrounding the defect will not relax the way they would if the defect was isolated.

is kept constant, the energy is overestimated and given by

$$E_{\mathrm{el},V} = E_{\mathrm{el}} \left[1 + \frac{3(1-\nu)}{2(1-2\nu)} \frac{r_c^3}{R^3} \right], \tag{8.2}$$

where E_{el} is the correct elastic energy, ν the Poisson ratio of the elastic medium and r_c the radius of the defect core. For zero-pressure calculations the energy is instead underestimated and the corresponding expression is

$$E_{\mathrm{el},p} = E_{\mathrm{el}} \left[1 - 3\frac{(1-\nu)}{(1+\nu)} \frac{r_c^3}{R^3} \right]. \tag{8.3}$$

The conclusions are that the error due to elastic interaction scales the same way as the second term of the Makov-Payne estimation of the electrostatic error in both constant-volume and zero-pressure calculations. Since the scaling is known, contributions from elastic interaction and multipole electrostatic interaction can in principle be suppressed by extrapolating formation energies obtained with different supercell sizes. In the present thesis the approach is to use constant-volume calculations, since the energetic overestimation of the elastic contribution leads to some cancellation of errors due to electrostatic interaction.

8.2 The band gap problem

As a rule, DFT based on LDA or GGA severely underestimates the band gap of semiconductors and insulators [95–99]. In barium zirconate, for instance, the band gap is calculated to 3.1 eV with GGA, in contrast to the experimental value 5.3 eV. In addition to the explicit misdescription of wide-bandgap ionic conductors as mixed or electronic conductors, the problem has direct bearings to the relative formation energy and concentration of defects in different charge states.

There are two aspects of the band gap problem. First, exchange-correlation approximations based on local electron densities (both LDA and GGA) suffer from an error in energy related to self-interaction of electrons. The error is most severe in densities characterized by strongly localized electron orbitals, primarily f- but also d-states.

The second aspect of the band gap problem is related to an inability of standard LDA- and GGA-based DFT to produce accurate total energies for systems where an electron has been placed in an otherwise unoccupied level (*cf.* conduction band), if that level is separated from the occupied ones (*cf.* valence band) by a gap in energy [95–97, 100, 101]. At the heart of the problem lies a discontinuity Δ_{xc} in the xc potential V_{xc} between that seen by the added electron and that seen by the others. This xc discontinuity *should* be present but vanishes in LDA and GGA. The procedure of adding an electron to otherwise unoccupied states to calculate total energies is characteristic not only for calculating band gaps but also the formation energies of point defects with occupied defect-associated electron states. The discontinuity grows with the size of the energy gap, on which account primarily the formation energy of neutral or not fully ionized states of donors with transitions levels close to the conduction band are prone to error [66].

Two things can be pointed out in particular.[1] First, even when adding (or removing) electrons to a supercell the aim is to calculate ground states. The issue with adding electrons is thus not explicitly related to excited states versus DFT being a ground-state theory[2]. Second, on a primary level the band gap discussed here should not be confused with the one seen in the band structure or density of states derived from the Kohn-Sham eigenvalues. The two however happen to coincide due to the vanishing xc discontinuity. The fundamental band gap is defined as the difference in ionization energy E_{I} and electron affinity E_{A},

$$E_{\mathrm{gap}} = E_{\mathrm{A}} - E_{\mathrm{I}}, \tag{8.4}$$

or in other words the difference between the energy related to adding an electron to the material and the energy related to removing one. The difference between the Kohn-Sham band gap $E_{\mathrm{gap}}^{\mathrm{KS}}$ and the fundamental gap is in fact the xc discontinuity:

$$E_{\mathrm{gap}} = E_{\mathrm{gap}}^{\mathrm{KS}} + \Delta_{\mathrm{xc}}.$$

In practical DFT calculations the band gap given by equation 8.4 is calculated from total energies with electrons properly added and removed:

$$E_{\mathrm{gap}} = \underbrace{\left[E^{\mathrm{tot}}(+1) - E^{\mathrm{tot}}(0)\right]}_{E_{\mathrm{CBM}}} - \underbrace{\left[E^{\mathrm{tot}}(0) - E^{\mathrm{tot}}(-1)\right]}_{E_{\mathrm{VBM}}}, \tag{8.5}$$

where $E^{\mathrm{tot}}(i)$ is used to denote the total energy with i electrons added to the system. Indicated are also the conduction band minimum E_{CBM} and valence band maximum E_{VBM} corresponding to the band gap.

[1] For details, see *eg.* references 95–97, 100, 102.

[2] Compare: Just like an atom, an ion also has a quantum-mechanical ground state.

Methods and approaches to best alleviate errors related to the band gap problem are areas of ongoing research. Nevertheless, a few techniques can be highlighted and summarized [100, 102].

