Predicting Vibrational Spectra of Condensed Phase Systems

HABILITATIONSSCHRIFT

zur Erlangung des akademischen Grades

DR. RER. NAT. HABIL.

vorgelegt der

MATHEMATISCH–NATURWISSENSCHAFTLICH–
TECHNISCHEN FAKULTÄT
der
MARTIN-LUTHER-UNIVERSITÄT
HALLE–WITTENBERG

von

Herrn Dr. rer. nat. Martin Brehm,

geboren am 04.09.1985 in Leipzig.

Gutachter:

1.) Prof. Dr. Daniel Sebastiani

2.) Prof. Dr. Thomas Kühne

3.) Prof. Dr. Johannes Kästner

Halle (Saale), 01.03.2023

Dedicated to free software and open science.

CP2K

ORCA

TRAVIS

bqb dpd
pdp qbq

LAMMPS

SANScript

VMD

Grace

Gnuplot

GNU/Linux

debian

GCC

LaTeX

Contents

1 Introduction **1**

1.0 Structure of this Thesis . . . 1

1.1 Vibrational Spectroscopy . . . 2

1.2 Predicting Vibrational Spectra . . . 4
  1.2.1 The Static–Harmonic Approach . . . 4
  1.2.2 The Time–Correlation Approach . . . 7

1.3 Atomistic Simulation Methods . . . 9

1.4 Computing Electromagnetic Moments . . . 13

1.5 Programs used within this Thesis . . . 16

2 Method Development and Theory **17**

2.1 Voronoi Integration . . . 17
  2.1.1 Voronoi Tessellation . . . 17
  2.1.2 Radical Voronoi Tessellation . . . 19
  2.1.3 Voronoi Integration . . . 22
  2.1.4 Advantages over Wannier Localization . . . 24

2.2 Lossless Compression of Volumetric Data . . . 29
  2.2.1 Reducing Information Entropy via Extrapolation . . . 31
  2.2.2 Entropy Encoding . . . 36
  2.2.3 The bqb File Format . . . 43
  2.2.4 Results . . . 45

2.3 Power Spectrum . . . 49

2.4 Infrared Spectroscopy . . . 52

2.5 Raman Spectroscopy . . . 54

2.6 Classical Approach to Electric Currents . . . 57

2.7 VCD Spectroscopy . . . 60

2.8 ROA Spectroscopy . . . 62

2.9 SFG Spectroscopy . . . 67

2.10 Resonance Raman Spectroscopy . . . 69

2.11 Normal Mode Analysis . . . 74

2.12 Signal Processing Techniques 77
    2.12.1 Fast Correlation 77
    2.12.2 Window Function and Zero Padding 77
    2.12.3 Finite Differences Correction 78
    2.12.4 Improved Sampling via Time Reversibility 79
    2.12.5 Correcting the Verlet Frequency Shift 80
2.13 The TRAVIS Program Package 81

3 Applications **83**
3.1 Infrared Spectroscopy 84
3.2 Raman Spectroscopy 86
3.3 VCD Spectroscopy 91
3.4 ROA Spectroscopy 93
3.5 Resonance Raman Spectroscopy 100
3.6 Normal Mode Analysis 107
3.7 Computational Details 111

4 Conclusions **113**

A Appendix

A.1 The Ergodic Theorem 115
A.2 The Wiener–Khinchin Theorem 118
A.3 The Frequency Shift of the Verlet Integrator 121
A.4 Explicit Reversible Nosé–Hoover Chain Integrator 128
A.5 Approximate Overtones from Classical Oscillators 130

Publication List of Martin Brehm **131**

Bibliography **137**

Author Index **173**

List of Figures **187**

List of Abbreviations **189**

Index **193**

Acknowledgments **199**

Curriculum Vitae of the Author **201**

Statement of Authorship **203**

Chapter 1.

Introduction

1.0. Structure of this Thesis

For the convenience of the reader, the structure of this thesis will be briefly described here.

After a short discussion of existing methods and the state of the art for predicting vibrational spectra in **Chapter 1**, the focus will be put on method development and theory in **Chapter 2**. Some *(but not all)* of the methods described there have been developed within the scope of this thesis. To demonstrate that these approaches are indeed able to predict vibrational spectra of complex liquid systems, some examples are presented in **Chapter 3**. The main part of this thesis ends with some conclusions in **Chapter 4**.

Subsequently, an **Appendix** which contains some mathematical derivations and algorithms can be found from page 115 on.

Two separate **bibliographies** follow. While the first one (starting on page 131) contains all publications to which M.B. contributed, as indicated by the "MB" prefix when citing,[MB1] the second one (beginning on page 137) contains all other publications which are cited in this document without prefix.[1] All references are clickable links which lead directly to the bibliography *(at least if you are reading an electronic version of this thesis)*.

On page 173, an **author index** can be found, which contains an alphabetical list of all authors which contributed to the references cited in this thesis, with a list of corresponding reference numbers *(which are—again—clickable links)* after each author's name.

After a **list of figures**, **list of abbreviations**, and an **index**, the **acknowledgments** follow. Finally, a short **curriculum vitae** of M.B. as well as a **statement of authorship** close this document.

1.1. Vibrational Spectroscopy

Vibrational spectroscopy has been an important field of chemistry and physics for a very long time. As this class of methods is concerned with molecular vibrations *(or phonons in solid state materials)* which are highly sensitive to the chemical environment of a certain molecule, vibrational spectra allow to study the interactions present in the sample in great detail. Due to these characteristics, there exist countless applications of vibrational spectroscopy in science, industry, and medicine.[1–10]

The first **"infrared spectrum"** *(albeit of the earth's atmosphere instead of a prepared sample)* was recorded by J. F. W. Herschel in 1840.[11] The Raman effect was predicted by A. Smekal around 100 years ago,[12] and first experimental applications followed several years later,[13,14] but only the advent of the laser[15,16] in the 1960s led to the final breakthrough of **Raman spectroscopy**. Despite their long history, infrared and Raman spectroscopy are still by far the most important methods of vibrational spectroscopy for practical applications.

The most prominent limitation of Raman-based spectroscopic methods is their low sensitivity. Out of one million incident photons, typically less than one undergoes Raman scattering, so that a very weak signal is obtained. In the 1960s, it was discovered that the intensity of the Raman scattering drastically increases if the incident laser wavelength is close to the energy of an electronic excitation in the sample, which was termed as the resonance Raman effect.[17–21] Apart from the overall increase in intensity, also the intensity ratio of the spectral bands changes – vibrational modes which displace atoms involved in the electronic excitation show larger gains of Raman intensity *(vibronic coupling)*. With the low sensitivity of Raman spectroscopy in mind, the discovery of the resonance Raman effect was an important advance in the field, and the increased signal intensity allowed for higher resolution and shorter acquisition times. Moreover, as the signal increase only occurs for molecules which are electronically excited by the laser, it became possible to measure very dilute solutions – a reasonable choice of laser wavelength amplifies the bands of the solute, but not those of the solvent. **Resonance Raman spectroscopy** was recently applied in the fields of medicine,[22] electrochemistry,[23] and to study drug binding,[24] solar cells,[25] photosynthesis,[26] nanoparticles,[27] and water splitting.[28]

More recently, the chiral variants of infrared and Raman spectroscopy have been introduced experimentally, namely **vibrational circular dichroism (VCD)**[29–32] and **Raman optical activity (ROA)**[33–36] spectroscopy. Based on circularly polarized light, these techniques derive their results as the difference between two infrared and Raman spectra, respectively. For a sample which consists of only non-chiral molecules or a racemate, the VCD and ROA spectra are empty. However, if the sample contains an enantiomeric excess, characteristic signals appear, and the sign of these signals allows for the determination of the absolute configuration of the chiral compound. This is a very important feature, as most other kinds of spectroscopies are *not* able to determine the absolute configuration of molecules.

Another very interesting method is **sum frequency generation (SFG) spectroscopy**. It is a non-linear laser spectroscopy, in which two laser beams *(visible pump and infrared probe)* mix at an interface and generate an output beam with a frequency equal to the sum of the two input frequencies. As such a sum frequency signal is only created in close proximity to the interface, this is a surface-selective spectroscopy, and thus very valuable. This technique has been developed in 1987 by the group of Shen,[37,38] and only a few years later it was used by Guyot–Sionnest to study electronic and vibrational dynamics at surfaces.[39–41] The information contained in a SFG spectrum is in a certain sense complementary to infrared and Raman spectroscopy.[42]

1.2. Predicting Vibrational Spectra

With the rise of computers in the second half of the 20th century, the field of computational chemistry was established, and it became possible to predict vibrational spectra by quantum chemical methods. This was an important step, as these predictions were helpful to complement and better understand experimentally measured spectra. In the beginning, these prediction were based on the so-called static–harmonic approximation, which models molecular vibrations as a system of harmonic oscillators, as it will be discussed in the following.

1.2.1. The Static–Harmonic Approach

Assuming a reasonably smooth potential energy surface, the potential energy $V(\mathbf{x})$ of a set of N atoms in close proximity of some point $\mathbf{x}^0$ with respect to the atom's $3N$ Cartesian coordinates $\mathbf{x} = (x_1, \dots, x_{3N})$ can be approximately expressed in terms of a second-order multi-dimensional Taylor expansion:

$$V(\mathbf{x}) \approx V(\mathbf{x}^0) + \sum_{i=1}^{3N} \left(\frac{\partial V}{\partial x_i}\right)_{\mathbf{x}^0} \left(x_i - x_i^0\right) + \frac{1}{2} \sum_{i=1}^{3N} \sum_{k=1}^{3N} \left(\frac{\partial^2 V}{\partial x_i \partial x_k}\right)_{\mathbf{x}^0} \left(x_i - x_i^0\right) \left(x_k - x_k^0\right) \tag{1.2.1}$$

If the point $\mathbf{x}^0$ is an energy minimum, the gradients $\left(\frac{\partial V}{\partial x_i}\right)_{\mathbf{x}^0}$ vanish, and only the second derivatives remain:

$$V(\mathbf{x}) \approx V(\mathbf{x}^0) + \frac{1}{2} \sum_{i=1}^{3N} \sum_{k=1}^{3N} \left(\frac{\partial^2 V}{\partial x_i \partial x_k}\right)_{\mathbf{x}^0} \left(x_i - x_i^0\right) \left(x_k - x_k^0\right) \tag{1.2.2}$$

Based on this approximation, the force F_i acting on coordinate x_i can be expressed as

$$F_i := -\left(\frac{\partial V}{\partial x_i}\right) = -\frac{1}{2} \sum_{k=1}^{3N} \left(\frac{\partial^2 V}{\partial x_i \partial x_k}\right)_{\mathbf{x}^0} \left(x_k - x_k^0\right) \tag{1.2.3}$$

By inserting Equation 1.2.3 into Newton's equation

$$F_i = m_i \, a_i = m_i \left(\frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2}\right), \qquad i = 1, \dots, 3N, \tag{1.2.4}$$

one obtains a system of equations of motion

$$m_i \left(\frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2}\right) = -\sum_{k=1}^{3N} \left(\frac{\partial^2 V}{\partial x_i \partial x_k}\right)_{\mathbf{x}^0} \left(x_k - x_k^0\right), \qquad i = 1, \dots, 3N \tag{1.2.5}$$

To express these equations more compactly, it is desirable to switch to the set of so-called *mass-weighted Cartesian coordinates* $\mathbf{q}=(q_1,\ldots,q_{3N})$ defined as

$$q_i := x_i\sqrt{m_i}, \tag{1.2.6}$$

in which Equation 1.2.5 now reads

$$\frac{\mathrm{d}^2 q_i}{\mathrm{d}t^2} = -\sum_{k=1}^{3N} H_{i,k}\, q_k, \qquad i = 1,\ldots,3N \tag{1.2.7}$$

with the short-hand notation $H_{i,k}$ for the mass-weighted Hessian matrix

$$H_{i,k} := \frac{1}{\sqrt{m_i m_k}}\left(\frac{\partial^2 V}{\partial x_i \partial x_k}\right)_{\mathbf{x}^0}. \tag{1.2.8}$$

As the above approximation defines a purely harmonic potential energy surface *(see Equation 1.2.2)*, one can assume that the motion of each coordinate q_i can be described by a sinusoidal time evolution

$$q_i(t) = q_i^0 + A\cdot\cos(\omega t) \tag{1.2.9}$$

with some amplitude A and angular frequency ω. Inserting this into Equation 1.2.7 yields

$$\omega^2 q_i = \sum_{k=1}^{3N} H_{i,k}\, q_k, \qquad i = 1,\ldots,3N, \tag{1.2.10}$$

which can be rewritten as matrix eigenvalue problem for the mass-weighted Hessian matrix $\mathbf{H}$ with the eigenvalues ω^2:

$$\omega^2\mathbf{q} = \mathbf{H}\mathbf{q}. \tag{1.2.11}$$

In other words, solving this eigenvalue problem directly yields the vibrational frequencies ω of the system of atoms as the square roots of the eigenvalues. Furthermore, the eigenvectors $\mathbf{q}$ of the matrix represent the directions along which the atoms are displaced within each normal mode. Therefore, a full set of normal modes—each with vibrational frequency and atom displacement vector—can be obtained via this approach. Additional care has to be taken to project out the invariants *(i. e., translational and rotational invariance due to the conservation of momentum* *and angular momentum)* out of the mass-weighted Hessian matrix, so that typically only $(3N-6)$ normal modes are obtained.

The method described above only yields vibrational frequencies of the modes. However, in order to predict real vibrational spectra, also the intensities of each mode need to be computed. This is typically performed by computing derivatives of some more or less complicated properties with respect to the atom displacements $\mathbf{q}$ obtained for each mode. For example, the infrared intensity of a certain mode is proportional to the change in electric dipole moment that occurs when the atoms are displaced along that mode. More details on computing the intensities for the different kinds of vibrational spectroscopy can be found in Section 2.

Static–harmonic infrared[43–52] and Raman[43,45,47–50,53–57] spectra have been available for several decades in quantum chemistry software packages such as Gaussian,[58] and static VCD[29,48,50,59–66] and ROA[45,49,50,53,57,67–72] spectra can also be computed since many years. Concerning resonance Raman spectra, the situation is similar.[73–82]

Despite still being the standard approach in the literature, computing spectra via the static–harmonic approach comes with several severe limitations:

- Due to the harmonic approximation of the potential energy surface, all anharmonic effects[83] are neglected. If the system possesses features such as strong hydrogen bonds or hindered rotations, the harmonic approximation of certain modes will be poor, and so will be the quality of the predicted spectrum.

- The spectrum can only be computed for one minimum energy structure at a time. If there exist several conformers of the molecule, they need to be considered separately. If the system can hardly at all be described by minimum energy structures *(such as bulk phase liquids)*, it will be hard to obtain reasonable spectra at all.

- The method works best for molecules or small clusters in vacuum. Solvent effects on the spectrum *(which can be very pronounced)* can be crudely approximated either via continuum solvation models[84–91] such as COSMO and PCM, or by means of microsolvation, but the solvent effect cannot be captured in whole.

- The approach only yields a discrete line spectrum; no line widths or band shapes can be obtained. To predict realistic spectra, empirical line broadening needs to be applied.

It should be noted that a lot of effort was put into more accurate and more efficient electron structure treatment[92–94] as well as into going beyond the harmonic approximation, and more complex approaches such as, *e.g.*, vibrational self consistent field theory (VSCF), 2nd order vibrational perturbation theory (VPT2), and vibrational configuration interaction (VCI) were introduced to account for certain anharmonic effects.[45,49,50,95–105]

1.2.2. The Time–Correlation Approach

Apart from employing the static–harmonic approximation, there exists the possibility to compute vibrational spectra directly from molecular dynamics (MD) simulations. In this approach, the spectra are obtained as the Fourier transform of some time correlation functions along the simulation trajectory—an idea which is at least 55 years old now,[106–112] and is called the time–correlation function (TCF) formalism. This approach comes with several advantages over the static–harmonic concept:

- Condensed-phase systems can be handled; it is possible to explicitly capture the effects of solvent and entropy on the spectrum.
- Some anharmonic effects such as line broadening, approximate overtones and combination bands[113] are reproduced.
- Realistic band shapes are obtained instead of a discrete line spectrum.
- Intrinsic conformer sampling takes place during the MD simulation.
- No minimum energy structure is required to compute the spectrum.

However, due to the very limited accuracy of early MD simulations, this method initially was of relatively little use.

With the advent of faster computers in the 1980s, it became possible to perform molecular dynamics (MD) simulations on the basis of a quantum chemical description of the electron structure, referred to as *ab initio* molecular dynamics (AIMD).[114,115] The pioneers on that field were Roberto Car and Michele Parrinello, which published their approach of Car–Parrinello molecular dynamics (CPMD) in 1985.[116] With this class of methods, periodic bulk phase systems can be treated natively, and even anharmonic effects are covered to a certain extent, as line shapes, overtones, and combination bands are reproduced in qualitatively correct manner.[113,117] The first predictions of infrared spectra based on AIMD followed in 1997 on the examples of liquid water[118] and amorphous silica,[119] while the first computed Raman spectrum from AIMD was published in 2002.[120] Since then, many more infrared[113,121–135,MB14,MB19,MB23] and Raman[113,130,136–140,MB14,MB19,MB23,MB26,MB36,MB49] spectra of condensed phase systems based on this methodology have been presented in the literature. Complementing experimental spectra with atomistic simulations has shown to be a powerful method to gain insights into directed interactions such as, *e. g.*, hydrogen bonds.[141–149]

However, with regard to the chiral variants of these two spectroscopies, namely VCD and ROA, a similar progress was not visible for a long time. Only in 2016, Thomas and Kirchner published the first predicted VCD spectrum of a liquid system,[150] based on a classical approach to obtain the magnetic moments that will be discussed in Section 2.6 of this thesis. A few months later, Scherrer and Sebastiani published another predicted VCD spectrum of a liquid[151] which was based on nuclear velocity perturbation theory (NVPT).

Concerning ROA spectra, the first predicted such spectrum of a periodic liquid system was published by our group in 2017,[MB26] based on an AIMD simulation. Our newly developed approach is discussed in Section 2.8. Only a few months earlier, another ROA spectrum derived from AIMD was published by Luber,[152] but only one single molecule in vacuum was investigated, and as some derivations have explicitly been conducted under the assumption of a non-periodic system, it is concluded there that the application to liquid systems remains an open project for the future.

Predicting resonance Raman spectra from AIMD simulations was not possible for a long time. The first such spectrum has been published by our group in 2019,[MB36] using a newly developed approach that will be presented in Section 2.10 of this thesis.

SFG spectra based on molecular dynamics simulations can be computed since some years and have been presented in the literature several times now.[153–165]

It should be noted that most standard AIMD simulations describe the nuclei classically, and this description is an approximation which sometimes fails to capture certain effects relevant for reproducing experimental spectra.[166] Several approaches[167] to overcome this limitation have been developed, among them the path integral molecular dynamics (PIMD)[168,169] approach, which is often applied in the variants of centroid Molecular Dynamics (CMD)[170–174] and ring polymer molecular dynamics (RPMD).[175,176] It has been shown that PIMD is able to reproduce certain nuclear quantum effects which are missing in simulations with classical nuclei.[177,178]

1.3. Atomistic Simulation Methods

The discipline of atomistic simulation is concerned with computer simulations of chemical systems on the level of individual atoms. It is a relatively recent field of science, which originated in the late 1950s, when the first computers were applied to investigate properties of matter.[179] Due to the exponential rise of available computer power within the last decades,[180] methods of molecular simulation are nowadays an important and widely used tool in the field of computational chemistry and physics. On modern computers, huge systems can be simulated, consisting of more than one trillion (10^{12}) of particles,[181] reaching system sizes in the multi-micrometer regime, which are almost "macroscopic". Methods of molecular simulation were shown to enable the accurate prediction of real-world properties.[115]

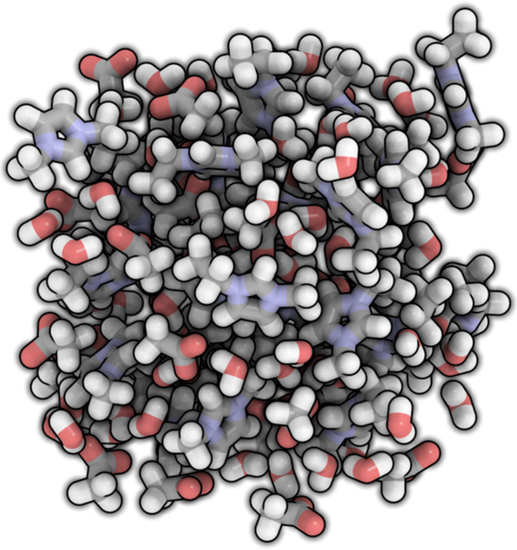


Figure 1.3.1.: A snapshot of an atomistic simulation containing the ionic liquid [EMIm][OAc] and water.[MB8,MB13]

Atomistic simulation methods can be categorized either by the sampling approach which is used to create new configurations *(i. e., sample the phase space)*, or by the method for the potential energy computation of a certain configuration.

The two most important sampling methods for atomistic simulation are molecular dynamics (MD) and Monte Carlo (MC). As the MC method is not frequently used to predict vibrational spectra, only MD shall be discussed in the scope of the thesis. In this approach, the simulation box is described as a dynamical system with a certain number of degrees of freedom for each simulated atom. Each such degree of freedom is described by its position and its velocity through the course of the simulation. The system is propagated in time by solving Newton's equations of motion for all degrees of freedom, given by

$$m_i \dot{\mathbf{v}}_i(t) = \mathbf{F}_i(\mathbf{r}_i) = -\nabla V(\mathbf{r}_i) \tag{1.3.1}$$

$$\dot{\mathbf{r}}_i(t) = \mathbf{v}_i(t)\,, \tag{1.3.2}$$

where $\mathbf{r}_i$, $\mathbf{v}_i$, and $\mathbf{F}_i$ denote the position, velocity, and force vector of particle i, respectively, m_i represents the particle's mass, $V(\mathbf{r})$ is the underlying potential energy surface, and ∇ depicts the gradient operator. This system of coupled ordinary differential equations (ODEs) of first order is solved by stepwise integration with a finite time step Δt, which is usually kept constant during the simulation. This yields a discrete trajectory with a steady progression of time, containing the particle's positions and velocities in each time step.

Of all different methods to numerically solve coupled systems of ODEs, the *symplectic integration schemes* for Hamiltonian systems attained special attention in the field of molecular dynamics. They possess a very good conservation of energy. In contrast to many other methods, they show a reasonable behavior when investigating the long-term evolution of chaotic Hamiltonian systems (like, *e.g.*, MD simulations). Three popular such symplectic integration schemes are the *Leapfrog* algorithm, the *Verlet* method,[182] and the *velocity Verlet* integrator. Despite their different names, they are very similar. It can be easily seen that the Verlet and velocity Verlet methods are algebraically equivalent *(by eliminating the velocities from the velocity Verlet algorithm)*, and it can be shown that, eventually, all three methods are identical.[183] All three methods are explicit integration methods with a global error of order 2, and therefore one order better than the semi-implicit Euler method, which is also a symplectic integration scheme. As the velocity Verlet algorithm is the only of these three methods which yields velocities and positions at the same points in time, many popular molecular dynamics packages use this integrator.

The general equations of the velocity Verlet scheme read

$$\mathbf{r}(t{+}\Delta t) \;:=\; \mathbf{r}(t) + \mathbf{v}(t)\,\Delta t + \frac{1}{2}\mathbf{a}(t)\,\Delta t^2, \tag{1.3.3}$$

$$\mathbf{v}(t{+}\Delta t) \;:=\; \mathbf{v}(r) + \frac{\mathbf{a}(t) + \mathbf{a}(t{+}\Delta t)}{2}\Delta t. \tag{1.3.4}$$

By inserting

$$\mathbf{a}_i(t) = \frac{\mathbf{F}_i(t)}{m_i}, \quad i = 1 \ldots N, \tag{1.3.5}$$

one arrives at the two-step method

$$\begin{aligned}
\mathbf{r}_i(t+\Delta t) \quad &:= \quad \mathbf{r}_i(t) + \mathbf{v}_i(t)\,\Delta t + \frac{\mathbf{F}_i(t)}{2m_i}\Delta t^2, \quad i = 1 \ldots N, &(1.3.6)\\
\mathbf{v}_i(t+\Delta t) \quad &:= \quad \mathbf{v}_i(t) + \frac{\mathbf{F}_i(t) + \mathbf{F}_i(t+\Delta t)}{2m_i}\Delta t, \quad i = 1 \ldots N, &(1.3.7)
\end{aligned}$$

which is implemented in most MD program packages.

One of the central aspects of atomistic simulations is the so-called *ergodicity*. Only if a system is ergodic, ensemble averages and temporal averages become equivalent, so that a simulation can be used to predict experimental quantities. Unfortunately, it is not possible to prove ergodicity for any system of practical relevance, so this has to remain an assumption in all work based on simulations. The mathematical aspects of ergodicity are discussed in Section A.1 in the Appendix.

Nosé–Hoover Chain Thermostat

Standard molecular dynamics simulations sample the NVE ensemble. Experiments, however, are often carried out at constant temperature, so that it would be more desirable to simulate in the NVT ensemble in order to reproduce experimental quantities. This can be achieved by the use of a so-called thermostat in the MD simulation. Many such thermostats have been developed over the years, and there is no "best" choice. One frequently applied method is the so-called Nosé–Hoover chain thermostat (NHC),[184] which combines multiple Nosé–Hoover thermostats[185–187] into a chain. It samples the NVT ensemble very well; however, it is only capable of maintaining adequate temperature control close to the thermodynamic equilibrium.[188] Chain lengths larger than 2 are recommended in general cases where there is a broad distribution of vibrational frequencies in the system.[189]

The adapted equations of motion[190,191] for the N particles with NHC thermostat and several corresponding virtual degrees of freedom ξ_j read

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i}, \tag{1.3.8}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \mathbf{p}_i \dot{\xi}_1, \tag{1.3.9}$$

where $\mathbf{r}_i$ and $\mathbf{p}_i$ are the position and momentum of the i-th particle, respectively, m_i is the particle's mass, and $\mathbf{F}_i$ is the external force acting on the particle. It can be seen that the velocity of the first virtual degree of freedom ξ_1 interacts with the motion of the real particles.

The motion of the M virtual degrees of freedom ξ_j belonging to a NHC thermostat with chain length M is governed by the set of differential equations

$$\ddot{\xi}_1 = \frac{1}{Q_1}\left(\sum_{i=1}^{N} m_i \mathbf{v}_i^2 - N_f k_\mathrm{B} T_\mathrm{set}\right) - \dot{\xi}_1 \dot{\xi}_2 \tag{1.3.10}$$

$$\ddot{\xi}_j = \frac{1}{Q_j}\left(Q_{j-1}\dot{\xi}_{j-1}^2 - k_\mathrm{B} T_\mathrm{set}\right) - \dot{\xi}_j \dot{\xi}_{j+1}, \qquad j = 2, \ldots, (M-1) \tag{1.3.11}$$

$$\ddot{\xi}_M = \frac{1}{Q_M}\left(Q_{M-1}\dot{\xi}_{M-1}^2 - k_\mathrm{B} T_\mathrm{set}\right), \tag{1.3.12}$$

where $\mathbf{v}_i$ is the velocity of the i-th real particle, N_f is the number of active degrees of freedom in the system of real particles, T_set is the thermostat's target temperature, and k_B is the Boltzmann constant.[190] The parameters Q_j are the masses of the virtual particles, which should be chosen[192] according to

$$Q_1 := N_f k_\mathrm{B} T_\mathrm{set} \tau^2, \tag{1.3.13}$$

$$Q_j := k_\mathrm{B} T_\mathrm{set} \tau^2, \qquad j = 2, \ldots, M, \tag{1.3.14}$$

where τ is the desired time constant of the Nosé–Hoover chain thermostat.

The conserved quantity of the simulation *(including both the real particles and the thermostat)* can be expressed as

$$E_\mathrm{NHC} := \frac{1}{2}\sum_{i=1}^{N} m_i \mathbf{v}_i^2 + E_\mathrm{pot}\left(\{\mathbf{r}_i\}\right) + \sum_{j=1}^{M} \frac{1}{2} Q_j \dot{\xi}_j^2 + N_f k_\mathrm{B} T_\mathrm{set} \xi_1 + \sum_{j=2}^{M} \xi_j k_\mathrm{B} T_\mathrm{set}, \tag{1.3.15}$$

where the first two terms are the kinetic and potential energy of the atomistic system, and the remaining terms represent the kinetic and potential energy of the thermostat, respectively.

The equations of motion given above can be solved iteratively using the Newton–Raphson procedure. However, there exist also direct (explicit) integration approaches, which are preferable. In these explicit integrators, a substantial improvement of the accuracy can be obtained by using a higher order integration algorithm, *e. g.* the Yoshida–Suzuki integration scheme,[192] combined with a multiple time step (MTS) approach.

An explicit integration algorithm to apply the Nosé–Hoover chain thermostat to a molecular dynamics simulation can be found in Section A.4 in the Appendix.

1.4. Computing Electromagnetic Moments

Computing vibrational spectra—both via the static–harmonic approach and via the time-correlation function (TCF) approach—requires knowledge of certain electromagnetic moments of the system *(e. g., the electric dipole moment)*. While in the former case, these moments are calculated for excursions of the minimum structure along the normal modes, they are computed for snapshots along the simulation trajectory in the latter case. For non-periodic systems, electric moments can be readily derived as expectation values from electron structure calculations by applying the corresponding moment operator to the converged wave function. Under periodic boundary conditions, however, the standard moment operators are ill-defined. This issue has been addressed by the so-called modern theory of polarization, which is based on a Berry phase,[193–196] and can provide the electric moments of the whole simulation cell under periodic boundary conditions. By combining this method with perturbation theory, it is also possible to calculate the polarizability.[197]

There are several reasons for considering molecular instead of system-wide electromagnetic moments. First, the sampling of the spectrum is improved and the spectrum contains less noise if cross-correlations of non-neighboring molecules are omitted *(see discussion of Equation 2.4.3 in Section 2.4)*. Secondly, it allows to separate the spectral contributions of the constituents of a mixture, so that, *e. g.*, the solvent spectrum can be suppressed.

One widely used approach to assign electric dipole moments to individual molecules is a *localization* of the molecular orbitals (MOs) in space. For non-periodic systems, there exist some well-known and computationally efficient methods such as the Boys–Foster localization[198] and the Pipek–Mezey localization,[199] among others.[200,201]

In periodic systems, localization of orbitals is considerably more involved. One commonly used method is the so-called **Wannier localization**.[202–207] It applies a unitary transformation **U** to the set of occupied Kohn–Sham orbitals $|\psi_i\rangle$ so that another set of molecular orbitals $|\tilde{\psi}_n\rangle$ is obtained, which are called Wannier orbitals or maximally localized Wannier functions (MLWFs):[208]

$$|\tilde{\psi}_n\rangle = \sum_i U_{i,n} |\psi_i\rangle . \tag{1.4.1}$$

The unitary transformation **U** is constructed in a way so that a so-called spread functional

$$\boldsymbol{\Omega} = \sum_n \sum_I f\left(|z_{I,n}|^2\right), \tag{1.4.2}$$

$$z_{I,n} = \left\langle \psi_n \middle| \mathbf{O}^I \middle| \psi_n \right\rangle, \tag{1.4.3}$$

is minimized.[208] Here, $\mathbf{O}^I$ is a class of suitable spread operators that are well-defined in periodic space, such as

$$\mathbf{O}^I = \exp(\mathrm{i}\mathbf{G}_I \cdot \mathbf{r}) \tag{1.4.4}$$

with the I reciprocal lattice vectors $\mathbf{G}_I$, and f is an appropriate function. Common choices for f are[208]

$$f_1\left(|z_{I,n}|^2\right) = \sqrt{|z_{I,n}|^2} = |z_{I,n}|\,, \text{[203]} \tag{1.4.5}$$

$$f_2\left(|z_{I,n}|^2\right) = \log\left(|z_{I,n}|^2\right), \text{[209]} \tag{1.4.6}$$

$$f_3\left(|z_{I,n}|^2\right) = |z_{I,n}|^2\,. \text{[210]} \tag{1.4.7}$$

Note that both the Boys–Foster localization[198] and the Pipek–Mezey localization[199] for non-periodic systems can be expressed in terms of the above equations with the choice of $f \equiv f_3$. For Boys–Foster, the operator $\mathbf{O}^I$ is simply defined as $\mathbf{O}^I = \mathbf{r}_I$ with the conventional position operator $\mathbf{r}_I$ for $I = x, y, z$.

One traditionally applied approach in quantum chemistry for localizing MOs is the method of two-by-two orbital rotations first introduced by Edmiston and Ruedenberg.[200] Unfortunately, the analytical expression for the optimal angle of these rotations can only be derived for the choice of $f \equiv f_3$,[208] *i. e.*, for the Silvestrelli–Marzari–Vanderbilt,[203,210] the Boys,[198] and the Pipek–Mezey functional.[199] For the choice of $f \equiv f_1$ which is commonly used for performing Wannier localization, one has to resort to iterative numerical methods such as a generalized Jacobi rotation scheme[211] or the so-called "crazy angle" algorithm by Joost VandeVondele in CP2k.[212–214] All these methods require considerable amounts of extra computer time for the localization and are not guaranteed to converge at all. A discussion of these issues can be found in Section 2.1.4.

The centroids of the Wannier orbitals are called Wannier centers or maximally localized Wannier centers (MLWCs); they can be seen as the positions of electron pairs in a simple picture. As those are located relatively closely to the atoms, it is well possible to assign Wannier centers to individual molecules. Based on these

Wannier centers, the molecular dipole moment $\boldsymbol{\mu}^{\mathrm{Mol}}$ can be expressed as

$$\boldsymbol{\mu}^{\mathrm{Mol}} = -2e\sum_{i=1}^{N}\mathbf{r}_i + e\sum_{j=1}^{M} Z_j\mathbf{R}_j, \tag{1.4.8}$$

where N is the number of Wannier centers in the molecule, $\mathbf{r}_i$ is the position of the i-th Wannier center, M is the number of atoms in the molecule, $\mathbf{r}_j$ and Z_j are the position and the nuclear charge of the j-th atom, respectively, and e is the elementary charge. The sum of all molecular dipole moments computed by this protocol is often a good approximation to the total dipole moment of the system.[206]

The Wannier localization approach can also be used to compute molecular electric polarizabilities via finite differences of external electric fields.[MB14] In linear approximation, the dipole moment $\boldsymbol{\mu}_{\mathrm{ind}}$ induced by an electric field $\mathbf{E}$ can be expressed as

$$\boldsymbol{\mu}_{\mathrm{ind}} = \boldsymbol{\alpha}\mathbf{E} \tag{1.4.9}$$

with the second-order electric polarizability tensor $\boldsymbol{\alpha}$. This leads to the central finite differences

$$\alpha_{i,j} = \frac{\mu_i^{j+} - \mu_i^{j-}}{\left|\mathbf{E}^{j+} - \mathbf{E}^{j-}\right|}, \quad i,j = x,y,z, \tag{1.4.10}$$

where $\mathbf{E}^{j+}$ and $\mathbf{E}^{j-}$ are the field vectors of the external electric field applied in positive and negative j direction, respectively, μ_i^{j+} and μ_i^{j-} are the i components of the molecular dipole moments *(obtained from Equation 1.4.8)* under the influence of these two fields, and $\alpha_{i,j}$ is the (i,j) component of the molecular polarizability tensor. By performing six additional SCF calculations with positive and negative fields in x, y, and z direction, the full polarizability tensor can thus be obtained. The strength of the electric field $|\mathbf{E}|$ needs to be chosen so that the molecule is still within the linear regime of polarizability *(i. e., Equation 1.4.9 is still a good approximation)*.

When molecular polarizabilities are computed according to Equation 1.4.10, the changes in the local electric field of a molecule by the polarization of the neighboring molecules are omitted. This effect can be captured by considering the dipole–dipole interaction tensor computed by Ewald summation[215] under periodic boundary conditions as explained in references 216 and 217. However, a recent study of water has shown that this has only a minor influence on the resulting spectra.[138]

There also exist alternative approaches to computing molecular polarizabilities: one by Partovi–Azar and Kühne based on the spread of the Wannier centers,[218,219] and one by Luber based on perturbation theory.[134]

1.5. Programs used within this Thesis

This section describes which programs have been used for creating this thesis.

- The molecular dynamics simulations which have been evaluated within this thesis have been performed with LAMMPS[220,221] and CP2k.[212–214] Initial configurations have been created with Packmol.[222,223]
- All trajectory analyses have been carried out with TRAVIS.[MB2,MB40]
- Plots of one-dimensional functions and spectra have been created with Xmgrace.[224]
- Contour plots and density plots have been obtained either from Mathematica[225] or from Gnuplot.[226]
- Snapshots of molecular structures as well as simulation boxes have been created with VMD,[227] using the Tachyon renderer.[228] Figures 1.3.1 and 3.6.4 have been rendered with Qutemol.[229]
- All calculations have been carried out on a local GNU/Linux compute cluster running Debian and Slurm, using OpenMPI for parallelization.
- This thesis has been created using LaTeX for typesetting, using the Palatino Linotype font.

Chapter 2.

Method Development and Theory

In this chapter, several methods for computing vibrational spectra from AIMD simulations *(and some other methods which yield quantities required to do so)* are described. While some of them have been developed in the scope of this thesis, others were developed by other groups and are introduced here because of their relevance.

2.1. Voronoi Integration

One of the central pieces of the approach presented herein to computing vibrational spectra is the radical Voronoi integration which we have developed.[MB23,MB40,MB48] In short, it is a method for obtaining atomic or molecular electromagnetic moments in periodic bulk phase simulations from the corresponding volumetric quantities on a grid. It will be described on the following pages.

2.1.1. Voronoi Tessellation

The Voronoi tessellation[230,231] is a mathematical tool which partitions an Euclidean space containing some points (*Voronoi sites*) into non-overlapping subsets. Each Voronoi site corresponds to exactly one such subset (called *Voronoi cell*), which contains all points from space which are closer to this Voronoi site than to any other Voronoi site. In mathematical form, this is written as

$$C_i := \left\{ \mathbf{x} \in \mathbb{R}^n \middle| \; \|\mathbf{x} - \mathbf{p}_i\| \leq \|\mathbf{x} - \mathbf{p}_j\| \;\; \forall j \in \{1 \dots k\}, \; j \neq i \right\}, \quad i \in \{1 \dots k\}, \tag{2.1.1}$$

where $\mathbb{R}^n$ stands for any Euclidean space with the norm $\|\cdot\|$, in which k Voronoi sites, each with position $\mathbf{p}_i \in \mathbb{R}^n$, are given, and the $C_i \subseteq \mathbb{R}^n$ are the resulting Voronoi cells.

By considering atoms in three-dimensional space as Voronoi sites, this concept has widely been applied in different fields of computational chemistry. To name a few advantages of the method, the Voronoi tessellation of a set of atoms is uniquely defined and can be calculated with moderate computational demands. The Voronoi tessellation can easily be adopted to systems with periodic boundary conditions, and is therefore well suited for bulk phase systems. Finally, the method does not possess any empirical parameters to tune, and therefore gives an unbiased and uniquely defined picture.

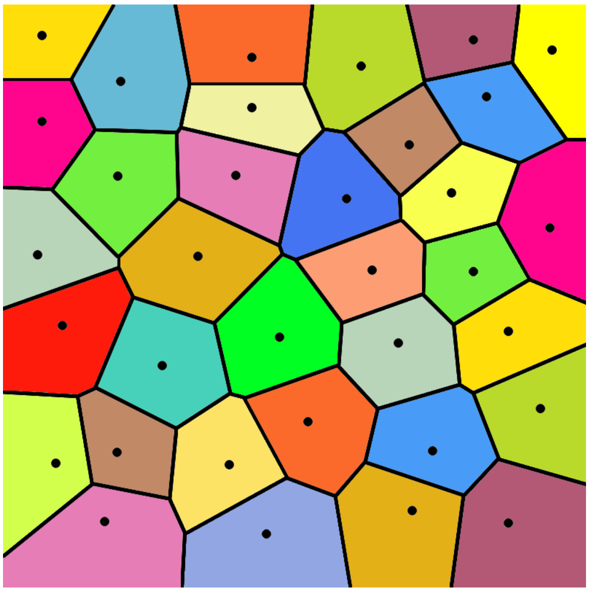


Figure 2.1.1.: Two-dimensional illustration of the Voronoi tessellation. Black dots correspond to Voronoi sites *(input)*, while colored polygons depict the resulting Voronoi cells.

Voronoi tessellation has already been used before to partition the total electron density, by placing a simple plane midway between two atoms.[232,233] This idea was extended later on to account for different atom sizes by shifting the boundary planes,[234–237] for example in Richards' "method B"[238] which found some applications.[236]

However, certain limitations do arise from the properties of the standard Voronoi tessellation. As all atoms are treated in the same way, Voronoi polyhedra of light atoms like hydrogen will on average have the same size as those around heavier atoms like iodine. From a mathematical point of view, this is not a problem, but from a chemical perspective, this is completely unreasonable. If, *e. g.*, the electron density within the Voronoi cell of a hydrogen atom is integrated, the hydrogen atom would always end up with a heavily negative partial charge, because way too much electron density would be considered as belonging to this hydrogen atom.

2.1.2. Radical Voronoi Tessellation

To overcome this problem, radii need to be introduced into the Voronoi tessellation, allowing to treat different atom types differently. Several ways to do so have been proposed. We employ a generalization in terms of the radical Voronoi tessellation (also known as *power diagram* in the two-dimensional case).[239] In this technique, a radius is assigned to each atom, allowing to model the sizes of the atoms in a chemically reasonable sense. Such radii have also been used in reference 236, but instead of the ratio of the radii, the difference between the squared radii determines the position of the cell face between two atoms here. Thus, in contrast to the aforementioned "method B"[236,238] and similar approaches, the radical Voronoi tessellation does not suffer from the "vertex error",[238,239] *i. e.*, it does not contain holes. When integrating electron density, this is important to keep the total charge of the system constant. As another advantage, the Voronoi sites around which the cells are constructed can be kept on the atoms and do not have to be shifted (as it was done in reference 235) to obtain a chemically reasonable partitioning. To the best of our knowledge, the radical Voronoi tessellation has not been used for the computation of molecular electromagnetic moments before.

The definition of the radical Voronoi tessellation as a generalization of the classical tessellation reads

$$C_i^{\mathrm{r}} := \left\{ \mathbf{x} \in \mathbb{R}^n \middle| \; \|\mathbf{x}-\mathbf{p}_i\|^2 - r_i^2 \leq \|\mathbf{x}-\mathbf{p}_j\|^2 - r_j^2 \quad \forall j \in \{1 \dots k\},\; j \neq i \right\}, \quad i \in \{1 \dots k\}, \tag{2.1.2}$$

with radius r_i for Voronoi site i. While in the classical case the face between two adjacent Voronoi cells is always placed in the middle between the corresponding Voronoi sites, its position is now determined by the difference of the squared radii. From Equation 2.1.2, it can be derived that the separation plane between two sites A and B with radii r_{A} and r_{B} will be located at a position

$$w := \left(\frac{1}{2} + \frac{r_{\mathrm{A}}^2 - r_{\mathrm{B}}^2}{2R_{\mathrm{AB}}^2} \right) R_{\mathrm{AB}}, \tag{2.1.3}$$

where R_{AB} is the distance between both sites, and w describes the distance of the separation plane from A – see Figure 2.1.2. It can be seen that the relative position of the plane depends on the distance between the sites: if the distance becomes large with respect to the radii, the plane will be located in the middle, even if the radii differ. In the other extreme case of a small inter-site distance when compared to the radii, w can even be outside of the interval $[0, R_{\mathrm{AB}}]$, which means that either one of the sites is no longer located inside of its Voronoi cell, or the Voronoi cell of this site is degenerate *(empty)*. However, both cases are not a problem if electron density shall be integrated within the cells. These effects are more pronounced if the differences between the radii become larger. If all radii are equal, the radical Voronoi tessellation becomes identical to the classical Voronoi tessellation, and those degeneracies cannot occur. A two-dimensional schematic illustration of the radical Voronoi tessellation in the case of benzene is shown in Figure 2.1.3. Please note that the term "radical" is not related to chemical radicals *(which possess unpaired electrons)*.

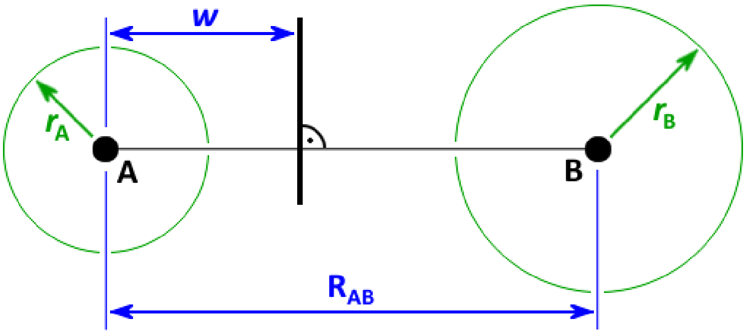


Figure 2.1.2.: Separation plane between two Voronoi sites A and B with radii r_A and r_B in the radical Voronoi tessellation, see Equation 2.1.3.

The crucial parameters in the radical Voronoi tessellation are the radii assigned to the atoms. We have recently shown[MB23] that van der Waals (vdW) radii[240–242] yield a reasonable separation of molecules in the bulk phase, and that the resulting molecular electromagnetic "Voronoi" moments can readily be used to calculate vibrational spectra of bulk phase systems from *ab initio* molecular dynamics (AIMD) simulations, including infrared,[MB23] Raman,[MB23,MB26,MB36] vibrational circular dichroism (VCD),[150] Raman optical activity (ROA),[MB26,MB51] and resonance Raman[MB36] spectra. Since van der Waals radii have been fitted to reproduce intermolecular distances, it can be expected that they lead to a suitable placement of the molecular boundaries in a radical Voronoi tessellation.

The definition of the radical Voronoi tessellation in Equation 2.1.2 shows that the tessellation will not change if the set of radii $\{r_i\}$ is transformed to a new set $\{r'_i\}$ by the map

$$r'_i := \sqrt{r_i^2 + C}, \quad i \in \{1 \ldots k\} \tag{2.1.4}$$

with some constant $C \in \mathbb{R}$. Due to this relation, the absolute value of the radii does not have a direct meaning.

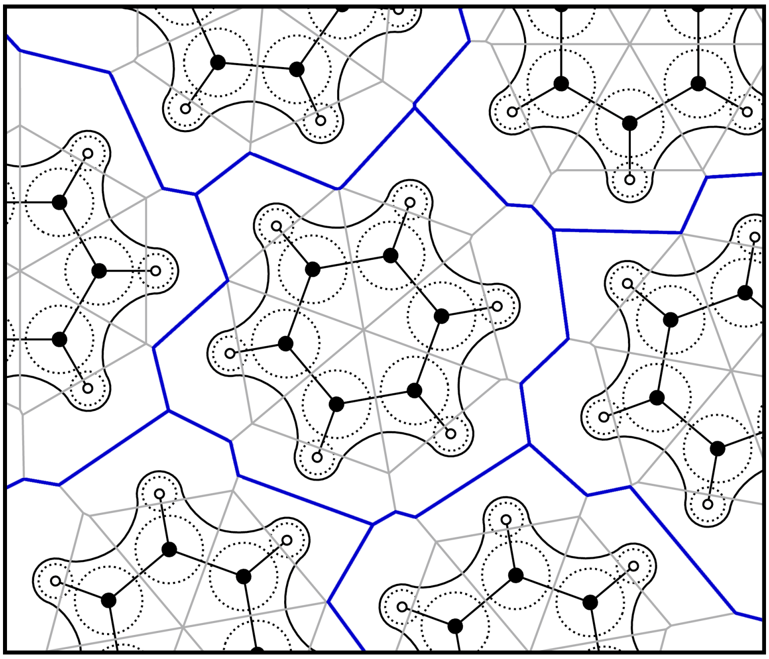


Figure 2.1.3.: Two-dimensional illustration of the radical Voronoi tessellation in liquid benzene. Solid black lines are iso-lines of the electron density, dashed circles indicate the atomic radii, radical Voronoi cells are shown as gray solid lines with resulting molecular boundaries drawn in blue.

In the TRAVIS[MB2,MB40] implementation of the method presented herein,[MB23] the Voro++ library[243,244] from Chris Rycroft is used to perform the radical Voronoi tessellation of periodic simulation cells, which may have the shape of any parallelepiped *(therefore not restricting our implementation to orthorhombic cells)*.

2.1.3. Voronoi Integration

After the construction of the Voronoi cells, the volumetric data *(e. g., the electron density)* needs to be integrated within each Voronoi cell to obtain the atomic electromagnetic moments. If molecular quantities are required, the integration runs over the Voronoi cell of the molecule instead, which is just defined as the union of all atomic Voronoi cells. The Voronoi integration of the total electron density yields the molecular charge q^{Mol}, the molecular dipole vector $\boldsymbol{\mu}^{\mathrm{Mol}}$, and the molecular trace-free quadrupole tensor $\mathbf{Q}^{\mathrm{Mol}}$ according to the following equations:

$$q^{\mathrm{Mol}} = \sum_{i=1}^{N_{\mathrm{Mol}}} q_i \qquad - \int\limits_{\mathrm{Mol}} \rho(\mathbf{s}) \;\; \mathrm{d}^3\mathbf{s}, \tag{2.1.5}$$

$$\boldsymbol{\mu}^{\mathrm{Mol}} = \sum_{i=1}^{N_{\mathrm{Mol}}} q_i \mathbf{r}_i \qquad - \int\limits_{\mathrm{Mol}} \rho(\mathbf{s})\, \mathbf{s} \;\; \mathrm{d}^3\mathbf{s}, \tag{2.1.6}$$

$$\mathbf{Q}_{jk}^{\mathrm{Mol}} = \sum_{i=1}^{N_{\mathrm{Mol}}} q_i \left(3\mathbf{r}_{i,j}\mathbf{r}_{i,k} - \|\mathbf{r}_i\|^2 \,\delta_{jk}\right) \qquad - \int\limits_{\mathrm{Mol}} \rho(\mathbf{s}) \left(3\mathbf{s}_j\mathbf{s}_k - \|\mathbf{s}\|^2 \,\delta_{jk}\right) \mathrm{d}^3\mathbf{s}, \tag{2.1.7}$$

where N_{Mol} is the number of atoms in the molecule, q_i and $\mathbf{r}_i$ are the core charge and position vector of the i-th atom of the molecule, respectively, $\rho(\mathbf{s})$ is the total electron density at position $\mathbf{s}$ in space, and δ is the Kronecker delta. The first part of the three equations is concerned with the charges of the nuclei, while the second part captures the electron density *(the actual Voronoi integration)*. The "Mol" integration volume denotes the molecular Voronoi cell. Please note that the molecular center of mass is chosen as the coordinate origin.

As the volumetric data in the simulation box is supplied on a grid, an efficient algorithm is required to traverse the grid points which are located inside a given Voronoi cell. A simplistic approach that checks for each grid point in which cell it is located would lead to very poor performance, as there are around 10 million grid points per typical AIMD snapshot. Instead, we have implemented another method: the three stride vectors of the grid are termed $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in the following. As non-orthorhombic simulation cells are permissible, these vectors do not need to be orthogonal to each other. At first, the maximum cross section of the Voronoi cell along the $\mathbf{v}_1$ direction is computed in the $\mathbf{v}_2$–$\mathbf{v}_3$ plane. A *(in the case of orthorhombic simulation cells)* rectangular bounding box in that plane is constructed around this section. For each grid coordinate pair within this bounding box in the $\mathbf{v}_2$–$\mathbf{v}_3$ plane, a ray is cast into $\mathbf{v}_1$ direction, and intersections between this ray and all Voronoi faces of the given Voronoi cell are probed. As Voronoi cells are always convex, there may be either zero or two such ray–face intersections, other combinations are not possible. With zero intersections, the ray misses the Voronoi cell, and no further action is taken. With two intersections, the entry and exit points of the ray through the Voronoi cell are known, and the grid points between the intersections can be summed up along the ray. This algorithm finally yields the sum over all grid points located within the given Voronoi cell. As each grid point is assigned to exactly one Voronoi cell by this algorithm, the total sum over all Voronoi cells is equal to the total sum over all grid points, which is important satisfy certain invariants *(e. g., to keep the total charge of the system fixed when integrating the electron density)*. This implementation has already been applied several times to obtain electromagnetic moments of molecules in bulk phase simulations.[MB23,MB26,MB36,MB48] Our approach is rather efficient – a full Voronoi integration of a bulk phase snapshot with around 1 000 atoms and 10 million grid points takes roughly 1 second on a single CPU core.

In real-world applications, the grid of the volumetric data is often relatively coarse in order to reduce the required storage space. Typical values are in the order of one grid point each 10 . . . 20 pm. As each grid point is completely assigned to exactly one Voronoi cell, infinitesimal changes in the radii may lead to grid points switching the cell they are assigned to. Therefore, the map from atomic radii to atomic electromagnetic moments is no longer continuous, or in other words, some amount of numerical noise is introduced, which would deteriorate the predicted spectra. To reduce the impact of this effect, we have developed and implemented an on-the-fly interpolation scheme for the volumetric data grid. During the integration pass, the grid can be refined by tri-linear interpolation. The smaller grid spacing which results from this procedure leads to a reduced amount of numerical noise. On the other hand, demands on storage system and core memory are not increased, as the interpolation is just performed on-the-fly while integrating. We call this approach *refinement*; it has been utilized in some applications of our method[MB48] with a refinement factor of 2 (*i. e.*, one grid point was interpolated to two grid points along each axis of the grid, yielding 8 grid points in total from each original grid point). Our implementation is not limited to a refinement factor of 2; higher values can be chosen on demand.

Since 2021, the Voronoi integration approach[MB23] is implemented directly in CP2k[212–214] *(starting from version 8.1)*, so that molecular electromagnetic moments can now directly be computed during AIMD runs.

Note that we have also applied the Voronoi integration approach to compute optimized atomic partial charges and atomic radii in condensed phase systems via minimization of the charge variance.[MB48]

2.1.4. Advantages over Wannier Localization

After we had developed the Voronoi integration approach to computing molecular electromagnetic moments, one of the first questions was if these moments are suitable to predict vibrational spectra. To investigate this, we used a simulation trajectory of liquid methanol to compute infrared and Raman spectra from molecular electric moments based on either Wannier localization or Voronoi integration.[MB23] The results are presented in Figure 2.1.4.

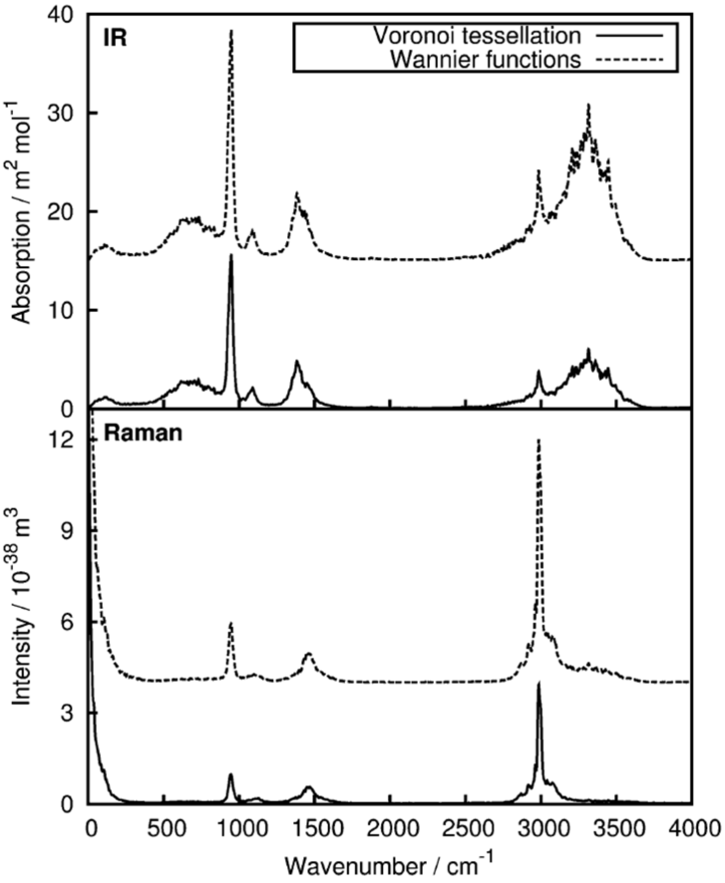


Figure 2.1.4.: Comparison of predicted infrared *(top panel)* and Raman *(bottom panel)* spectra of liquid methanol, using Wannier localization *(dashed lines)* and Voronoi integration *(solid lines)* to compute the molecular electric moments.[MB23]

It can be seen that apart from minor differences in the intensities, both the infrared and the Raman spectrum are almost identically predicted by the two approaches. The differences result from the fact that the Wannier localization enforces strictly neutral molecules *(as Wannier centers bear integer charge)*, while the Voronoi integration allows for a certain amount of charge transfer between the molecules. It can be concluded that molecular electromagnetic moments based on Voronoi integration are well suitable to predict bulk phase vibrational spectra.

Timing and Convergence

As described in Section 1.4, all known methods to perform a Wannier localization of a periodic system are *iterative*, which means that they converge towards the desired solution, but actually never reach the exact solution. Unfortunately, it can not even be guaranteed that the used algorithms always converge. In other words, it can happen that the localization procedure fails for a particular frame along a simulation, so that the electromagnetic moments are missing for that frame. Predicting vibrational spectra relies on computing time-correlation functions, where a missing frame would be highly problematic.

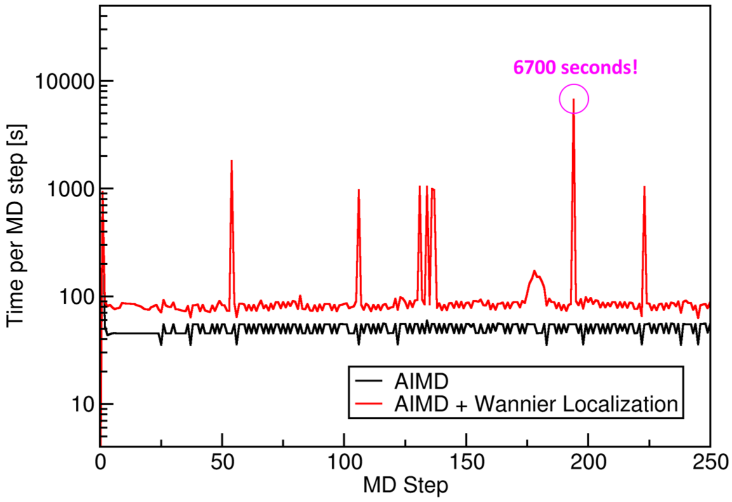


Figure 2.1.5.: Logarithmic frame times for a standard BOMD simulation *(black curve)* and a BOMD with Wannier localization *(red curve)* of a liquid phase system with ≈ 1000 atoms.

To give an example, consider Figure 2.1.5, where the frame times of a standard BOMD simulation with CP2k[212–214] *(black curve)* are compared to those with additional Wannier localization in each step *(red curve)*. The system is in the liquid phase and contains 936 atoms *(cubic cell size* ≈ 20 *Angstrom)*. Please note the logarithmic vertical axis. The average frame time of the standard BOMD is 47.9 s, while the frame time with additional Wannier localization is 139.3 s on average. This means that 65% of the total computer time are spent for the localization step, which is certainly not satisfactory. Apart from that, the frame times are highly irregular with Wannier localization. The reason is that CP2k uses the very efficient "crazy angle" algorithm for the localization by default. If this algorithm does not converge, iterative Jacobi diagonalization[211] is employed as a fallback. The latter is slightly more robust but considerably slower, so that frame times of several thousand seconds can be observed if the fallback is activated. But even if considering only the frames for which the fast "crazy angle" algorithm did converge, the time for the localization alone is still $\approx 30\,\mathrm{s}$ per frame, which is still a considerable amount of the total computer time.

Our Voronoi integration approach, in contrast, is not an iterative method and does not need to converge. This means that there do not exist cases in which no electromagnetic moments can be obtained for a particular frame. Furthermore, our approach is significantly faster than the Wannier localization—it requires only 2.0 s per frame for the same system on a single CPU core. Therefore, more than a factor 2 of total computer time can be saved by utilizing Voronoi integration instead of Wannier localization for each frame of the simulation.

Issues with Aromatic Systems

Another disadvantage of the Wannier localization approach are certain issues with aromatic systems. Please consider Figure 2.1.6, where we have predicted the infrared spectrum of liquid benzene based on Wannier localization *(dashed line)* and Voronoi integration *(solid line)*.[MB23] In the Wannier-based spectrum, artificial peaks appear between 1200 and 1350 cm^{-1}. These peaks are neither present in the experimental spectrum, nor if the spectrum is predicted via total cell dipole moment or Voronoi integration.

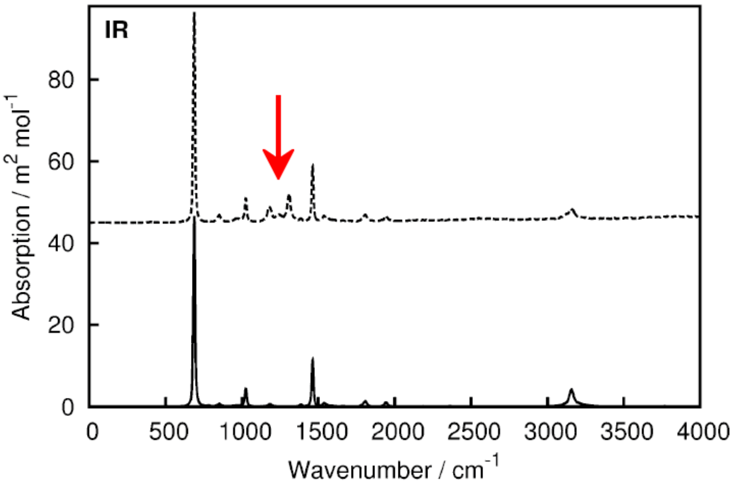


Figure 2.1.6.: Predicted infrared spectrum for liquid benzene based on Wannier localization *(dashed line)* and Voronoi integration *(solid line)*. The Wannier-based spectrum possesses artificial peaks between 1200 and 1350 cm^{-1}.[MB23]

After some investigation, we were able to identify the cause of these artificial peaks. When performing a Wannier localization of benzene, the aromatic electrons need to be localized, so that an alternating single bond/double bond pattern results—see the left panel of Figure 2.1.7. When considering all vibrational normal modes of benzene, one of them deforms the molecule towards cyclohexatriene *(see right panel of Figure 2.1.7)*. As this mode possesses an inversion symmetry, it cannot alter the total dipole moment of the molecule, and is therefore invisible in infrared spectroscopy. However, if this mode is active, the six ring bonds no longer have identical bond lengths, and there appear preferred positions where the Wannier centers of the aromatic electrons should be localized. As a result, the single bond/double bond pattern of the aromatic electrons *flips* with the frequency of this vibration. Due to numerical inaccuracies *(an iterative localization can never be fully converged)*, this leads to a small jump in the molecular dipole moment, which appears as an artificial peak in the infrared spectrum at the frequency of the cyclohexatriene deformation mode which should be invisible.

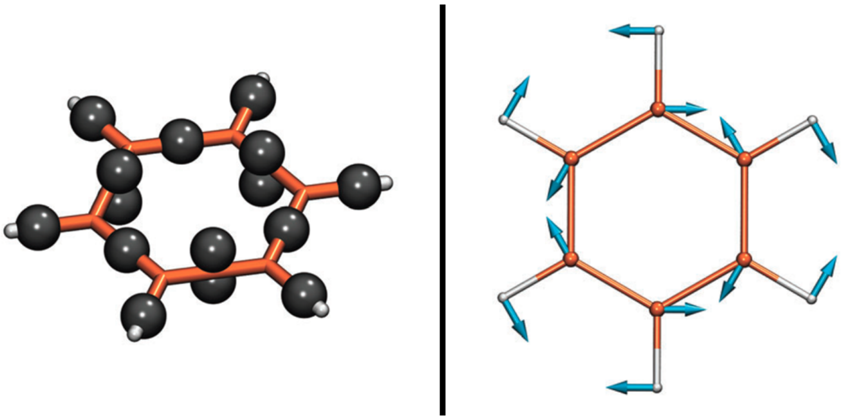


Figure 2.1.7.: Wannier localization for one benzene molecule leads to an alternating single bond/double bond pattern *(left panel)*; normal mode of benzene which deforms the molecule towards cyclohexatriene *(right panel)*.[MB23]

The situation becomes even worse if one tries to compute a Raman spectrum based on molecular polarizabilities resulting from Wannier localization and external field finite differences. Depending on the direction of the external field, the preferred localization of single bonds and double bonds in the ring differs, so that different such patterns can occur in the two calculations for the finite differences. This introduces an amount of noise which is actually so large that the Raman spectrum completely vanishes in the noise. For an example, see Figure 2.1.8, where the Raman spectrum of liquid benzene was predicted based on Wannier localization *(dashed curve)* and on Voronoi integration *(solid line)*. As described above, the Wannier-based spectrum consists almost exclusively of noise, while the Voronoi-based result gives a good prediction of the Raman spectrum.

