Development of a Coupled Molecular Dynamics / Lattice Monte Carlo Scheme

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II
Chemie, Physik und Mathematik

der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt von

Herrn Gabriel Kabbe
geb. am 15.09.1986

Erstgutachter/Betreuer

Prof. Dr. Daniel Sebastiani
MLU Halle-Wittenberg
Institut für Physikalische Chemie
Von-Danckelmann-Platz 4
06120 Halle (Saale)

Zweitgutachter

Prof. Dr. Gotthard Seifert
Technische Universität Dresden
Fakultät Chemie und Lebensmittelchemie
Professur für Theoretische Chemie
König-Bau, Bergstr. 66c
01062 Dresden

Verteidigungsdatum: 2. Juli 2018

Hiermit versichere ich, die vorliegende Arbeit selbst und ohne weitere Hilfsmittel, außer den angegebenen, angefertigt zu haben. Diese Arbeit wurde nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Halle, den

Publication List

Kabbe G., Wehmeyer C. and Sebastiani D.

A Coupled Molecular Dynamics/Kinetic Monte Carlo Approach for Protonation Dynamics in Extended Systems

J. Chem. Theory. Comput. **2014**, 10, 4221–4228

For this publication, I developed a coupled Molecular Dynamics / Lattice Monte Carlo scheme which models proton diffusion in phosphonic acid-functionalized molecular systems. I analyzed *ab initio* trajectories of Hexakis(*p*-phosphonatophenyl)benzene (*p*-6PA-HPB) to determine proton transfer rates, which were then used as input for the Lattice Monte Carlo scheme. Furthermore, I analyzed the proton dynamics of the system in non-equilibrium settings. Christoph Wehmeyer provided the *ab initio* trajectories and gave me advice on the manuscript and the method development. Daniel Sebastiani supervised the project giving me valuable advice during the development and while I wrote the manuscript.

Kabbe G., Dreßler C. and Sebastiani D.

Toward Realistic Transfer Rates within the Coupled Molecular Dynamics/Lattice Monte Carlo Approach

J. Phys. Chem. C **2016**, 120, 19905–19912

For this publication, I continued development on the coupled Molecular Dynamics / Lattice Monte Carlo scheme by refining several aspects. I derived proton transfer rates from separate quantum chemical calculations, which I calculated. Furthermore, I extended the functional form of the proton jump rates used within the coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) scheme to include an additional parameter for the hydrogen bond angle. Christian Dreßler provided the *ab initio* trajectories of CsH_2PO_4. Daniel Sebastiani supervised the project giving me valuable advice during the development and while I wrote the manuscript.

Dreßler C., Kabbe G. and Sebastiani, D.

Proton Conductivity in Hydrogen Phosphate/Sulfates from a Coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach

J. Phys. Chem. C **2016**, 120, 19913–19922

For this publication, the proton conductivity of two solid acids was investigated by means of *ab initio* simulations and the cMD/LMC approach. Christian Dreßler provided the *ab initio* trajectories and analyzed the systems. I adapted the Lattice Monte Carlo scheme to the new molecular systems, and analyzed the proton transfer rates from the *ab initio* trajectories.

Dreßler C., Kabbe G. and Sebastiani, D.
Insight from Atomistic Simulations of Protonation Dynamics at the Nanoscale
Fuel Cells **2016**, 16, 682–694
This publication constitutes a review of theoretical findings with regard to proton transfer in aqueous media and water-free proton conductors. My focus lay mainly on the review of articles about proton transfer in Nafion-like membranes and within nano-sized pores. C. Dreßler provided reviews about solid acids and molecular liquids.

Kabbe G., Dreßler C. and Sebastiani, D.
Proton Mobility in Aqueous Systems: Combining ab initio Accuracy with Millisecond Timescales
Phys Chem Chem Phys **2017**, 19, 28604–28609
For this publication, I investigated the applicability of the cMD/LMC scheme on highly fluctuating systems. I calculated force field trajectories of water, while Christian Dreßler provided *ab initio* trajectories of H_3O^+ in water. Based on my analysis of H_3O^+ excess charge diffusion in water, I incorporated the dielectric relaxation effects which can be observed within the first solvation shell of H_3O^+ into the cMD/LMC scheme. Daniel Sebastiani supervised the project giving me valuable advice during the development and while I wrote the manuscript.

Contents

1 Introduction 3

2 Theory 7
- 2.1 Molecular Dynamics 7
  - 2.1.1 Ensemble Averages 7
  - 2.1.2 Molecular Dynamics based on Force Fields 8
  - 2.1.3 Verlet Algorithm 9
  - 2.1.4 Velocity-Verlet Algorithm 10
  - 2.1.5 Thermostats 10
- 2.2 Electronic Structure Methods 12
  - 2.2.1 Hartree Fock 14
  - 2.2.2 Roothaan Equations and Self-Consistent Field Approach 15
  - 2.2.3 Homogeneous Electron Gas 16
  - 2.2.4 Density Functional Theory 18
- 2.3 Monte Carlo Methods 21
  - 2.3.1 Basic Principles 21
  - 2.3.2 Metropolis-Hastings Algorithm 24
  - 2.3.3 Random Walk 25
  - 2.3.4 Inverse Transform Sampling 27
  - 2.3.5 Asymmetric Simple Exclusion Process (ASEP) 28
  - 2.3.6 Kinetic Monte Carlo 29

3 Algorithmic Description 31
- 3.1 Coupling of Molecular Dynamics and Lattice Monte Carlo Scheme . . . 31
- 3.2 Coupled MD/LMC Scheme using the Asymmetric Exclusion Process . . 32
- 3.3 Excess Charge Kinetic Monte Carlo Algorithm 33

4 Publications 37
- 4.1 Publication Overview 37
  - 4.1.1 Proton Transport in Solid Acids 37

        4.1.2 Water-Free Proton Transport 38
        4.1.3 Water-Mediated Proton Transport 39
    4.2 A coupled Molecular Dynamics / kinetic Monte Carlo Approach for Protonation Dynamics in Extended Systems 41
    4.3 Toward Realistic Transfer Rates within the Coupled Molecular Dynamics/Lattice Monte Carlo Approach 50
    4.4 Proton Conductivity in Hydrogen Phosphate/Sulfates from a Coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach . . . 59
    4.5 Proton Mobility in Aqueous Systems: Combining ab initio Accuracy with Millisecond Timescales 71
    4.6 Insight from Atomistic Simulations of Protonation Dynamics at the Nanoscale 94

5 Academic Curriculum Vitae **95**

6 Conclusion **99**

7 Acknowledgements **101**

8 References **103**

Introduction

Proton transfer is a ubiquitious phenomenon and it lies at the core of many scientific problems. One such field is biology, where proton transport across membrane proteins is frequently encountered and is considered one of the fundamental building blocks of nerves, muscles and synapses. [1–3] Another important field is the technical sector, where fuel cell technology takes advantage of proton transfer processes to efficiently convert energy stored in a fuel, such as hydrogen, into electrical energy. This is done by dividing the exothermic redox reaction, in which hydrogen reacts with oxygen to water, into two separated processes: At the anode, an oxidation reaction takes place, dissociating the incoming hydrogen into protons and electrons, whereas at the cathode protons react with oxygen and electrons to water. These two reactions are separated spatially by an electrolyte, which only permits protons to pass, and thus forces the electrons to move through a connected electric circuit in order to reach the cathode.

For an efficient operation of the fuel cell, the membrane material of the electrolyte needs to meet several conditions: High ionic conductivity (preferably in low humidity conditions), high thermal and oxidative stability, and cost effectiveness. [4] In search of materials fulfilling these demands, many studies have been conducted both experimentally and theoretically. [5–36]

One of the most commonly used water-based membrane materials is Nafion. It consists of perfluorinated vinyl ether groups attached to a tetrafluoroethylene backbone. [37] While the backbone is hydrophobic, the hydrophilic SO_3^- groups lead to the formation of water channels under hydrated conditions. Therefore, the proton conduction properties of Nafion can be mostly attributed to the proton conductivity of bulk water. This, however, limits the possible operating temperature of Nafion membranes, as temperatures above 100 °C lead to water loss and thus a decrease in conductivity.

Another group of membrane materials which are particularly in the focus of this thesis, are phosphonic acid- and phosphoric acid-functionalized compounds. In contrast to Nafion-based compounds, materials based on phosphonic or phosphoric acid groups are a suitable choice for low-humidity electrolytes in fuel cells. [38, 39] One of them is Hexakis(*p*-phosphonatophenyl)benzene (*p*-6PA-HPB) [40, 41], an aromatic, disk-shaped compound with six phosphonic acid groups per molecule. In contrast to Nafion, its proton conductivity stays constant over a temperature range from 100 °C to 180 °C. [40] X-ray diffraction experiments show that *p*-6PA-HPB molecules arrange in a columnar supramolecular structure. Because of this anisotropic arrangement, proton transfer takes place preferably along the direction of the columns.

Solid acids are another candidate for fuel cell membrane materials as they exhibit proton conductivity under low-humidity conditions and temperature stability. [42, 43] In this thesis, the focus will be on solid acids of the form CsH_2PO_4 and $CsHSO_4$. Solid acids have the interesting property to undergo a superprotonic phase transition, during which the proton conductivity increases by several orders of magnitude. [44–48] In the low temperature phase their hydrogen bond network is highly ordered, which leads to a decrease in proton conductivity. Higher temperatures lead to a phase transition (around 400 K for $CsHSO_4$ and 500 K for CsH_2PO_4) resulting in a disordered hydrogen bond network and a strong increase in proton conductivity.

On the microscopic scale, two main mechanisms have been found to be responsible for proton conductivity: The *vehicle mechanism* describes the diffusion of protons bonded to a carrier such as H_2O or NH_3. [49] In this case, the proton conductivity is limited by the diffusivity of the carrier molecule. On the other hand, the *Grotthuss mechanism* describes charge transfer by means of a hopping process of charge carriers between host molecules without the requirement of a diffusive motion of the host.[50] Although it was proposed more than 200 years ago, only in recent years have Molecular Dynamics (MD) studies revealed its finer details. [23, 36]

Simulation with Molecular Dynamics In the field of theoretical chemistry, molecular dynamics simulations have become an invaluable tool to model chemical processes on an atomistic level. By numerically calculating the temporal evolution of atomic systems they are able to reveal both structural and dynamic characteristica of complex chemical compounds. The core idea of molecular dynamics methods is simple: given an ensemble of interacting atoms, the corresponding equations of motion are integrated, and thus the phase space of the atomic system is sampled via its time evolution. By calculation of ensemble averages it is then possible to determine macroscopic quantities from the microscopic molecular dynamics trajectories. In combination with a multitude of methods which enable the control of external variables such as pressure or temperature, molecular dynamics allow to carry out “in-silico” experiments in which the external variables are adjusted to experimental conditions, while also giving insight into the microscopic picture of chemical processes or structural features.

The quality of a molecular dynamics simulation depends strongly on the accuracy of the interatomic potentials. One common approach is the use of empirical potentials which describe the potential energy surface landscape of molecular systems by means of parameterized functions. These functions can take geometric parameters such as bond lengths or angles between bonds into account, and are often fitted to quantumchemical and experimental data with the aim to reproduce selected properties (for example internal geometries, vibrations or conformational energies [51]). The advantage of parameterized approaches lies in the fast function evaluation which allows system sizes of up to several nanometers and timescales up to milliseconds. Their main drawback, on the other hand, is the fixed bond topology, which does not allow the simulation of bond breaking reactions.

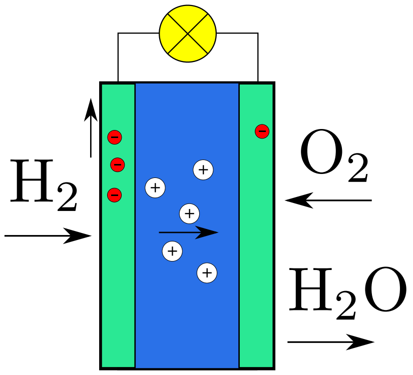

Figure 1.1: Schematic depiction of the oxidation-reduction process inside a fuel cell. Hydrogen is split into protons and electrons at the anode. While the protons travel through the electrolyte, the electrons move through an external electric circuit.

An alternative approach is to explicitly solve the Schrödinger equation of the electronic subsystem and thus calculate chemical bonds and chemical reactions. The methods employing this approach are subsumed under the term *ab initio*. For a fixed atomic configuration, the energy eigenvalue of the electronic Schrödinger equation represents the interatomic interaction potential. This potential in turn allows the determination of interatomic forces, and therefore the time-evolution of a chemical system by explicit integration of the equations of motions. In this approach, the molecular bond topology is the result of the calculation, not the input as for classical MD . Electronic structure-based MD simulations are a well-suited choice for the investigation of proton transport processes. In general, however, the time scales reached by *ab initio* methods are several orders of magnitude lower than those of parameterized approaches.

Scale-Bridging with Kinetic Monte Carlo Monte Carlo is an umbrella term for a large array of algorithmic approaches which use random sampling in order to calculate numerical results. The focus in this thesis lies especially on so-called “Kinetic Monte Carlo” methods. [52–54] This class of Monte Carlo methods is commonly used to model time-dependent processes based on predetermined transfer rates. They are commonly used to model phenomena such as adsorption processes[55–59], charge transport[60–62], oxidation reactions[63–65], hydrogenation[66–68], deposition processes[69–71], crystal growth[72–74] or diffusion processes. [75–78] In this thesis, kinetic Monte Carlo is used to model proton transport/diffusion in various compounds.

cMD/LMC Scheme The scheme which has been developed in the context of this work combines Molecular Dynamics with a kinetic Monte Carlo like algorithm. The translational diffusion of protons is modeled by the kinetic Monte Carlo scheme, whereas the temporal evolution of all heavier atoms (and molecules) is determined via MD . The unique feature of this combined approach developed is that the combination of MD and kinetic Monte Carlo allows to calculate the proton diffusion on time scales only accessible to kinetic Monte Carlo while still taking into account the accurate temporal evolution of the proton pathways as determined by MD .