- *The GW approximation* [103]. A many-body approach based on Green's-function techniques. Provides accurate band gaps but has currently no feasible way of obtaining total energies for the size of supercells needed for realistic point defect investigations.
- *LDA/GGA+U* [104]. An explicit, repulsive, coulomb energy (U) is added to localized orbitals in the valence band. This can counteract the self-interaction error and partly improve the band gap.
- *Hartree-Fock* [105]. Separate method, not related to DFT. Many-body treatment including exact exchange evaluation (Fock exchange). Yields too large band gaps.
- *Exact exchange* (EXX) [106]. Exact exchange is calculated from the Kohn-Sham orbitals instead of using the Hartree-Fock method.
- *Hybrid functionals* [107, 108]. Motivated by the overestimation of band gaps by the Hartree-Fock method and the underestimation by LDA/GGA, an interpolating amount of Fock exchange is mixed in with normal LDA or GGA. High computational costs, primarily due to the long range of the Fock exchange. By only treating short-range exchange exactly, performance can sometimes be improved while retaining improved band gaps [109, 110].

Application of hybrid exchange-correlation functionals is becoming increasingly popular as a means to obtaining accurate band gaps. Compared to standard GGA, such calculations are however typically 50–100 times more demanding in terms of computing time, which is close to unfeasible in combination with the large supercells required to manage the finite-size errors in point defect calculations. In Paper V an approach suggested by Alkauskas and Pasquarello [111] to overcoming this dilemma was taken. With basis in the often close overlap of wavefunctions derived from GGA and corresponding hybrid functionals, the wavefunctions of converged and optimized GGA (PBE [112]) calculations were used as input to "oneshot" hybrid (PBE0 [113]) computations, in which only one iteration of the electronic problem were performed. Excellent agreement between formation energies resulting from this approach and fully self-consistent hybrid calculations was obtained. Furthermore the calculated band gap of 5.25 eV was in gratifying agreement with the experimental value 5.3 eV.

Chapter 9

Summary of the papers

Paper I

The stable proton sites in LZO are pinpointed and an uninterrupted migration pathway connecting these sites is identified. Furthermore the proton trapping energies of four common dopant species are evaluated. The two most energetically favorable proton positions are found adjacent to oxygen in the octahedra enclosing Zr (*cf.* figure 3.4(a)) with occupation of the other oxygen site (O') unlikely on energetic grounds. Based on these two positions a pathway is identified, consisting of alternating jumps between equivalent positions, and transitions between inequivalent positions by rotation of the proton around its host oxide ion. According to the calculations the migration barrier amounts to 0.3 eV while trapping energies as large as 1 eV are seen. In comparison with experimental activation energies of 0.7 eV in Ca-doped samples, the energies are in reasonable agreement although slightly underestimated. Out of the investigated dopants Ba, Mg, Sr, Ca, the latter two show the most modest trapping energies.

Paper II

Effects of twelve different species of di- and trivalent acceptor dopants on the charge-compensating concentration and trapping of oxygen vacancies in LZO is evaluated. It is found that substitution of La or Zr for aliovalent dopants results in defects with nominal charge. As a consequence the determining factor for the charge state of a given dopant is whether it replaces La^{3+} or Zr^{4+}. The results show a clear correlation between replacement preference and dopant size, with dopants having small ionic radii occupying the smaller Zr site and larger dopants energetically preferring the larger La site. This behavior is explained by minimization of lattice strain and is in line with conventional wisdom that size-matching between substitutional dopants and the host cation intended for replacement is an important factor.

It turns out that elastic effects are significant also in the formation of oxygen vacancies and in the pair interaction between dopants and vacancies. First, oxygen vacancies are seen to form only on one of the two inequivalent oxygen sites in the

material, almost entirely due to greater lattice relaxation. The most stable vacancy site is the same as the one found energetically preferred by protons in Paper I. Secondly, while no decisive correlation between dopant charge and trapping energy is noticed, the pair interaction varies greatly with dopant size.

Taken together, the results of Papers I and II point to Ca and Sr being the most promising of the investigated dopant species by exhibiting a good combination of charge-compensating concentration enhancement and low trapping energy.

Paper III

Investigations are initiated into the segregation of oxygen vacancies to grain boundaries in BZO and its effects on proton conductivity due to space-charge depletion in the grain boundary region. Calculations show a segregation energy of -1.25 eV of oxygen vacancies to the core of a prototype grain boundary. On the basis of a simple space charge model it is demonstrated that significant space charge effects can be expected from vacancy segregation energies of this magnitude and results in electrostatic potential barriers comparable to estimates based on measured conductivities. Furthermore it is shown that the core vacancies are not expected to be extensively replaced by hydroxide ions in hydrated samples, but persist well below the hydration temperature of bulk oxygen vacancies. It is noted that the high stability of the oxygen vacancy position found at the interface is related to charge mismatch and lattice relaxation during vacancy formation.

Paper IV

DFT calculations of oxygen vacancy segregation to two grain boundaries are performed, and used as comparison for classical model-potential computations of segregation energies in a total of eight different tilt grain boundary structures. Segregation energies ranging between -0.5 and -2 eV are found. In this work an energetic grain boundary core is defined, based on the one-dimensional profile of the segregation energies calculated in the grain boundaries. A continuous space charge model with a numeric solution scheme is constructed, which shows that the calculated segregation energies give rise to electrostatic potential barriers between 0.2 and 0.8 V. By evaluation of the ratio between grain boundary and bulk proton conductivity, it is demonstrated that the potential barriers cause grain boundary blocking effects which essentially span the range seen in a collection of experimental data. The conclusion is that the grain boundary character of the material is determinant for the severity of grain boundary blocking effects caused by oxygen vacancies.