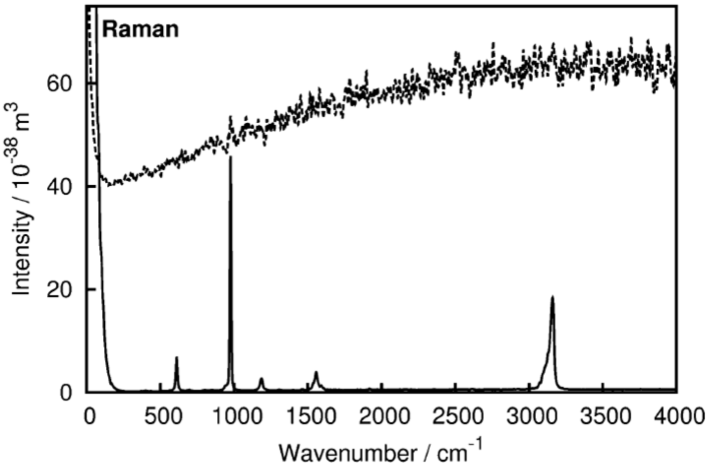


Figure 2.1.8.: Predicted Raman spectrum for liquid benzene based on Wannier localization *(dashed line)* and Voronoi integration *(solid line)*.[MB23] The Wannier-based prediction shows only noise.

Computing Higher Multipoles

Wannier localization is frequently used in the literature to compute molecular electric dipole moments. However, certain types of vibrational spectroscopy *(such as Raman optical activity—see Section 2.8)* require also the electric quadrupole moment.[68] To the best of our knowledge, it is not possible to compute higher molecular electric multipoles via Wannier localization. Based on the Voronoi integration, on the other hand, this is not an issue, as shown in Equation 2.8.1.

Conclusions

To conclude this section, we find the following four advantages of Voronoi integration over Wannier localization:

- Voronoi integration requires considerably less computer time for medium-sized and large systems. More than a factor of 2 can be saved.

- Iterative Wannier localization is not guaranteed to converge at all; Voronoi integration is non-iterative and always yields results.

- Wannier localization has severe issues with aromatic systems *(artificial bands in the infrared spectrum, large amounts of noise in the Raman spectrum)*, while Voronoi integration has no problems with such systems.

- In contrast to Wannier localization, Voronoi integration can also compute higher multipole moments such as the electric quadrupole tensor, which is, *e. g.*, required to predict ROA spectra.

2.2. Lossless Compression of Volumetric Data

Many of the methods for predicting vibrational spectra presented here rely on the total electron density on a three-dimensional grid along the simulation trajectory. Such volumetric data sets require significant storage space even for a single frame, and in particular if required along a full trajectory. To give an example, consider the Raman optical activity (ROA) spectrum shown in Figure 3.4.1 on page 93. The electron density grid has a resolution of 160 × 160 × 160, so that one frame requires 52 MiB storage space in Gaussian Cube file format.[245] To compute the spectrum, 4 × 65 000 such frames are necessary, leading to the enormous amount of 13 Terabyte raw data for this single spectrum. Such a storage requirement is completely unacceptable, even for large compute centers.

There are two possible ways how to overcome this difficulty. First, one could perform all necessary calculations *on the fly*, without storing the raw electron density data at all. However, this is problematic as long as some steps in the computational protocol are still in development, because any attempt to improve on one of these steps would require to re-run the whole simulation. Furthermore, the electron density contains much more information on the system than just the spectral intensities, so it might be worthwhile to be stored.

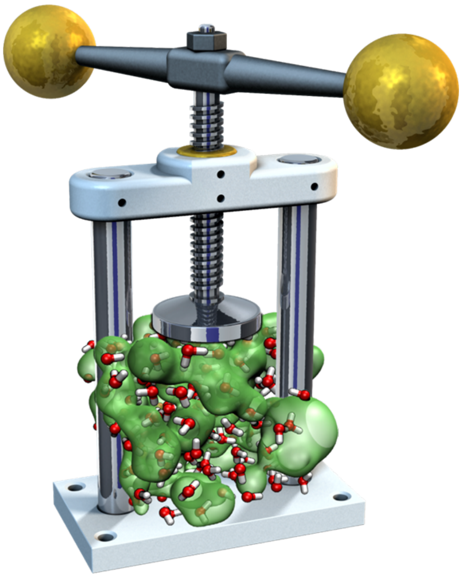


Figure 2.2.1.: Schematic illustration of compressing volumetric data.

The other way to overcome the issue would be to compress the volumetric data in order to reduce the required storage space. As long as the impact of limited accuracy on the spectra has not been studied in detail, it would be desirable to use a *lossless* compression algorithm here,[*] so that a degradation of the results' quality can be excluded. To achieve that purpose, we tried common lossless compression algorithms such as bzip2, which is able to reach a compression ratio of around 4.5 : 1 relatively to the Gaussian Cube file, so that the raw data still requires $\approx$ 3 Terabyte. But the main issue here is that the compression takes almost as much time as the AIMD simulations themselves. In terms of computational resource use, this is not an option. It seemed that there was no other option than to develop a tailor-made lossless compression algorithm which exploits the special structure and characteristics of volumetric electron density trajectories—namely, continuity both in space and time. The results of this endeavor are described in this section; they have been published in 2018.[MB34]

The basic principle of our approach is as follows. The continuity of the data allows to *extrapolate* the first unknown element based on all previously known elements. Such an extrapolation will never be exact. However, it opens the possibility to store the *residuals* (difference between the true value and the extrapolation) instead of the original data, which possess significantly less information entropy. In a second step, this set of residuals is compressed by an entropy encoding method such as Huffman coding to reduce its size. The details are described in the following two sections, concluded by a presentation of the obtained results.

Together with our compression algorithm, we have published a file format to store the data, which is called the bqb format *("binary cube")*. It is very size-efficient and robust *(due to checksums at multiple levels)*. The frames along the trajectory are stored independently in the file, but an index frame at the end allows for quick random access to certain frames, and allows to append to the bqb file without modifying the existing part of the file *(similar to the approach found in PDF files)*. More aspects of the newly developed file format are discussed in Section 2.2.3. Details on how to work with the bqb format and a free tool to compress and decompress trajectories *(both of atom position or volumetric data)* can be found on the following website:

`https://brehm-research.de/bqb`

---

[*] In the sense that the input data can be bit-wise restored without any changes. All numbers stored in computers are of limited precision already before compression.

2.2.1. Reducing Information Entropy via Extrapolation

The central idea of our lossless compression algorithm is to exploit the continuity of data resulting from molecular simulation. First, the simulation proceeds in relatively small discrete time steps, so that the change in any property between two subsequent time steps is often small and smooth. Secondly, volumetric data such as the total electron density cannot possess discontinuities *(gaps, jumps, etc.)*, and is therefore smooth in space. Both properties allow to *extrapolate* the next data element *(both in time and space)* based on all previously known elements. Such an extrapolation will never be exact. However, it opens the possibility to store the *residuals* (difference between the true value and the extrapolation) instead of the original data, which possess significantly less information entropy.

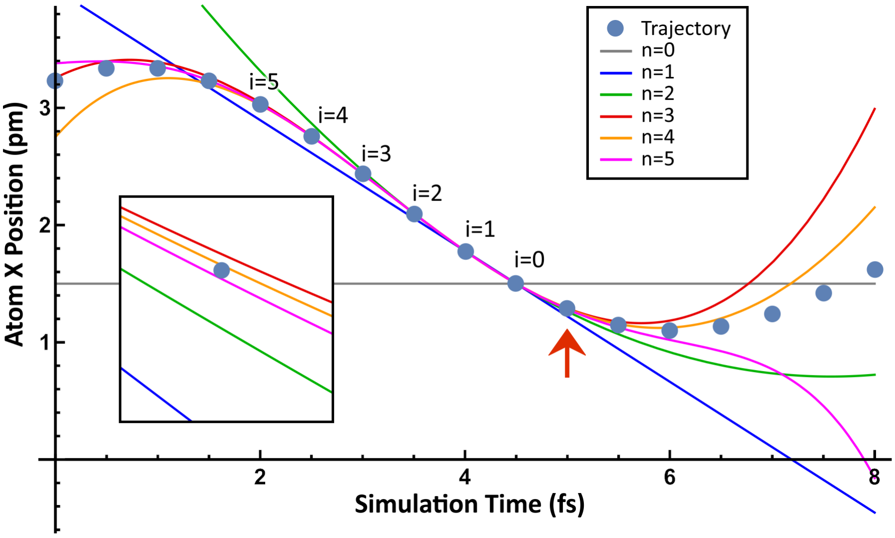


Figure 2.2.2.: Schematic illustration of extrapolation polynomials of different degree n (lines) fitted to the trajectory of one atom from a simulation (circles). Red arrow depicts position to be extrapolated.

An exemplary illustration of this concept is presented in Figure 2.2.2, where the X position of one selected atom along a MD simulation is shown as gray dots, with the simulation time on the horizontal axis. The value at "$i = 0$" is the last known position, and the value above the red arrow shall be extrapolated. Several different extrapolation polynomials which include terms up to different orders are optimized to match all known data points. The number of known data points taken into consideration depends on the degree of the polynomial. For example, the red curve with $n = 3$ is a cubic polynomial and is optimized to the last four previously known points. The inset shows how the quality of the extrapolation depends on the polynomial's degree. The best extrapolation is obtained with $n = 4$. Polynomials of both lower or higher degree lead to worse predictions. Which polynomial degree works best depends on the nature of the data set and should be determined before compression *(our algorithm does this automatically)*.[MB34]

In the case of volumetric data, extrapolation in three-dimensional space is required, which is slightly more involved. When the grid points are traversed in a row-by-row fashion, there is always one current point which shall be extrapolated and a set of already known points from the current as well as previous rows. A two-dimensional illustration of this situation is presented in Figure 2.2.3, where the red point shall be extrapolated, the blue and green points are already known, and the green points are used to optimize the multi-dimensional extrapolation polynomial.

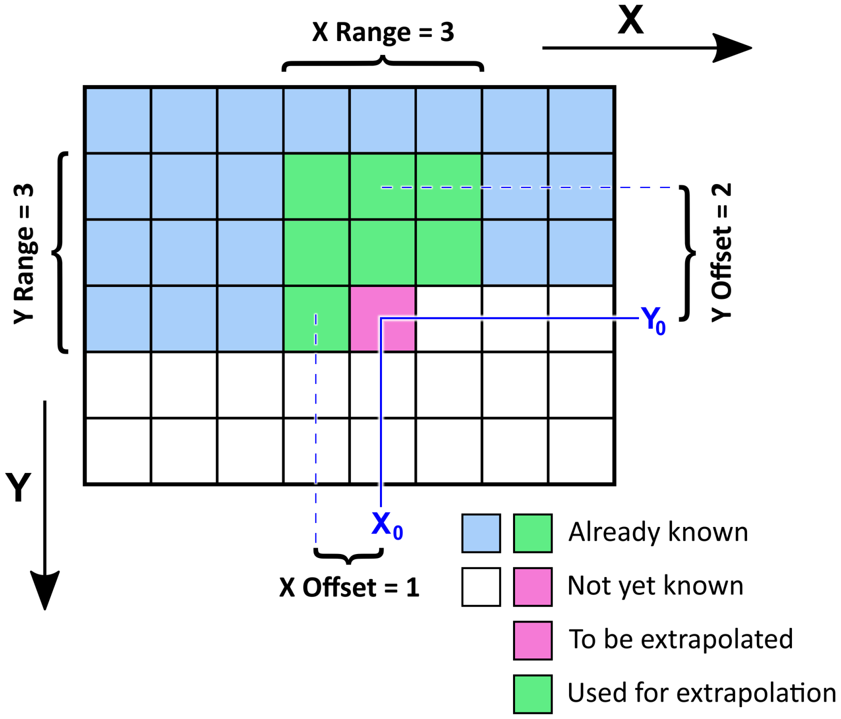


Figure 2.2.3.: Schematic two-dimensional illustration of grid points in the volumetric extrapolation.

To show the mathematical derivation on how to efficiently optimize such an extrapolation polynomial in the multi-dimensional case, let's consider an exemplary *ansatz*, which is a function F in the two variables x and y, defined as a polynomial up to second order (only in x):

$$F(x,y) = c_0 + c_x \cdot x + c_y \cdot y + c_{xy} \cdot xy + c_{x^2} \cdot x^2 \tag{2.2.1}$$

In order to choose the coefficient values c to fit the known data points *(green squares in Figure 2.2.3)* as close as possible, the following system of linear equations needs to be solved:

$$
\begin{aligned}
F(x_0-1, y_0-2) &= c_0 + c_x(x_0-1) + c_y(y_0-2) + c_{xy}(x_0-1)(y_0-2) + c_{x^2}(x_0-1)^2 \\
F(x_0, \quad y_0-2) &= c_0 + c_x\, x_0 + c_y(y_0-2) + c_{xy}\, x_0 \, (y_0-2) + c_{x^2}\, x_0^2 \\
F(x_0+1, y_0-2) &= c_0 + c_x(x_0+1) + c_y(y_0-2) + c_{xy}(x_0+1)(y_0-2) + c_{x^2}(x_0+1)^2 \\
F(x_0-1, y_0-1) &= c_0 + c_x(x_0-1) + c_y(y_0-1) + c_{xy}(x_0-1)(y_0-1) + c_{x^2}(x_0-1)^2 \\
F(x_0, \quad y_0-1) &= c_0 + c_x\, x_0 + c_y(y_0-1) + c_{xy}\, x_0 \, (y_0-1) + c_{x^2}\, x_0^2 \\
F(x_0+1, y_0-1) &= c_0 + c_x(x_0+1) + c_y(y_0-1) + c_{xy}(x_0+1)(y_0-1) + c_{x^2}(x_0+1)^2 \\
F(x_0-1, y_0) &= c_0 + c_x(x_0-1) + c_y\, y_0 + c_{xy}(x_0-1)\, y_0 + c_{x^2}(x_0-1)^2
\end{aligned}
\tag{2.2.2}
$$

Such a system of linear equations is typically *(formally)* solved by converting it to matrix form and then inverting the coefficient matrix A:

$$
\begin{aligned}
\mathbf{A} \cdot \mathbf{x} &= \mathbf{b} \\
\mathbf{x} &= \mathbf{A}^{-1} \cdot \mathbf{b}
\end{aligned}
\tag{2.2.3}
$$

By doing so with the system in Equation 2.2.2 and assuming without loss of generality that $x_0 = 0$ and $y_0 = 0$, we obtain

$$
\begin{pmatrix}
1 & -1 & -2 & 2 & 1 \\
1 & 0 & -2 & 0 & 0 \\
1 & 1 & -2 & -2 & 1 \\
1 & -1 & -1 & 1 & 1 \\
1 & 0 & -1 & 0 & 0 \\
1 & 1 & -1 & -1 & 1 \\
1 & -1 & 0 & 0 & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
c_0 \\ c_x \\ c_y \\ c_{xy} \\ c_{x^2}
\end{pmatrix}
=
\begin{pmatrix}
F(-1,-2) \\ F(0,-2) \\ F(1,-2) \\ F(-1,-1) \\ F(0,-1) \\ F(1,-1) \\ F(-1,0)
\end{pmatrix} .
\tag{2.2.4}
$$

As the coefficient matrix is a non-square matrix, it cannot be inverted, so that Equation 2.2.3 cannot be used to solve the system. This is the case because the system of equations is over-determined *(more equations than unknowns)*. Luckily, such a system can still be solved approximately *(least-squares solution)* by employing the so-called **Moore–Penrose inverse** $\mathbf{A}^+$ of $\mathbf{A}$.[246,247] The Moore–Penrose inverse of $\mathbf{A}$ is the matrix $\mathbf{A}^+$ which fulfills the following four relations:

$$
\begin{aligned}
\mathbf{A}\mathbf{A}^+\mathbf{A} &= \mathbf{A} \\
\mathbf{A}^+\mathbf{A}\mathbf{A}^+ &= \mathbf{A}^+ \\
\left(\mathbf{A}\mathbf{A}^+\right)^* &= \mathbf{A}\mathbf{A}^+ \\
\left(\mathbf{A}^+\mathbf{A}\right)^* &= \mathbf{A}^+\mathbf{A}
\end{aligned}
\tag{2.2.5}
$$

It can be shown that such a matrix $\mathbf{A}^+$ always exists and is uniquely defined for any $\mathbf{A} \in \mathbb{R}^{m \times n}$. In a certain sense, the Moore–Penrose inverse is the generalized matrix inverse. For matrices which possess and inverse, the Moore–Penrose inverse equals the traditional inverse.

Once $\mathbf{A}^+$ is known, the system from Equation 2.2.3 can be approximately solved via

$$\begin{aligned} \mathbf{A} \cdot \mathbf{x} &= \mathbf{b} \\ \mathbf{x} &= \mathbf{A}^+ \cdot \mathbf{b}. \end{aligned} \tag{2.2.6}$$

It remains the question how the Moore–Penrose inverse can be efficiently *(numerically)* computed in practice. This can be achieved by the **Singular Value Decomposition (SVD).**[248] Any real matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ can be written as the product

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}, \tag{2.2.7}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, $\boldsymbol{\Sigma}$ is a diagonal matrix *(with all non-diagonal elements being zero)*, and $\cdot^{\mathrm{T}}$ denotes the matrix transpose. The singular value decomposition can be efficiently and accurately computed with a numerical algorithm developed by Gene Golub *et al.*[249,250] Once the decomposition is known, the Moore–Penrose inverse $\mathbf{A}^+$ of $\mathbf{A}$ can be expressed as the simple product

$$\mathbf{A}^+ = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^{\mathrm{T}} \tag{2.2.8}$$

where $\boldsymbol{\Sigma}^+$ is formed by forming the reciprocal of all non-zero values of $\boldsymbol{\Sigma}$ and then transposing the resulting matrix *(see Equation 2.2.5)*.

Coming back to the example from Equation 2.2.4, the approximate least-squares solution *(determined by applying Equations 2.2.6 and 2.2.8)* reads

$$\begin{pmatrix} -0.29 & 0.14 & -0.57 & 0 & 0.86 & 0.57 & 0.29 \\ 0.14 & -0.24 & -0.38 & -0.17 & 0.24 & 0.88 & -0.48 \\ -0.19 & -0.24 & -0.38 & 0 & 0.24 & 0.38 & 0.19 \\ 0.29 & -0.14 & -0.43 & 0 & 0.14 & 0.43 & -0.29 \\ 0.29 & -0.48 & 0.24 & 0.27 & -0.52 & 0.26 & 0.048 \end{pmatrix} \cdot \begin{pmatrix} F(-1,-2) \\ F(0,-2) \\ F(1,-2) \\ F(-1,-1) \\ F(0,-1) \\ F(1,-1) \\ F(-1,0) \end{pmatrix} = \begin{pmatrix} c_0 \\ c_x \\ c_y \\ c_{xy} \\ c_{x^2} \end{pmatrix}. \tag{2.2.9}$$

As we only want to extrapolate the value $F(0,0)$, we do not need to compute the coefficients c of the polynomial at all. We can simply write

$$\begin{aligned} F(0,0) &= c_0 + c_x \cdot 0 + c_y \cdot 0 + c_{xy} \cdot 0 + c_{x^2} \cdot 0^2 \\ &= c_0, \end{aligned} \tag{2.2.10}$$

and the value $F(0,0)$ is therefore obtained as a simple dot product

$$F(0,0) = c_0 = \begin{pmatrix} -0.29 & 0.14 & -0.57 & 0 & 0.86 & 0.57 & 0.29 \end{pmatrix} \cdot \begin{pmatrix} F(-1,-2) \\ F(0,-2) \\ F(1,-2) \\ F(-1,-1) \\ F(0,-1) \\ F(1,-1) \\ F(-1,0) \end{pmatrix}. \tag{2.2.11}$$

This means that the full process of fitting a multi-dimensional extrapolation polynomial to the known data points *(multi-dimensional non-linear regression)* and then using this polynomial to extrapolate the first unknown point has been broken down to a dot product between two vectors. The first vector contains coefficients which only depend on the geometric position of the known data points relatively to the point which shall be extrapolated. These coefficients are identical for all such extrapolations and can therefore be pre-computed before the run. The second vector contains the values of the known data points. This approach is very computationally efficient and allows to perform millions of polynomial regressions and extrapolations *(as required for compressing volumetric data)* within milliseconds.

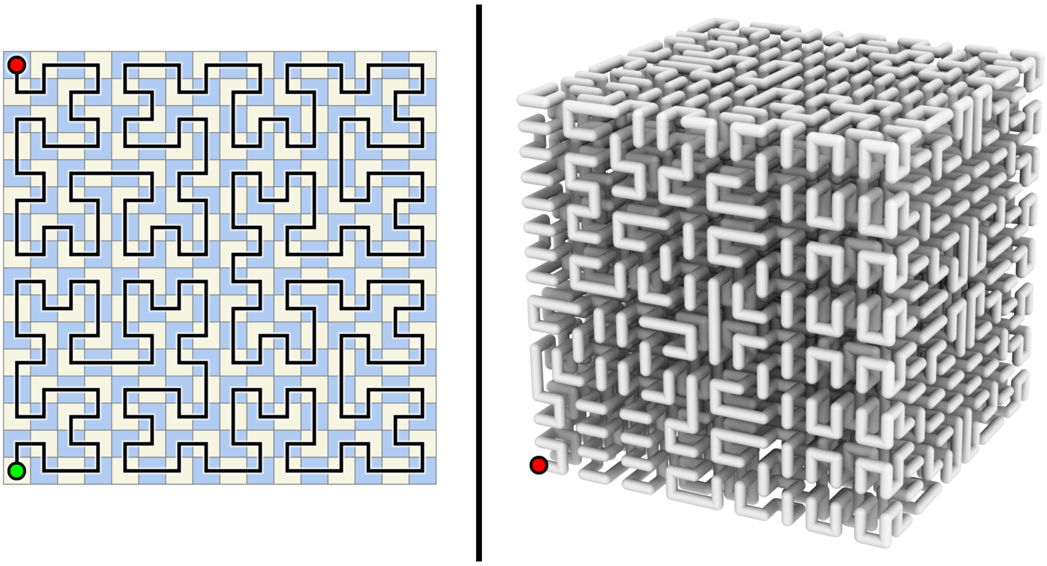


Figure 2.2.4.: Illustration of grid traversal by a space-filling Hilbert curve in two dimensions *(left panel)* and three dimensions *(right panel)*; the locality of neighboring grid elements is preserved.[MB34]

To allow for a high compression ratio in the entropy encoding algorithm, *data locality* in the sequence of residuals is very important. In a three-dimensional grid, locality means that neighboring grid entries are also close to each other in the one-dimensional sequence of residuals. When traversing the grid in canonical order (with three nested loops for X, Y, Z index), locality is very bad. Neighbors in X direction are still neighbors in the sequence, but neighbors in Y and especially Z direction are very far separated from each other. To overcome this, non-canonical traversal schemes can be applied to preserve some degree of locality. A good choice in such a case is the space-filling Hilbert curve.[251] It is guaranteed to touch every grid point exactly once (and therefore is a valid traversal path), and it preserves locality. Two illustrations of such a Hilbert curve are presented in Figure 2.2.4 in two dimensions *(left panel)* and three dimensions *(right panel)*. It is visible that neighboring grid points are often *(not always)* close to each other on the traversal path. Our implementation uses a fast algorithm (without recursion) to construct three-dimensional Hilbert curves.[252] In a typical example application *(as presented in Figure 2.2.8)*, the Hilbert curve traversal reduces the output file size by around 10 %.[MB34]

2.2.2. Entropy Encoding

The sequence of transformations presented below is inspired by the bzip2 compression tool,[253] which was developed by Julian Seward. The source code of bzip2 is freely available in the internet under a BSD-like license. No parts of the bzip2 source code have been copied when the implementation presented here was written; only the general concept is similar. The following discussion closely follows our original article.[MB34]

Alphabet Creation

The first step in our compression approach is the creation of an alphabet. The use of an alphabet transforms the symbol space from all possible integers to a consecutive range of non-negative integers. If n different symbols are present in the input integer stream, then the output stream will contain the values $0 \ldots (n-1)$. The alphabet is created in such a way that the symbols are sorted in ascending order. This means that a zero in the output stream corresponds to the smallest integer from the input stream, and a $(n-1)$ corresponds to the largest integer. To accelerate the transformation of the input stream by the alphabet (especially for very large n), a hash table is utilized.

Burrows–Wheeler Transformation

After the alphabet has been applied, the resulting list of consecutive integer numbers can be treated with a Burrows–Wheeler transformation.[254] This is optional and can be switched on/off by the user. The Burrows–Wheeler transformation is the central part of bzip2.

A descriptive explanation of the Burrows–Wheeler transformation is given in Figure 2.2.5, and explained in the following. The input data is a list of symbols, in this example it is `^ANANAS$` *("Ananas" is the German term for pineapple)*. All cyclic permutations of the input sequence are created (*i. e.*, for all n, shift the sequence to the right n places, such that symbols which fall out on the right side enter on the left side). The list of all these permutations is then sorted lexicographically *(i. e., comparing the first symbol of both sequences, and if equal, go to the next symbol)*. From the sorted list of permutations, the last column is extracted as the output sequence. It becomes clear that all the symbols from the input sequence also appear in the output sequence, but in a different order. This is the essence of the Burrows–Wheeler transformation: it re-orders the sequence of symbols such that equal symbols have a higher probability to follow after each other *(depending on the structure of the input data)*. Additionally, the index of the original *(non-permuted)* input sequence in the sorted list needs to be stored, which is 7 in this example.

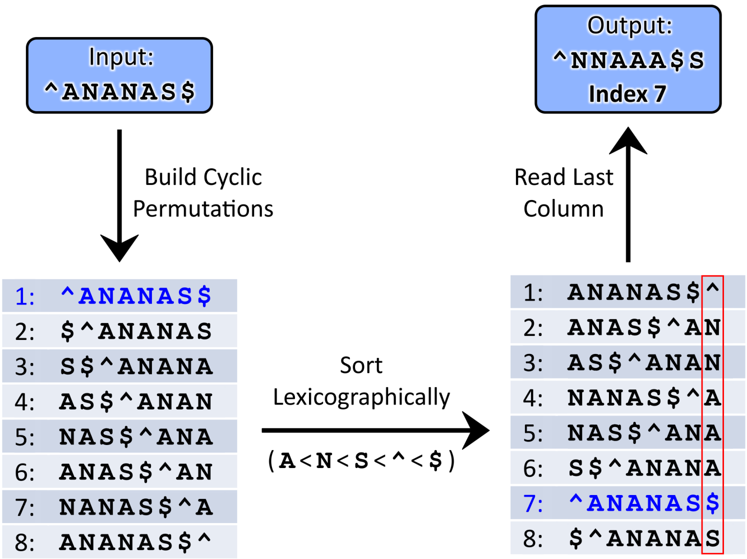
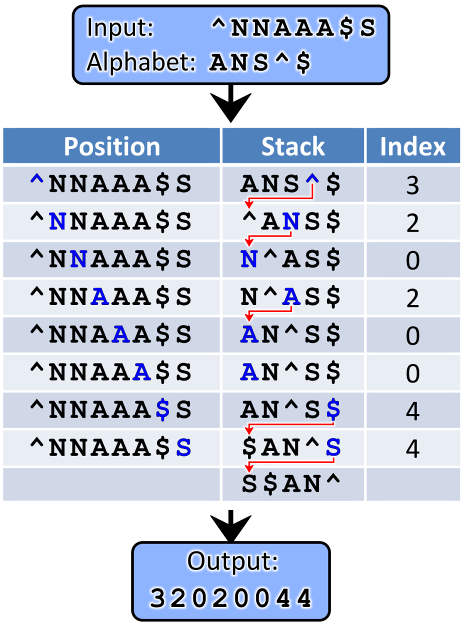

Input:
^ANANAS$
Build Cyclic Permutations
1: ^ANANAS$
2: $^ANANAS
3: S$^ANANA
4: AS$^ANAN
5: NAS$^ANA
6: ANAS$^AN
7: NANAS$^A
8: ANANAS$^
Sort Lexicographically
(A<N<S<^<$)
1: ANANAS$^
2: ANAS$^AN
3: AS$^ANAN
4: NANAS$^A
5: NAS$^ANA
6: S$^ANANA
7: ^ANANAS$
8: $^ANANAS
Read Last Column
Output:
^NNAAA$S
Index 7

Figure 2.2.5.: Descriptive example of the Burrows–Wheeler transformation.

An efficient implementation is, however, more involved than the description given above. For long sequences of symbols, it would require a large amount of memory to explicitly construct all permutations. Instead of doing so, the sorting algorithm can directly operate on the input sequence *(periodically wrapped around at the end)*, which does not require any additional memory and is very efficient in terms of cache utilization. Furthermore, parts of the sequences which consist of only one repeated symbol are detected and handled separately, as those would significantly reduce the efficiency of sorting algorithms such as Quicksort.[255] Even for large input sequences of $> 10^7$ symbols, this algorithm only takes a few seconds on a single modern CPU core.

Move-To-Front Transformation

Directly after a Burrows–Wheeler transformation, often a move-to-front transformation[256] follows. While the aim of the Burrows–Wheeler transformation is to bring as many identical symbols as possible into consecutive sequences, the aim of the move-to-front transformation is to replace these sequences by runs of zeros. In the implementation presented here, the move-to-front transformation is optional and can be switched on/off by the user. A descriptive example of the move-to-front transformation is given in Figure 2.2.6.



Input: ^NNAAA$S
Alphabet: ANS^$

Position	Stack	Index
^NNAAA$S	ANS^$	3
^NNAAA$S	^ANS$	2
^NNAAA$S	N^AS$	0
^NNAAA$S	N^AS$	2
^NNAAA$S	AN^S$	0
^NNAAA$S	AN^S$	0
^NNAAA$S	AN^S$	4
^NNAAA$S	$AN^S	4
	S$AN^	

Output:
32020044

Figure 2.2.6.: Descriptive example of the Move-to-Front transformation. Input data is taken from the output of Figure 2.2.5.

The input data for the transformation given as a sequence of symbols, together with an alphabet. The alphabet is copied to a temporary stack in some well-defined order *(e. g., sorted by symbol values)*. To create the output data, the input sequence is read starting from the beginning, one symbol at a time. The index at which the input symbol is currently found on the alphabet stack is pushed to the output. Subsequently, the symbol on the stack is deleted at its original place and instead inserted at the front position ("move-to-front"), shifting all other symbols one position further down the stack. This process is repeated for all input symbols, as shown in the figure. It becomes clear that sequences of identical symbols lead to sequences of zeros in the output data, because this symbol is already found at the front of the stack when it is encountered for the second time.

Run-Length Encoding

The final step after a Burrows–Wheeler transform and a move-to-front transform is often a run-length encoding[257] for zero runs. As the aim of Burrows–Wheeler and move-to-front is to create such zero runs, this is a reasonable choice. In the implementation presented here, the run-length encoding is optional and can be switched on/off by the user.

In bzip2 as well as our implementation presented here, the run-length encoding works as follows. The alphabet contains two additional symbols, which are typically termed as RUNA and RUNB, and which do not appear in the input data. Every occurrence of one or more zeros in the input data is replaced by sequences of RUNA and RUNB, which encode the number of consecutive zeros. Such a sequence ends when another symbol apart from RUNA and RUNB is found. The encoding works similar to base-2 bijective numeration. Let $x_i \in \{\text{RUNA}, \text{RUNB}\}, i = 1 \ldots n$ be a sequence of n RUNA and RUNB symbols. A RUNA symbol at index i in this sequence has a value of 2^{i-1}, a RUNB symbol at this index has a value of 2^i. All those symbol values in the sequence are added up and give the number of consecutive zeros that were originally found in the input data. For example, an isolated RUNA symbol encodes a single zero, whereas the sequence RUNA, RUNB, RUNA, RUNA, RUNB encodes $1 + 4 + 4 + 8 + 32 = 49$ consecutive zeros.

Paired with an ideal entropy encoding algorithm *(which truly maximizes the entropy)*, run-length encoding would be of no use. However, in the general case there exists no perfect entropy encoder if single symbols are to be encoded. The compression efficiency of the Huffman coding described below suffers if the frequency of a symbol *(typically of "zero" if move-to-front was used)* becomes very high. This is mitigated by applying run-length encoding before.

Huffman Coding

The most important step for size reduction is the Huffman coding.[258] Both bzip2 and our implementation presented herein strongly profit from an advanced multi-table Huffman coding. Before this will be covered below, Huffman coding with a single table will be explained.

Huffman coding is a widely used entropy encoding method (*e. g.*, in the JPEG image format and MP3 audio format). It aims at reducing the size of a stream of symbols by maximizing the entropy. This is achieved by creating a variable-length prefix code where each symbol from the input corresponds to a variable-length sequence of bits. According to Shannon's source coding theorem,[259] such a code is optimal if the length of each bit sequence is approximately proportional to $-\log_2(p)$,[260] where p is the probability of finding the corresponding symbol within the input sequence. In simple words, the symbols which appear most often will have the shortest corresponding bit sequences. This is not always strictly fulfilled, but it can be shown that Huffman coding produces an optimal code among all methods that encode each symbol separately.[261]

An efficient algorithm to construct a Huffman coding table makes use of a binary tree structure. First, a list of 2–tuples is created where each tuple contains one symbol from the alphabet and the frequency of that symbol, *i. e.*, the total count of this symbol in the input data. This list is sorted in ascending order by frequency, such that it starts with the least frequent symbols. The first two tuples are taken and removed from the list. A new tree node is created, with these two tuples as children. The new tree node is also represented as tuple, with its frequency being the sum of the two children's frequency. This newly created node is now inserted into the list of tuples on the correct position, such that the list remains sorted with respect to the tuple frequencies. Then, the process is repeated by taking and removing the first two tuples from the list, and so on. The algorithm stops if there is only one tuple left in the list. This is the root of a binary tree. All symbols from the alphabet represent leaves in this tree. To find the Huffman code of a given symbol, the tree is traversed starting from the root until the leaf is reached which corresponds to this symbol. During this traversal, a "0" bit is produced each time the left child of a node is entered, and a "1" bit each time the right child of a node is accessed. The resulting sequence of bits is the Huffman code word for the symbol in the leaf. Following from this, the length of the code word for some symbol is equivalent to the depth of this symbol within the binary tree.

Decoding of Huffman-encoded data is simple if the binary tree is known *(or can be reconstructed)*. Then, bits are taken from the input bit stream, and depending on the bit value, the left or right child of a tree node is entered, starting at the root node. As soon as a leaf is encountered, the end of the code word is reached *(due to the prefix property of the code)*. The leaf's symbol value is written to the output, and the position in the tree is reset to the root node. However, it is inefficient to store the whole Huffman tree in the compressed file. To reconstruct the tree at decode time, the exact symbol frequencies would be required, and their storage would also reduce efficiency. A more efficient solution to store the Huffman table is the so-called canonical Huffman code, which is used in both bzip2 and our implementation.

Canonical Huffman Coding

Canonical Huffman coding[262] takes the original Huffman code (*i. e.*, a table which assigns a code word of bits to each symbol) created above as an input, but subsequently modifies the code. This modification is performed as follows. The Huffman code table is sorted first in ascending order by code word length, and secondly by symbol. All code words preserve their bit lengths, but the bits are overwritten. The first code word from the sorted list is overwritten by all zero bits *(keeping the length)*. Subsequent code words are overwritten by the next binary number in sequence, such that the code word is larger in value than all preceding code words. If a code word with larger bit length is reached, the required amount of zero bits is appended at the right side *after* increasing the binary number. This procedure yields a canonical Huffman code table. All symbols are encoded by code words with the same bit length as in the non-canonical Huffman code, and therefore, the compression efficiency is not influenced. However, the canonical Huffman table itself can be stored with much less information, as explained in the following.

For storage, the canonical Huffman table is first sorted in ascending order by symbol, so that the code words of the symbols will be stored in alphabetical order. For each symbol, one requires to store three pieces of information: the symbol itself, the bit length of the code word, and the code word itself. However, as the symbols are stored in alphabetical order, the symbol itself can be omitted. As this is a canonical Huffman code, we have the additional information that an *(alphabetically)* later code word will always be higher in value than an earlier one of the same bit length. Taking this together, it is only necessary to store the bit length of each symbol's code word in order to unambiguously reconstruct the canonical Huffman table. This is a large saving when compared to the storage requirements of a non-canonical Huffman tree.

Multi-Table Huffman Coding

Huffman coding possesses disadvantages when the input data features some degree of locality. A symbol which infrequently occurs in the complete input data will be assigned to a long code word. If there is a short block in the input data with very high frequency of this symbol, nothing will change. One way to take advantage of such situations is to use multiple Huffman tables. The input data is separated into short blocks, and each such block is encoded with the Huffman table that leads to the smallest output size. This raises the question how to construct different Huffman tables such that many blocks can be encoded efficiently with one of them. In the bzip2 source code, an elegant solution is found, which will be explained in the following.

As an example, three independent Huffman tables shall be constructed for an input data stream with known symbol frequencies. In the beginning, the three Huffman tables are constructed in such a way that each symbol is contained in exactly one Huffman table, and each Huffman table contains symbols which account for approximately $\frac{1}{3}$ of the total input data. To achieve this, the alphabet is sorted in descending order by symbol frequency, so that the most frequent symbol is found in the beginning of the list. Then, symbols from the beginning of the list are assigned to the first Huffman table, until the sum of symbol frequencies exceeds $\frac{1}{3}$ of the total symbol count. The following symbols are assigned to the second table, until the sum of symbol frequencies also in this table exceeds $\frac{1}{3}$ of the total symbol count. All remaining symbols are assigned to the third table. The three Huffman tables are now constructed, based on the symbols assigned to them and their respective frequencies.

The input data is separated into small blocks. For each block, it is decided which of the three Huffman tables is most efficient in encoding the block data. If a symbol from the block is already contained in the Huffman table, it is clear how many bits its code word would occupy. If a symbol is not yet contained in this table, it is computed how much additional storage would be required by adding this symbol to the table *(including the effect on all preceding blocks that have been encoded with this table)*. After all blocks of the input data have been processed, each of the blocks is assigned to one of the Huffman tables, and the Huffman tables were augmented with many additional symbols *(many symbols now appear in more than one Huffman table)*. The whole process is iteratively repeated, starting again from the first input data block, and deciding for each block which table would be most efficient to encode it. The total required storage size reduces with each iteration. Specialized Huffman tables will form, which are very efficient for specific types of data blocks. The iteration is stopped if the total size no longer decreases or if a maximum cycle count is reached.

2.2.3. The bqb File Format

The output data from the compression algorithm is a stream of bits. Such a raw sequence is impractical to handle, and needs to be encapsulated in a suitable file format for practical use. To this aim, we have developed the bqb file format.[MB34] Initially, it was considered as "binary cube" file format *(therefore the name bqb)*, but in the end it was designed as a multi-purpose file format which can carry many different kinds of payload data related to trajectories. It does not offer the degree of flexibility known from formats such as HDF5; however, in contrast to HDF5, it considers a high compression ratio as central design goal. Even the headers and control structures are compressed, and the format is built as bit stream *(i. e., not respecting byte boundaries)*, so that not a single bit of storage space is wasted.

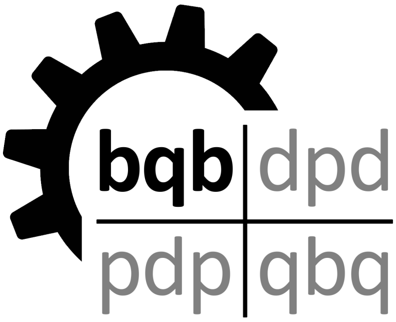


Figure 2.2.7.: The logo of the bqb file format.

The bqb file format is defined on a per-frame basis. A bqb frame starts with a frame header, which contains a "magic number", followed by a frame type specifier, the total payload size in bytes, and a CRC-32 code[263] of the frame payload. Depending on the frame type, a frame can either be a "short frame" or a "long frame". A long frame can contain many different chunks of data, which can encode different types of information in one single frame. A short frame contains a fixed type of data for a specific application. There exist, *e. g.*, short frames for storing compressed position data *(when a position trajectory is compressed)* and for storing both position and volumetric data in the same frame *(when a volumetric trajectory is to be compressed)*. Both frame formats can optionally contain information on the *(possibly non-orthorhombic)* cell vectors and the atom types in the system, which often are not required in every single frame. Apart from that, also a short frame format for compressing arbitrary binary files is available *(similar to bzip2)*.

Most bqb files possess a special frame type at the end of the file, which is an index frame. Index frames in the middle of a bqb file are simply ignored. This brings the following advantage. If additional frames are written to an existing bqb file, an updated index frame is simply appended at the end, and the existing file data is not modified in any way. The index frame contains a list of all preceding frames in the file, together with their respective frame types and file offsets. This enables fast random access to arbitrary frames. Such an index frame is optional, it can be left out if it is not required. An application can probe the existence of an index frame by considering the last bytes of a bqb file. Index frames end with a special signature. If this signature is found, then the length of the index frame is stored in the bytes directly in front of this signature, so that the start of the index frame can be reconstructed without scanning the whole file. This index concept is inspired by the PDF document file format, where it is called a “catalog” and has proven to be suitable for many applications.

A bqb trajectory is simply a sequence of bqb frames. No data structures exist above the bqb frame level. This means that several bqb frames / trajectories can be concatenated by simply concatenating the underlying raw files. If required, a new index frame can be created for the concatenated file and appended at the very end. bqb files can be also concatenated by creating a so-called list file, which is a text file that contains absolute file names to bqb files in each of its rows. The access to the individual bqb files is handled transparently, such that the list file behaves as a standard bqb file.

When trajectory data is compressed with polynomial extrapolation, the exact data from the previous n frames is required to decompress one frame. This means that to decompress a single frame in the middle of the bqb file, one needs to start at the first frame, and decompress all frames up to the desired frame in sequential order. In many applications, trajectories are read sequentially anyway, and this might not be a concern. If random access to decompressed frames is desired, one can take a compromise: every, *e. g.*, 100 frames, the temporal polynomial extrapolation order is reset to zero, so that one frame appears which is not dependent on any preceding frame. All following frames increase the extrapolation order by 1 each, up to the desired global extrapolation order. This concept is very similar to “key frames” in modern video encoding. If a random frame shall be decompressed, only 99 preceding frames need to be decompressed first in the worst case. The key frame interval can be adapted to the specific application. Key frames possess a special frame type, and can therefore be easily identified in the bqb index.

For more details on the bqb file format as well as tools to compress and decompress trajectories, please visit

`https://brehm-research.de/bqb`

2.2.4. Results

To benchmark the performance of our newly developed lossless compression algorithm for volumetric data trajectories,[MB34] we applied it to a simulation of the ionic liquid 1-ethyl-3-methylimidazolium acetate,[MB8,MB13,MB19,MB31] also known as [EMIm][OAc], for which we stored the total electron density on a grid with $216 \times 216 \times 216$ grid points in each simulation step ($\Delta t = 0.5\,\mathrm{fs}$). In Figure 2.2.8, we compare our method to several other existing storage formats. "Cube" corresponds to the Gaussian Cube file format,[245] which is a simple text format for volumetric data with 5 significant digits. "FP32" is the hypothetical size of storing every grid entry in a 32 bit *(single precision)* floating point number, such as performed in the gOpenMol[264] PLT file format. "bzip2" and "xz" are command-line compression utilities,[253,265] which were used in highest compression settings to compress the Cube files *(as xz is very slow, this took many days to complete)*. "bqb" depicts the resulting size of the bqb trajectory, obtained with the algorithms presented here. All data was averaged over 1 000 volumetric data frames. In all formats compared here, 5 significant digits were stored for the volumetric data. Strictly speaking, this is a *lossy* storage. However, as the original Cube files written by the electron structure program also have 5 significant digits, the compression is lossless in the sense that it can bit-wise reproduce the original input data.

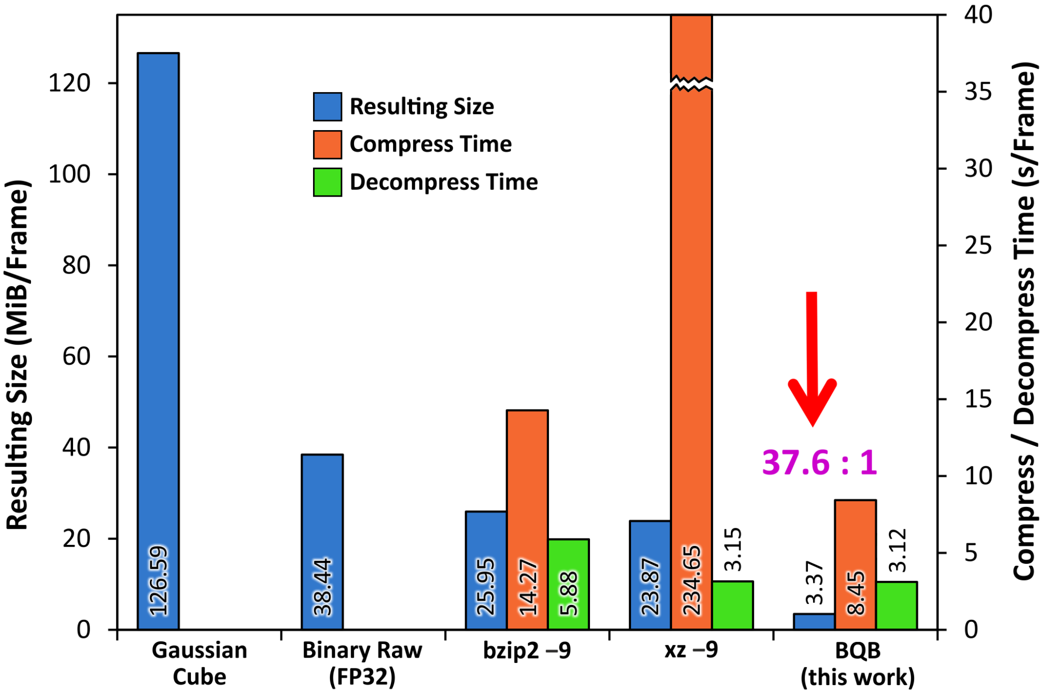


Figure 2.2.8.: Comparison of compression ratio and timing for a typical volumetric data trajectory *(liquid phase simulation,* $216 \times 216 \times 216$ *grid points)* with $\Delta t = 0.5\,\mathrm{fs}$ and 5 significant digits. Averaged over 1 000 frames. The algorithm presented here is on the far right.[MB34]

The blue bars in Figure 2.2.8 depict the compressed average frame size in MiB, while the orange and green bars correspond to the average computer time required to compress and decompress one frame, respectively. At first sight, the very high compression rate of the bqb format can be seen, reaching almost 40 : 1. The other formats can only achieve around 5 : 1 here. Even for single volumetric data frames *(where no temporal continuity can be exploited)*, a compression ratio of around 20 : 1 is still reached. Apart from the by far best compression ratio, the newly developed format also has the shortest compression and decompression times, which are 8.5 s and 3.1 s per frame, respectively *(all measured on a single CPU core "Intel Xeon E5–2609" at 2.5 GHz)*. When considering the size of an uncompressed frame in Gaussian Cube format, this corresponds to a compression and decompression data rate of 15.0 MiB s^{-1} and 40.6 MiB s^{-1}, respectively, on a single core.

To come back to the initial question described in the beginning of Section 2.2, the 13 Terabyte of raw data required to compute a single ROA spectrum can be compressed to just 350 Gigabyte of data in bqb format without losing any information. Thus, the challenge is completely solved now, as it is not an issue to store a few hundred Gigabytes of raw data.

Position Trajectories

After having successfully applied the extrapolation-based approach to compressing volumetric data trajectories, we considered to utilize the same approach to the lossless compression of standard position trajectories. And indeed it turned out that a high compression ratio can be reached also for this type of input data. As above, we have applied the method to a simulation trajectory of the ionic liquid[266] 1-ethyl-3-methylimidazolium acetate,[MB8,MB13,MB19,MB31] also known as [EMIm][OAc], which contains around 1 000 atoms, and the frames are stored with a time step of $\Delta t = 0.5$ fs. The results are presented in Figure 2.2.9. As the position extrapolation is carried out independently for each atom, the compression ratio is almost independent on the total atom count, so that the system size is not of importance here. "XYZ" is a simple text file format[267] which contains one atom position per row; no unnecessary characters *(multiple whitespaces, more digits than required, etc.)* have been written to ensure the smallest possible file size. "FP32" refers to storing atom coordinates in single-precision floating point numbers (*i. e.*, 32 bits per coordinate). "bzip2" and "xz" refer to command line compression utilities,[253,265] which have been used *(in highest compression settings)* to compress the XYZ text file. Please note that "xz" gives high compression ratio, but is extremely slow at compressing data. "XTC" is a binary position trajectory format used by the Gromacs program package.[268–270] "bqb" corresponds to the compressed files created with the algorithms described herein. All numbers are averaged over 1 000 consecutive trajectory frames.

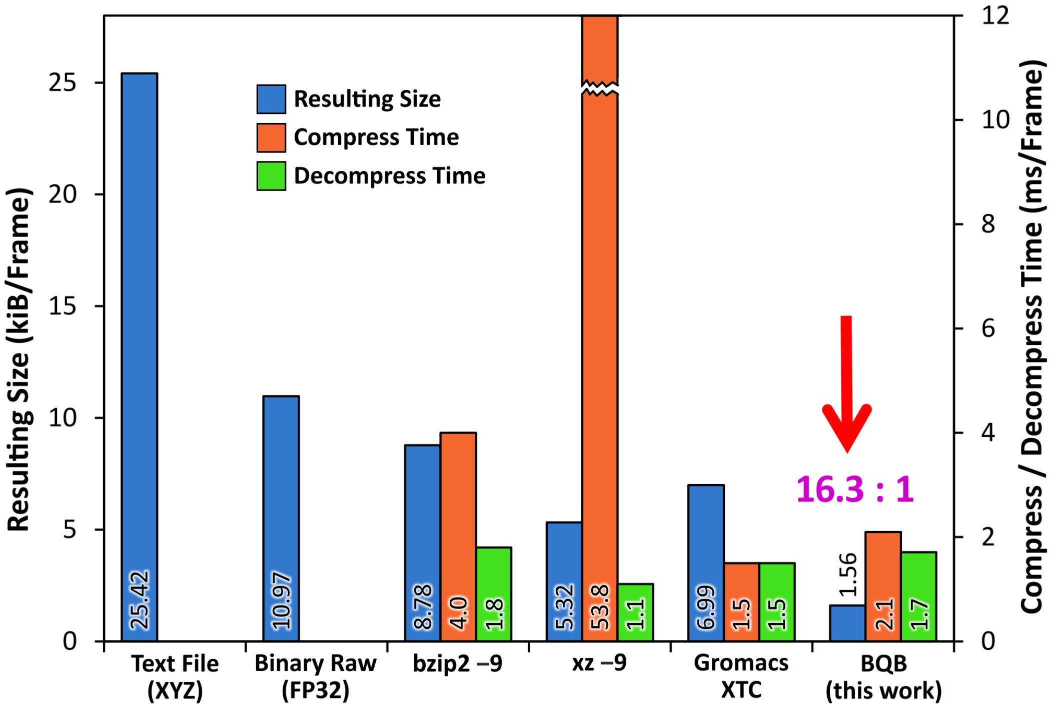


Figure 2.2.9.: Comparison of compression ratio and timing for a typical position trajectory *([EMIm][OAc] trajectory, 936 atoms)* with $\Delta t = 0.5\,\mathrm{fs}$ and a precision of 10^{-3} pm. Averaged over 1 000 frames. The algorithm presented here is on the far right.[MB34]

The blue bars in Figure 2.2.9 indicate the average compressed frame size in kiB, while the orange and green bars depict the average computer time required to compress and decompress one frame, respectively. At first sight, the very high compression rate of the bqb format can be seen, reaching 16 : 1. The other formats can only achieve around 5 : 1 here. Apart from the best compression ratio, the bqb file format also has very reasonable compression and decompression times of 2.1 ms and 1.7 ms per frame, respectively *(measured on a single CPU core "Intel Xeon E5–2609" at 2.5 GHz)*. This corresponds to a compression and decompression data rate of around $12\,\mathrm{MiB\,s^{-1}}$ on a single core.

To obtain more insight into the dependence of the compression ratio on the trajectory stride and coordinate accuracy, some further benchmarks have been performed, see Table 2.2.1. All tests have been performed with the same liquid phase [EMIm][OAc] simulation trajectory which was also used in Figures 2.2.8 and 2.2.9. All numbers are averaged over 1 000 trajectory frames. The first column in the table depicts the trajectory stride, *i. e.*, the temporal distance between two successive frames. "shuffle" means that the average was computed over 1 000 randomly

Table 2.2.1.: Comparison of compressed size of [EMIm][OAc] trajectory *(936 atoms)* for different trajectory strides Δt and coordinate precisions, averaged over 1 000 steps each. "shuffle" means randomly selecting 1 000 time steps from a long trajectory. Last column depicts XYZ to bqb compression ratio.[MB34]

Δt (fs)	Precision (pm)	Size (kiB/Frame)					bqb Ratio
		XYZ	bzip2 –9	xz –9	XTC	bqb	
0.5	1	17.13	3.19	1.15	3.59	0.52	32.94
	0.1	19.87	5.63	2.76	4.73	0.78	25.47
	0.01	22.61	7.12	4.06	5.89	1.11	20.37
	0.001	25.36	8.78	5.32	6.99	1.56	16.26
1.0	0.001	25.35	8.83	5.81	7.00	2.69	9.42
2.0		25.36	8.85	6.42	7.00	4.73	5.36
4.0		25.37	8.87	7.32	7.00	5.58	4.55
10.0		25.39	8.88	8.58	7.00	6.07	4.18
40.0		25.56	8.91	9.52	7.00	6.79	3.76
shuffle		25.56	8.95	9.68	7.00	7.96	3.21

selected frames *(in random order)* from a very long simulation trajectory, so that no continuity in time exists which might be exploited by the compression algorithm. In the second column, the absolute numerical precision of the atomic coordinates is specified. The following columns depict the average frame size for the different formats in kiB, while the last column indicates the compression ratio of the bqb format when compared to the XYZ format.

For the XYZ and XTC formats, it can be seen that the frame size depends on the number of decimal places stored for the coordinates, but not on the trajectory stride, as the frames are stored independently of each other. The bzip2 format shows only a slight dependence of the trajectory stride. Interestingly, this effect is significantly larger for xz, despite both are general compression algorithms and don't recognize any temporal continuity. The strongest dependence on the trajectory stride is found for the bqb format, which particularly aims at exploiting the temporal continuity. For a typical trajectory with $\Delta t = 0.5\,\text{fs}$ and 10^{-3} pm precision, a lossless compression ratio of 16 : 1 with respect to XYZ format is reached, which is almost a factor 4 smaller than the closest competitor *(xz compression)*. With increasing Δt, the compression ratio of the bqb format is significantly decreased, but even for $\Delta t = 40\,\text{fs}$, it is still the most size-effective format among all competitors. Only for the "shuffle" case with randomly arranged frames, the Gromacs XTC format beats bqb by a few percent.

It can be concluded that the bqb format is a very efficient lossless storage format also for position trajectories as long as some temporal continuity is present *(as it always is the case in molecular dynamics simulations)*.

2.3. Power Spectrum

The concept of the power spectrum originally stems from signal processing, where a power spectrum can identify how much *(electrical)* power is contained in individual frequency ranges of a signal. Such a power spectrum can also be computed for dynamical systems such as particles in motion, as it is known in the literature for a long time,[271–274] where it is also termed as velocity spectrum[275] or vibrational density of states (vDOS).[276] In this context, a power spectrum captures all motions of the system, but it is not bound to certain selection rules, so that the intensity of each spectral band only depends on the amplitude of a particular motion, but not on the change of some electromagnetic quantity *(such as dipole moment)*. Therefore, there exists no experimental analogue, and a power spectrum of molecular vibrations cannot be measured. Despite this, power spectra can give useful information on atomistic simulations, as it will be discussed below.

One possible definition of a power spectrum for a one-dimensional oscillator reads

$$P(\omega) = m\omega^2 \int \Big\langle \mathbf{r}(\tau) \cdot \mathbf{r}(t+\tau) \Big\rangle_\tau \, \exp\big(-\mathrm{i}\omega t\big) \, \mathrm{d}t, \tag{2.3.1}$$

where ω is the angular frequency, m is the reduced mass of the oscillator, τ and t denote the time, $\mathbf{r}(\tau)$ is the position of the oscillator at time τ, and $\langle\cdot\rangle_\tau$ stands for averaging over all values of τ. From the properties of the Fourier transform, it is known that forming a derivative in the time domain is equivalent to a multiplication with the independent variable in the frequency domain, so that

$$P(\omega) = m \int \Big\langle \dot{\mathbf{r}}(\tau) \cdot \dot{\mathbf{r}}(t+\tau) \Big\rangle_\tau \, \exp\big(-\mathrm{i}\omega t\big) \, \mathrm{d}t \tag{2.3.2}$$

is formally equivalent to the first equation, using $\dot{\mathbf{r}}(\tau)$ as the velocity of the oscillator at time τ. Despite being equivalent, the latter equation offers certain numerical advantages, so that almost only this form is used in practice. This also coincides with the verbal definition, stating that the power spectrum is "the Fourier transform of the velocity autocorrelation function".

One important aspect in computing power spectra is the mass weighting, as performed in the two equations from above by introducing the reduced mass m of the oscillator. The mass weighting ensures that the power spectrum is obtained in the correct physical units so that the integral over the spectrum possesses the unit of temperature. This means that the temperature of a certain degree of freedom in the system can be directly read by considering the integral over the corresponding band in the power spectrum. According to the equipartition theorem,[277–280] all degrees of freedom in a well-equilibrated system carry the same amount of energy on average, so that the power spectrum can even be used to assess whether a simulation is well equilibrated by comparing the integrals of bands in the spectrum corresponding to different degrees of freedom in the system.

Going beyond one-dimensional oscillators, the power spectrum of a more complex system *(such as an atomistic simulation)* can be simply expressed as sum over the individual atoms' power spectra

$$P(\omega) = \sum_{i=1}^{N} \left[m_i \int \left\langle \dot{\mathbf{r}}_i(\tau) \cdot \dot{\mathbf{r}}_i(t+\tau) \right\rangle_\tau \exp\left(-\mathrm{i}\omega t\right) \, \mathrm{d}t \right], \qquad (2.3.3)$$

where N is the total number of atoms in the system, m_i is the mass of atom i, and $\dot{\mathbf{r}}_i(\tau)$ is the velocity of atom i at time τ.[MB14,MB19,MB40]

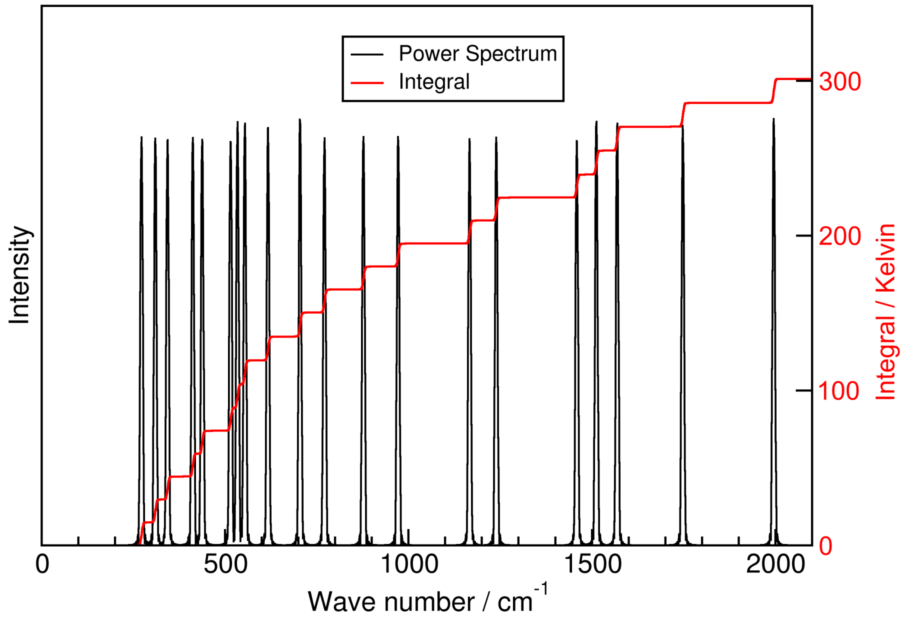


Figure 2.3.1.: Power spectrum of a system containing 20 different harmonic diatomic molecules with different atom masses and force constants *(black curve)* together with integral depicting the temperature of each normal mode *(red curve)*.[MB40]

An illustrative example is presented in Figure 2.3.1, where the power spectrum of an atomistic simulation containing 20 different harmonic diatomic molecules with different atom masses and force constants is shown as black curve. The red curve depicts the integral *(see temperature units on the right vertical axis)*. It can be seen that the integral of all spectral bands is approximately equal, and therefore the equipartition theorem is well fulfilled in this simulation.

Please note that the autocorrelation in Equations 2.3.1 to 2.3.3 can be efficiently computed by utilizing the Wiener–Khinchin theorem[281,282] that is described in Section 2.12.1 as well as in Section A.2 in the Appendix. This approach is also used in the TRAVIS program package[MB2,MB40] to save computational time.

The power spectrum is the foundation of the normal mode decomposition approach[MB19] that will be discussed in Section 2.11.

2.4. Infrared Spectroscopy

Infrared spectra based on the static–harmonic approach can be computed since many decades[43–52,283–286] using popular quantum chemistry program packages such as Gaussian,[58] ADF,[287], Turbomole,[288,289] and ORCA.[290–292] When predicting infrared spectra in the static–harmonic approach, the integral infrared absorption coefficient $\mathcal{A}_{\tilde{\nu}}$ of normal mode p is given by[43]

$$\mathcal{A}_{\tilde{\nu}} = \frac{1}{4\pi\epsilon_0} \cdot \frac{N_\text{A}\pi}{3c^2} \cdot \left(\frac{\partial \boldsymbol{\mu}}{\partial Q_p}\right)^2, \tag{2.4.1}$$

where Q_p denotes the mass-weighted normal coordinate corresponding to normal mode p, N_A is Avogadro's constant, c is the speed of light in vacuum, and ϵ_0 represents the vacuum permittivity. The derivative of the molecular dipole vector $\boldsymbol{\mu}$ with respect to the normal mode is evaluated at the equilibrium geometry.

Apart from employing the static–harmonic approximation, there exists the possibility to compute vibrational spectra directly from molecular dynamics (MD) simulations. In this approach, the spectra are obtained as the Fourier transform of some time correlation functions along the simulation trajectory—an idea which is at least 55 years old now,[106–112] and is called the time–correlation function (TCF) formalism. As discussed in Section 1.2.2, this approach comes with several advantages over the static–harmonic concept. However, practical applications to condensed phase systems were not possible before *ab initio* molecular dynamics (AIMD) simulations[114,115] could be routinely performed for such systems. The pioneers of that field were Roberto Car and Michele Parrinello, which published their approach of Car–Parrinello molecular dynamics (CPMD) in 1985.[116] With this class of methods, periodic bulk phase systems can be treated natively, and even anharmonic effects are covered to a certain extent, as line shapes, overtones, and combination bands are reproduced in qualitatively correct manner.[113,117]

The first predictions of condensed phase infrared spectra based on AIMD were published in 1997 on the examples of liquid water[118] and amorphous silica.[119] Since then, many more infrared spectra of condensed phase systems based on this methodology have been presented in the literature.[113,121–135,293–298,MB14,MB19,MB23]

For the prediction of vibrational spectra based on AIMD simulations, the derivatives with respect to the normal modes need to be replaced by suitable time-correlation functions. In the case of infrared spectroscopy, the derivative of the molecular dipole moment $\boldsymbol{\mu}$ in Equation 2.4.1 is replaced with an autocorrelation function of the molecular dipole moment along the trajectory and a subsequent

Fourier transform, so that the infrared absorption spectrum $A(\tilde{\nu})$ is obtained as

$$A(\tilde{\nu}) = \frac{N_\mathrm{A}}{12\epsilon_0 c k_\mathrm{B} T} \int \langle \dot{\mu}(\tau) \cdot \dot{\mu}(t+\tau) \rangle_\tau \exp\Big(-2\pi \mathrm{i} c \tilde{\nu} t\Big)\, \mathrm{d}t, \tag{2.4.2}$$

where $\tilde{\nu}$ is the wavenumber, T denotes the temperature at which the observation is performed, N_A is Avogadro's number, k_B is Boltzmann's constant, ϵ_0 stands for the vacuum permittivity, c denotes the speed of light in vacuum, and $\langle \cdot \rangle_\tau$ stands for the temporal average over the simulation time τ. Note that instead of the molecular dipole moment μ the time derivative $\dot{\mu}$ enters the expression, which is equivalent to multiplying the final spectrum with ω^2, but offers some numerical advantages as discussed in Section 2.3.

There arises the interesting question whether the dipole moments μ in Equation 2.4.2 should be molecular quantities, or rather those of the whole system. In the latter case, a treatment in the framework of the modern theory of polarization *(based on a Berry phase)*[193,195,196] is required, because the traditional dipole moment operator is ill-defined for periodic systems as discussed in Section 1.4.

When assuming that the sum of the molecular dipole moments equals the total dipole moment *(which is an approximation in the periodic case, but the deviation is small*[206]*)*, it can be shown that using the total dipole moment is mathematically equivalent to using all autocorrelations of individual molecules' dipole moments together with all cross-correlations between different molecules. For the case of two molecules A and B with dipole moments μ_A and μ_B with the assumption $\mu_\mathrm{A}+\mu_\mathrm{B} = \mu_\mathrm{Total}$, we find that

$$\begin{gathered}
\Big\langle (\mu_\mathrm{A}+\mu_\mathrm{B})(\tau) \cdot (\mu_\mathrm{A}+\mu_\mathrm{B})(\tau+t) \Big\rangle_\tau = \\
\Big\langle [\mu_\mathrm{A}(\tau) \cdot \mu_\mathrm{A}(\tau+t)] + [\mu_\mathrm{A}(\tau) \cdot \mu_\mathrm{B}(\tau+t)] + [\mu_\mathrm{B}(\tau) \cdot \mu_\mathrm{A}(\tau+t)] + [\mu_\mathrm{B}(\tau) \cdot \mu_\mathrm{B}(\tau+t)] \Big\rangle_\tau = \\
\Big\langle \mu_\mathrm{A}(\tau) \cdot \mu_\mathrm{A}(\tau+t) \Big\rangle_\tau + \Big\langle \mu_\mathrm{A}(\tau) \cdot \mu_\mathrm{B}(\tau+t) \Big\rangle_\tau + \Big\langle \mu_\mathrm{B}(\tau) \cdot \mu_\mathrm{A}(\tau+t) \Big\rangle_\tau + \Big\langle \mu_\mathrm{B}(\tau) \cdot \mu_\mathrm{B}(\tau+t) \Big\rangle_\tau ,
\end{gathered} \tag{2.4.3}$$

which is equivalent to the above statement for the special case of two molecules. It is often argued that cross-correlations between molecules which are not neighbors in space only introduce noise to the spectrum, as the motions of such molecules are not at all correlated. Therefore, we decided to consider only the molecular autocorrelation functions and to neglect all cross-correlation terms between different molecules. We use our Voronoi integration approach[MB23] *(see Section 2.1)* to compute the molecular dipole moments from the total electron density.