Theory

2.1 Molecular Dynamics

Molecular Dynamics (MD) is a numerical simulation technique which is commonly used to determine macroscopic and dynamical properties of systems of interacting particles. This is done by explicitly integrating the Newtonian equations of motion of each particle in the system. Because of its versatility, it has become a common tool to tackle various problems regarding condensed matter systems.

MD can be roughly divided into two large categories: Force-Field MD (or classical MD) and *ab initio* MD . While the former uses parameterized interatomic potentials to determine the forces acting within the system, the latter actually solves the Schrödinger equation of the electronic subsystem in order to determine the forces acting on each atom. The advantages of classical MD lie in the low computational costs of the force-field calculations. This allows the simulation of large systems such as proteins or even viral structures. [79]

The use of *ab initio* MD on the other hand allows the simulation of chemical reactions without heuristic approaches or parameterizations. The term *ab initio* is a collective name for a variety of different methods which solve the electronic Schrödinger equation by means of different approaches and with varying degrees of accuracy. The focus in this thesis will be restricted to Hartree-Fock, which represents a wave function-based approach, and Density Functional Theory, which minimizes the energy as a functional of the electron density.

2.1.1 Ensemble Averages

Molecular dynamics simulations can be seen as a virtual experiment carried out on the computer. At each time step in the simulation, the positions and momenta of the atoms are propagated. This way, each time step yields a point in the system's phase space.

In order to arrive at the macroscopic value of some observable a, one would in theory have to integrate over the whole phase space:

$$\langle a \rangle = \int \rho\left(\mathbf{r}_1,\ldots,\mathbf{r}_N,\mathbf{p}_1,\ldots,\mathbf{p}_N\right) a(\mathbf{r}_1,\ldots,\mathbf{r}_N,\mathbf{p}_1,\ldots,\mathbf{p}_N) d\mathbf{r}_1 \ldots d\mathbf{r}_N d\mathbf{p}_1 \ldots d\mathbf{p}_N \tag{2.1}$$

where ρ is the probability density to find the system at the corresponding point in phase space.

If the system is ergodic, i.e. if given enough time, it will visit all points in phase space eventually, then the integration over phase space can be replaced by an integral over time.

$$\langle a \rangle = \lim_{T\to\infty} \frac{1}{T} \int\limits_0^T a\left(\mathbf{r}_1(t),\ldots \mathbf{r}_N(t),\mathbf{p}_1(t),\ldots \mathbf{p}_N(t)\right) dt \tag{2.2}$$

In this form, Molecular Dynamics allows the evaluation of $\langle a \rangle$ by propagating the molecular system in discrete time steps and approximating the integral of equation 2.2 by a finite sum.

$$\langle a \rangle \approx \frac{1}{N_{\text{steps}}} \sum_i^{N_{\text{steps}}} a\left(\mathbf{r}_1(i\Delta t),\ldots \mathbf{r}_N(i\Delta t),\mathbf{p}_1(i\Delta t),\ldots \mathbf{p}_N(i\Delta t)\right) \tag{2.3}$$

2.1.2 Molecular Dynamics based on Force Fields

One approach to calculate the interatomic potentials within a Molecular Dynamics simulation is by means of parameterized functions which take geometric parameters of the system, such as bond lengths or bond angles, as arguments.

In this approach, the total energy of the system is composed of a set of potential functions. These include stretching (U_{bond}), bending (U_{angle}) and torsional ($U_{\text{torsional}}$ and U_{improper}) interaction energies, plus a van-der-Waals term for non-bonded interactions, and a Coulomb term for electrostatic interactions. [80]

$$U_{\text{total}} = U_{\text{bond}} + U_{\text{angle}} + U_{\text{torsional}} + U_{\text{improper}} + U_{\text{vdW}} + U_{\text{Coulomb}} \tag{2.4}$$

The stretching potential describes each covalent bond as a parabolic function characterized by an equilibrium distance r_{0i} and a constant k_i^{bond} which determines the bond strength. The internal coordinate r_i is the distance between the two atoms forming the ith bond.

$$U_{\text{bond}} = \sum_i k_i^{\text{bond}} (r_i - r_{0i})^2 \tag{2.5}$$

The bending potential is defined by the same functional form. The internal coordinate Θ_i is the angle between the ith pair of covalent bonds, and Θ_{0i} describes the equilibrium angle.

$$U_{\text{angle}} = \sum_i k_i^{\text{angle}} (\Theta_i - \Theta_{0i})^2 \tag{2.6}$$

The torsional potential is described by a sinusoidal function. The internal coordinate Φ_i describes for three consecutive covalent bonds the angle between the first and the last bond. The integer n describes the periodicity of the potential.

$$U_{\text{torsional}} = \sum_i k_i^{\text{dihed}} \left[1 + \cos\left(n_i \Phi_i - \gamma_i\right)\right] \tag{2.7}$$

Furthermore, an improper potential is added with the aim to keep selected groups planar or to conserve chirality.

The van-der-Waals potential is described by a Lennard-Jones potential

$$\begin{aligned} U_{\text{vdW}} &= \frac{1}{2} \sum_{i \neq j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \\ U_{\text{Coulomb}} &= \frac{1}{2} \sum_{i \neq j} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \end{aligned} \tag{2.8}$$

The parameter ϵ_{ij} determines the well depth of the potential between atom i and atom j. σ_{ij} determines the equilibrium distance.

Thanks to the explicit analytic form of the interatomic potentials, force field MD can be computed very effectively on computers. This allows the simulation of large systems up to several nanometers and on millisecond timescales.

2.1.3 Verlet Algorithm

If we assume that the potential acting on the particles is conservative, i.e. it depends only on their positions, the forces acting on each particle are given by

$$\mathbf{F}_i = -\nabla_i V(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \tag{2.9}$$

In order to arrive at an expression for a numerical integration scheme, one starts with a Taylor expansion of the position vector up to second order of the ith particle in the system:

$$\mathbf{r}_i(t + \Delta t) \approx \mathbf{r}_i(t) + \frac{d\mathbf{r}_i}{dt} \Delta t + \frac{1}{2} \frac{d^2\mathbf{r}_i}{dt^2} \Delta t^2 \tag{2.10}$$

Using Newton's second law, the acceleration can be replaced by the acting forces.

$$\mathbf{r}_i(t + \Delta t) \approx \mathbf{r}_i(t) + \frac{d\mathbf{r}_i}{dt} \Delta t + \frac{\mathbf{F}_i}{2m_i} \Delta t^2 \tag{2.11}$$

In order to eliminate the derivative of **r** in the equation, we calculate $\mathbf{r}_i(t+\Delta t)+\mathbf{r}_i(t-\Delta t)$ using equation 2.11.

$$\mathbf{r}_i(t+\Delta t) \approx 2\mathbf{r}_i(t) - \mathbf{r}_i(t-\Delta t) + \frac{\mathbf{F}_i}{m_i}\Delta t^2 \tag{2.12}$$

This equation is also known as the Verlet algorithm. [81]

In order to determine the velocities, an equation for $\mathbf{r}(t-\Delta t)$ can be determined in analogy to equation 2.11. Subtracting it then from equation 2.11 yields:

$$\mathbf{v}_i(t) = \frac{d\mathbf{r}_i(t)}{dt} = \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t-\Delta t)}{2\Delta t} \tag{2.13}$$

2.1.4 Velocity-Verlet Algorithm

The Velocity-Verlet algorithm represents a modification of the standard Verlet algorithm and offers explicit formulas for both positions and velocities. For this, starting from time $t_{\text{next}} = t + \Delta t$, a Taylor expansion is done for $\mathbf{r}_i(t_{\text{next}} - \Delta t)$:

$$\mathbf{r}_i(t) = \mathbf{r}_i(t_{\text{next}} - \Delta t) \approx \mathbf{r}_i(t+\Delta t) - \frac{d\mathbf{r}_i(t+\Delta t)}{dt}\Delta t + \frac{\mathbf{F}_i(t+\Delta t)}{2m_i}\Delta t^2 \tag{2.14}$$

Inserting the above equation in equation 2.11, yields the expression

$$\mathbf{v}_i(t+\Delta t) = \mathbf{v}_i(t) + \frac{\mathbf{F}_i(t) + \mathbf{F}_i(t+\Delta t)}{2m_i}\Delta t \tag{2.15}$$

This way, together with equation 2.11, this allows for the simultaneous calculation of positions and velocities. The direct access to positions and velocities via equations 2.11 and 2.15 is especially useful when employing constant temperature or constant pressure algorithms. [82]

2.1.5 Thermostats

By combining one of the Verlet integration algorithms with a suitable potential, it is possible to compute a trajectory of a molecular system. As the system is fully isolated (there is no external system it could exchange energy with), and the equations of motion conserve the energy, the energy of the system will be constant over time. Given that the number of particles N and the system volume V is fixed, this corresponds to an NVE-Ensemble. Oftentimes, however, it is desirable to sample from other ensembles, such as a constant temperature (NVT) or constant pressure (NPT) ensemble, as they allow to run simulations under “experimental conditions”, where a fully isolated system cannot be realized. While it is possible to relate the system temperature to the average kinetic energy of the particles, this only allows to determine the temperature in retrospect, but ideally the temperature should be a controllable parameter.

To address this problem, a multitude of algorithms were developed which simulate the interaction of the atomic system with a surrounding heat bath. [83] The discussion here will be restricted to the Nosé-Hoover thermostat.

Nosé-Hoover Thermostat

In the canonical ensemble (NVT), the system is coupled to a large heat bath, with which it can exchange heat. Consequently, the system will assume the temperature of the heat bath. In order to achieve the same effect in molecular dynamics simulations, the Nosé-Hoover thermostat modifies the equations of motions by adding an additional term [84, 85]:

$$\frac{dq_i}{dt} = \frac{p_i}{m_i} \tag{2.16}$$

$$\frac{dp_i}{dt} = -\frac{\partial V}{\partial q_i} - \xi p_i \tag{2.17}$$

$$\frac{Q}{2}\frac{d\xi}{dt} = \sum_i \frac{p_i^2}{2m_i} - \frac{3}{2}NkT \tag{2.18}$$

In the above equations, q_i denotes the position, p_i the momentum, and m_i the mass of the ith particle. V is the potential acting on the particles. ξp_i acts similarly to a friction term, but as the change in ξ depends on the difference between the system temperature and the target temperature, it will counteract too large deviations from the target temperature. The new degrees of freedom introduced by the additional terms, can be thought of as a fictitious particle with mass Q and velocity ξ. Therefore, one can analogously define

$$p_\xi = Q\xi \tag{2.19}$$

The change of the equations of motions give rise to a new conserved quantity

$$H = \sum_i \frac{p_i^2}{2m_i} + V(\mathbf{q}) + \frac{Q}{2}\xi^2 + 3NkT\ln(s) \tag{2.20}$$

with

$$\frac{d\ln(s)}{dt} = \xi \tag{2.21}$$

It was shown that these equations produce a canonical distribution only if the system is ergodic. [85] To remedy this shortcoming, the Nosé-Hoover method was extended to include a set of M thermostats which are coupled as follows:

$$Q_1\dot{\xi}_1 = \sum_i \frac{p_i^2}{m_i} - 3NkT - \xi_2 p_{\xi_1} \tag{2.22}$$

$$Q_j\dot{\xi}_j = \frac{p_{\xi_{j-1}}^2}{Q_{j-1}} - 3kT - \xi_{j+1} p_j \tag{2.23}$$

$$Q_M\dot{\xi}_M = \frac{p_{\xi_{M-1}}^2}{Q_{M-1}} - 3kT \tag{2.24}$$

with the corresponding Hamiltonian

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + V(\mathbf{q}) + \sum_{i=1}^{M} \frac{p_{\xi_i}^2}{Q_i} + 3NkT\ln(s_1) + \sum_{i=2}^{M} kT\ln(s_i) \tag{2.25}$$

The choice of sensible values for the Q_i is essential, as for very large masses Q_i, the distribution will approach the microcanonical distribution, whereas too large masses may lead to the inhibition of the momenta. [85]

2.2 Electronic Structure Methods

The essence of electronic structure methods is the quantum mechanical treatment of molecular systems by solving their corresponding Schrödinger equation. The Hamiltonian of any chemical system can be written in the following form:

$$\hat{H}_{\text{mol.}} = \hat{T}_{\text{nuc.}} + \hat{T}_{\text{el.}} + \hat{V}_{\text{ee}} + \hat{V}_{\text{ne}} + \hat{V}_{\text{nn}} \tag{2.26}$$

consisting of (from left to right) the kinetic energy operators of the nuclei and the kinetic energy operator of the electrons, as well as the interaction potentials describing electron-electron, nucleus-electron and nucleus-nucleus interactions. By solving the time-independent Schrödinger equation

$$\hat{H}_{\text{mol.}} \left|\Psi_{\text{mol.}}\right\rangle = E \left|\Psi_{\text{mol.}}\right\rangle \tag{2.27}$$

it is in principle possible to determine all properties of a chemical system. This is a daunting task, however, as a full quantum mechanical treatment involves the determination of the full many-particle wave function containing degrees of freedom for both nuclei and electrons. In order to reduce the complexity of equation 2.26, it is thus advisable to make a few approximations.

The most prominent one is the *Born-Oppenheimer approximation* (also known as *adiabatic approximation*), which assumes that the full wave function $\Psi(\mathbf{r},\mathbf{R})$ can be expressed by the ansatz

$$\Psi(\mathbf{r},\mathbf{R}) = \Psi_{\text{nuc.}}(\mathbf{R})\Psi_{\text{el.}}(\mathbf{r},\mathbf{R}) \tag{2.28}$$

where **r** denotes the degrees of freedom of the electronic subsystem, and **R** denotes the degrees of freedom of the nuclear subsystem.

The qualitative justification for this ansatz lies in the fact that the masses of nuclei and electrons differ by around three orders of magnitude. Thus, the electrons can be considered to adapt instantaneously to the position of the slowly moving cores.