Similar to the observation made in Paper III, correlations are noticed between segregation energy and areas of lattice strain and charge mismatch. Particularly in low-angle grain boundaries, regions of high vacancy stability are found near dislocation cores. Tendencies of increasing magnitude of the segregation energies with increasing grain boundary energy are also seen.

Paper V

A wider and more general use of the formalism developed in Paper IV for evaluating the space charge effects arising from calculated segregation energies, is demonstrated. The segregation energies of all relevant charge states of an oxygen vacancy in the prototype grain boundary from Paper III, are calculated. By further performing self-consistent calculations of bulk defect concentrations within the thermodynamic stability range of BZO, and use them as input to the space charge model, the equilibrium defect structure in bulk and grain boundaries of a material characterized by oxygen-deficient grain boundary cores is obtained. The bending of the valence and conduction bands in the grain boundary due to the electrostatic potential, as well as the associated change in concentrations of conduction electrons and valence holes, are taken into account. A perturbative application of a hybrid exchange-correlation functional is used for efficient calculations with a proper description of the band gap.

Unlike DFT results produced using standard GGA functionals, use of the hybrid exchange-correlation functional shows that the oxygen vacancy in bulk BZO is a relatively deep donor defect, with the transition between the $+2/+1$ charge state 1.4 eV below the conduction band. Furthermore, it is seen that the most popular dopant species in BZO, Y, incorporates in an ideal manner as an acceptor on the Zr site under Ba-rich conditions, but can partly self-compensate by forming unintended donor defects on the Ba site in Zr-rich conditions. Finally, with the possible influence on grain boundary conductivity in mind, the Y segregation to the grain boundary due to the positive electrostatic potential is studied as function of dopant concentration, thermochemical conditions and sintering temperature. It is demonstrated that the space charge effects caused by oxygen vacancy segregation is a strong driving force for Y segregation to the grain boundary during sintering. Low temperatures and Ba-rich conditions amplify the segregation.

Chapter 10

Conclusion and outlook

The microscopic mechanisms which govern the ability of solid oxide materials to conduct protons is a fascinating subject. We have seen in this thesis that this ability can be directly enhanced by adding foreign atomic species (dopants) to the material, which increases the amount of protons but may in the same stroke cause entrapment of the very protons which were gained. Oxygen, which is inherent as a building block of the material actively chosen, is cunningly first removed in reaction to the added foreign atoms, and then put back again with protons following along. The oxygen however prefers not to return to the boundaries between grains inside the material, and oxygen vacancies instead blocks the passage of protons.

While it may seem as if nature stubbornly refuses to fully comply with our attempts to optimize the materials candidating for use as proton-conducting electrolytes, it is hopefully just an expression of limited understanding on our behalf. In the research presented in this thesis, theoretical and computational methods aimed at raising our level of understanding a bit further. In Papers I and II efficient dopant species with modest trapping energies in lanthanum zirconate were for instance identified, and in Papers III–V insight into the behavior and effects of oxygen vacancies in grain boundaries of barium zirconate was provided. Continuing this research by similarly investigating the energetic properties of protons in grain boundaries comes across as relevant. In Papers I and II it was found that protons in lanthanum zirconate are more stable at the oxygen sites preferred also by oxygen vacancies, and analogous tendencies have been seen in comparison between different perovskite oxides [114]. It is in other words not unlikely that protons, like oxygen vacancies, display energetic segregation preferences to grain boundaries in barium zirconate. High core concentrations of protons, greater space-charge potentials and wider depletion regions would be the results. Conductivity measurements support this conjecture, with indications of greater space-charge effects in hydrated samples than in dry samples [22].

In this context one may reflect on the possibility for qualitatively different behavior of grain boundary hydrogen compared to bulk. Hydrogen is for instance often amphoteric in semiconductors, with both stable donor- and acceptor states [71]. The possible existence and effects of acceptor states of hydrogen, in particular incorporated as negative hydride ions on vacant oxygen sites, in perovskite oxides have been

discussed before [115,116] and were recently revisited in a DFT investigation of barium titanate [117]. As was illustrated for the Y dopant in Paper V, acceptor states are stabilized compared to donor states in the presence of positive space charge potentials at grain boundaries. Further considering the elevated concentrations of oxygen vacancies we have predicted at the grain boundary cores, the possibility for anomalous effects in terms of concentration and conduction mechanisms, caused by negative hydrogen states, can at least be kept in mind. On a related note, it should not be forgotten that increased proton migration barriers have been shown in at least one grain boundary structure [118].

Equally important, or maybe even more so, as the behavior of protons in grain boundaries, seems to be the behavior of acceptor dopants. It was noted in Papers IV and V that attraction of negative acceptor dopants by the positive core oxygen vacancies does not cause significant neutralization of the grain boundary cores. The vacancy concentration increases in response, and the net electrostatic result is negligible. At variance with this behavior, a decrease of the blocking effect in grain boundaries as a result of increasing dopant concentration has been noted in measurements, and understanding this effect might be a route towards devising schemes to counteract the grain boundary problem.

Finally, the range of oxygen vacancy segregation energies we found for the different grain boundary structures in Paper IV, raised questions regarding the grain boundary character in polycrystalline samples. In this light it seems useful to look toward ways to produce samples with well-characterized interfaces, in order to enable more direct comparisons between measurements and models. The rapid progression in thin-film techniques will probably make this possible in a near future [119].