Some examples for infrared spectra predicted via this formalism are shown in Section 3.1.

2.5. Raman Spectroscopy

Static–harmonic Raman spectra based on normal mode analysis can be computed since several decades now.[43,45,47–50,53,55–57,299–303] For computing Raman intensities, the scattering geometry needs to be taken into account. Here, we consider only the typical setup of an incident beam propagating in the y direction with polarization along the x axis, and a scattering angle of 90°. As described in the literature,[304,MB14] the Raman scattering intensity $I_{\|}$ of molecules fixed in space for detection of x polarized light in this case reads

$$I_{\|} = \frac{\pi^2 h N \mathcal{I}}{8\pi^2 c \epsilon_0^2} \cdot \frac{\left(\tilde{\nu}_{\text{in}} - \tilde{\nu}_p\right)^4}{\tilde{\nu}_p} \cdot \frac{1}{1 - \exp\left(-\frac{hc\tilde{\nu}_p}{k_{\text{B}}T}\right)} \cdot \left(\frac{\partial \alpha_{xx}}{\partial Q_p}\right)^2, \tag{2.5.1}$$

where Q_p denotes the mass-weighted normal coordinate corresponding to normal mode p, $\tilde{\nu}_{\text{in}}$ is the wavenumber of the incident light, $\tilde{\nu}_p$ is the wavenumber of normal mode p, N is the number of molecules, $\mathcal{I}$ is the irradiance of the incident light, h is Planck's constant, k_{B} is Boltzmann's constant, and T is the temperature. As in the case of infrared intensities, the derivative of the molecular polarizability tensor $\boldsymbol{\alpha}$ is evaluated at the equilibrium geometry.

In practical applications, molecules are often oriented arbitrarily in space, so that it is advantageous to drop the requirement of a fixed molecular orientation. Averaging over random molecular orientation leads to the expressions[304,MB14]

$$I_{\|} = \frac{\pi^2 h N \mathcal{I}}{8\pi^2 c \epsilon_0^2} \cdot \frac{\left(\tilde{\nu}_{\text{in}} - \tilde{\nu}_p\right)^4}{\tilde{\nu}_p} \cdot \frac{1}{1 - \exp\left(-\frac{hc\tilde{\nu}_p}{k_{\text{B}}T}\right)} \cdot \frac{45a_p^2 + 4\gamma_p^2}{45} \tag{2.5.2}$$

for detection of x polarized light (*parallel* to the incident light polarization plane), and

$$I_{\perp} = \frac{\pi^2 h N \mathcal{I}}{8\pi^2 c \epsilon_0^2} \cdot \frac{\left(\tilde{\nu}_{\text{in}} - \tilde{\nu}_p\right)^4}{\tilde{\nu}_p} \cdot \frac{1}{1 - \exp\left(-\frac{hc\tilde{\nu}_p}{k_{\text{B}}T}\right)} \cdot \frac{3\gamma_p^2}{45} \tag{2.5.3}$$

for detection of y polarized light (*perpendicular* to the incident light polarization plane). These expressions make use of the two rotationally averaged Raman invariants a_p and γ_p for normal mode p, where the isotropic polarizability a_p is defined as

$$a_p := \frac{1}{3}\left(\frac{\partial \alpha_{xx}}{\partial Q_p} + \frac{\partial \alpha_{yy}}{\partial Q_p} + \frac{\partial \alpha_{zz}}{\partial Q_p}\right), \tag{2.5.4}$$

and the square of the anisotropy γ_p is given by

$$\begin{aligned}\gamma_p^2 :=& \frac{1}{2}\left(\frac{\partial \alpha_{xx}}{\partial Q_p} - \frac{\partial \alpha_{yy}}{\partial Q_p}\right)^2 + \frac{1}{2}\left(\frac{\partial \alpha_{yy}}{\partial Q_p} - \frac{\partial \alpha_{zz}}{\partial Q_p}\right)^2 + \frac{1}{2}\left(\frac{\partial \alpha_{zz}}{\partial Q_p} - \frac{\partial \alpha_{xx}}{\partial Q_p}\right)^2 \\ &+ 3\left(\frac{\partial \alpha_{xy}}{\partial Q_p}\right)^2 + 3\left(\frac{\partial \alpha_{yz}}{\partial Q_p}\right)^2 + 3\left(\frac{\partial \alpha_{zx}}{\partial Q_p}\right)^2 .\end{aligned} \tag{2.5.5}$$

The first prediction of a Raman spectrum from an AIMD simulation—based on the time-correlation function (TCF) formalism—was published in 2002 by Putrino *et al.*[120] Since then, many more examples of computed Raman spectra based on this approach appeared in the literature.[113,130,136–140,218,219,293,298,305,MB14,MB19,MB23,MB26,MB36,MB49] As shown before for infrared spectroscopy, the derivatives with respect to the normal modes need to be replaced by corresponding time-correlation functions. For predicting Raman spectra, this needs to be done for the two Raman invariants in Equations 2.5.4 and 2.5.5.

From the time series of the polarizability tensor $\boldsymbol{\alpha}(t)$ along the trajectory, these two invariants can be computed via cross-correlation and subsequent Fourier transformations,[MB14] as given by

$$a(\tilde{\nu}) := \int\limits_{-\infty}^{\infty} \left\langle \frac{\dot{\alpha}_{xx}(\tau) + \dot{\alpha}_{yy}(\tau) + \dot{\alpha}_{zz}(\tau)}{3} \cdot \frac{\dot{\alpha}_{xx}(\tau{+}t) + \dot{\alpha}_{yy}(\tau{+}t) + \dot{\alpha}_{zz}(\tau{+}t)}{3} \right\rangle_{\tau} \exp(-2\pi \mathrm{i} c \tilde{\nu} t)\, \mathrm{d}t \tag{2.5.6}$$

for the isotropic contribution, and

$$\begin{aligned}\gamma(\tilde{\nu}) := \int\limits_{-\infty}^{\infty} \Bigg[& \frac{1}{2}\Big\langle \big(\dot{\alpha}_{xx}(\tau) - \dot{\alpha}_{yy}(\tau)\big) \cdot \big(\dot{\alpha}_{xx}(\tau{+}t) - \dot{\alpha}_{yy}(\tau{+}t)\big) \Big\rangle_{\tau} \\ & + \frac{1}{2}\Big\langle \big(\dot{\alpha}_{yy}(\tau) - \dot{\alpha}_{zz}(\tau)\big) \cdot \big(\dot{\alpha}_{yy}(\tau{+}t) - \dot{\alpha}_{zz}(\tau{+}t)\big) \Big\rangle_{\tau} \\ & + \frac{1}{2}\Big\langle \big(\dot{\alpha}_{zz}(\tau) - \dot{\alpha}_{xx}(\tau)\big) \cdot \big(\dot{\alpha}_{zz}(\tau{+}t) - \dot{\alpha}_{xx}(\tau{+}t)\big) \Big\rangle_{\tau} \\ & + \frac{3}{2}\Big\langle \dot{\alpha}_{xy}(\tau) \cdot \dot{\alpha}_{xy}(\tau{+}t) \Big\rangle_{\tau} + \frac{3}{2}\Big\langle \dot{\alpha}_{yz}(\tau) \cdot \dot{\alpha}_{yz}(\tau{+}t) \Big\rangle_{\tau} \\ & + \frac{3}{2}\Big\langle \dot{\alpha}_{zx}(\tau) \cdot \dot{\alpha}_{zx}(\tau{+}t) \Big\rangle_{\tau} \Bigg] \cdot \exp(-2\pi \mathrm{i} c \tilde{\nu} t)\, \mathrm{d}t, \end{aligned} \tag{2.5.7}$$

for the anisotropic contribution, where $\dot{\alpha}_{xx}$ denotes the time derivative of the xx component of the polarizability tensor $\boldsymbol{\alpha}$, and $\langle \cdot \rangle_{\tau}$ stands for the temporal average over the simulation time τ. Note that instead of the molecular polarizabilities $\boldsymbol{\alpha}$ the time derivatives $\dot{\boldsymbol{\alpha}}$ enter these expressions, which is equivalent to multiplying the final spectrum with ω^2, but offers some numerical advantages as discussed in

Section 2.3.

Based on these invariants, the Raman intensities for a scattering angle of 90° can be computed for different observed polarizations.[MB14,MB26] If the polarization planes of incident light and observed signal are parallel *("parallel Raman spectrum")*, the intensity is given by

$$I_{\parallel}(\tilde{\nu}) = \frac{h}{8\epsilon_0^2 c k_\mathrm{B} T} \cdot \frac{(\tilde{\nu}_\mathrm{in} - \tilde{\nu})^4}{\tilde{\nu}} \cdot \frac{1}{1 - \exp\left(-\frac{hc\tilde{\nu}}{k_\mathrm{B}T}\right)} \cdot \frac{45a(\tilde{\nu}) + 4\gamma(\tilde{\nu})}{45}. \tag{2.5.8}$$

If the polarization planes of incident light and observed signal are perpendicular to each other *("orthogonal Raman spectrum")*, the intensity is obtained as

$$I_{\perp}(\tilde{\nu}) = \frac{h}{8\epsilon_0^2 c k_\mathrm{B} T} \cdot \frac{(\tilde{\nu}_\mathrm{in} - \tilde{\nu})^4}{\tilde{\nu}} \cdot \frac{1}{1 - \exp\left(-\frac{hc\tilde{\nu}}{k_\mathrm{B}T}\right)} \cdot \frac{3\gamma(\tilde{\nu})}{45}. \tag{2.5.9}$$

If the polarization plane of the scattered light is ignored and all scattered light is considered *("unpolarized/depolarized Raman spectrum")*, the intensity is simply expressed as the sum of $I_{\parallel}(\tilde{\nu})$ and $I_{\perp}(\tilde{\nu})$:

$$I(\tilde{\nu}) = \frac{h}{8\epsilon_0^2 c k_\mathrm{B} T} \cdot \frac{(\tilde{\nu}_\mathrm{in} - \tilde{\nu})^4}{\tilde{\nu}} \cdot \frac{1}{1 - \exp\left(-\frac{hc\tilde{\nu}}{k_\mathrm{B}T}\right)} \cdot \frac{45a(\tilde{\nu}) + 7\gamma(\tilde{\nu})}{45}. \tag{2.5.10}$$

Due to the lengthy pre-factor, the intensity I is obtained in the correct physical units of length^2. In all three cases, $\tilde{\nu}$ is the wavenumber of the signal, $\tilde{\nu}_\mathrm{in}$ is the wavenumber of the incident laser light, T denotes the temperature at which the observation is performed, k_B is Boltzmann's constant, ϵ_0 stands for the vacuum permittivity, and c denotes the speed of light in vacuum.

Finally, the depolarization ratio[306,307] ρ of the Raman spectrum can be computed as the quotient of the orthogonal and parallel Raman spectrum via

$$\rho(\tilde{\nu}) = \frac{I_{\perp}(\tilde{\nu})}{I_{\parallel}(\tilde{\nu})}. \tag{2.5.11}$$

It can have values < 0.75 for totally symmetric normal modes, but is 0.75 in all other cases.[MB14]

A few examples for Raman spectra predicted by this method can be found in Section 3.2.

2.6. Classical Approach to Electric Currents

The prediction of VCD and ROA spectra from AIMD simulation requires magnetic moments, which in turn require electric currents to be computed along the trajectory. However, most AIMD simulations are performed in the Born–Oppenheimer approximation *(BOMD simulations)*, which includes that no electric currents can be described. There are several ways how to overcome this difficulty, such as using nuclear velocity perturbation theory (NVPT)[62,308,309] to compute the electric currents. However, this requires a significant amount of additional computer time, and is hard to implement in a generic way.

In 2016, Thomas and Kirchner developed a purely classical approach to deduce electric currents from BOMD simulations.[150] In simple words, this approach considers the total electron density in subsequent steps of the MD trajectory, and then asks how the current could have been flown to cause the observed change in total electron density. This idea is coupled with the assumption that current should never flow through the vacuum, but only along paths where some electron density is located.

The derivation[150] starts with the continuity equation

$$\frac{\partial\rho(\mathbf{r})}{\partial t} + \nabla\cdot\mathbf{j}(\mathbf{r}) = 0, \tag{2.6.1}$$

where $\rho(\mathbf{r})$ denotes the total electron density at position $\mathbf{r}$ in space, and $\mathbf{j}(\mathbf{r})$ is the electric current vector at $\mathbf{r}$. The electric current is then approximated as the product of the total electron density and a conservative vector field $\mathbf{v}(\mathbf{r})$, which can be written as the gradient of some auxiliary scalar field $\alpha(\mathbf{r})$:

$$\mathbf{j}(r) = \rho(\mathbf{r})\,\mathbf{v}(\mathbf{r}) = -\rho(\mathbf{r})\,\nabla\alpha(\mathbf{r}) \tag{2.6.2}$$

By combining Equations 2.6.1 and 2.6.2, one obtains

$$\begin{align}\frac{\partial\rho(\mathbf{r})}{\partial t} &= -\nabla\cdot\mathbf{j}(\mathbf{r}) \tag{2.6.3}\\ &= -\nabla\cdot\Big(-\rho(\mathbf{r})\,\nabla\alpha(\mathbf{r})\Big) \tag{2.6.4}\\ &= \nabla\rho(\mathbf{r})\,\nabla\alpha(\mathbf{r}) + \rho(\mathbf{r})\,\Delta\alpha(\mathbf{r})\,, \tag{2.6.5}\end{align}$$

where $\Delta = \nabla^2$ represents the Laplace operator. After solving Equation 2.6.5, the electric current can be obtained via Equation 2.6.2.

This partial differential equation is now discretized on a Cartesian grid with $n_x \times n_y \times n_z$ grid points. The derivatives are replaced by second-order central finite differences, *i. e.*

$$(\mathrm{D}_x\alpha)_{i,j,k} = \frac{\alpha_{i+1,j,k} - \alpha_{i-1,j,k}}{2h_x}, \tag{2.6.6}$$

$$(\mathrm{D}_y\alpha)_{i,j,k} = \frac{\alpha_{i,j+1,k} - \alpha_{i,j-1,k}}{2h_y}, \tag{2.6.7}$$

$$(\mathrm{D}_z\alpha)_{i,j,k} = \frac{\alpha_{i,j,k+1} - \alpha_{i,j,k-1}}{2h_z} \tag{2.6.8}$$

and

$$\begin{aligned}(\Delta\alpha)_{i,j,k} =& \frac{\alpha_{i+1,j,k} - 2\alpha_{i,j,k} + \alpha_{i-1,j,k}}{h_x^2} + \frac{\alpha_{i,j+1,k} - 2\alpha_{i,j,k} + \alpha_{i,j-1,k}}{h_y^2} + \\ & \frac{\alpha_{i,j,k+1} - 2\alpha_{i,j,k} + \alpha_{i,j,k-1}}{h_z^2},\end{aligned} \tag{2.6.9}$$

where i, j, k depict the grid index in x, y, z direction, and h_x, h_y, h_z indicate the grid spacing in these directions, respectively. The periodic boundary conditions are taken into account by letting $i \in \mathbb{Z}/n_x\mathbb{Z}$, $j \in \mathbb{Z}/n_y\mathbb{Z}$, $k \in \mathbb{Z}/n_z\mathbb{Z}$, so that *e. g.* the successor of $i = (n_x - 1)$ is $i = 0$.

This leads to a system of linear equations

$$\mathbf{A} \cdot \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_{n_x n_y n_z - 1} \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial\rho}{\partial t}\right)_0 \\ \vdots \\ \left(\frac{\partial\rho}{\partial t}\right)_{n_x n_y n_z - 1} \end{pmatrix} \tag{2.6.10}$$

with the coefficient matrix $\mathbf{A} \in \mathbb{R}^{n_x n_y n_z \times n_x n_y n_z}$. In typical practical applications we find that $n_x, n_y, n_z \approx 200$, so that this system contains $\approx 10^7$ equations and unknowns. Such a huge system of linear equations cannot be solved by classical approaches such as LU decomposition, and not even the coefficient matrix $\mathbf{A}$ can be represented in memory *(5 800 Terabyte in double precision)*. Luckily, $\mathbf{A}$ is very sparse, so that space-optimized representations for sparse matrices can be utilized. One way to solve such a large system of linear equations is the iterative BiCGstab(l) *("biconjugate gradient stabilized method")* approach developed by van der Vorst.[310–313] The convergence of the algorithm can be improved by using an incomplete LU factorization[314] of $\mathbf{A}$ for preconditioning.

A simple two-dimensional illustration of the approach is presented in Figure 2.6.1. In the left-hand panel, the total "electron density" is shown. The middle panel represents the change of electron density between two subsequent steps. To cause this change, electron density needs to be transported from the black circle to the white circle. Based on these inputs, the partial differential equation from Equation 2.6.5 is solved. As described above, the equation contains the assumption that current can only flow along paths where electron density is located *(not through vacuum)*. The resulting vector field of the current is shown in the right-hand panel. It can be seen that indeed the current follows the path of the total electron density, as one could expect.

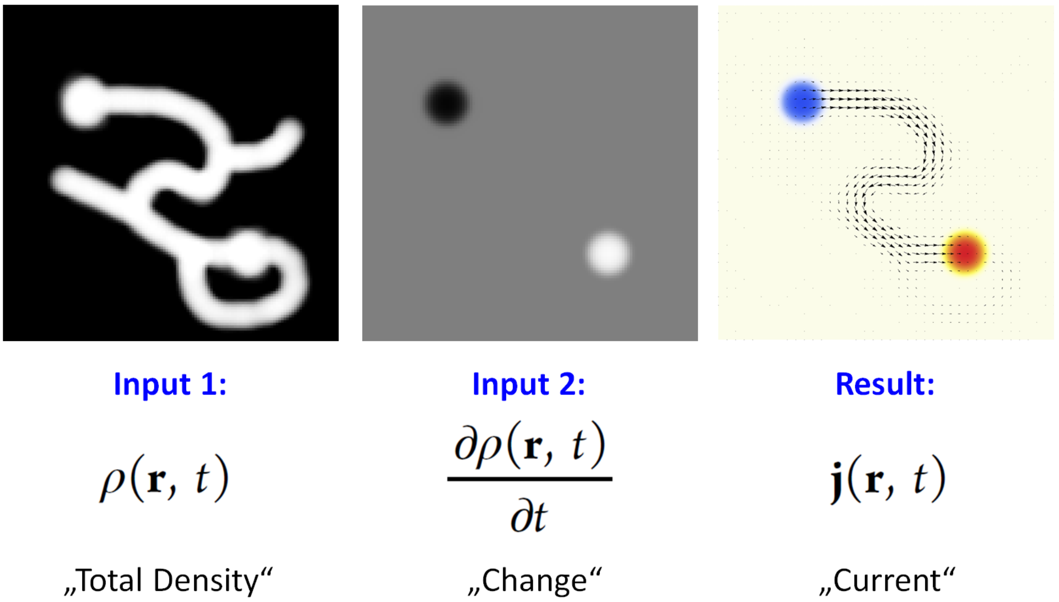


Figure 2.6.1.: Schematic example of the classical approach to obtain electric currents. From the inputs of total electron density *(left panel)* and change in electron density *(middle panel)*, the vector field of the current is obtained from solving Equation 2.6.5 *(right panel)*.

One advantage of the approach presented above is that it only requires the total electron density of subsequent MD steps on a grid, and can therefore easily be applied to a wide variety of electron structure methods without any modifications to the quantum chemistry program package. Promising candidates are hybrid DFT functionals such as PBE0,[315] or even post-Hartree–Fock methods such as MP2, for which an implementation of NVPT would pose a significant challenge.

2.7. VCD Spectroscopy

Static–harmonic VCD spectra based on normal mode analysis can be computed routinely since several decades,[29,48,50,59–66,316,317] and the predicted spectra often agree reasonably well with the experiments.[318–325] The most commonly used approach to predict such spectra is the magnetic field perturbation theory (MFPT)[29,61,63,64,326] as implemented in Gaussian,[58] ADF,[287] and Turbomole.[288,289] But there exist also alternative theories. For example, it was shown that the calculation of the required magnetic transition dipole moment needs non-Born–Oppenheimer contributions in the wave function.[59,60] In the nuclear velocity perturbation theory (NVPT),[62,308,309] this is achieved by the perturbative calculation of the complete adiabatic (CA) correction to the Born–Oppenheimer ground-state orbitals.[327–330] This provides a closely related yet conceptually different route for the calculation of VCD spectra, and has been first implemented by Scherrer *et al.* into the CPMD code[331] in 2013.[65] A different derivation of NVPT based on the exact factorization of the electron–nuclear wave function instead of CA was published later.[66]

When computing static–harmonic VCD spectra, the VCD intensity of a normal mode k is proportional to the rotational strength R_k of that mode, and

$$R_k \propto \frac{\partial \langle \hat{\mathbf{m}} \rangle}{\partial \dot{q}_k} \cdot \frac{\partial \langle \hat{\boldsymbol{\mu}} \rangle}{\partial \dot{q}_k} \tag{2.7.1}$$

with the time derivative of the electric dipole moment $\langle \hat{\boldsymbol{\mu}} \rangle$ and the magnetic dipole moment $\langle \hat{\mathbf{m}} \rangle$, while $\dot{q}_k$ is the velocity along normal mode k.[66] The two derivatives are known as the atomic polar tensor (APT) $\mathcal{P}$ and the atomic axial tensor (AAT) $\mathcal{M}$:

$$\mathcal{P}_{\alpha\beta} \equiv \frac{\partial \left\langle \hat{\dot{\boldsymbol{\mu}}}_\beta \right\rangle}{\partial \dot{R}_\alpha}, \qquad \mathcal{M}_{\alpha\beta} \equiv \frac{\partial \langle \hat{\mathbf{m}}_\beta \rangle}{\partial \dot{R}_\alpha} \tag{2.7.2}$$

It has been shown in the literature[332–336] that VCD spectra can be computed via the time-correlation function (TCF) formalism[106–112] by Fourier transforming the cross-correlation of electric dipole moment $\boldsymbol{\mu}$ and magnetic dipole moment $\mathbf{m}$ from a MD simulation. The VCD spectrum is obtained as[150]

$$\Delta A(\tilde{\nu}) = \frac{N_\mathrm{A}}{12\epsilon_0 c^2 k_\mathrm{B} T} \int\limits_{-\infty}^{\infty} \Big(\langle \dot{\boldsymbol{\mu}}(\tau) \cdot \dot{\mathbf{m}}(\tau{+}t) \rangle_\tau - \langle \dot{\mathbf{m}}(\tau) \cdot \dot{\boldsymbol{\mu}}(\tau{+}t) \rangle_\tau \Big) \exp(-2\pi \mathrm{i} c \tilde{\nu} t)\, \mathrm{d}t, \tag{2.7.3}$$

where $\tilde{\nu}$ denotes the wavenumber, t and τ represent the time, N_A is the Avogadro constant, ϵ_0 the vacuum permittivity, c the speed of light in vacuum, k_B the Boltzmann constant, T the average simulation temperature, and $\langle \dot{\boldsymbol{\mu}}(\tau) \cdot \dot{\mathbf{m}}(\tau{+}t) \rangle_\tau$ stands for the cross-correlation function of the time derivatives of the electric and magnetic dipole moment. Note that the difference between the two correlation functions inside of the integral is a special case of the more general "commutator trick" that has been derived in the scope of this thesis—see Equation 2.12.6 in Section 2.12.4.

The first *ab initio* prediction of a liquid phase VCD spectrum based on the TCF approach via molecular dynamics was published by Thomas and Kirchner in 2016,[150] using a purely classical approach to compute the electric current and the magnetic dipole moment *(see Section 2.6)*. Only a few months later, Scherrer and Sebastiani presented another *ab initio* prediction of a VCD spectrum via MD,[151] based on nuclear velocity perturbation theory (NVPT) to obtain the magnetic dipole moments. Since then, only a few more VCD spectra based on AIMD simulations have appeared in the literature.[296,297,337,338] The former approach is implemented in the TRAVIS program package and will be shortly described here.

One of the strengths of this approach is that it only requires a standard Born–Oppenheimer molecular dynamics (BOMD) simulation as an input. The molecular dipole moments $\boldsymbol{\mu}^{\mathrm{Mol}}$ are obtained from the total electron density in each frame by using our Voronoi integration approach[MB23,MB48] *(see Section 2.1)* as

$$\boldsymbol{\mu}^{\mathrm{Mol}} = \sum_{i=1}^{N_{\mathrm{Mol}}} q_i \mathbf{r}_i - \int_{\mathrm{Mol}} \rho(\mathbf{s})\, \mathbf{s}\, \mathrm{d}^3\mathbf{s}, \tag{2.7.4}$$

where q_i and $\mathbf{r}_i$ depict the nuclear charge and position of atom i, respectively, $\rho(\mathbf{s})$ is the total electron density in space, and the integration limits "Mol" indicate integration over the molecular Voronoi cell. The electric current $\mathbf{j}(s)$ is obtained on a grid by solving the classical partial differential equation discussed in Section 2.6 with the total electron densities of three subsequent steps as the input. Once $\mathbf{j}(s)$ is known, the molecular magnetic dipole moment $\mathbf{m}^{\mathrm{Mol}}$ is obtained—again—via Voronoi integration as

$$\mathbf{m}^{\mathrm{Mol}} = \frac{1}{2} \sum_{i=1}^{N_{\mathrm{Mol}}} q_i \left(\mathbf{r}_i \times \mathbf{v}_i \right) - \frac{1}{2} \int_{\mathrm{Mol}} \mathbf{s} \times \mathbf{j}(\mathbf{s})\ \mathrm{d}^3\mathbf{s}, \tag{2.7.5}$$

with $\mathbf{v}_i$ representing the velocity vector of atom i. In both integrals, the molecular center of mass is used as the coordinate origin. Based on these two quantities, the VCD spectrum of the system can be readily computed via Equation 2.7.3 by averaging over all molecular sub-spectra. Some technical details which are important to obtain high-quality spectra are discussed in Section 2.12.

An example for a VCD spectrum computed following this approach can be found in Section 3.3.

2.8. ROA Spectroscopy

Static–harmonic ROA spectra based on the normal mode analysis of a molecule can be computed routinely since more than two decades,[45,49,50,53,57,63,69,70,72,339–345] for example by using the Gaussian program package.[58] However—similar to VCD spectroscopy—a predicted ROA spectrum based on molecular dynamics simulations via the time–correlation function (TCF) approach was not seen for a long time. The first such spectrum was published in 2017 by Luber,[152] but only one single molecule in vacuum was investigated, and as some derivations have explicitly been conducted under the assumption of a non-periodic system, it is concluded there that the application to liquid systems remains an open project for the future. A few months later, we published the first predicted ROA spectrum of a periodic liquid system,[MB26] based on a novel approach through AIMD simulations.

When computing spectra in the TCF formalism, time-correlation functions of certain quantities have to be formed. In case of ROA, this is not straight-forward, as the relevant properties are three different polarizability tensors of the system *(also known as optical activity tensors)*.[304] These are the electric dipole–electric dipole polarizability tensor α, the electric quadrupole–electric dipole polarizability tensor A, and the magnetic dipole–electric dipole polarizability tensor G'. The main challenge in computing ROA spectra from MD simulations is to obtain these tensors in each simulation time step in order to compute the required time correlation functions. This can be achieved, *e. g.*, by applying perturbation theory[152], which requires, however, significant amounts of computer resources to solve the resulting perturbation equations.

Our approach is based on the purely classical approximation to the electric current as presented in Section 2.6. Only the coordinates $\mathbf{r}_i$ of the nuclei and the total electron density $\rho(\mathbf{r})$ on a three-dimensional grid are required as input in each time step. Based on this information, we compute the molecular dipole moment vector $\boldsymbol{\mu}^{\mathrm{Mol}}$ and the trace-free molecular quadrupole moment[68] tensor $\mathbf{Q}^{\mathrm{Mol}}$ for each molecule by the classical expressions

$$\boldsymbol{\mu}^{\mathrm{Mol}} = \sum_{i=1}^{N_{\mathrm{Mol}}} q_i \mathbf{r}_i - \int_{\mathrm{Mol}} \rho(\mathbf{s})\, \mathbf{s}\ \mathrm{d}^3\mathbf{s}, \tag{2.8.1}$$

$$\mathbf{Q}^{\mathrm{Mol}}_{jk} = \sum_{i=1}^{N_{\mathrm{Mol}}} q_i \left(3\mathbf{r}_{i,j}\mathbf{r}_{i,k} - \left\|\mathbf{r}_i\right\|^2 \delta_{jk}\right) - \int_{\mathrm{Mol}} \rho(\mathbf{s}) \left(3\mathbf{s}_j\mathbf{s}_k - \left\|\mathbf{s}\right\|^2 \delta_{jk}\right) \mathrm{d}^3\mathbf{s}, \tag{2.8.2}$$

where $\mathbf{r}_i$ and q_i denote the position and core charge of the i-th nucleus, respectively, $\mathbf{s}$ denotes points in space at which the electron density is evaluated, $\mathbf{s}_j$ is the j-th component of vector $\mathbf{s}$, and δ is the Kronecker delta. The volume in space over which ρ is integrated for each molecule is determined by a radical Voronoi tessellation, as we have proposed before *(see Section 2.1)*.[MB23,MB48] In the general case (*i. e.*, charged molecules), $\boldsymbol{\mu}^{\mathrm{Mol}}$ is dependent on the choice of the coordinate origin; similarly, $\mathbf{Q}^{\mathrm{Mol}}$ depends on the origin if the dipole moment is non-zero. For each molecule, we use the center of mass as the coordinate origin when performing these computations.

After the electric current $\mathbf{j}$ has been calculated by the classical differential equation, the molecular magnetic dipole moment $\mathbf{m}^{\mathrm{Mol}}$ is computed via the expression

$$\mathbf{m}^{\mathrm{Mol}} = \frac{1}{2}\sum_{i=1}^{N_{\mathrm{Mol}}} q_i \left(\mathbf{r}_i \times \mathbf{v}_i\right) - \frac{1}{2}\int\limits_{\mathrm{Mol}} \mathbf{s} \times \mathbf{j}(\mathbf{s})\ \mathrm{d}^3\mathbf{s}, \tag{2.8.3}$$

where $\mathbf{v}_i$ denotes the velocity of the i-th nucleus, $\mathbf{j}(\mathbf{s})$ is the volumetric electric current in space, and the molecular center of mass is used as coordinate origin again.

To obtain the corresponding polarizabilities from $\boldsymbol{\mu}^{\mathrm{Mol}}$, $\mathbf{Q}^{\mathrm{Mol}}$, and $\mathbf{m}^{\mathrm{Mol}}$, we apply an external homogeneous electric field $\mathbf{E}$ in each spatial direction to the simulation frame under consideration. From a central finite differences scheme, we obtain the molecular electric dipole–electric dipole polarizability tensor α^{Mol}, the molecular electric dipole–electric quadrupole polarizability tensor $\mathcal{A}^{\mathrm{Mol}}$, and the molecular electric dipole–magnetic dipole polarizability tensor $\mathcal{G}'^{\mathrm{Mol}}$, and by using the following identities from the literature,[304] finally the required polarizabilities A^{Mol} and G'^{Mol}:

$$\alpha^{\mathrm{Mol}} = \frac{\mathrm{d}}{\mathrm{d}\mathbf{E}}\boldsymbol{\mu}^{\mathrm{Mol}} \tag{2.8.4}$$

$$A^{\mathrm{Mol}} = \mathcal{A}^{\mathrm{Mol}} = \frac{\mathrm{d}}{\mathrm{d}\mathbf{E}}\mathbf{Q}^{\mathrm{Mol}} \tag{2.8.5}$$

$$G'^{\mathrm{Mol}} = -\left(\mathcal{G}'^{\mathrm{Mol}}\right)^T = -\left(\frac{\mathrm{d}}{\mathrm{d}\mathbf{E}}\mathbf{m}^{\mathrm{Mol}}\right)^T \tag{2.8.6}$$

A schematic illustration of this workflow to obtain the required electromagnetic moments is presented in Figure 2.8.1,[MB26,MB40] where the rows represent successive time steps of the standard BOMD simulation, the columns stand for volumetric and molecular properties, and the red font depicts quantities under the influence of an external electric field. The $^{\mathrm{Mol}}$ superscript will be omitted in the following, but all quantities are still molecular quantities.

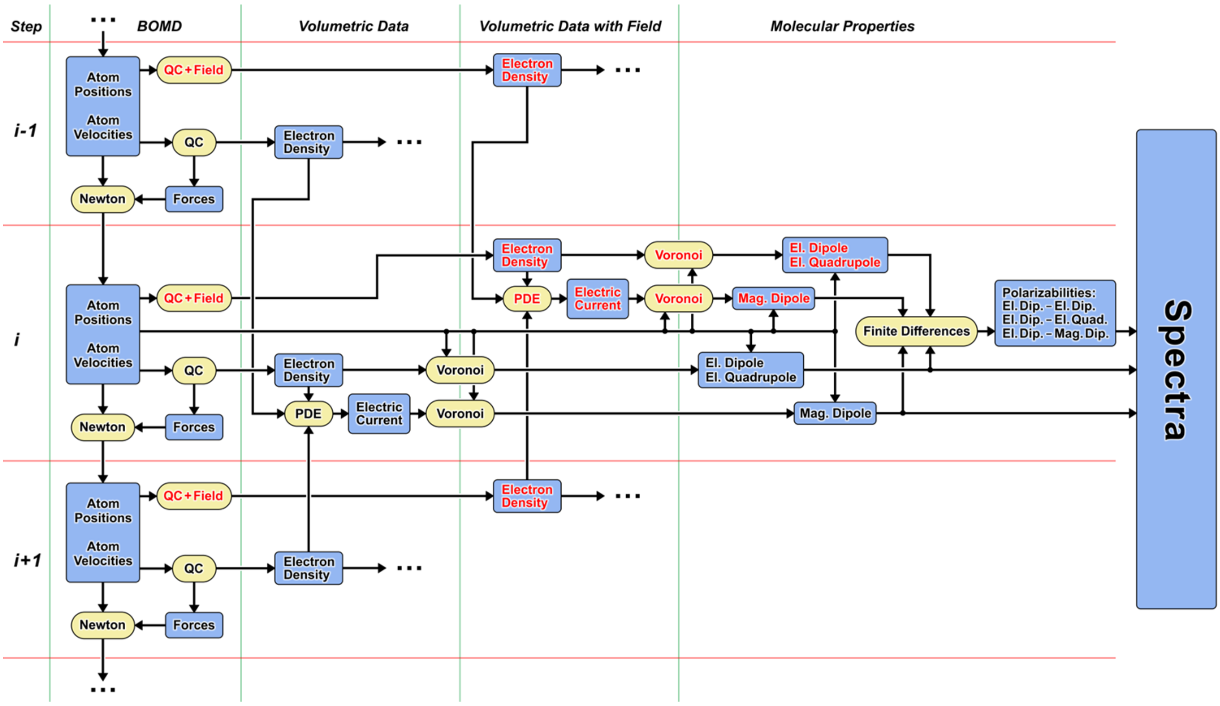

Step
BOMD
Volumetric Data
Volumetric Data with Field
Molecular Properties
i-1
i
i+1
Atom Positions
Atom Velocities
QC+Field
QC
Newton
Forces
Electron Density
PDE
Electric Current
Voronoi
El. Dipole
El. Quadrupole
Mag. Dipole
Finite Differences
Polarizabilities:
El. Dip. – El. Dip.
El. Dip. – El. Quad.
El. Dip. – Mag. Dip.
Spectra

Figure 2.8.1.: Schematic approach to compute the electromagnetic moments for ROA spectra from standard BOMD simulations.[MB26,MB40] Rows represent successive BOMD time steps. Red font depicts quantities under external electric field.

The ROA spectrum consists of three invariants, aG', $\gamma^2_{G'}$, and γ^2_A. In static ROA calculations, these are computed from the derivatives of the polarizability tensor elements along each normal mode.[304] In our new approach, we calculate the invariants as the Fourier transform of cross-correlation functions between the time series of certain molecular quantities along the trajectory. These quantities are the tensor G' as well as the time derivatives $\dot{\alpha}$ and $\dot{A}$ of the tensors defined above. Only quantities from one molecule at a time are considered; cross-terms between different molecules are not included. Following this approach, the ROA invariants are functions of the wave number $\tilde{\nu}$ rather than scalar values as in the static case:

$$\begin{aligned} aG'(\tilde{\nu}) = 2\pi c\tilde{\nu}_{\text{in}} \int_{-\infty}^{\infty} & \left\langle \frac{\dot{\alpha}_{xx}(\tau)+\dot{\alpha}_{yy}(\tau)+\dot{\alpha}_{zz}(\tau)}{3} \, \frac{G'_{xx}(\tau{+}t)+G'_{yy}(\tau{+}t)+G'_{zz}(\tau{+}t)}{3} \right\rangle_{\tau} \\ & \cdot \exp(-2\pi \mathrm{i} c\tilde{\nu}t)\, \mathrm{d}t, \end{aligned} \tag{2.8.7}$$

$$\begin{aligned} \gamma^2_{G'}(\tilde{\nu}) = 2\pi c\tilde{\nu}_{\text{in}} \int_{-\infty}^{\infty} & \Bigg[\frac{1}{2} \left\langle \left(\dot{\alpha}_{xx}(\tau)-\dot{\alpha}_{yy}(\tau)\right) \left(G'_{xx}(\tau{+}t)-G'_{yy}(\tau{+}t)\right) \right\rangle_{\tau} \\ & + \frac{1}{2} \left\langle \left(\dot{\alpha}_{yy}(\tau)-\dot{\alpha}_{zz}(\tau)\right) \left(G'_{yy}(\tau{+}t)-G'_{zz}(\tau{+}t)\right) \right\rangle_{\tau} \\ & + \frac{1}{2} \left\langle \left(\dot{\alpha}_{zz}(\tau)-\dot{\alpha}_{xx}(\tau)\right) \left(G'_{zz}(\tau{+}t)-G'_{xx}(\tau{+}t)\right) \right\rangle_{\tau} \\ & + \frac{3}{2} \left\langle \dot{\alpha}_{xy}(\tau) \left(G'_{xy}(\tau{+}t)+G'_{yx}(\tau{+}t)\right) \right\rangle_{\tau} \\ & + \frac{3}{2} \left\langle \dot{\alpha}_{yz}(\tau) \left(G'_{yz}(\tau{+}t)+G'_{zy}(\tau{+}t)\right) \right\rangle_{\tau} \\ & + \frac{3}{2} \left\langle \dot{\alpha}_{zx}(\tau) \left(G'_{zx}(\tau{+}t)+G'_{xz}(\tau{+}t)\right) \right\rangle_{\tau} \Bigg] \exp(-2\pi \mathrm{i} c\tilde{\nu}t)\, \mathrm{d}t, \end{aligned} \tag{2.8.8}$$

$$\begin{aligned} \gamma^2_{A}(\tilde{\nu}) = \pi c\tilde{\nu}_{\text{in}} \int_{-\infty}^{\infty} & \Bigg[\left\langle \left(\dot{\alpha}_{yy}(\tau)-\dot{\alpha}_{xx}(\tau)\right) \dot{A}_{z,xy}(\tau{+}t) \right\rangle_{\tau} \\ & + \left\langle \left(\dot{\alpha}_{xx}(\tau)-\dot{\alpha}_{zz}(\tau)\right) \dot{A}_{y,zx}(\tau{+}t) \right\rangle_{\tau} + \left\langle \left(\dot{\alpha}_{zz}(\tau)-\dot{\alpha}_{yy}(\tau)\right) \dot{A}_{x,yz}(\tau{+}t) \right\rangle_{\tau} \\ & + \left\langle \dot{\alpha}_{xy}(\tau) \left(\dot{A}_{y,yz}(\tau{+}t)-\dot{A}_{z,yy}(\tau{+}t)+\dot{A}_{z,xx}(\tau{+}t)-\dot{A}_{x,xz}(\tau{+}t)\right) \right\rangle_{\tau} \\ & + \left\langle \dot{\alpha}_{yz}(\tau) \left(\dot{A}_{z,zx}(\tau{+}t)-\dot{A}_{x,zz}(\tau{+}t)+\dot{A}_{x,yy}(\tau{+}t)-\dot{A}_{y,yx}(\tau{+}t)\right) \right\rangle_{\tau} \\ & + \left\langle \dot{\alpha}_{zx}(\tau) \left(\dot{A}_{y,zz}(\tau{+}t)-\dot{A}_{z,zy}(\tau{+}t)+\dot{A}_{x,xy}(\tau{+}t)-\dot{A}_{y,xx}(\tau{+}t)\right) \right\rangle_{\tau} \Bigg] \\ & \cdot \exp(-2\pi \mathrm{i} c\tilde{\nu}t)\, \mathrm{d}t, \end{aligned} \tag{2.8.9}$$

where $\tilde{\nu}_{\text{in}}$ denotes the wave number of the incident laser light. Note that the "commutator trick" can be applied when computing the cross-correlation functions to improve the sampling quality—see Equations 2.12.5 and 2.12.6 in Section 2.12.4.

Table 2.8.1.: Coefficients in Equation 2.8.10 for ICP, SCP, DCPin operating modes and several scattering geometries; polarization relative to scattering plane; 0°=forward scattering, 180°=backward scattering.

Scattering angle	Polarization	X	Y	Z
0°	$\Delta I^{\perp} = \Delta I^{\parallel}$	360	8	-8
0°	ΔI	720	16	-16
90°	$\Delta I^{\perp}$	180	28	4
90°	$\Delta I^{\parallel}$	0	24	-8
90°	ΔI	180	52	-4
180°	$\Delta I^{\perp} = \Delta I^{\parallel}$	0	48	16
180°	ΔI	0	96	32

The final ROA spectrum $\Delta I(\tilde{\nu})$ is assembled as a linear combination of these three invariants:[304]

$$\Delta I(\tilde{\nu}) = \frac{h}{8\varepsilon_0^2 c k_{\mathrm{B}} T} \cdot \frac{(\tilde{\nu}_{\mathrm{in}} - \tilde{\nu})^4}{\tilde{\nu}\left(1 - \exp\left(-\frac{hc\tilde{\nu}}{k_{\mathrm{B}}T}\right)\right)} \cdot \frac{1}{90}\Big(X \cdot aG'(\tilde{\nu}) + Y \cdot \gamma_{G'}^2(\tilde{\nu}) + Z \cdot \gamma_A^2(\tilde{\nu})\Big), \tag{2.8.10}$$

where the values of the coefficients X, Y, and Z depend on the experimental setup for which the spectrum shall be predicted. The values of the coefficients for common scattering geometries are taken from the literature[304] and shown in Table 2.8.1. These values are applicable to the following experimental operating modes: Incident circularly polarized radiation (ICP), scattered circularly polarized radiation (SCP), and in-phase dual circularly polarized radiation (DCPin). In the case of out-of-phase dual circularly polarized radiation (DCPout), no signal is obtained within this approximation. The frequency-dependent pre-factor in Equation 2.8.10 ensures that the correct absolute intensity values are obtained. Some technical details which are important to obtain high-quality spectra are discussed in Section 2.12.

Some examples for condensed phase ROA spectra predicted by our method are presented in Section 3.4.

2.9. SFG Spectroscopy

As described in the introduction, sum frequency generation (SFG) spectroscopy is a non-linear laser spectroscopy which was developed in 1987 by the group of Shen.[37,38] As a surface-selective spectroscopy, it is very valuable, and has been applied to many interfacial system since then.[39–42,346–352] The information contained in a SFG spectrum is in a certain sense complementary to infrared and Raman spectroscopy.[42]

Within the electric dipole approximation,[159] SFG is a second-order non-linear process with an intensity I_{SFG} directly proportional to the square of the second-order effective polarization $P^{(2)}_{\mathrm{eff}}$.[353–355] This process is induced by the non-linear second-order effective susceptibility $\chi^{(2)}_{\mathrm{eff}}$ in the presence of the two incident laser beams *(visible pump and infrared probe)*. Thus, it is found that

$$I_{\mathrm{SFG}} \propto \left| P^{(2)}_{\mathrm{eff}} \right|^2 = \left| \chi^{(2)}_{\mathrm{eff}} \mathbf{E}_{\mathrm{vis}} \mathbf{E}_{\mathrm{IR}} \right|^2 \tag{2.9.1}$$

with $\mathbf{E}_{\mathrm{vis}}$ and $\mathbf{E}_{\mathrm{IR}}$ being the electric fields of the incident visible and IR laser beams, respectively. $\chi^{(2)}_{\mathrm{eff}}$ is a third-rank tensor which is composed of tensorial Fresnel factors at the interface and the macroscopic second-order non-linear susceptibility $\chi^{(2)}_{ijk}$. Based on that, the SFG intensity can be expressed[159] as

$$I_{\mathrm{SFG}} \propto \left| \sum_i \sum_{j,k} F_{ijk} \chi^{(2)}_{ijk} \mathbf{E}_{\mathrm{vis}} \mathbf{E}_{\mathrm{IR}} \right|^2 , \qquad i,j,k \in \{x,y,z\}, \tag{2.9.2}$$

where F_{ijk} are combined Fresnel coefficients which contain the non-linear Fresnel coefficient for a SFG field originating from a monolayer at the interface between the two media as well as the Fresnel coefficients of visible and IR incoming beams.

The macroscopic susceptibility $\chi^{(2)}_{ijk}$ can be expressed as the sum of two contributions. $\chi^{(2)}_{\mathrm{NR}}$ is a near-resonant susceptibility which is used for describing SFG intensity contributions from the substrate. $\chi^{(2)}_{\mathrm{R}}$, on the other hand, is the second-order resonant susceptibility, containing orientation information as well as the second-order hyperpolarizability of the molecule on the surface. There are a total of 27 elements in the latter tensor, but due to surface symmetry considerations, the number of independent elements is significantly smaller. For example, in C_{2v} surface symmetry such as found on TiO_2 (110), there exist the following five non-vanishing unique components:[346]

$$\chi^{(2)}_{xxz}, \qquad \chi^{(2)}_{yyz}, \qquad \chi^{(2)}_{zzz}, \qquad \chi^{(2)}_{xzx} = \chi^{(2)}_{zxx}, \qquad \chi^{(2)}_{yzy} = \chi^{(2)}_{zyy}. \tag{2.9.3}$$

When the molecular coordinate system is aligned with the laboratory coordinate system and the plane of incidence is set to the x–z plane, it is convenient to regroup the macroscopic susceptibility tensor elements according to the distinct polarization combinations of the three optical fields *(SFG, visible, and IR)*. At the surface, the light can be expressed as a combination of two linear polarizations, s and p. The s-polarized and p-polarized lights have electric field vectors perpendicular and parallel to the plane of incidence, respectively. As a result, s-polarized light contains the y component, and p-polarized light contains the x and z components of the optical field. Thus, there are a total of eight possible unique combinations of polarizations in a SFG experiment. The effective resonant susceptibility tensor elements related to the experimental polarization combinations on a C_{2v} surface are as follows:[159]

$$\begin{aligned}
\chi^{(2)}_{\mathrm{eff,R},sss} &= \chi^{(2)}_{\mathrm{eff,R},spp} = \chi^{(2)}_{\mathrm{eff,R},psp} = \chi^{(2)}_{\mathrm{eff,R},pps} = 0 \\
\chi^{(2)}_{\mathrm{eff,R},ssp} &= F_{yyz}\chi^{(2)}_{yyz} \\
\chi^{(2)}_{\mathrm{eff,R},sps} &= F_{yzy}\chi^{(2)}_{yzy} \\
\chi^{(2)}_{\mathrm{eff,R},pss} &= F_{zyy}\chi^{(2)}_{zyy} \\
\chi^{(2)}_{\mathrm{eff,R},ppp} &= -F_{xxz}\chi^{(2)}_{xxz} - F_{xzx}\chi^{(2)}_{xzx} + F_{zxx}\chi^{(2)}_{zxx} + F_{zzz}\chi^{(2)}_{zzz}
\end{aligned} \tag{2.9.4}$$

As shown here, only four of the polarization combinations, namely ssp, sps, pss, and ppp, contribute to the SFG signal on a C_{2v} surface.

The resonant part of the non-linear second-order effective susceptibility tensor $\chi^{(2)}_{\mathrm{eff}}$ can be computed from molecular dynamics simulations by the time-dependent approach introduced by Morita *et al.*,[356–358] based on the time-correlation function (TCF) formalism. Based on this approach, many SFG spectra have already been predicted from MD simulations of interfaces.[153–165,293,359] In Morita's approach, $\chi^{(2)}_{\mathrm{eff,R}}$ can be computed as

$$\chi^{(2)}_{\mathrm{eff,R},pqr}(\omega) = \frac{-\mathrm{i}}{k_\mathrm{B}T\omega} \sum_{i=1}^{N} \int \left\langle \dot{\alpha}^{i}_{pq}(\tau) \cdot \dot{\mu}^{i}_{r}(t+\tau) \right\rangle_{\tau} \exp(\mathrm{i}\omega t)\ \mathrm{d}t, \tag{2.9.5}$$

where $\dot{\mu}^{i}_{r}(\tau)$ and $\dot{\alpha}^{i}_{pq}(\tau)$ are the time derivatives of the molecular dipole moment and the molecular polarizability tensor of molecule i at time τ, respectively, N is the number of molecules, k_B is Boltzmann's constant, and $p, q, r \in \{x, y, z\}$ in the laboratory coordinate frame, so that all 27 tensor elements can be obtained. The molecular dipole moments and polarizabilities can be conveniently computed via our Voronoi integration approach[MB23,MB48]—see Section 2.1.

2.10. Resonance Raman Spectroscopy

As described in the introduction, the most prominent limitation of Raman-based spectroscopic methods is their low sensitivity. Out of one million incident photons, typically less than one undergoes Raman scattering, so that a very weak signal is obtained. In the 1960s, it was discovered that the intensity of the Raman scattering drastically increases if the incident laser wavelength is close to the energy of an electronic excitation in the sample, which was termed as the resonance Raman effect.[17–21] It followed the development of the resonance Raman spectroscopy, which is a very important experimental technique today.[22–28]

Predicting resonance Raman spectra turned out to be quite intricate. As the resonance Raman effect involves electronic excitation of the sample, it is in general not sufficient to solve the time-independent Schrödinger equation to obtain the scattering intensities. Several approaches to compute these intensities have been presented in the literature. Many of them are based on the vibronic theory of Albrecht and co-workers[360–362] or on the time-dependent formalism of Heller and co-workers.[363–366] Another method, based on linear response time-dependent density functional theory (LR-TDDFT),[367] was published by Jensen and Schatz.[368,369] Scattering intensities obtained from these methods have been used in many studies in the literature.[73–82]

Recently, a different approach which uses real-time time-dependent density functional theory (RT-TDDFT)[370–374] to obtain the dynamic polarizability of the sample appeared in the literature.[375] In contrast to the methods mentioned before, the real-time approach offers the advantage of including all electronic excitations into the calculation, such that the full frequency range is covered and no subset of low-lying excitations needs to be selected. Furthermore, this approach intrinsically includes non-linear effects which are neglected in perturbative methods such as LR-TDDFT. Within the last years, a few studies on computing resonance Raman scattering intensities from RT-TDDFT were published.[376–378] However, all spectra presented in these studies were based on static–harmonic frequency calculations of isolated molecules in vacuum, and therefore lack solvent influence and anharmonic effects as discussed above. To the best of our knowledge, a predicted resonance Raman spectrum of a bulk phase system has not been published before.

Our newly proposed method[MB36] is based on a Born–Oppenheimer molecular dynamics (BOMD) simulation of the bulk phase system of interest. Along the BOMD trajectory, snapshots of the system are stored in equidistant intervals. For each of those snapshots, a separate RT-TDDFT calculation with fixed geometry is started, as schematically shown in Figure 2.10.1.

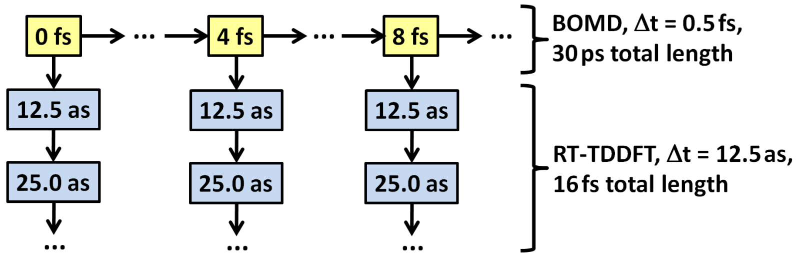


Figure 2.10.1.: Schematic simulation protocol of our approach to resonance Raman spectra. The horizontal direction depicts the BOMD simulation, while the vertical direction shows the RT-TDDFT calculation of each BOMD snapshot.[MB36]

The initial wave function is optimized under the influence of an external periodic electric field, which is switched off in the beginning of the RT-TDDFT run, so that the electron density starts to fluctuate (step response). During the RT-TDDFT run, the temporal development of the total electron density is stored in compressed bqb format[MB34] *(see Section 2.2)*, and subsequently processed with our Voronoi integration scheme[MB23,MB48] *(see Section 2.1)* to yield time series of molecular electric dipole vectors $\boldsymbol{\mu}(\tau, t)$ with BOMD time t and RT-TDDFT time τ. The Fourier transform of the three dipole vector components yields three entries of the molecular dynamic polarizability tensor, as shown for an example in Figure 2.10.2.

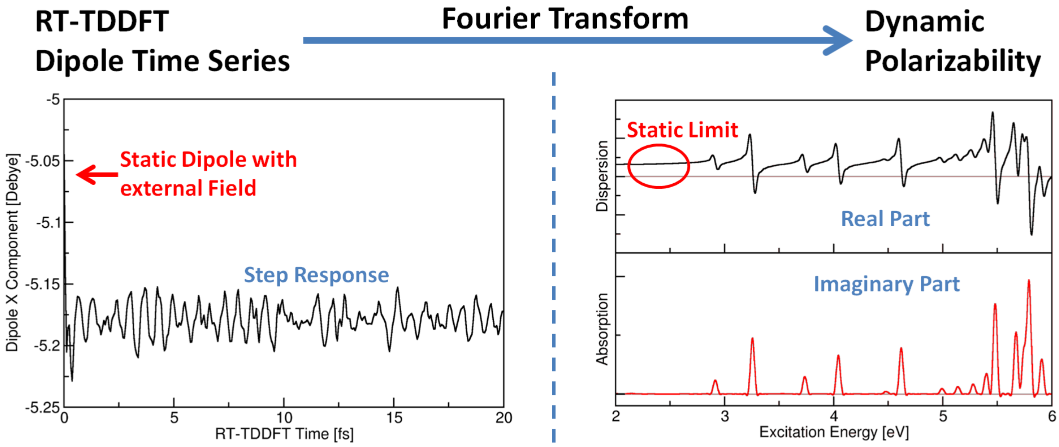


Figure 2.10.2.: The Fourier transform of the dipole moment step response in the RT-TDDFT simulation yields the real and imaginary part of the dynamic polarizability tensor.[MB36]

To obtain the full dynamic polarizability tensor $\alpha_{ij}(\omega, t)$ for each molecule (with ω the incident laser frequency and t the time along the BOMD trajectory), three RT-TDDFT runs are performed from initial wave functions optimized under external fields in X, Y, and Z direction:

$$\alpha_{ij}(\omega,t)=\frac{1}{|\mathbf{E}|}\int\limits_{0}^{T}\Big(\mu_i(\tau,t)-\mu_i^{0,\mathrm{j}}(t)\Big)\exp\Big(\big(-c\frac{\tau}{T}\big)^2\Big)\exp\Big(-\mathrm{i}\omega\tau\Big)\,\mathrm{d}\tau \qquad (2.10.1)$$

Please note that we use a Gaussian window function with parameter c for the Fourier transform of the RT-TDDFT time series. T is the total RT-TDDFT simulation time, $\boldsymbol{\mu}^{0,\mathrm{j}}(t)$ denotes the initial molecular dipole moment at BOMD time t after wave function optimization under external electric field in j direction, and $|\mathbf{E}|$ is the absolute value of the external electric field. Please also note that the dynamic polarizability tensor obtained from Equation 2.10.1 is complex-valued, with dispersion as real part and absorption as imaginary part. An example of the dynamic polarizability is shown in Figure 2.10.3.

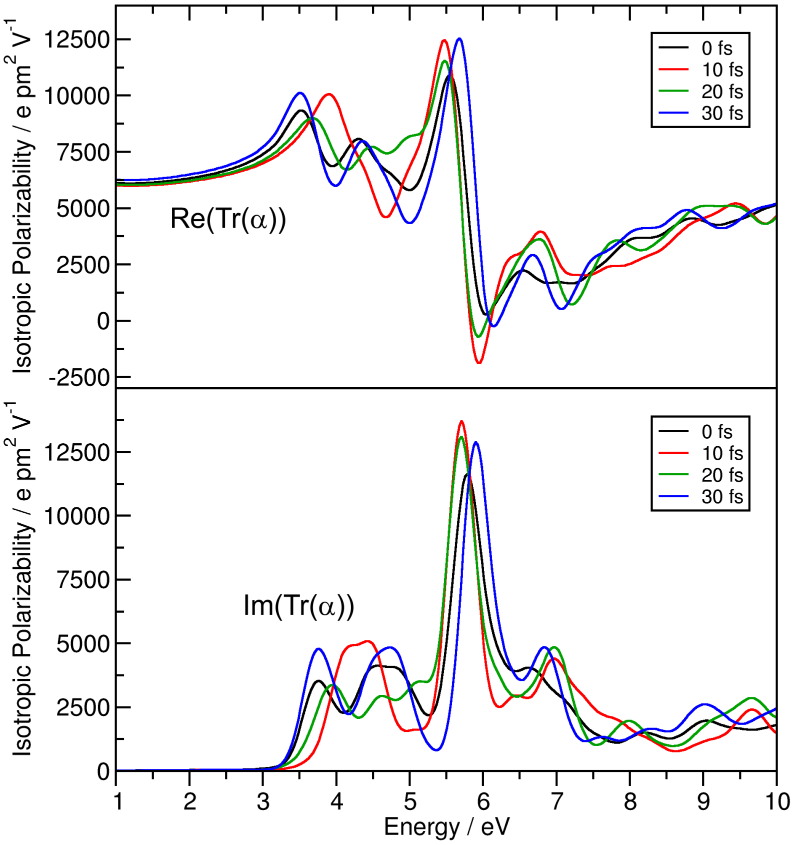


Figure 2.10.3.: Real part *(dispersion; upper panel)* and imaginary part *(absorption; lower panel)* of the trace of the dynamic polarizability tensor from a bulk phase simulation at four different BOMD simulation times.[MB36]

Based on the dynamic polarizability tensor, the resonance Raman spectrum for any incident laser wavelength can be computed as Fourier transform of autocorrelation functions along the BOMD time series, as discussed in the literature for non-resonant Raman spectra before.[MB14] We obtain the two Raman invariants, namely the isotropic polarizability $a_p^2\left(\tilde{\nu},\omega\right)$ and the anisotropy $\gamma_p^2\left(\tilde{\nu},\omega\right)$, as follows:

$$\begin{aligned} a_p^2\left(\tilde{\nu},\omega\right) = \int\limits_0^{\hat{\mathbf{T}}} \frac{1}{9}\Big\langle &\left(\dot{\alpha}_{xx}(\omega,\tau) + \dot{\alpha}_{yy}(\omega,\tau) + \dot{\alpha}_{zz}(\omega,\tau)\right) \\ &\cdot\left(\dot{\alpha}_{xx}(\omega,\tau{+}t) + \dot{\alpha}_{yy}(\omega,\tau{+}t) + \dot{\alpha}_{zz}(\omega,\tau{+}t)\right)\Big\rangle_\tau \exp\Big(-2\pi\mathrm{i}c\tilde{\nu}t\Big)\,\mathrm{d}t \end{aligned} \tag{2.10.2}$$

$$\begin{aligned} \gamma_p^2\left(\tilde{\nu},\omega\right) = \int\limits_0^{\hat{\mathbf{T}}} \Bigg[&\frac{1}{2}\Big\langle\left(\dot{\alpha}_{xx}(\omega,\tau) - \dot{\alpha}_{yy}(\omega,\tau)\right)\cdot\left(\dot{\alpha}_{xx}(\omega,\tau{+}t) - \dot{\alpha}_{yy}(\omega,\tau{+}t)\right)\Big\rangle_\tau \\ &+\frac{1}{2}\Big\langle\left(\dot{\alpha}_{yy}(\omega,\tau) - \dot{\alpha}_{zz}(\omega,\tau)\right)\cdot\left(\dot{\alpha}_{yy}(\omega,\tau{+}t) - \dot{\alpha}_{zz}(\omega,\tau{+}t)\right)\Big\rangle_\tau \\ &+\frac{1}{2}\Big\langle\left(\dot{\alpha}_{zz}(\omega,\tau) - \dot{\alpha}_{xx}(\omega,\tau)\right)\cdot\left(\dot{\alpha}_{zz}(\omega,\tau{+}t) - \dot{\alpha}_{xx}(\omega,\tau{+}t)\right)\Big\rangle_\tau \\ &+3\Big\langle\dot{\alpha}_{xy}(\omega,\tau)\cdot\dot{\alpha}_{xy}(\omega,\tau{+}t)\Big\rangle_\tau + 3\Big\langle\dot{\alpha}_{xz}(\omega,\tau)\cdot\dot{\alpha}_{xz}(\omega,\tau{+}t)\Big\rangle_\tau \\ &+3\Big\langle\dot{\alpha}_{yz}(\omega,\tau)\cdot\dot{\alpha}_{yz}(\omega,\tau{+}t)\Big\rangle_\tau\Bigg] \exp\Big(-2\pi\mathrm{i}c\tilde{\nu}t\Big)\,\mathrm{d}t \end{aligned} \tag{2.10.3}$$

The integrals run over the total BOMD simulation time $\hat{\mathbf{T}}$. Taking the time derivative $\dot{\boldsymbol{\alpha}}$ of the dynamic polarizability tensor $\boldsymbol{\alpha}$ is equivalent to a factor of $\tilde{\nu}^2$ outside of the Fourier transform.[MB14] Following from the definition

$$\left\langle f(\tau)\cdot f(\tau{+}t)\right\rangle_\tau := \int f(\tau{+}t)\cdot\overline{f(\tau)}\,\mathrm{d}\tau, \tag{2.10.4}$$

the autocorrelation of a complex time series $f(t)$ which is required here can be computed from autocorrelations and cross-correlations of real time series in the following way:[MB36]

$$\begin{aligned} \Big\langle f(t)\cdot f(t{+}\tau)\Big\rangle_\tau = &\Big\langle\mathrm{Re}\big(f(t)\big)\cdot\mathrm{Re}\big(f(t{+}\tau)\big)\Big\rangle_\tau + \Big\langle\mathrm{Im}\big(f(t)\big)\cdot\mathrm{Im}\big(f(t{+}\tau)\big)\Big\rangle_\tau \\ &+ i\left[\Big\langle\mathrm{Re}\big(f(t)\big)\cdot\mathrm{Im}\big(f(t{+}\tau)\big)\Big\rangle_\tau - \Big\langle\mathrm{Im}\big(f(t)\big)\cdot\mathrm{Re}\big(f(t{+}\tau)\big)\Big\rangle_\tau\right] \end{aligned} \tag{2.10.5}$$

Finally, the resonance Raman spectrum $I(\tilde{\nu},\omega)$ at laser frequency ω can be expressed as a linear combination of the two invariants $a_p^2(\tilde{\nu},\omega)$ and $\gamma_p^2(\tilde{\nu},\omega)$, where the coefficients X and Y depend on the scattering geometry and polarization and can be found in the literature. Common choices for $(X;Y)$ are $(0;3)$ for orthogonal polarization, $(45;4)$ for parallel polarization, and $(45;7)$ for unpolarized/depolarized setups *(with a scattering angle of 90° in all three cases)*.[304]

$$I(\tilde{\nu},\omega) = \frac{h}{8\varepsilon_0^2 c k_\mathrm{B} T} \cdot \frac{(\omega-\tilde{\nu})^4}{\tilde{\nu}\left(1-\exp\left(-\frac{hc\tilde{\nu}}{k_\mathrm{B}T}\right)\right)} \cdot \frac{1}{45}\left(X \cdot a_p^2(\tilde{\nu},\omega) + Y \cdot \gamma_p^2(\tilde{\nu},\omega)\right) \tag{2.10.6}$$

As the RT-TDDFT run yields the dynamic polarizability tensor over the full frequency range *(see Equation 2.10.1)*, we obtain the resonance Raman spectra for all possible laser wavelengths ω from Equation 2.10.6 in one pass—see *e. g.* Figure 3.5.3 on page 102. Such a two-dimensional spectrum—also known as "excitation profile"—can reveal interesting details on the vibronic coupling in the system. This is a clear advantage over other methods, where the resonance Raman intensities are often computed for a single laser wavelength per pass. The low-frequency part of the polarizability is correctly reproduced and matches the static polarizability obtained via finite electric field differences. Therefore, also the non-resonant Raman spectrum is correctly obtained from our approach. We implemented the methodology in our freely available open-source program package TRAVIS.[MB2,MB40]

A few examples for resonance Raman spectra predicted via this approach are shown in Section 3.5.