In the context of this work, the typical form of the Schrödinger equation usually consists of the electronic Hamiltonian alone. For an atomistic/molecular system, it has the form

$$\hat{H}_{\text{el.}} = \hat{T}_{\text{el.}} + \hat{V}_{\text{ee}} + \hat{V}_{\text{ext}} \tag{2.29}$$

consisting of the electronic kinetic energy operator

$$\hat{T}_{\text{el.}} = -\frac{1}{2}\nabla^2 \tag{2.30}$$

the electron-electron interaction

$$\hat{V}_{\text{ee}} = \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.31}$$

and an external potential

$$\hat{V}_{\text{ext}} = -\sum_{i}^{N_{\text{el.}}} \sum_{j}^{N_{\text{nuc.}}} \frac{1}{2} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \tag{2.32}$$

which includes electron-nucleus interactions, but can be extended to include interactions with external fields, as well.

According to the *Hellman-Feynman* theorem, the atomic forces can be obtained from the electronic Hamiltonian via

$$\mathbf{F}_i = -\frac{\partial}{\partial \mathbf{R}} \langle \Psi_0 | H_{\text{el.}} | \Psi_0 \rangle \tag{2.33}$$

Thus, an approach for an *ab initio* simulation can be devised by determining the ground state energy $|\Psi_0\rangle$ of the Hamiltonian $H_{\text{el.}}(\mathbf{R})$ at each MD time step, and propagating the positions and momenta of the nuclei according to the forces obtained by the Hellman-Feynman equation.

Different approaches have been developed to find numerical solutions of equation 2.29. The focus of this thesis will be on the *Hartree-Fock* method and *Density Functional Theory*. While the former constructs an ansatz for the many-particle wave function from a Slater determinant of single-particle wave functions, the latter treats the energy of the electronic system as a functional of the electron density. In the following, a summary of the different methods will be given.

2.2.1 Hartree Fock

The Hartree-Fock method is one of the central approximations of theoretical chemistry, and serves as a basis for more advanced methods. The following discussion mostly follows the explanation by Szabo and Ostlund. [86]

The starting point of Hartree-Fock is the construction of the electron wave function as a Slater determinant of one-particle spin orbitals $\chi_i(\mathbf{x})$ where the vector **x** contains both positional coordinates **r** and a spin coordinate σ.

$$|\Psi_0\rangle = \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \dots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \dots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \dots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.34}$$

By construction the wave function $|\Psi_0\rangle$ is antisymmetric with respect to exchange of any two coordinates, therefore fulfilling the Pauli exclusion principle.

In order to find the optimal wave function within this ansatz, one minimizes (according to the variational principle)

$$E[\Psi_0] = \langle\Psi_0| \hat{H}_{\text{el.}} |\Psi_0\rangle \tag{2.35}$$

under the constraint that the spin orbitals are orthonormal

$$\langle\chi_i|\chi_j\rangle = \delta_{ij} \tag{2.36}$$

The minimization can be achieved by means of the method of Lagrange multipliers, and the result is a new eigenvalue equation for each orbital.

$$\hat{f}\chi_i = \epsilon_i\chi_i \tag{2.37}$$

The Fock operator $\hat{f}$ consists of three components shown below:

$$\hat{f} = \hat{h} + \sum_b \mathcal{J}_b - \mathcal{K}_b \tag{2.38}$$

The operator $\hat{h}$ is simply the Hamiltonian of a single electron subject to the Coulomb potential due to the nuclei.

$$\hat{h}\chi_a = \left[-\frac{1}{2}\nabla^2 - \sum_{\alpha} \frac{Z_\alpha}{|\mathbf{r}_a - \mathbf{R}_\alpha|}\right] \chi_a(\mathbf{x}_1) \tag{2.39}$$

The operator $\hat{\mathcal{J}}$ describes the Coulomb potential acting on electron one caused by the averaged position of electron two. The summation in equation 2.38 can thus be interpreted as the total Coulomb energy between a single electron and the averaged charge of the remaining electrons.

$$\hat{\mathcal{J}}_b\chi_a = \left[\int d\mathbf{x}_2 \frac{\chi_b(\mathbf{x}_2)^*\chi_b(\mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\right] \chi_a(\mathbf{x}_1) \tag{2.40}$$

The operator $\hat{\mathcal{K}}$ on the other hand has no classical interpretation. As can be seen in equation 2.41, the calculation of $\hat{\mathcal{K}}\chi_a$ contains an integral which includes χ_a at every point in space. $\hat{\mathcal{K}}$ is therefore called non-local.

$$\hat{\mathcal{K}}_b\chi_a = \left[\int d\mathbf{x}_2 \frac{\chi_b^*(\mathbf{x}_2)\chi_a(\mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\right] \chi_b(\mathbf{x}_1) \tag{2.41}$$

In the special case of restricted close-shell Hartree-Fock, only systems with closed shells are considered, i.e. each orbital is occupied by two electrons. Furthermore, each spin orbital has the form

$$\chi_i(\mathbf{x}) = \begin{cases} \Psi_j(\mathbf{r})\alpha(\sigma) \\ \Psi_j(\mathbf{r})\beta(\sigma) \end{cases} \tag{2.42}$$

i.e. the spatial orbital Ψ_j is the same for α and β spin. In this case, after integrating over the spin coordinate, the Fock operator can be expressed as

$$\hat{f} = 2\hat{h} + \sum_b 2\mathcal{J}_b - \mathcal{K}_b \tag{2.43}$$

and the Hartree-Fock equations become

$$\hat{f}\,|\Psi_a\rangle = \epsilon_a\,|\Psi_a\rangle \tag{2.44}$$

2.2.2 Roothaan Equations and Self-Consistent Field Approach

In order to achieve a numerical solution of the previously discussed Hartree-Fock equations, it is helpful to convert them to matrix form by expanding the orbitals in a basis set of K orbitals.

$$|\Psi_a\rangle = \sum_i^K c_{ia} |\phi_i\rangle \tag{2.45}$$

Inserting the expansion into the Hartree-Fock equations (equation 2.44) and multiplying with $\langle\phi_j|$ from the left yields

$$\sum_i c_{ia} \langle\phi_j| \hat{f} |\phi_i\rangle = \sum_i c_{ia}\epsilon_a \langle\phi_j|\phi_i\rangle \tag{2.46}$$

By defining the overlap matrix **S**

$$S_{ji} = \langle\phi_j|\phi_i\rangle \tag{2.47}$$

the Fock matrix **F**

$$F_{ji} = \langle\phi_j| \hat{f} |\phi_i\rangle \tag{2.48}$$

and the coefficient matrix **C**

$$C_{ia} = c_{ia} \tag{2.49}$$

the Hartree-Fock equation can be written as

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon} \tag{2.50}$$

In the case of an orthonormal basis set, **S** reduces to unity, and the eigenvalue problem can be solved by diagonalization of **F**. It is important to note that the Fock operator, and therefore the Fock matrix, depend on the electron density due to the included operators $\hat{J}$ and $\hat{K}$, and therefore on the coefficient matrix **C**. In order to obtain a solution of equation 2.50, an iterative scheme is used. Based on a first guess of the orbital coefficients, the Fock operator can be calculated, and thus the coefficient matrix **C** can be found. This, in turn allows a further calculation of **F**. This scheme is repeated until the change in the electron density is below a specified threshold.

2.2.3 Homogeneous Electron Gas

The Thomas-Fermi model is often seen as the predecessor of modern density functional theory. It builds on the assumptions that electrons behave like a free electron gas within small volume elements, and that the interaction energies within an atom can be determined by the electron density alone. Based on these approximations, an explicit formula for the energy as a functional of the electron density can be found.

The derivations here are based on the explanation given by Yang and Parr. [87] The goal is to find an analytic expression for the kinetic energy of an electron gas as a functional of the electron density. We start with the assumption that electrons behave like free particles within a small Volume $\Delta V = l^3$. In this case, the Schrödinger equation can be solved directly, and the resulting energy levels are

$$E(n_x, n_y, n_z) = \frac{h^2}{8ml^2}\left(n_x^2 + n_y^2 + n_z^2\right) \tag{2.51}$$

where n_x, n_y and n_z are positive integers. The number of states with an energy smaller or equal to ϵ can therefore be approximated by counting the number of states within a sphere of radius $\left(n_x^2 + n_y^2 + n_z^2\right)^{1/2}$. As only positive integers are allowed, only the volume within the first octant is considered.

$$N = \frac{1}{8}\frac{4}{3}\pi\left(n_x^2 + n_y^2 + n_z^2\right)^{3/2} = \frac{\pi}{6}\left(\frac{8ml^2\epsilon}{h^2}\right)^{3/2} \tag{2.52}$$

The density of states $g(\epsilon)$ is therefore:

$$g(\epsilon) = \frac{dN}{d\epsilon} = \frac{\pi}{4}\left(\frac{8ml^2}{h^2}\right)^{3/2}\epsilon^{1/2} \tag{2.53}$$

The total energy (of a volume element) can then be found by integrating over all energies and considering that electrons are subject to the Fermi distribution:

$$f(\epsilon) = \frac{1}{\exp[(\epsilon - \epsilon_F)/kT] + 1} \tag{2.54}$$

In the limit $T \to 0$K, the Fermi distribution $f(\epsilon)$ becomes a step function. Thus, by integrating up to the Fermi energy ϵ_F, one obtains

$$\Delta E = 2\int^{\epsilon_F} g(\epsilon)\epsilon d\epsilon = \frac{8\pi}{5}\left(\frac{2m}{h^2}\right)^{3/2} l^3\epsilon_F^{5/2} \tag{2.55}$$

Together with the total number of electrons

$$\Delta N = 2\int^{\epsilon_F} g(\epsilon)d\epsilon = \frac{8\pi}{3}\left(\frac{2m}{h^2}\right)^{3/2} l^3\epsilon_F^{3/2} \tag{2.56}$$

and by integrating over all volume elements, one obtains the full kinetic energy functional:

$$T_{\text{T-F}}[\rho] = C_F\int \rho^{5/3}(\mathbf{r})d\mathbf{r} \tag{2.57}$$

with

$$C_F = \frac{3}{10}(3\pi^2)^{2/3} \tag{2.58}$$

Likewise is it possible to express the electron-electron interactions and the Coulomb interactions between electrons and cores in terms of the electron density. For a single atom, this gives rise to the total energy functional

$$E_{\mathrm{TF}}[\rho] = T_{\mathrm{TF}}[\rho] + \frac{1}{2}\int\int d\mathbf{r}d\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - Z\int d\mathbf{r}\frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}|} \tag{2.59}$$

Following physical intuition (or mimicking the procedure shown in section 2.2.1) one might be tempted to minimize the energy functional with respect to the electron density. But it should be noted that this has not been proven yet (for the justification of this procedure, see section 2.2.4). As before, the minimization should be constrained such that the total number of electrons stays constant.

$$\delta\left\{E[\rho] - \lambda\left[\int d\mathbf{r}\rho(\mathbf{r}) - N\right]\right\} = 0 \tag{2.60}$$

Calculating the functional derivative above yields an Euler-Lagrange equation:

$$\frac{5}{3}C_F\rho^{2/3} + \frac{1}{2}\int d\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \frac{Z}{|\mathbf{r}-\mathbf{R}|} = \lambda \tag{2.61}$$

By solving this equation together with the condition

$$\int d\mathbf{r}\rho(\mathbf{r}) = N \tag{2.62}$$

the electron density can be determined, and using the energy functional from equation 2.59, the total energy can be calculated.

The homogeneous electron gas, although conceptually simple, has been shown to yield reasonable numerical results for various systems. [88, 89] On the other hand, molecular bonding cannot be described, as the energy of a neutral molecule is always higher than that of the isolated atoms. [90] Furthermore, electron interactions are only treated classically.

2.2.4 Density Functional Theory

Density Functional Theory continues the idea of the Thomas Fermi model that an electronic system can be described by an energy functional of the electron density. However, it is shown that such an approach is not only an approximate technique, but can, in principle, be exact.

Hohenberg-Kohn Theorems

The Hohenberg-Kohn theorems[91] have formed the foundation of modern Density Functional Theory.

Theorem 1: The external potential $V_{\text{ext}}(\mathbf{r})$ is uniquely determined (apart from a constant) by the ground state electron density $n_0(\mathbf{r})$

The consequence of this theorem is that the electron density therefore also completely determines the Hamiltonian of the system.

Theorem 2: The total energy of the system can be written as a functional of n, whose minimization yields the ground state density for any external potential.

$$E\left[n\right] = \int V_{\text{ext}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + F\left[n\right] \tag{2.63}$$

with

$$F\left[n\right] = \langle\Psi|\,\hat{T} + \hat{V}\,|\Psi\rangle \tag{2.64}$$

Note that the choice of n is not completely free. Instead, the Hohenberg-Kohn Theorems are only valid for densities which stem from antisymmetric ground state wave functions. This condition is called *v-representability.*

Levy-Lieb Formulation

From a practical point of view, the Hohenberg-Kohn theorems do not provide an actual approach how to minimize the energy functional with respect to the electron density. The reason for this is the *v-representability*, as it is in general not possible to decide for a given density, whether it is v-representable or not.

Levy and Lieb[92, 93] provide a similar functional as Hohenberg and Kohn, which however does not require the density to be v-representable. Instead, a weaker condition is required: the *N-representability*, which only requires that the density stems from an antisymmetric wave function for N electrons.