In conclusion, further understanding of the blocking effect in grain boundaries of BZO is a route worth pursuing, and grain boundaries offer interesting dimensions to the defect thermochemistry of proton conductors. In a wider perspective, the micro- or nanostructure of solid oxide materials represents a degree of freedom which might be taken advantage of to improve the proton conductivity. For instance, perhaps high defect concentrations at interfaces can, in the future, be turned from a problem to an asset, and lead to electrolytes with tailored nanoscale designs having conductivities of a completely different magnitude than today [120–123]. To make this possible, theoretical and computational methods will surely prove valuable complements to experiments. It is a good thing that the mind is free and computers get cheaper and cheaper!

Acknowledgements

With the completion of this thesis near at hand, I am grateful and in debt to the many people who made it possible. First and most of all i am thankful to my supervisor and examiner Professor Göran Wahnström for his patient direction and encouraging support throughout all parts of my work. His combination of skill and humility has been truly inspirational. The stimulating environment generated by my talented friends and, in some cases former, colleagues in the Materials and Surface Theory group has also been a great source of inspiration. Thank you all for providing fond memories of MST! Special appreciation is extended to Mårten Björketun for contributions to the present work and more importantly for kind guidance during my first months at MST, and to Dan Fors for patient Vasp support and rewarding discussions during most of the remaining months. Finally, Edit Helgee, Anders Lindman and Paul Erhart are explicitly acknowledged and especially thanked for their hard work on the grain boundary project. Without their support and dedication, the quality and content of this thesis would be far from what it is. I owe you, guys.

Many people from my pre-graduate period at Chalmers could be acknowledged for making me embark on the present undertaking in the first place. I am particularly happy that Calle Borgentun and Anders Ström stuck in there all the way and kept me company to, and through, graduate school. Thank you for being plain and simply awesome people, and for making especially the first years at Chalmers a pure joy! My MSc supervisor Martin Nilsson Jacobi should also be mentioned, for unknowingly making me consider a PhD degree at all.

A heap of gratitude is directed toward my family and the rest of my friends, for suitably distracting me from work on a regular basis. A special thought is sent to my father (mom, I think of you as well) but I will not mention the rest of you. You know who you are, and if not I will remind you more often.

Last and most. Rimma: Thanks, and props to you, for the remarkable feat of proofreading this thesis (yes, it is pronounced zir-co-nate)! More than that, thank you for always being there, in and throughout everything, for being you, and for being *my* everything.

Bibliography

[1] B. C. H. Steele, *Fuel-cell technology: Running on natural gas*, Nature **400**, 619 (1999).

[2] T. Norby, *The promise of protonics*, Nature **410**, 877 (2001).

[3] A. Boudghene Stambouli and E. Traversa, *Fuel cells, an alternative to standard sources of energy*, Renewable and Sustainable Energy Reviews **6**, 295 (2002).

[4] M. Momirlan and T. Veziroglu, *The properties of hydrogen as fuel tomorrow in sustainable energy system for a cleaner planet*, International Journal of Hydrogen Energy **30**, 795 (2005).

[5] M. Granovskii, I. Dincer, and M. A. Rosen, *Exergetic life cycle assessment of hydrogen production from renewables*, Journal of Power Sources **167**, 461 (2007).

[6] L. Hernandez and V. Kafarov, *Use of bioethanol for sustainable electrical energy production*, International Journal of Hydrogen Energy **34**, 7041 (2009).

[7] I. Dincer and M. A. Rosen, *Sustainability aspects of hydrogen and fuel cell systems*, Energy for Sustainable Development **15**, 137 (2011).

[8] E. Fabbri, D. Pergolesi, and E. Traversa, *Materials challenges toward proton-conducting oxide fuel cells: a critical review*, Chemical Society Reviews **39**, 4355 (2010).

[9] B. C. H. Steele and A. Heinzel, *Materials for fuel-cell technologies*, Nature **414**, 345 (2001).

[10] L. Carrette, K. A. Friedrich, and U. Stimming, *Fuel Cells: Principles, Types, Fuels, and Applications*, ChemPhysChem **1**, 162 (2000).

[11] I. EG&G Technical Services, *Fuel cell handbook*, 7 ed. (U.S. Department of Energy, Office of Fossil Energy, National Energy Technology Laboratory, Morgantown, West Virginia, 2004).

[12] S. M. Haile, *Materials for fuel cells*, Materials Today **6**, 24 (2003).

[13] A. J. Jacobson, *Materials for Solid Oxide Fuel Cells*, Chemistry of Materials **22**, 660 (2009).

[14] F. Lefebvre-Joud, G. Gauthier, and J. Mougin, *Current status of proton-conducting solid oxide fuel cells development*, Journal of Applied Electrochemistry **39**, 535 (2009).

[15] A. Kirubakaran, S. Jain, and R. Nema, *A review on fuel cell technologies and power electronic interface*, Renewable and Sustainable Energy Reviews **13**, 2430 (2009).

[16] A. Orera and P. R. Slater, *New Chemical Systems for Solid Oxide Fuel Cells*, Chemistry of Materials **22**, 675 (2009).