2.11. Normal Mode Analysis

When computing spectra in the static–harmonic framework, each spectral band is obtained together with a corresponding normal mode *(see Section 1.2.1)* so that the spectral features can be easily assigned to specific molecular vibrations. When computing spectra from AIMD simulations via the time-correlation function (TCF) formalism, this cannot be easily achieved, because the spectrum is obtained as a superposition of all vibrational modes in the system without additional information on specific molecular motions that contribute to a certain spectral feature.

To overcome this limitation, several approaches for the extraction of normal modes from molecular dynamics simulations have been reported in the literature. These approaches include the instantaneous normal mode analysis (INMA) where the Hessian of the system is calculated in certain time steps along the trajectory.[379–382] Moreover, it includes the principal mode analysis (PMA) where an eigenvalue problem—derived from cross-correlation functions of particle positions and velocities—is solved.[122,276,380,381,383–388] For our purpose, we have adopted the generalized normal coordinate scheme of Mathias *et al.*,[389,390] which is similar to the PMA approach, but does not require the equipartition theorem to be fulfilled. A short summary of the approach will be given in the following paragraphs, closely following the description in our publication.[MB19] The methodology is based on computing power spectra which are discussed in Section 2.3.

The central object in the approach is the matrix $\mathbf{P}(\omega)$ of velocity cross-correlation spectra *(closely related to the tensorial vibrational density of states in Ref. 389)*. For a molecule consisting of n atoms, this is a $3n \times 3n$ matrix with the elements

$$P_{ij}(\omega) := \sqrt{m_i m_j}\ \mathrm{Re}\left(\int \left\langle \dot{x}_i(\tau)\, \dot{x}_j(\tau+t) \right\rangle_\tau \exp(-\mathrm{i}\omega t)\ \mathrm{d}t \right) \tag{2.11.1}$$

where t and τ denote the time and ω is the frequency. The indices i and j run over the $3n$ Cartesian coordinates of all atoms in the molecule. $\dot{x}_i$ denotes the time derivative *(velocity)* of coordinate i and m_i is the mass of the atom corresponding to coordinate i. $\left\langle \dot{x}_i(\tau)\, \dot{x}_j(\tau+t) \right\rangle_\tau$ denotes the cross-correlation function of the velocities corresponding to coordinates i and j. Taking the real part of the Fourier transform corresponds to averaging over the time forward and the time backward trajectory.[389] The main-diagonal elements $P_{ii}(\omega)$ are equivalent to atom specific power spectra along one Cartesian axis, so the trace of $\mathbf{P}(\omega)$ gives the total power spectrum of the system as described in the literature.[MB14] The off-diagonal elements $P_{ij}(\omega)$ are the Fourier transforms of velocity cross-correlation functions, which characterize, in a simplified view, the overlap of the atom specific power spectra $P_{ii}(\omega)$ and $P_{jj}(\omega)$.

Since the Cartesian coordinates of the atoms are used, the atom specific power spectra $P_{ii}(\omega)$ usually range over a broad frequency region with contributions at all molecular vibrations involving atom i. The main idea is to find a coordinate transformation such that the main-diagonal elements $P_{ii}(\omega)$ are sharply located in frequency space. This is equivalent to minimizing the off-diagonal elements $P_{ij}(\omega)$ describing the overlaps. The new coordinates are considered as the generalized normal coordinates of the system.[389]

If $\mathbf{P}(\omega)$ was a regular matrix, the minimization of the off-diagonal elements could be achieved by a simple diagonalization. However, $\mathbf{P}(\omega)$ depends on the frequency ω so that standard diagonalization techniques cannot be used. Still it is possible to minimize the integral over the off-diagonal elements by a modified Jacobi algorithm,[211,389] leading to an orthogonal transformation matrix $\mathbf{C}$ that gives the new matrix

$$\mathbf{P}'(\omega) = \mathbf{C}\mathbf{P}(\omega)\,\mathbf{C}^{-1} \tag{2.11.2}$$

with minimized cross-correlation spectra $P'_{ij}(\omega)$. The matrix $\mathbf{C}$ consists of the set of new coordinate vectors $\mathbf{q}_k$ characterizing the normal modes. Since the trace of a matrix is not changed by an orthogonal transformation, the total power spectrum of the system is not modified by this procedure.

The concept of normal modes generally relies on the assumption that the molecule undertakes small oscillations around a fixed reference structure, which is usually a minimum on the potential energy surface. However, in a molecular dynamics trajectory, there will also be translational and rotational motion, which needs to be removed prior to the normal coordinate analysis. For this purpose, the trajectories are transformed to the Eckart frame of reference[391] by finding a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{T}$ in each time step such that the mass-weighted root-mean-square distance

$$d := \sqrt{\sum_{k=1}^{n} m_k \left(\mathbf{R}\mathbf{r}_k(t) + \mathbf{T} - \mathbf{r}_k^0\right)^2} \tag{2.11.3}$$

is minimized, with $\mathbf{r}_k(t)$ and $\mathbf{r}_k^0$ being the position vectors of atom k in the trajectory and the reference structure, respectively.

The assumption of small oscillations around one single reference structure breaks down if there are conformational changes in the trajectory. In this context, the procedure needs to be extended to several reference structures.[390] Either these can differ only in the ordering of equivalent atoms *(e. g., if a methyl group rotates where the hydrogen atoms are indistinguishable)* or they belong to structurally different minima on the potential energy surface *(e. g., if a butyl group changes between trans and gauche conformation)*. In the former case, all conformations can be mapped to a single minimum by considering the permutations of equivalent atoms, whereas in the latter case, the normal coordinate analysis results in independent normal modes for all minima.

If there are several reference structures, a probability $p_m(t)$ for being in structure m has to be defined in each time step. A possible choice[390] is of the general form

$$p_m(t) := N \exp\left(-\frac{d\left(\mathbf{r}(t)\,,\mathbf{r}^{0,m}\right)}{2\sigma^2}\right), \tag{2.11.4}$$

where $d\left(\mathbf{r}(t)\,,\mathbf{r}^{0,m}\right)$ is a function measuring the distance between the Eckart frame coordinates from the trajectory and reference structure m, σ is a parameter determining the width of the switching region between two reference structures, and N is a proper normalization factor to ensure that the sum of all probabilities is 1. The distance function $d\left(\mathbf{r}(t)\,,\mathbf{r}^{0,m}\right)$ can either be the mass-weighted root-mean-square distance analogous to Equation 2.11.3 or a root-mean-square deviation in certain internal coordinates.[390] With the probabilities $p_m(t)$, the cross-correlation function $\left\langle \dot{x}_i(\tau)\,\dot{x}_j(\tau+t)\right\rangle_\tau$ in Equation 2.11.1 is replaced by $\left\langle p_m(\tau)\,\dot{x}_i(\tau)\,\dot{x}_j(\tau+t)\right\rangle_\tau$, resulting in a cross-correlation matrix $\mathbf{P}_m(\tau)$ for each reference structure. Properly accounting for the permutation operations, the matrices $\mathbf{P}_m(\tau)$ of reference structures just differing by the ordering of equivalent atoms can be added, finally leading to a single set of normal modes for the corresponding minimum on the potential energy surface.[390]

An example for a normal mode decomposition in a bulk phase simulation is discussed in Section 3.6.

2.12. Signal Processing Techniques

This section describes some general techniques used in the scope of this thesis (and implemented in TRAVIS[MB2,MB40]) which are relevant for the calculation of all types of vibrational spectra. While most of these techniques are basic knowledge of signal processing, they are rarely discussed in articles on computer simulations.

2.12.1. Fast Correlation

All prediction methods for vibrational spectra from molecular dynamics simulations are based on either autocorrelation or cross-correlation functions of some time series along the trajectory. Typically, these correlations are computed for molecular properties, not for properties of the whole simulation box. The reason is that correlating properties of the whole system is equivalent to taking into account all cross-correlations between different molecules. However, the motion of very distant molecules is not correlated at all, so that these cross-correlations mostly introduce noise into the spectrum. Computing all the molecular correlation functions requires much computational power, because such a calculation scales with $\mathcal{O}(m \cdot n)$, where n is the length of the time series, and m is the chosen correlation depth. Luckily, there exists the Wiener–Khinchin theorem,[281,282]—see also Section A.2 in the Appendix—which states that computing the autocorrelation $C(\tau)$ of a time series $f(t)$ is equivalent to Fourier transforming the time series, taking the absolute square, and transforming back:

$$C(\tau) = \int_{-\infty}^{\infty} \overline{f(t)} \cdot f(t+\tau) \; \mathrm{d}t = \mathcal{F}^{-1}\left(\left| \mathcal{F}\Big(f(\tau) \Big) \right|^2 \right) \tag{2.12.1}$$

As the Fourier transform of a discrete data set of length n can be efficiently computed by the fast Fourier transform (FFT)[392,393] algorithm which only scales with $\mathcal{O}(n \cdot \log n)$, this saves a lot of computational time. Equation 2.12.1 is only a special case of the cross-correlation theorem, so that a similar approach can also be used to speed up the calculation of cross-correlation functions via FFT. TRAVIS computes all autocorrelation and cross-correlation functions based on these approaches, using the KISS FFT library[394] *(integrated in the* TRAVIS *source code).*

2.12.2. Window Function and Zero Padding

When considering Fourier transforms of data, it is very important to choose a suitable window function. There exists no Fourier transform without a window function—using no window function means using a rectangular window, which is a very poor choice. Applying a window function simply means element-wise multiplication of the discrete input data with the window function before Fourier transforming. By default, TRAVIS uses the Hann window function[395] for computing spectra, which is given by

$$W_n := \cos^2\left(\frac{\pi n}{2\,(N-1)}\right), \qquad n = 0, \ldots, (N-1)\,, \tag{2.12.2}$$

but other window functions (exponential, Gaussian) are also implemented. The total integral of the Fourier transform of a discrete data set a_i depends only on the first value a_0. Therefore, any window function which fulfills the condition $W_0 = 1$ does not modify the total integral at all. In other words, window functions alter the peak shapes of the spectrum, but keep the integral intensities of the bands invariant.

By adding zeroes to the tail of the correlation function before Fourier transforming, the resolution of the resulting spectrum can be increased, which is called zero padding. No new spectral information is gained by this technique, it is just an interpolation *(so-called "trigonometric interpolation")* between the data points of the original spectrum, *i. e.*, a cosmetic measure. By default, TRAVIS uses a zero padding factor of 4.

2.12.3. Finite Differences Correction

For most spectra, the time derivative instead of the quantity itself enters the correlation functions. For a cosine function $\cos(\omega t+\varphi)$, it can be shown[113] that the second-order central finite difference derivative is given by

$$\begin{aligned} \mathrm{D}\cos(\omega t+\varphi) &= \frac{\cos\big(\omega(t+\Delta t)+\varphi\big) - \cos\big(\omega(t-\Delta t)+\varphi\big)}{2\Delta t} \\ &= -\omega\sin(\omega t+\varphi)\,\frac{\sin(\omega\Delta t)}{\omega\Delta t}, \end{aligned} \tag{2.12.3}$$

i. e., the exact value of the derivative is multiplied by a *sinus cardinalis* (sinc) function $\frac{\sin(\omega\Delta t)}{\omega\Delta t}$. Since the product of two such time derivatives is formed in the correlation functions, the final spectra need to be divided by a factor of $\left(\frac{\sin(\omega\Delta t)}{\omega\Delta t}\right)^2$ to obtain the correct intensities. This correction is implemented in TRAVIS and active by default. Due to the correction, high-quality spectra with accurate intensities on the full spectral range (up to $3500\,\mathrm{cm}^{-1}$) can be obtained even if the electromagnetic moments only were computed every 4.0 fs, which saves a lot of computer time.

2.12.4. Improved Sampling via Time Reversibility

Similar to most other properties, vibrational spectra should be invariant under reversal of the time direction: the trajectory with reversed time samples the same thermodynamic ensemble as the forward trajectory. And as vibrational spectra are ensemble averages, they cannot differ for these two cases. Consider the cross-correlation of two time series $a(t)$ and $b(t)$ along the trajectory. Let $\tilde{a}(t) := a(T-t)$ and $\tilde{b}(t) := b(T-t)$ be the two time series of the trajectory with reversed time, where T is the total trajectory length. We find that

$$\begin{aligned} \left\langle \tilde{a}(t+\tau) \cdot \tilde{b}(t) \right\rangle_t &= \int \tilde{a}(t+\tau) \cdot \tilde{b}(t) \ \mathrm{d}t \\ &= \int a(T-t-\tau) \cdot b(T-t) \ \mathrm{d}t \\ &= \int a(\tilde{t}) \cdot b(\tilde{t}+\tau) \ \mathrm{d}\tilde{t} \\ &= \left\langle b(\tilde{t}+\tau) \cdot a(\tilde{t}) \right\rangle_{\tilde{t}} \end{aligned} \qquad (2.12.4)$$

with the substitution $\tilde{t} := T-t-\tau$. Thus, the reversal of the time direction is equivalent to swapping the first and second argument of the cross-correlation. If the spectra are invariant under time reversal, also the underlying cross-correlation functions should be invariant. Therefore, swapping the two arguments will yield identical correlation functions for infinite trajectory length. However, in practice, sampling is not perfect, and the two correlation functions differ due to statistical noise. By simply taking the arithmetic average of both correlation functions *(with forward time and reversed time)*, see Equation 2.12.5, the sampling is effectively increased by a factor of two at no additional computational cost. We call this approach the "commutator trick".[MB40]

$$\mathrm{CCF}(a,b) := \frac{1}{2}\Big(\left\langle a(t+\tau) \cdot b(t) \right\rangle_t + \left\langle b(t+\tau) \cdot a(t) \right\rangle_t \Big) \qquad (2.12.5)$$

It only improves sampling if $a(t)$ and $b(t)$ are different time series, *i. e.*, in the case of true cross-correlation. If a and b are identical *(i. e. autocorrelation, as with correlating the dipole moment for infrared spectra)*, nothing is gained.

Another subtlety arises if one of both properties a and b is the magnetic dipole moment *(which is required for VCD and ROA spectra; see below)*. This quantity depends linearly on the atomic velocities. A time reversal inverts all atomic velocities, and therefore swaps the sign of the magnetic dipole moment. In this case, taking the average of both correlation functions would only yield noise, as both contributions would cancel out. The sign of the correlation function with reversed time has to be changed to obtain the desired effect, see Equation 2.12.6.

$$\mathrm{CCF}^*(a,b) := \frac{1}{2}\Big(\big\langle a(t+\tau)\cdot b(t)\big\rangle_t - \big\langle b(t+\tau)\cdot a(t)\big\rangle_t\Big) \tag{2.12.6}$$

Please note that this approach to enhance the sampling quality has infrequently been applied before in the literature in a less general formulation for specific spectroscopic prediction methods such as for VCD spectra.[150]

2.12.5. Correcting the Verlet Frequency Shift

Describing the dynamics of a system with finite time steps does unfortunately not leave the vibrational frequencies of harmonic oscillators invariant. If a harmonic oscillator with "true" frequency ω_0 is solved by a Verlet integration scheme[182] with integrator time step Δt, it can be derived that the observed frequency $\tilde{\omega}$ is given by

$$\tilde{\omega} = \frac{1}{\Delta t}\arccos\big(1 - \frac{1}{2}\omega_0^2\Delta t^2\big), \tag{2.12.7}$$

as demonstrated in Section A.3 in the Appendix. Luckily, the effect is rather small. When simulating an oscillator with $\omega_0 = 3000.0\,\mathrm{cm}^{-1}$ and $\Delta t = 0.5\,\mathrm{fs}$, the observed frequency is $\tilde{\omega} = 2990.0\,\mathrm{cm}^{-1}$, which is a deviation of around 0.3%. When computing spectra from a trajectory, the observed frequencies $\tilde{\omega}$ are given, and the true frequencies ω_0 are sought after. The inverse of Equation 2.12.7 can be easily written as

$$\omega_0 = \frac{\sqrt{2 - 2\cos\left(\tilde{\omega}\Delta t\right)}}{\Delta t}. \tag{2.12.8}$$

By using Equation 2.12.8, the frequency axis of the spectra can be corrected for the frequency shift of the Verlet integrator. The correction is implemented in TRAVIS. For a mathematical derivation, please see Section A.3 in the Appendix. Some typical frequency deviations for common time steps Δt are shown in Table A.1 in the Appendix.

2.13. The TRAVIS Program Package

Starting in 2009, a program package by the name of TRAVIS *("trajectory analyzer and visualizer")* for analyzing atomistic simulations has been developed within the scope of M.B.'s PhD work in the group of Barbara Kirchner. Later, Martin Thomas and Sascha Gehrke joined the developer team. Since the original publication of the package in 2011,[MB2] the article has already been cited more than 660 times, which indicates a good adoption of TRAVIS in the scientific community. A more recent article from 2020 shows a subset of the functions and abilities which TRAVIS has gained over the time.[MB40]

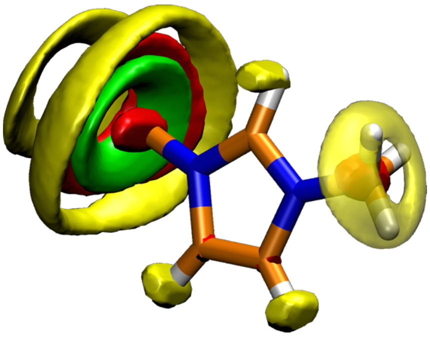
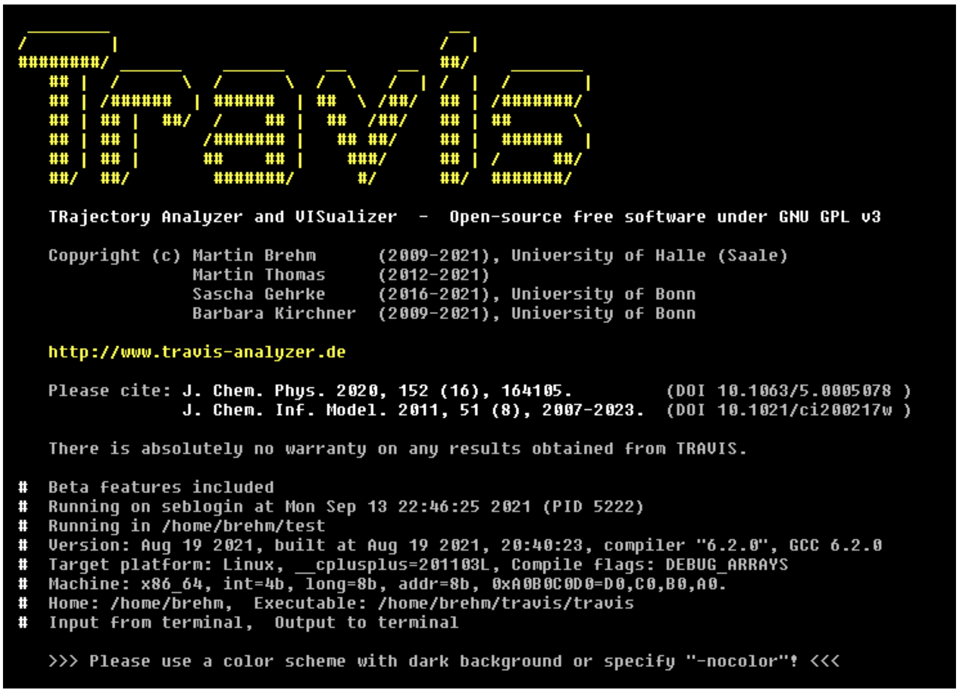


Figure 2.13.1.: Logo of the TRAVIS program package.

TRAVIS is open-source free software, written in C++ *(currently around 300 000 lines of source code)*, platform independent, and does not depend on any external libraries. It possesses an interactive command line interface *(see Figure 2.13.2)* which guides the user through the analysis so that no input file needs to be prepared.



```
 ________                                 __
/        |                               /  |
########/ ______    ______    __     __  ##/    _______
   ## |  /      \  /      \  /  \   /  | /  |  /       |
   ## | /######  | ######  | ##  \ /##/  ## | /#######/
   ## | ## |  ##/  /    ## |  ##  /##/   ## | ##      \
   ## | ## |      /####### |   ## ##/    ## |  ######  |
   ## | ## |      ##    ## |    ###/     ## | /     ##/
   ##/  ##/        #######/      #/      ##/  #######/

   TRajectory Analyzer and VISualizer  -  Open-source free software under GNU GPL v3

   Copyright (c) Martin Brehm      (2009-2021), University of Halle (Saale)
                 Martin Thomas     (2012-2021)
                 Sascha Gehrke     (2016-2021), University of Bonn
                 Barbara Kirchner  (2009-2021), University of Bonn

   http://www.travis-analyzer.de

   Please cite: J. Chem. Phys. 2020, 152 (16), 164105.         (DOI 10.1063/5.0005078 )
                J. Chem. Inf. Model. 2011, 51 (8), 2007-2023.  (DOI 10.1021/ci200217w )

   There is absolutely no warranty on any results obtained from TRAVIS.

#  Beta features included
#  Running on seblogin at Mon Sep 13 22:46:25 2021 (PID 5222)
#  Running in /home/brehm/test
#  Version: Aug 19 2021, built at Aug 19 2021, 20:40:23, compiler "6.2.0", GCC 6.2.0
#  Target platform: Linux, __cplusplus=201103L, Compile flags: DEBUG_ARRAYS
#  Machine: x86_64, int=4b, long=8b, addr=8b, 0xA0B0C0D0=D0,C0,B0,A0.
#  Home: /home/brehm,  Executable: /home/brehm/travis/travis
#  Input from terminal,  Output to terminal

   >>> Please use a color scheme with dark background or specify "-nocolor"! <<<
```

Figure 2.13.2.: The interactive command line interface of TRAVIS.

All methods which are described in Sections 2.1 to 2.12 of this chapter are implemented in Travis, so that they are freely available to the scientific community. Apart from spectroscopic functions, Travis contains many more functions for analyzing trajectories. After performing a simulation of a complex system, a clear and descriptive visualization of the results is equally important, and Travis offers a set of tools to do so. One example are so-called Sankey diagrams which can be used to illustrate the competition between different hydrogen bond donors and acceptors in a mixture—see Figure 2.13.3 for the example of cellulose dissolved in an [EMIm][OAc] / water mixture.[MB40]

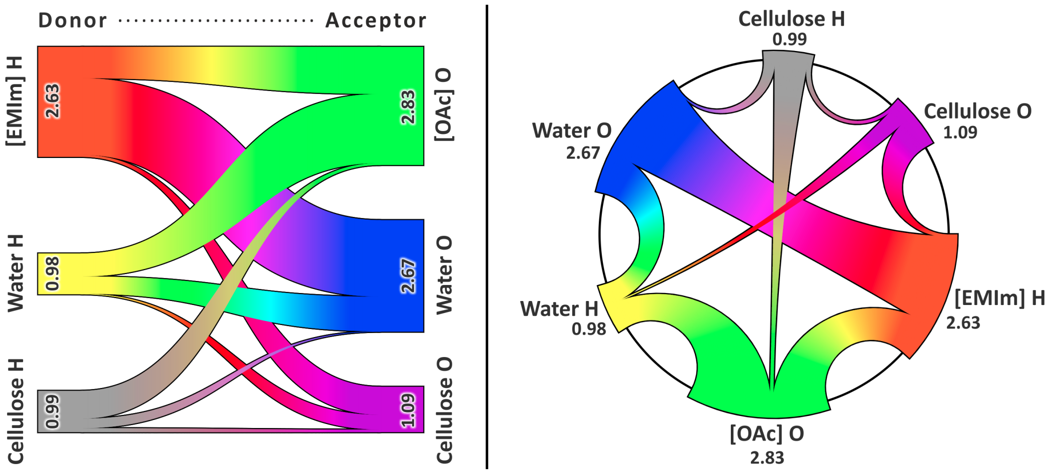


Figure 2.13.3.: An example of a Sankey diagram created with Travis, illustrating the competition between different hydrogen bonds in a complex liquid mixture.[MB40]

The Travis source code distribution as well as some pre-compiled executables for Windows can be downloaded from

`http://www.travis-analyzer.de`

For a step-by-step tutorial on how to predict bulk phase vibrational spectra with CP2k and Travis, please visit

`https://brehm-research.de/spectroscopy`

Chapter 3.

Applications

In this chapter, some vibrational spectra which have been predicted using the approaches from Chapter 2 are presented. After discussing infrared and Raman spectra in Sections 3.1 and 3.2, we turn to the chiral variants of these two—namely VCD and ROA spectroscopy—in Sections 3.3 and 3.4, respectively. Subsequently, a few predicted resonance Raman spectra are shown in Section 3.5. Finally, an example for the bulk phase normal mode analysis is presented in Section 3.6. In Section 3.7, the general computational protocol *(including all relevant literature references)* is outlined.

All spectra in this chapter have been computed with the TRAVIS program package,[MB2,MB40] and the underlying AIMD simulations were carried out with the CP2k software[212–214] in all cases. The full workflow from the idea to the final spectrum involves only open-source free software. For a step-by-step tutorial on how to predict bulk phase vibrational spectra with CP2k and TRAVIS, please visit

`https://brehm-research.de/spectroscopy`

It should be emphasized that all predicted spectra presented below are shown directly as computed. No empirical corrections *(scaling, shifting, line broadening, etc.)* have been applied. As the protocol for computing spectra is based on *ab initio* molecular dynamics simulations, the method is truly predictive and can be applied to predicting spectra of substances never measured *(or even synthesized)* before.

3.1. Infrared Spectroscopy

Methanol

The calculation of infrared spectra from AIMD simulations via time-correlation of molecular dipole moments is well established in the literature. Just one example will be shown here. In Figure 3.1.1, the infrared spectra of methanol in the gas phase *(top panel)* and liquid phase *(bottom panel)* are presented.[MB14] The black curves depict the experimental spectra, while the blue curves are the AIMD-based predictions. Note that the "gas phase" experimental spectrum was measured in CCl_4 to eliminate hydrogen bonding *(similar to gas phase)*, because a true gas phase spectrum was not available.

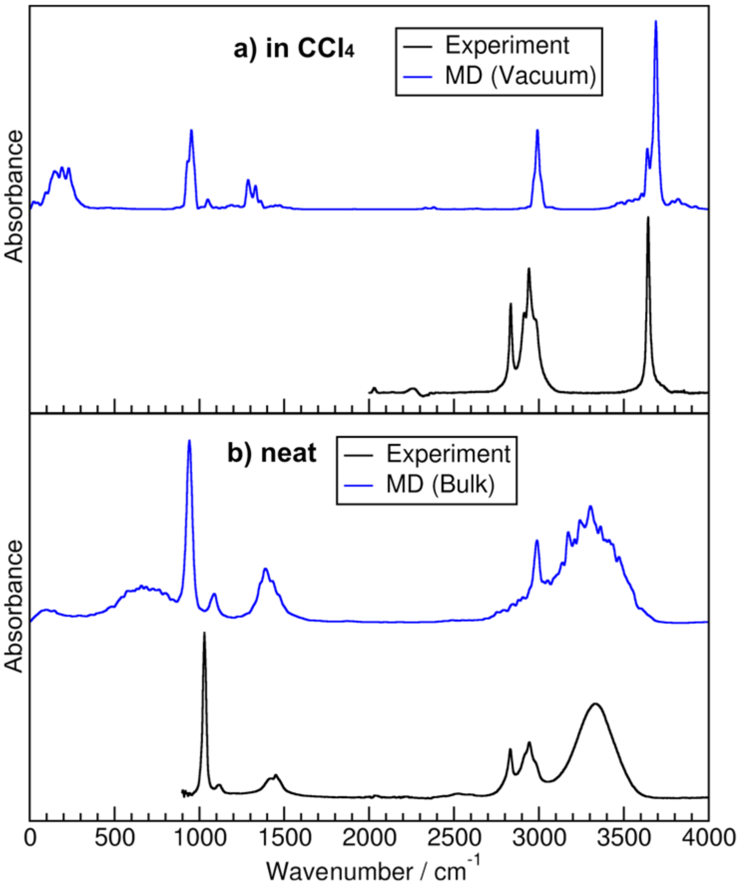


Figure 3.1.1.: AIMD-based predictions *(blue curves)* and experimental *(black curves)* infrared spectra of methanol in gas phase *(upper panel)* and liquid phase *(lower panel)*.[MB14] The "gas phase" experimental spectrum was measured in CCl_4 to eliminate hydrogen bonding.

As it can be seen, both the gas phase and the liquid phase infrared spectra are reproduced well by the AIMD predictions. Most importantly, the drastic differences between both experimental spectra are quantitatively captured by the predicted spectra. Due to the strong hydrogen bonding in the liquid phase of methanol, the

O–H stretching mode, found at around 3700 cm^{-1} in the gas phase, is shifted to lower wavenumbers and significantly broadened due to the strong anharmonicity. Another interesting effect *(which is unfortunately not covered in the experimental spectrum)* is the appearance of a new band at around 650 cm^{-1} which corresponds to intermolecular interactions between methanol molecules.[MB14] On the other hand, some other spectral bands which are not directly influenced by hydrogen bonding retain their narrow shape when going from the gas phase to the liquid phase, as it is also well reproduced by the AIMD simulations.

The computational protocol was as follows. For the liquid phase AIMD simulation, 16 methanol molecules were placed inside a cubic box *(which is quite a low number, but seemingly sufficient to reproduce the bulk phase behavior with respect to vibrational spectra)*. The simulation temperature was set to 400 K. More computational details can be found in Section 3.7.

3.2. Raman Spectroscopy

Liquid Propylene Oxide

One system to which we applied our approach to computing condensed phase Raman spectra is liquid *(R)*-propylene oxide, also known as methyl oxirane, which possesses a molecular structure as shown in Figure 3.2.1.

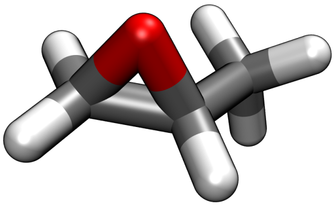


Figure 3.2.1.: Molecular structure of *(R)*-propylene oxide.

Based on AIMD simulations, we have predicted the depolarized Raman spectrum of liquid *(R)*-propylene oxide, and compare it to an experimental spectrum in Figure 3.2.2.[MB26] In order to emphasize the important fingerprint region, we show only a subset of the full frequency range here. It can be seen that the experiment is reproduced by the simulation very well. All features of the experimental spectrum can be found in the predicted spectrum, and also the intensity ratios match very well for most bands. Most peaks appear at lower frequencies in the computed spectrum. This effect is well-known and was observed in comparable Raman studies on propylene oxide before.[139] It is therefore not related to our method for computing spectra presented herein, and could be cured by using a higher level of theory for

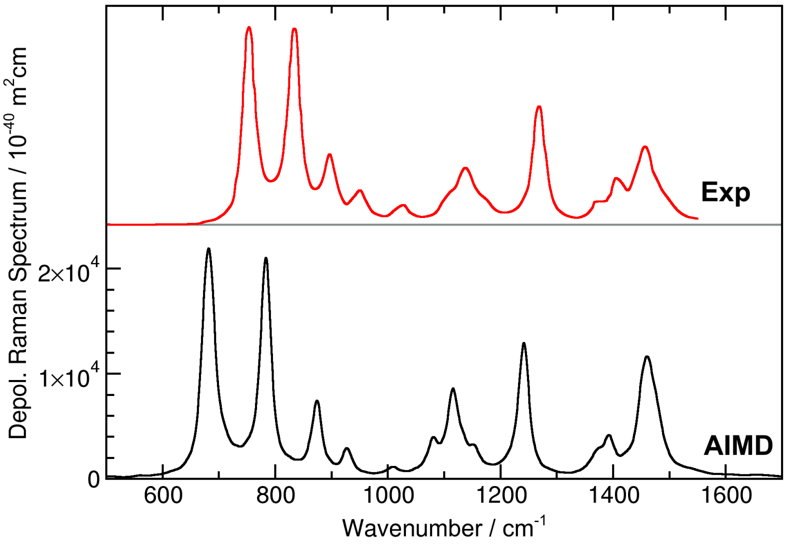


Figure 3.2.2.: Depolarized Raman spectrum of liquid propylene oxide predicted from AIMD *(black curve)*[MB26] compared to experimental spectrum *(red curve)*.[396]

performing the electron structure calculations in the AIMD simulation. As already mentioned, our approach can, in principle, be coupled to any electron structure method which is able to provide the total electron density.[MB26]

Apart from the Raman spectrum itself, another very interesting property which can be predicted based on the methods shown in Section 3.2 is the Raman depolarization ratio. As discussed there, the depolarization ratio $\rho(\tilde{\nu})$ is defined as the quotient of the orthogonal Raman spectrum $I_\perp(\tilde{\nu})$ and the parallel Raman spectrum $I_\parallel(\tilde{\nu})$. For normal modes with low symmetry, the value of ρ is always 0.75 (so-called *"depolarized band"*), but for totally symmetric modes, it can be lower (so-called *"polarized band"*). For example, the breathing mode in methane possesses a depolarization ratio of $\rho = 0$. Based on this rationale, the depolarization ratio can help to assign spectral bands to vibrational modes in a molecule.

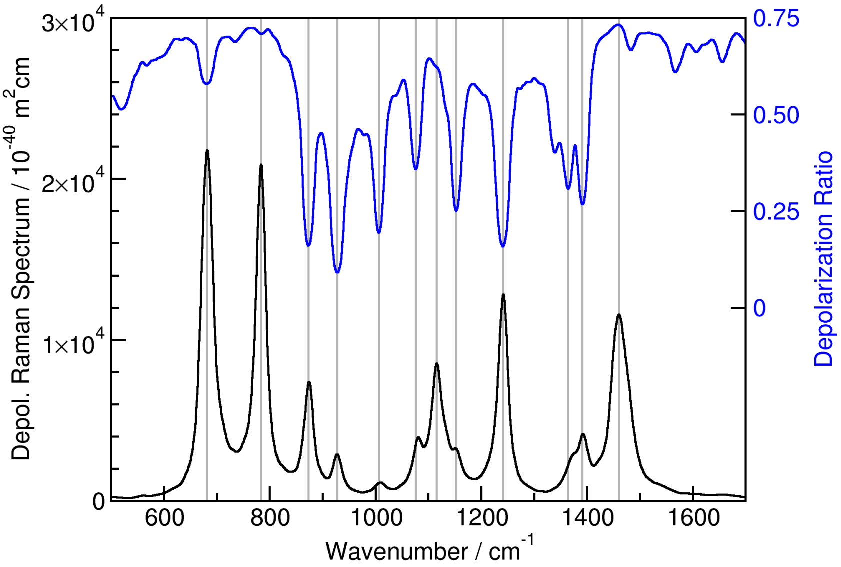


Figure 3.2.3.: Depolarized Raman spectrum of liquid propylene oxide predicted from AIMD *(black curve)*[MB26] together with predicted depolarization ratio *(blue curve)*.

The depolarization ratio of liquid propylene oxide is presented as blue curve in Figure 3.2.3, together with the depolarized Raman spectrum *(black curve)* that was already shown on the last page. The grey vertical bars help to assign the depolarization ratio to each spectral band. As it can be seen, some intense bands *(e. g., 790 cm^{-1} and 1460 cm^{-1})* possess a depolarization ratio close to 0.75 *("depolarized bands")* and are therefore resulting from asymmetric vibrational modes, while some other bands *(e. g., 930 cm^{-1} and 1240 cm^{-1})* show a depolarization value close to zero *("polarized bands")* and therefore correspond to highly symmetric modes.

The AIMD simulation contained 32 molecules of (*R*)-propylene oxide in a cubic simulation cell with 1558 pm edge length, resulting in a density of 0.816 $\mathrm{g\,cm^{-3}}$, and was conducted with the CP2k program package[212–214] in the NVT ensemble at a temperature of 300 K. A snapshot of the simulation cell is shown in Figure 3.2.4. After 20 ps of equilibration with and without massive thermostats, a production run of 32.5 ps followed. In order to obtain the polarizability via finite differences, additional simulations were performed with static external electric fields in X/Y/Z direction. The field strength was set to $5.0 \cdot 10^{-3}$ a.u. $= 2.57 \cdot 10^{9}\,\mathrm{V\,m^{-1}}$. More computational details can be found in Section 3.7.

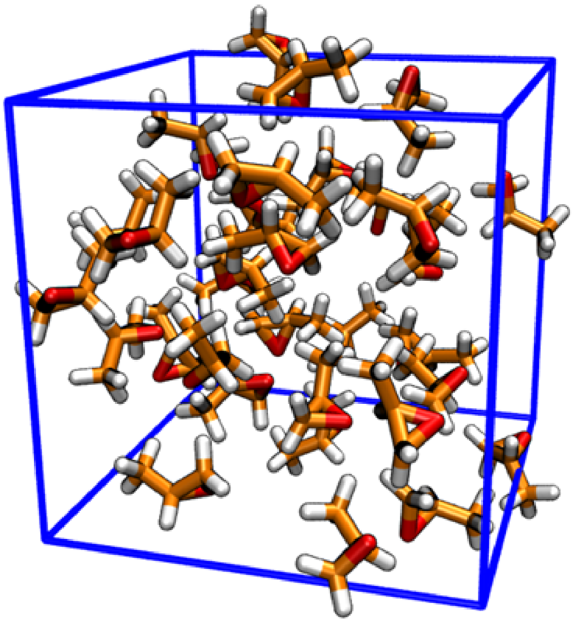


Figure 3.2.4.: Snapshot of the liquid phase propylene oxide simulation cell.

$MgCl_2$ / KCl Melt

Within the scope of a collaborative project with the group of Vyacheslav Bryantsev,[MB49] we predicted unpolarized Raman spectra of inorganic melts consisting of $MgCl_2$ and KCl based on AIMD simulations. The results are presented in Figure 3.2.5 with the experimental spectra in the top panel and the predictions with the PBE–D3, PBE–D2, and revPBE–D2 functionals in the remaining panels, respectively. Black curves correspond to pure $MgCl_2$, while red curves depict the spectra of 50/50 mol% mixtures of $MgCl_2$ and KCl.

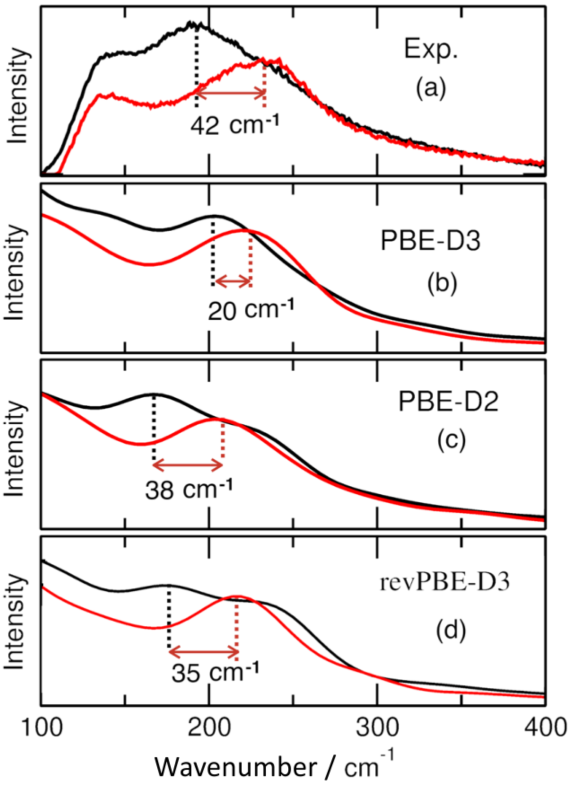


Figure 3.2.5.: Experimental *(top panel)* and predicted *(other panels)* unpolarized Raman spectra for pure $MgCl_2$ *(black curves)* and a 50/50 mol% mixture of $MgCl_2$ and KCl *(red curves)* at 1073 K.[MB49]

It can be seen that the prediction based on AIMD simulations with PBE–D3 reproduces the frequencies and intensity ratio of the experiment very well, while the other approaches are slightly worse in one or more aspects. Given such a level of agreement between the experimental and simulated Raman spectra, we are confident that the underlying structural information at the atomic resolution is faithfully represented in the AIMD simulation with PBE–D3.[MB49]

As the Voronoi integration approach is currently only suitable for the calculation of molecular electromagnetic moments, the total cell dipole moment from the Berry phase operator was used here instead for the computation of the spectra.

The computational protocol was as follows.[MB49] 149 Mg^{2+} and 298 Cl^{-} ions *(pure system)* or 100 Mg^{2+}, 100 K^{+}, and 300 Cl^{-} ions *(mixture)* were pre-equilibrated in NpT ensemble for 1 ns at 1073 K and 1 bar via force field molecular dynamics, using a polarizable ion model (PIM)[397–399] to accurately model the interactions. Structures with the average equilibrated densities ($1.612\,\mathrm{g\,cm^{-3}}$ for pure $MgCl_2$ and $1.573\,\mathrm{g\,cm^{-3}}$ for the mixture) were extracted as initial configurations for the AIMD simulations. Note that these densities are in very good agreement with the experimental densities ($1.65\,\mathrm{g\,cm^{-3}}$ for pure $MgCl_2$ and $1.56\,\mathrm{g\,cm^{-3}}$ for the mixture).[400,401] A snapshot of the pure $MgCl_2$ simulation cell is shown in Figure 3.2.6.

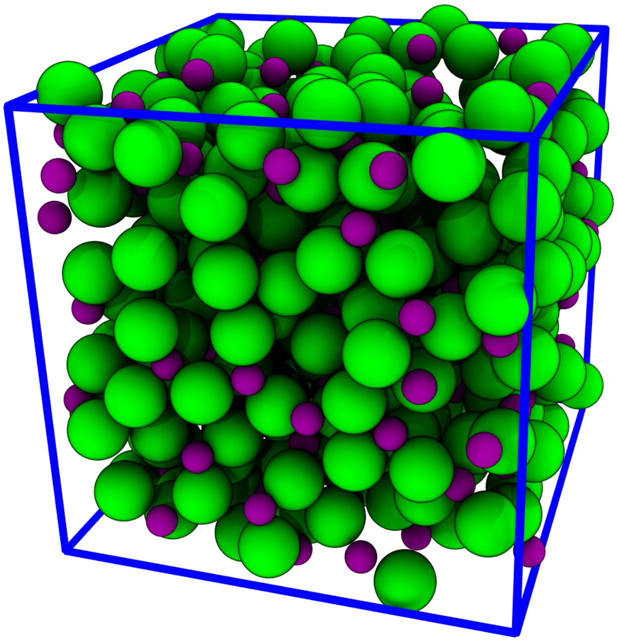


Figure 3.2.6.: Snapshot of the simulation cell of the pure $MgCl_2$ simulation *(green: chloride, purple: magnesium)*.

The AIMD simulations have been performed with the CP2k program package,[212–214] employing the Quickstep (QS) method[402] and the orbital transformation (OT)[403,404] approach with the FULL_ALL preconditioner[405] and a conjugate gradient (CG) minimizer.[406–408] The electronic structure was calculated with the PBE–D3, PBE–D2, and revPBE–D2 methods, which are based on the PBE[409] and revPBE[410] functionals and employ either the D2[411] or the D3[412,413] empirical dispersion correction from Grimme. Basis sets of the kind MOLOPT–DZVP[414] and GTH pseudopotentials[415,416] were applied to all atoms. The plane wave cutoff was set to 600 Ry, and the time step was chosen to be 1.0 fs in all simulations. The simulation temperature was kept at 1073 K by a Nosé–Hoover chain thermostat[184–186] with a thermostat time constant $\tau = 1000\,\mathrm{fs}$ (*i. e.*, NVT ensemble). After an equilibration interval of 20 ps, a production run was performed for 60 ps.

3.3. VCD Spectroscopy

Liquid Propylene Oxide

A worthwhile target for the prediction of VCD spectra is propylene oxide. It is one of the smallest chiral organic molecules that exist, and therefore it is relatively easy to perform an AIMD simulation of the liquid phase with reasonable sampling. VCD spectra of propylene oxide have been predicted before,[417–422] but only for single isolated molecules *(mostly based on the static–harmonic approximation)* and therefore not including any explicit solvent effects which are present in the liquid phase. We have predicted[†] the VCD spectrum of liquid *(R)*-propylene oxide with the AIMD-based approach described in Section 2.7; the resulting spectrum is presented in Figure 3.3.1 as black curve, together with an experimentally measured spectrum *(red curve)*.[417] In order to emphasize the important fingerprint region, we show only a subset of the full frequency range here.

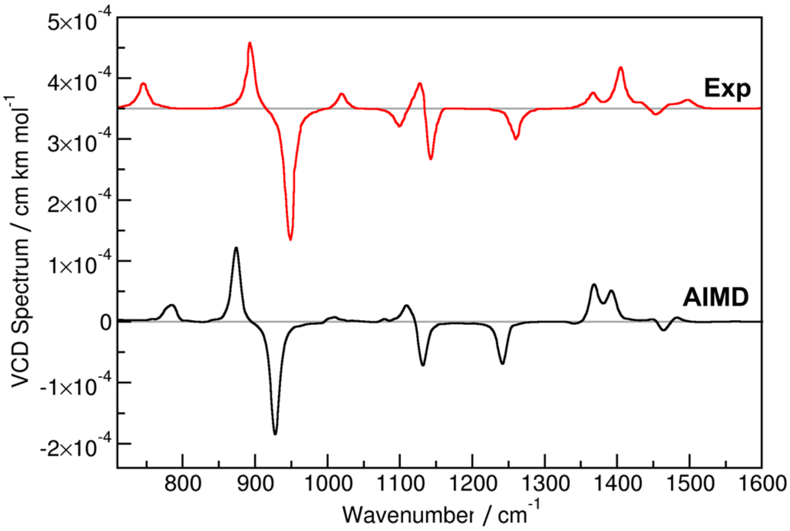


Figure 3.3.1.: Predicted VCD spectrum of liquid *(R)*-propylene oxide *(black curve)*[†] together with experimental spectrum.[417]

It can be seen that our prediction reproduces the experimental VCD spectrum very well. All signals observed in the experiment are also found in the simulated spectrum, and also the sign and intensity ratio of the signals is in good agreement. We would like to emphasize again that the predicted spectrum is directly presented as computed—no empirical corrections *(scaling, shift, line broadening, etc.)* were applied. Thus, the method is truly predictive, and can help to identify absolute configurations.

[†] Unpublished result.

As already discussed in Section 3.2 for the same system, some peaks appear at slightly different frequencies in the computed spectrum, which is a well-known effect and was observed in comparable studies on propylene oxide before.[139] It is therefore not related to our method for computing spectra presented herein, and could be cured by using a higher level of theory for performing the electron structure calculations in the AIMD simulation.

The computations were performed as follows. During the AIMD simulation, the total electron density was written to disk as volumetric data in Gaussian Cube format in each time step. A setting of STRIDE=1 was used to write these files, which resulted in a grid of $160 \times 160 \times 160$ data points (*i. e.*, 9.7 pm grid spacing). The electric current was obtained from the purely classical approach[150] introduced in Section 2.6 by solving the partial differential equation described there, using a relative convergence threshold for the BiCGStab method of 10^{-3} and a background density of $10^{-6}\,\text{a.u.} = 7 \cdot 10^{-3}\,\text{e}\,\text{nm}^{-3}$. In the radical Voronoi tessellation scheme, van der Waals atom radii[240–242] were employed, as these have been shown to reproduce spectroscopic properties very well.[MB23] More computational details can be found in Section 3.7.

Details on the system setup can be found in Section 3.2, where also a snapshot of the simulation cell is presented in Figure 3.2.4 on page 88.

3.4. ROA Spectroscopy

Liquid Propylene Oxide

As already discussed in Section 3.3, a worthwhile target for the prediction of ROA spectra is propylene oxide, which is one of the smallest chiral organic molecules that exist. Not surprisingly, ROA spectra of propylene oxide have been predicted before,[50,152,423,424] but only for single isolated molecules *(mostly based on the static–harmonic approximation)* and therefore not including any explicit solvent effects which are present in the liquid phase. In 2017, we have predicted[MB26] the ROA spectrum of liquid *(R)*-propylene oxide with the novel AIMD-based approach described in Section 2.8; the resulting spectrum is presented in Figure 3.4.1 as black curve, together with an experimentally measured spectrum *(red curve)*.[425] This was the first *ab initio* prediction of a liquid phase ROA spectrum in the literature. In order to emphasize the important fingerprint region, we show only a subset of the full frequency range here.

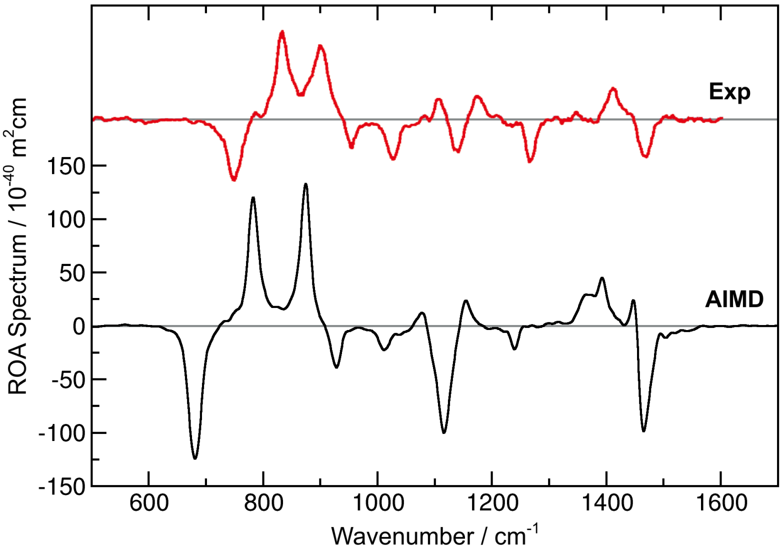


Figure 3.4.1.: Predicted ROA spectrum of liquid *(R)*-propylene oxide[MB26] together with experimental spectrum.[425]

As it can be seen, the simulated ROA bands are in very good agreement with the experimentally observed pattern of positive and negative peaks. For each signal in the experiment, the simulation shows a corresponding counterpart, and also the relative intensity ratios agree very well. The calculated spectrum would easily allow to identify the correct enantiomer of propylene oxide, since the other enantiomer possesses all bands with opposite sign. We would like to emphasize again that the predicted spectrum is directly presented as computed—no empirical corrections *(scaling, shift, line broadening, etc.)* were applied. Thus, the method is truly predictive.

As already discussed in Section 3.2 for the same system, some peaks appear at slightly different frequencies in the computed spectrum, which is a well-known effect and was observed in comparable studies on propylene oxide before.[139] It is therefore not related to our method for computing spectra presented herein, and could be cured by using a higher level of theory for performing the electron structure calculations in the AIMD simulation.

The computations were performed as follows. During the AIMD simulation, the total electron density was written to disk as volumetric data in Gaussian Cube format in each time step. A setting of STRIDE=1 was used to write these files, which resulted in a grid of $160 \times 160 \times 160$ data points (*i. e.*, 9.7 pm grid spacing). The electric current was obtained from the purely classical approach[150] introduced in Section 2.6 by solving the partial differential equation described there, using a relative convergence threshold for the BiCGStab method of 10^{-3} and a background density of $10^{-6}\,\mathrm{a.u.} = 7 \cdot 10^{-3}\,\mathrm{e\,nm^{-3}}$. In the radical Voronoi tessellation scheme, van der Waals atom radii[240–242] were employed, as these have been shown to reproduce spectroscopic properties very well.[MB23]

In order to obtain the polarizabilities required for ROA, additional simulations were performed with static external electric fields in X/Y/Z direction. The field strength was set to $5.0 \cdot 10^{-3}\,\mathrm{a.u.} = 2.57 \cdot 10^{9}\,\mathrm{V\,m^{-1}}$. The external field strength used here was increased by an order of magnitude with respect to our previous work.[MB14,MB19,MB23] This was done to reduce the noise level in the computation of the electric dipole–magnetic dipole polarizabilities via finite differences. We made sure that the polarization effects are still within the linear regime despite the increased field strength. More computational details can be found in Section 3.7.

Compared to other approaches, the total computational time used in this project is moderate; it accounts to 2350 ch (core hours; Intel Xeon "Haswell" @ 2.4 GHz) for the equilibration, 25 500 ch for the production run of 4 trajectories with/without external field, and 2 000 ch to solve the partial differential equation for the electric current in all time steps (see supporting information for computational details). This means that the spectrum presented in Figure 3.4.1 can be obtained within 3 weeks on a 64 core small desktop server, which is available for less than 10 000 USD today.

Details on the system setup can be found in Section 3.2, where also a snapshot of the simulation cell is presented in Figure 3.2.4 on page 88.

N-Acetyl-L-cysteine

Within the scope of an ongoing collaboration with the group of Yunjie Xu,[MB51] the ROA spectrum of *N*-acetyl-L-cysteine (NALC) *(see Figure 3.4.2)* in water was predicted from AIMD simulations using the approach described in Section 2.8. As NALC is a very flexible molecule with several different low-energy conformers, and as it possesses several strong hydrogen bond donor and acceptor sites which tightly integrate into the hydrogen bond network of the solvent, it is challenging to obtain a good prediction for the ROA spectrum via the static–harmonic approach without explicit solvent. Therefore, we considered this molecule a worthwhile target for the AIMD-based approach to predicting ROA spectra.

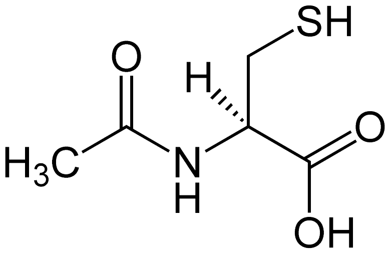


Figure 3.4.2.: Molecular structure of *N*-acetyl-L-cysteine (NALC).

The resulting predicted ROA backscattering spectrum of NALC in water is presented in Figure 3.4.3 *(black curve)* together with the experimentally measured spectrum *(red curve)*. Apart from the strong positive peak at 1325 cm^{-1} which is not captured in the prediction, and apart from several slightly shifted bands, the overall agreement between experiment and prediction is reasonable, which is a remarkable achievement for such a complex molecule.

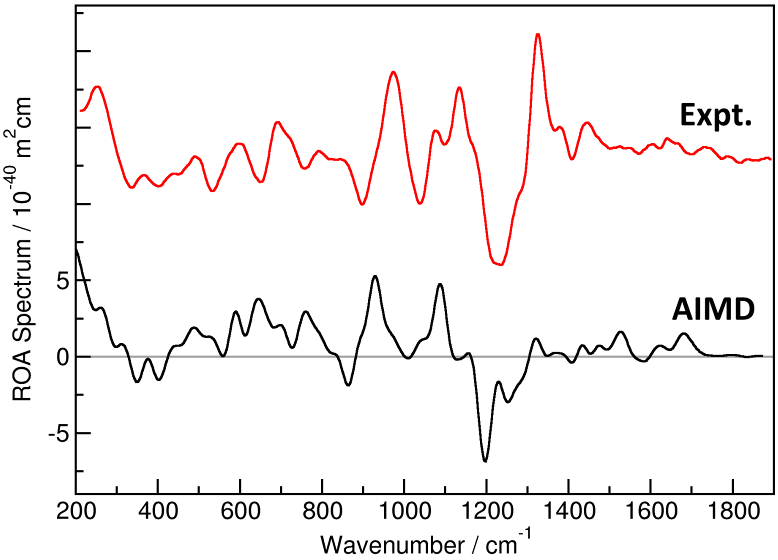


Figure 3.4.3.: Predicted ROA backscattering spectrum of *N*-acetyl-L-cysteine (NALC) in water *(black curve)* together with experimental spectrum *(red curve)*.[MB51]

To compute the spectrum, one molecule of NALC was placed in a cubic box of $\approx 1300\,\mathrm{pm}$ edge length together with 64 water molecules. A snapshot of the simulation cell is shown in Figure 3.4.4. After force field pre-equilibration using the TIP4P–EW force field[426] for water and OPLS–AA[427] combined with optimized RESP charges[428,429] for NALC, the last frame was used to set up AIMD simulations in CP2k[212–214] with a total length of 40 ps, of which the last 25 ps were taken as production runs to compute the spectra with TRAVIS.[MB2,MB40] The simulation temperature was set to 350 K. To obtain the required polarizabilities via finite differences, an external field strength of $E = 5.0 \cdot 10^{-3}\,\mathrm{a.u.} = 2.57 \cdot 10^{9}\,\mathrm{V\,m^{-1}}$ was applied. A general description of the computational protocol can be found in Section 3.7.

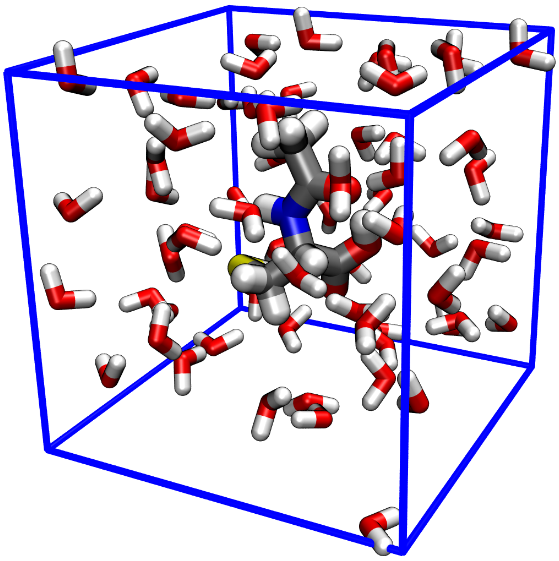


Figure 3.4.4.: Snapshot of the AIMD simulation cell containing one *N*-acetyl-L-cysteine (NALC) molecule in 64 molecules of water *(1300 pm cubic cell size)*.

As described above, one difficulty with NALC is the flexible character of the molecule, which is responsible for the existence of several different low-energy conformers which all contribute to the spectrum. Six such relevant conformers are depicted in Figure 3.4.5. One advantage of the AIMD-based approach to computing spectra when compared to the static–harmonic approach is the intrinsic conformer sampling due to the dynamics; however, the NALC conformers in solution are relatively stable, so that an AIMD simulation is significantly too short to sample all relevant conformers in one run.

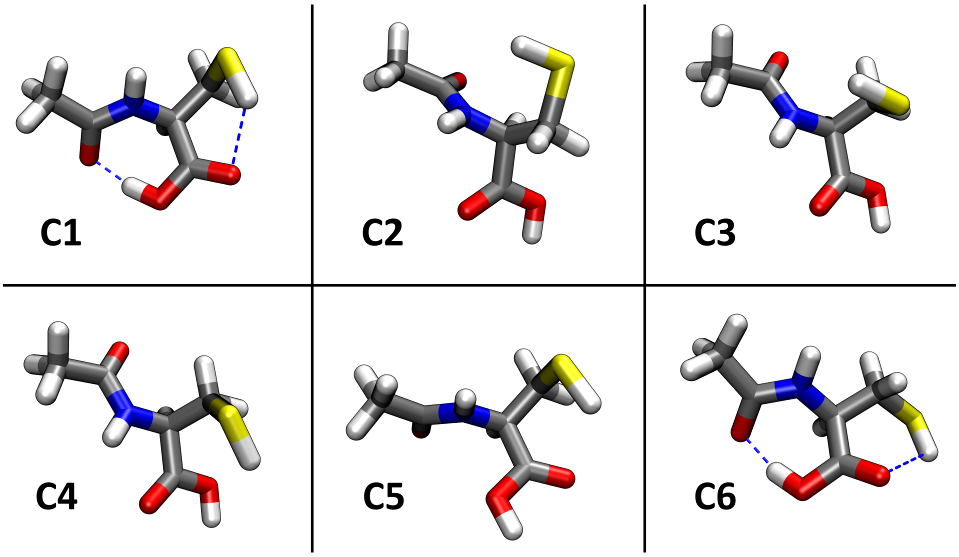


Figure 3.4.5.: Six relvant low-energy conformers of *N*-acetyl-L-cysteine (NALC) for which independent AIMD calculations were performed.

To overcome this, we decided to perform six independent AIMD simulations for the six conformers shown in Figure 3.4.5 in water, and to predict vibrational spectra for each conformer separately. To avoid interconversion of the conformers during each AIMD simulation, we decided to use harmonic restraints on relevant dihedral angles in order to keep the backbone geometry intact. First, we identified the relevant dihedral angles which distinguish the conformers, which we refer to as $\phi_1 \dots \phi_6$, and which are illustrated in Figure 3.4.6. The values of these dihedral angles for the six conformers C1. . .C6 are given in Table 3.4.1.

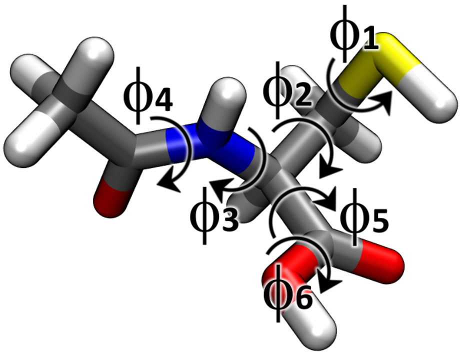


Figure 3.4.6.: Dihedral angles in *N*-acetyl-L-cysteine (NALC) which were restrained to keep the conformers from Figure 3.4.5 intact.[MB51]

Table 3.4.1.: Target values of the restraints on the dihedral angles $\phi_1 \ldots \phi_6$ in *N*-acetyl-L-cysteine (NALC) *(see Figure 3.4.6)* to keep the six conformers C1…C6 intact *(see Figure 3.4.5)*.[MB51]

Dihedral	C1	C2	C3	C4	C5	C6
ϕ_1	64.7	65.2	−64.9	78.6	73.3	−71.0
ϕ_2	52.9	−64.6	65.4	62.6	59.1	−166.6
ϕ_3	159.6	128.7	80.9	80.7	127.3	159.0
ϕ_4	−172.9	179.2	178.3	179.9	−179.4	−171.8
ϕ_5	−178.0	−68.9	−58.3	−58.6	150.0	−174.4
ϕ_6	−1.7	177.9	179.4	178.4	177.0	−5.7

In the AIMD simulations for the six conformers, harmonic restraints with a spring constant of $K = 0.005\,\text{Hartree}\,\text{rad}^{-2}$ were placed on the six dihedral angles $\phi_1 \ldots \phi_6$, using the values in Table 3.4.1 as equilibrium values. In contrast to constraints, restraints are not rigid—they allow a certain motion of the restrained quantity around the equilibrium value, but avoid large elongations. This keeps each of the six conformers in place without altering the backbone dynamics too much, so that the spectra are not significantly influenced. The concept of restraining the backbone dihedral angles is visualized in Figure 3.4.7, where an overlay of the conformer C1 over the whole AIMD production run is shown. It can be seen that despite the restraints, the backbone is not completely rigid and retains a certain degree of freedom, but the overall shape of the conformer is maintained.

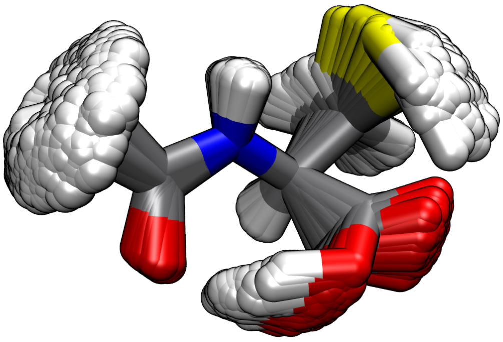


Figure 3.4.7.: Overlay of *N*-acetyl-L-cysteine (NALC) structures along the trajectory for conformer C1, which is kept intact by dihedral angle restraints *(see Figure 3.4.6)*.[MB51]

To see how significant the influence of the conformer geometry on the ROA spectrum is, please consider Figure 3.4.8, where the individual predicted ROA spectra for the six NALC conformers are shown. While some bands are predicted similarly for some of the different conformers, the differences are very large, and even the sign of some peaks differs between the conformers. The experimental ROA spectrum is a superposition of these *(and probably some more)* conformer spectra, with weights determined by the free enthalpy ΔG of each conformer in solution via the corresponding Boltzmann factor $\exp\left(\frac{-\Delta G}{k_\mathrm{B}T}\right)$. In principle, the free enthalpy values can be determined by AIMD simulations; however, this would require significantly longer simulation times for the sampling to converge. Therefore, the final predicted ROA spectrum shown in Figure 3.4.3 was created as a linear combination of the six conformer spectra with empirically determined coefficients.

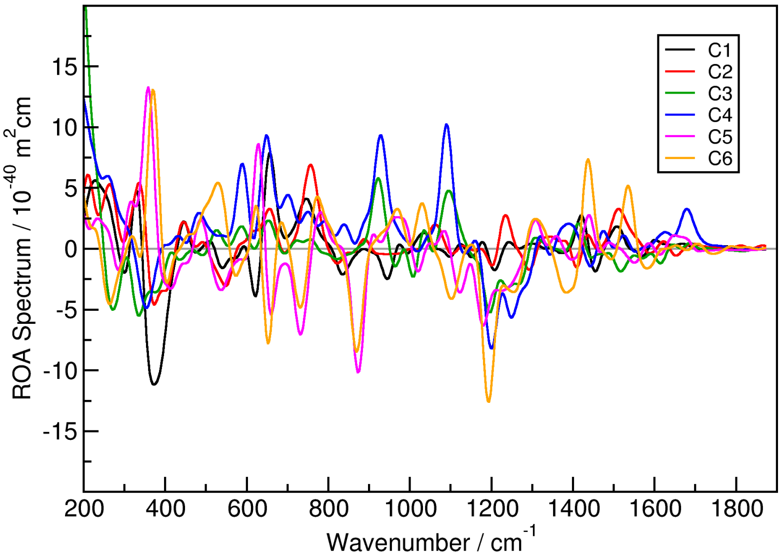


Figure 3.4.8.: Predicted ROA backscattering spectra for the six different *N*-acetyl-L-cysteine (NALC) conformers *(see Figure 3.4.5)* in water.