The conditions for N-representability are:

$$n(\mathbf{r}) \geq 0 \tag{2.65}$$

$$\int n(\mathbf{r})d\mathbf{r} = N \tag{2.66}$$

$$\left|\nabla n(\mathbf{r})^{1/2}\right|^2 < \infty \tag{2.67}$$

A connection to the Hohenberg-Kohn theorems can be made as follows:

According to the variation principle, a wave function $|\Psi_{n_0}\rangle$ which yields the ground state density n_0 will fulfil the inequality

$$\langle\Psi_{n_0}|\,\hat{H}\,|\Psi_{n_0}\rangle \geq \langle\Psi_0|\,\hat{H}\,|\Psi_0\rangle \tag{2.68}$$

with the ground state wave function $|\Psi_0\rangle$. This is equal to

$$\langle\Psi_{n_0}|\,\hat{T}+\hat{V}\,|\Psi_{n_0}\rangle + \int v_{\text{ext}}(\mathbf{r})n_0(\mathbf{r})d\mathbf{r} \geq \langle\Psi_0|\,\hat{T}+\hat{V}\,|\Psi_0\rangle + \int v_{\text{ext}}(\mathbf{r})n_0(\mathbf{r})d\mathbf{r} \tag{2.69}$$

Identifying $\langle\Psi_0|\,\hat{T}+\hat{V}\,|\Psi_0\rangle$ as equation 2.64, one arrives at

$$\langle\Psi_{n_0}|\,\hat{T}+\hat{V}\,|\Psi_{n_0}\rangle \geq F[n_0] \tag{2.70}$$

Therefore, $F[n_0]$ can be found by minimizing $F[n]$ over all wave functions which yield the ground state density n_0:

$$F[n_0] = \min_{\Psi\rightarrow n_0} \langle\Psi|\,\hat{T}+\hat{V}\,|\Psi\rangle \tag{2.71}$$

Kohn-Sham Ansatz

The Kohn-Sham ansatz provides a practical approach to the determination of the electron density. By replacing the many-body system with a system of N non-interacting electrons, a direct calculation is made possible.

In the Kohn-Sham approach, the energy functional has the form

$$E_{\text{KS}}[n] = T_s[n] + \int v_{\text{ext}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + E_H[n] + E_{XC}[n] \tag{2.72}$$

T_s is simply the sum of kinetic energies of the N independent electrons

$$T_s = -\frac{\hbar^2}{2m_e}\sum_\sigma\sum_i^{N_\sigma}\langle\Psi_i^\sigma|\,\nabla^2\,|\Psi_i^\sigma\rangle \tag{2.73}$$

whereas $E_H[n]$ is the hartree energy

$$E_H[n] = \frac{1}{2}\int d\mathbf{r}d\mathbf{r}'\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.74}$$

The functional $E_{XC}[n]$ contains the difference between the independent-electron system, and the real system of interacting electrons. An accurate description of $E_{XC}[n]$ will therefore lead to an accurate ground state density.

In order to find the density which minimizes $E_H[n]$, one solves

$$\frac{\delta E_H}{\delta\Psi_i^{\sigma*}} \overset{!}{=} 0 \tag{2.75}$$

with the constraint of orthonormal orbitals.

This results in N equations similar to the Schrödinger equation

$$\left(-\frac{\hbar^2}{2m_e}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_H(\mathbf{r}) + v_{XC}(\mathbf{r})\right)\Psi_i^\sigma(\mathbf{r}) = \varepsilon_i \Psi_i^\sigma(\mathbf{r}) \tag{2.76}$$

where the Kohn Sham potential v_{KS} is defined as

$$v_{KS} = v_{\text{ext}}(\mathbf{r}) + v_H(\mathbf{r}) + v_{XC}(\mathbf{r}) \tag{2.77}$$

Analogously to section 2.2.2 these equations can be written in matrix form.

The above equations allow the determination of the electron density by means of a self-consistent iteration scheme. Starting from a guessed initial density, v_{KS} can be calculated, which then allows solving the Schrödinger equation for the single particle wave functions. The newly obtained wave functions, in return, allow an update of the density. This iteration can then be repeated until the electron density converges.

2.3 Monte Carlo Methods

Monte Carlo is an umbrella term for a variety of algorithms which use stochastic techniques in order to solve numerical problems. Its first prominent application was during Second World War in the Manhattan Project. Since then, the method has gained traction because of its straightforward use and the continuously increasing performance of computers. In contrast to molecular dynamics which deterministically creates trajectories of molecular systems, Monte Carlo methods statistically sample possible configurations of the system. Besides calculations of high-dimensional integrals, Monte Carlo methods are commonly used to determine equilibrium structures of molecules or dynamical features such as adsorption processes and diffusion.

2.3.1 Basic Principles

The discussion in this section mostly follows the explanations given by Kalos. [94]

Probabilities and Random Variables

In order to be able to talk unambiguously about random processes such as Monte Carlo, a few mathematical terms should be defined first which are important in the context of random events. A *random variable* is a variable which can assume a set of values which are the result of a random process. For example when rolling a six-sided dice, the random variable $X \in \{1, 2, 3, 4, 5, 6\}$ could be defined as the resulting number.

Each possible outcome of a random event E_i can be assigned a probability $p_i = P(E_i)$. The probabilities of the random events are defined such that they satisfy the

Kolmogorov axioms. [95]

1. The probability of an event E_i is a real number greater than or equal to zero for each event in the event space.
$$P(E_i) \geq 0 \tag{2.78}$$
2. The probability that at least one elementary element of the whole sample space occurs is one
$$P(\Omega) = 1 \tag{2.79}$$
3. For a set of mutually exclusive events, the following applies
$$\bigcup_i E_i = \sum_i P(E_i) \tag{2.80}$$

From these three axioms all other properties of probability can be derived.

In the case of the dice example, one quickly sees that the six possible outcomes are mutually exclusive. From 3) and 2) we can then deduce that

$$P(\{1,2,3,4,5,6\}) = \sum_{i=1}^{6} P(X = i) \overset{!}{=} 1 \tag{2.81}$$

Assuming that each possible outcome has the same probability, one arrives at $P(X = i) = \frac{1}{6}$.

Continuous Random Variables

In many cases, the random variables which describe a random process are not discrete, but can assume any real value within an interval. For example, the decay of an atomic nucleus can occur at any time τ between zero (immediate decay) and infinity (no decay). The probability to find a continuous random variable X within an interval $[a, b]$ is described by a probability density function $p(x)$.

$$P(a \leq X \leq b) = \int\limits_a^b p(x)dx \tag{2.82}$$

Furthermore, the probability density is normalized such that

$$\int\limits_{-\infty}^{\infty} p(x)dx = 1 \tag{2.83}$$

Expectation Values

The expectation value of a discrete random variable is defined as

$$\langle X \rangle = \sum_i P(X = x_i) x_i \tag{2.84}$$

In the case of a continuous variable, one has to replace the sum by an integral

$$\langle X \rangle = \int\limits_{-\infty}^{\infty} p(x) x dx \tag{2.85}$$

Monte Carlo Integration

The idea of Monte Carlo integration is to replace an (often high-dimensional) integral by an approximation where N samples X_i are drawn from a uniform distribution.

$$\int\limits_V f(x) dx \approx \frac{V}{N} \sum_i^N f(X_i) \tag{2.86}$$

A common example to demonstrate the use of random numbers to obtain a numerical result is the determination of the area of a quarter-circle. Figure 2.1 on the left side visualizes the approximation described in equation 2.86. For large sample sizes, the approximation will converge to the true integral.

There is another way the integral can be solved, however. By interpreting the quarter-circle as a two-dimensional function

$$f(x, y) = \begin{cases} 1 & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.87}$$

it is possible to determine an approximation by sampling from two-dimensional space.

$$\int\limits_V f(x, y) dx dy \approx \frac{V}{N} \sum_i^N f(X_i, Y_i) \tag{2.88}$$

A useful property of the Monte Carlo integration is the fact that it is straightforward to extend the integration method to arbitrary dimensions.

It should be noted that uniform distribution can be replaced by some other distribution g(x). In this case the integrand needs to be weighted by $\frac{1}{g(x)}$.

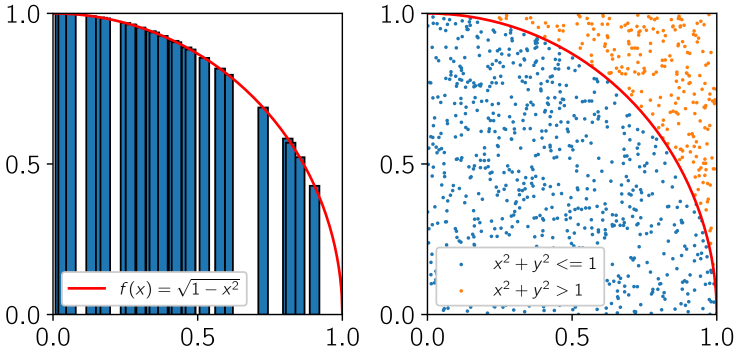

Figure 2.1: By repeatedly drawing random numbers, it is possible to approximate the area of the quarter circle. The integral can be approximated either by Monte Carlo integration in one dimension (left side) or integration over two dimensions (right side).

2.3.2 Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm is a method whose purpose is to draw samples from a probability distribution $f(x)$ which cannot be sampled directly. Initially presented in 1953,[96] it was first used to determine thermodynamically probable configurations of particles which were part of a canonical ensemble. Calculating an ensemble average of some observable in the canonical ensemble involves calculating integrals of the form

$$\langle a \rangle = \int d\mathbf{r}_1 \dots d\mathbf{r}_N a(\mathbf{r}_1, \dots, \mathbf{r}_N) \exp(-\beta U(\mathbf{r}_1, \dots, \mathbf{r}_N)) / \int d\mathbf{r}_1 \dots d\mathbf{r}_N \exp(-\beta U(\mathbf{r}_1, \dots, \mathbf{r}_N)) \tag{2.89}$$

This is done by constructing a Markov chain whose stationary distribution equals the desired probability distribution. In order to guarantee that there is a stationary distribution of the Markov chain, it is sufficient (although not a necessary condition) to show that it fulfills *detailed balance*, i.e. the probability that the system is in state x and transitions to x' is equal to the probability to be in x' and transition to x.

$$p(x)p(x'|x) = p(x')p(x|x') \tag{2.90}$$

If one writes $p(x'|x)$ as a product of a probability to move $q(x, x')$ and an acceptance probability $a(x'|x)$ with

$$a(x|x') = \min\left\{1, \frac{f(x')}{f(x)} \frac{q(x|x')}{q(x'|x)}\right\} \tag{2.91}$$

it can be shown that detailed balance is fulfilled. [97] Furthermore, if $q(x, x')$ is symmetric, equation 2.91 further simplifies, and the second fraction becomes one. Following from the above discussion, the algorithm to sample from a distribution f becomes:

Given a position x_i

1. Draw a new position in the vicinity of $x_{\text{trial}} = x_i + r\xi$ (where ξ may be a uniform random number between -1 and 1, and r the vicinity radius)
2. Set $x_{i+1} = x_{\text{trial}}$ with probability $\min\left\{1, \frac{f(x_{\text{trial}})}{f(x)}\right\}$
   Otherwise set $x_{i+1} = x_i$

In order to assure that the Markov Chain samples from the stationary distribution, so-called burn-in periods are usually calculated before the actual Monte Carlo run, in which no samples are taken from the Markov Chain.

2.3.3 Random Walk

An example commonly used to illustrate the concept of a random walk is that of a drunk person staggering from street light to street light at night. Because of their intoxicated state, they do not remember which steps they have taken before arriving at the current street light, and thus the direction of each new step toward the next street light only depends on the position they currently have. Surprisingly, this simple model can be used to make predictions of very different real-world problems such as Brownian motion [98, 99] or even the fluctuation of stock prices. [100–102]

In order to arrive at a more formal description of the random walk, we can express the trajectory of the random walker as a sequence of states $\mathbf{x}_i$ which are reached at time t_i. As the process is stochastic, both the positions $\mathbf{x}_i$ and the times t_i are random variables, whose outcome is not deterministic. Therefore, it is only possible to give a probability that the random walker walks a certain trajectory.

$$P(\mathbf{x}_N, t_N; \mathbf{x}_{N-1}, t_{N-1}; ...; \mathbf{x}_1, t_1) \tag{2.92}$$

For the characterization of the random walk, it is helpful to consider conditional probabilities, in other words the probability that the walker reaches state $\mathbf{x}_N$ at time t_N given his previous trajectory. Without any further assumptions, the conditional probability is

$$P(\mathbf{x}_N, t_N | \mathbf{x}_{N-1}, t_{N-1}; ...; \mathbf{x}_1, t_1) = \frac{P(\mathbf{x}_N, t_N; \mathbf{x}_{N-1}, t_{N-1}; ...; \mathbf{x}_1, t_1)}{P(\mathbf{x}_{N-1}, t_{N-1}; ...; \mathbf{x}_1, t_1)} \tag{2.93}$$

In this general form, the conditional probability may well depend on the complete history of the trajectory. The important property of a random walk, however, is the fact that the probability to reach a new state only depends on the current state, also referred to as *Markov property*. This simplifies equation 2.93 to

$$P(\mathbf{x}_N, t_N | \mathbf{x}_{N-1}, t_{N-1}; ...; \mathbf{x}_1, t_1) = P(\mathbf{x}_N, t_N | \mathbf{x}_{N-1}, t_{N-1}) \tag{2.94}$$

Given a state **x**, and using the Markov property, the following expression emerges for the probability to find the system in state **x** at time $t + \Delta t$:

$$P(\mathbf{x}, t + \Delta t) = \sum_{\mathbf{y} \neq \mathbf{x}} P(\mathbf{y}, t) P(\mathbf{x}, t + \Delta t | \mathbf{y}, t) + P(\mathbf{x}, t) P(\mathbf{x}, t + \Delta t | \mathbf{x}, t) \tag{2.95}$$

In this formula, the first summand describes the probability that the system is in some state other than **x** at time t, and transitions to **x** in the next time step. The second summand describes the probability that the system is already in state **x** at time t, and stays there.