[17] C. Xia and M. Liu, *A Simple and Cost-Effective Approach to Fabrication of Dense Ceramic Membranes on Porous Substrates*, Journal of the American Ceramic Society **84**, 1903 (2001).

[18] S. M. Haile, *Fuel cell materials and components*, Acta Materialia **51**, 5981 (2003).

[19] M. W. Chase, *JANAF Thermochemical Tables*, 3 ed. (American Chemical Society and the American Institute of Physics, New York, 1986).

[20] L. Malavasi, C. A. J. Fisher, and M. S. Islam, *Oxide-ion and proton conducting electrolyte materials for clean energy applications: structural and mechanistic features*, Chem. Soc. Rev. **39**, 4370 (2010).

[21] H. Iwahara, *Technological challenges in the application of proton conducting ceramics*, Solid State Ionics **77**, 289 (1995).

[22] C. Kjolseth, H. Fjeld, O. Prytz, P. I. Dahl, C. Estournes, R. Haugsrud, and T. Norby, *Space-charge theory applied to the grain boundary impedance of proton conducting BaZr0.9Y0.1O3-delta*, Solid State Ionics **181**, 268 (2010).

[23] S. Duval, P. Holtappels, U. Vogt, E. Pomjakushina, K. Conder, U. Stimming, and T. Graule, *Electrical conductivity of the proton conductor BaZr0.9Y0.1O3-delta obtained by high temperature annealing*, Solid State Ionics **178**, 1437 (2007).

[24] P. Babilo, T. Uda, and S. M. Haile, *Processing of Yttrium-Doped Barium Zirconate for High Proton Conductivity*, Journal of Materials Research **22**, 1322 (2007).

[25] H. Iwahara, T. Esaka, H. Uchida, and N. Maeda, *Proton conduction in sintered oxides and its application to steam electrolysis for hydrogen production*, Solid State Ionics **3-4**, 359 (1981).

[26] H. Iwahara, H. Uchida, and S. Tanaka, *High temperature type proton conductor based on SrCeO3 and its application to solid electrolyte fuel cells*, Solid State Ionics **9-10, Part 2**, 1021 (1983).

[27] H. Iwahara, *Proton Conduction in Sintered Oxides Based on BaCeO3*, Journal of The Electrochemical Society **135**, 529 (1988).

[28] H. Iwahara, T. Yajima, T. Hibino, K. Ozaki, and H. Suzuki, *Protonic conduction in calcium, strontium and barium zirconates*, Solid State Ionics **61**, 65 (1993).

[29] P. I. Dahl, H. L. Lein, Y. Yu, J. Tolchard, T. Grande, M. Einarsrud, C. Kjolseth, T. Norby, and R. Haugsrud, *Microstructural characterization and electrical properties of spray pyrolyzed conventionally sintered or hot-pressed BaZrO3 and BaZr0.9Y0.1O3-delta*, Solid State Ionics **182**, 32 (2011).

[30] J. Tolchard and T. Grande, *Chemical compatibility of candidate oxide cathodes for BaZrO3 electrolytes*, Solid State Ionics **178**, 593 (2007).

[31] K. D. Kreuer, *Proton-conducting oxides*, Annual Review of Materials Research **33**, 333 (2003).

[32] T. Norby and Y. Larring, *Concentration and transport of protons in oxides*, Current Opinion in Solid State and Materials Science **2**, 593 (1997).

[33] X. Guo and R. Waser, *Space charge concept for acceptor-doped zirconia and ceria and experimental evidences*, Solid State Ionics **173**, 63 (2004).

[34] F. Iguchi, N. Sata, T. Tsurui, and H. Yugami, *Microstructures and grain boundary conductivity of BaZr1-xYxO3 (x = 0.05, 0.10, 0.15) ceramics*, Solid State Ionics **178**, 691 (2007).

[35] M. Shirpour, R. Merkle, and J. Maier, *Evidence for space charge effects in Y-doped BaZrO3 from reduction experiments*, Solid State Ionics .

[36] F. Iguchi, C. Chen, H. Yugami, and S. Kim, *Direct evidence of potential barriers at grain boundaries in Y-doped BaZrO3 from dc-bias dependence measurements*, J. Mater. Chem. **21**, 16517 (2011).

[37] T. Shimura, M. Komori, and H. Iwahara, *Ionic conduction in pyrochlore-type oxides containing rare earth elements at high temperature*, Solid State Ionics **86-88, Part 1**, 685 (1996).

[38] A. Ota, Y. Matsumura, M. Yoshinaka, K. Hirota, and O. Yamaguchi, *Formation and sintering of 8 mol% Y2O3-substituted La2Zr2O7 by the hydrazine method*, Journal of materials science letters **17**, 199201 (1998).

[39] A. Chen, J. R. Smith, K. L. Duncan, R. T. DeHoff, K. S. Jones, and E. D. Wachsman, *Effect of La2Zr2O7 on Interfacial Resistance in Solid Oxide Fuel Cells*, Journal of The Electrochemical Society **157**, B1624 (2010).

[40] S. M. Haile, D. L. West, and J. Campbell, *The role of microstructure and processing on the proton conducting properties of gadolinium-doped barium cerate*, Journal of Materials Research **13**, 1576 (1998).