In a certain sense, one could argue that fixing a molecule's backbone geometry in an AIMD simulation by restraining certain dihedral angles is not a sensible way to go, as one loses one of the advantages of the AIMD approach—namely, the intrinsic conformer sampling. However, this conformer sampling is only effective if the energy barrier between all relevant conformers is sufficiently small. If this is not the case *(such as for NALC)*, the AIMD simulation time will be too short to obtain a meaningful average over the relevant conformers, and the spectrum will heavily depend on the choice of the initial conformer, which should not be the case in a MD simulation with reasonable sampling. In such cases, it can be advantageous to consciously sacrifice the conformer sampling and run independent simulations for each conformer, which can then be averaged according to the conformers' free enthalpies.

3.5. Resonance Raman Spectroscopy

Uracil in Aqueous Solution

To demonstrate our newly developed approach[MB36] for predicting bulk phase resonance Raman spectra *(see Section 2.10)*, uracil in water was chosen as a model system for several reasons:

- Uracil is an important building block of RNA and ubiquitous in biological systems.
- Uracil possesses both hydrogen bond donors and acceptors and tightly integrates into the hydrogen bond network of water—see Figure 3.5.1. A strong solvent effect on the spectrum can be expected.
- Uracil possesses an extended π electron system and shows electronic excitations in near UV, so that these can be easily excited by a laser.

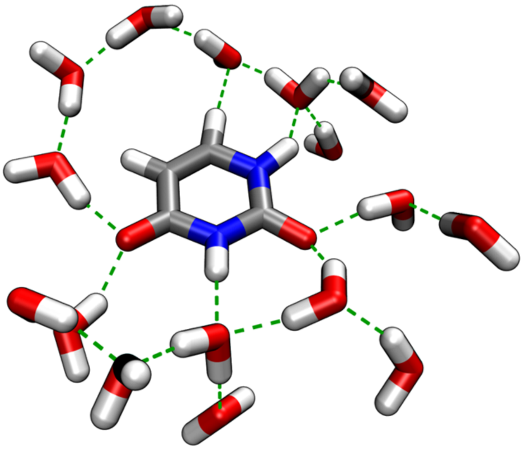


Figure 3.5.1.: Uracil integrates into the hydrogen bond network of liquid water.

To emphasize the importance of the solvent effect in this system, we also predicted the spectra for uracil in gas phase without solvent.

In the top panel of Figure 3.5.2, we compare our non-resonant AIMD-based spectra of uracil in gas phase and aqueous solution to a non-resonant Raman experiment.[430] While the gas phase simulation *(green curve)* deviates strongly from the experimental spectrum *(red curve)*, the bulk phase prediction *(black curve)* shows a nice agreement with experiment. Even if resonance effects do not play a role, it is therefore mandatory to perform bulk phase simulations in order to correctly predict the Raman spectrum of uracil.

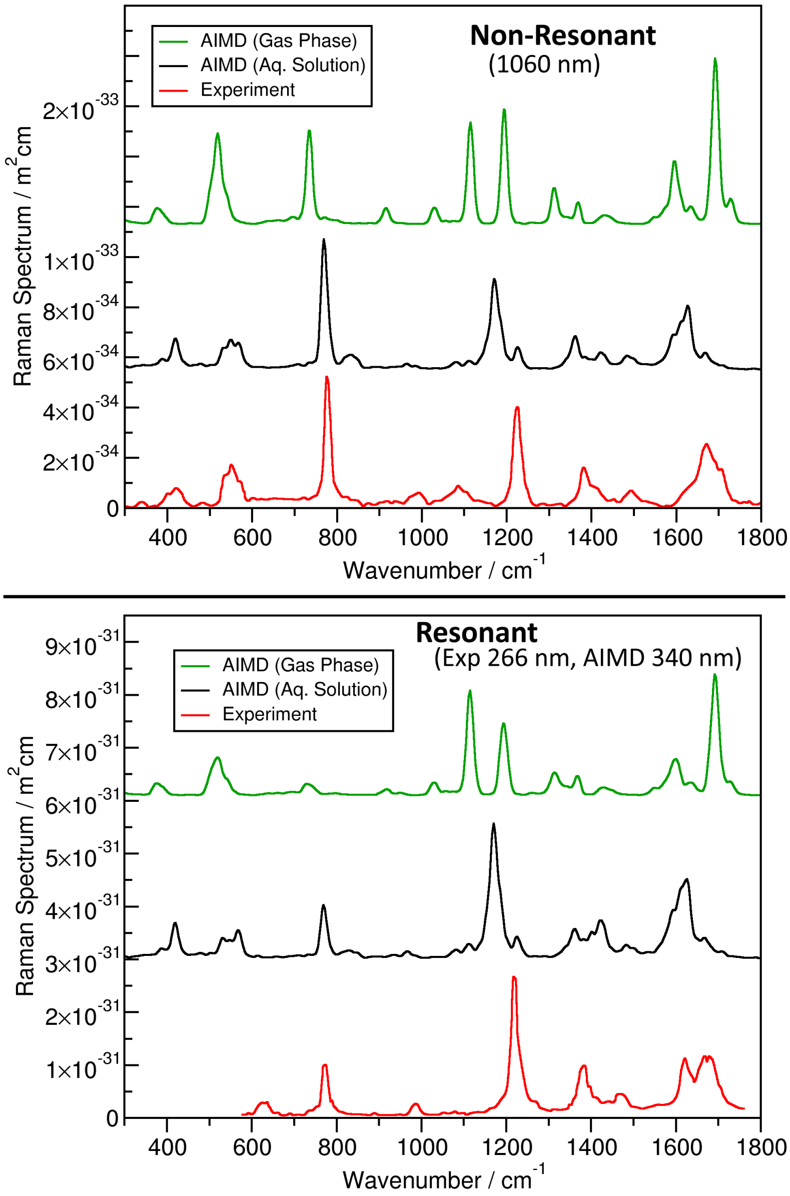


Figure 3.5.2.: Upper panel: Computed non-resonant Raman spectrum of Uracil in gas phase *(green)* and water *(black)* compared to experiment[430] *(red)*; Lower panel: Computed resonance Raman spectrum of Uracil in gas phase *(green)* and water *(black)* at 340 nm compared to experiment[431] at 266 nm *(red)*.[MB36]

As the next step, we discuss the resonance Raman spectrum. In the bottom panel of Figure 3.5.2, we show the AIMD-based spectra of uracil in gas phase and aqueous solution together with an experimental resonance Raman spectrum.[431] While the experiment was conducted with a laser energy of 4.66 eV / 266 nm, we used a value of 3.65 eV / 340 nm in our prediction. This difference can be explained by the fact that GGA functionals often underestimate electronic excitation energies, and is therefore not related to our method. Our approach works also with hybrid functionals, where the agreement between experimental and computed excitation energies is known to be far better.

Similarly to the non-resonant case above, the predicted spectrum of the gas phase system *(green curve in bottom panel of Figure 3.5.2)* shows some significant deviations from experiment *(red curve)*, while the bulk phase simulation *(black curve)* is in very good agreement with the experimental spectrum. This was the first *ab initio* prediction of a liquid phase resonance Raman spectrum in the literature. We conclude that for both non-resonant and resonance Raman spectra, it is definitely required to take the solvent influence explicitly into account in the case of uracil. When comparing the ordinate axes in both panels of Figure 3.5.2, it is visible that the absolute signal intensity is increased by around three orders of magnitude due to the resonance Raman effect, as also observed in the experiment.[431]

As described in Section 2.10, our computational approach does no only yield the resonance Raman spectrum for one laser wavelength at a time—it predicts all resonance Raman spectra for all possible laser wavelengths in one pass. Based on this information, an excitation profile can be created. This is a contour plot with the vibrational frequency on the horizontal axis and the laser wavelength on the vertical axis—see Figure 3.5.3.[MB36,MB40]

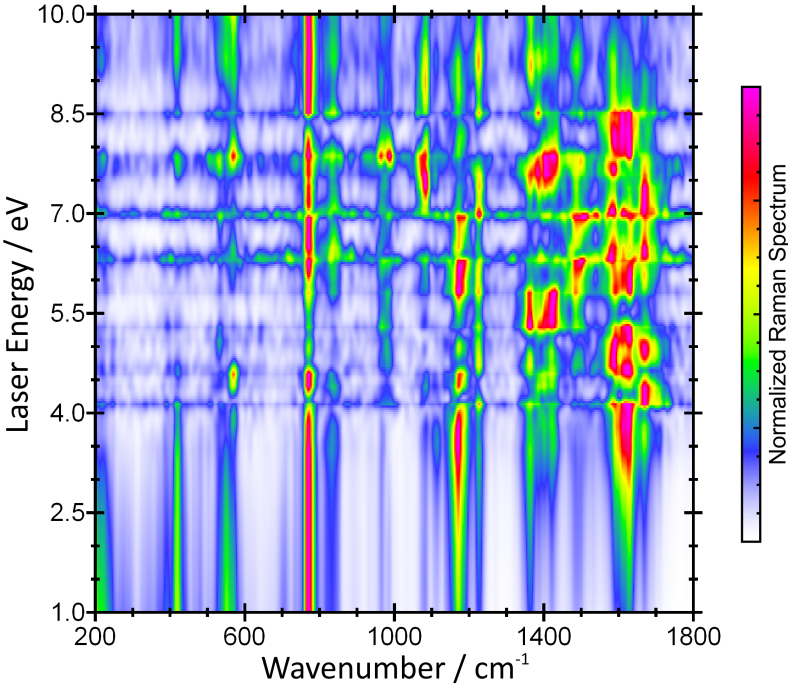


Figure 3.5.3.: Predicted excitation profile of uracil in water, *i. e.*, resonance Raman spectra for all possible laser wavelengths *(vertical axis)*.[MB36,MB40] Rows are normalized to show relative intensity ratio.

Such an excitation profile is not only helpful to understand the coupling between vibrational modes and electronic excitations *(vibronic coupling)* in the system investigated, but can even help to design new interesting experiments by picking laser wavelengths at which interesting resonance effects can be expected. In the lower part of the contour plot, the non-resonant Raman spectrum is reproduced, while the intensity ratio between the spectral bands changes drastically in the resonant upper part of the plot.

We can conclude that—by using our novel approach—it is now possible to directly compute resonance Raman spectra of bulk phase systems. This has not been achieved before to the best of our knowledge, therefore constituting an important contribution to the field of computational vibrational spectroscopy. In contrast to existing methods, our approach includes the full solvent influence and some anharmonic effects. We have computed the resonance Raman spectrum of an aqueous solution of uracil and find that it is in very good agreement with the experiment.

The computational protocol for the prediction of the spectra was as follows. For the bulk phase simulation, one uracil molecule was placed in a cubic box together with 32 water molecules *(cell size* $\approx 1050\,pm$*)*, and a force field pre-equilibration in NpT ensemble was carried out to converge the density. A snapshot of the simulation cell is shown in Figure 3.5.4. For the gas phase simulation, these steps were skipped. Subsequently, BOMD simulations of both systems at a temperature of 300 K were started with CP2k. After another equilibration interval, production runs of 20 ps were performed. More computational details for these steps are described in Section 3.7 below.

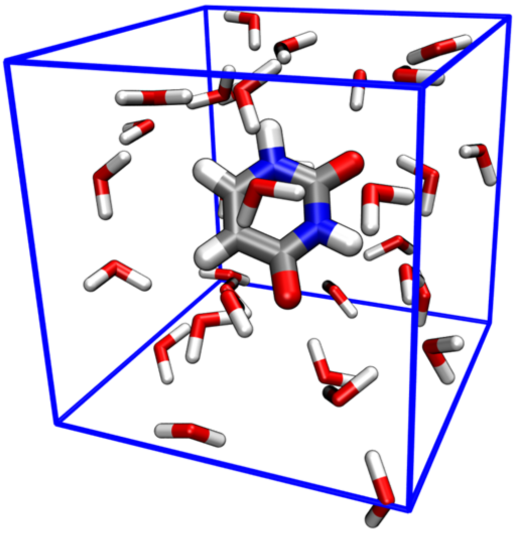


Figure 3.5.4.: Snapshot of the simulation cell for predicting the bulk phase resonance Raman spectrum of uracil in water.[MB36]

From these BOMD production trajectories, snapshots were taken every 2.5 fs (*i. e.*, every 5 steps), so that 8 000 snapshots per system resulted. For each of these snapshots, a real-time propagation run (RTP)[372] was started with CP2k.[212–214] The initial wave function for the propagation was optimized under the influence of an external periodic electric field in X, Y, and Z direction. For each field direction, a separate RTP run was performed. The absolute value of the electric field amounted to $|\mathbf{E}| = 5.0 \cdot 10^{-4}\,\text{a.u.} = 2.57 \cdot 10^{8}\,\text{V}\,\text{m}^{-1}$. Directly in the beginning of the RTP runs, the electric field was switched off (step response). The propagation time step was set to 0.0125 fs, and 1 280 steps were performed (*i. e.*, 16 fs of total physical time). In the RTP runs, we chose EPS_DEFAULT to 10^{-10} and EPS_ITER to 10^{-6}. Every 0.0625 fs (*i. e.*, every 5 propagation steps), the total electron density was written to disk in Gaussian Cube file format, so that 256 frames per BOMD snapshot resulted. The spatial resolution of the volumetric grid was $108 \times 108 \times 108$ for all three systems. The Cube files were compressed to bqb format[MB34] directly after each RTP run. The computational cost was 1 350 core hours for the BOMD production run, 230 000 core hours for the RTP runs, and 430 core hours for compressing the electron density and performing the Voronoi integration. The resonance Raman spectra were calculated as described in Section 2.10 above.

ortho-Nitrophenol in Gas Phase

In addition to uracil, we computed resonance Raman spectra for *o*-nitrophenol in the gas phase. Resonance Raman spectra of this system have been investigated in the literature before,[76,376] in particular also by means of RT-TDDFT within the static–harmonic approximation.[376] In Figure 3.5.5, we compare the spectra from our approach *(black curves)* to these previously reported results *(red bars)* for two different laser wavelengths. The top and bottom panels correspond to the non-resonant and resonant regime, respectively. As expected, there are some differences because the AIMD-based spectra take into account some anharmonic effects which are completely missing in the static spectra. This leads both to shifts in band positions and to line broadening which alters the peak heights. Apart from these effects, we find that the change in relative band intensities and the total increase in intensity due to the resonance Raman effect is captured by our approach very well.

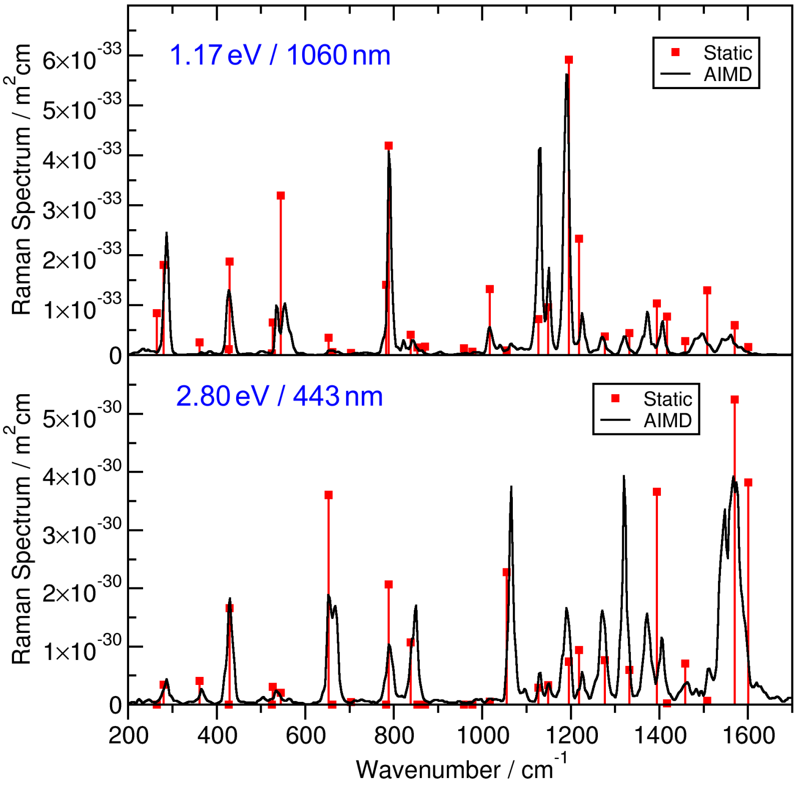


Figure 3.5.5.: Resonance Raman spectra of *o*-nitrophenol in the gas phase predicted from AIMD[MB36] *(black curves)* compared to results from static–harmonic calculations[376] *(red bars)* for two different laser energies *(see panels)*. Note the change in intensity due to the resonance effect.

As explained above, our method does not require a set of laser energies as input, but yields the resonance Raman spectra for all possible laser energies in one pass. In Figure 3.5.6, we present the set of all such spectra for *o*-nitrophenol, with the laser energy on the ordinate axis, and the vibrational frequency shown on the abscissa. Each spectrum (*i. e.*, each row of the plot) has been normalized to uniform maximum band height, because otherwise the non-resonant spectra would not be visible at all due to the strong increase in intensity caused by the resonance Raman effect. It is clearly visible how bands which are almost invisible in the non-resonant Raman spectrum become very intense at certain laser energies (*e. g.*, the bands at 850, 1050, and 1600 cm^{-1}). Such an increase in intensity only happens if the spectral band involves movement of atoms which take part in the electronic excitation at a given laser energy *(vibronic coupling)*.

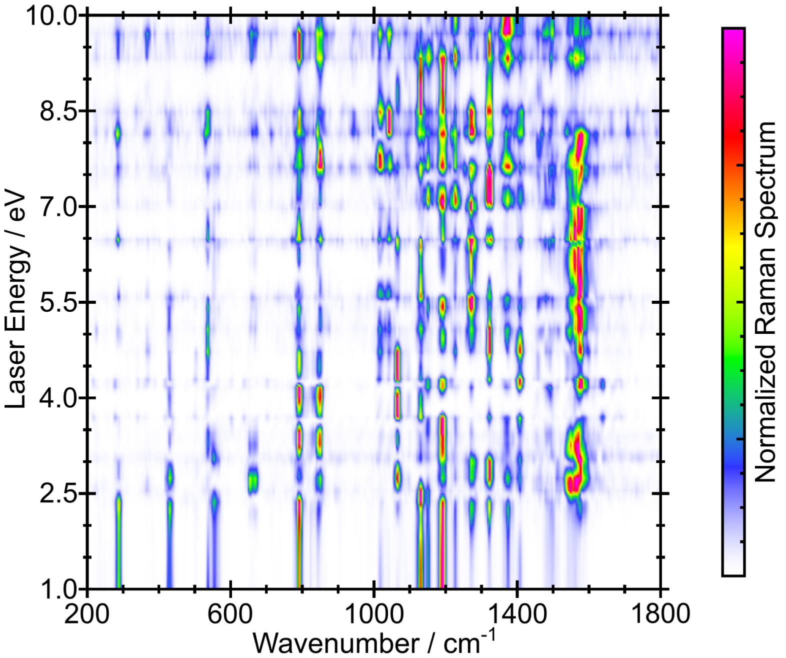


Figure 3.5.6.: Predicted excitation profile of *o*-nitrophenol in the gas phase, *i. e.*, resonance Raman spectra for all possible laser wavelengths *(vertical axis)*.[MB36] Rows are normalized to show relative intensity ratio.

3.6. Normal Mode Analysis

[EMIm][OAc] / Water Mixture

Based on the approach by Mathias *et al.*[389,390] *(see Section 2.11)* which we have implemented in the TRAVIS program package,[MB19] we have computed normal modes from an AIMD simulation of the ionic liquid 1-ethyl-3-methylimidazolium acetate,[MB8,MB13,MB19,MB31] also known as [EMIm][OAc], with a molecular structure as depicted in Figure 3.6.1.

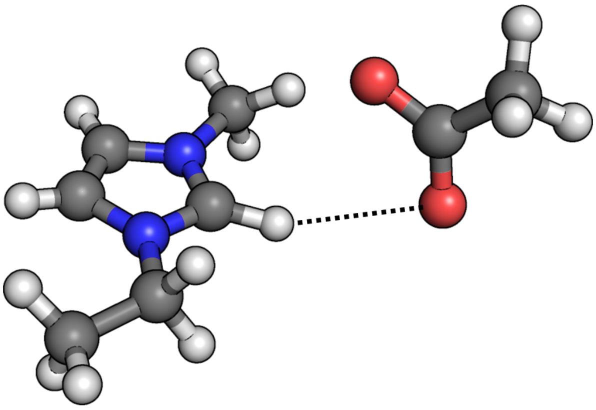


Figure 3.6.1.: Molecular structure of the ionic liquid 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]) with strong hydrogen bond between anion and cation indicated by dashed line.

One prominent feature of this ionic liquid is the strong hydrogen bond between cation and anion,[432,MB8] which is indicated in the figure by the dashed line. It is an interesting question how the water content of an ionic liquid influences the hydrogen bond strength between cation and anion. This can be achieved by vibrational spectroscopy, because it is well known that hydrogen bonding leads to a red shift and line broadening of the C–H stretching mode.[433] However, these effects cannot be directly observed in experiment, because there exist many different kinds of C–H bonds *(see Figure 3.6.1)*, and a slight shift of a single such band will vanish in the superposition of all signals in that frequency range. A bulk phase normal mode analysis is an extremely helpful approach to identify such an effect anyways.

In Figure 3.6.2, the results of such a normal mode analysis for AIMD simulations of liquid [EMIm][OAc] with and without water are shown. The left panel shows the mode spectra for all C–H stretching modes in the system, where the three stretching modes for the $[\text{EMIm}]^+$ ring protons are highlighted by red, green, and blue color. The upper part of the panel corresponds to the simulation of pure [EMIm][OAc], while the lower part depicts the [EMIm][OAc] / water mixture. The

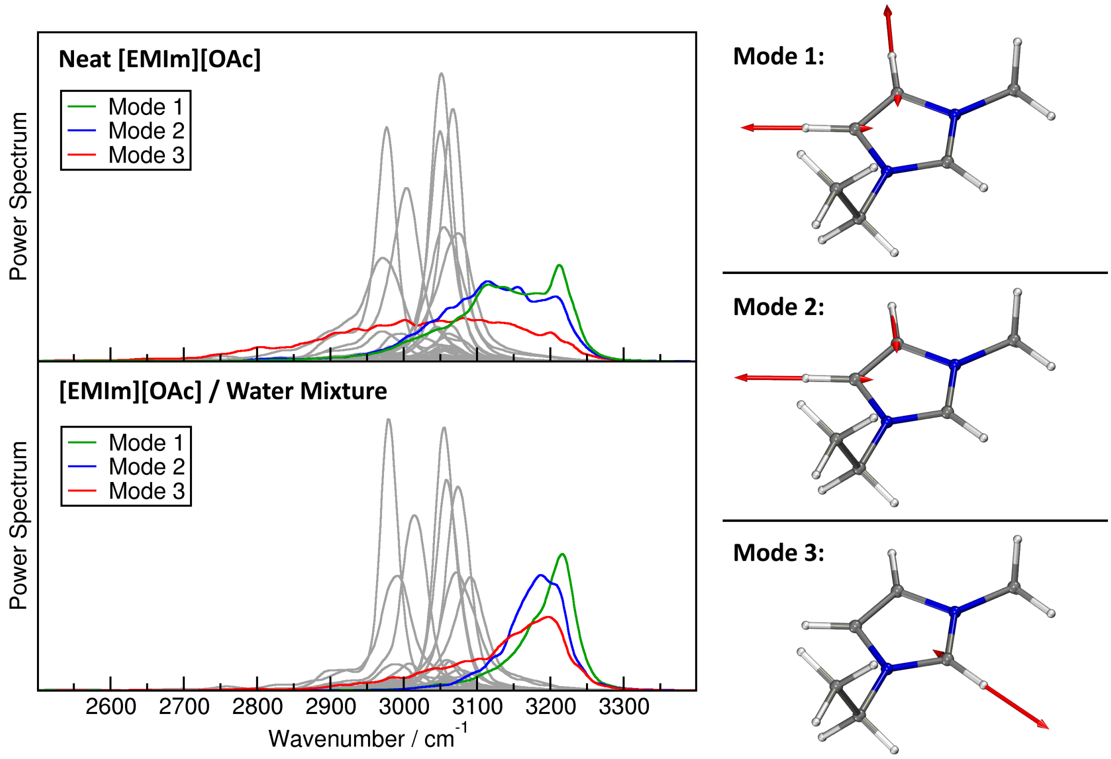


Figure 3.6.2.: Left panel: Normal mode spectra of C–H stretching modes in [EMIm][OAc] *(upper panel)* and [EMIm][OAc] / water mixture *(lower panel)*; Right panel: Three C–H stretching normal modes of the ring protons.[MB19,MB40]

right panel of the Figure visualizes the displacement vectors of the three highlighted modes. While "Mode 1" and "Mode 2" *(green and blue)* are the stretching modes of the two neighboring rear ring protons, "Mode 3" *(red color)* belongs to the stretching motion of the isolated front ring proton.

It can be seen that in the pure system, the C–H stretching modes for all three ring protons have a very large line width as a consequence of the strong hydrogen bonds in which these protons are involved. The spectral contribution ranges from 3300 cm^{-1} *(which is close to the excepted frequency in vacuum)* down to 2700 cm^{-1}, which is a significant red shift. Mode 3 has more contributions at low frequencies, which corresponds to a stronger hydrogen bond of the isolated ring proton when compared to the other two, as it has been reported in the literature before.[MB8]

In the aqueous system, on the other hand, the broadening of the three lines is significantly less pronounced, and the three bands have a lower frequency range of around 2900 cm^{-1}. This indicates a slightly weaker but still distinct hydrogen bond. It can be concluded that the addition of water weakens the hydrogen bond between the acetate anion and the three ring protons of the cation, while the hydrogen bond formed by the isolated ring proton is still the most important one.

As described in Section 2.11, the normal mode analysis requires one or more reference structures for each molecule type present in the simulation. The molecules in the trajectory are then projected onto the closest matching reference structure *(according to the RMSD)*. If the molecule is flexible or has rotating groups *(as it is the case with [EMIm]$^+$)*, a single reference structure will not suffice. For methyl groups, our implementation in Travis can automatically create all permutations of the protons, so that all relevant reference structures for the methyl rotamers are generated on-the-fly. However, for the ethyl group in [EMIm]$^+$, this is not easily possible. Therefore, we have supplied two reference structures for different ethyl group orientations as shown in Figure 3.6.3.

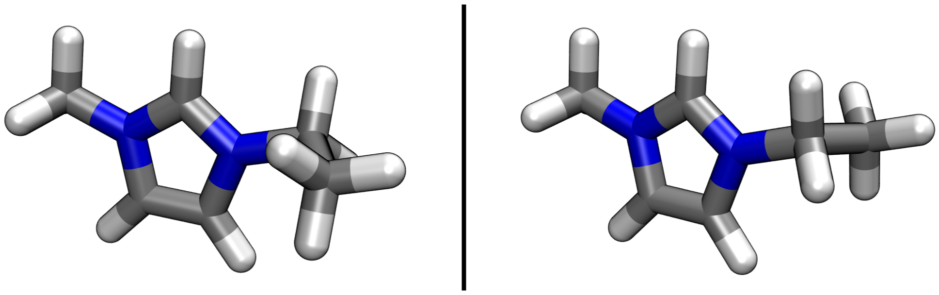


Figure 3.6.3.: Two reference structures of [EMIm]$^+$ which have been used for the normal mode analysis in Figure 3.6.2.

For each of these structures, the permutations for the two methyl groups were automatically generated, so that a total of $2 \times 3 \times 3 = 18$ reference structures resulted for the normal mode analysis of [EMIm]$^+$. In [OAc]$^-$, only one methyl group is present, so that the on-the-fly generation of the three relevant permutations was sufficient. The resulting normal mode spectra were obtained individually for each reference structure and averaged before plotting.

The computational details of the simulations were as follows. For the pure ionic liquid trajectory, 36 ion pairs of [EMIm][OAc] where placed in a cubic cell of 2121 pm edge length, while the mixture contained 27 ion pairs of [EMIm][OAc] together with 81 water molecules in a cubic cell of 2158 pm edge length. Figure 3.6.4 shows a snapshot of the mixture. All AIMD simulations were performed at 350 K. Note that no electromagnetic moments were computed here, because the normal mode analysis works on the power spectrum of the system, which is not connected to any selection rules. More computational details can be found in Section 3.7.

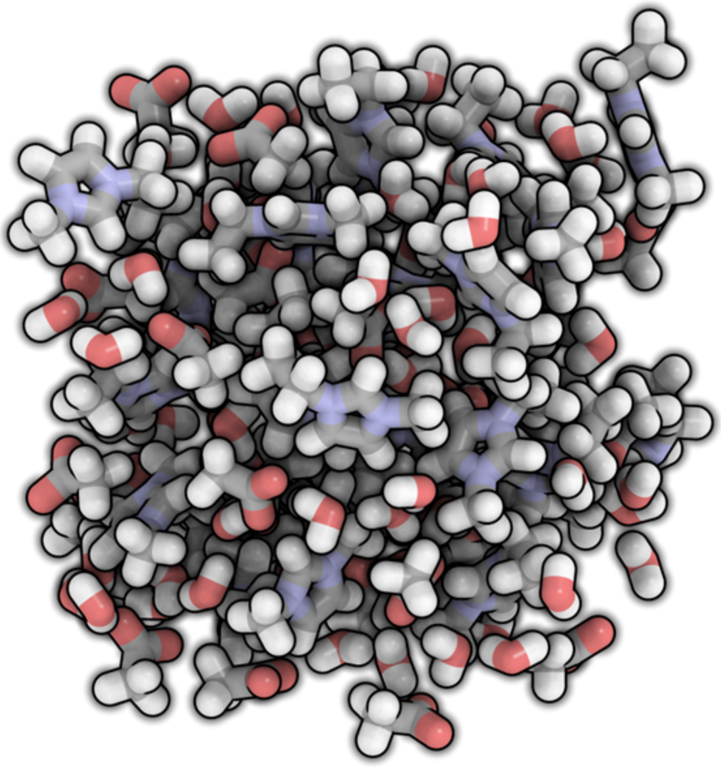


Figure 3.6.4.: Snapshot of the [EMIm][OAc]/water mixture simulation cell.[MB8,MB13,MB19,MB31]

3.7. Computational Details

In this section, general computational details for all spectra in Chapter 3 are given. Unless otherwise noted in each section, the calculations have been performed as stated below.

For bulk phase simulations, all molecules which constitute the system were placed inside a cubic simulation cell with the Packmol software.[222,223] A multi-step pre-equilibration protocol follows, performed with the LAMMPS program package.[220] A first equilibration was performed for 2.5 ps in an NVE ensemble at a temperature of 500 K using a Berendsen thermostat[434] with a time constant of 1.0 fs. Afterwards, the temperature was ramped down to the target temperature during a time interval of 50 ps. Later, a simulation of 100 ps in an NpT ensemble using a Nosé–Hoover chain thermostat[184–186] with a coupling constant of 100 fs and a Nosé–Hoover barostat with a coupling constant of 2000 fs was conducted. The acoustic shock waves resulting from the size change of the simulation box were damped in a subsequent run using a Langevin thermostat[435,436] with a coupling constant of 100 fs for a duration of 50 ps. A further simulation step without the Langevin thermostat for a duration of 2.5 ns followed. In an another NpT simulation lasting 7.5 ns using the Nosé-Hoover thermostat with a coupling constant of 100 fs, the average value of the volume of the simulation box was computed and the resulting density and box size were noted. Subsequently, the volume of the simulation cell was changed to the previously obtained average value during a 100 ps run. The resulting shock waves were again damped in another simulation using the Langevin thermostat with a coupling constant of 100 fs and a simulation time of 250 ps. To equilibrate the system, a simulation in a NVT ensemble lasting 2.5 ns was performed by using a Nosé–Hoover chain thermostat with a coupling constant of 100 fs. The final production run was conducted in NVT ensemble using a Nosé–Hoover thermostat with a coupling constant of 100 fs and a simulation time of 10 ns. The integrator time step was set to 0.5 fs in all steps. If water was present in the system, it was modeled by the TIP4P–EW model,[426] and the bonds and angles in water were kept rigid via the RATTLE algorithm.[437,438] No other bonds or angles were constrained. We used a Coulomb and Lennard-Jones cutoff radius of 800 pm and a PPPM long-range Coulomb solver *(as implemented in LAMMPS)*.[220] Finally, the last frame of the force field simulation was used to set up *ab initio* molecular dynamics (AIMD) simulations. For gas phase systems, the steps described above were skipped.

For the AIMD simulations, we used the program package CP2k,[212–214] employing the Quickstep method[402] and orbital transformation (OT)[403,404] with the FULL_KINETIC preconditioner[405] and the direct inversion in iterative subspace (DIIS) minimizer[439–441] for faster convergence. The electronic structure was calculated with density functional theory (DFT),[442,443] utilizing the BLYP functional[444,445] together with the recent re-parametrization[446] of Grimme's D3 dispersion correction[412,413] with Becke–Johnson damping.[447–449] The initial guess in each step was extrapolated from previous converged wave functions by using the always stable predictor–corrector (ASPC) method with an order of 3.[450,451] Basis sets of the kind MOLOPT–DZVP–SR–GTH[414] and GTH pseudopotentials[415,416] were applied to all atoms. The plane wave cutoff was set to 350 Ry, and an SCF convergence criterion of 10^{-6} was used. The time step was chosen to be 0.5 fs in all simulations. The simulation temperature was adjusted to 300 K by a Nosé–Hoover chain thermostat[184–186] (*i. e.*, NVT ensemble). After two more equilibration intervals (1 ps with massive thermostating and thermostat time constant $\tau = 10$ fs, then 10 ps with global thermostating and thermostat time constant $\tau = 100$ fs), a production run with the thermostat settings of the latter equilibration interval was performed for 30.0 ps (*i. e.*, 60 000 time steps).

During the AIMD production runs, the total electron density was exported in Gaussian Cube file format by CP2k and subsequently compressed to bqb format using the bqbtool.[MB34] To obtain the molecular electromagnetic moments from the total electron density, Voronoi integration[MB23,MB48] was performed using TRAVIS,[MB2,MB40] employing the atomic van der Waals radii[240–242] from literature as Voronoi radii. The calculation of the spectra from the electromagnetic moments was also performed with TRAVIS.

The full workflow from the idea to the final spectrum involves only open-source free software. A tutorial on computing vibrational spectra with CP2k and TRAVIS can be found on M.B.'s homepage:

`https://brehm-research.de/spectroscopy`

Chapter 4.

Conclusions

This habilitation thesis is focused on methods for predicting vibrational spectra of condensed phase systems. After an overview of existing methods is given in the introduction, several novel methods which have been developed within the scope of this thesis are presented in Chapter 2 and then applied to a few example systems in Chapter 3.

In some cases, these methods were able to correctly predict spectra that could not be computed before. For example, we published the first *ab initio* prediction of a liquid phase ROA spectrum[MB26] in 2017—our novel approach is described in Section 2.8, while the spectrum and the computational details are discussed in Section 3.4. Two years later, we were able to present the first *ab initio* prediction of a condensed phase resonance Raman spectrum[MB36] with a newly developed protocol that is explained in Section 2.10 *(for details on the results, see Section 3.5)*.

Other types of spectra such as infrared and Raman spectra could be predicted for liquid phase systems since several decades, but we were still able to improve on the existing protocols. For example, our Voronoi integration approach[MB23] saves up to a factor 2 of total computer time for the prediction of infrared and Raman spectra when compared to the widely used Wannier localization, and works flawlessly also for aromatic systems *(see Section 2.1.4 for a discussion)*.

As a by-product of our protocol to derive vibrational spectra from the total electron density along an AIMD simulation, we have developed a lossless compression algorithm for trajectories of volumetric data or atom positions,[MB34] which is able to reach a very high compression ratio with moderate compression and decompression times, so that it is well suitable for archiving simulation trajectories *(see Section 2.2)*.

All methods and approaches described in Chapter 2 of this thesis have been implemented in the publicly available TRAVIS program package,[MB2,MB40] which was used to compute all spectra which are shown in this thesis. TRAVIS is open source free software under the GNU GPL license and can be freely obtained from

`http://www.travis-analyzer.de`

The underlying AIMD simulations for the prediction of the spectra were carried out with the CP2k software in all cases, which is also free software. Therefore, the full workflow from the idea to the final spectrum involves only open-source free software, which is an important aspect. For a step-by-step tutorial on how to predict bulk phase vibrational spectra with CP2k and TRAVIS, please visit

`https://brehm-research.de/spectroscopy`

It should be emphasized that all predicted spectra presented above are shown directly as computed. No empirical corrections *(scaling, shifting, line broadening, etc.)* have been applied. As the protocol for computing spectra is based on *ab initio* molecular dynamics simulations, the method is truly predictive, and can be applied to predicting spectra of substances never measured *(or even synthesized)* before.

A. Appendix

A.1. The Ergodic Theorem

The probably most crucial point when analyzing trajectories from molecular simulation is the ergodic theorem. In a very simplified form, it states that for any quantity which can be defined within a system, the *temporal average* equals the *ensemble average* if the *propagator* of the simulation (i. e., the function which propagates one step into the following step) is *ergodic*. Only if this holds, it is allowed to perform a simulation over a certain number of steps, and then analyze the results without explicitly discussing the initial configuration. If this did not hold, any result would strongly depend on the choice of the initial configuration (even for *very* long simulation times), which would render all analyses useless.

As this concept is of such great importance, it will be discussed in some more detail here. This is done on an abstract level, utilizing mathematical formalisms from *measure theory*. The central object in the following part is the *finite measure space* (X, Σ, μ). X is the *phase space* of the dynamical system – a set which contains all different states the system can exhibit. If N (distinguishable) particles are simulated, X is a $6N$-dimensional space, as each particle's position and velocity are considered. Σ is a σ-algebra over X, which intuitively (and simplified) means that Σ is a set which contains all possible subsets of X. Finally, μ is a measure on X, i. e., a map which assigns real numbers to all elements of Σ:

$$\mu : \Sigma \to \mathbb{R}, \quad E \mapsto \mu(E) \ \forall E \in \Sigma. \tag{A.1.1}$$

In our case, this can be imagined as simply *counting* the number of different configurations within a subset of the phase space X (actually, μ measures the *phase space volume* of a subset of the phase space). The predicate *finite* in case of a measure space (X, Σ, μ) requires that all subsets of the space X have a measure $< \infty$ with regard to μ:

$$\mu(E) < \infty \ \forall E \in \Sigma. \tag{A.1.2}$$

Another important element of this theory is the *propagator function* $T : X \to X$. This function maps one certain configuration from the phase space to the consecutive configuration within the simulation. To give an example, in the case of molecular dynamics simulation this would mean to calculate the forces and to perform a time integration over a finite time step. The subsequent application of the propagator function, i. e., performing multiple steps, is denoted by a superscript index:

$$T^n(a) = T\Big(\cdots T(a) \cdots\Big), \quad a \in X. \tag{A.1.3}$$

For the case of *conservative* systems, which are considered here, the *Liouville theorem* requires that T is a *measure-preserving transformation*, which keeps the phase space volume constant with regard to μ:

$$\mu(A) = \mu\Big(T(A)\Big) \;\forall A \in \Sigma. \tag{A.1.4}$$

The Poincaré Recurrence Theorem

By using the terminology introduced above, an important theorem can be expressed, the Poincaré recurrence theorem,[452] published in 1890:

> Let (X, Σ, μ) be a finite measure space, and $T : X \to X$ a measure-preserving transformation. Then, for any $A \in \Sigma$, the subset of points $x \in A$ which fulfill $T^n(x) \notin A \;\; \forall n > 0$ has a measure of zero. Intuitively speaking, μ-almost every point within the subset of the phase space A returns into A at some later time.

This theorem already has some far-reaching implications: A simulation which is started from a certain initial configuration will very likely return to a configuration which is *close* to its initial configuration. There exist only very few initial configurations (a *zero measure set* within the overall phase space) for which this statement does not hold.

Definition of Ergodicity

> Let $T : X \to X$ be a measure-preserving transformation on a measure space (X, Σ, μ) with $\mu(X) = 1$. If for every $A \in \Sigma$ with $T^{-1}(A) = A$ either $\mu(A) = 0$ or $\mu(A) = 1$ holds, then the transformation T is called *ergodic*.

The picture behind this rather abstract definition is the following: A measure-preserving transformation T is called ergodic if and only if the only two subsets of the phase space which are invariant under T are the empty set and the complete phase space. This means, any true subset A of the phase space X will be modified by the transformation T, while at the same time $\mu(A) = \mu\Big(T(A)\Big)$ holds (T preserves the measure).

An even more intuitive picture can be constructed. Imagine the set X as a bowl filled with water. Now, a tiny droplet of dye is dropped into the water. Then, successive iterations of the inverse of an ergodic transformation (e. g., stirring) will distribute the dye over all the water – it cannot happen that the dye remains within a certain sub-region of the bowl. This is what ergodicity states. At the same time, the water will neither be compressed nor dilated in any sub-region of the bowl, because the transformation preserves the measure, which is the density in this case.

Birkhoff's Ergodic Theorem

> Let $T : X \to X$ be a measure-preserving transformation on a measure space (X, Σ, μ) with $\mu(X) = 1$. Let $f : X \to \mathbb{R}$ a μ-integrable function. Define the temporal average of f for initial value x as
>
> $$\hat{f}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f\Big(T^k(x)\Big) \qquad \text{(A.1.5)}$$
>
> and the ensemble average of f as
>
> $$\overline{f} = \int_X f \, d\mu. \qquad \text{(A.1.6)}$$
>
> If T is an ergodic transformation, then both averages are equal, and
>
> $$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f\Big(T^k(x)\Big) = \int_X f \, d\mu \qquad \text{(A.1.7)}$$
>
> holds for μ-almost all initial values x.

This is Birkhoff's ergodic theorem,[453,454] which was first published in 1931. If f is considered as the function which should be determined from the trajectory, and if the simulation is indeed ergodic (which is very hard to proof, though), this theorem allows to evaluate the temporal average instead of the ensemble average (which cannot be obtained directly), because both are equal. Therefore, strictly speaking, all predictions of spectra from MD simulations are based on this theorem.

A.2. The Wiener–Khinchin Theorem

The autocorrelation of a discrete set of n scalar values $v(i)$, $i = \{1, ..., n\}$, is given by the following equation:

$$C(\tau) = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \Big(v(t)\ v(t+\tau) \Big) \tag{A.2.1}$$

Unfortunately, the computational cost to evaluate $C(\tau)$ for all values $t = 1, ..., n$ scales with $\mathcal{O}(n^2)$, which slows down the calculation drastically for large data sets. Typical applications in TRAVIS require autocorrelations of sets with 10^5 or more values to be calculated, demanding amounts of computer time which are no longer acceptable for the user.

Autocorrelation functions are required by many approaches to compute vibrational spectra in the time-correlation function formalism. Therefore, it is very important to find a method which circumvents the problem depicted above.

Fortunately, there exists a way to compute the autocorrelation function by utilizing the Fourier transform, and this can be done by utilizing the Fast Fourier transform (FFT)[392,393] algorithm with a time complexity of $\mathcal{O}\Big(n \log(n)\Big)$. The computation is based on the Wiener–Khinchin theorem,[272,273,455] which has been independently discovered by N. Wiener in 1930[281] and by A. Khinchin in 1934:[282]

> "The autocorrelation function of a wide-sense-stationary random process has a spectral decomposition given by the power spectrum of that process."

TRAVIS utilizes this methodology, which strongly accelerates the overall process. The theorem will be derived in the following.

The autocorrelation of a complex-valued continuous function $f(t)$ is defined by

$$C(\tau) \equiv \int_{-\infty}^{\infty} \overline{f}(t)\; f(t+\tau)\; dt. \tag{A.2.2}$$

The Fourier transform $F(\nu)$ of $f(t)$, and its inverse $\mathcal{F}^{-1}$, on the other hand, are given by

$$F(\nu) \;=\; \mathcal{F}\Big(f(t)\Big) \;=\; \int_{-\infty}^{\infty} f(t)\; e^{-i\nu t}\, dt \tag{A.2.3}$$

$$f(t) \;=\; \mathcal{F}^{-1}\Big(F(\nu)\Big) \;=\; \int_{-\infty}^{\infty} F(\nu)\; e^{i\nu t}\, d\nu \tag{A.2.4}$$

For the complex conjugate $\overline{f}(t)$ of $f(t)$, this leads to

$$\overline{f}(t) = \int_{-\infty}^{\infty} \overline{F}(\nu)\; e^{-i\nu t}\, d\nu \tag{A.2.5}$$

By substituting these expressions into the autocorrelation integral, we obtain

$$C(\tau) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \overline{F}(\nu)\; e^{-i\nu t}\, d\nu \right) \left(\int_{-\infty}^{\infty} F\big(\nu'\big)\; e^{i\nu'(t+\tau)}\, d\nu' \right) dt \tag{A.2.6}$$

$$C(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \overline{F}(\nu)\; F\big(\nu'\big)\; e^{i(\nu-\nu')t}\, e^{i\nu'\tau}\, d\nu\, d\nu'\, dt \tag{A.2.7}$$

$$C(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \overline{F}(\nu)\; F\big(\nu'\big)\; \delta\big(\nu-\nu'\big)\; e^{i\nu'\tau}\, d\nu\, d\nu' \tag{A.2.8}$$

$$C(\tau) = \int_{-\infty}^{\infty} \overline{F}(\nu)\; F(\nu)\; e^{i\nu\tau}\, d\nu \tag{A.2.9}$$

$$C(\tau) = \int_{-\infty}^{\infty} |F(\nu)|^2\; e^{i\nu\tau}\, d\nu \tag{A.2.10}$$

$$C(\tau) = \mathcal{F}^{-1}\left(\left| \mathcal{F}\Big(f(\tau)\Big) \right|^2 \right) \tag{A.2.11}$$

This derivation shows that the autocorrelation of function $f(t)$ can be calculated by taking the Fourier transform of f, calculating its absolute square in frequency domain, and finally applying the back-transform to time domain.

There is also an intuitive picture behind this relation: Consider the Fourier transform of a time series of real values, it is a complex function in general. A complex number z can be expressed as $z = r\,\mathrm{e}^{i\varphi}$ with absolute value r and complex argument φ. In the Fourier transform, the absolute value contains the information on the amplitudes of certain frequencies, whereas the complex argument contains the phase information of these frequencies. If one takes the square of the absolute value of the Fourier transform, the complex argument vanishes, and only the absolute value remains: $|z|^2 = r^2$. When this function is transformed back with the inverse Fourier transform, the phase information has completely been erased, all frequency contributions are now in the same phase. This is exactly what the autocorrelation does: Removing the phase information of some time series and putting all frequency contributions into the same phase.

Finally, it should be noted that the Wiener–Khinchin theorem is a special case of the cross-correlation theorem:

> Let $f \star g$ denote the cross-correlation of two functions $f(t)$ and $g(t)$. Then
>
> $$\mathcal{F}(f \star g) = \mathcal{F}(\overline{f})\,\mathcal{F}(g) \tag{A.2.12}$$
>
> holds, which is equivalent to
>
> $$f \star g = \mathcal{F}^{-1}\Big(\mathcal{F}(\overline{f})\,\mathcal{F}(g)\Big). \tag{A.2.13}$$

In the case $f \equiv g$, and by using the equality

$$f \star f = \int\limits_{-\infty}^{\infty} \overline{f}(t)\; f(t{+}\tau)\; dt \equiv \mathrm{C}(\tau)\,, \tag{A.2.14}$$

Equation A.2.13 becomes identical to Equation A.2.11 on the preceding page.

If the complex conjugation of f in Equation A.2.13 is removed, the *convolution theorem* is yielded instead.

A.3. The Frequency Shift of the Verlet Integrator

One of the most common integrators in molecular dynamics simulations is probably the Verlet algorithm,[182] including the variants which are known as Leapfrog algorithm and velocity Verlet algorithm. It can be easily shown that all three algorithms produce identical trajectories, and are therefore considered to be equivalent.[183]

The general equations of the one-dimensional velocity Verlet integrator are given by

$$x(t+\Delta t) = x(t) + \Delta t v(t) + \frac{\Delta t^2}{2} a\big(x\,(t)\big), \tag{A.3.1}$$

$$v(t+\Delta t) = v(t) + \frac{\Delta t}{2}\Big(a\big(x(t)\big) + a\big(x(t+\Delta t)\big)\Big), \tag{A.3.2}$$

where $x(t)$ denotes the position of the system at time t, $v(t)$ stands for the velocity at time t, $a(x)$ gives the (time-independent) acceleration at position x, and Δt is a finite time step.

Equation A.3.2 requires knowledge of the new position $x(t+\Delta t)$ in order to determine the new velocity, which is disadvantageous in the further derivation. However, this can be easily overcome by substituting Equation A.3.1 into Equation A.3.2, yielding

$$v(t+\Delta t) = v(t) + \frac{\Delta t}{2}\left(a\big(x(t)\big) + a\Big(x(t) + \Delta t v(t) + \frac{\Delta t^2}{2} a\big(x(t)\big)\Big)\right). \tag{A.3.3}$$

The state of the system can be completely described by knowing its position $x(t)$ and velocity $v(t)$ at time t. Both these quantities can be stored together as a complex number $z(t)$, where

$$\mathrm{Re}\big(z(t)\big) = x(t)\,, \tag{A.3.4}$$

$$\mathrm{Im}\big(z(t)\big) = v(t)\,. \tag{A.3.5}$$

This enables to write the equations of the velocity Verlet integrator as an operator $T_{\Delta t} : \mathbb{C} \rightarrow \mathbb{C}$ by

$$\begin{aligned} T_{\Delta t}(z) =& \mathrm{Re}(z) + \Delta t\,\mathrm{Im}(z) + \frac{\Delta t^2}{2} a\big(\mathrm{Re}(z)\big) + \\ & i\left[\mathrm{Im}(z) + \frac{\Delta t}{2}\left(a\big(\mathrm{Re}(z)\big) + a\Big(\mathrm{Re}(z) + \Delta t\,\mathrm{Im}(z) + \frac{\Delta t^2}{2} a\big(\mathrm{Re}(z)\big)\Big)\right)\right]. \end{aligned} \tag{A.3.6}$$

This operator will be termed *Verlet propagator* in the following.

The differential equation of a one-dimensional harmonic oscillator reads

$$-\frac{\partial V\big(x(t)\big)}{\partial x(t)} = -F\big(x(t)\big) = m\frac{\partial^2 x(t)}{\partial t^2} \tag{A.3.7}$$

with a harmonic potential V defined by

$$V(x) = \frac{k}{2}x^2. \tag{A.3.8}$$

The well-known solution of this differential equation reads

$$x(t) = C_1 \cos\Big(\sqrt{\frac{k}{m}}t\Big) + C_2 \sin\Big(\sqrt{\frac{k}{m}}t\Big), \tag{A.3.9}$$

where usually the substitution $\omega = \sqrt{\frac{k}{m}}$ is performed, yielding

$$x(t) = C_1 \cos\Big(\omega t\Big) + C_2 \sin\Big(\omega t\Big). \tag{A.3.10}$$

The corresponding velocity $v(t)$ then is obtained as

$$v(t) = \frac{\partial}{\partial t}x(t) = -C_1\omega \sin\Big(\omega t\Big) + C_2\omega \cos\Big(\omega t\Big). \tag{A.3.11}$$

By letting $C_1 = 1$, $C_2 = 0$, it becomes clear that the trajectories of this system are ellipses in the $x - v$ phase space with the two radii 1 and ω. Only in the case $\omega = 1$ (*i.e.*, $k = m$), the trajectories are circles.

For the further derivation, it is desirable for all such trajectories to be circles. Therefore a *transformed* velocity $v'(t)$ is defined by

$$v'(t) = \frac{v(t)}{\omega}. \tag{A.3.12}$$

It follows directly from this definition that in the $x - v'$ space all trajectories which obey Equation A.3.7 are circles. In the following, only the transformed velocity will be considered.

Applying the equations of the Verlet integrator from Equations A.3.1 and A.3.3 to the harmonic oscillator by substituting

$$a\big(x(t)\big) = -\frac{1}{m}\frac{\partial V\big(x(t)\big)}{\partial x(t)} = -\frac{k}{m}x(t) = -\omega^2 x(t) \tag{A.3.13}$$

gives

$$x(t+\Delta t) = x(t) + \Delta t v(t) - \frac{\Delta t^2 \omega^2}{2} x(t)\,, \tag{A.3.14}$$

$$v(t+\Delta t) = v(t) - \frac{\Delta t \omega^2}{2}\left(2x(t) + \Delta t v(t) - \frac{\Delta t^2 \omega^2}{2} x(t)\right). \tag{A.3.15}$$

with the velocity finally being substituted by the transformed velocity v' in the second equation:

$$v'(t+\Delta t) = v'(t) - \frac{\Delta t \omega}{2}\left(2x(t) + \Delta t \omega v'(t) - \frac{\Delta t^2 \omega^2}{2} x(t)\right). \tag{A.3.16}$$

Combining Equations A.3.14 and A.3.16 into the form of a complex Verlet propagator such as in Equation A.3.6 yields

$$\begin{aligned} T_{\Delta t}(z) =& \mathrm{Re}(z) + \Delta t \omega \mathrm{Im}(z) - \frac{\Delta t^2 \omega^2}{2} \mathrm{Re}(z) \\ &+ i\left[\mathrm{Im}(z) - \frac{\Delta t \omega}{2}\left(2\mathrm{Re}(z) + \Delta t \omega \mathrm{Im}(z) - \frac{\Delta t^2 \omega^2}{2} \mathrm{Re}(z)\right)\right] \end{aligned} \tag{A.3.17}$$

with

$$\mathrm{Re}\big(z(t)\big) = x(t)\,, \tag{A.3.18}$$

$$\mathrm{Im}\big(z(t)\big) = v'(t)\,. \tag{A.3.19}$$

Please note that the parameters Δt and ω always appear with equal exponents in the above equation. The propagator is therefore invariant under any parameter change which leaves the product $\Delta t\, \omega$ invariant. This is in line with intuition, as doubling the frequency ω and halving the time step Δt is just a re-parametrization of time and should not change anything else.

As already stated above, exact solution trajectories of any harmonic oscillator are circles in the $x - v'$ space, i.e. the absolute value $|z|$ remains constant along these trajectories.

Claim 1: The Verlet propagator from Equation A.3.17 keeps $|z|$ approximately constant, and the residuum vanishes quickly as Δt becomes small.

Note: As

$$|z|^2 = x^2 + v'^2 = x^2 + \frac{v^2}{\omega^2} = \frac{2}{k} E_{pot} + \frac{2}{m} E_{kin}, \tag{A.3.20}$$

Claim 1 is equivalent to the statement that the Verlet propagator approximately conserves the total energy of the system.

Proof: Switching to the trigonometric representation of z by substituting $z = r\big(\cos(\varphi) + i\sin(\varphi)\big)$ on the right hand side of the following equation gives

$$\begin{aligned} T_{\Delta t}\big(z\big) &= r\cos(\varphi) + \Delta t\omega r\sin(\varphi) - \frac{\Delta t^2\omega^2}{2} r\cos(\varphi) \\ &\quad + i\left[r\sin(\varphi) - \frac{\Delta t\omega}{2}\left(2r\cos(\varphi) + \Delta t\omega r\sin(\varphi) - \frac{\Delta t^2\omega^2}{2} r\cos(\varphi)\right)\right] \end{aligned} \tag{A.3.21}$$

and

$$\begin{aligned} \Big|T_{\Delta t}(z)\Big|^2 &= \left(r\cos(\varphi) + \Delta t\omega r\sin(\varphi) - \frac{\Delta t^2\omega^2}{2} r\cos(\varphi)\right)^2 \\ &\quad + \left(r\sin(\varphi) - \frac{\Delta t\omega}{2}\left(2r\cos(\varphi) + \Delta t\omega r\sin(\varphi) - \frac{\Delta t^2\omega^2}{2} r\cos(\varphi)\right)\right)^2 \\ &= r^2\left(1 + \frac{1}{4}\Delta t^3\omega^3\sin(2\varphi) - \frac{1}{4}\Delta t^4\omega^4\cos(2\varphi) - \frac{1}{8}\Delta t^5\omega^5\sin(2\varphi) + \frac{1}{32}\Delta t^6\omega^6\Big(1 + \cos(2\varphi)\Big)\right), \end{aligned} \tag{A.3.22}$$

thus

$$\Big|T_{\Delta t}(z)\Big| = |z|\sqrt{1 + \frac{1}{4}\Delta t^3\omega^3\sin(2\varphi) - \frac{1}{4}\Delta t^4\omega^4\cos(2\varphi) - \frac{1}{8}\Delta t^5\omega^5\sin(2\varphi) + \frac{1}{32}\Delta t^6\omega^6\Big(1 + \cos(2\varphi)\Big)}. \tag{A.3.23}$$

It can be seen that the Verlet propagator keeps the absolute value of the argument approximately constant, and the residue is of the order of $\mathcal{O}\big(\Delta t^3\omega^3\big)$ (to see this, please note that $\sqrt{1+x^3} \approx 1 + \frac{x^3}{2} + \ldots$). A time step $\Delta t \ll \omega$ will lead to quickly vanishing deviations. Apart from that, any real simulation will uniformly sample the angle φ, such that all terms which contain $\sin(2\varphi)$ or $\cos(2\varphi)$ will cancel out on average. Depending on the angle, some propagator steps will slightly enlarge $|z|$, while others will slightly reduce it, causing a fluctuation of $|z|$ around its exact value. The only residual term in Equation A.3.23 which does not fluctuate around zero is of the order of $\mathcal{O}\big(\Delta t^6\omega^6\big)$. This term is always positive, meaning that $|z|$ (and therefore the total energy of the system) will always increase over the long term. However, a sufficiently small time step Δt will make this effect almost vanish.

As already discussed above, the exact solution trajectory of any harmonic oscillator constitutes a circle in the $x - v'$ space. Therefore it can be assumed that the discrete Verlet propagator rotates the argument by a specific angle within the complex plane in order to approximately resemble this exact trajectory.

Claim 2: The angle by which the Verlet propagator rotates its argument around the origin within the $x - v'$ space is approximately independent on the choice of the argument itself.

This would mean that successive time steps of the Verlet propagator always rotate by approximately the same angle in the $x - v'$ space, and this angle only depends on Δt and ω.

Proof: The angle between two complex numbers can be determined by utilizing the dot product:

$$\sphericalangle(z_1, z_2) = \arccos\left(\frac{\operatorname{Re}(z_1)\operatorname{Re}(z_2) + \operatorname{Im}(z_1)\operatorname{Im}(z_2)}{|z_1| \cdot |z_2|}\right). \tag{A.3.24}$$

To obtain the angle by which the Verlet propagator rotates its argument, write

$$\sphericalangle\Big(z, T_{\Delta t}(z)\Big) = \arccos\left(\frac{\operatorname{Re}(z)\operatorname{Re}\big(T_{\Delta t}(z)\big) + \operatorname{Im}(z)\operatorname{Im}\big(T_{\Delta t}(z)\big)}{|z| \cdot \big|T_{\Delta t}(z)\big|}\right). \tag{A.3.25}$$

Following from Claim 1, we can safely approximate $\big|T_{\Delta t}(z)\big| \approx |z|$, and therefore $|z| \cdot \big|T_{\Delta t}(z)\big| \approx |z|^2$. Substituting $z = r\big(\cos(\varphi) + i\sin(\varphi)\big)$ yields

$$\sphericalangle\Big(z, T_{\Delta t}(z)\Big) = \arccos\left(\frac{r\cos(\varphi)\operatorname{Re}\big(T_{\Delta t}(z)\big) + r\sin(\varphi)\operatorname{Im}\big(T_{\Delta t}(z)\big)}{r^2}\right). \tag{A.3.26}$$

Inserting the terms from Equation A.3.21 for $\operatorname{Re}\big(T_{\Delta t}(z)\big)$ and $\operatorname{Im}\big(T_{\Delta t}(z)\big)$ and dividing by r^2 leads to

$$\begin{aligned}\sphericalangle\Big(z, T_{\Delta t}(z)\Big) &= \arccos\Bigg(\cos^2(\varphi) - \frac{\Delta t^2\omega^2}{2}\cos^2(\varphi) + \sin^2(\varphi) \\ &\qquad - \frac{\Delta t^2\omega^2}{2}\sin^2(\varphi) + \frac{\Delta t^3\omega^3}{4}\sin(\varphi)\cos(\varphi)\Bigg) \qquad \text{(A.3.27)} \\ &= \arccos\left(1 - \frac{\Delta t^2\omega^2}{2} + \frac{\Delta t^3\omega^3}{4}\sin(\varphi)\cos(\varphi)\right). \qquad \text{(A.3.28)}\end{aligned}$$

Keeping in mind that $\arccos\big(1 - x^2\big) \approx x + \mathcal{O}\big(x^3\big)$, it is visible that the rotation angle is approximately linearly dependent of the time step Δt for fixed ω, which is completely in line with the expectations, as the Verlet propagator need to traverse larger pieces of the trajectory circle with larger time steps. The term which depends on φ is of higher order with respect to $\Delta t\omega$, and therefore vanishes quickly. Apart from that, like already discussed in the proof of claim 1, the angles φ will be uniformly sampled in a real simulation, canceling out the residual terms on average over long runs, as $\int_0^{2\pi} \sin(\varphi)\cos(\varphi)\,\mathrm{d}\varphi = 0$. Some of the time steps will possess a rotation angle above average, others below.

The final approximation for the angle (which neglects the residual terms) therefore reads

$$\sphericalangle\Big(z, T_{\Delta t}(z)\Big) \approx \arccos\Big(1 - \frac{\Delta t^2 \omega^2}{2}\Big). \tag{A.3.29}$$

Derivation of the Frequency Shift

Based on the angle α which one time step covers in the $x - v'$ space derived in Equation A.3.29, the cycle duration of the discrete system can be determined by

$$\begin{aligned} \tau_{\text{cycle}} &= \frac{2\pi}{\alpha}\Delta t & \text{(A.3.30)} \\ &= \frac{2\pi\Delta t}{\arccos\big(1 - \frac{\Delta t^2\omega^2}{2}\big)}. & \text{(A.3.31)} \end{aligned}$$

Applying $\omega = \frac{2\pi}{\tau_{\text{cycle}}}$, one then obtains

$$\omega_{\text{verlet}} = \frac{\arccos\big(1 - \frac{\Delta t^2 \omega_{\text{exact}}^2}{2}\big)}{\Delta t}. \tag{A.3.32}$$

In practical applications, it might be more useful to compute the exact frequency based on the approximate one. Fortunately, the inverse function of the above is easily given by

$$\omega_{\text{exact}} = \frac{\sqrt{2 - 2\cos\left(\Delta t\, \omega_{\text{verlet}}\right)}}{\Delta t}. \tag{A.3.33}$$

Results

In the following, some results with parameters from typical simulations are shown. For a typically used time step of 0.5 fs, the deviation of the frequency is found to be around 10 cm^{-1} for a C–H vibration at 3000 cm^{-1}. At $\Delta t = 1.0\,\mathrm{fs}$, the deviation even amounts to 40 cm^{-1}.

Table A.1.: Frequency shift of Verlet integrator for typical vibrations and time steps Δt.

ω_{exact} / cm^{-1}	ω_{verlet}		
	$\Delta t = 0.1\,\mathrm{fs}$	$\Delta t = 0.5\,\mathrm{fs}$	$\Delta t = 1.0\,\mathrm{fs}$
10	10.000 000 014 8	10.000 000 369	10.000 001 48
100	100.000 014 8	100.000 369	100.001 48
1000	1000.0148	1000.370	1001.483
2000	2000.118	2002.967	2012.013
3000	3000.399	3010.064	3041.396
4000	4000.946	4024.025	4101.159
10 000	10 014.834	10 411.945	13 033.591

A.4. Explicit Reversible Nosé–Hoover Chain Integrator

Below, an explicit and reversible algorithm to apply a Nosé–Hoover chain thermostat to a molecular simulation is presented.[191] The Yoshida–Suzuki integrator[192] and a multiple-time-step scheme are used. Note that this is a half-step and needs to be applied **twice**, before and after the velocity Verlet integration of the particles.