The probability that the system transitions from **x** to **x** can also be expressed by its complementary event:

$$P(\mathbf{x}, t + \Delta t | \mathbf{x}, t) = 1 - \sum_{\mathbf{y} \neq \mathbf{x}} P(\mathbf{y}, t + \Delta t | \mathbf{x}, t) \tag{2.96}$$

Inserting this into equation 2.95, and subtracting $P(\mathbf{x}, t)$ from both sides yields

$$P(\mathbf{x}, t + \Delta t) - P(\mathbf{x}, t) = \sum_{\mathbf{y} \neq \mathbf{x}} P(\mathbf{y}, t) P(\mathbf{x}, t + \Delta t | \mathbf{y}, t) - P(\mathbf{x}, t) P(\mathbf{y}, t + \Delta t | \mathbf{x}, t) \tag{2.97}$$

In the last step, the conditional probabilities for a jump within a time interval Δt are expressed as transition rates. This assumes that for small time steps Δt the transfer probabilities grow linearly in time.

$$P(\mathbf{x}, t + \Delta t | \mathbf{y}, t) = \omega(\mathbf{y} \to \mathbf{x}) \Delta t \tag{2.98}$$

Thus, equation 2.97 becomes

$$P(\mathbf{x}, t + \Delta t) - P(\mathbf{x}, t) = \sum_{\mathbf{y} \neq \mathbf{x}} P(\mathbf{y}, t) \omega(\mathbf{y} \to \mathbf{x}) \Delta t - P(\mathbf{x}, t) \omega(\mathbf{x} \to \mathbf{y}) \Delta t \tag{2.99}$$

Dividing by Δt and taking the limit, we arrive at a differential equation for $P(\mathbf{x}, t)$ whose solution tells us the temporal evolution of the probability to find the system in state **x**.

$$\frac{d}{dt} P(\mathbf{x}, t) = \sum_{\mathbf{y} \neq \mathbf{x}} P(\mathbf{y}, t) \omega(\mathbf{y} \to \mathbf{x}) - P(\mathbf{x}, t) \omega(\mathbf{x} \to \mathbf{y}) \tag{2.100}$$

This differential equation is called *Master equation*. It describes the dynamics of the system as a differential equation of probabilities.

If the probabilities become stationary, i.e. they stay constant in time, the time derivative becomes zero

$$\frac{d}{dt}P(\mathbf{x},t) \overset{!}{=} 0 \tag{2.101}$$

In this case, the following equation applies:

$$\sum_{\mathbf{y}\neq\mathbf{x}} P(\mathbf{y},t)\omega(\mathbf{y}\to\mathbf{x}) = \sum_{\mathbf{y}\neq\mathbf{x}} P(\mathbf{x},t)\omega(\mathbf{x}\to\mathbf{y}) \tag{2.102}$$

which is the condition for detailed balance (see section 2.3.2).

2.3.4 Inverse Transform Sampling

The inverse transform sampling is a technique commonly used to generate random numbers with a desired probability density. Given a continuous random variable, assume that a random variable X has the probability distribution function $p_X(x)$. Its cumulative distribution function is defined as

$$F_X(x) = \int_{-\infty}^{x} p(t)dt \tag{2.103}$$

If a random variable Y has a uniform distribution on $[0,1]$, then the random variable $F_X^{-1}(Y)$ has the probability distribution $p_X(x)$. [103]

As an example to illustrate the way the inverse transform sampling works, let us try to simulate a biased dice, whose probability to show a three is three times more likely than for any other number. The upper-left picture in figure 2.2 shows the corresponding probability distribution. In order to draw numbers from this distribution, one could do the following:

1. Stack the bars on top of each other
2. Draw a (uniformly distributed) random number between one and zero
3. Its value then determines the height at which a value is drawn from the stack

In the above example it is easy to see that the number three will be selected more often as it constitutes a larger part of the total stack.

In the case of a continuous variable, the “stacking” of discrete probabilities has to be replaced by the cumulative distribution function. A value from the desired probability distribution $p_X(x)$ is then sampled by drawing a random number r between zero and one, and determining the x-value of the intersection between r and the cumulative distribution function $F_X(x)$

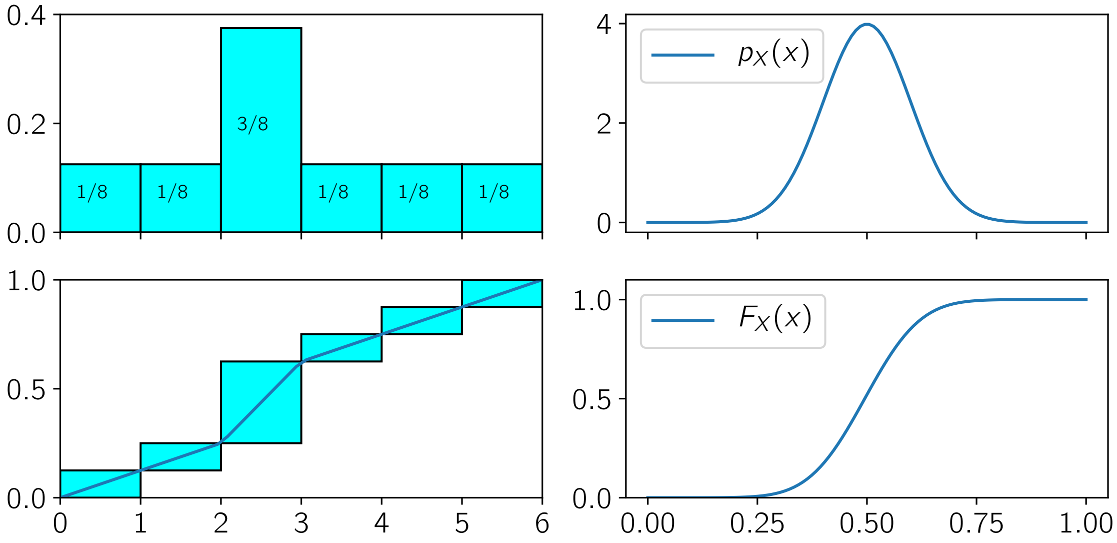

Figure 2.2: Upper-left: Probability distribution of a biased dice. The probability of getting a three is three times more likely than for any other number.
Lower-left: The behavior of the dice can be simulated by choosing a (uniformly) random number between zero and one on the y-axis and determining with which bar it intersects.
In the case of a density of a continuous random variable (upper-right), the sum over probabilities needs to be replaced by an integral over a probability distribution (lower-right).

2.3.5 Asymmetric Simple Exclusion Process (ASEP)

The Asymmetric Simple Exclusion Process (ASEP) is a simple model which realizes a Markov process of interacting particles moving on a discrete lattice. Each site of the lattice can be occupied by at most one particle. An update procedure moves these particles in discrete time steps over the lattice. Due to its simple set of rules the (one-dimensional) ASEP has been used for the modeling of non-equilibrium systems. [54, 104–107] Applications are, among others, traffic flow simulations [108] and protein synthesis. [109]

In figure 2.3, the example of a one-dimensional chain is shown. In this case, a particle sitting at a site i in the middle of the chain can jump with probability $p_{i,i+1}$ to the right, and with probability $p_{i,i-1}$ to the left.

For the one-dimensional case, a linear chain consisting of L sites, different update procedures exist: [54]

(a) **Random-Sequential Update:** At each time step, an edge $(i, i+1)$ is randomly chosen and updated

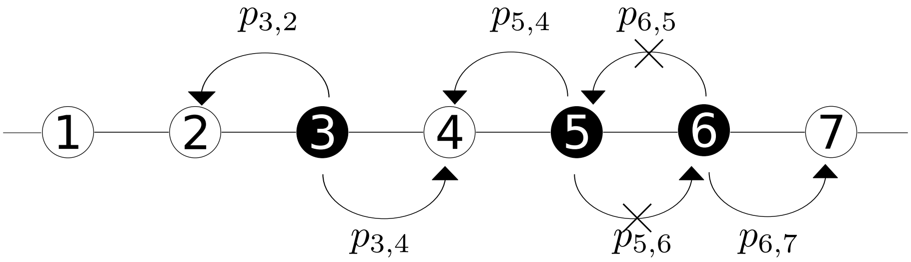

Figure 2.3: Illustration of the Asymmetric Exclusion Process using the example of a one-dimensional chain. Each lattice site (numbered from 1 to 7) can be either occupied or unoccupied. At each step, an edge (i, j) between two sites is chosen, and a particle will jump with probability $p_{i,j}$, only if site i is occupied, and site j free.

(b) **Sublattice-Parallel Update:** In the first step, the condition for a jump is tested at the end points 1 and L. Then, all even pairs $(2i, 2i + 1)$ are updated. Finally all uneven pairs $(i, i + 1)$ are updated.

(c) **Ordered-Sequential Update:** In this update scheme, subsequent lattice sites are updated (either from left to right or from right to left)

2.3.6 Kinetic Monte Carlo

The Kinetic Monte Carlo scheme (also referred to as Stochastic Simulation Algorithm) is a method to determine the temporal evolution of a system by randomly sampling the transitions between possible states of the system.

Starting from a system in state i at time $t = 0$, we want to determine the probability that it is still in the same state at a later time t. Using the master equation

$$\frac{d}{dt}P_i(t) = \sum_{i\neq j} P_j(t)\omega_{j,i} - P_i(t)\omega_{i,j} \tag{2.104}$$

one can see that the probability to stay in state i is given by

$$\frac{dP_i(t)}{dt} = -\sum_j P_i(t)\omega_{ij}(t) \tag{2.105}$$

Integration of this differential equation for $p_i(t)$ gives the probability to still find the system in state i after a time τ has passed:

$$P_i(\tau) = \exp\left(-\int\limits_0^\tau \Omega_i(t')dt'\right) \tag{2.106}$$

with

$$\Omega_i(t) = \sum_j \omega_{ij}(t) \tag{2.107}$$

Equation 2.106 yields a lifetime distribution of state i. In the case of time-independent rates, $\Omega(t)$ becomes a constant, and equation 2.106 simplifies to an exponential distribution. The Kinetic Monte Carlo scheme samples from this distribution in order to determine the waiting time before the next state transition occurs. More specifically, a random number x in the range (0, 1] is drawn, and the waiting time τ is determined, for which equation 2.108 is fulfilled:

$$-\ln(x) = \int\limits_0^\tau \Omega_{ij}(t')dt' \tag{2.108}$$

In the time-independent case the equation can be explicitly solved for τ:

$$\tau = -\frac{\ln(x)}{\Omega} \tag{2.109}$$

whereas in equation 2.108 τ must in general be determined numerically.

Once the lifetime τ has passed, the algorithm determines the target site for the excess charge by drawing another random number x from the interval $[0, \sum\limits_j \omega_{i_0,j}(t)]$ and choosing the second state with index j_0 such that

$$\sum_{j=0}^{j_0-1} \omega_{i_0,j}(t) \leq x < \sum_{j=0}^{j_0} \omega_{i_0,j}(t) \tag{2.110}$$

Note that both the determination of τ and the determination of the target site are just the application of the inverse transform sampling (section 2.3.4).

Algorithmic Description

The coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) algorithm is a stochastic scheme which models proton transfer on mesoscopic time scales by means of a kinetic Monte Carlo like algorithm which is coupled to a molecular dynamics trajectory. By taking into account the temporal change of the atomic positions, the cMD/LMC scheme is able to accurately react to changes of the hydrogen bond network. In this chapter, a detailed explanation of the scheme is given.

3.1 Coupling of Molecular Dynamics and Lattice Monte Carlo Scheme

Given a precomputed Molecular Dynamics (MD) trajectory of a proton conducting system, the cMD/LMC scheme translates donor/acceptor pairs found in the trajectory into a lattice of proton hosting sites. Each site can be in either of two states: occupied by a proton, or unoccupied. Within the Monte Carlo scheme, the protons are then moved stochastically between lattice sites according to transfer rates which

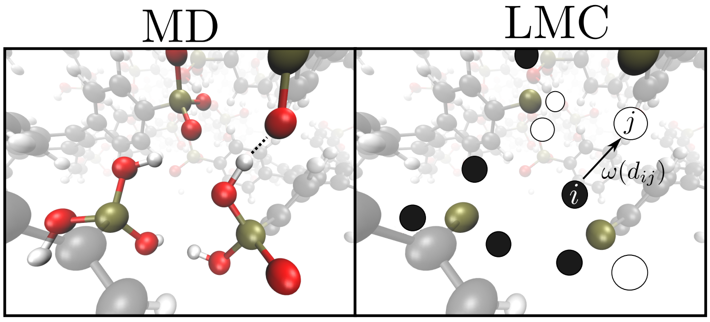

Figure 3.1: Illustration of the translation from an atomic structure to a discrete lattice of proton hosting sites. The proton transfer rate $\omega(d_{ij})$ between two sites i and j determines the dynamics of the system.

are defined for each pair of sites. In order to compute chemically realistic proton diffusion, acurate values for the transfer rates need to be found. The ansatz in this work consists of two steps: Firstly, the proton transfer rate is expressed as a function of geometric parameters with respect to donor and acceptor. In the most basic approach this can be a simple distance dependency, but further parameters such as hydrogen bond angles can be taken into account as well. Secondly, the temporal change of the molecular structure obtained from MD is taken into account for the construction of the transfer rates. More specifically, for a number of N donor/acceptor atoms, a distance matrix is determined with matrix elements d_{ij} denoting the distance between atom i and atom j.

$$d_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \tag{3.1}$$

Given that $\omega(d)$ is the functional form of the proton transfer rates with respect to the donor/acceptor distance, a transfer rate matrix can be constructed via

$$\omega_{ij} = \omega(d_{ij}) \tag{3.2}$$

Over the course of the MD trajectory, the pair-wise distances $d_{ij}(t)$ vary with time, and thus the transfer rates are also time-dependent.

The determination of the rate function is discussed in detail in the appended papers. In this thesis, it was determined via statistical analysis of *ab initio* MD trajectories, and from energy surface scans of the proton transfer reaction path.

3.2 Coupled MD/LMC Scheme using the Asymmetric Exclusion Process

In the case of proton diffusion in a neutral system, the cMD/LMC scheme uses a variation of the Asymmetric Simple Exclusion Process (ASEP) (see section 2.3.5) for the stochastic propagation of the protons. This scheme uses a discrete lattice, where each lattice site represents a donor/acceptor atom. Each lattice site can be either occupied or unoccupied, and the transfer rate matrix ω_{ij} contains the proton transfer rate between each pair of lattice sites i and j.