[41] J. A. Kilner, *Fast oxygen transport in acceptor doped oxides*, Solid State Ionics **129**, 13 (2000).

[42] K. D. Kreuer, in *Perovskite Oxide for Solid Oxide Fuel Cells*, *Fuel Cells and Hydrogen Energy*, edited by T. Ishihara (Springer US, Boston, MA, 2009), pp. 261–272.

[43] T. Norby, in *Perovskite Oxide for Solid Oxide Fuel Cells*, edited by T. Ishihara (Springer US, Boston, MA, 2009), pp. 217–241.

[44] W. Muench, G. Seifert, K. D. Kreuer, and J. Maier, *A quantum molecular dynamics study of proton conduction phenomena in BaCeO3*, Solid State Ionics **86-88, Part 1**, 647 (1996).

[45] W. Muench, K. D. Kreuer, G. Seifertli, and J. Maier, *A quantum molecular dynamics study of proton diffusion in SrTiO3 and CaTiO3*, Solid State Ionics **125**, 39 (1999).

[46] J. A. Kilner, A. Berenov, and J. Rossiny, in *Perovskite Oxide for Solid Oxide Fuel Cells*, edited by T. Ishihara (Springer US, Boston, MA, 2009), pp. 95–116.

[47] N. F. Mott and R. W. Gurney, *Electronic processes in ionic crystals* (Clarendon, Oxford, 1940).

[48] W. Kohn, *Solid State Physics V* (Academic Press Inc., New York, 1957), p. 252.

[49] J. Maier, *Physical Chemistry of Ionic Materials* (John Wiley & Sons, Chichester, West Sussex, 2004).

[50] I. Denk, J. Claus, and J. Maier, *Electrochemical Investigations of SrTiO3 Boundaries*, Journal of The Electrochemical Society **144**, 3526 (1997).

[51] X. Guo and R. Waser, *Electrical properties of the grain boundaries of oxygen ion conductors: Acceptor-doped zirconia and ceria*, Progress in Materials Science **51**, 151 (2006).

[52] R. O. Jones and O. Gunnarsson, *The density functional formalism, its applications and prospects*, Reviews of Modern Physics **61**, 689 (1989).

[53] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, *Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients*, Reviews of Modern Physics **64**, 1045 (1992).

[54] W. Kohn, *Nobel Lecture: Electronic structure of matterwave functions and density functionals*, Reviews of Modern Physics **71**, 1253 (1999).

[55] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).

[56] J. Kohanoff, *Electronic Structure Calculations for Solids and Molecules: Theory and Computational Methods* (Cambridge University Press, Cambridge, 2006).

[57] M. Born and R. Oppenheimer, *Zur Quantentheorie der Molekeln*, Annalen der Physik **389**, 457 (1927).

[58] P. Hohenberg and W. Kohn, *Inhomogeneous Electron Gas*, Physical Review **136**, B864 (1964).

[59] W. Kohn and L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Physical Review **140**, A1133 (1965).

[60] D. M. Ceperley and B. J. Alder, *Ground State of the Electron Gas by a Stochastic Method*, Physical Review Letters **45**, 566 (1980).

[61] F. Bloch, *Ueber die Quantenmechanik der Elektronen in Kristallgittern*, Zeitschrift fuer Physik **52**, 555 (1929).

[62] R. P. Feynman, *Forces in Molecules*, Physical Review **56**, 340 (1939).

[63] S. B. Zhang and J. E. Northrup, *Chemical potential dependence of defect formation energies in GaAs: Application to Ga self-diffusion*, Physical Review Letters **67**, 2339 (1991).

[64] D. B. Laks, C. G. Van de Walle, G. F. Neumark, P. E. Bloechl, and S. T. Pantelides, *Native defects and self-compensation in ZnSe*, Physical Review B **45**, 10965 (1992).

[65] C. G. Van de Walle, D. B. Laks, G. F. Neumark, and S. T. Pantelides, *First-principles calculations of solubilities and doping limits: Li, Na, and N in ZnSe*, Physical Review B **47**, 9425 (1993).

[66] A. F. Kohan, G. Ceder, D. Morgan, and C. G. Van de Walle, *First-principles study of native point defects in ZnO*, Physical Review B **61**, 15019 (2000).

[67] C. Stampfl, M. Veronica Ganduglia-Pirovano, K. Reuter, and M. Scheffler, *Catalysis and corrosion: the theoretical surface-science context*, Surface Science **500**, 368 (2002).

[68] K. Reuter and M. Scheffler, *Composition, structure, and stability of RuO_2(110) as a function of oxygen pressure*, Physical Review B **65**, 035406 (2001).

[69] C. G. Van de Walle and J. Neugebauer, *First-principles calculations for defects and impurities: Applications to III-nitrides*, Journal of Applied Physics **95**, 3851 (2004).

[70] A. Janotti and C. G. Van de Walle, *Hydrogen multicentre bonds*, Nat Mater **6**, 44 (2007).

[71] C. G. Van de Walle and J. Neugebauer, *Universal alignment of hydrogen levels in semiconductors, insulators and solutions*, Nature **423**, 626 (2003).

[72] P. Erhart and K. Albe, *Thermodynamics of mono- and di-vacancies in barium titanate*, Journal of Applied Physics **102**, 084111 (2007).