Algorithm 1 NHC Integrator Half-Step[191]

```
$akin := \sum_{i=1}^{N} m_i \mathbf{v}_i^2$
for $k = 1 \ldots n_c$ do
  for $j = 1 \ldots m$ do
    $\delta t_s := w_j \delta t / n_c$
    # Calculate Thermostat Acceleration
    $\ddot{\xi}_1 := \left(akin - N_f k_\mathrm{B} T_\mathrm{set}\right) / Q_1$
    for $i = 2 \ldots M$ do
      $\ddot{\xi}_i := \left(Q_{i-1} \dot{\xi}_{i-1}^2 - k_\mathrm{B} T_\mathrm{set}\right) / Q_i$
    end for
    # Update Thermostat Velocities
    $\dot{\xi}_M := \dot{\xi}_M + \delta t_s / 4 \cdot \ddot{\xi}_M$
    for $i = 1 \ldots M - 1$ do
      $aa := \exp\left(-\delta t_s / 8 \cdot \dot{\xi}_{M+1-i}\right)$
      $\dot{\xi}_{M-i} := \dot{\xi}_{M-i} \cdot aa^2 + \delta t_s / 4 \cdot \ddot{\xi}_{M-i} \cdot aa$
    end for
    # Update Particle Velocities
    $aa := \exp\left(-\delta t_s / 2 \cdot \dot{\xi}_1\right)$
    for $i = 1 \ldots N$ do
      $\mathbf{v}_i := \mathbf{v}_i \cdot aa$
    end for
    # Update Thermostat Positions
    for $i = 1 \ldots M$ do
      $\xi_i := \xi_i + \delta t_s / 2 \cdot \dot{\xi}_i$
    end for
    # Update Thermostat Accelerations
    $akin := akin \cdot aa^2$
    $\ddot{\xi}_1 := \left(akin - N_f k_\mathrm{B} T_\mathrm{set}\right) / Q_1$
    for $i = 2 \ldots M$ do
      $\ddot{\xi}_i := \left(Q_{i-1} \dot{\xi}_{i-1}^2 - k_\mathrm{B} T_\mathrm{set}\right) / Q_i$
    end for
    # Update Thermostat Velocities
    for $i = 1 \ldots M - 1$ do
      $aa := \exp\left(-\delta t_s / 8 \cdot \dot{\xi}_{i+1}\right)$
      $\dot{\xi}_i := \dot{\xi}_i \cdot aa^2 + \delta t_s / 4 \cdot \ddot{\xi}_i \cdot aa$
      $\ddot{\xi}_{i+1} := \left(Q_i \dot{\xi}_i^2 - k_\mathrm{B} T_\mathrm{set}\right) / Q_{i+1}$
    end for
    $\dot{\xi}_M := \dot{\xi}_M + \delta t_s / 4 \cdot \ddot{\xi}_M$
  end for
end for
```

In this algorithm, M is the number of thermostats (NHC chain length), $\xi_1 \dots \xi_M$, $\dot{\xi}_1 \dots \dot{\xi}_M$, $\ddot{\xi}_1 \dots \ddot{\xi}_M$, and $Q_1 \dots Q_M$ are the positions, velocities, accelerations, and masses of the M virtual thermostat particles, respectively, N is the number of real particles (atoms), $m_1 \dots m_N$ are the particle masses, $v_1 \dots v_N$ are the particle velocities, n_c is the number of multiple time steps used, m is the Yoshida–Suzuki integrator order, N_f depicts the active degrees of freedom of the real particle system, k_B is the Boltzmann constant, T_set is the thermostat's target temperature, and $akin$ as well as aa are scalar auxiliary variables.

Chain lengths M larger than 2 are recommended in general cases where there is a broad distribution of vibrational frequencies in the system.[189] Typical values used in real simulations are $M = 3$, $n_c = 2$, and $m = 3$.

As noted in Section 1.3, the virtual particle masses $Q_1 \dots Q_M$ should be chosen[191,192] according to

$$Q_1 := N_f k_\mathrm{B} T_\mathrm{set} \tau^2, \tag{A.4.1}$$

$$Q_i := k_\mathrm{B} T_\mathrm{set} \tau^2, \qquad i = 2, \dots, M, \tag{A.4.2}$$

where τ is the desired time constant of the Nosé–Hoover chain thermostat.

The coefficients w_i for the Yoshida–Suzuki integrator are shown in Table A.2.[191,192]

Table A.2.: Coefficients w_j for Yoshida–Suzuki integrators of different order m.[191,192]

m	w_j	
1	$w_1 =$	1
3	$w_1 = w_3 =$	1.351 207 191 959 65
	$w_2 =$	−1.702 414 383 919 31
5	$w_1 = w_5 =$	0.414 490 771 794 37
	$w_2 = w_4 =$	0.414 490 771 794 37
	$w_3 =$	−0.657 963 087 177 50
7	$w_1 = w_7 =$	−1.177 679 984 178 87
	$w_2 = w_6 =$	0.235 573 213 359 35
	$w_3 = w_5 =$	0.784 513 610 477 56
	$w_4 =$	1.315 186 320 683 92

A.5. Approximate Overtones from Classical Oscillators

In vibrational spectroscopy, both overtones and combination bands are important effects which are essential to the understanding of experimental spectra. To give an example, liquid water appears blue only because a four-quantum overtone transition at 698 nm absorbs parts of the red light.[456] Overtones and combination bands are purely quantum mechanical phenomena and do not have classical equivalents. Thus, one would not expect them in spectra predicted from AIMD simulations, where the nuclei are treated classically. However, interestingly, a similar phenomenon also can be observed in purely classical oscillators as soon as the potential is anharmonic, as it has been shown in the literature.[113]

A simple illustration of this effect is presented in Figure A.1, where the potential functions *(upper panel)* and the power spectra *(lower panel)* of a harmonic *(left)* and a Morse oscillator[457,458] *(right)* are shown. The different lines in the power spectra correspond to different temperatures *(amplitudes of the oscillators)*. While in the harmonic case, all power spectra possess only a single peak at the fundamental frequency of the oscillator, the position of this peak is shifted to lower wavenumbers in the Morse oscillator when the temperature is increased. Furthermore, additional bands similar to overtones appear at multiples of the wavenumber of the fundamental mode. This is in contrast to quantum mechanical overtones, which appear at slightly less than full multiples of the fundamental frequency.

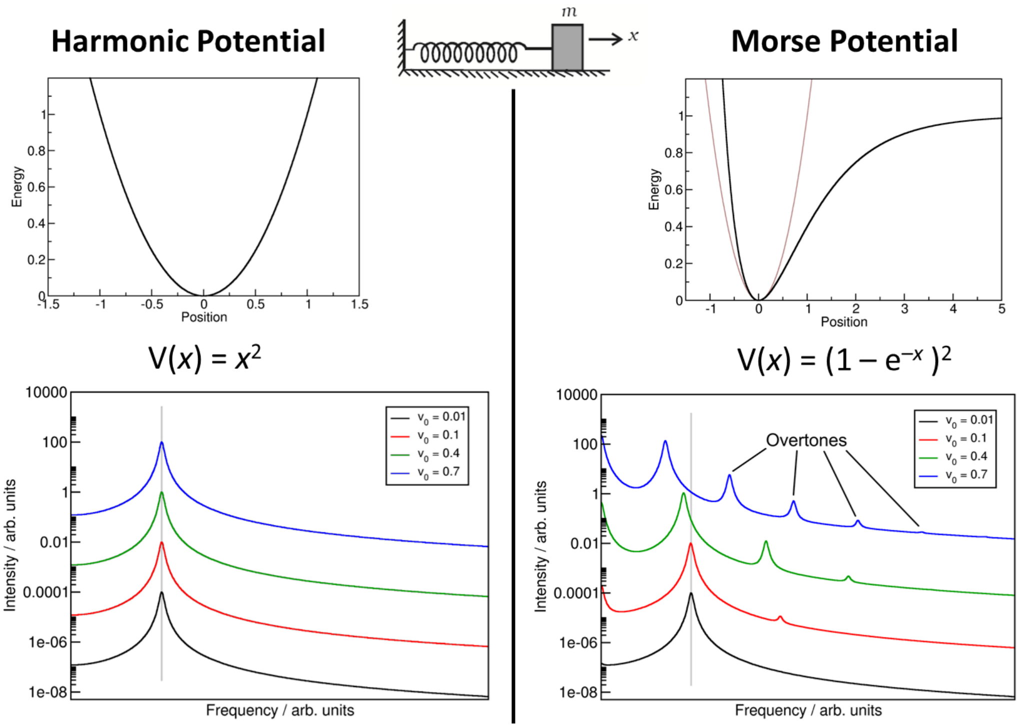


Figure A.1.: One-dimensional oscillator with harmonic potential *(left)* and Morse potential[457,458] *(right)*, together with corresponding power spectra for different temperatures (amplitudes) in the lower part. Morse oscillator shows approximate overtones at higher temperatures.

Publication List of Martin Brehm

56 articles in peer-reviewed journals *(in chronological order)*.

h-index: 26

Total citations: 3048

ORCID: 0000–0002–6861–459X

[MB1] J. Thar, M. Brehm, A. P. Seitsonen, B. Kirchner, “Unexpected Hydrogen Bond Dynamics in Imidazolium-Based Ionic Liquids”, *J. Phys. Chem. B* **2009**, *113 (46)*, 15129–15132, doi:`10.1021/jp908110j`. `(cited on page 1)`

[MB2] M. Brehm, B. Kirchner, “TRAVIS – A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories”, *J. Chem. Inf. Model.* **2011**, *51 (8)*, 2007–2023, doi:`10.1021/ci200217w`. `(cited on pages 16, 21, 51, 73, 77, 81, 83, 96, 112, and 114)`

[MB3] M. Kohagen, M. Brehm, J. Thar, W. Zhao, F. Müller-Plathe, B. Kirchner, “Performance of Quantum Chemically Derived Charges and Persistence of Ion Cages in Ionic Liquids. A Molecular Dynamics Simulations Study of 1-*n*-Butyl-3-Methylimidazolium Bromide”, *J. Phys. Chem. B* **2011**, *115 (4)*, 693–702, doi:`10.1021/jp109612k`.

[MB4] M. Brüssel, M. Brehm, T. Voigt, B. Kirchner, “Ab initio Molecular Dynamics Simulations of a Binary System of Ionic Liquids”, *Phys. Chem. Chem. Phys.* **2011**, *13*, 13617–13620, doi:`10.1039/C1CP21550G`.

[MB5] P. J. di Dio, M. Brehm, B. Kirchner, “Singular Value Decomposition for Analyzing Temperature- and Pressure-Dependent Radial Distribution Functions: Decomposition into Grund RDFs (GRDFs)”, *J. Chem. Theory Comput.* **2011**, *7 (10)*, 3035–3039, doi:`10.1021/ct2003385`.

[MB6] M. Kohagen, M. Brehm, Y. Lingscheid, R. Giernoth, J. Sangoro, F. Kremer, S. Naumov, C. Iacob, J. Kärger, R. Valiullin, B. Kirchner, “How Hydrogen Bonds Influence the Mobility of Imidazolium-Based Ionic Liquids. A Combined Theoretical and Experimental Study of 1-*n*-Butyl-3-Methylimidazolium Bromide”, *J. Phys. Chem. B* **2011**, *115 (51)*, 15280–15288, doi:`10.1021/jp206974h`.

[MB7] A. S. Pensado, M. Brehm, J. Thar, A. P. Seitsonen, B. Kirchner, “Effect of Dispersion on the Structure and Dynamics of the Ionic Liquid 1-Ethyl-3-Methylimidazolium Thiocyanate”, *ChemPhysChem* **2012**, *13 (7)*, 1845–1853, doi:`10.1002/cphc.201100917`.

[MB8] M. Brehm, H. Weber, A. S. Pensado, A. Stark, B. Kirchner, “Proton Transfer and Polarity Changes in Ionic Liquid–Water Mixtures: A Perspective on Hydrogen Bonds from ab initio Molecular Dynamics at the Example of 1-Ethyl-3-Methylimidazolium Acetate–Water Mixtures—Part 1”, *Phys. Chem. Chem. Phys.* **2012**, *14*, 5030–5044, doi:`10.1039/C2CP23983C`. `(cited on pages 9, 45, 46, 107, 108, and 110)`

[MB9] K. Wendler, M. Brehm, F. Malberg, B. Kirchner, L. Delle Site, "Short Time Dynamics of Ionic Liquids in AIMD-Based Power Spectra", *J. Chem. Theory Comput.* **2012**, *8 (5)*, 1570–1579, doi:10.1021/ct300152t.

[MB10] M. Brüssel, M. Brehm, A. S. Pensado, F. Malberg, M. Ramzan, A. Stark, B. Kirchner, "On the Ideality of Binary Mixtures of Ionic Liquids", *Phys. Chem. Chem. Phys.* **2012**, *14*, 13204–13215, doi:10.1039/C2CP41926B.

[MB11] M. Brüssel, E. Perlt, M. von Domaros, M. Brehm, B. Kirchner, "A One-Parameter Quantum Cluster Equilibrium Approach", *J. Chem. Phys.* **2012**, *137*, 164107, doi:10.1063/1.4759154.

[MB12] O. Hollóczki, D. S. Firaha, J. Friedrich, M. Brehm, R. Cybik, M. Wild, A. Stark, B. Kirchner, "Carbene Formation in Ionic Liquids: Spontaneous, Induced, or Prohibited?", *J. Phys. Chem. B* **2013**, *117 (19)*, 5898–5907, doi:10.1021/jp4004399.

[MB13] M. Brehm, H. Weber, A. S. Pensado, A. Stark, B. Kirchner, "Liquid Structure and Cluster Formation in Ionic Liquid / Water Mixtures – An Extensive ab initio Molecular Dynamics Study on 1-Ethyl-3-Methylimidazolium Acetate/Water Mixtures – Part 2", *Z. Phys. Chem.* **2013**, *227*, 177–203, doi:10.1524/zpch.2012.0327. (cited on pages 9, 45, 46, 107, and 110)

[MB14] M. Thomas, M. Brehm, R. Fligg, P. Vöhringer, B. Kirchner, "Computing Vibrational Spectra from ab initio Molecular Dynamics", *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608–6622, doi:10.1039/C3CP44302G. (cited on pages 8, 15, 50, 52, 54, 55, 56, 72, 74, 84, 85, and 94)

[MB15] F. Malberg, M. Brehm, O. Hollóczki, A. S. Pensado, B. Kirchner, "Understanding the Evaporation of Ionic Liquids using the Example of 1-Ethyl-3-Methylimidazolium Ethylsulfate", *Phys. Chem. Chem. Phys.* **2013**, *15*, 18424–18436, doi:10.1039/C3CP52966E.

[MB16] M. Thomas, M. Brehm, O. Hollóczki, B. Kirchner, "How Can a Carbene be Active in an Ionic Liquid?", *Chem. Eur. J* **2014**, *20 (6)*, 1622–1629, doi:10.1002/chem.201303329.

[MB17] S. Zahn, M. Brehm, M. Brüssel, O. Hollóczki, M. Kohagen, S. B. C. Lehmann, F. Malberg, A. S. Pensado, M. Schöppke, H. Weber, B. Kirchner, "Understanding Ionic Liquids from Theoretical Methods", *J. Mol. Liq.* **2014**, *192*, 71–76, doi:10.1016/j.molliq.2013.08.015.

[MB18] R. Giernoth, A. Bröhl, M. Brehm, Y. Lingscheid, "Interactions in Ionic Liquids probed by in situ NMR Spectroscopy", *J. Mol. Liq.* **2014**, *192*, 55–58, doi:10.1016/j.molliq.2013.07.010.

[MB19] M. Thomas, M. Brehm, O. Hollóczki, Z. Kelemen, L. Nyulászi, T. Pasinszki, B. Kirchner, "Simulating the Vibrational Spectra of Ionic Liquid Systems: 1-Ethyl-3-Methylimidazolium Acetate and its Mixtures", *J. Chem. Phys.* **2014**, *141*, 024510, doi:10.1063/1.4887082. (cited on pages 8, 45, 46, 50, 51, 52, 55, 74, 94, 107, 108, and 110)

[MB20] A. Stark, M. Brehm, M. Brüssel, S. B. C. Lehmann, A. S. Pensado, M. Schöppke, B. Kirchner, "A Theoretical and Experimental Chemist's Joint View on Hydrogen Bonding in Ionic Liquids and Their Binary Mixtures", *Top. Curr. Chem.* **2014**, *351*, 149–187, doi:`10.1007/128_2013_485`.

[MB21] O. Hollóczki, M. Macchiagodena, H. Weber, M. Thomas, M. Brehm, A. Stark, O. Russina, A. Triolo, B. Kirchner, "Triphilic Ionic-Liquid Mixtures: Fluorinated and Non-Fluorinated Aprotic Ionic-Liquid Mixtures", *ChemPhysChem* **2015**, *16 (15)*, 3325–3333, doi:`10.1002/cphc.201500473`.

[MB22] M. Brehm, H. Weber, M. Thomas, O. Hollóczki, B. Kirchner, "Domain Analysis in Nanostructured Liquids: A Post-Molecular Dynamics Study at the Example of Ionic Liquids", *ChemPhysChem* **2015**, *16 (15)*, 3271–3277, doi:`10.1002/cphc.201500471`.

[MB23] M. Thomas, M. Brehm, B. Kirchner, "Voronoi Dipole Moments for the Simulation of Bulk Phase Vibrational Spectra", *Phys. Chem. Chem. Phys.* **2015**, *17*, 3207–3213, doi:`10.1039/C4CP05272B`. `(cited on pages 8, 17, 20, 21, 23, 24, 26, 27, 28, 52, 53, 55, 61, 63, 68, 70, 92, 94, 112, and 113)`

[MB24] M. Cooper, A. Wagner, D. Wondrousch, F. Sonntag, A. Sonnabend, M. Brehm, G. Schüürmann, L. Adrian, "Anaerobic Microbial Transformation of Halogenated Aromatics and Fate Prediction Using Electron Density Modeling", *Environ. Sci. Technol.* **2015**, *49 (10)*, 6018–6028, doi:`10.1021/acs.est.5b00303`.

[MB25] C. Slawik, C. Rickmeyer, M. Brehm, A. Böhme, G. Schüürmann, "Glutathione Adduct Patterns of Michael-Acceptor Carbonyls", *Environ. Sci. Technol.* **2017**, *51 (7)*, 4018–4026, doi:`10.1021/acs.est.6b04981`.

[MB26] M. Brehm, M. Thomas, "Computing Bulk Phase Raman Optical Activity Spectra from ab initio Molecular Dynamics Simulations", *J. Phys. Chem. Lett.* **2017**, *8 (14)*, 3409–3414, doi:`10.1021/acs.jpclett.7b01616`. `(cited on pages 8, 20, 23, 55, 56, 62, 63, 64, 86, 87, 93, and 113)`

[MB27] M. Brehm, G. Saddiq, T. Watermann, D. Sebastiani, "Influence of Small Fluorophilic and Lipophilic Organic Molecules on Dipalmitoylphosphatidylcholine Bilayers", *J. Phys. Chem. B* **2017**, *121 (35)*, 8311–8321, doi:`10.1021/acs.jpcb.7b06520`.

[MB28] C. Peschel, M. Brehm, D. Sebastiani, "Polyphilic Interactions as Structural Driving Force Investigated by Molecular Dynamics Simulation (Project 7)", *Polymers* **2017**, *9 (9)*, 445, doi:`10.3390/polym9090445`.

[MB29] M. Brehm, A. Kafka, M. Bamler, R. Kühne, G. Schüürmann, L. Sikk, J. Burk, P. Burk, T. Tamm, K. Tämm, S. Pokhrel, L. Mädler, A. Kahru, V. Aruoja, M. Sihtmäe, J. Scott-Fordsmand, P. B. Sorensen, L. Escorihuela, C. P. Roca, A. Fernández, F. Giralt, R. Rallo, "An Integrated Data-Driven Strategy for Safe-by-Design Nanoparticles: The FP7 MODERN Project.", *Adv. Exp. Med. Biol.* **2017**, *947*, 257–301, doi:`10.1007/978-3-319-47754-1_9`.

[MB30] S. Gehrke, M. von Domaros, R. Clark, O. Hollóczki, M. Brehm, T. Welton, A. Luzar, B. Kirchner, "Structure and Lifetimes in Ionic Liquids and their Mixtures", *Faraday Discuss.* **2018**, *206*, 219–245, doi:`10.1039/C7FD00166E`.

[MB31] M. Brehm, D. Sebastiani, “Simulating Structure and Dynamics in Small Droplets of 1-Ethyl-3-Methylimidazolium Acetate”, *J. Chem. Phys.* **2018**, *148*, 193802, doi:`10.1063/1.5010342`. `(cited on pages 45, 46, 107, and 110)`

[MB32] S. Pylaeva, M. Brehm, D. Sebastiani, “Salt Bridge in Aqueous Solution: Strong Structural Motifs but Weak Enthalpic Effects”, *Sci. Rep.* **2018**, *8*, 13626, doi:`10.1038/s41598-018-31935-z`.

[MB33] U. Cerajewski, J. Träger, S. Henkel, A. H. Roos, M. Brehm, D. Hinderberger, “Nanoscopic Structures and Molecular Interactions Leading to a Dystectic and two Eutectic Points in [EMIm][Cl] / Urea Mixtures”, *Phys. Chem. Chem. Phys.* **2018**, *20*, 29591–29600, doi:`10.1039/C8CP04912B`.

[MB34] M. Brehm, M. Thomas, “An Efficient Lossless Compression Algorithm for Trajectories of Atom Positions and Volumetric Data”, *J. Chem. Inf. Model.* **2018**, *58 (10)*, 2092–2107, doi:`10.1021/acs.jcim.8b00501`. `(cited on pages 30, 32, 35, 36, 43, 45, 47, 70, 104, 112, and 113)`

[MB35] M. Brehm, M. Pulst, J. Kressler, D. Sebastiani, “Triazolium-Based Ionic Liquids – A Novel Class of Cellulose Solvents”, *J. Phys. Chem. B* **2019**, *123 (18)*, 3994–4003, doi:`10.1021/acs.jpcb.8b12082`.

[MB36] M. Brehm, M. Thomas, “Computing Bulk Phase Resonance Raman Spectra from ab initio Molecular Dynamics and Real-Time TDDFT”, *J. Chem. Theory Comput.* **2019**, *15 (7)*, 3901–3905, doi:`10.1021/acs.jctc.9b00512`. `(cited on pages 8, 20, 23, 55, 69, 70, 71, 72, 100, 101, 102, 103, 105, 106, and 113)`

[MB37] L. Scarbath-Evers, R. Hammer, D. Golze, M. Brehm, D. Sebastiani, W. Widdra, “From Flat to Tilted: Gradual Interfaces in Organic Thin Film Growth”, *Nanoscale* **2020**, *12*, 3834–3845, doi:`10.1039/C9NR06592J`.

[MB38] C. Dreßler, G. Kabbe, M. Brehm, D. Sebastiani, “Dynamical Matrix Propagator Scheme for Large-Scale Proton Dynamics Simulations”, *J. Chem. Phys.* **2020**, *152 (11)*, 114114, doi:`10.1063/1.5140635`.

[MB39] C. Dreßler, G. Kabbe, M. Brehm, D. Sebastiani, “Exploring Non-Equilibrium Molecular Dynamics of Mobile Protons in the Solid Acid CsH_2PO_4 on the Micrometer and Microsecond Scale”, *J. Chem. Phys.* **2020**, *152 (16)*, 164110, doi:`10.1063/5.0002167`.

[MB40] M. Brehm, M. Thomas, S. Gehrke, B. Kirchner, “TRAVIS – A Free Analyzer for Trajectories from Molecular Simulation”, *J. Chem. Phys.* **2020**, *152 (16)*, 164105, doi:`10.1063/5.0005078`. `(cited on pages 16, 17, 21, 50, 51, 63, 64, 73, 77, 79, 81, 82, 83, 96, 102, 108, 112, and 114)`

[MB41] M. Brehm, J. Radicke, M. Pulst, F. Shaabani, D. Sebastiani, J. Kressler, “Dissolving Cellulose in 1,2,3-Triazolium- and Imidazolium-Based Ionic Liquids with Aromatic Anions”, *Molecules* **2020**, *25 (15)*, 3539, doi:`10.3390/molecules25153539`.

[MB42] J. Hunold, J. Eisermann, M. Brehm, D. Hinderberger, "Characterization of Aqueous Lower Polarity Solvation Shells Around Amphiphilic TEMPO Radicals in Water", *J. Phys. Chem. B* **2020**, *124 (39)*, 8601–8609, doi:10.1021/acs.jpcb.0c04863.

[MB43] M. Weiß, M. Brehm, "Exploring Free Energy Profiles of Enantioselective Organocatalytic Aldol Reactions under Full Solvent Influence", *Molecules* **2020**, *25 (24)*, 5861, doi:10.3390/molecules25245861.

[MB44] M. Mukherjee, D. Tripathi, M. Brehm, C. Riplinger, A. K. Dutta, "Efficient EOM-CC-Based Protocol for the Calculation of Electron Affinity of Solvated Nucleobases: Uracil as a Case Study", *J. Chem. Theory Comput.* **2021**, *17 (1)*, 105–116, doi:10.1021/acs.jctc.0c00655.

[MB45] E. Roos, M. Brehm, "A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 1: [EMIm][OAc] / Water Mixtures", *Phys. Chem. Chem. Phys.* **2021**, *23*, 1242–1253, doi:10.1039/D0CP04537C.

[MB46] A. Triolo, F. L. Celso, M. Brehm, V. D. Lisio, O. Russina, "Liquid Structure of a Choline Chloride-Water Natural Deep Eutectic Solvent: A Molecular Dynamics Characterization", *J. Mol. Liq.* **2021**, *331*, 115750, doi:10.1016/j.molliq.2021.115750.

[MB47] M.-A. Codescu, M. Weiß, M. Brehm, O. Kornilov, D. Sebastiani, E. T. J. Nibbering, "Switching Between Proton Vacancy and Excess Proton Transfer Pathways in the Reaction Between 7-Hydroxyquinoline and Formate", *J. Phys. Chem. A* **2021**, *125 (9)*, 1845–1859, doi:10.1021/acs.jpca.0c10191.

[MB48] M. Brehm, M. Thomas, "Optimized Atomic Partial Charges and Radii Defined by Radical Voronoi Tessellation of Bulk Phase Simulations", *Molecules* **2021**, *26 (7)*, 1875, doi:10.3390/molecules26071875. `(cited on pages 17, 23, 61, 63, 68, 70, and 112)`

[MB49] S. Roy, M. Brehm, S. Sharma, F. Wu, D. Maltsev, P. Halstenberg, L. C. Gallington, S. M. Mahurin, S. Dai, A. S. Ivanov, C. J. Margulis, V. Bryantsev, "Unraveling Local Structure of Molten Salts via X-Ray Scattering, Raman Spectroscopy, and ab initio Molecular Dynamics", *J. Phys. Chem. B* **2021**, *125 (22)*, 5971–5982, doi:10.1021/acs.jpcb.1c03786. `(cited on pages 8, 55, 89, and 90)`

[MB50] A. Triolo, M. E. D. Pietro, A. Mele, F. L. Celso, M. Brehm, V. D. Lisio, A. Martinelli, P. Chater, O. Russina, "Liquid Structure and Dynamics in the Choline Acetate:Urea 1:2 Deep Eutectic Solvent", *J. Chem. Phys.* **2021**, *154*, 244501, doi:10.1063/5.0054048.

[MB51] Y. Yang, J. Cheramy, M. Brehm, Y. Xu, "Raman Optical Activity of *N*-Acetyl-L-Cysteine in Water and in Methanol: The »Clusters-in-a-Liquid« Model and ab initio Molecular Dynamics Simulations", *ChemPhysChem* **2022**, *23* (11), e202200161, doi:10.1002/cphc.202200161R1. `(cited on pages 20, 95, 97, and 98)`

[MB52] S. Taherivardanjani, J. Blasius, M. Brehm, R. Dötzer, B. Kirchner, "Conformer Weighting and Differently Sized Cluster Weighting for Nicotine and its Phosphorus Derivatives", *J. Phys. Chem. A* **2022**, *126 (40)*, 7070–7083, doi:`10.1021/acs.jpca.2c03133`.

[MB53] T. Frömbgen, J. Blasius, V. Alizadeh, A. Chaumont, M. Brehm, B. Kirchner, "Cluster Analysis in Liquids: A Novel Tool in TRAVIS", *J. Chem. Inf. Model.* **2022**, *62 (22)*, 5634–5644, doi:`10.1021/acs.jcim.2c01244`.

[MB54] R. Chahal, S. Roy, M. Brehm, S. Banerjee, V. Bryantsev, S. Lam, "Transferable Deep Learning Potential Reveals Intermediate-Range Ordering Effects in LiF–NaF–ZrF_4 Molten Salt", *JACS Au* **2022**, *2 (12)*, 2693–2702, doi:`10.1021/jacsau.2c00526`.

[MB55] J. Radicke, E. Roos, D. Sebastiani, M. Brehm, J. Kressler, "Lactate-Based Ionic Liquids as Chiral Solvents for Cellulose", *J. Polym. Sci.* **2023**, *61 (5)*, 372–384, doi:`10.1002/pol.20220687`.

[MB56] E. Roos, D. Sebastiani, M. Brehm, "A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 2: Cellulose in [EMIm][OAc] / Water Mixtures", *Phys. Chem. Chem. Phys.* **2023**, *accepted*, doi:`10.1039/D2CP05636D`.

Bibliography

[1] A. B. S. Elliott, R. Horvath, K. C. Gordon, "Vibrational Spectroscopy as a Probe of Molecule-Based Devices", *Chem. Soc. Rev.* **2012**, *41* (5), 1929–1946, doi:10.1039/C1CS15208D. `(cited on pages 1 and 2)`

[2] C. K. Muro, K. C. Doty, J. Bueno, L. Halámková, I. K. Lednev, "Vibrational Spectroscopy: Recent Developments to Revolutionize Forensic Science", *Anal. Chem.* **2014**, *87* (1), 306–327, doi:10.1021/ac504068a. `(cited on page 2)`

[3] M. López-López, C. García-Ruiz, "Infrared and Raman Spectroscopy Techniques Applied to Identification of Explosives", *TrAC Trends in Analytical Chemistry* **2014**, *54*, 36–44, doi:10.1016/j.trac.2013.10.011. `(cited on page 2)`

[4] T. D. Fridgen, "Infrared Consequence Spectroscopy of Gaseous Protonated and Metal Ion Cationized Complexes", *Mass Spectrom. Rev.* **2009**, *28* (4), 586–607, doi:10.1002/mas.20224. `(cited on page 2)`

[5] D. Cialla, A. März, R. Böhme, F. Theil, K. Weber, M. Schmitt, J. Popp, "Surface-Enhanced Raman Spectroscopy (SERS): Progress and Trends", *Anal. Bioanal. Chem.* **2011**, *403* (1), 27–54, doi:10.1007/s00216-011-5631-x. `(cited on page 2)`

[6] M. Manley, "Near-Infrared Spectroscopy and Hyperspectral Imaging: Non-Destructive Analysis of Biological Materials", *Chem. Soc. Rev.* **2014**, *43* (24), 8200–8214, doi:10.1039/C4CS00062E. `(cited on page 2)`

[7] S. Pahlow, A. März, B. Seise, K. Hartmann, I. Freitag, E. Kämmer, R. Böhme, V. Deckert, K. Weber, D. Cialla, J. Popp, "Bioanalytical Application of Surface- and Tip-Enhanced Raman Spectroscopy", *Eng. Life Sci.* **2012**, *12* (2), 131–143, doi:10.1002/elsc.201100056. `(cited on page 2)`

[8] T. Deckert-Gaudig, V. Deckert, "Nanoscale Structural Analysis Using Tip-Enhanced Raman Spectroscopy", *Curr. Opin. Chem. Biol.* **2011**, *15* (5), 719–724, doi:10.1016/j.cbpa.2011.06.020. `(cited on page 2)`

[9] C. Krafft, J. Popp, "The Many Facets of Raman Spectroscopy for Biomedical Analysis", *Anal. Bioanal. Chem.* **2014**, *407* (3), 699–717, doi:10.1007/s00216-014-8311-9. `(cited on page 2)`

[10] M. Olschewski, S. Knop, J. Lindner, P. Vöhringer, "From Single Hydrogen Bonds to Extended Hydrogen-Bond Wires: Low-Dimensional Model Systems for Vibrational Spectroscopy of Associated Liquids", *Angew. Chem. - Int. Ed.* **2013**, *52* (37), 9634–9654, doi:10.1002/anie.201210009. `(cited on page 2)`

[11] J. F. W. Herschel, "I. On the Chemical Action of the Rays of the Solar Spectrum on Preparations of Silver and Other Substances, Both Metallic and Non-Metallic, and on Some Photographic Processes", *Philos. Trans. R. Soc. Lond.* **1840**, *130*, 1–59, doi:10.1098/rstl.1840.0002. `(cited on page 2)`

[12] A. Smekal, "Zur Quantentheorie der Dispersion", *Naturwissenschaften* **1923**, *11* (43), 873–875, doi:10.1007/BF01576902. `(cited on page 2)`

[13] C. V. Raman, K. S. Krishnan, "A New Type of Secondary Radiation", *Nature* **1928**, *121*, 501–502, doi:10.1038/121501c0. `(cited on page 2)`

[14] K. S. Krishnan, "The Raman Effect in Crystals", *Nature* **1928**, *122*, 477–478, doi:10.1038/122477a0. (cited on page 2)

[15] A. L. Schawlow, C. H. Townes, "Infrared and Optical Masers", *Phys. Rev.* **1958**, *112*, 1940–1949, doi:10.1103/PhysRev.112.1940. (cited on page 2)

[16] T. H. Maiman, "Stimulated Optical Radiation in Ruby", *Nature* **1960**, *187*, 493–494, doi:10.1016/b978-0-08-013320-1.50016-6. (cited on page 2)

[17] D. G. Rea, "On the Theory of the Resonance Raman Effect", *J. Mol. Spectrosc.* **1960**, *4* (6), 499–506, doi:10.1016/0022-2852(60)90113-2. (cited on pages 2 and 69)

[18] L. N. Ovander, "Resonance Raman Scattering in Crystals", *Sov. Phys. – Sol. State* **1962**, *4* (6), 1081–1083. (cited on pages 2 and 69)

[19] R. Loudon, "Theory of Resonance Raman Effect in Crystals", *J. Phys.* **1965**, *26* (11), 677–683, doi:10.1051/jphys:019650026011067700. (cited on pages 2 and 69)

[20] T. M. Ivanova, L. A. Yanovska, P. P. Shorygin, "Resonance Raman Spectra of Polyene Compounds with Structured Absorption Bands", *Opt. Spectrosc.* **1965**, *18* (2), 115. (cited on pages 2 and 69)

[21] M. Rokni, S. Yatsiv, "Resonance Raman Effect in Free Atoms of Potassium", *Phys. Lett. A* **1967**, *24* (5), 277–278, doi:10.1016/0375-9601(67)90435-5. (cited on pages 2 and 69)

[22] N. A. Owens, A. Pinter, M. D. Porter, "Surface-Enhanced Resonance Raman Scattering for the Sensitive Detection of a Tuberculosis Biomarker in Human Serum", *J. Raman Spectrosc.* **2019**, *50* (1), 15–25, doi:10.1002/jrs.5500. (cited on pages 2 and 69)

[23] C. Li, T. Ahmed, M. Ma, T. Edvinsson, J. Zhu, "A Facile Approach to ZnO / CdS Nanoarrays and Their Photocatalytic and Photoelectrochemical Properties", *Appl. Catal. B* **2013**, *138-139*, 175–183, doi:10.1016/j.apcatb.2013.02.042. (cited on pages 2 and 69)

[24] S. M. Derayea, H. Tsujino, Y. Oyama, Y. Ishikawa, T. Yamashita, T. Uno, "Investigation on Drug-Binding in Heme Pocket of Cyp2C19 with UV-Visible and Resonance Raman Spectroscopies", *Spectrochim. Acta A* **2019**, *209*, 209–216, doi:10.1016/j.saa.2018.10.045. (cited on pages 2 and 69)

[25] Y. Gao, T. P. Martin, A. K. Thomas, J. K. Grey, "Resonance Raman Spectroscopic- and Photocurrent Imaging of Polythiophene / Fullerene Solar Cells", *J. Phys. Chem. Lett.* **2010**, *1* (1), 178–182, doi:10.1021/jz900038c. (cited on pages 2 and 69)

[26] L. Zedler, J. Guthmuller, I. Rabelo de Moraes, S. Kupfer, S. Krieck, M. Schmitt, J. Popp, S. Rau, B. Dietzek, "Resonance-Raman Spectro-Electrochemistry of Intermediates in Molecular Artificial Photosynthesis of Bimetallic Complexes", *Chem. Commun.* **2014**, *50*, 5227–5229, doi:10.1039/C3CC47487A. (cited on pages 2 and 69)

[27] A. Faust, Y. Amit, U. Banin, "Phonon-Plasmon Coupling and Active Cu Dopants in Indium Arsenide Nanocrystals Studied by Resonance Raman Spectroscopy", *J. Phys. Chem. Lett.* **2017**, *8* (11), 2519–2525, doi:10.1021/acs.jpclett.7b00661. (cited on pages 2 and 69)

[28] D. Friebel, M. W. Louie, M. Bajdich, K. E. Sanwald, Y. Cai, A. M. Wise, M.-J. Cheng, D. Sokaras, T.-C. Weng, R. Alonso-Mori, R. C. Davis, J. R. Bargar, J. K. Nørskov, A. Nilsson, A. T. Bell, "Identification of Highly Active Fe Sites in (Ni,Fe)OOH for Electrocatalytic Water Splitting", *J. Am. Chem. Soc.* **2015**, *137* (3), 1305–1313, doi:10.1021/ja511559d. (cited on pages 2 and 69)

[29] P. J. Stephens, "Theory of Vibrational Circular Dichroism", *J. Phys. Chem.* **1985**, *89*, 748–752, doi:10.1021/j100251a006. (cited on pages 2, 6, and 60)

[30] K. Nakanishi, N. Berova, R. W. Woody, *Circular Dichroism: Principles and Applications*, Wiley-VCH: Weinheim, Germany, **2000**, ISBN 978-0-471-33003-5, doi:10.1021/jm010040t. (cited on page 2)

[31] E. Castiglioni, P. Biscarini, S. Abbate, "Experimental Aspects of Solid State Circular Dichroism", *Chirality* **2009**, *21* (1E), E28–E36, doi:10.1002/chir.20770. (cited on page 2)

[32] S. Abbate, G. Longhi, F. Lebon, E. Castiglioni, S. Superchi, L. Pisani, F. Fontana, F. Torricelli, T. Caronna, C. Villani, R. Sabia, M. Tommasini, A. Lucotti, D. Mendola, A. Mele, D. A. Lightner, "Helical Sense-Responsive and Substituent-Sensitive Features in Vibrational and Electronic Circular Dichroism, in Circularly Polarized Luminescence, and in Raman Spectra of Some Simple Optically Active Hexahelicenes", *J. Phys. Chem. C* **2014**, *118* (3), 1682–1695, doi:10.1021/jp4105695. (cited on page 2)

[33] P. W. Atkins, L. D. Barron, "Rayleigh Scattering of Polarized Photons by Molecules", *Mol. Phys.* **1969**, *16*, 453–466, doi:10.1080/00268976900100501. (cited on page 3)

[34] L. D. Barron, A. D. Buckingham, "Rayleigh and Raman Scattering from Optically Active Molecules", *Mol. Phys.* **1971**, *20*, 1111–1119, doi:10.1080/00268977100101091. (cited on page 3)

[35] L. D. Barron, M. P. Bogaard, A. D. Buckingham, "Raman Scattering of Circularly Polarized Light by Optically Active Molecules", *J. Am. Chem. Soc.* **1973**, *95*, 603–605, doi:10.1021/ja00783a058. (cited on page 3)

[36] W. Hug, S. Kint, G. Bailey, J. R. Scherer, "Raman Circular Intensity Differential Spectroscopy. Spectra of (−)-α-Pinene and (+)-α-Phenylethylamine", *J. Am. Chem. Soc.* **1975**, *97*, 5589–5590, doi:10.1021/ja00852a049. (cited on page 3)

[37] X. D. Zhu, H. Suhr, Y. R. Shen, "Surface Vibrational Spectroscopy by Infrared-Visible Sum Frequency Generation", *Phys. Rev. B* **1987**, *35* (6), 3047–3050, doi:10.1103/physrevb.35.3047. (cited on pages 3 and 67)

[38] J. H. Hunt, P. Guyot-Sionnest, Y. R. Shen, "Observation of C–H Stretch Vibrations of Monolayers of Molecules Optical Sum-Frequency Generation", *Chem. Phys. Lett.* **1987**, *133* (3), 189–192, doi:10.1016/0009-2614(87)87049-5. (cited on pages 3 and 67)

[39] P. Guyot-Sionnest, P. Dumas, Y. J. Chabal, G. S. Higashi, "Lifetime of an Adsorbate–Substrate Vibration: H on Si(111)", *Phys. Rev. Lett.* **1990**, *64* (18), 2156–2159, doi:10.1103/physrevlett.64.2156. (cited on pages 3 and 67)

[40] P. Guyot-Sionnest, "Coherent Processes at Surfaces: Free-Induction Decay and Photon Echo of the Si–H Stretching Vibration for H/Si(111)", *Phys. Rev. Lett.* **1991**, *66* (11), 1489–1492, doi:10.1103/physrevlett.66.1489. (cited on pages 3 and 67)

[41] P. Guyot-Sionnest, "Two-Phonon Bound State for the Hydrogen Vibration on the H/Si(111) Surface", *Phys. Rev. Lett.* **1991**, *67* (17), 2323–2326, doi:10.1103/physrevlett.67.2323. (cited on pages 3 and 67)

[42] Y. R. Shen, "Surface Properties Probed by Second-Harmonic and Sum-Frequency Generation", *Nature* **1989**, *337* (6207), 519–525, doi:10.1038/337519a0. (cited on pages 3 and 67)

[43] J. Neugebauer, M. Reiher, C. Kind, B. A. Hess, "Quantum Chemical Calculation of Vibrational Spectra of Large Molecules—Raman and IR Spectra for Buckminsterfullerene", *J. Comput. Chem.* **2002**, *23* (9), 895–910, doi:10.1002/jcc.10089. (cited on pages 6, 52, and 54)

[44] F. Pawłowski, A. Halkier, P. Jørgensen, K. L. Bak, T. Helgaker, W. Klopper, "Accuracy of Spectroscopic Constants of Diatomic Molecules from ab initio Calculations", *J. Chem. Phys.* **2003**, *118* (6), 2539–2549, doi:10.1063/1.1533032. (cited on pages 6 and 52)

[45] P. Daněček, J. Kapitán, V. Baumruk, L. Bednárová, V. Kopecký Jr., P. Bouř, "Anharmonic Effects in IR, Raman, and Raman Optical Activity Spectra of Alanine and Proline Zwitterions", *J. Chem. Phys.* **2007**, *126* (22), 224513, doi:10.1063/1.2738065. (cited on pages 6, 7, 52, 54, and 62)

[46] R. I. Ovsyannikov, W. Thiel, S. N. Yurchenko, M. Carvajal, P. Jensen, "PH_3 Revisited: Theoretical Transition Moments for the Vibrational Transitions Below", *J. Mol. Spectrosc.* **2008**, *252* (2), 121–128, doi:10.1016/j.jms.2008.07.005. (cited on pages 6 and 52)

[47] C. R. Jacob, M. Reiher, "Localizing Normal Modes in Large Molecules", *J. Chem. Phys.* **2009**, *130* (8), 084106, doi:10.1063/1.3077690. (cited on pages 6, 52, and 54)

[48] F. Neese, "Prediction of Molecular Properties and Molecular Spectroscopy with Density Functional Theory: From Fundamental Theory to Exchange-Coupling", *Coord. Chem. Rev.* **2009**, *253* (5-6), 526–563, doi:10.1016/j.ccr.2008.05.014. (cited on pages 6, 52, 54, and 60)

[49] T. Weymuth, M. P. Haag, K. Kiewisch, S. Luber, S. Schenk, C. R. Jacob, C. Herrmann, J. Neugebauer, M. Reiher, "MOVIPAC: Vibrational Spectroscopy with a Robust Meta-Program for Massively Parallel Standard and Inverse Calculations", *J. Comput. Chem.* **2012**, *33* (27), 2186–2198, doi:10.1002/jcc.23036. (cited on pages 6, 7, 52, 54, and 62)

[50] J. Bloino, M. Biczysko, V. Barone, "Anharmonic Effects on Vibrational Spectra Intensities: Infrared, Raman, Vibrational Circular Dichroism, and Raman Optical Activity", *J. Phys. Chem. A* **2015**, *119* (49), 11862–11874, doi:10.1021/acs.jpca.5b10067. (cited on pages 6, 7, 52, 54, 60, 62, and 93)

[51] R. R. Oliveira, G. Molpeceres, F. Fantuzzi, H. M. Quitián-Lara, H. M. Boechat-Roberty, J. Kästner, "Gas-Phase Spectroscopic Characterization of Neutral and Ionic Polycyclic Aromatic Phosphorus Heterocycles (PAPHs)", *Mon. Not. R. Astron Soc.* **2020**, *500* (2), 2564–2576, doi:10.1093/mnras/staa3460. (cited on pages 6 and 52)

[52] L. Böselt, R. Dötzer, S. Steiner, M. Stritzinger, S. Salzmann, S. Riniker, "Determining the Regiochemistry and Relative Stereochemistry of Small and Druglike Molecules Using an Alignment Algorithm for Infrared Spectra", *Anal. Chem.* **2020**, *92* (13), 9124–9131, doi:10.1021/acs.analchem.0c01399. (cited on pages 6 and 52)

[53] M. Pecul, A. Rizzo, J. Leszczynski, "Vibrational Raman and Raman Optical Activity Spectra of D-Lactic Acid, D-Lactate, and D-Glyceraldehyde: Ab initio Calculations", *J. Phys. Chem. A* **2002**, *106* (46), 11008–11016, doi:10.1021/jp021030m. (cited on pages 6, 54, and 62)

[54] C. Herrmann, M. Reiher, "First-Principles Approach to Vibrational Spectroscopy of Biomolecules", in "Atomistic Approaches in Modern Biology", pages 85–132, Springer Berlin Heidelberg, **2007**, doi:10.1007/128_2006_082. (cited on page 6)

[55] M. A. Mroginski, F. Mark, W. Thiel, P. Hildebrandt, "Quantum Mechanics / Molecular Mechanics Calculation of the Raman Spectra of the Phycocyanobilin Chromophore in α-C-Phycocyanin", *Biophys. J.* **2007**, *93*, 1885–1894, doi:10.1529/biophysj.107.108878. (cited on pages 6 and 54)

[56] J. J. Lee, S. Höfener, W. Klopper, T. N. Wassermann, M. A. Suhm, "Origin of the Argon Nanocoating Shift in the OH Stretching Fundamental of *n*-Propanol: A Combined Experimental and Quantum Chemical Study", *J. Phys. Chem. C* **2009**, *113* (25), 10929–10938, doi:10.1021/jp902194h. (cited on pages 6 and 54)

[57] S. Luber, M. Reiher, "Intensity-Carrying Modes in Raman and Raman Optical Activity Spectroscopy", *ChemPhysChem* **2009**, *10* (12), 2049–2057, doi:10.1002/cphc.200900255. (cited on pages 6, 54, and 62)

[58] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, "Gaussian 16 Revision C.01", **2016**. (cited on pages 6, 52, 60, and 62)

[59] L. A. Nafie, T. B. Freedman, "Vibronic Coupling Theory of Infrared Vibrational Transitions", *J. Chem. Phys.* **1983**, *78* (12), 7108–7116, doi:10.1063/1.444741. (cited on pages 6 and 60)

[60] L. A. Nafie, "Adiabatic Molecular Properties beyond the Born–Oppenheimer Approximation. Complete Adiabatic Wave Functions and Vibrationally Induced Electronic Current Density", *J. Chem. Phys.* **1983**, *79* (10), 4950–4957, doi:10.1063/1.445588. (cited on pages 6 and 60)

[61] P. J. Stephens, "Gauge Dependence of Vibrational Magnetic Dipole Transition Moments and Rotational Strengths", *J. Phys. Chem.* **1987**, *91* (7), 1712–1715, doi:10.1021/j100291a009. (cited on pages 6 and 60)

[62] A. D. Buckingham, P. W. Fowler, P. A. Galwas, "Velocity-Dependent Property Surfaces and the Theory of Vibrational Circular Dichroism", *Chem. Phys.* **1987**, *112* (1), 1–14, doi:10.1016/0301-0104(87)85017-6. (cited on pages 6, 57, and 60)

[63] J. R. Cheeseman, M. J. Frisch, F. J. Devlin, P. J. Stephens, "Ab initio Calculation of Atomic Axial Tensors and Vibrational Rotational Strengths Using Density Functional Theory", *Chem. Phys. Lett.* **1996**, *252* (3), 211–220, doi:10.1016/0009-2614(96)00154-6. (cited on pages 6, 60, and 62)

[64] V. P. Nicu, J. Neugebauer, S. K. Wolff, E. J. Baerends, "A Vibrational Circular Dichroism Implementation within a Slater-Type-Orbital based Density Functional Framework and its Application to Hexa- and Hepta-Helicenes", *Theor. Chem. Acc.* **2008**, *119* (1), 245–263, doi:10.1007/s00214-006-0234-x. (cited on pages 6 and 60)

[65] A. Scherrer, R. Vuilleumier, D. Sebastiani, "Nuclear Velocity Perturbation Theory of Vibrational Circular Dichroism", *J. Chem. Theory Comput.* **2013**, *9* (12), 5305–5312, doi:10.1021/ct400700c. (cited on pages 6 and 60)

[66] A. Scherrer, F. Agostini, D. Sebastiani, E. K. U. Gross, R. Vuilleumier, "Nuclear Velocity Perturbation Theory for Vibrational Circular Dichroism: An Approach Based on the Exact Factorization of the Electron-Nuclear Wave Function", *J. Chem. Phys.* **2015**, *143* (7), 074106, doi:10.1063/1.4928578. (cited on pages 6 and 60)

[67] C. Herrmann, K. Ruud, M. Reiher, "Importance of Backbone Angles Versus Amino Acid Configurations in Peptide Vibrational Raman Optical Activity Spectra", *Chem. Phys.* **2008**, *343* (2-3), 200–209, doi:10.1016/j.chemphys.2007.06.002. (cited on page 6)

[68] S. Luber, C. Herrmann, M. Reiher, "Relevance of the Electric-Dipole–Electric-Quadrupole Contribution to Raman Optical Activity Spectra", *J. Phys. Chem. B* **2008**, *112* (7), 2218–2232, doi:10.1021/jp0756404. (cited on pages 6, 28, and 62)

[69] N. S. Bieler, M. P. Haag, C. R. Jacob, M. Reiher, "Analysis of the Cartesian Tensor Transfer Method for Calculating Vibrational Spectra of Polypeptides", *J. Chem. Theory Comput.* **2011**, *7* (6), 1867–1881, doi:10.1021/ct2001478. (cited on pages 6 and 62)

[70] J. R. Cheeseman, M. S. Shaik, P. L. A. Popelier, E. W. Blanch, "Calculation of Raman Optical Activity Spectra of Methyl-β-D-Glucose Incorporating a Full Molecular Dynamics Simulation of Hydration Effects", *J. Am. Chem. Soc.* **2011**, *133* (13), 4991–4997, doi:10.1021/ja110825z. (cited on pages 6 and 62)

[71] V. P. Nicu, E. J. Baerends, P. L. Polavarapu, "Understanding Solvent Effects in Vibrational Circular Dichroism Spectra: [1,1′-Binaphthalene]-2,2′-diol in Dichloromethane, Acetonitrile, and Dimethyl Sulfoxide Solvents", *J. Phys. Chem. A* **2012**, *116* (32), 8366–8373, doi:10.1021/jp303891x. (cited on page 6)

[72] S. Luber, “Solvent Effects in Calculated Vibrational Raman Optical Activity Spectra of α-Helices”, *J. Phys. Chem. A* **2013**, *117* (13), 2760–2770, doi:10.1021/jp400105u. (cited on pages 6 and 62)

[73] C. Herrmann, J. Neugebauer, M. Presselt, U. Uhlemann, M. Schmitt, S. Rau, J. Popp, M. Reiher, “The First Photoexcitation Step of Ruthenium-Based Models for Artificial Photosynthesis Highlighted by Resonance Raman Spectroscopy”, *J. Phys. Chem. B* **2007**, *111* (21), 6078–6087, doi:10.1021/jp071692h. (cited on pages 6 and 69)

[74] K. Kiewisch, J. Neugebauer, M. Reiher, “Selective Calculation of High-Intensity Vibrations in Molecular Resonance Raman Spectra”, *J. Chem. Phys.* **2008**, *129* (20), doi:10.1063/1.3013351. (cited on pages 6 and 69)

[75] S. K. Panda, C. Jacob, “Surface Enhanced Raman Scattering and Photoluminescence Properties of Catalytic Grown ZnO Nanostructures”, *Appl. Phys. A Mater. Sci. Process.* **2009**, *96* (4), 805–811, doi:10.1007/s00339-009-5309-9. (cited on pages 6 and 69)

[76] J. Guthmuller, “Assessment of TD-DFT and CC2 Methods for the Calculation of Resonance Raman Intensities: Application to *o*-Nitrophenol”, *J. Chem. Theory Comput.* **2011**, *7*, 1082–1089, doi:10.1021/ct200010p. (cited on pages 6, 69, and 105)

[77] J. M. Boereboom, M. C. van Hemert, J. Neugebauer, “The Resonance Raman Spectra of Spheroidene Revisited with a First-Principles Approach”, *ChemPhysChem* **2011**, *12* (17), 3157–3169, doi:10.1002/cphc.201100545. (cited on pages 6 and 69)

[78] A. Kubas, J. Noak, A. Trunschke, R. Schlögl, F. Neese, D. Maganas, “A Combined Experimental and Theoretical Spectroscopic Protocol for Determination of the Structure of Heterogeneous Catalysts: Developing the Information Content of the Resonance Raman Spectra of M1 MoVOx”, *Chem. Sci.* **2017**, *8* (9), 6338–6353, doi:10.1039/c7sc01771e. (cited on pages 6 and 69)

[79] Y. Ma, S. Knecht, M. Reiher, “Multiconfigurational Effects in Theoretical Resonance Raman Spectra”, *ChemPhysChem* **2017**, *18* (4), 384–393, doi:10.1002/cphc.201601072. (cited on pages 6 and 69)

[80] C. Römelt, S. Ye, E. Bill, T. Weyhermüller, M. van Gastel, F. Neese, “Electronic Structure and Spin Multiplicity of Iron Tetraphenylporphyrins in Their Reduced States As Determined by a Combination of Resonance Raman Spectroscopy and Quantum Chemistry”, *Inorg. Chem.* **2018**, *57* (4), 2141–2148, doi:10.1021/acs.inorgchem.7b03018. (cited on pages 6 and 69)

[81] B. de Souza, G. Farias, F. Neese, R. Izsák, “Efficient Simulation of Overtones and Combination Bands in Resonant Raman Spectra”, *J. Chem. Phys.* **2019**, *150* (21), 214102, doi:10.1063/1.5099247. (cited on pages 6 and 69)

[82] A. Haack, J. Hillenbrand, M. van Gastel, A. Fürstner, F. Neese, “Spectroscopic and Theoretical Study on Siloxy-Based Molybdenum and Tungsten Alkylidyne Catalysts for Alkyne Metathesis”, *ACS Catalysis* **2021**, *11* (15), 9086–9101, doi:10.1021/acscatal.1c01587. (cited on pages 6 and 69)

[83] J. Lindner, D. Cringus, M. S. Pshenichnikov, P. Vöhringer, “Anharmonic Bend–Stretch Coupling in Neat Liquid Water”, *Chem. Phys.* **2007**, *341* (1-3), 326–335, doi:10.1016/j.chemphys.2007.07.051. (cited on page 6)

[84] A. Klamt, G. Schüürmann, "COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient", *J. Chem. Soc., Perkin Trans. 2* **1993**, *5*, 799–805, doi:10.1039/p29930000799. (cited on page 6)

[85] B. Mennucci, E. Cancès, J. Tomasi, "Evaluation of Solvent Effects in Isotropic and Anisotropic Dielectrics and in Ionic Solutions with a Unified Integral Equation Method: Theoretical Bases, Computational Implementation, and Numerical Applications", *J. Phys. Chem. B* **1997**, *101* (49), 10506–10517, doi:10.1021/jp971959k. (cited on page 6)

[86] V. Barone, M. Cossi, "Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model", *J. Phys. Chem. A* **1998**, *102* (11), 1995–2001, doi:10.1021/jp9716997. (cited on page 6)

[87] D. M. York, M. Karplus, "A Smooth Solvation Potential Based on the Conductor-Like Screening Model", *J. Phys. Chem. A* **1999**, *103* (50), 11060–11079, doi:10.1021/jp992097l. (cited on page 6)

[88] B. Mennucci, J. Tomasi, R. Cammi, J. R. Cheeseman, M. J. Frisch, F. J. Devlin, S. Gabriel, P. J. Stephens, "Polarizable Continuum Model (PCM) Calculations of Solvent Effects on Optical Rotations of Chiral Molecules", *J. Phys. Chem. A* **2002**, *106* (25), 6102–6113, doi:10.1021/jp020124t. (cited on page 6)

[89] M. Cossi, N. Rega, G. Scalmani, V. Barone, "Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model", *J. Comput. Chem.* **2003**, *24* (6), 669–681, doi:10.1002/jcc.10189. (cited on page 6)

[90] H. M. Senn, P. M. Margl, R. Schmid, T. Ziegler, P. E. Blöchl, "Ab initio Molecular Dynamics with a Continuum Solvation Model", *J. Chem. Phys.* **2003**, *118* (3), 1089–1100, doi:10.1063/1.1528890. (cited on page 6)

[91] J. Tomasi, B. Mennucci, R. Cammi, "Quantum Mechanical Continuum Solvation Models", *Chem. Rev.* **2005**, *105* (8), 2999–3094, doi:10.1021/cr9904009. (cited on page 6)

[92] F. Neese, F. Wennmohs, A. Hansen, "Efficient and Accurate Local Approximations to Coupled-Electron Pair Approaches: An Attempt to Revive the Pair Natural Orbital Method", *J. Chem. Phys.* **2009**, *130* (11), 114108, doi:10.1063/1.3086717. (cited on page 7)

[93] F. Neese, F. Wennmohs, A. Hansen, U. Becker, "Efficient, Approximate and Parallel Hartree–Fock and Hybrid DFT Calculations. A »Chain-of-Spheres« Algorithm for the Hartree–Fock Exchange", *Chem. Phys.* **2009**, *356* (1-3), 98–109, doi:10.1016/j.chemphys.2008.10.036. (cited on page 7)

[94] F. Neese, A. Hansen, F. Wennmohs, S. Grimme, "Accurate Theoretical Chemistry with Coupled Pair Models", *Acc. Chem. Res.* **2009**, *42* (5), 641–648, doi:10.1021/ar800241t. (cited on page 7)

[95] J. Neugebauer, B. A. Hess, "Fundamental Vibrational Frequencies of Small Polyatomic Molecules from Density-Functional Calculations and Vibrational Perturbation Theory", *J. Chem. Phys.* **2003**, *118* (16), 7215–7225, doi:10.1063/1.1561045. (cited on page 7)

[96] T. Hrenar, H.-J. Werner, G. Rauhut, "Accurate Calculation of Anharmonic Vibrational Frequencies of Medium Sized Molecules Using Local Coupled Cluster Methods", *J. Chem. Phys.* **2007**, *126* (13), 134108, doi:`10.1063/1.2718951`. `(cited on page 7)`

[97] M. Kumarasiri, C. Swalina, S. Hammes-Schiffer, "Anharmonic Effects in Ammonium Nitrate and Hydroxylammonium Nitrate Clusters", *J. Phys. Chem. B* **2007**, *111* (18), 4653–4658, doi:`10.1021/jp065569m`. `(cited on page 7)`

[98] S. Heislbetz, G. Rauhut, "Vibrational Multiconfiguration Self-Consistent Field Theory: Implementation and Test Calculations", *J. Chem. Phys.* **2010**, *132* (12), 124102, doi:`10.1063/1.3364861`. `(cited on page 7)`

[99] P. T. Panek, C. R. Jacob, "Efficient Calculation of Anharmonic Vibrational Spectra of Large Molecules with Localized Modes", *ChemPhysChem* **2014**, *15* (15), 3365–3377, doi:`10.1002/cphc.201402251`. `(cited on page 7)`

[100] R. Ramakrishnan, G. Rauhut, "Semi-Quartic Force Fields Retrieved from Multi-Mode Expansions: Accuracy, Scaling Behavior, and Approximations", *J. Chem. Phys.* **2015**, *142* (15), 154118, doi:`10.1063/1.4918587`. `(cited on page 7)`

[101] P. T. Panek, C. R. Jacob, "Anharmonic Theoretical Vibrational Spectroscopy of Polypeptides", *J. Phys. Chem. Lett.* **2016**, *7* (16), 3084–3090, doi:`10.1021/acs.jpclett.6b01451`. `(cited on page 7)`

[102] B. Ziegler, G. Rauhut, "Accurate Vibrational Configuration Interaction Calculations on Diborane and its Isotopologues", *J. Phys. Chem. A* **2019**, *123* (15), 3367–3373, doi:`10.1021/acs.jpca.9b01604`. `(cited on page 7)`

[103] S. Brünken, F. Lipparini, A. Stoffels, P. Jusko, B. Redlich, J. Gauss, S. Schlemmer, "Gas-Phase Vibrational Spectroscopy of the Hydrocarbon Cations l-C_3H^+, HC_3H^+, and c-$C_3H_2^+$: Structures, Isomers, and the Influence of Ne-Tagging", *J. Phys. Chem. A* **2019**, *123* (37), 8053–8062, doi:`10.1021/acs.jpca.9b06176`. `(cited on page 7)`

[104] D. Witsch, V. Lutter, A. A. Breier, K. M. T. Yamada, G. W. Fuchs, J. Gauss, T. F. Giesen, "Infrared Spectroscopy of Disilicon-Carbide, Si_2C: The ν_3 Fundamental Band", *J. Phys. Chem. A* **2019**, *123* (19), 4168–4177, doi:`10.1021/acs.jpca.9b01605`. `(cited on page 7)`

[105] H. R. Larsson, "Computing Vibrational Eigenstates with Tree Tensor Network States (TTNS)", *J. Chem. Phys.* **2019**, *151* (20), 204102, doi:`10.1063/1.5130390`. `(cited on page 7)`

[106] R. G. Gordon, "Molecular Motion in Infrared and Raman Spectra", *J. Chem. Phys.* **1965**, *43* (4), 1307–1312, doi:`10.1063/1.1696920`. `(cited on pages 7, 52, and 60)`

[107] H. Shimizu, "Time-Correlation Function of Molecular Random Motion and Shape of Spectral Bands", *J. Chem. Phys.* **1965**, *43* (7), 2453–2465, doi:`10.1063/1.1697145`. `(cited on pages 7, 52, and 60)`

[108] R. G. Gordon, "Correlation Functions for Molecular Motion", in "Advances in Magnetic and Optical Resonance", pages 1–42, Elsevier, **1968**, doi:`10.1016/b978-1-4832-3116-7.50008-4`. `(cited on pages 7, 52, and 60)`

[109] H. Shimizu, "Dependence of the Time-Correlation Functions of Molecular Random Motions on the Intermolecular Potential", *J. Chem. Phys.* **1968**, *48* (6), 2494–2501, doi:`10.1063/1.1669473`. `(cited on pages 7, 52, and 60)`

[110] A. Cabana, R. Bardoux, A. Chamberland, “Infrared Spectra, Rotational Correlation Functions, Band Moments, and Intermolecular Mean Squared Torques of Methane Dissolved in Liquid Noble Gases”, *Can. J. Chem.* **1969**, *47* (16), 2915–2920, doi:10.1139/v69-488. (cited on pages 7, 52, and 60)

[111] K. D. Hänsel, “On the Spectra of Highly Vibrationally Excited Polyatomic Molecules”, *Chem. Phys.* **1978**, *33* (1), 35–43, doi:10.1016/0301-0104(78)87068-2. (cited on pages 7, 52, and 60)

[112] R. Kubo, M. Toda, N. Hashitsume, *Statistical Physics, 2nd Ed.*, volume II, Springer, Berlin, **1991**. (cited on pages 7, 52, and 60)

[113] M. Thomas, *Theoretical Modeling of Vibrational Spectra in the Liquid Phase*, Springer International Publishing, **2017**, doi:10.1007/978-3-319-49628-3. (cited on pages 7, 8, 52, 55, 78, and 130)

[114] D. Marx, J. Hutter, *Ab initio Molecular Dynamics*, Cambridge University Press, **2009**, doi:10.1017/cbo9780511609633. (cited on pages 7 and 52)

[115] B. Kirchner, P. J. di Dio, J. Hutter, “Real-World Predictions from ab initio Molecular Dynamics Simulations”, *Top. Curr. Chem.* **2012**, *307*, 109–154, doi:10.1007/128_2011_195. (cited on pages 7, 9, and 52)

[116] R. Car, M. Parrinello, “Unified Approach for Molecular Dynamics and Density-Functional Theory”, *Phys. Rev. Lett.* **1985**, *55* (22), 2471–2474, doi:10.1103/PhysRevLett.55.2471. (cited on pages 7 and 52)

[117] H. Wang, N. Agmon, “Reinvestigation of the Infrared Spectrum of the Gas-Phase Protonated Water Tetramer”, *J. Phys. Chem. A* **2017**, *121* (16), 3056–3070, doi:10.1021/acs.jpca.7b01856. (cited on pages 7 and 52)

[118] P. L. Silvestrelli, M. Bernasconi, M. Parrinello, “Ab initio Infrared Spectrum of Liquid Water”, *Chem. Phys. Lett.* **1997**, *277* (5-6), 478–482, doi:10.1016/s0009-2614(97)00930-5. (cited on pages 7 and 52)

[119] A. Pasquarello, R. Car, “Dynamical Charge Tensors and Infrared Spectrum of Amorphous SiO_2”, *Phys. Rev. Lett.* **1997**, *79* (9), 1766–1769, doi:10.1103/physrevlett.79.1766. (cited on pages 7 and 52)

[120] A. Putrino, M. Parrinello, “Anharmonic Raman Spectra in High-Pressure Ice from ab initio Simulations”, *Phys. Rev. Lett.* **2002**, *88* (17), 176401, doi:10.1103/physrevlett.88.176401. (cited on pages 7 and 55)

[121] Z. Zhu, M. E. Tuckerman, “Ab initio Molecular Dynamics Investigation of the Concentration Dependence of Charged Defect Transport in Basic Solutions via Calculation of the Infrared Spectrum”, *J. Phys. Chem. B* **2002**, *106* (33), 8009–8018, doi:10.1021/jp020866m. (cited on pages 8 and 52)

[122] M.-P. Gaigeot, M. Sprik, “Ab initio Molecular Dynamics Computation of the Infrared Spectrum of Aqueous Uracil”, *J. Phys. Chem. B* **2003**, *107* (38), 10344–10358, doi:10.1021/jp034788u. (cited on pages 8, 52, and 74)