Within the ASEP scheme, the proton transfer rates ω_{ij} between lattice sites need to be converted to jump probabilities p_{ij} via

$$p_{ij} = \omega_{ij}\Delta t \tag{3.3}$$

At each Monte Carlo step, a connection between an oxygen O_i and a second oxygen O_j is chosen randomly. If site i is occupied by a proton, and site j is empty, the proton is moved to site j with probability p_{ij}.

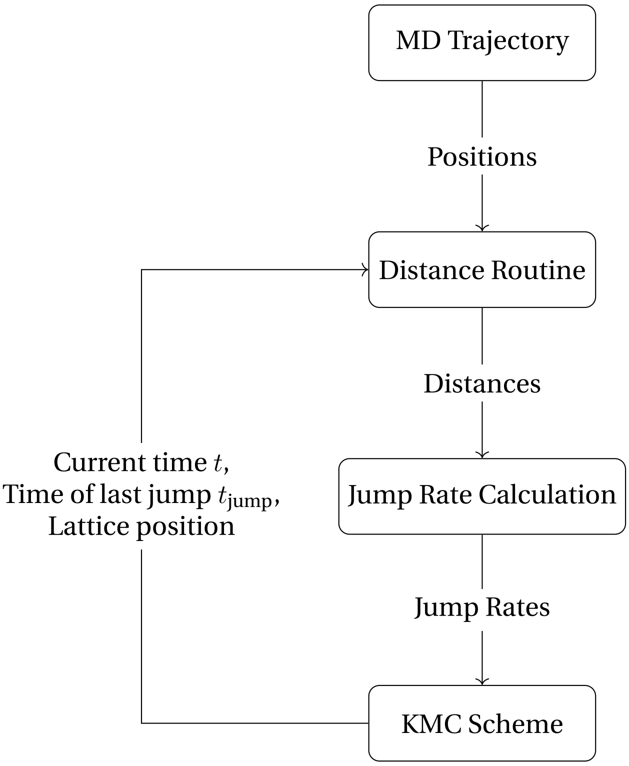
MD Trajectory
Positions
Distance Routine
Distances
Jump Rate Calculation
Jump Rates
KMC Scheme
Current time t,
Time of last jump t_{jump},
Lattice position

Figure 3.2: Flowchart diagram of the feedback loop within the cMD/LMC scheme for excess charge diffusion.

Given a total number of lattice sites N, there are a total number of $N(N-1)$ connections between lattice sites. Thus, at each Monte Carlo step a particular connection (i,j) has the probability $\frac{1}{N(N-1)}$ to be chosen. Therefore, $N(N-1)$ Monte Carlo steps (also referred to as one sweep) are associated with the MD time step Δt. When coupled with an MD trajectory, Δt in equation 3.3 is chosen such that it corresponds to the MD time between updates of the atomic positions.

3.3 Excess Charge Kinetic Monte Carlo Algorithm

The cMD/LMC scheme for excess charge diffusion represents a continuation of the algorithm presented in section 3.2 with the aim to extend its applicability to aqueous media. In order to adequately describe such highly fluctuating systems, the cMD/LMC scheme requires several modifications.

As water shows a strong structural response to the presence of an excess charge, the cMD/LMC scheme needs to be aware of these effects. By design, the molecular dynamics trajectory runs independently from the LMC scheme. This prevents the implementation of a direct structural response of the MD structure to the excess charge state within the LMC scheme. However, an *ab initio* analysis of the first solvation shell of H_3O^+ shows that the H_3O^+-H_2O distance distribution can be obtained by means of a straightforward transformation of the neutral H_2O-H_2O distance distribution. Thus, a distance-rescaling routine is implemented into the cMD/LMC scheme, which receives atomic positions of neutral bulk water as input, and modifies the first solvation shell around the LMC excess charge accordingly. In order to take into account the relaxation time of the first solvation shell, an additional parameter controls the speed of the structural relaxation. Finally, in order to take the hydrogen bond topology of H_3O^+ within water into account, the number of lattice connections from the excess charge donor to neighboring excess charge acceptors is restricted to the three closest neighboring molecules.

Figure 3.2 illustrates the concept. The distance generating routine receives the current LMC state consisting of the current time t, the time of the last excess charge transfer t_{jump}, and the lattice position of the excess charge. Based on this input, the corresponding donor-acceptor distances are calculated and passed on to the jump rate generator. The jump rate generator in turn passes the jump rates to the LMC scheme, which then determines the time of the next excess charge transfer. The determination of the donor-acceptor distances, on the other hand, involves several steps within the distance generating routine. Figure 3.3 illustrates these steps. Within a specified time interval after an excess charge transfer has occurred, interpolated values for the distances are computed which lie between the unrescaled/neutral and relaxed H_3O^+-H_2O distances.

$$d^{\text{OO}}(t) = d^{\text{OO}}_{\text{neutral}}(t) + \frac{t - t_0}{t_{\text{relax}}} \left[d^{\text{OO}}_{\text{hydronium}}(t) - d^{\text{OO}}_{\text{neutral}}(t) \right] \tag{3.4}$$

In contrast to the ASEP scheme, the time is not incremented in equal steps, but is a random variable whose distribution depends on the proton transfer rates (see section 2.3.6).

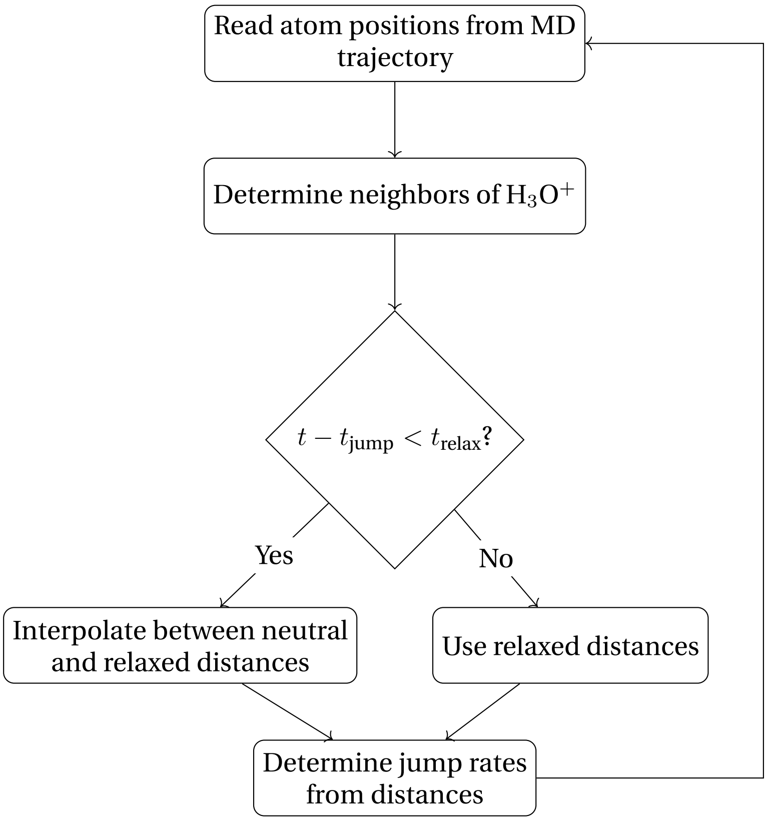
Read atom positions from MD trajectory
Determine neighbors of H_3O^+
$t - t_{\text{jump}} < t_{\text{relax}}$?
Yes
No
Interpolate between neutral and relaxed distances
Use relaxed distances
Determine jump rates from distances

Figure 3.3: Flowchart of the distance scheme. After receiving the neutral H_2O-H_2O distances from the molecular dynamics trajectory, the distance scheme decides whether to interpolate between neutral and relaxed distances, thus imitating the relaxation process, or to use the relaxed distances.

Publications

4.1 Publication Overview

For the design of fuel cell membrane materials, it is essential to examine proton transport processes on an atomistic level, as the overall diffusivity is determined both by structural features of the molecular assembly and the proton hopping dynamics. While *ab initio* methods are able to offer a high level of accuracy with respect to molecular structure and proton dynamics, this comes at the price of high computational demands for the necessary molecular dynamics simulations. The cMD/LMC scheme developed in this thesis strives to determine proton dynamics at moderate computational costs while still maintaining a high accuracy by parameterizing *ab initio* results.

4.1.1 Proton Transport in Solid Acids

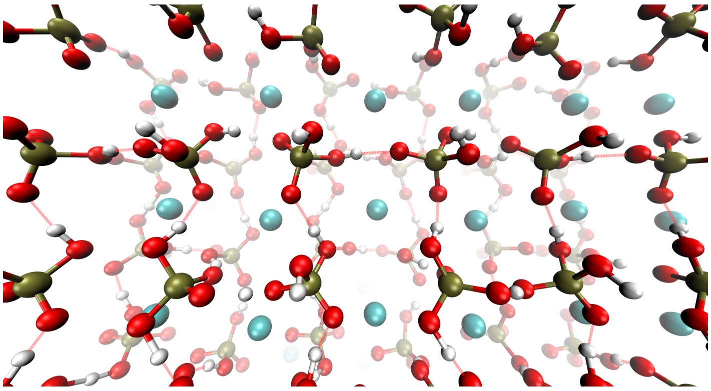

Figure 4.1: Cubic phase of the solid acid CsH_2PO_4

In the scope of this project, two solid acid materials, CsH_2PO_4 and $CsHSO_4$ were investigated by means of *ab initio* MD simulations and using the cMD/LMC scheme.

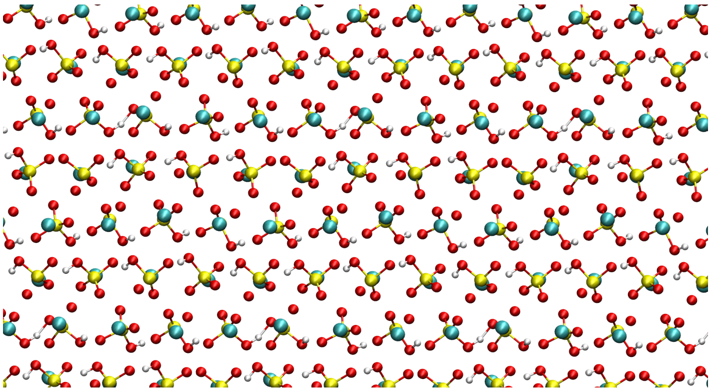

Figure 4.2: Tetragonal phase of the solid acid $CsHSO_4$

In particular the superprotonic phase transition of CsH_2PO_4, during which the proton conductivity increases by several orders of magnitude, was investigated. Using the cMD/LMC scheme the proton conductivity could be decomposed into contributions of proton hopping and structural reorientation. For CsH_2PO_4, it turns out that the decisive factor for the drastic change in proton conductivity between the low-temperature phase and high-temperature phase is due to an increased anion rotation, which promotes long-range proton transport. On the other hand, a chemical substitution (P → S) leads to a decrease in proton conductivity which can mainly be attributed to a decline of the proton transfer rates.

4.1.2 Water-Free Proton Transport

An important topic with respect to fuel cell materials are chemical compounds which allow proton transport at low humidity conditions. This way, operation temperatures above 100 °C can be realized without the risk of dehydration of the fuel cell membrane. On a microscopic scale, the connectivity of the hydrogen bond network plays a crucial role, as it needs to compensate for the lack of H_2O as a carrier molecule.

In the context of this work, a supramolecular assembly of disk-shaped molecules functionalized with phosphonic acid groups (Hexakis(*p*-phosphonatophenyl)benzene (*p*-6PA-HPB)) was investigated at various temperatures. As figure 4.3 shows, the disk arrangement forms channels of hydrogen bonded phosphonic acid groups along which proton transport is possible. Investigations of the proton conductivity of *p*-6PA-HPB by means of the cMD/LMC scheme confirmed that the anisotropic arrangement of the molecules leads to a preference direction of the proton conductivity. Furthermore, it was investigated how structural reorientation, specifically the rotation of the phosphonic groups influence the overall proton conductivity. For this, the cMD/LMC scheme was enhanced by an additional stochastic “rotation event” in

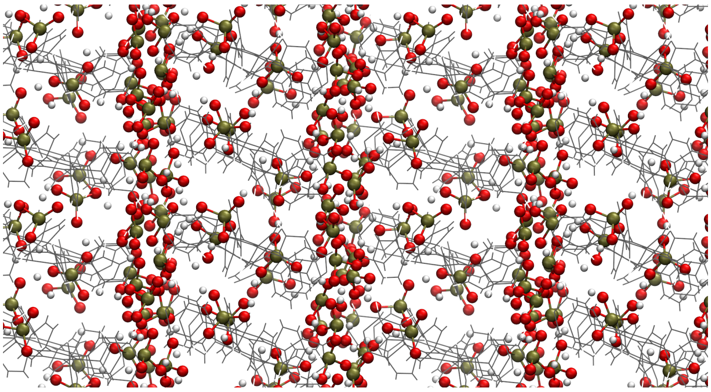

Figure 4.3: The supramolecular assembly of the phosphonic acid-functionalized disks leads to a columnar structure. Channels of hydrogen bonded phosphonic acid groups between the columns enable proton transport.

which protons residing at the same phosphonic group are swapped in cyclic order, thus emulating a partial rotation.

4.1.3 Water-Mediated Proton Transport

Nafion-based proton exchange membranes continue to be used both in research and commercially. Nafion, which consists of perfluorinated vinyl ether groups, has been shown to yield high proton conductivity under hydrated conditions. In order to understand the proton conduction properties of such exchange membranes, it is therefore essential to investigate proton transport processes in bulk water.