[73] P. Erhart and K. Albe, *Modeling the electrical conductivity in BaTiO3 on the basis of first-principles calculations*, Journal of Applied Physics **104**, 044315 (2008).

[74] G. Qian, R. M. Martin, and D. J. Chadi, *First-principles study of the atomic reconstructions and energies of Ga- and As-stabilized GaAs(100) surfaces*, Physical Review B **38**, 7649 (1988).

[75] D. Neamen, *Semiconductor Physics And Devices: Basic Principles*, 3 ed. (McGraw-Hill, New York, 2003).

[76] M. Lannoo and J. Bourgoin, *Semiconductors I: Theoretical Aspects* (Springer, Berlin, 1981).

[77] M. Lannoo and J. Bourgoin, *Semiconductors II: Experimental Aspects* (Springer, Berlin, 1983).

[78] P. M. Mooney, in *Identification of Defects in Semiconductors*, edited by M. Stavola (Academic Press, San Diego, 1999), Vol. 51B.

[79] L. E. Reichl, *Statistical Physics*, 2 ed. (Wiley, New York, 1998).

[80] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, *Phonons and related crystal properties from density-functional perturbation theory*, Reviews of Modern Physics **73**, 515 (2001).

[81] M. Leslie and N. J. Gillan, *The energy and elastic dipole tensor of defects in ionic crystals calculated by the supercell method*, Journal of Physics C: Solid State Physics **18**, 973 (1985).

[82] G. Makov and M. C. Payne, *Periodic boundary conditions in ab initio calculations*, Physical Review B **51**, 4014 (1995).

[83] J. Shim, E. Lee, Y. J. Lee, and R. M. Nieminen, *Density-functional calculations of defect formation energies using supercell methods: Defects in diamond*, Physical Review B **71**, 035206 (2005).

[84] A. F. Wright and N. A. Modine, *Comparison of two methods for circumventing the Coulomb divergence in supercell calculations for charged point defects*, Physical Review B **74**, 235209 (2006).

[85] C. W. M. Castleton, A. Hoeglund, and S. Mirbt, *Managing the supercell approximation for charged defects in semiconductors: Finite-size scaling, charge correction factors, the band-gap problem, and the ab initio dielectric constant*, Physical Review B **73**, 035215 (2006).

[86] P. Carloni, P. E. Bloechl, and M. Parrinello, *Electronic structure of the Cu, Zn superoxide dismutase active site and its interactions with the substrate*, The Journal of Physical Chemistry **99**, 1338 (1995).

[87] D. Segev and S. Wei, *Design of Shallow Donor Levels in Diamond by Isovalent-Donor Coupling*, Physical Review Letters **91**, 126406 (2003).

[88] P. A. Schultz, *Local electrostatic moments and periodic boundary conditions*, Physical Review B **60**, 1551 (1999).

[89] P. A. Schultz, *Charged Local Defects in Extended Systems*, Physical Review Letters **84**, 1942 (2000).

[90] L. N. Kantorovich, *Elimination of the long-range dipole interaction in calculations with periodic boundary conditions*, Physical Review B **60**, 15476 (1999).

[91] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Fully Ab Initio Finite-Size Corrections for Charged-Defect Supercell Calculations*, Physical Review Letters **102**, 016402 (2009).

[92] I. Dabo, B. Kozinsky, N. E. Singh-Miller, and N. Marzari, *Electrostatics in periodic boundary conditions and real-space corrections*, Physical Review B **77**, 115139 (2008).

[93] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, *Exact Coulomb cutoff technique for supercell calculations*, Physical Review B **73**, 205119 (2006).

[94] Y. Mishin, M. R. Sorensen, and A. F. Voter, *Calculation of point-defect entropy in metals*, Philosophical Magazine A **81**, 2591 (2001).

[95] J. P. Perdew and M. Levy, *Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities*, Physical Review Letters **51**, 1884 (1983).

[96] L. J. Sham and M. Schlueter, *Density-Functional Theory of the Energy Gap*, Physical Review Letters **51**, 1888 (1983).

[97] L. J. Sham and M. Schlueter, *Density-functional theory of the band gap*, Physical Review B **32**, 3883 (1985).

[98] G. Ortiz, *Gradient-corrected pseudopotential calculations in semiconductors*, Physical Review B **45**, 11328 (1992).

[99] C. Filippi, D. J. Singh, and C. J. Umrigar, *All-electron local-density and generalized-gradient calculations of the structural properties of semiconductors*, Physical Review B **50**, 14947 (1994).

[100] C. Van de Walle, *Computational studies of conductivity in wide-band-gap semiconductors and oxides*, Journal of Physics: Condensed Matter **20**, 064230 (2008).

[101] R. W. Godby, M. Schlueter, and L. J. Sham, *Self-energy operators and exchange-correlation potentials in semiconductors*, Physical Review B **37**, 10159 (1988).

[102] R. M. Nieminen, in *Theory of Defects in Semiconductors*, Vol. 104 of *Topics in Applied Physics*, edited by D. A. Drabold and S. K. Estreicher (Springer, Berlin, Heidelberg, 2007), pp. 29–68.

[103] M. S. Hybertsen and S. G. Louie, *Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies*, Physical Review B **34**, 5390 (1986).