[123] R. Iftimie, M. E. Tuckerman, “Decomposing Total IR Spectra of Aqueous Systems into Solute and Solvent Contributions: A Computational Approach Using Maximally Localized Wannier Orbitals”, *J. Chem. Phys.* **2005**, *122* (21), 214508, doi:10.1063/1.1908950. (cited on pages 8 and 52)

[124] M.-P. Gaigeot, R. Vuilleumier, M. Sprik, D. Borgis, "Infrared Spectroscopy of *N*-Methylacetamide Revisited by ab initio Molecular Dynamics Simulations", *J. Chem. Theory Comput.* **2005**, *1* (5), 772–789, doi:`10.1021/ct050029z`. `(cited on pages 8 and 52)`

[125] J. P. Devlin, V. Buch, F. R. Mohamed, M. Parrinello, "The Amorphous Analogs of the Crystalline Monohydrate of HCl: Structures and Spectra", *Chem. Phys. Lett.* **2006**, *432* (4-6), 462–467, doi:`10.1016/j.cplett.2006.11.001`. `(cited on pages 8 and 52)`

[126] R. Iftimie, M. E. Tuckerman, "The Molecular Origin of the »Continuous« Infrared Absorption in Aqueous Solutions of Acids: A Computational Approach", *Angew. Chem. - Int. Ed.* **2006**, *45* (7), 1144–1147, doi:`10.1002/anie.200502259`. `(cited on pages 8 and 52)`

[127] M.-P. Gaigeot, M. Martinez, R. Vuilleumier, "Infrared Spectroscopy in the Gas and Liquid Phase from First Principle Molecular Dynamics Simulations: Application to Small Peptides", *Mol. Phys.* **2007**, *105* (19-22), 2857–2878, doi:`10.1080/00268970701724974`. `(cited on pages 8 and 52)`

[128] M.-P. Gaigeot, "Alanine Polypeptide Structural Fingerprints at Room Temperature: What Can Be Gained from Non-Harmonic Car–Parrinello Molecular Dynamics Simulations", *J. Phys. Chem. A* **2008**, *112* (51), 13507–13517, doi:`10.1021/jp807550j`. `(cited on pages 8 and 52)`

[129] W. Chen, M. Sharma, R. Resta, G. A. Galli, R. Car, "Role of Dipolar Correlations in the Infrared Spectra of Water and Ice", *Phys. Rev. B* **2008**, *77* (24), 245114, doi:`10.1103/physrevb.77.245114`. `(cited on pages 8 and 52)`

[130] M. Pagliai, C. Cavazzoni, G. Cardini, G. Erbacci, M. Parrinello, V. Schettino, "Anharmonic Infrared and Raman Spectra in Car–Parrinello Molecular Dynamics Simulations", *J. Chem. Phys.* **2008**, *128* (22), 224514, doi:`10.1063/1.2936988`. `(cited on pages 8, 52, and 55)`

[131] M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith, D. Marx, "Dissecting the THz Spectrum of Liquid Water from First Principles via Correlations in Time and Space", *Proc. Natl. Acad. Sci.* **2010**, *107* (27), 12068–12073, doi:`10.1073/pnas.0914885107`. `(cited on pages 8 and 52)`

[132] M.-P. Gaigeot, N. A. Besley, J. D. Hirst, "Modeling the Infrared and Circular Dichroism Spectroscopy of a Bridged Cyclic Diamide", *J. Phys. Chem. B* **2011**, *115* (18), 5526–5535, doi:`10.1021/jp111140f`. `(cited on pages 8 and 52)`

[133] N. Sieffert, M. Bühl, M.-P. Gaigeot, C. A. Morrison, "Liquid Methanol from DFT and DFT/MM Molecular Dynamics Simulations", *J. Chem. Theory Comput.* **2013**, *9* (1), 106–118, doi:`10.1021/ct300784x`. `(cited on pages 8 and 52)`

[134] S. Luber, "Local Electric Dipole Moments for Periodic Systems via Density Functional Theory Embedding", *J. Chem. Phys.* **2014**, *141* (23), 234110, doi:`10.1063/1.4903828`. `(cited on pages 8, 15, and 52)`

[135] J. Sun, G. Niehues, H. Forbert, D. Decka, G. Schwaab, D. Marx, M. Havenith, "Understanding THz Spectra of Aqueous Solutions: Glycine in Light and Heavy Water", *J. Am. Chem. Soc.* **2014**, *136* (13), 5031–5038, doi:`10.1021/ja4129857`. `(cited on pages 8 and 52)`

[136] I. Souza, J. Íñiguez, D. Vanderbilt, “First-Principles Approach to Insulators in Finite Electric Fields”, *Phys. Rev. Lett.* **2002**, *89* (11), 117602, doi:10.1103/physrevlett.89.117602. (cited on pages 8 and 55)

[137] P. Umari, A. Pasquarello, “Ab initio Molecular Dynamics in a Finite Homogeneous Electric Field”, *Phys. Rev. Lett.* **2002**, *89* (15), 157602, doi:10.1103/physrevlett.89.157602. (cited on pages 8 and 55)

[138] Q. Wan, L. Spanu, G. A. Galli, F. Gygi, “Raman Spectra of Liquid Water from ab initio Molecular Dynamics: Vibrational Signatures of Charge Fluctuations in the Hydrogen Bond Network”, *J. Chem. Theory Comput.* **2013**, *9* (9), 4124–4130, doi:10.1021/ct4005307. (cited on pages 8, 15, and 55)

[139] S. Luber, M. Iannuzzi, J. Hutter, “Raman Spectra from ab initio Molecular Dynamics and Its Application to Liquid *(S)*-Methyloxirane”, *J. Chem. Phys.* **2014**, *141* (9), 094503, doi:10.1063/1.4894425. (cited on pages 8, 55, 86, 92, and 94)

[140] T. Ikeda, “Infrared Absorption and Raman Scattering Spectra of Water under Pressure via First Principles Molecular Dynamics”, *J. Chem. Phys.* **2014**, *141* (4), 044501, doi:10.1063/1.4890369. (cited on pages 8 and 55)

[141] M. Petrov, B. Minofar, L. Vrbka, P. Jungwirth, P. Koelsch, H. Motschmann, “Aqueous Ionic and Complementary Zwitterionic Soluble Surfactants: Molecular Dynamics Simulations and Sum Frequency Generation Spectroscopy of the Surfaces”, *Langmuir* **2006**, *22* (6), 2498–2505, doi:10.1021/la0532095. (cited on page 8)

[142] C. Roth, A. Appelhagen, N. Jobst, R. Ludwig, “Microheterogeneities in Ionic-Liquid–Methanol Solutions Studied by FTIR Spectroscopy, DFT Calculations and Molecular Dynamics Simulations”, *ChemPhysChem* **2012**, *13* (7), 1708–1717, doi:10.1002/cphc.201101022. (cited on page 8)

[143] C. Roth, S. Chatzipapadopoulos, D. Kerlé, F. Friedriszik, M. Lütgens, S. Lochbrunner, O. Kühn, R. Ludwig, “Hydrogen Bonding in Ionic Liquids Probed by Linear and Nonlinear Vibrational Spectroscopy”, *New J. Phys.* **2012**, *14* (10), 105026, doi:10.1088/1367-2630/14/10/105026. (cited on page 8)

[144] E. Pluhařová, M. D. Baer, C. J. Mundy, B. Schmidt, P. Jungwirth, “Aqueous Cation–Amide Binding: Free Energies and IR Spectral Signatures by ab initio Molecular Dynamics”, *J. Phys. Chem. Lett.* **2014**, *5* (13), 2235–2240, doi:10.1021/jz500976m. (cited on page 8)

[145] K. Fumino, V. Fossog, P. Stange, K. Wittler, W. Polet, R. Hempelmann, R. Ludwig, “Ion Pairing in Protic Ionic Liquids Probed by Far-Infrared Spectroscopy: Effects of Solvent Polarity and Temperature”, *ChemPhysChem* **2014**, *15* (12), 2604–2609, doi:10.1002/cphc.201402205. (cited on page 8)

[146] T. Zentel, V. Overbeck, D. Michalik, O. Kühn, R. Ludwig, “Hydrogen Bonding in Protic Ionic Liquids: Structural Correlations, Vibrational Spectroscopy, and Rotational Dynamics of Liquid Ethylammonium Nitrate”, *J. Phys. B: At. Mol. Opt. Phys.* **2018**, *51* (3), 034002, doi:10.1088/1361-6455/aa9a9e. (cited on page 8)

[147] T. Niemann, A. Strate, R. Ludwig, H. J. Zeng, F. S. Menges, M. A. Johnson, “Spectroscopic Evidence for an Attractive Cation–Cation Interaction in Hydroxy-Functionalized Ionic Liquids: A Hydrogen-Bonded Chain-Like Trimer”, *Angew. Chem. - Int. Ed.* **2018**, *57* (47), 15364–15368, doi:10.1002/anie.201808381. (cited on page 8)

[148] D. H. Zaitsau, J. Neumann, T. Niemann, A. Strate, D. Paschek, S. P. Verevkin, R. Ludwig, "Isolating the Role of Hydrogen Bonding in Hydroxyl-Functionalized Ionic Liquids by Means of Vaporization Enthalpies, Infrared Spectroscopy and Molecular Dynamics Simulations", *Phys. Chem. Chem. Phys.* **2019**, *21* (36), 20308–20314, doi:`10.1039/c9cp04337c`. `(cited on page 8)`

[149] D. M. de Oliveira, S. R. Zukowski, V. Palivec, J. Hénin, H. Martinez-Seara, D. Ben-Amotz, P. Jungwirth, E. Duboué-Dijon, "Binding of Divalent Cations to Acetate: Molecular Simulations Guided by Raman Spectroscopy", *Phys. Chem. Chem. Phys.* **2020**, *22* (41), 24014–24027, doi:`10.1039/d0cp02987d`. `(cited on page 8)`

[150] M. Thomas, B. Kirchner, "Classical Magnetic Dipole Moments for the Simulation of Vibrational Circular Dichroism by ab initio Molecular Dynamics", *J. Phys. Chem. Lett.* **2016**, *7* (3), 509–513, doi:`10.1021/acs.jpclett.5b02752`. `(cited on pages 8, 20, 57, 60, 61, 80, 92, and 94)`

[151] A. Scherrer, R. Vuilleumier, D. Sebastiani, "Vibrational Circular Dichroism from ab initio Molecular Dynamics and Nuclear Velocity Perturbation Theory in the Liquid Phase", *J. Chem. Phys.* **2016**, *145* (8), 084101, doi:`10.1063/1.4960653`. `(cited on pages 8 and 61)`

[152] S. Luber, "Raman Optical Activity Spectra from Density Functional Perturbation Theory and Density-Functional-Theory-Based Molecular Dynamics", *J. Chem. Theory Comput.* **2017**, *13* (3), 1254–1262, doi:`10.1021/acs.jctc.6b00820`. `(cited on pages 8, 62, and 93)`

[153] A. Morita, J. T. Hynes, "A Theoretical Analysis of the Sum Frequency Generation Spectrum of the Water Surface", *Chem. Phys.* **2000**, *258* (2-3), 371–390, doi:`10.1016/s0301-0104(00)00127-0`. `(cited on pages 8 and 68)`

[154] A. Perry, H. Ahlborn, B. Space, P. B. Moore, "A Combined Time Correlation Function and Instantaneous Normal Mode Study of the Sum Frequency Generation Spectroscopy of the Water / Vapor Interface", *J. Chem. Phys.* **2003**, *118* (18), 8411–8419, doi:`10.1063/1.1565994`. `(cited on pages 8 and 68)`

[155] A. Perry, C. Neipert, C. R. Kasprzyk, T. Green, B. Space, P. B. Moore, "A Theoretical Description of the Polarization Dependence of the Sum Frequency Generation Spectroscopy of the Water / Vapor Interface", *J. Chem. Phys.* **2005**, *123* (14), 144705, doi:`10.1063/1.2046630`. `(cited on pages 8 and 68)`

[156] T. Ishiyama, A. Morita, "Intermolecular Correlation Effect in Sum Frequency Generation Spectroscopy of Electrolyte Aqueous Solution", *Chem. Phys. Lett.* **2006**, *431* (1-3), 78–82, doi:`10.1016/j.cplett.2006.09.024`. `(cited on pages 8 and 68)`

[157] D. S. Walker, D. K. Hore, G. L. Richmond, "Understanding the Population, Coordination, and Orientation of Water Species Contributing to the Nonlinear Optical Spectroscopy of the Vapor–Water Interface through Molecular Dynamics Simulations", *J. Phys. Chem. B* **2006**, *110* (41), 20451–20459, doi:`10.1021/jp063063y`. `(cited on pages 8 and 68)`

[158] V. Buch, T. Tarbuck, G. L. Richmond, H. Groenzin, I. Li, M. J. Shultz, "Sum Frequency Generation Surface Spectra of Ice, Water, and Acid Solution Investigated by an Exciton Model", *J. Chem. Phys.* **2007**, *127* (20), 204710, doi:`10.1063/1.2790437`. `(cited on pages 8 and 68)`

[159] J. H. Jang, F. Lydiatt, R. Lindsay, S. Baldelli, "Quantitative Orientation Analysis by Sum Frequency Generation in the Presence of Near-Resonant Background Signal: Acetonitrile on Rutile TiO_2 (110)", *J. Phys. Chem. A* **2013**, *117* (29), 6288–6302, doi:10.1021/jp401019p. (cited on pages 8, 67, and 68)

[160] S. Luber, "Sum Frequency Generation of Acetonitrile on a Rutile (110) Surface from Density Functional Theory-Based Molecular Dynamics", *J. Phys. Chem. Lett.* **2016**, *7* (24), 5183–5187, doi:10.1021/acs.jpclett.6b02530. (cited on pages 8 and 68)

[161] M. Sulpizi, M. Salanne, M. Sprik, M.-P. Gaigeot, "Vibrational Sum Frequency Generation Spectroscopy of the Water Liquid–Vapor Interface from Density Functional Theory-Based Molecular Dynamics Simulations", *J. Phys. Chem. Lett.* **2012**, *4* (1), 83–87, doi:10.1021/jz301858g. (cited on pages 8 and 68)

[162] S. Pezzotti, D. R. Galimberti, M.-P. Gaigeot, "2D H-Bond Network As the Topmost Skin to the Air–Water Interface", *J. Phys. Chem. Lett.* **2017**, *8* (13), 3133–3141, doi:10.1021/acs.jpclett.7b01257. (cited on pages 8 and 68)

[163] S. Pezzotti, D. R. Galimberti, M.-P. Gaigeot, "Deconvolution of BIL–SFG and DL–SFG Spectroscopic Signals Reveals Order / Disorder of Water at the Elusive Aqueous Silica Interface", *Phys. Chem. Chem. Phys.* **2019**, *21* (40), 22188–22202, doi:10.1039/c9cp02766a. (cited on pages 8 and 68)

[164] D. Ojha, N. K. Kaliannan, T. D. Kühne, "Time-Dependent Vibrational Sum-Frequency Generation Spectroscopy of the Air–Water Interface", *Commun. Chem.* **2019**, *2* (1), doi:10.1038/s42004-019-0220-6. (cited on pages 8 and 68)

[165] D. Ojha, T. D. Kühne, "»On-the-Fly« Calculation of the Vibrational Sum-Frequency Generation Spectrum at the Air–Water Interface", *Molecules* **2020**, *25* (17), 3939, doi:10.3390/molecules25173939. (cited on pages 8 and 68)

[166] Y. Litman, J. Behler, M. Rossi, "Temperature Dependence of the Vibrational Spectrum of Porphycene: A Qualitative Failure of Classical-Nuclei Molecular Dynamics", *Faraday Discuss.* **2020**, *221*, 526–546, doi:10.1039/c9fd00056a. (cited on page 8)

[167] A. D. Bandrauk, S. Chelkowski, D. J. Diestler, J. Manz, K.-J. Yuan, "Quantum Simulation of High-Order Harmonic Spectra of the Hydrogen Atom", *Phys. Rev. A* **2009**, *79* (2), 023403, doi:10.1103/physreva.79.023403. (cited on page 8)

[168] D. Marx, M. Parrinello, "Ab initio Path Integral Molecular Dynamics: Basic Ideas", *J. Chem. Phys.* **1996**, *104* (11), 4077–4082, doi:10.1063/1.471221. (cited on page 8)

[169] J. Manz, "The Relation of Chemical Potentials and Reactivity Studied by a State Path Sum", *Mol. Phys.* **1974**, *28* (2), 399–413, doi:10.1080/00268977400102931. (cited on page 8)

[170] J. Cao, G. A. Voth, "The Formulation of Quantum Statistical Mechanics Based on the Feynman Path Centroid Density. I. Equilibrium Properties", *J. Chem. Phys.* **1994**, *100* (7), 5093–5105, doi:10.1063/1.467175. (cited on page 8)

[171] J. Cao, G. A. Voth, "The Formulation of Quantum Statistical Mechanics Based on the Feynman Path Centroid Density. II. Dynamical Properties", *J. Chem. Phys.* **1994**, *100* (7), 5106–5117, doi:10.1063/1.467176. (cited on page 8)

[172] S. Jang, G. A. Voth, "A Derivation of Centroid Molecular Dynamics and Other Approximate Time Evolution Methods for Path Integral Centroid Variables", *J. Chem. Phys.* **1999**, *111* (6), 2371–2384, doi:`10.1063/1.479515`. `(cited on page 8)`

[173] R. Ramírez, T. López-Ciudad, "The Schrödinger Formulation of the Feynman Path Centroid Density", *J. Chem. Phys.* **1999**, *111* (8), 3339–3348, doi:`10.1063/1.479666`. `(cited on page 8)`

[174] E. A. Polyakov, A. P. Lyubartsev, P. N. Vorontsov-Velyaminov, "Centroid Molecular Dynamics: Comparison with Exact Results for Model Systems", *J. Chem. Phys.* **2010**, *133* (19), 194103, doi:`10.1063/1.3484490`. `(cited on page 8)`

[175] I. R. Craig, D. E. Manolopoulos, "Quantum Statistics and Classical Mechanics: Real Time Correlation Functions from Ring Polymer Molecular Dynamics", *J. Chem. Phys.* **2004**, *121* (8), 3368–3373, doi:`10.1063/1.1777575`. `(cited on page 8)`

[176] B. J. Braams, D. E. Manolopoulos, "On the Short-Time Limit of Ring Polymer Molecular Dynamics", *J. Chem. Phys.* **2006**, *125* (12), 124105, doi:`10.1063/1.2357599`. `(cited on page 8)`

[177] J. Kessler, H. Elgabarty, T. Spura, K. Karhan, P. Partovi-Azar, A. A. Hassanali, T. D. Kühne, "Structure and Dynamics of the Instantaneous Water / Vapor Interface Revisited by Path-Integral and ab initio Molecular Dynamics Simulations", *J. Phys. Chem. B* **2015**, *119* (31), 10079–10086, doi:`10.1021/acs.jpcb.5b04185`. `(cited on page 8)`

[178] T. Spura, H. Elgabarty, T. D. Kühne, "»On-the-Fly« Coupled Cluster Path-Integral Molecular Dynamics: Impact of Nuclear Quantum Effects on the Protonated Water Dimer", *Phys. Chem. Chem. Phys.* **2015**, *17* (22), 14355–14359, doi:`10.1039/c4cp05192k`. `(cited on page 8)`

[179] B. J. Alder, T. E. Wainwright, "Studies in Molecular Dynamics. I. General Method", *J. Chem. Phys.* **1959**, *31 (2)*, 459, doi:`10.1063/1.1730376`. `(cited on page 9)`

[180] G. E. Moore, "Cramming More Components onto Integrated Circuits", *Electronics Magazine* **1965**, page 4, doi:`10.1109/jproc.1998.658762`. `(cited on page 9)`

[181] W. Eckhardt, A. Heinecke, R. Bader, M. Brehm, N. Hammer, H. Huber, H.-G. Kleinhenz, J. Vrabec, H. Hasse, M. Horsch, M. Bernreuther, C. W. Glass, C. Niethammer, A. Bode, H.-J. Bungartz, "591 TFLOPs Multi-Trillion Particles Simulation on Supermuc", *Lecture Notes in Comput. Sci.* **2013**, *7905*, 1–12, doi:`10.1007/978-3-642-38750-0_1`. `(cited on page 9)`

[182] L. Verlet, "Computer »Experiments« on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules", *Phys. Rev.* **1967**, *159*, 98–103, doi:`10.1103/physrev.159.98`. `(cited on pages 10, 80, and 121)`

[183] C. Lubich, E. Hairer, G. Wanner, *Geometric Numerical Integration*, Springer-Verlag GmbH, **2006**, ISBN 3540306668, doi:`10.4171/owr/2011/16`. `(cited on pages 10 and 121)`

[184] G. J. Martyna, M. L. Klein, M. E. Tuckerman, "Nosé–Hoover Chains: The Canonical Ensemble via Continuous Dynamics", *J. Chem. Phys.* **1992**, *97*, 2635–2643, doi:`10.1063/1.463940`. `(cited on pages 11, 90, 111, and 112)`

[185] S. Nose, “A Unified Formulation of the Constant Temperature Molecular Dynamics Methods”, *J. Chem. Phys.* **1984**, *81*, 511–519, doi:10.1063/1.447334. `(cited on pages 11, 90, 111, and 112)`

[186] S. Nose, “A Molecular Dynamics Method for Simulations in the Canonical Ensemble”, *Mol. Phys.* **1984**, *52*, 255–268, doi:10.1080/00268978400101201. `(cited on pages 11, 90, 111, and 112)`

[187] W. G. Hoover, “Canonical Dynamics: Equilibrium Phase-Space Distributions”, *Phys. Rev. A* **1985**, *31* (3), 1695–1697, doi:10.1103/physreva.31.1695. `(cited on page 11)`

[188] Y. Liu, M. E. Tuckerman, “Generalized Gaussian Moment Thermostatting: A New Continuous Dynamical Approach to the Canonical Ensemble”, *J. Chem. Phys.* **2000**, *112* (4), 1685–1700, doi:10.1063/1.480769. `(cited on page 11)`

[189] S. Jang, G. A. Voth, “Simple Reversible Molecular Dynamics Algorithms for Nosé–Hoover Chain Dynamics”, *J. Chem. Phys.* **1997**, *107* (22), 9514–9526, doi:10.1063/1.475247. `(cited on pages 11 and 129)`

[190] M. E. Tuckerman, M. Parrinello, “Integrating the Car–Parrinello Equations. I. Basic Integration Techniques”, *J. Chem. Phys.* **1994**, *101* (2), 1302–1315, doi:10.1063/1.467823. `(cited on pages 11 and 12)`

[191] R. Windiks, “http://homepage.swissonline.ch/windiks/documents/thermostats.pdf , accessed on October 14 2018”, . `(cited on pages 11, 128, and 129)`

[192] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, M. L. Klein, “Explicit Reversible Integrators for Extended Systems Dynamics”, *Mol. Phys.* **1996**, *87* (5), 1117–1157, doi:10.1080/00268979600100761. `(cited on pages 12, 128, and 129)`

[193] R. D. King-Smith, D. Vanderbilt, “Theory of Polarization of Crystalline Solids”, *Phys. Rev. B* **1993**, *47* (3), 1651–1654, doi:10.1103/physrevb.47.1651. `(cited on pages 13 and 53)`

[194] R. Resta, “Macroscopic Electric Polarization As a Geometric Quantum Phase”, *Europhysics Letters (EPL)* **1993**, *22* (2), 133–138, doi:10.1209/0295-5075/22/2/010. `(cited on page 13)`

[195] R. Resta, “Macroscopic Polarization in Crystalline Dielectrics: The Geometric Phase Approach”, *Rev. Mod. Phys.* **1994**, *66* (3), 899–915, doi:10.1103/revmodphys.66.899. `(cited on pages 13 and 53)`

[196] R. Resta, “Quantum-Mechanical Position Operator in Extended Systems”, *Phys. Rev. Lett.* **1998**, *80* (9), 1800–1803, doi:10.1103/physrevlett.80.1800. `(cited on pages 13 and 53)`

[197] A. Putrino, D. Sebastiani, M. Parrinello, “Generalized Variational Density Functional Perturbation Theory”, *J. Chem. Phys.* **2000**, *113* (17), 7102–7109, doi:10.1063/1.1312830. `(cited on page 13)`

[198] S. F. Boys, “Construction of Some Molecular Orbitals to be Approximately Invariant for Changes from One Molecule to Another”, *Rev. Mod. Phys.* **1960**, *32* (2), 296–299, doi:10.1103/revmodphys.32.296. `(cited on pages 13 and 14)`

[199] J. Pipek, P. G. Mezey, "A Fast Intrinsic Localization Procedure Applicable for ab initio and Semiempirical Linear Combination of Atomic Orbital Wave Functions", *J. Chem. Phys.* **1989**, *90* (9), 4916–4926, doi:`10.1063/1.456588`. `(cited on pages 13 and 14)`

[200] C. Edmiston, K. Ruedenberg, "Localized Atomic and Molecular Orbitals", *Rev. Mod. Phys.* **1963**, *35* (3), 457–464, doi:`10.1103/revmodphys.35.457`. `(cited on pages 13 and 14)`

[201] W. Niessen, "Density Localization of Atomic and Molecular Orbitals", *Theor. Chim. Acta* **1972**, *27* (1), 9–23, doi:`10.1007/bf00528221`. `(cited on page 13)`

[202] G. H. Wannier, "The Structure of Electronic Excitation Levels in Insulating Crystals", *Phys. Rev.* **1937**, *52*, 191–197, doi:`10.1103/physrev.52.191`. `(cited on page 13)`

[203] N. Marzari, D. Vanderbilt, "Maximally Localized Generalized Wannier Functions for Composite Energy Bands", *Phys. Rev. B* **1997**, *56*, 12847–12865, doi:`10.1103/physrevb.56.12847`. `(cited on pages 13 and 14)`

[204] P. L. Silvestrelli, M. Parrinello, "Water Molecule Dipole in the Gas and in the Liquid Phase", *Phys. Rev. Lett.* **1999**, *82* (16), 3308–3311, doi:`10.1103/physrevlett.82.3308`. `(cited on page 13)`

[205] P. L. Silvestrelli, M. Parrinello, "Structural, Electronic, and Bonding Properties of Liquid Water from First Principles", *J. Chem. Phys.* **1999**, *111* (8), 3572–3580, doi:`10.1063/1.479638`. `(cited on page 13)`

[206] B. Kirchner, J. Hutter, "Solvent Effects on Electronic Properties from Wannier Functions in a Dimethyl Sulfoxide / Water Mixture", *J. Chem. Phys.* **2004**, *121* (11), 5133–5142, doi:`10.1063/1.1785780`. `(cited on pages 13, 15, and 53)`

[207] R. Iftimie, J. W. Thomas, M. E. Tuckerman, "On-the-Fly Localization of Electronic Orbitals in Car–Parrinello Molecular Dynamics", *J. Chem. Phys.* **2004**, *120* (5), 2169–2181, doi:`10.1063/1.1636697`. `(cited on page 13)`

[208] G. Berghold, C. J. Mundy, A. H. Romero, J. Hutter, M. Parrinello, "General and Efficient Algorithms for Obtaining Maximally Localized Wannier Functions", *Phys. Rev. B* **2000**, *61* (15), 10040–10048, doi:`10.1103/physrevb.61.10040`. `(cited on pages 13 and 14)`

[209] R. Resta, S. Sorella, "Electron Localization in the Insulating State", *Phys. Rev. Lett.* **1999**, *82* (2), 370–373, doi:`10.1103/physrevlett.82.370`. `(cited on page 14)`

[210] P. L. Silvestrelli, "Maximally Localized Wannier Functions for Simulations with Supercells of General Symmetry", *Phys. Rev. B* **1999**, *59* (15), 9703–9706, doi:`10.1103/physrevb.59.9703`. `(cited on page 14)`

[211] C. Jacobi, "Über ein Leichtes Verfahren die in der Theorie der Säcularstörungen Vorkommenden Gleichungen Numerisch Aufzulösen", *J. Reine Angew. Math.* **1846**, *1846* (30), 51–94, doi:`10.1515/crll.1846.30.51`. `(cited on pages 14, 25, and 75)`

[212] "The CP2k Developers Group, 2022. CP2k is freely available from `http://www.cp2k.org/` (accessed 05th May 2022).", . `(cited on pages 14, 16, 23, 25, 83, 88, 90, 96, 104, and 112)`

[213] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, "CP2k: Atomistic Simulations of Condensed Matter Systems", *WIREs Comput. Mol. Sci.* **2013**, *4* (1), 15–25, doi:10.1002/wcms.1159. (cited on pages 14, 16, 23, 25, 83, 88, 90, 96, 104, and 112)

[214] T. D. Kühne, M. Iannuzzi, M. D. Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, J. Hutter, "CP2k: An Electronic Structure and Molecular Dynamics Software Package – Quickstep: Efficient and Accurate Electronic Structure Calculations", *J. Chem. Phys.* **2020**, *152* (19), 194103, doi:10.1063/5.0007045. (cited on pages 14, 16, 23, 25, 83, 88, 90, 96, 104, and 112)

[215] P. P. Ewald, "Die Berechnung Optischer Und Elektrostatischer Gitterpotentiale", *Ann. Phys. (Berlin)* **1921**, *369* (3), 253–287, doi:10.1002/andp.19213690304. (cited on page 15)

[216] R. J. Heaton, P. A. Madden, S. J. Clark, S. Jahn, "Condensed Phase Ionic Polarizabilities from Plane Wave Density Functional Theory Calculations", *J. Chem. Phys.* **2006**, *125* (14), 144104, doi:10.1063/1.2357151. (cited on page 15)

[217] M. Salanne, R. Vuilleumier, P. A. Madden, C. Simon, P. Turq, B. Guillot, "Polarizabilities of Individual Molecules and Ions in Liquids from First Principles", *J. Phys.: Condens. Matter* **2008**, *20* (49), 494207, doi:10.1088/0953-8984/20/49/494207. (cited on page 15)

[218] P. Partovi-Azar, T. D. Kühne, "Efficient »On-the-Fly« Calculation of Raman Spectra from ab-initio Molecular Dynamics: Application to Hydrophobic / Hydrophilic Solutes in Bulk Water", *J. Comput. Chem.* **2015**, *36* (29), 2188–2192, doi:10.1002/jcc.24198. (cited on pages 15 and 55)

[219] P. Partovi-Azar, T. D. Kühne, "Full Assignment of ab-initio Raman Spectra at Finite Temperatures Using Wannier Polarizabilities: Application to Cyclohexane Molecule in Gas Phase", *Micromachines* **2021**, *12* (10), 1212, doi:10.3390/mi12101212. (cited on pages 15 and 55)

[220] "LAMMPS Molecular Dynamics Simulator, see http://lammps.sandia.gov", . (cited on pages 16 and 111)

[221] S. Plimpton, "Fast Parallel Algorithms for Short-Range Molecular Dynamics", *J. Comput. Chem.* **1995**, *117*, 1–19, doi:10.1006/jcph.1995.1039. (cited on page 16)

[222] J. M. Martínez, L. Martínez, "Packing Optimization for Automated Generation of Complex System's Initial Configurations for Molecular Dynamics and Docking", *J. Comput. Chem.* **2003**, *24* (7), 819–825, doi:10.1002/jcc.10216. (cited on pages 16 and 111)

[223] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, "Packmol: A Package for Building Initial Configurations for Molecular Dynamics Simulations", *J. Comput. Chem.* **2009**, *30* (13), 2157–2164, doi:10.1002/jcc.21224. (cited on pages 16 and 111)

[224] "Grace, (c) 1996-2008 Grace Development Team, see `http://plasma-gate.weizmann.ac.il/grace`", . `(cited on page 16)`

[225] "Wolfram Research, Inc., Mathematica, Version 8.0, Champaign, Il (2010)", . `(cited on page 16)`

[226] T. Williams, C. Kelley, many others, "Gnuplot 4.4: An Interactive Plotting Program", `http://gnuplot.sourceforge.net/`, **2010**. `(cited on page 16)`

[227] W. Humphrey, A. Dalke, K. Schulten, "VMD – Visual Molecular Dynamics", *J. Mol. Graphics* **1996**, *14*, 33–38, doi:`10.1016/0263-7855(96)00018-5`. `(cited on page 16)`

[228] J. E. Stone, *An Efficient Library for Parallel Ray Tracing and Animation*, Master's thesis, University of Missouri-Rolla, **1998**. `(cited on page 16)`

[229] M. Tarini, P. Cignoni, C. Montani, "Ambient Occlusion and Edge Cueing for Enhancing Real Time Molecular Visualization", *IEEE T. Vis. Comput. Gr.* **2006**, *12* (5), 1237–1244, doi:`10.1109/TVCG.2006.115`. `(cited on page 16)`

[230] G. Voronoi, "Nouvelles Applications Des Paramètres Continus à La Théorie Des Formes Quadratiques. Deuxième Mémoire. Recherches Sur Les Parallélloèdres Primitifs.", *J. Reine Angew. Math.* **1908**, *134*, 198, doi:`10.1515/crll.1908.134.198`. `(cited on page 17)`

[231] N. N. Medvedev, "The Algorithm for Three-Dimensional Voronoi Polyhedra", *J. Comput. Phys.* **1986**, *67* (1), 223–229, doi:`10.1016/0021-9991(86)90123-3`. `(cited on page 17)`

[232] M. Pollak, R. Rein, "Molecular-Orbital Studies of Intermolecular Interaction Energies. II. Approximations Concerned with Coulomb Interactions and Comparison of the Two London Schemes", *J. Chem. Phys.* **1967**, *47* (6), 2045–2052, doi:`10.1063/1.1712236`. `(cited on page 18)`

[233] A. D. Becke, "A Multicenter Numerical Integration Scheme for Polyatomic Molecules", *J. Chem. Phys.* **1988**, *88* (4), 2547–2553, doi:`10.1063/1.454033`. `(cited on page 18)`

[234] P. Politzer, R. R. Harris, "Properties of Atoms in Molecules. I. Proposed Definition of the Charge on an Atom in a Molecule", *J. Am. Chem. Soc.* **1970**, *92* (22), 6451–6454, doi:`10.1021/ja00725a011`. `(cited on page 18)`

[235] E. R. Batista, S. S. Xantheas, H. Jónsson, "Multipole Moments of Water Molecules in Clusters and Ice Ih from First Principles Calculations", *J. Chem. Phys.* **1999**, *111* (13), 6011–6015, doi:`10.1063/1.479897`. `(cited on pages 18 and 19)`

[236] B. Rousseau, A. Peeters, C. V. Alsenoy, "Atomic Charges from Modified Voronoi Polyhedra", *J. Mol. Struct. THEOCHEM* **2001**, *538* (1-3), 235–238, doi:`10.1016/s0166-1280(00)00692-8`. `(cited on pages 18 and 19)`

[237] M. Swart, P. T. van Duijnen, "Atomic Radii in Molecules for Use in a Polarizable Force Field", *Int. J. Quantum Chem.* **2011**, *111* (7-8), 1763–1772, doi:`10.1002/qua.22855`. `(cited on page 18)`

[238] F. M. Richards, "The Interpretation of Protein Structures: Total Volume, Group Volume Distributions and Packing Density", *J. Mol. Biol.* **1974**, *82* (1), 1–14, doi:`10.1016/0022-2836(74)90570-1`. `(cited on pages 18 and 19)`

[239] B. J. Gellatly, J. L. Finney, "Calculation of Protein Volumes: An Alternative to the Voronoi Procedure", *J. Mol. Biol.* **1982**, *161* (2), 305–322, doi:10.1016/0022-2836(82)90155-3. (cited on page 19)

[240] A. Bondi, "Van der Waals Volumes and Radii", *J. Phys. Chem.* **1964**, *68* (3), 441–451, doi:10.1021/j100785a001. (cited on pages 20, 92, 94, and 112)

[241] R. S. Rowland, R. Taylor, "Intermolecular Nonbonded Contact Distances in Organic Crystal Structures: Comparison with Distances Expected from van der Waals Radii", *J. Phys. Chem.* **1996**, *100* (18), 7384–7391, doi:10.1021/jp953141+. (cited on pages 20, 92, 94, and 112)

[242] M. Mantina, A. C. Chamberlin, R. Valero, C. J. Cramer, D. G. Truhlar, "Consistent van der Waals Radii for the Whole Main Group", *J. Phys. Chem. A* **2009**, *113* (19), 5806–5812, doi:10.1021/jp8111556. (cited on pages 20, 92, 94, and 112)

[243] C. H. Rycroft, G. S. Grest, J. W. Landry, M. Z. Bazant, "Analysis of Granular Flow in a Pebble-Bed Nuclear Reactor", *Phys. Rev. E* **2006**, *74*, 021306, doi:10.1103/physreve.74.021306. (cited on page 21)

[244] C. H. Rycroft, "Voro++: A Three-Dimensional Voronoi Cell Library in C++", *Chaos* **2009**, *19* (4), 041111, doi:10.2172/946741. (cited on page 21)

[245] "Gaussian Cube File Format; see http://paulbourke.net/dataformats/cube/ (accessed 24th July 2018)", . (cited on pages 29 and 45)

[246] E. H. Moore, "On the Reciprocal of the General Algebraic Matrix", *Bull. Am. Math. Soc.* **1920**, *26* (9), 394–395, doi:10.1090/s0002-9904-1920-03322-7. (cited on page 33)

[247] R. Penrose, "A Generalized Inverse for Matrices", *Math. Proc. Cambridge Philos. Soc.* **1955**, *51* (3), 406–413, doi:10.1017/s0305004100030401. (cited on page 33)

[248] C. Eckart, G. Young, "The Approximation of one Matrix by Another of Lower Rank", *Psychometrika* **1936**, *1* (3), 211–218, doi:10.1007/bf02288367. (cited on page 34)

[249] G. H. Golub, W. Kahan, "Calculating the Singular Values and Pseudo-Inverse of a Matrix", *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis* **1965**, *2* (2), 205–224, doi:10.1137/0702016. (cited on page 34)

[250] G. H. Golub, C. Reinsch, "Singular Value Decomposition and Least Squares Solutions", *Numer. Math.* **1970**, *14* (5), 403–420, doi:10.1007/bf02163027. (cited on page 34)

[251] D. Hilbert, "Über die Stetige Abbildung einer Linie auf ein Flächenstück", *Math. Annal.* **1891**, *38*, 459–460, doi:10.1007/978-3-662-25726-5_1. (cited on page 36)

[252] A. R. Butz, "Alternative Algorithm for Hilbert's Space-Filling Curve", *IEEE Trans. Comp.* **1971**, *20*, 424–426, doi:10.1109/t-c.1971.223258. (cited on page 36)

[253] "Bzip2 Program, see http://www.bzip.org/ (accessed 24th July 2018)", . (cited on pages 36, 45, and 46)

[254] M. Burrows, D. J. Wheeler, "A Block-Sorting Lossless Data Compression Algorithm", Technical Report SRC-RR-124, Palo Alto, California, **1994**. (cited on page 37)

[255] C. A. R. Hoare, "Algorithm 64: Quicksort", *Commun. ACM* **1961**, *4* (7), 321, doi:10.1145/366622.366644. (cited on page 38)

[256] Tech Correspondence, Corporate, "Technical Correspondence", *Commun. ACM* **1987**, *30* (9), 792–796, doi:10.1145/30401.315747. (cited on page 38)

[257] A. H. Robinson, C. Cherry, "Results of a Prototype Television Bandwidth Compression Scheme", *Proc. IEEE* **1967**, *55* (3), 356–364, doi:10.1109/PROC.1967.5493. (cited on page 39)

[258] D. A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes", *Proc. IRE* **1952**, *40* (9), 1098–1101, doi:10.1109/JRPROC.1952.273898. (cited on page 40)

[259] C. E. Shannon, "A Mathematical Theory of Communication", *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656, doi:10.1002/j.1538-7305.1948.tb00917.x. (cited on page 40)

[260] D. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, Cambridge, England, **2003**. (cited on page 40)

[261] E. N. Gilbert, E. F. Moore, "Variable-Length Binary Encodings", *Bell Syst. Tech. J.* **1959**, *38* (4), 933–967, doi:10.1002/j.1538-7305.1959.tb01583.x. (cited on page 40)

[262] S. T. Klein, "Space- and Time-Efficient Decoding with Canonical Huffman Trees", in A. Apostolico, J. Hein (Editors), "Combinatorial Pattern Matching", Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-540-69214-0, **1997** pages 65–75, doi:10.1007/3-540-63220-4_50. (cited on page 41)

[263] W. W. Peterson, D. T. Brown, "Cyclic Codes for Error Detection", *Proc. IRE* **1961**, *49* (1), 228–235, doi:10.1109/jrproc.1961.287814. (cited on page 43)

[264] L. Laaksonen, *gOpenMol: Program for the Display and Analysis of Molecular Structures*, Centre for Scientific Computing, Espoo, Finland. (cited on page 45)

[265] "XZ File Format, see https://tukaani.org/xz/format.html (accessed 24th July 2018)", . (cited on pages 45 and 46)

[266] F. Dommert, K. Wendler, B. Qiao, L. Delle Site, C. Holm, "Generic Force Fields for Ionic Liquids", *J. Mol. Liq.* **2014**, *192*, 32–37, doi:10.1016/j.molliq.2013.09.001. (cited on page 46)

[267] "XYZ File Format, see http://www.ccl.net/chemistry/resources/messages/1996/10/21.005-dir/index.html (accessed July 24th 2018)", . (cited on page 46)

[268] "XTC File Format, see http://manual.gromacs.org/online/xtc.html (accessed 24th July 2018)", . (cited on page 46)

[269] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, "Gromacs: Fast, Flexible, and Free", *J. Comput. Chem.* **2005**, *26* (16), 1701–1718, doi:10.1002/jcc.20291. (cited on page 46)

[270] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, "Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation", *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447, doi:10.1021/ct700301q. (cited on page 46)

[271] J. M. Dickey, A. Paskin, “Computer Simulation of the Lattice Dynamics of Solids”, *Phys. Rev.* **1969**, *188* (3), 1407–1418, doi:10.1103/physrev.188.1407. (cited on page 49)

[272] R. P. Futrelle, D. J. McGinty, “Calculation of Spectra and Correlation Functions from Molecular Dynamics Data Using the Fast Fourier Transform”, *Chem. Phys. Lett.* **1971**, *12*, 285–287, doi:10.1016/0009-2614(71)85065-0. (cited on pages 49 and 118)

[273] E. Kestemont, J. van Craen, “On the Computation of Correlation Functions in Molecular Dynamics Experiments”, *J. Comput. Phys.* **1976**, *22*, 451–458, doi:10.1016/0021-9991(76)90043-7. (cited on pages 49 and 118)

[274] D. W. Noid, M. L. Koszykowski, R. A. Marcus, “A Spectral Analysis Method of Obtaining Molecular Spectra from Classical Trajectories”, *J. Chem. Phys.* **1977**, *67* (2), 404–408, doi:10.1063/1.434901. (cited on page 49)

[275] P. H. Berens, D. H. J. Mackay, G. M. White, K. R. Wilson, “Thermodynamics and Quantum Corrections from Molecular Dynamics for Liquid Water”, *J. Chem. Phys.* **1983**, *79* (5), 2375–2389, doi:10.1063/1.446044. (cited on page 49)

[276] M.-P. Gaigeot, M. Martinez, R. Vuilleumier, “Infrared Spectroscopy in the Gas and Liquid Phase from First Principle Molecular Dynamics Simulations: Application to Small Peptides”, *Mol. Phys.* **2007**, *105* (19-22), 2857–2878, doi:10.1080/00268970701724974. (cited on pages 49 and 74)

[277] J. J. Waterston, “On the Physics of Media That Are Composed of Free and Perfectly Elastic Molecules in a State of Motion”, *Abstracts of the Papers Communicated to the Royal Society of London* **1851**, *5*, 604–604, doi:10.1098/rspl.1843.0077. (cited on page 50)

[278] L. Boltzmann, “Einige Allgemeine Sätze über Wärmegleichgewicht”, *Wiener Berichte* **1871**, *63*, 679–711. (cited on page 50)

[279] L. Boltzmann, “Über die Natur der Gasmoleküle”, *Wiener Berichte* **1876**, *74*, 553–560, doi:10.1002/andp.18772360120. (cited on page 50)

[280] J. J. Waterston, J. W. Strutt, “I. On the Physics of Media That Are Composed of Free and Perfectly Elastic Molecules in a State of Motion”, *Philos. Trans. R. Soc. Lond.* **1892**, *183*, 1–79, doi:10.1098/rsta.1892.0001. (cited on page 50)

[281] N. Wiener, “Generalized Harmonic Analysis”, *Acta Math.* **1930**, *55*, 117–258, doi:10.1007/bf02546511. (cited on pages 51, 77, and 118)

[282] A. Khintchine, “Korrelationstheorie der Stationären Stochastischen Prozesse”, *Math. Annal.* **1934**, *109 (1)*, 604–615, doi:10.1007/bf01449156. (cited on pages 51, 77, and 118)

[283] M. Perić, S. D. Peyerimhoff, R. J. Buenker, “Analysis and Predictions of the Vibronic Spectrum of the Ethynyl Radical C_2H by ab initio Methods”, *Z. Phys. D: At., Mol. Clusters* **1992**, *24* (2), 177–198, doi:10.1007/bf01426704. (cited on page 52)

[284] H. Fliegl, A. Köhn, C. Hättig, R. Ahlrichs, “Ab initio Calculation of the Vibrational and Electronic Spectra of *trans*- and *cis*-Azobenzene”, *J. Am. Chem. Soc.* **2003**, *125* (32), 9821–9827, doi:10.1021/ja034433o. (cited on page 52)

[285] A. Fielicke, A. Kirilyuk, C. Ratsch, J. Behler, M. Scheffler, G. von Helden, G. Meijer, "Structure Determination of Isolated Metal Clusters via Far-Infrared Spectroscopy", *Phys. Rev. Lett.* **2004**, *93* (2), 023401, doi:10.1103/physrevlett.93.023401. (cited on page 52)

[286] C. Ratsch, A. Fielicke, A. Kirilyuk, J. Behler, G. von Helden, G. Meijer, M. Scheffler, "Structure Determination of Small Vanadium Clusters by Density-Functional Theory in Comparison with Experimental Far-Infrared Spectra", *J. Chem. Phys.* **2005**, *122* (12), 124302, doi:10.1063/1.1862621. (cited on page 52)

[287] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, T. Ziegler, "Chemistry with ADF", *J. Comput. Chem.* **2001**, *22* (9), 931–967, doi:10.1002/jcc.1056. (cited on pages 52 and 60)

[288] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, F. Weigend, "Turbomole", *WIREs Comput. Mol. Sci.* **2013**, *4* (2), 91–100, doi:10.1002/wcms.1162. (cited on pages 52 and 60)

[289] S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. M. Khah, S. K. Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. van Wüllen, V. K. Voora, F. Weigend, A. Wodyński, J. M. Yu, "TURBOMOLE: Modular Program Suite for ab initio Quantum-Chemical and Condensed-Matter Simulations", *J. Chem. Phys.* **2020**, *152* (18), 184107, doi:10.1063/5.0004635. (cited on pages 52 and 60)

[290] F. Neese, "The ORCA Program System", *WIREs Comput. Mol. Sci.* **2011**, *2* (1), 73–78, doi:10.1002/wcms.81. (cited on page 52)

[291] F. Neese, "Software Update: The ORCA Program System, Version 4.0", *WIREs Comput. Mol. Sci.* **2017**, *8* (1), doi:10.1002/wcms.1327. (cited on page 52)

[292] F. Neese, F. Wennmohs, U. Becker, C. Riplinger, "The ORCA Quantum Chemistry Program Package", *J. Chem. Phys.* **2020**, *152* (22), 224108, doi:10.1063/5.0004608. (cited on page 52)

[293] F. Perakis, L. D. Marco, A. Shalit, F. Tang, Z. R. Kann, T. D. Kühne, R. Torre, M. Bonn, Y. Nagata, "Vibrational Spectroscopy and Dynamics of Water", *Chem. Rev.* **2016**, *116* (13), 7590–7607, doi:10.1021/acs.chemrev.5b00640. (cited on pages 52, 55, and 68)

[294] M. Gastegger, J. Behler, P. Marquetand, "Machine Learning Molecular Dynamics for the Simulation of Infrared Spectra", *Chem. Sci.* **2017**, *8* (10), 6924–6935, doi:10.1039/c7sc02267k. (cited on page 52)

[295] V. Quaranta, M. Hellström, J. Behler, J. Kullgren, P. D. Mitev, K. Hermansson, "Maximally Resolved Anharmonic OH Vibrational Spectrum of the Water / ZnO($10\bar{1}0$) Interface from a High-Dimensional Neural Network Potential", *J. Chem. Phys.* **2018**, *148* (24), 241720, doi:10.1063/1.5012980. (cited on page 52)

[296] B. Kirchner, J. Blasius, L. Esser, W. Reckien, "Predicting Vibrational Spectroscopy for Flexible Molecules and Molecules with Non-Idle Environments", *Adv. Theory Simul.* **2020**, *4* (4), 2000223, doi:10.1002/adts.202000223. (cited on pages 52 and 61)

[297] J. Blasius, R. Elfgen, O. Hollóczki, B. Kirchner, "Glucose in Dry and Moist Ionic Liquid: Vibrational Circular Dichroism, IR, and Possible Mechanisms", *Phys. Chem. Chem. Phys.* **2020**, *22* (19), 10726–10737, doi:`10.1039/c9cp06798a`. `(cited on pages 52 and 61)`

[298] S. Taherivardanjani, R. Elfgen, W. Reckien, E. Suarez, E. Perlt, B. Kirchner, "Benchmarking the Computational Costs and Quality of Vibrational Spectra from ab initio Simulations", *Adv. Theory Simul.* **2021**, *5* (1), 2100293, doi:`10.1002/adts.202100293`. `(cited on pages 52 and 55)`

[299] R. Geitner, J. Kötteritzsch, M. Siegmann, R. Fritzsch, T. W. Bocklitz, M. D. Hager, U. S. Schubert, S. Gräfe, B. Dietzek, M. Schmitt, J. Popp, "Molecular Self-Healing Mechanisms between C60-Fullerene and Anthracene Unveiled by Raman and Two-Dimensional Correlation Spectroscopy", *Phys. Chem. Chem. Phys.* **2016**, *18* (27), 17973–17982, doi:`10.1039/c6cp03464k`. `(cited on page 54)`

[300] F. Latorre, S. Kupfer, T. W. Bocklitz, D. Kinzel, S. Trautmann, S. Gräfe, V. Deckert, "Spatial Resolution of Tip-Enhanced Raman Spectroscopy – DFT Assessment of the Chemical Effect", *Nanoscale* **2016**, *8* (19), 10229–10239, doi:`10.1039/c6nr00093b`. `(cited on page 54)`

[301] S. Oswald, N. A. Seifert, F. Bohle, M. Gawrilow, S. Grimme, W. Jäger, Y. Xu, M. A. Suhm, "The Chiral Trimer and a Metastable Chiral Dimer of Achiral Hexafluoroisopropanol: A Multi-Messenger Study", *Angew. Chem.* **2019**, *131* (15), 5134–5138, doi:`10.1002/ange.201813881`. `(cited on page 54)`

[302] R. Geitner, A. Gurinov, T. Huang, S. Kupfer, S. Gräfe, B. M. Weckhuysen, "Reaction Mechanism of Pd-Catalyzed »CO-Free« Carbonylation Reaction Uncovered by in situ Spectroscopy: The Formyl Mechanism", *Angew. Chem. - Int. Ed.* **2020**, *60* (7), 3422–3427, doi:`10.1002/anie.202011152`. `(cited on page 54)`

[303] K. Fiederling, M. Abasifard, M. Richter, V. Deckert, S. Gräfe, S. Kupfer, "The Chemical Effect Goes Resonant – a Full Quantum Mechanical Approach on TERS", *Nanoscale* **2020**, *12* (11), 6346–6359, doi:`10.1039/c9nr09814c`. `(cited on page 54)`

[304] D. A. Long, *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules*, John Wiley & Sons Ltd., **2002**, ISBN 978-0-471-49028-9, doi:`10.1002/0470845767`. `(cited on pages 54, 62, 63, 65, 66, and 73)`

[305] J. Ibaceta-Jaña, R. Muydinov, P. Rosado, H. Mirhosseini, M. Chugh, O. Nazarenko, D. N. Dirin, D. Heinrich, M. R. Wagner, T. D. Kühne, B. Szyszka, M. V. Kovalenko, A. Hoffmann, "Vibrational Dynamics in Lead Halide Hybrid Perovskites Investigated by Raman Spectroscopy", *Phys. Chem. Chem. Phys.* **2020**, *22* (10), 5604–5614, doi:`10.1039/c9cp06568g`. `(cited on page 55)`

[306] D. A. Long, "Intensities in Raman Spectra I. A Bond Polarizability Theory", *Proc. Math. Phys. Eng. Sci.* **1953**, *217* (1129), 203–221, doi:`10.1098/rspa.1953.0057`. `(cited on page 56)`

[307] C. D. Allemand, "Depolarization Ratio Measurements in Raman Spectrometry", *Appl. Spectrosc.* **1970**, *24* (3), 348–353, doi:`10.1366/000370270774371552`. `(cited on page 56)`

[308] L. A. Nafie, "Velocity-Gauge Formalism in the Theory of Vibrational Circular Dichroism and Infrared Absorption", *J. Chem. Phys.* **1992**, *96* (8), 5687–5702, doi:`10.1063/1.462668`. `(cited on pages 57 and 60)`

[309] L. A. Nafie, "Theory of Vibrational Circular Dichroism and Infrared Absorption: Extension to Molecules with Low-Lying Excited Electronic States", *J. Phys. Chem. A* **2004**, *108* (35), 7222–7231, doi:10.1021/jp0499124. `(cited on pages 57 and 60)`

[310] H. A. van der Vorst, "Bi-CGstab: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems", *SIAM Journal on Scientific and Statistical Computing* **1992**, *13* (2), 631–644, doi:10.1137/0913035. `(cited on page 58)`

[311] G. L. G. Sleijpen, D. R. Fokkema, "BiCGstab(l) for Linear Equations Involving Unsymmetric Matrices with Complex Spectrum", *Electron. Trans. Numer. Anal.* **1993**, *1*, 11–32. `(cited on page 58)`

[312] G. L. G. Sleijpen, H. A. van der Vorst, "Maintaining Convergence Properties of BiCGstab Methods in Finite Precision Arithmetic", *Numer. Algorithms* **1995**, *10* (2), 203–223, doi:10.1007/bf02140769. `(cited on page 58)`

[313] G. L. G. Sleijpen, H. A. van der Vorst, "Reliable Updated Residuals in Hybrid Bi-CG Methods", *Computing* **1996**, *56* (2), 141–163, doi:10.1007/bf02309342. `(cited on page 58)`

[314] J. A. Meijerink, H. A. van der Vorst, "An Iterative Solution Method for Linear Systems of Which the Coefficient Matrix is a Symmetric m-Matrix", *Math. Comput.* **1977**, *31* (137), 148–148, doi:10.1090/s0025-5718-1977-0438681-4. `(cited on page 58)`

[315] C. Adamo, V. Barone, "Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model", *J. Chem. Phys.* **1999**, *110* (13), 6158–6170, doi:10.1063/1.478522. `(cited on page 59)`

[316] V. P. Nicu, E. Debie, W. Herrebout, B. V. der Veken, P. Bultinck, E. J. Baerends, "A VCD Robust Mode Analysis of Induced Chirality: The Case of Pulegone in Chloroform", *Chirality* **2009**, *21* (1E), E287–E297, doi:10.1002/chir.20817. `(cited on page 60)`

[317] V. P. Nicu, J. Autschbach, E. J. Baerends, "Enhancement of IR and VCD Intensities Due to Charge Transfer", *Phys. Chem. Chem. Phys.* **2009**, *11* (10), 1526, doi:10.1039/b816151h. `(cited on page 60)`

[318] T. B. Freedman, X. Cao, R. K. Dukor, L. A. Nafie, "Absolute Configuration Determination of Chiral Molecules in the Solution State Using Vibrational Circular Dichroism", *Chirality* **2003**, *15* (9), 743–758, doi:10.1002/chir.10287. `(cited on page 60)`

[319] G. Magyarfalvi, G. Tarczay, E. Vass, "Vibrational Circular Dichroism", *WIREs Comput. Mol. Sci.* **2011**, *1* (3), 403–425, doi:10.1002/wcms.39. `(cited on page 60)`

[320] P. J. Stephens, F. J. Devlin, J.-J. Pan, "The Determination of the Absolute Configurations of Chiral Molecules Using Vibrational Circular Dichroism (VCD) Spectroscopy", *Chirality* **2008**, *20* (5), 643–663, doi:10.1002/chir.20477. `(cited on page 60)`

[321] C. Merten, T. P. Golub, N. M. Kreienborg, "Absolute Configurations of Synthetic Molecular Scaffolds from Vibrational CD Spectroscopy", *J. Org. Chem.* **2019**, *84* (14), 8797–8814, doi:10.1021/acs.joc.9b00466. `(cited on page 60)`

[322] L. Böselt, D. Sidler, T. Kittelmann, J. Stohner, D. Zindel, T. Wagner, S. Riniker, "Determination of Absolute Stereochemistry of Flexible Molecules Using a Vibrational Circular Dichroism Spectra Alignment Algorithm", *J. Chem. Inf. Model.* **2019**, *59* (5), 1826–1838, doi:`10.1021/acs.jcim.8b00789`. `(cited on page 60)`

[323] C. Merten, "Recent Advances in the Application of Vibrational Circular Dichroism Spectroscopy for the Characterization of Asymmetric Catalysts", *Eur. J. Org. Chem.* **2020**, *2020* (37), 5892–5900, doi:`10.1002/ejoc.202000876`. `(cited on page 60)`

[324] K. Scholten, E. Engelage, C. Merten, "Basis Set Dependence of S–O Stretching Frequencies and its Consequences for IR and VCD Spectra Predictions", *Phys. Chem. Chem. Phys.* **2020**, *22* (48), 27979–27986, doi:`10.1039/d0cp05420h`. `(cited on page 60)`

[325] D. P. Demarque, M. Kemper, C. Merten, "VCD Spectroscopy Reveals That a Water Molecule Determines the Conformation of Azithromycin in Solution", *Chem. Commun.* **2021**, *57* (33), 4031–4034, doi:`10.1039/d1cc00932j`. `(cited on page 60)`

[326] V. P. Nicu, E. J. Baerends, "On the Origin Dependence of the Angle Made by the Electric and Magnetic Vibrational Transition Dipole Moment Vectors", *Phys. Chem. Chem. Phys.* **2011**, *13* (36), 16126, doi:`10.1039/c1cp21442j`. `(cited on page 60)`

[327] T. B. Freedman, M.-L. Shih, E. Lee, L. A. Nafie, "Electron Transition Current Density in Molecules. 3. Ab initio Calculations for Vibrational Transitions in Ethylene and Formaldehyde", *J. Am. Chem. Soc.* **1997**, *119* (44), 10620–10626, doi:`10.1021/ja9701568`. `(cited on page 60)`

[328] T. B. Freedman, X. Gao, M.-L. Shih, L. A. Nafie, "Electron Transition Current Density in Molecules. 2. Ab initio Calculations for Electronic Transitions in Ethylene and Formaldehyde", *J. Phys. Chem. A* **1998**, *102* (19), 3352–3357, doi:`10.1021/jp972345i`. `(cited on page 60)`

[329] A. Abedi, N. T. Maitra, E. K. U. Gross, "Exact Factorization of the Time-Dependent Electron-Nuclear Wave Function", *Phys. Rev. Lett.* **2010**, *105* (12), 123002, doi:`10.1103/physrevlett.105.123002`. `(cited on page 60)`

[330] S. Patchkovskii, "Electronic Currents and Born–Oppenheimer Molecular Dynamics", *J. Chem. Phys.* **2012**, *137* (8), 084109, doi:`10.1063/1.4747540`. `(cited on page 60)`

[331] J. Hutter, M. Iannuzzi, "CPMD: Car–Parrinello Molecular Dynamics", *Z. Kristallogr. – Cryst. Mater.* **2005**, *220* (5-6), 549–551, doi:`10.1524/zkri.220.5.549.65080`. `(cited on page 60)`

[332] S. Abbate, G. Longhi, K. Kwon, A. Moscowitz, "The Use of Cross-Correlation Functions in the Analysis of Circular Dichroism Spectra", *J. Chem. Phys.* **1998**, *108* (1), 50–62, doi:`10.1063/1.475364`. `(cited on page 60)`

[333] J.-H. Choi, S. Hahn, M. Cho, "Amide I IR, VCD, and 2D IR Spectra of Isotope-Labeled α-Helix in Liquid Water: Numerical Simulation Studies", *Int. J. Quantum Chem.* **2005**, *104* (5), 616–634, doi:`10.1002/qua.20543`. `(cited on page 60)`

[334] J. Horníček, P. Kaprálová, P. Bouř, "Simulations of Vibrational Spectra from Classical Trajectories: Calibration with ab initio Force Fields", *J. Chem. Phys.* **2007**, *127* (8), 084502, doi:`10.1063/1.2756837`. `(cited on page 60)`

[335] K. Kwac, K.-K. Lee, J. B. Han, K.-I. Oh, M. Cho, "Classical and Quantum Mechanical / Molecular Mechanical Molecular Dynamics Simulations of Alanine Dipeptide in Water: Comparisons with IR and Vibrational Circular Dichroism Spectra", *J. Chem. Phys.* **2008**, *128* (10), 105106, doi:`10.1063/1.2837461`. `(cited on page 60)`

[336] S. Yang, M. Cho, "Direct Calculations of Vibrational Absorption and Circular Dichroism Spectra of Alanine Dipeptide Analog in Water: Quantum Mechanical / Molecular Mechanical Molecular Dynamics Simulations", *J. Chem. Phys.* **2009**, *131* (13), 135102, doi:`10.1063/1.3243078`. `(cited on page 60)`

[337] S. Jähnigen, A. Scherrer, R. Vuilleumier, D. Sebastiani, "Chiral Crystal Packing Induces Enhancement of Vibrational Circular Dichroism", *Angew. Chem. - Int. Ed.* **2018**, *57* (40), 13344–13348, doi:`10.1002/anie.201805671`. `(cited on page 61)`

[338] S. Jähnigen, D. Sebastiani, R. Vuilleumier, "The Important Role of Non-Covalent Interactions for the Vibrational Circular Dichroism of Lactic Acid in Aqueous Solution", *Phys. Chem. Chem. Phys.* **2021**, *23* (32), 17232–17241, doi:`10.1039/d1cp03106f`. `(cited on page 61)`

[339] T. Helgaker, K. Ruud, K. L. Bak, P. Jørgensen, J. Olsen, "Vibrational Raman Optical Activity Calculations Using London Atomic Orbitals", *Faraday Discuss.* **1994**, *99*, 165–180, doi:`10.1039/fd9949900165`. `(cited on page 62)`

[340] R. K. Dukor, L. A. Nafie, "Vibrational Optical Activity of Pharmaceuticals and Biomolecules", **2000**, doi:`10.1002/9780470027318.a0215`. `(cited on page 62)`

[341] K. Ruud, T. Helgaker, P. Bouř, "Gauge-Origin Independent Density-Functional Theory Calculations of Vibrational Raman Optical Activity", *J. Phys. Chem. A* **2002**, *106* (32), 7448–7455, doi:`10.1021/jp026037i`. `(cited on page 62)`

[342] A. J. Thorvaldsen, K. Ruud, K. Kristensen, P. Jørgensen, S. Coriani, "A Density Matrix-Based Quasienergy Formulation of the Kohn–Sham Density Functional Response Theory Using Perturbation- and Time-Dependent Basis Sets", *J. Chem. Phys.* **2008**, *129* (21), 214108, doi:`10.1063/1.2996351`. `(cited on page 62)`

[343] J. R. Cheeseman, M. J. Frisch, "Basis Set Dependence of Vibrational Raman and Raman Optical Activity Intensities", *J. Chem. Theory Comput.* **2011**, *7* (10), 3323–3334, doi:`10.1021/ct200507e`. `(cited on page 62)`

[344] G. Li, J. Kessler, J. Cheramy, T. Wu, M. R. Poopari, P. Bouř, Y. Xu, "Transfer and Amplification of Chirality within the »Ring of Fire« Observed in Resonance Raman Optical Activity Experiments", *Angew. Chem. - Int. Ed.* **2019**, *58* (46), 16495–16498, doi:`10.1002/anie.201909603`. `(cited on page 62)`

[345] T. Wu, G. Li, J. Kapitán, J. Kessler, Y. Xu, P. Bouř, "Two Spectroscopies in One: Interference of Circular Dichroism and Raman Optical Activity", *Angew. Chem. - Int. Ed.* **2020**, *59* (49), 21895–21898, doi:`10.1002/anie.202011146`. `(cited on page 62)`

[346] P. Guyot-Sionnest, W. Chen, Y. R. Shen, "General Considerations on Optical Second-Harmonic Generation from Surfaces and Interfaces", *Phys. Rev. B* **1986**, *33* (12), 8254–8263, doi:`10.1103/physrevb.33.8254`. `(cited on page 67)`

[347] Q. Du, R. Superfine, E. Freysz, Y. R. Shen, "Vibrational Spectroscopy of Water at the Vapor / Water Interface", *Phys. Rev. Lett.* **1993**, *70* (15), 2313–2316, doi:`10.1103/physrevlett.70.2313`. `(cited on page 67)`

[348] N. Ji, V. Ostroverkhov, C. S. Tian, Y. R. Shen, “Characterization of Vibrational Resonances of Water–Vapor Interfaces by Phase-Sensitive Sum-Frequency Spectroscopy”, *Phys. Rev. Lett.* **2008**, *100* (9), 096102, doi:10.1103/physrevlett.100.096102. (cited on page 67)