In the context of this work, the cMD/LMC scheme was developed further to allow the calculation of excess charge diffusivity in water. As water is a highly unordered system, the cMD/LMC scheme needs to be modified in order to take into account structural relaxation effects. Based on an *ab initio* trajectory of an excess charge in water, the influence of an excess charge on the first solvation shell of its host H_3O^+ molecule is investigated, and incorporated into the cMD/LMC scheme. As the presence of an excess charge influences the distance of the H_3O^+ ion to the molecules in its proximity, the cMD/LMC scheme needs to take into account this dielectric relaxation as well. For this, a relaxation scheme is constructed such that the donor-acceptor distances between neutral water molecules are mapped to typical distances found between H_3O^+ and H_2O. An additional parameter t_{relax} controls the time scale of the relaxation process within the cMD/LMC scheme. With these new adjustments, the cMD/LMC scheme is able to yields excess charge dynamics which are in good agreement with *ab initio* simulations.

A coupled Molecular Dynamics / kinetic Monte Carlo Approach for Protonation Dynamics in Extended Systems

Gabriel Kabbe, Christoph Wehmeyer, Daniel Sebastiani

J. Chem. Theory. Comput., **2014**, 10, 4221–4228

`http://dx.doi.org/10.1021/ct500482k`

Abstract: We propose a multi-scale simulation scheme that combines first-principles Molecular Dynamics (MD) and kinetic Monte Carlo (kMC) simulations to describe ion transport processes. On the one hand, the molecular dynamics trajectory provides an accurate atomistic structure and its temporal evolution, and on the other hand, the Monte Carlo part models the long-time motion of the acidic protons. Our hybrid approach defines a coupling scheme between the MD and kMC simulations that allows the kMC topology to adapt continuously to the propagating atomistic microstructure of the system. On the example of a fuel cell membrane material, we validate our model by comparing its results with those of the pure MD simulation. We show that the hybrid scheme with an evolving topology results in a better description of proton diffusion than a conventional approach with a static kMC transfer rate matrix. Furthermore, we show that our approach can incorporate additional dynamical features such as the coupling of the rotation of a side group in the molecular building blocks. In the present implementation, we focus on ion conduction, but it is straightforward to generalize our approach to other transport phenomena such as electronic conduction or spin diffusion.

Toward Realistic Transfer Rates within the Coupled Molecular Dynamics/Lattice Monte Carlo Approach

Gabriel Kabbe, Christian Dreßler, Daniel Sebastiani

J. Phys. Chem. C, **2016**, 120, 19905–19912

`http://dx.doi.org/10.1021/acs.jpcc.6b05821`

Abstract: We refine our recently developed coupled Molecular Dynamics / Lattice Monte Carlo scheme for the simulation of protonation dynamics in complex hydrogen-bonded solids in view of improving the resulting transport processes. The distance dependency of the proton jump rate between lattice sites and its dependence on additional geometric criteria (bond angles) are derived in a systematic and consistent way. The distance dependency follows an accurate potential energy surface scan from quantum chemical calculations. The novel geometric criterion takes into account that proton hopping occurs almost exclusively along linear hydrogen bonds. We illustrate the capabilities and the versatility of our scheme on the example of two chemically quite different condensed phase systems: a crystalline solid acid compound and a liquid crystal. Surprisingly, we find that our coupled Molecular Dynamics / Lattice Monte Carlo scheme yields converged mobility parameters even when based on underlying *ab initio* molecular dynamics trajectories which themselves are not fully converged. Our method yields more accurate values for the mean square displacement, the OH bond autocorrelation function and the proton jump frequencies in agreement with both reference *ab initio* molecular dynamics simulations and experimental values.

Proton Conductivity in Hydrogen Phosphate/Sulfates from a Coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) Approach

Christian Dreßler, Gabriel Kabbe, Daniel Sebastiani

J. Phys. Chem. C, **2016**, 120, 19913–19922

`http://dx.doi.org/10.1021/acs.jpcc.6b05822`

Abstract: The ionic conductivity of solid acids of CsH_nXO_4 type (X=P, S; n=2,1) varies upon chemical substitution P↔S and between different crystal structures ($T_c = 503$ K for CsH_2PO_4). We apply a recently developed coupled Molecular Dynamics/Lattice Monte-Carlo simulation approach to explain both the phosphate / sulfate and temperature / phase-related variations of the proton conductivity on a molecular level. Our simulation method elucidates the relative importance of the two key components of the Grotthuss-type proton conduction mechanism, proton hopping and structural reorientation, as a function of the chemical/thermodynamical conditions. We find that the chemical substitution leads to a substantial change in the proton hopping rate, which however results only in a modest variation of the proton diffusivity. The variation of the temperature of CsH_2PO_4 results in a significant response of the anion rotation frequency, which turns out to be the rate-limiting process for proton conduction. In particular, the dramatic conductivity response to the phase transition can be explained by a large change of the rotation frequency. In contrast to this, our simulations show that for $CsHSO_4$, the local proton hopping rate is the decisive mechanism which controls long-range proton transport. These findings illustrate that the actually rate limiting factor of proton conduction in such solid acids is

clearly system-dependent.Our simulated results for the proton conductivities agree almost quantitatively with experimental values, providing further evidence for the high predictive capabilities of our scale-bridging cMD/LMC simulation approach.

Proton Mobility in Aqueous Systems: Combining ab initio Accuracy with Millisecond Timescales

Gabriel Kabbe, Christian Dreßler, Daniel Sebastiani

Phys Chem Chem Phys, **2017**, 19, 28604–28609

`http://dx.doi.org/10.1039/C7CP05632J`

Abstract: We present a multiscale simulation of proton transport in liquid water, combining *ab initio* molecular dynamics simulations with force- field ensemble averaging and kinetic Monte-Carlo simulations. This unique Ansatz allows for *ab initio* accuracy incorporating the femtosecond dielectric relaxation dynamics of the aqueous hydrogen bonding network, and bridges the time-scale gap towards the explicit simulation of millisecond diffusion dynamics.

Insight from Atomistic Simulations of Protonation Dynamics at the Nanoscale

Christian Dreßler, Gabriel Kabbe, Daniel Sebastiani

Fuel Cells, **2016**, 16, 682–694

`http://dx.doi.org/10.1002/fuce.201500217`

Abstract: In this paper, we give an overview of the role of Molecular Dynamics (MD) simulations in the field of proton exchange membranes. It focuses on structural and dynamical findings regarding the topology of hydrogen bond networks and proton diffusion. On the one hand, findings about water-containing PEM fuel cell materials, such as Nafion and liquid containing pore materials are discussed. On the other hand, proton conduction in water-free systems is elucidated. Here, the focus lies on phosphonic acids, which possess a rigid structure, and polymer structures such as Poly(vinyl phosphonic acid) and Hexakis-(p-phosphonatophenyl)benzene.

4.6. 81-94

Academic Curriculum Vitae

Gabriel Kabbe

Adresse Am Steintor 17,
06112 Halle
Geburtstag 15. September 1986

Telefon 01522 316 32 18
E-Mail gabriel.kabbe@mail.de

Ausbildung

2013-Heute Promotion Theoretische Chemie, *Martin-Luther Universität*, Halle

Thema: *Development of a coupled Molecular Dynamics / Lattice Monte Carlo Scheme*

2010-2013 Master of Science, *Freie Universität Berlin*

Masterarbeit: *Simulating Proton Transfer in Hexakis(p-phosphonatophenyl)benzene*

2007-2010 Bachelor of Science, *Freie Universität Berlin*

Bachelorarbeit: *Optische Rotationsdispersion und zirkularer Dichroismus an chiralen Lösungen*

1997-2006 Abitur, *Carl-Benz Gynmnasium*, Ladenburg

Berufliche Erfahrung

2013-Heute Martin-Luther Universität Halle-Wittenberg
Lehre

- Vorlesung Physikalische Chemie (Tutor/Dozent)
- Vorlesung Theoretische Chemie (Tutor/Dozent)
- Vorlesung Python für Naturwissenschaftler (Dozent)

2013-Heute Martin-Luther Universität Halle-Wittenberg
System-Administration

- Linux-Server Administration
- Systemweite Linux-Installationen mit FAI
- Backup-Verwaltung
- Virtualisierung mit Docker, VirtualBox und Vagrant

2017 Freiberuflich für die Berlin-Brandenburgische Akademie der Wissenschaften
Software-Entwicklung

Konvertierung historischer Dokumente von LaTeX nach XML

Programmier- und IT-Kenntnisse

- **Programmiersprachen**

  Python
  C, C++
  Haskell
  Java

- **Data Science**

  NumPy, SciPy - Numerik
  SymPy - Symbolisches Rechnen
  Pandas - Datenanalyse
  Tensorflow, scikit-learn - Machine Learning

- **Verschiedenes**

  Docker, VirtualBox, Vagrant - Virtualisierung
  CherryPy - Web Framework in Python
  git - Versionskontrolle

Sprachen

- **Englisch**
- **Spanisch**

Hobbies

- **Klavier, Gitarre**
- **Reisen**
- **Kraftsport, Fahrrad**

Conclusion

The central objective of this thesis project was the development, implementation and benchmarking of a coupled Molecular Dynamics/Lattice Monte Carlo (cMD/LMC) scheme for the efficient simulation of ion diffusion in condensed phase compounds. The natural subsequent objective was the application of the cMD/LMC scheme to fuel cell membrane materials.

In the first part of this thesis, the focus lies on the development of the cMD/LMC scheme with the aim to simulate proton transfer on extended time scales while still maintaining structural and dynamical accuracy as provided by *ab initio* simulations. In order to capture the dynamical features of the proton hopping mechanism, a parameterized proton jump rate function is determined and fitted to *ab initio* data. In combination with trajectories of the molecular structure, it is possible to determine a time-dependent proton transfer rate matrix, which is used as an input to the Lattice Monte Carlo scheme. The predictive accuracy of the cMD/LMC scheme with regard to proton dynamics is examined using the example of Hexakis(*p*-phosphonatophenyl)benzene (*p*-6PA-HPB) , a disk-shaped molecule functionalized with six phosphonic acid groups. Its supramolecular assembly has a columnar structure which forms an intercolumnar hydrogen bond network. Proton mean squared displacement and the temporal evolution of covalent OH bonds are analyzed and shown to be in good agreement with *ab initio* results. The activation energy of a proton transfer is determined as a function of the donor-acceptor distance by means of a temperature analysis of the proton jump rate. Finally, directed proton flux through several layers of *p*-6PA-HPB is simulated by connecting the start and end points of the assembly to an artificial proton reservoir and sink. In this way, it is possible to determine the anisotropy of the proton conductivity tensor for such compounds.

The second part represents a refinement of the cMD/LMC method in view of enhancing the flexibility of the Monte-Carlo transfer rate topology in order to achieve a better transferability of the approach to different chemical systems. First, the accuracy of the proton transfer rate is improved by means of quantum chemical calculations. Furthermore, the correlation between proton jump rate and the linearity of the hydrogen bonds is included. the improved generality and accuracy of the cMD/LMC scheme is demonstrated with respect to proton mean squared displacement and OH bond autocorrelation on the example of a liquid crystalline proton conductor (*p*-6PA-HPB) and a crystalline solid acid compound (CsH_2PO_4). Further, the convergence of the dynamical properties calculated within the LMC scheme corresponds to the convergence of the structural subvolume of the phase space sampled via *ab initio* Molecular Dynamics (MD) .

The last part of this thesis constitutes a proof of concept for the applicability of the cMD/LMC in highly fluctuating systems. On the example of an excess proton in water, the cMD/LMC scheme is further enhanced to realistically describe the complex proton transfer mechanisms caused by the strongly fluctuating hydrogen bonds and the dielectrical relaxation effects in water. The enhancement consists of two parts: firstly, the LMC scheme needs to be able to mimic the shortening of the hydrogen bonds caused by the electrostatic effects between the positively charged H_3O^+ and its first solvation shell. Secondly, the time scale of this dielectrical relaxation needs to be accounted for in order to realistically model the response of the hydrogen network after an excess charge transfer. The calculations here show that both effects are represented quantitatively.

In conclusion, the cMD/LMC scheme has been shown to be an effective tool for the simulation of proton diffusion on mesoscopic time scales. The combination of molecular dynamics with a kinetic Monte Carlo scheme allows accurate calculations with low computational effort. This makes it an attractive tool for postprocessing of existing *ab initio* trajectories in cases where the heavy atom structure has converged, but the motion of the light hydrogen atoms has not been sufficiently sampled yet.

In the near future, the cMD/LMC scheme will be further extended and refined. One natural extension is the description of OH^- diffusion. It will be interesting to see whether the diffusion mechanisms of a solvated hydroxide ion can be approximated in a similar way as for hydronium ions. Furthermore, Lithium ion transport, which is currently being investigated by means of molecular dynamics simulations, might be another field of applications. The author is therefore optimistic that the cMD/LMC scheme will be a useful tool for future studies.

Acknowledgements

An dieser Stelle möchte ich meinem Betreuer Daniel Sebastiani herzlich danken, der während meiner gesamten Zeit als Doktorand immer ein offenes Ohr für Probleme jeder Art hatte und mir mit hilfreichen Ratschlägen zur Seite stand. Außerdem möchte ich besonders meinem Kollegen Christian Dreßler danken, der regelmäßig für angeregte wissenschaftliche (sowie nicht wissenschaftliche) Diskussionen sorgte, und mich mit großem Teamgeist und chemischem Know-How während der Promotion unterstützte. Ich danke auch Tobias Watermann und Arne Scherrer, die dem täglichen Kampf mit der IT stets gelassen ins Auge sahen, und mit denen die gemeinsame Systemadministration gleich viel entspannter verlief.

Im privaten Kreis danke ich vor allem meiner Freundin Hannah, die mich in den letzten Monaten seelisch und moralisch unterstützt hat. Außerdem möchte ich mich bei meinen Eltern für die liebevolle Unterstützung bedanken.

References

[1] B. Hille, vol. 507, 2001.

[2] M. H. Stowell, vol. 276, pp. 812–816, Jan. 1997.

[3] D. M. Kramer, T. J. Avenson, and G. E. Edwards, *Trends Plant Sci*, vol. 9, pp. 349–357, Jul. 2004.