[104] V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein, *First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA + U method*, Journal of Physics: Condensed Matter **9**, 767 (1997).

[105] V. Fock, *Naeherungsmethode zur Loesung des quantenmechanischen Mehrkoerperproblems*, Zeitschrift fuer Physik **61**, 126 (1930).

[106] M. Staedele, J. A. Majewski, P. Vogl, and A. Goerling, *Exact Kohn-Sham Exchange Potential in Semiconductors*, Physical Review Letters **79**, 2089 (1997).

[107] A. D. Becke, *Density-Functional Thermochemistry. III. The Role of Exact Exchange*, Journal of Chemical Physics **98**, 5648 (1993).

[108] C. Adamo and V. Barone, *Toward reliable density functional methods without adjustable parameters: The PBE0 model*, The Journal of Chemical Physics **110**, 6158 (1999).

[109] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *Hybrid functionals based on a screened Coulomb potential*, The Journal of Chemical Physics **118**, 8207 (2003).

[110] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *Erratum: Hybrid functionals based on a screened Coulomb potential [J. Chem. Phys. 118, 8207 (2003)]*, The Journal of Chemical Physics **124**, 219906 (2006).

[111] A. Alkauskas and A. Pasquarello, *Effect of improved band-gap description in density functional theory on defect energy levels in alpha-quartz*, Physica B: Condensed Matter **401-402**, 670 (2007).

[112] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Physical Review Letters **77**, 3865 (1996).

[113] J. P. Perdew, M. Ernzerhof, and K. Burke, *Rationale for mixing exact exchange with density functional approximations*, The Journal of Chemical Physics **105**, 9982 (1996).

[114] T. S. Bjorheim, A. Kuwabara, I. Ahmed, R. Haugsrud, S. Stolen, and T. Norby, *A combined conductivity and DFT study of protons in PbZrO3 and alkaline earth zirconate perovskites*, Solid State Ionics **181**, 130 (2010).

[115] S. Steinsvik, Y. Larring, and T. Norby, *Hydrogen ion conduction in iron-substituted strontium titanate, SrTi1xFexO3x/2 (0x0.8)*, Solid State Ionics **143**, 103 (2001).

[116] P. Finn W., *Speculations on the existence of hydride ions in proton conducting oxides*, Solid State Ionics **145**, 387 (2001).

[117] Y. Iwazaki, T. Suzuki, and S. Tsuneyuki, *Negatively charged hydrogen at oxygen-vacancy sites in BaTiO3: Density-functional calculation*, Journal of Applied Physics **108**, 083705 (2010).

[118] D. Kim, B. Kim, and Y. Kim, *(In press) Energy barriers for proton migration in yttrium-doped barium zirconate super cell with Sigma5 (310)/[001] tilt grain boundary*, Solid State Ionics .

[119] D. C. van der Laan, T. J. Haugan, P. N. Barnes, D. Abraimov, F. Kametani, D. C. Larbalestier, and M. W. Rupich, *The effect of strain on grains and grain boundaries in YBa 2 Cu 3 O 7 coated conductors*, Superconductor Science and Technology **23**, 014004 (2010).

[120] N. Sata, K. Eberman, K. Eberl, and J. Maier, *Mesoscopic fast ion conduction in nanometre-scale planar heterostructures*, Nature **408**, 946948 (2000).

[121] J. Maier, *Nanoionics: ion transport and electrochemical storage in confined systems*, Nature Materials **4**, 805815 (2005).

[122] J. Garcia-Barriocanal, A. Rivera-Calzada, M. Varela, Z. Sefrioui, E. Iborra, C. Leon, S. J. Pennycook, and J. Santamaria, *Colossal Ionic Conductivity at Interfaces of Epitaxial ZrO2:Y2O3/SrTiO3 Heterostructures*, Science **321**, 676 (2008).

[123] H. J. Avila-Paredes, E. Barrera-Calva, H. U. Anderson, R. A. D. Souza, M. Martin, Z. A. Munir, and S. Kim, *Room-temperature protonic conduction in nanocrystalline films of yttria-stabilized zirconia*, J. Mater. Chem. **20**, 6235 (2010).

Paper I

Protonic defects in pure and doped $La_2Zr_2O_7$ pyrochlore oxide
M. E. Björketun, C. S. Knee, B. J. Nyman and G. Wahnström
Solid State Ionics **178**, 1642 (2008).

Paper II

Substitutional doping and oxygen vacancies in $La_2Zr_2O_7$ pyrochlore oxide
B. J. Nyman, M. E. Björketun and G. Wahnström
Solid State Ionics **189**, 19 (2011)

Paper III

Oxygen vacancy segregation and space-charge effects in grain boundaries of dry and hydrated $BaZrO_3$

B. J. Nyman, E. E. Helgee and G. Wahnström

Applied Physics Letters, doi:10.1063/1.3681169 (2012)

Paper IV

Grain boundaries in BaZrO3: Oxygen vacancy segregation and space charge effects from first principles and atomistic simulations
E. E. Helgee, B. J. Nyman, A. Lindman and G. Wahnström
(in manuscript)

Paper V

Equilibrium defect structure of yttrium-doped barium zirconate with oxygen deficient grain boundary cores
B. J. Nyman, P. Erhart and G. Wahnström
(in manuscript)