[4] S. J. Paddison and R. Paul, vol. 4, pp. 1158–1163, Mar. 2002.

[5] T. E. Decoursey, *Physiol Rev*, vol. 83, pp. 475–579, Apr. 2003.

[6] G. Bekçioğlu, C. Allolio, and D. Sebastiani, *J Phys Chem B*, vol. 119, pp. 4053–4060, Mar. 2015.

[7] H. Lapid, N. Agmon, M. K. Petersen, and G. A. Voth, *J Chem Phys*, vol. 122, p. 14 506, Jan. 2005.

[8] S. Meiboom, *J Chem Phys*, vol. 34, pp. 375–388, Feb. 1961.

[9] O. F. Mohammed, D. Pines, J. Dreyer, E. Pines, and E. T. J. Nibbering, *Science*, vol. 310, pp. 83–86, Oct. 2005.

[10] M. Rini, B.-Z. Magnes, E. Pines, and E. T. J. Nibbering, *Science*, vol. 301, pp. 349–352, Jul. 2003.

[11] H. Dong, G. Fiorin, W. F. DeGrado, and M. L. Klein, *J Phys Chem B*, vol. 118, pp. 12 644–12 651, Nov. 2014.

[12] G. Geneste, A. Ottochian, J. Hermet, and G. Dezanneau, *Phys Chem Chem Phys*, vol. 17, pp. 19 104–19 118, Jul. 2015.

[13] A. Roudgar, S. P. Narasimachary, and M. Eikerling, vol. 457, pp. 337–341, May 2008.

[14] J. Lobaugh and G. A. Voth, vol. 104, pp. 2056–2069, Feb. 1996.

[15] J. Niskanen, C. J. Sahle, I. Juurinen, J. Koskelo, S. Lehtola, R. Verbeni, H. Müller, M. Hakala, and S. Huotari, *J Phys Chem B*, vol. 119, pp. 11 732–11 739, Sep. 2015.

[16] R. Vuilleumier and D. Borgis, vol. 284, pp. 71–77, Feb. 1998.

[17] D. E. Sagnella and M. E. Tuckerman, vol. 108, pp. 2073–2083, Feb. 1998.

[18] G. A. Voth, *Acc Chem Res*, vol. 39, pp. 143–150, Feb. 2006.

[19] S. Walbran and A. A. Kornyshev, vol. 114, pp. 10 039–10 048, Jun. 2001.

[20] W. L. Cavalcanti, D. F. Portaluppi, and J.-O. Joswig, *J Chem Phys*, vol. 133, p. 104 703, Sep. 2010.

[21] J.-O. Joswig and G. Seifert, *J Phys Chem B*, vol. 113, pp. 8475–8480, Jun. 2009.

[22] S. Lammers, S. Lutz, and M. Meuwly, *J Comput Chem*, vol. 29, pp. 1048–1063, May 2008.

[23] D. Marx, *ChemPhysChem*, vol. 7, pp. 1848–1870, Sep. 2006.

[24] B. Prod'hom, D. Pietrobon, and P. Hess, *Nature*, vol. 329, pp. 243–246, Jan. 1987.

[25] K. Chen, J. Hirst, R. Camba, C. A. Bonagura, C. D. Stout, B. K. Burgess, and F. A. Armstrong, *Nature*, vol. 405, pp. 814–817, Jun. 2000.

[26] C. Fang, R. R. Frontiera, R. Tran, and R. A. Mathies, *Nature*, vol. 462, pp. 200–204, Nov. 2009.

[27] C. A. Wraight, *Biochim Biophys Acta*, vol. 1757, pp. 886–912, Aug. 2006.

[28] C. Wang and S. J. Paddison, *Phys. Chem. Chem. Phys.*, vol. 12, pp. 970–981, Jan. 2010.

[29] S. J. Paddison, *Annu. Rev. Mater. Res.*, vol. 33, pp. 289–319, Aug. 2003.

[30] R. Devanathan, N. Idupulapati, M. D. Baer, C. J. Mundy, and M. Dupuis, *J. Phys. Chem. B*, vol. 117, pp. 16 522–16 529, Dec. 2013.

[31] B. F. Habenicht, S. J. Paddison, and M. E. Tuckerman, *Phys. Chem. Chem. Phys.*, vol. 12, pp. 8728–8732, Aug. 2010.

[32] X. Ke and I. Tanaka, *Phys. Rev. B*, vol. 69, p. 165 114, Apr. 2004.

[33] X. Ke and I. Tanaka, *Solid State Ionics*, vol. 172, pp. 145–148, Jan. 2004.

[34] X. Ke and I. Tanaka, *Phys. Rev. B*, vol. 69, p. 165 114, Apr. 2004.

[35] X. Ke and I. Tanaka, *Solid State Ionics*, vol. 172, pp. 145–148, Jan. 2004.

[36] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, *Nature*, vol. 397, pp. 601–604, Feb. 1999.

[37] K. A. Mauritz and R. B. Moore, vol. 104, pp. 4535–4586, Jan. 2004.

[38] M. Schuster, T. Rager, A. Noda, K. D. Kreuer, and J. Maier, vol. 5, pp. 355–365, Jan. 2005.

[39] A. Schechter, vol. 147, pp. 181–187, Jan. 2002.

[40] L. Jiménez-García, A. Kaltbeitzel, W. Pisula, J. S. Gutmann, M. Klapper, and K. Müllen, *Angew. Chem. Int. Ed. Engl.*, vol. 48, pp. 9951–9953, Jan. 2009.

[41] L. Jiménez-García, A. Kaltbeitzel, V. Enkelmann, J. S. Gutmann, M. Klapper, and K. Müllen, *Adv. Funct. Mater.*, vol. 21, pp. 2216–2224, Jan. 2011.

[42] D. A. Boysen, T. Uda, C. R. I. Chisholm, and S. M. Haile, *Science*, vol. 303, pp. 68–70, Jan. 2004.

[43] S. M. Haile, D. A. Boysen, C. R. Chisholm, and R. B. Merle, *Nature*, vol. 410, pp. 910–913, Apr. 2001.

[44] B. Merinov, *Solid State Ionics*, vol. 213, pp. 72–75, Jan. 2012.

[45] K. Yamada, *Solid State Ionics*, vol. 175, pp. 557–562, Jan. 2004.

[46] E. Ortiz, R. A. Vargas, and B.-E. Mellander, *J. Chem. Phys.*, vol. 110, pp. 4847–4853, Mar. 1999.

[47] C. R. I. Chisholm, Y. H. Jang, S. M. Haile, and W. A. Goddard, *Phys. Rev. B*, vol. 72, p. 134 103, Oct. 2005.

[48] H.-S. Lee and M. E. Tuckerman, *J. Phys. Chem. C*, vol. 112, pp. 9917–9930, Jan. 2008.

[49] K.-D. Kreuer, A. Rabenau, and W. Weppner, vol. 21, pp. 208–209, Jan. 1982.

[50] N. Agmon, *Chem. Phys. Lett.*, vol. 244, pp. 456–462, Oct. 1995.

[51] A. D. MacKerell, *J. Comput. Chem.*, vol. 25, pp. 1584–1604, 2004.

[52] D. T. Gillespie, *J. Phys. Chem.*, vol. 81, pp. 2340–2361, Jan. 1977.

[53] Y. Cao, H. Li, and L. Petzold, *J Chem Phys*, vol. 121, pp. 4059–4067, Sep. 2004.

[54] N. Rajewsky, L. Santen, A. Schadschneider, and M. Schreckenberg, vol. 92, p. 151, Jul. 1998.

[55] A. Chatterjee and D. G. Vlachos, *Journal of Computer-Aided Materials Design, Volume 14, Issue 2, pp.253-308*, vol. 14, p. 253, Jul. 2007.

[56] G. Henkelman and H. Jónsson, vol. 115, pp. 9657–9666, Dec. 2001.

[57] C. H. Claassens, J. J. Terblans, M. J. H. Hoffman, and H. C. Swart, vol. 37, pp. 1021–1026, Jan. 2005.

[58] P. Nath and M. Ranganathan, vol. 606, pp. 1450–1457, Sep. 2012.

[59] E. Antoshchenkova, M. Hayoun, F. Finocchi, and G. Geneste, vol. 606, pp. 605–614, Mar. 2012.

[60] S. Kerisit and K. M. Rosso, *J Chem Phys*, vol. 127, p. 124 706, Sep. 2007.

[61] M. V. Fischetti and S. E. Laux, vol. 38, pp. 9721–9745, Nov. 1988.

[62] C. Groves, vol. 6, p. 3202, Jan. 2013.

[63] K. Reuter and M. Scheffler, vol. 73, p. 045 433, Jan. 2006.

[64] J. Rogal, K. Reuter, and M. Scheffler, vol. 77, p. 155 410, Apr. 2008.

[65] M. Rieger, J. Rogal, and K. Reuter, *Phys Rev Lett*, vol. 100, p. 016 105, Jan. 2008.

[66] E. W. Hansen and M. Neurock, vol. 196, pp. 241–252, Jan. 2000.

[67] D. Mei, M. Neurock, and C. M. Smith, vol. 268, pp. 181–195, Jan. 2009.

[68] D. MEI, P. SHETH, M. NEUROCK, and C. SMITH, vol. 242, p. 1, Jan. 2006.

[69] C. C. Battaile, D. J. Srolovitz, and J. E. Butler, *J. Appl. Phys.*, vol. 82, pp. 6293–6300, Dec. 1997.

[70] J. M. Pomeroy, J. Jacobsen, C. C. Hill, B. H. Cooper, and J. P. Sethna, vol. 66, p. 235 412, Dec. 2002.

[71] G. Mazaleyrat, A. Estève, L. Jeloaica, and M. Djafari-Rouhani, vol. 33, pp. 74–82, Jan. 2005.

[72] M. Rak, M. Izdebski, and A. Brozi, vol. 138, pp. 250–263, Aug. 2001.

[73] S. X. M. Boerrigter, G. P. H. Josten, J. van de Streek, F. F. A. Hollander, J. Los, H. M. Cuppen, P. Bennema, and H. Meekes, vol. 108, pp. 5894–5902, Jul. 2004.

[74] P. Zhu and R. W. Smith, vol. 40, pp. 683–692, Jan. 1992.

[75] B. Sadigh, T. J. Lenosky, S. K. Theiss, M.-J. Caturla, T. Diaz de La Rubia, and M. A. Foad, vol. 83, pp. 4341–4344, Nov. 1999.

[76] F. M. Bulnes, V. D. Pereyra, and J. L. Riccardo, vol. 58, pp. 86–92, Jul. 1998.

[77] C. Saravanan and S. M. Auerbach, vol. 107, pp. 8132–8137, Nov. 1997.

[78] M. A. Katsoulakis and D. G. Vlachos, vol. 119, pp. 9412–9427, Nov. 2003.

[79] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson, and K. Schulten, *Structure*, vol. 14, pp. 437–449, Mar. 2006.

[80] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, *J. Comput. Chem.*, vol. 26, p. 1781, 2005.

[81] L. Verlet, vol. 159, p. 98, Jul. 1967.

[82] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, *J Chem Theory Comput*, vol. 4, pp. 435–447, Mar. 2008.

[83] P. H. Hünenberger, in. 2005, pp. 105–149.

[84] S. Nose, vol. 2, SA115–SA119, Dec. 1990.

[85] G. J. Martyna, M. L. Klein, and M. Tuckerman, *J. Chem. Phys.*, vol. 97, pp. 2635–2643, Aug. 1992.

[86] A. Szabo and N. S. Ostlund. 2012.

[87] R. G. Parr, S. R. Gadre, and L. J. Bartolotti, ser. Proceedings of the National Academy of Sciences. 1979, vol. 76, pp. 2522–2526.

[88] R. Latter, vol. 99, pp. 510–519, Jul. 1955.

[89] W. D. Myers and W. J. Swiatecki, vol. 601, pp. 141–167, Feb. 1996.

[90] E. Teller, vol. 34, pp. 627–630, Oct. 1962.

[91] P. Hohenberg and W. Kohn, *Phys. Rev.*, vol. 136, pp. 864–871, Nov. 1964.

[92] M. Levy, *Proc Natl Acad Sci U S A*, vol. 76, pp. 6062–6065, Dec. 1979.

[93] M. Levy, vol. 26, pp. 1200–1208, Sep. 1982.

[94] M. H. Kalos and P. A. Whitlock. 2008.

[95] E. T. Jaynes. 2003.

[96] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, vol. 21, pp. 1087–1092, Jun. 1953.

[97] S. Chib and E. Greenberg, vol. 49, pp. 327–335, Jan. 1995.

[98] B. J. Legg and M. R. Raupach, *Boundary-Layer Meteorology*, vol. 24, p. 3, 1982.

[99] A. Dixit, *Journal of economic dynamics and control*, vol. 15, pp. 657–673, 1991.

[100] M. R. Hassan, *Neurocomputing*, vol. 72, pp. 3439–3446, 2009.

[101] M. R. Hassan and B. Nath, 2005, pp. 192–196.

[102] C. M. Turner, R. Startz, and C. R. Nelson, *Journal of Financial Economics*, vol. 25, p. 3, 1989.

[103] L. Devroye, pp. 260–265, Jan. 1986.

[104] O. Golinelli and K. Mallick, *Journal of Physics A: Mathematical and General*, vol. 39, p. 12 679, 2006.

[105] T. Chou, K. Mallick, and R. K. P. Zia, *Reports on progress in physics*, vol. 74, p. 116 601, 2011.

[106] Y. Hieida, *Journal of the Physical Society of Japan*, vol. 67, pp. 369–372, 1998.

[107] B. Derrida, *Physics Reports*, vol. 301, pp. 65–83, 1998.

[108] M. Schreckenberg, A. Schadschneider, K. Nagel, and N. Ito, vol. 51, pp. 2939–2949, Apr. 1995.

[109] L. B. Shaw, R. K. P. Zia, and K. H. Lee, *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 68, p. 021 910, Aug. 2003.