[349] M. Sovago, R. K. Campen, G. W. H. Wurpel, M. Müller, H. J. Bakker, M. Bonn, “Vibrational Response of Hydrogen-Bonded Interfacial Water is Dominated by Intramolecular Coupling”, *Phys. Rev. Lett.* **2008**, *100* (17), 173901, doi:10.1103/physrevlett.100.173901. (cited on page 67)

[350] S. Nihonyanagi, T. Ishiyama, T.-K. Lee, S. Yamaguchi, M. Bonn, A. Morita, T. Tahara, “Unified Molecular View of the Air / Water Interface Based on Experimental and Theoretical $\chi^{(2)}$ Spectra of an Isotopically Diluted Water Surface”, *J. Am. Chem. Soc.* **2011**, *133* (42), 16875–16880, doi:10.1021/ja2053754. (cited on page 67)

[351] A. Adhikari, “Accurate Determination of Complex $\chi^{(2)}$ Spectrum of the Air / Water Interface”, *J. Chem. Phys.* **2015**, *143* (12), 124707, doi:10.1063/1.4931485. (cited on page 67)

[352] M. von Domaros, Y. Liu, J. L. Butman, E. Perlt, F. M. Geiger, D. J. Tobias, “Molecular Orientation at the Squalene / Air Interface from Sum Frequency Generation Spectroscopy and Atomistic Modeling”, *J. Phys. Chem. B* **2021**, *125* (15), 3932–3941, doi:10.1021/acs.jpcb.0c11158. (cited on page 67)

[353] S. Yamaguchi, K. Shiratori, A. Morita, T. Tahara, “Electric Quadrupole Contribution to the Nonresonant Background of Sum Frequency Generation at Air / Liquid Interfaces”, *J. Chem. Phys.* **2011**, *134* (18), 184705, doi:10.1063/1.3586811. (cited on page 67)

[354] P. Guyot-Sionnest, Y. R. Shen, “Local and Nonlocal Surface Nonlinearities for Surface Optical Second-Harmonic Generation”, *Phys. Rev. B* **1987**, *35* (9), 4420–4426, doi:10.1103/physrevb.35.4420. (cited on page 67)

[355] X. Wei, S.-C. Hong, A. I. Lvovsky, H. Held, Y. R. Shen, “Evaluation of Surface vs Bulk Contributions in Sum-Frequency Vibrational Spectroscopy Using Reflection and Transmission Geometries”, *J. Phys. Chem. B* **2000**, *104* (14), 3349–3354, doi:10.1021/jp9933929. (cited on page 67)

[356] A. Morita, J. T. Hynes, “A Theoretical Analysis of the Sum Frequency Generation Spectrum of the Water Surface. II. Time-Dependent Approach”, *J. Phys. Chem. B* **2001**, *106* (3), 673–685, doi:10.1021/jp0133438. (cited on page 68)

[357] G. R. Medders, F. Paesani, “Dissecting the Molecular Structure of the Air / Water Interface from Quantum Simulations of the Sum-Frequency Generation Spectrum”, *J. Am. Chem. Soc.* **2016**, *138* (11), 3912–3919, doi:10.1021/jacs.6b00893. (cited on page 68)

[358] A. Morita, T. Ishiyama, “Recent Progress in Theoretical Analysis of Vibrational Sum Frequency Generation Spectroscopy”, *Phys. Chem. Chem. Phys.* **2008**, *10* (38), 5801, doi:10.1039/b808110g. (cited on page 68)

[359] N. K. Kaliannan, A. H. Aristizabal, H. Wiebeler, F. Zysk, T. Ohto, Y. Nagata, T. D. Kühne, “Impact of Intermolecular Vibrational Coupling Effects on the Sum-Frequency Generation Spectra of the Water / Air Interface”, *Mol. Phys.* **2019**, *118* (4), 1620358, doi:10.1080/00268976.2019.1620358. (cited on page 68)

[360] A. C. Albrecht, “On the Theory of Raman Intensities”, *J. Chem. Phys.* **1961**, *34* (5), 1476–1484, doi:10.1063/1.1701032. (cited on page 69)

[361] J. Tang, A. C. Albrecht, “Studies in Raman Intensity Theory”, *J. Chem. Phys.* **1968**, *49* (3), 1144–1154, doi:10.1063/1.1670202. (cited on page 69)

[362] A. C. Albrecht, M. C. Hutley, “On the Dependence of Vibrational Raman Intensity on the Wavelength of Incident Light”, *J. Chem. Phys.* **1971**, *55* (9), 4438–4443, doi:10.1063/1.1676771. (cited on page 69)

[363] S.-Y. Lee, E. J. Heller, “Time-Dependent Theory of Raman Scattering”, *J. Chem. Phys.* **1979**, *71* (12), 4777–4788, doi:10.1063/1.438316. (cited on page 69)

[364] E. J. Heller, “The Semiclassical Way to Molecular Spectroscopy”, *Acc. Chem. Res.* **1981**, *14* (12), 368–375, doi:10.1021/ar00072a002. (cited on page 69)

[365] E. J. Heller, R. Sundberg, D. J. Tannor, “Simple Aspects of Raman Scattering”, *J. Phys. Chem.* **1982**, *86* (10), 1822–1833, doi:10.1021/j100207a018. (cited on page 69)

[366] D. J. Tannor, E. J. Heller, “Polyatomic Raman Scattering for General Harmonic Potentials”, *J. Chem. Phys.* **1982**, *77* (1), 202–218, doi:10.1063/1.443643. (cited on page 69)

[367] W. Liu, V. Settels, P. H. P. Harbach, A. Dreuw, R. F. Fink, B. Engels, “Assessment of TD-DFT- and TD-HF-Based Approaches for the Prediction of Exciton Coupling Parameters, Potential Energy Curves, and Electronic Characters of Electronically Excited Aggregates”, *J. Comput. Chem.* **2011**, *32* (9), 1971–1981, doi:10.1002/jcc.21781. (cited on page 69)

[368] L. Jensen, L. L. Zhao, J. Autschbach, G. C. Schatz, “Theory and Method for Calculating Resonance Raman Scattering from Resonance Polarizability Derivatives”, *J. Chem. Phys.* **2005**, *123* (17), 174110, doi:10.1063/1.2046670. (cited on page 69)

[369] L. Jensen, J. Autschbach, G. C. Schatz, “Finite Lifetime Effects on the Polarizability within Time-Dependent Density-Functional Theory”, *J. Chem. Phys.* **2005**, *122* (22), 224115, doi:10.1063/1.1929740. (cited on page 69)

[370] T. Kunert, R. Schmidt, “Non-Adiabatic Quantum Molecular Dynamics: General Formalism and Case Study H_2^+ in Strong Laser Fields”, *Eur. Phys. J. D* **2003**, *25* (1), 15–24, doi:10.1140/epjd/e2003-00086-8. (cited on page 69)

[371] F. Uhlig, O. Marsalek, P. Jungwirth, “Unraveling the Complex Nature of the Hydrated Electron”, *J. Phys. Chem. Lett.* **2012**, *3* (20), 3071–3075, doi:10.1021/jz301449f. (cited on page 69)

[372] S. Andermatt, J. Cha, F. Schiffmann, J. VandeVondele, “Combining Linear-Scaling DFT with Subsystem DFT in Born–Oppenheimer and Ehrenfest Molecular Dynamics Simulations: From Molecules to a Virus in Solution”, *J. Chem. Theory Comput.* **2016**, *12* (7), 3214–3227, doi:10.1021/acs.jctc.6b00398. (cited on pages 69 and 104)

[373] M. R. Provorse, C. M. Isborn, “Electron Dynamics with Real-Time Time-Dependent Density Functional Theory”, *Int. J. Quantum Chem.* **2016**, *116* (10), 739–749, doi:10.1002/qua.25096. (cited on page 69)

[374] J. J. Goings, P. J. Lestrange, X. Li, "Real-Time Time-Dependent Electronic Structure Theory", *WIREs Comput. Mol. Sci.* **2018**, *8* (1), e1341, doi:`10.1002/wcms.1341`. `(cited on page 69)`

[375] H. Chen, J. M. Mcmahon, M. A. Ratner, G. C. Schatz, "Classical Electrodynamics Coupled to Quantum Mechanics for Calculation of Molecular Optical Properties: A RT-TDDFT / FDTD Approach", *J. Phys. Chem. C* **2010**, *114* (34), 14384–14392, doi:`10.1021/jp1043392`. `(cited on page 69)`

[376] M. Thomas, F. Latorre, P. Marquetand, "Resonance Raman Spectra of *ortho*-Nitrophenol Calculated by Real-Time Time-Dependent Density Functional Theory", *J. Chem. Phys.* **2013**, *138* (4), 044101, doi:`10.1063/1.4776218`. `(cited on pages 69 and 105)`

[377] M. Thomas, S. Mühlig, T. Deckert-Gaudig, C. Rockstuhl, V. Deckert, P. Marquetand, "Distinguishing Chemical and Electromagnetic Enhancement in Surface-Enhanced Raman Spectra: The Case of *para*-Nitrothiophenol", *J. Raman Spectrosc.* **2013**, *44* (11), 1497–1505, doi:`10.1002/jrs.4377`. `(cited on page 69)`

[378] J. Mattiat, S. Luber, "Efficient Calculation of (Resonance) Raman Spectra and Excitation Profiles with Real-Time Propagation", *J. Chem. Phys.* **2018**, *149* (17), 174108, doi:`10.1063/1.5051250`. `(cited on page 69)`

[379] M. Nonella, G. Mathias, P. Tavan, "Infrared Spectrum of *p*-Benzoquinone in Water Obtained from a QM/MM Hybrid Molecular Dynamics Simulation", *J. Phys. Chem. A* **2003**, *107* (41), 8638–8647, doi:`10.1021/jp027747r`. `(cited on page 74)`

[380] M. Schmitz, P. Tavan, "Vibrational Spectra from Atomic Fluctuations in Dynamics Simulations. I. Theory, Limitations, and a Sample Application", *J. Chem. Phys.* **2004**, *121* (24), 12233, doi:`10.1063/1.1822914`. `(cited on page 74)`

[381] M. Schmitz, P. Tavan, "Vibrational Spectra from Atomic Fluctuations in Dynamics Simulations. II. Solvent-Induced Frequency Fluctuations at Femtosecond Time Resolution", *J. Chem. Phys.* **2004**, *121* (24), 12247, doi:`10.1063/1.1822915`. `(cited on page 74)`

[382] L. Pejov, D. Spångberg, K. Hermansson, "Using MD Snapshots in ab initio and DFT Calculations: OH Vibrations in the First Hydration Shell around Li^+(aq)", *J. Phys. Chem. A* **2005**, *109* (23), 5144–5152, doi:`10.1021/jp047395j`. `(cited on page 74)`

[383] R. A. Wheeler, H. Dong, S. E. Boesch, "Quasiharmonic Vibrations of Water, Water Dimer, and Liquid Water from Principal Component Analysis of Quantum or QM/MM Trajectories", *ChemPhysChem* **2003**, *4* (4), 382–384, doi:`10.1002/cphc.200390066`. `(cited on page 74)`

[384] R. A. Wheeler, H. Dong, "Optimal Spectrum Estimation in Statistical Mechanics", *ChemPhysChem* **2003**, *4* (11), 1227–1230, doi:`10.1002/cphc.200300750`. `(cited on page 74)`

[385] A. Strachan, "Normal Modes and Frequencies from Covariances in Molecular Dynamics or Monte Carlo Simulations", *J. Chem. Phys.* **2004**, *120* (1), 1–4, doi:`10.1063/1.1635364`. `(cited on page 74)`

[386] M. Martinez, M.-P. Gaigeot, D. Borgis, R. Vuilleumier, "Extracting Effective Normal Modes from Equilibrium Dynamics at Finite Temperature", *J. Chem. Phys.* **2006**, *125* (14), 144106, doi:10.1063/1.2346678. `(cited on page 74)`

[387] P. Carbonniere, A. Dargelos, I. Ciofini, C. Adamo, C. Pouchan, "Vibrational Analysis of Glycine Radical: A Comparative ab initio Static and Dynamic Study", *Phys. Chem. Chem. Phys.* **2009**, *11* (21), 4375, doi:10.1039/b817468g. `(cited on page 74)`

[388] F. Calvo, P. Parneix, N.-T. Van-Oanh, "Finite-Temperature Infrared Spectroscopy of Polycyclic Aromatic Hydrocarbon Molecules. II. Principal Mode Analysis and Self-Consistent Phonons", *J. Chem. Phys.* **2010**, *133* (7), 074303, doi:10.1063/1.3465554. `(cited on page 74)`

[389] G. Mathias, M. D. Baer, "Generalized Normal Coordinates for the Vibrational Analysis of Molecular Dynamics Simulations", *J. Chem. Theory Comput.* **2011**, *7* (7), 2028–2039, doi:10.1021/ct2001304. `(cited on pages 74, 75, and 107)`

[390] G. Mathias, S. D. Ivanov, A. Witt, M. D. Baer, D. Marx, "Infrared Spectroscopy of Fluxional Molecules from (ab initio) Molecular Dynamics: Resolving Large-Amplitude Motion, Multiple Conformations, and Permutational Symmetries", *J. Chem. Theory Comput.* **2011**, *8* (1), 224–234, doi:10.1021/ct2006665. `(cited on pages 74, 76, and 107)`

[391] C. Eckart, "Some Studies Concerning Rotating Axes and Polyatomic Molecules", *Phys. Rev.* **1935**, *47* (7), 552–558, doi:10.1103/physrev.47.552. `(cited on page 75)`

[392] G. Danielson, C. Lanczos, "Some Improvements in Practical Fourier Analysis and Their Application to X-Ray Scattering from Liquids", *J. Franklin Inst.* **1942**, *233* (4), 365–380, doi:10.1016/s0016-0032(42)90767-1. `(cited on pages 77 and 118)`

[393] J. W. Cooley, J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series", *Math. Comput.* **1965**, *19* (90), 297–301, doi:10.1090/s0025-5718-1965-0178586-1. `(cited on pages 77 and 118)`

[394] M. Borgerding, "KISS FFT Library, see `http://sourceforge.net/projects/kissfft/` ", . `(cited on page 77)`

[395] P. Kahlig, *Some Aspects of Julius Von Hann's Contribution to Modern Climatology*, pages 1–7, American Geophysical Union (AGU), ISBN 9781118666593, **2013**. `(cited on page 78)`

[396] P. L. Polavarapu, L. Hecht, L. D. Barron, "Vibrational Raman Optical Activity in Substituted Oxiranes", *J. Phys. Chem.* **1993**, *97* (9), 1793–1799, doi:10.1021/j100111a014. `(cited on page 86)`

[397] M. C. C. Ribeiro, M. Wilson, P. A. Madden, "The Nature of the »Vibrational Modes« of the Network-Forming Liquid $ZnCl_2$", *J. Chem. Phys.* **1998**, *109* (22), 9859–9869, doi:10.1063/1.477655. `(cited on page 90)`

[398] N. Ohtori, M. Salanne, P. A. Madden, "Calculations of the Thermal Conductivities of Ionic Materials by Simulation with Polarizable Interaction Potentials", *J. Chem. Phys.* **2009**, *130* (10), 104507, doi:10.1063/1.3086856. `(cited on page 90)`

[399] F. Wu, S. Roy, A. S. Ivanov, S. K. Gill, M. Topsakal, E. Dooryhee, M. Abeykoon, G. Kwon, L. C. Gallington, P. Halstenberg, B. Layne, Y. Ishii, S. M. Mahurin, S. Dai, V. S. Bryantsev, C. J. Margulis, "Elucidating Ionic Correlations Beyond Simple Charge Alternation in Molten $MgCl_2$–KCl Mixtures", *J. Phys. Chem. Lett.* **2019**, *10* (24), 7603–7610, doi:`10.1021/acs.jpclett.9b02845`. `(cited on page 90)`

[400] G. J. Janz, R. P. T. Tomkins, C. B. Allen, J. R. Downey, G. L. Garner, U. Krebs, S. K. Singer, "Molten Salts: Volume 4, Part 2, Chlorides and Mixtures—Electrical Conductance, Density, Viscosity, and Surface Tension Data", *J. Phys. Chem. Ref. Data* **1975**, *4* (4), 871–1178, doi:`10.1063/1.555527`. `(cited on page 90)`

[401] G. J. Janz, *Thermodynamic and Transport Properties for Molten Salts: Correlation Equations for Critically Evaluated Density, Surface Tension, Electrical Conductance, and Viscosity Data*, American Chemical Societyand the American Institute of Physics for the National Bureau of Standards, **1998**. `(cited on page 90)`

[402] J. VandeVondele, M. Krack, F. R. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, "Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach", *Comput. Phys. Commun.* **2005**, *167*, 103–128, doi:`10.1016/j.cpc.2004.12.014`. `(cited on pages 90 and 112)`

[403] J. VandeVondele, J. Hutter, "An Efficient Orbital Transformation Method for Electronic Structure Calculations", *J. Chem. Phys.* **2003**, *118*, 4365–4369, doi:`10.1063/1.1543154`. `(cited on pages 90 and 112)`

[404] V. Weber, J. VandeVondele, J. Hutter, A. M. N. Niklasson, "Direct Energy Functional Minimization under Orthogonality Constraints", *J. Chem. Phys.* **2008**, *128* (8), 084113, doi:`10.1063/1.2841077`. `(cited on pages 90 and 112)`

[405] F. Schiffmann, J. VandeVondele, "Efficient Preconditioning of the Electronic Structure Problem in Large Scale ab initio Molecular Dynamics Simulations", *J. Chem. Phys.* **2015**, *142* (24), 244117, doi:`10.1063/1.4922988`. `(cited on pages 90 and 112)`

[406] R. Fletcher, C. Reeves, "Function Minimization by Conjugate Gradients", *Comput. J.* **1964**, *7*, 149–154. `(cited on page 90)`

[407] J. Kiefer, "Sequential Minimax Search for a Maximum", *Proc. Amer. Math. Soc.* **1953**, *4 (3)*, 502–506. `(cited on page 90)`

[408] M. Avriel, D. J. Wilde, "Optimality Proof for the Symmetric Fibonacci Search Technique", *Fibonacci Quart* **1966**, *4*, 265–269. `(cited on page 90)`

[409] J. P. Perdew, K. Burke, M. Ernzerhof, "Generalized Gradient Approximation Made Simple", *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868, doi:`10.1103/physrevlett.77.3865`. `(cited on page 90)`

[410] Y. Zhang, W. Yang, "Comment on »Generalized Gradient Approximation Made Simple«", *Phys. Rev. Lett.* **1998**, *80* (4), 890–890, doi:`10.1103/physrevlett.80.890`. `(cited on page 90)`

[411] S. Grimme, "Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction", *J. Comput. Chem.* **2006**, *27* (15), 1787–1799, doi:`10.1002/jcc.20495`. `(cited on page 90)`

[412] S. Grimme, J. Antony, S. Ehrlich, S. Krieg, "A Consistent and Accurate ab initio Parametrization of Density Functional Dispersion Correction (DFT–D) for the 94 Elements H–Pu", *J. Chem. Phys.* **2010**, *132*, 154104, doi:`10.1063/1.3382344`. `(cited on pages 90 and 112)`

[413] S. Grimme, S. Ehrlich, L. Goerigk, "Effect of the Damping Function in Dispersion Corrected Density Functional Theory", *J. Comput. Chem.* **2011**, *32* (7), 1456–1465, doi:`10.1002/jcc.21759`. `(cited on pages 90 and 112)`

[414] J. VandeVondele, J. Hutter, "Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases", *J. Chem. Phys.* **2007**, *127*, 114105, doi:`10.1063/1.2770708`. `(cited on pages 90 and 112)`

[415] S. Goedecker, M. Teter, J. Hutter, "Separable Dual-Space Gaussian Pseudopotentials", *Phys. Rev. B* **1996**, *54*, 1703–1710, doi:`10.1103/physrevb.54.1703`. `(cited on pages 90 and 112)`

[416] C. Hartwigsen, S. Goedecker, J. Hutter, "Relativistic Separable Dual-Space Gaussian Pseudopotentials from H to Rn", *Phys. Rev. B* **1998**, *58*, 3641–3662, doi:`10.1103/physrevb.58.3641`. `(cited on pages 90 and 112)`

[417] R. W. Kawiecki, F. J. Devlin, P. J. Stephens, R. D. Amos, N. C. Handy, "Vibrational Circular Dichroism of Propylene Oxide", *Chem. Phys. Lett.* **1988**, *145* (5), 411–417, doi:`10.1016/0009-2614(88)80200-8`. `(cited on page 91)`

[418] R. W. Kawiecki, F. J. Devlin, P. J. Stephens, R. D. Amos, "Vibrational Circular Dichroism of Propylene Oxide", *J. Phys. Chem.* **1991**, *95* (24), 9817–9831, doi:`10.1021/j100177a040`. `(cited on page 91)`

[419] M. Losada, P. Nguyen, Y. Xu, "Solvation of Propylene Oxide in Water: Vibrational Circular Dichroism, Optical Rotation, and Computer Simulation Studies", *J. Phys. Chem. A* **2008**, *112* (25), 5621–5627, doi:`10.1021/jp801996m`. `(cited on page 91)`

[420] C. Merten, J. Bloino, V. Barone, Y. Xu, "Anharmonicity Effects in the Vibrational CD Spectra of Propylene Oxide", *J. Phys. Chem. Lett.* **2013**, *4* (20), 3424–3428, doi:`10.1021/jz401854y`. `(cited on page 91)`

[421] D. Kröner, "Laser-Driven Electron Dynamics for Circular Dichroism in Mass Spectrometry: From One-Photon Excitations to Multiphoton Ionization", *Phys. Chem. Chem. Phys.* **2015**, *17* (29), 19643–19655, doi:`10.1039/c5cp02193f`. `(cited on page 91)`

[422] N. M. Kreienborg, J. Bloino, T. Osowski, C. H. Pollok, C. Merten, "The Vibrational CD Spectra of Propylene Oxide in Liquid Xenon: A Proof-of-Principle CryoVCD Study That Challenges Theory", *Phys. Chem. Chem. Phys.* **2019**, *21* (12), 6582–6587, doi:`10.1039/c9cp00537d`. `(cited on page 91)`

[423] T. D. Crawford, K. Ruud, "Coupled-Cluster Calculations of Vibrational Raman Optical Activity Spectra", *ChemPhysChem* **2011**, *12* (17), 3442–3448, doi:`10.1002/cphc.201100547`. `(cited on page 93)`

[424] A. S. Perera, J. Cheramy, C. Merten, J. Thomas, Y. Xu, "IR, Raman, and Vibrational Optical Activity Spectra of Methyl Glycidate in Chloroform and Water: The Clusters-in-a-Liquid Solvation Model", *ChemPhysChem* **2018**, *19* (17), 2234–2242, doi:`10.1002/cphc.201800309`. `(cited on page 93)`

[425] J. Šebestík, P. Bouř, “Raman Optical Activity of Methyloxirane Gas and Liquid”, *J. Phys. Chem. Lett.* **2011**, *2* (5), 498–502, doi:`10.1021/jz200108v`. `(cited on page 93)`

[426] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, T. Head-Gordon, “Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P–EW”, *J. Chem. Phys.* **2004**, *120* (20), 9665–9678, doi:`10.1063/1.1683075`. `(cited on pages 96 and 111)`

[427] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, “Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids”, *J. Am. Chem. Soc.* **1996**, *118* (45), 11225–11236, doi:`10.1021/ja9621760`. `(cited on page 96)`

[428] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, “A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model”, *J. Phys. Chem.* **1993**, *97* (40), 10269–10280, doi:`10.1021/j100142a004`. `(cited on page 96)`

[429] D. Golze, J. Hutter, M. Iannuzzi, “Wetting of Water on Hexagonal Boron Nitride@Rh(111): A QM/MM Model Based on Atomic Charges Derived for Nano-Structured Substrates”, *Phys. Chem. Chem. Phys.* **2015**, *17* (22), 14307–14316, doi:`10.1039/c4cp04638b`. `(cited on page 96)`

[430] A. Aamouche, M. Ghomi, C. Coulombeau, H. Jobic, L. Grajcar, M. H. Baron, V. Baumruk, P. Y. Turpin, C. Henriet, G. Berthier, “Neutron Inelastic Scattering, Optical Spectroscopies and Scaled Quantum Mechanical Force Fields for Analyzing the Vibrational Dynamics of Pyrimidine Nucleic Acid Bases. 1. Uracil”, *J. Phys. Chem.* **1996**, *100* (13), 5224–5234, doi:`10.1021/jp952485x`. `(cited on pages 100 and 101)`

[431] T. Bourova, G. Ten, S. Andreeva, V. Berezin, “Theoretical Analysis of Raman and Resonance Raman Spectra of Simple Bases of Nucleic Acids”, *J. Raman Spectrosc.* **2000**, *31* (8-9), 827–836, doi:`10.1002/1097-4555(200008/09)31:8/9<827::aid-jrs593>3.0.co;2-t`. `(cited on pages 101 and 102)`

[432] A. K. Abdul-Sada, A. M. Greenway, P. B. Hitchcock, T. J. Mohammed, K. R. Seddon, J. A. Zora, “Upon the Structure of Room Temperature Halogenoaluminate Ionic Liquids”, *J. Chem. Soc., Chem. Commun.* **1986**, (24), 1753, doi:`10.1039/c39860001753`. `(cited on page 107)`

[433] A. Wulf, K. Fumino, D. Michalik, R. Ludwig, “IR and NMR Properties of Ionic Liquids: Do They Tell Us the Same Thing?”, *ChemPhysChem* **2007**, *8* (15), 2265–2269, doi:`10.1002/cphc.200700508`. `(cited on page 107)`

[434] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, “Molecular Dynamics with Coupling to an External Bath”, *J. Chem. Phys.* **1984**, *81* (8), 3684–3690, doi:`10.1063/1.448118`. `(cited on page 111)`

[435] T. Schneider, E. Stoll, “Molecular-Dynamics Study of a Three-Dimensional One-Component Model for Distortive Phase Transitions”, *Phys. Rev. B* **1978**, *17* (3), 1302–1322, doi:`10.1103/physrevb.17.1302`. `(cited on page 111)`

[436] B. Dünweg, W. Paul, “Bownian Dynamics Simulations without Gaussian Random Numbers”, *Int. J. Mod. Phys. C* **1991**, *02* (03), 817–827, doi:`10.1142/s0129183191001037`. `(cited on page 111)`

[437] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, "Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes", *J. Comput. Phys.* **1977**, *23* (3), 327 – 341. `(cited on page 111)`

[438] H. C. Andersen, "Rattle: A »Velocity« Version of the Shake Algorithm for Molecular Dynamics Calculations", *J. Comput. Phys.* **1983**, *52* (1), 24 – 34, doi:`10.1016/0021-9991(83)90014-1`. `(cited on page 111)`

[439] P. Pulay, "Convergence Acceleration of Iterative Sequences. The Case of SCF Iteration", *Chem. Phys. Lett.* **1980**, *73* (2), 393–398, doi:`10.1016/0009-2614(80)80396-4`. `(cited on page 112)`

[440] P. Pulay, "Improved SCF Convergence Acceleration", *J. Comput. Chem.* **1982**, *3* (4), 556–560, doi:`10.1002/jcc.540030413`. `(cited on page 112)`

[441] R. Shepard, M. Minkoff, "Some Comments on the DIIS Method", *Mol. Phys.* **2007**, *105* (19-22), 2839–2848, doi:`10.1080/00268970701691611`. `(cited on page 112)`

[442] P. Hohenberg, W. Kohn, "Inhomogeneous Electron Gas", *Phys. Rev. B* **1964**, *136*, 864, doi:`10.1103/physrev.136.b864`. `(cited on page 112)`

[443] W. Kohn, L. Sham, "Self-Consistent Equations Including Exchange and Correlation Effect", *Phys. Rev.* **1965**, *140*, 1133, doi:`10.1103/physrev.140.a1133`. `(cited on page 112)`

[444] A. D. Becke, "Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior", *Phys. Rev. A* **1988**, *38*, 3098–3100, doi:`10.1103/physreva.38.3098`. `(cited on page 112)`

[445] C. Lee, W. Yang, R. Parr, "Development of the Colle–Salvetti Correlation Energy Formula into a Functional of the Electron Density", *Phys. Rev. B* **1988**, *37*, 785–789, doi:`10.1103/physrevb.37.785`. `(cited on page 112)`

[446] D. G. A. Smith, L. A. Burns, K. Patkowski, C. D. Sherrill, "Revised Damping Parameters for the D3 Dispersion Correction to Density Functional Theory", *J. Phys. Chem. Lett.* **2016**, *7* (12), 2197–2203, doi:`10.1021/acs.jpclett.6b00780`. `(cited on page 112)`

[447] A. D. Becke, E. R. Johnson, "A Density-Functional Model of the Dispersion Interaction", *J. Chem. Phys.* **2005**, *123* (15), 154101, doi:`10.1063/1.2065267`. `(cited on page 112)`

[448] E. R. Johnson, A. D. Becke, "A Post-Hartree–Fock Model of Intermolecular Interactions", *J. Chem. Phys.* **2005**, *123* (2), 024101, doi:`10.1063/1.1949201`. `(cited on page 112)`

[449] E. R. Johnson, A. D. Becke, "A Post-Hartree–Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections", *J. Chem. Phys.* **2006**, *124* (17), 174104, doi:`10.1063/1.2190220`. `(cited on page 112)`

[450] J. Kolafa, "Time-Reversible Always Stable Predictor-Corrector Method for Molecular Dynamics of Polarizable Molecules", *J. Comput. Chem.* **2003**, *25* (3), 335–342, doi:`10.1002/jcc.10385`. `(cited on page 112)`

[451] T. D. Kühne, M. Krack, F. R. Mohamed, M. Parrinello, "Efficient and Accurate Car–Parrinello-like Approach to Born–Oppenheimer Molecular Dynamics", *Phys. Rev. Lett.* **2007**, *98* (6), 066401, doi:`10.1103/physrevlett.98.066401`. `(cited on page 112)`

[452] H. Poincaré, "Sur le problème des trois corps et les équations de la dynamique", *Acta Math.* **1890**, *13*, 1–270. `(cited on page 116)`

[453] G. D. Birkhoff, "Proof of the Ergodic Theorem", *Proc. Natl. Acad. Sci. USA* **1931**, *17 (12)*, 656–660, doi:`10.1073/pnas.17.2.656`. `(cited on page 117)`

[454] G. D. Birkhoff, "What Is the Ergodic Theorem?", *Am. Math. Mon.* **1942**, *49 (4)*, 222–226, doi:`10.1080/00029890.1942.11991212`. `(cited on page 117)`

[455] W. Smith, "An Introduction to the Discrete Fourier Transform", *CCP5 Quarterly* **1982**, *5*, 34–41. `(cited on page 118)`

[456] G. Herzberg, *Infrared and Raman Spectra*, D. Van Nostrand, Princeton, New York, **1945**. `(cited on page 130)`

[457] P. M. Morse, "Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels", *Phys. Rev.* **1929**, *34* (1), 57–64, doi:`10.1103/physrev.34.57`. `(cited on page 130)`

[458] J. P. Dahl, M. Springborg, "The Morse Oscillator in Position Space, Momentum Space, and Phase Space", *J. Chem. Phys.* **1988**, *88* (7), 4535–4547, doi:`10.1063/1.453761`. `(cited on page 130)`

Author Index

This section contains an alphabetical list of all 1162 authors which contributed to the publications cited in this thesis *(except for M.B.)*. After each author name, the reference numbers to which this author contributed are listed. The numbers are links which directly lead to the corresponding entry in the bibliography. Indices starting with "MB" refer to the Section "Publication List of Martin Brehm".

Aamouche, A.: 430
Abasifard, M.: 303
Abbate, S.: 31, 32, 332
Abdul-Sada, A. K.: 432
Abedi, A.: 329
Abeykoon, M.: 399
Adamo, C.: 58, 315, 387
Adhikari, A.: 351
Adrian, L.: MB24
Agmon, N.: 117
Agostini, F.: 66
Ahlborn, H.: 154
Ahlrichs, R.: 284, 288
Ahmed, T.: 23
Albrecht, A. C.: 360, 361, 362
Alder, B. J.: 179
Alizadeh, V.: MB53
Allemand, C. D.: 307
Allen, C. B.: 400
Alonso-Mori, R.: 28
Alsenoy, C. V.: 236
Amit, Y.: 27
Amos, R. D.: 417, 418
Andermatt, S.: 214, 372
Andersen, H. C.: 438
Andrade, R.: 223
Andreeva, S.: 431
Antony, J.: 412
Appelhagen, A.: 142
Aristizabal, A. H.: 359
Aruoja, V.: MB29
Atkins, P. W.: 33
Autschbach, J.: 317, 368, 369
Avriel, M.: 408

Bader, R.: 181
Baer, M. D.: 144, 389, 390
Baerends, E. J.: 64, 71, 287, 316, 317, 326
Bailey, G.: 36
Bajdich, M.: 28
Bak, K. L.: 44, 339
Bakker, H. J.: 349
Balasubramani, S. G.: 289
Baldelli, S.: 159
Bamler, M.: MB29
Bandrauk, A. D.: 167
Banerjee, S.: MB54
Bani-Hashemian, M. H.: 214
Banin, U.: 27
Bardoux, R.: 110
Bargar, J. R.: 28
Baron, M. H.: 430
Barone, V.: 50, 58, 86, 89, 315, 420
Barron, L. D.: 33, 34, 35, 396
Batista, E. R.: 235
Baumruk, V.: 45, 430
Bayly, C. I.: 428
Bazant, M. Z.: 243
Bearpark, M. J.: 58
Becke, A. D.: 233, 444, 447, 448, 449
Becker, U.: 93, 292
Bednárová, L.: 45
Behler, J.: 166, 285, 286, 294, 295
Bell, A. T.: 28
Belleflamme, F.: 214
Ben, M. D.: 214
Ben-Amotz, D.: 149
Berendsen, H. J. C.: 269, 434, 437
Berens, P. H.: 275
Berezin, V.: 431
Berghold, G.: 208
Bernasconi, M.: 118
Bernreuther, M.: 181
Berova, N.: 30
Berthier, G.: 430
Besley, N. A.: 132
Bethune, I.: 214

Bickelhaupt, F. M.: 287
Biczysko, M.: 50
Bieler, N. S.: 69
Bill, E.: 80
Birgin, E. G.: 223
Birkhoff, G. D.: 453, 454
Biscarini, P.: 31
Blanch, E. W.: 70
Blasius, J.: 296, 297, MB52, MB53
Blöchl, P. E.: 90
Bloino, J.: 50, 58, 420, 422
Bocklitz, T. W.: 299, 300
Bode, A.: 181
Boechat-Roberty, H. M.: 51
Boereboom, J. M.: 77
Boesch, S. E.: 383
Bogaard, M. P.: 35
Bohle, F.: 301
Böhme, A.: MB25
Böhme, R.: 5, 7
Boltzmann, L.: 278, 279
Bondi, A.: 240
Bonn, M.: 293, 349, 350
Borgerding, M.: 394
Borgis, D.: 124, 386
Borštnik, U.: 214
Böselt, L.: 52, 322
Bouř, P.: 45, 334, 341, 344, 345, 425
Bourova, T.: 431
Boys, S. F.: 198
Braams, B. J.: 176
Brehm, M.: 181
Breier, A. A.: 104
Bröhl, A.: MB18
Brothers, E. N.: 58
Brown, D. T.: 263
Brünken, S.: 103
Brüssel, M.: MB4, MB10, MB11, MB17, MB20
Bryantsev, V.: MB49, MB54
Bryantsev, V. S.: 399
Buch, V.: 125, 158
Buckingham, A. D.: 34, 35, 62
Buenker, R. J.: 283
Bueno, J.: 2
Bühl, M.: 133
Bultinck, P.: 316
Bungartz, H.-J.: 181
Burant, J. C.: 58
Burk, J.: MB29
Burk, P.: MB29
Burke, K.: 409
Burns, L. A.: 446
Burrows, M.: 254
Bussy, A.: 214
Butman, J. L.: 352
Butz, A. R.: 252

Cabana, A.: 110
Cai, Y.: 28
Calvo, F.: 388
Cammi, R.: 58, 88, 91
Campen, R. K.: 349
Cancès, E.: 85
Cao, J.: 170, 171
Cao, X.: 318
Car, R.: 116, 119, 129
Carbonniere, P.: 387
Cardini, G.: 130
Caricato, M.: 58
Caronna, T.: 32
Carvajal, M.: 46
Castiglioni, E.: 31, 32
Cavazzoni, C.: 130
Celso, F. L.: MB46, MB50
Cerajewski, U.: MB33
Cha, J.: 372
Chabal, Y. J.: 39
Chahal, R.: MB54
Chamberland, A.: 110
Chamberlin, A. C.: 242
Chassaing, T.: 402
Chater, P.: MB50
Chatzipapadopoulos, S.: 143
Chaumont, A.: MB53
Cheeseman, J. R.: 58, 63, 70, 88, 343
Chelkowski, S.: 167
Chen, G. P.: 289
Chen, H.: 375
Chen, W.: 129, 346
Cheng, M.-J.: 28
Cheramy, J.: 344, 424, MB51
Cherry, C.: 257
Cho, M.: 333, 335, 336
Choi, J.-H.: 333
Chugh, M.: 305

Chulkov, S.: 214
Cialla, D.: 5, 7
Ciccotti, G.: 437
Cieplak, P.: 428
Cignoni, P.: 229
Ciofini, I.: 387
Clark, R.: MB30
Clark, S. J.: 216
Codescu, M.-A.: MB47
Cooley, J. W.: 393
Cooper, M.: MB24
Coriani, S.: 289, 342
Cornell, W.: 428
Cossi, M.: 58, 86, 89
Coulombeau, C.: 430
Craig, I. R.: 175
Cramer, C. J.: 242
Crawford, T. D.: 423
Cringus, D.: 83
Cybik, R.: MB12

Dahl, J. P.: 458
Dai, S.: 399, MB49
Dalke, A.: 227
Daněček, P.: 45
Danielson, G.: 392
Dargelos, A.: 387
Davis, R. C.: 28
de Oliveira, D. M.: 149
de Souza, B.: 81
Debie, E.: 316
Decka, D.: 135
Deckert, V.: 7, 8, 300, 303, 377
Deckert-Gaudig, T.: 8, 377
Delle Site, L.: 266, MB9
Demarque, D. P.: 325
der Veken, B. V.: 316
Derayea, S. M.: 24
Devlin, F. J.: 63, 88, 320, 417, 418
Devlin, J. P.: 125
di Dio, P. J.: 115, MB5
Dick, T. J.: 426
Dickey, J. M.: 271
Diedenhofen, M.: 289
Diestler, D. J.: 167
Dietzek, B.: 26, 299
Ding, F.: 58
DiNola, A.: 434
Dirin, D. N.: 305
Dommert, F.: 266
Dong, H.: 383, 384
Dooryhee, E.: 399
Doty, K. C.: 2
Dötzer, R.: 52, MB52
Downey, J. R.: 400
Dreßler, C.: MB38, MB39
Dreuw, A.: 367
Du, Q.: 347
Duboué-Dijon, E.: 149
Dukor, R. K.: 318, 340
Dumas, P.: 39
Dünweg, B.: 436
Dutta, A. K.: MB44

Eckart, C.: 248, 391
Eckhardt, W.: 181
Edmiston, C.: 200
Edvinsson, T.: 23
Egidi, F.: 58
Ehara, M.: 58
Ehrlich, S.: 412, 413
Eisermann, J.: MB42
Elfgen, R.: 297, 298
Elgabarty, H.: 177, 178
Elliott, A. B. S.: 1
Engelage, E.: 324
Engels, B.: 367
Erbacci, G.: 130
Ernzerhof, M.: 409
Escorihuela, L.: MB29
Esser, L.: 296
Ewald, P. P.: 215

Fantuzzi, F.: 51
Farias, G.: 81
Farkas, O.: 58
Faust, A.: 27
Fernández, A.: MB29
Fiederling, K.: 303
Fielicke, A.: 285, 286
Fink, R. F.: 367
Finney, J. L.: 239
Firaha, D. S.: MB12
Fletcher, R.: 406
Fliegl, H.: 284
Fligg, R.: MB14

Fokkema, D. R.: 311
Fonseca Guerra, C.: 287
Fontana, F.: 32
Forbert, H.: 131, 135
Foresman, J. B.: 58
Fossog, V.: 145
Fowler, P. W.: 62
Fox, D. J.: 58
Frank, M. S.: 289
Franzke, Y. J.: 289
Freedman, T. B.: 59, 318, 327, 328
Freitag, I.: 7
Freysz, E.: 347
Fridgen, T. D.: 4
Friebel, D.: 28
Friedrich, J.: MB12
Friedriszik, F.: 143
Frisch, M. J.: 58, 63, 88, 343
Fritzsch, R.: 299
Frömbgen, T.: MB53
Fuchs, G. W.: 104
Fukuda, R.: 58
Fumino, K.: 145, 433
Funkner, S.: 131
Furche, F.: 288, 289
Fürstner, A.: 82
Futrelle, R. P.: 272

Gabriel, S.: 88
Gaigeot, M.-P.: 122, 124, 127, 128, 132, 133, 161, 162, 163, 276, 386
Galimberti, D. R.: 162, 163
Galli, G. A.: 129, 138
Gallington, L. C.: 399, MB49
Galwas, P. A.: 62
Gao, J.: 58
Gao, X.: 328
Gao, Y.: 25
García-Ruiz, C.: 3
Garner, G. L.: 400
Gastegger, M.: 294
Gauss, J.: 103, 104
Gawrilow, M.: 301
Gehrke, S.: MB30, MB40
Geiger, F. M.: 352
Geitner, R.: 299, 302
Gellatly, B. J.: 239
Ghomi, M.: 430
Giernoth, R.: MB6, MB18
Giesen, T. F.: 104
Gilbert, E. N.: 261
Gill, S. K.: 399
Giralt, F.: MB29
Glass, C. W.: 181
Glöß, A.: 214
Goedecker, S.: 415, 416
Goerigk, L.: 413
Goings, J. J.: 58, 374
Golub, G. H.: 249, 250
Golub, T. P.: 321
Golze, D.: 214, 429, MB37
Gomperts, R.: 58
Gordon, K. C.: 1
Gordon, R. G.: 106, 108
Gräfe, S.: 299, 300, 302, 303
Grajcar, L.: 430
Green, T.: 155
Greenway, A. M.: 432
Grest, G. S.: 243
Grey, J. K.: 25
Grimme, S.: 94, 301, 411, 412, 413
Groenhof, G.: 269
Groenzin, H.: 158
Gross, E. K. U.: 66, 329
Grotjahn, R.: 289
Guidon, M.: 214
Guillot, B.: 217
Gurinov, A.: 302
Guthmuller, J.: 26, 76
Guyot-Sionnest, P.: 38, 39, 40, 41, 346, 354
Gygi, F.: 138

Haack, A.: 82
Haag, M. P.: 49, 69
Haak, J. R.: 434
Hada, M.: 58
Hager, M. D.: 299
Hahn, S.: 333
Hairer, E.: 183
Halámková, L.: 2
Halkier, A.: 44
Halstenberg, P.: 399, MB49
Hammer, N.: 181
Hammer, R.: MB37
Hammes-Schiffer, S.: 97

Han, J. B.: 335
Handy, N. C.: 417
Hänsel, K. D.: 111
Hansen, A.: 92, 93, 94
Harbach, P. H. P.: 367
Harding, M. E.: 289
Harris, R. R.: 234
Hartmann, K.: 7
Hartwigsen, C.: 416
Hasegawa, J.: 58
Hashitsume, N.: 112
Hassanali, A. A.: 177
Hasse, H.: 181
Hättig, C.: 284, 288, 289
Havenith, M.: 131, 135
Head-Gordon, T.: 426
Heaton, R. J.: 216
Hecht, L.: 396
Hehn, A.: 214
Heinecke, A.: 181
Heinrich, D.: 305
Heislbetz, S.: 98
Held, H.: 355
Helgaker, T.: 44, 339, 341
Heller, E. J.: 363, 364, 365, 366
Hellström, M.: 295
Hellweg, A.: 289
Helmich-Paris, B.: 289
Hempelmann, R.: 145
Henderson, T.: 58
Hénin, J.: 149
Henkel, S.: MB33
Henriet, C.: 430
Hermansson, K.: 295, 382
Herrebout, W.: 316
Herrmann, C.: 49, 54, 67, 68, 73
Herschel, J. F. W.: 11
Herzberg, G.: 456
Hess, B.: 269, 270
Hess, B. A.: 43, 95
Heyd, J. J.: 58
Heyden, M.: 131
Higashi, G. S.: 39
Hilbert, D.: 251
Hildebrandt, P.: 55
Hillenbrand, J.: 82
Hinderberger, D.: MB33, MB42
Hirst, J. D.: 132
Hitchcock, P. B.: 432
Hoare, C. A. R.: 255
Höfener, S.: 56
Hoffmann, A.: 305
Hohenberg, P.: 442
Hollóczki, O.: 297, MB12, MB15, MB16, MB17, MB19, MB21, MB22, MB30
Holm, C.: 266
Holmberg, N.: 214
Holzer, C.: 289
Honda, Y.: 58
Hong, S.-C.: 355
Hoover, W. G.: 187
Hore, D. K.: 157
Horn, H. W.: 426
Horníček, J.: 334
Horsch, M.: 181
Horvath, R.: 1
Hratchian, H. P.: 58
Hrenar, T.: 96
Huang, T.: 302
Huber, H.: 181
Huffman, D. A.: 258
Hug, W.: 36
Humphrey, W.: 227
Huniar, U.: 289
Hunold, J.: MB42
Hunt, J. H.: 38
Hura, G. L.: 426
Hutley, M. C.: 362
Hutter, J.: 114, 115, 139, 206, 208, 213, 214, 331, 402, 403, 404, 414, 415, 416, 429
Hynes, J. T.: 153, 356

Iacob, C.: MB6
Iannuzzi, M.: 139, 213, 214, 331, 429
Ibaceta-Jaña, J.: 305
Iftimie, R.: 123, 126, 207
Ikeda, T.: 140
Íñiguez, J.: 136
Isborn, C. M.: 373
Ishida, M.: 58
Ishii, Y.: 399
Ishikawa, Y.: 24
Ishiyama, T.: 156, 350, 358
Ivanov, A. S.: 399, MB49
Ivanov, S. D.: 390

Ivanova, T. M.: 20
Iyengar, S. S.: 58
Izmaylov, A. F.: 58
Izsák, R.: 81

Jacob, C.: 75
Jacob, C. R.: 47, 49, 69, 99, 101
Jacobi, C.: 211
Jäger, W.: 301
Jahn, S.: 216
Jähnigen, S.: 337, 338
Jakobovits, A. S.: 214
Janesko, B. G.: 58
Jang, J. H.: 159
Jang, S.: 172, 189
Janz, G. J.: 400, 401
Jensen, L.: 368, 369
Jensen, P.: 46
Ji, N.: 348
Jobic, H.: 430
Jobst, N.: 142
Johnson, E. R.: 447, 448, 449
Johnson, M. A.: 147
Jónsson, H.: 235
Jorgensen, W. L.: 427
Jørgensen, P.: 44, 339, 342
Jungwirth, P.: 141, 144, 149, 371
Jusko, P.: 103

Kabbe, G.: MB38, MB39
Kafka, A.: MB29
Kahan, W.: 249
Kahlig, P.: 395
Kahru, A.: MB29
Kaliannan, N. K.: 164, 359
Kämmer, E.: 7
Kann, Z. R.: 293
Kapitán, J.: 45, 345
Kaprálová, P.: 334
Kärger, J.: MB6
Karhan, K.: 177
Karplus, M.: 87
Kasprzyk, C. R.: 155
Kästner, J.: 51
Kaupp, M.: 289
Kawiecki, R. W.: 417, 418
Keith, T. A.: 58
Kelemen, Z.: MB19
Kelley, C.: 226
Kemper, M.: 325
Kerlé, D.: 143
Kessler, J.: 177, 344, 345
Kestemont, E.: 273
Khah, A. M.: 289
Khaliullin, R. Z.: 214
Khani, S. K.: 289
Khintchine, A.: 282
Kiefer, J.: 407
Kiewisch, K.: 49, 74
Kind, C.: 43
King-Smith, R. D.: 193
Kint, S.: 36
Kinzel, D.: 300
Kirchner, B.: 115, 150, 206, 296, 297, 298, MB1, MB2, MB3, MB4, MB5, MB6, MB7, MB8, MB9, MB10, MB11, MB12, MB13, MB14, MB15, MB16, MB17, MB19, MB20, MB21, MB22, MB23, MB30, MB40, MB52, MB53
Kirilyuk, A.: 285, 286
Kitao, O.: 58
Kittelmann, T.: 322
Klamt, A.: 84
Klein, M. L.: 184, 192
Klein, S. T.: 262
Kleinhenz, H.-G.: 181
Klene, M.: 58
Klopper, W.: 44, 56, 288
Knecht, S.: 79
Knop, S.: 10
Kobayashi, R.: 58
Koelsch, P.: 141
Kohagen, M.: MB3, MB6, MB17
Köhn, A.: 284
Kohn, W.: 442, 443
Kolafa, J.: 450
Kollman, P. A.: 428
Kopecký Jr., V.: 45
Kornilov, O.: MB47
Koszykowski, M. L.: 274
Kötteritzsch, J.: 299
Kovalenko, M. V.: 305
Krack, M.: 214, 402, 451
Krafft, C.: 9
Krebs, U.: 400

Kreienborg, N. M.: 321, 422
Kremer, F.: MB6
Kressler, J.: MB35, MB41, MB55
Krieck, S.: 26
Krieg, S.: 412
Krishnan, K. S.: 13, 14
Kristensen, K.: 342
Kröner, D.: 421
Kubas, A.: 78
Kubo, R.: 112
Kudin, K. N.: 58
Kühn, O.: 143, 146
Kühne, R.: MB29
Kühne, T. D.: 164, 165, 177, 178, 214, 218, 219, 293, 305, 359, 451
Kullgren, J.: 295
Kumarasiri, M.: 97
Kunert, T.: 370
Kupfer, S.: 26, 300, 302, 303
Kutzner, C.: 270
Kwac, K.: 335
Kwon, G.: 399
Kwon, K.: 332

Laaksonen, L.: 264
Laino, T.: 214
Lam, S.: MB54
Lanczos, C.: 392
Landry, J. W.: 243
Larsson, H. R.: 105
Lass, M.: 214
Latorre, F.: 300, 376
Layne, B.: 399
Lazzaro, A.: 214
Lebon, F.: 32
Lednev, I. K.: 2
Lee, C.: 445
Lee, E.: 327
Lee, J. J.: 56
Lee, K.-K.: 335
Lee, S.-Y.: 363
Lee, T.-K.: 350
Lehmann, S. B. C.: MB17, MB20
Lestrange, P. J.: 374
Leszczynski, J.: 53
Li, C.: 23
Li, G.: 344, 345
Li, I.: 158
Li, X.: 58, 374
Liang, W.: 58
Lightner, D. A.: 32
Lindahl, E.: 269, 270
Lindner, J.: 10, 83
Lindsay, R.: 159
Lingscheid, Y.: MB6, MB18
Lipparini, F.: 58, 103
Lisio, V. D.: MB46, MB50
Litman, Y.: 166
Liu, W.: 367
Liu, Y.: 188, 352
Lochbrunner, S.: 143
Long, D. A.: 304, 306
Longhi, G.: 32, 332
López-Ciudad, T.: 173
López-López, M.: 3
Losada, M.: 419
Loudon, R.: 19
Louie, M. W.: 28
Luber, S.: 49, 57, 68, 72, 134, 139, 152, 160, 378
Lubich, C.: 183
Lucotti, A.: 32
Ludwig, R.: 142, 143, 145, 146, 147, 148, 433
Lütgens, M.: 143
Lutter, V.: 104
Luzar, A.: MB30
Lvovsky, A. I.: 355
Lydiatt, F.: 159
Lyubartsev, A. P.: 174

Ma, M.: 23
Ma, Y.: 79
Macchiagodena, M.: MB21
Mack, F.: 289
MacKay, D.: 260
Mackay, D. H. J.: 275
Madden, P. A.: 216, 217, 397, 398
Mädler, L.: MB29
Madura, J. D.: 426
Maganas, D.: 78
Magyarfalvi, G.: 319
Mahurin, S. M.: 399, MB49
Maiman, T. H.: 16
Maitra, N. T.: 329
Malberg, F.: MB9, MB10, MB15, MB17

Maltsev, D.: MB49
Manley, M.: 6
Manolopoulos, D. E.: 175, 176
Mantina, M.: 242
Manz, J.: 167, 169
Marco, L. D.: 293
Marcus, R. A.: 274
Marenich, A. V.: 58
Margl, P. M.: 90
Margulis, C. J.: 399, MB49
Mark, A. E.: 269
Mark, F.: 55
Marquetand, P.: 294, 376, 377
Marsalek, O.: 371
Martin, R. L.: 58
Martin, T. P.: 25
Martinelli, A.: MB50
Martínez, J. M.: 222, 223
Martínez, L.: 222, 223
Martinez, M.: 127, 276, 386
Martinez-Seara, H.: 149
Martyna, G. J.: 184, 192
Marx, D.: 114, 131, 135, 168, 390
März, A.: 5, 7
Marzari, N.: 203
Mathias, G.: 131, 379, 389, 390
Mattiat, J.: 378
Maxwell, D. S.: 427
McGinty, D. J.: 272
Mcmahon, J. M.: 375
Medders, G. R.: 357
Medvedev, N. N.: 231
Meijer, G.: 285, 286
Meijerink, J. A.: 314
Mele, A.: 32, MB50
Mendola, D.: 32
Menges, F. S.: 147
Mennucci, B.: 58, 85, 88, 91
Merten, C.: 321, 323, 324, 325, 420, 422, 424
Mezey, P. G.: 199
Michalik, D.: 146, 433
Millam, J. M.: 58
Minkoff, M.: 441
Minofar, B.: 141
Mirhosseini, H.: 305
Mitev, P. D.: 295
Mohamed, F. R.: 125, 402, 451
Mohammed, T. J.: 432
Molpeceres, G.: 51
Montani, C.: 229
Montgomery Jr., J. A.: 58
Moore, E. F.: 261
Moore, E. H.: 246
Moore, G. E.: 180
Moore, P. B.: 154, 155
Morita, A.: 153, 156, 350, 353, 356, 358
Morokuma, K.: 58
Morrison, C. A.: 133
Morse, P. M.: 457
Moscowitz, A.: 332
Motschmann, H.: 141
Mroginski, M. A.: 55
Mühlig, S.: 377
Mukherjee, M.: MB44
Müller, M.: 349
Müller, T.: 214, 289
Müller-Plathe, F.: MB3
Mundy, C. J.: 144, 208, 214
Muro, C. K.: 2
Muydinov, R.: 305

Nafie, L. A.: 59, 60, 308, 309, 318, 327, 328, 340
Nagata, Y.: 293, 359
Nakai, H.: 58
Nakajima, T.: 58
Nakanishi, K.: 30
Nakatsuji, H.: 58
Naumov, S.: MB6
Nazarenko, O.: 305
Neese, F.: 48, 78, 80, 81, 82, 92, 93, 94, 290, 291, 292
Neipert, C.: 155
Neugebauer, J.: 43, 49, 64, 73, 74, 77, 95
Neumann, J.: 148
Nguyen, B. D.: 289
Nguyen, P.: 419
Nibbering, E. T. J.: MB47
Nicu, V. P.: 64, 71, 316, 317, 326
Niehues, G.: 135
Niemann, T.: 147, 148
Niessen, W.: 201
Niethammer, C.: 181
Nihonyanagi, S.: 350
Niklasson, A. M. N.: 404

Nilsson, A.: 28
Noak, J.: 78
Noid, D. W.: 274
Nonella, M.: 379
Normand, J.: 58
Nose, S.: 185, 186
Nørskov, J. K.: 28
Nyulászi, L.: MB19

Ochterski, J. W.: 58
Ogliaro, F.: 58
Oh, K.-I.: 335
Ohto, T.: 359
Ohtori, N.: 398
Ojha, D.: 164, 165
Oliveira, R. R.: 51
Olschewski, M.: 10
Olsen, J.: 339
Ortiz, J. V.: 58
Osowski, T.: 422
Ostroverkhov, V.: 348
Oswald, S.: 301
Ovander, L. N.: 18
Overbeck, V.: 146
Ovsyannikov, R. I.: 46
Owens, N. A.: 22
Oyama, Y.: 24

Pabst, H.: 214
Paesani, F.: 357
Pagliai, M.: 130
Pahlow, S.: 7
Palivec, V.: 149
Pan, J.-J.: 320
Panda, S. K.: 75
Panek, P. T.: 99, 101
Parker, S. M.: 289
Parneix, P.: 388
Parr, R.: 445
Parrinello, M.: 116, 118, 120, 125, 130, 168, 190, 197, 204, 205, 208, 402, 451
Partovi-Azar, P.: 177, 218, 219
Paschek, D.: 148
Pasinszki, T.: MB19
Paskin, A.: 271
Pasquarello, A.: 119, 137
Patchkovskii, S.: 330
Patkowski, K.: 446
Paul, W.: 436
Pawłowski, F.: 44
Pecul, M.: 53
Peeters, A.: 236
Pejov, L.: 382
Peng, B.: 58
Penrose, R.: 247
Pensado, A. S.: MB7, MB8, MB10, MB13, MB15, MB17, MB20
Perakis, F.: 293
Peralta, J. E.: 58
Perdew, J. P.: 409
Perera, A. S.: 424
Perić, M.: 283
Perlt, E.: 289, 298, 352, MB11
Perry, A.: 154, 155
Peschel, C.: MB28
Peterson, W. W.: 263
Petersson, G. A.: 58
Petrone, A.: 58
Petrov, M.: 141
Peyerimhoff, S. D.: 283
Pezzotti, S.: 162, 163
Pietro, M. E. D.: MB50
Pinter, A.: 22
Pipek, J.: 199
Pisani, L.: 32
Pitera, J. W.: 426
Plessl, C.: 214
Plimpton, S.: 221
Pluhařová, E.: 144
Poincaré, H.: 452
Pokhrel, S.: MB29
Polavarapu, P. L.: 71, 396
Polet, W.: 145
Politzer, P.: 234
Pollak, M.: 232
Pollok, C. H.: 422
Polyakov, E. A.: 174
Poopari, M. R.: 344
Popelier, P. L. A.: 70
Popp, J.: 5, 7, 9, 26, 73, 299
Porter, M. D.: 22
Postma, J. P. M.: 434
Pouchan, C.: 387
Presselt, M.: 73
Provorse, M. R.: 373

Pshenichnikov, M. S.: 83
Pulay, P.: 439, 440
Pulst, M.: MB35, MB41
Putrino, A.: 120, 197
Pylaeva, S.: MB32

Qiao, B.: 266
Quaranta, V.: 295
Quitián-Lara, H. M.: 51

Rabelo de Moraes, I.: 26
Radicke, J.: MB41, MB55
Raghavachari, K.: 58
Rallo, R.: MB29
Ramakrishnan, R.: 100
Raman, C. V.: 13
Ramírez, R.: 173
Ramzan, M.: MB10
Ranasinghe, D.: 58
Rappoport, D.: 289
Ratner, M. A.: 375
Ratsch, C.: 285, 286
Rau, S.: 26, 73
Rauhut, G.: 96, 98, 100, 102
Rea, D. G.: 17
Reckien, W.: 296, 298
Redlich, B.: 103
Reeves, C.: 406
Rega, N.: 58, 89
Reiher, M.: 43, 47, 49, 54, 57, 67, 68, 69, 73, 74, 79
Rein, R.: 232
Reinsch, C.: 250
Reiter, K.: 289
Rendell, A. P.: 58
Resta, R.: 129, 194, 195, 196, 209
Ribeiro, M. C. C.: 397
Richards, F. M.: 238
Richmond, G. L.: 157, 158
Richter, M.: 303
Rickmeyer, C.: MB25
Riniker, S.: 52, 322
Riplinger, C.: 292, MB44
Rizzo, A.: 53
Robb, M. A.: 58
Robinson, A. H.: 257
Roca, C. P.: MB29
Rockstuhl, C.: 377
Rokni, M.: 21
Römelt, C.: 80
Romero, A. H.: 208
Roos, A. H.: MB33
Roos, E.: MB45, MB55, MB56
Rosado, P.: 305
Rossi, M.: 166
Roth, C.: 142, 143
Rousseau, B.: 236
Rowland, R. S.: 241
Roy, S.: 289, 399, MB49, MB54
Rückert, M.: 289
Ruedenberg, K.: 200
Russina, O.: MB21, MB46, MB50
Ruud, K.: 67, 339, 341, 342, 423
Rybkin, V. V.: 214
Ryckaert, J.-P.: 437
Rycroft, C. H.: 243, 244

Sabia, R.: 32
Saddiq, G.: MB27
Salanne, M.: 161, 217, 398
Salzmann, S.: 52
Sangoro, J.: MB6
Sanwald, K. E.: 28
Scalmani, G.: 58, 89
Scarbath-Evers, L.: MB37
Schade, R.: 214
Schatz, G. C.: 368, 369, 375
Schawlow, A. L.: 15
Scheffler, M.: 285, 286
Schenk, S.: 49
Schenter, G. K.: 214
Scherer, J. R.: 36
Scherrer, A.: 65, 66, 151, 337
Schettino, V.: 130
Schiffmann, F.: 213, 214, 372, 405
Schlegel, H. B.: 58
Schlemmer, S.: 103
Schlögl, R.: 78
Schmid, R.: 90
Schmidt, B.: 144
Schmidt, R.: 370
Schmitt, M.: 5, 26, 73, 299
Schmitz, G.: 289
Schmitz, M.: 380, 381

Schneider, T.: 435
Scholten, K.: 324
Schöppke, M.: MB17, MB20
Schubert, U. S.: 299
Schulten, K.: 227
Schütt, O.: 214
Schüürmann, G.: 84, MB24, MB25, MB29
Schwaab, G.: 135
Scott-Fordsmand, J.: MB29
Scuseria, G. E.: 58
Sebastiani, D.: 65, 66, 151, 197, 337, 338, MB27, MB28, MB31, MB32, MB35, MB37, MB38, MB39, MB41, MB47, MB55, MB56
Šebestík, J.: 425
Seddon, K. R.: 432
Seewald, P.: 214
Seifert, N. A.: 301
Seise, B.: 7
Seitsonen, A. P.: MB1, MB7
Senn, H. M.: 90
Settels, V.: 367
Shaabani, F.: MB41
Shaik, M. S.: 70
Shalit, A.: 293
Sham, L.: 443
Shannon, C. E.: 259
Sharma, M.: 129
Sharma, S.: MB49
Shen, Y. R.: 37, 38, 42, 346, 347, 348, 354, 355
Shepard, R.: 441
Sherrill, C. D.: 446
Shih, M.-L.: 327, 328
Shimizu, H.: 107, 109
Shiratori, K.: 353
Shorygin, P. P.: 20
Shultz, M. J.: 158
Sidler, D.: 322
Sieffert, N.: 133
Siegmann, M.: 299
Sierka, M.: 288, 289
Sihtmäe, M.: MB29
Sikk, L.: MB29
Silvestrelli, P. L.: 118, 204, 205, 210
Simon, C.: 217
Singer, S. K.: 400
Slawik, C.: MB25
Sleijpen, G. L. G.: 311, 312, 313
Smekal, A.: 12
Smith, D. G. A.: 446
Smith, W.: 455
Snijders, J. G.: 287
Sokaras, D.: 28
Sonnabend, A.: MB24
Sonnenberg, J. L.: 58
Sonntag, F.: MB24
Sorella, S.: 209
Sorensen, P. B.: MB29
Souza, I.: 136
Sovago, M.: 349
Space, B.: 154, 155
Spångberg, D.: 382
Spanu, L.: 138
Sprik, M.: 122, 124, 161
Springborg, M.: 458
Spura, T.: 177, 178
Stange, P.: 145
Stark, A.: MB8, MB10, MB12, MB13, MB20, MB21
Staroverov, V. N.: 58
Stein, F.: 214
Steiner, S.: 52
Stephens, P. J.: 29, 61, 63, 88, 320, 417, 418
Stoffels, A.: 103
Stohner, J.: 322
Stoll, E.: 435
Stone, J. E.: 228
Strachan, A.: 385
Strate, A.: 147, 148
Stritzinger, M.: 52
Strutt, J. W.: 280
Suarez, E.: 298
Suhm, M. A.: 56, 301
Suhr, H.: 37
Sulpizi, M.: 161
Sun, J.: 131, 135
Sundberg, R.: 365
Superchi, S.: 32
Superfine, R.: 347
Swalina, C.: 97
Swart, M.: 237
Swope, W. C.: 426
Szyszka, B.: 305

Tabacchi, G.: 214
Tahara, T.: 350, 353
Taherivardanjani, S.: 298, MB52
Taillefumier, M.: 214
Tämm, K.: MB29
Tamm, T.: MB29
Tang, F.: 293
Tang, J.: 361
Tannor, D. J.: 365, 366
Tapavicza, E.: 289
Tarbuck, T.: 158
Tarczay, G.: 319
Tarini, M.: 229
Tavan, P.: 379, 380, 381
Taylor, R.: 241
te Velde, G.: 287
Ten, G.: 431
Teter, M.: 415
Tew, D. P.: 289
Thar, J.: MB1, MB3, MB7
Theil, F.: 5
Thiel, W.: 46, 55
Thomas, A. K.: 25
Thomas, J.: 424
Thomas, J. W.: 207
Thomas, M.: 113, 150, 376, 377, MB14, MB16, MB19, MB21, MB22, MB23, MB26, MB34, MB36, MB40, MB48
Thorvaldsen, A. J.: 342
Throssell, K.: 58
Tian, C. S.: 348
Tirado-Rives, J.: 427
Tobias, D. J.: 192, 352
Toda, M.: 112
Tomasi, J.: 58, 85, 88, 91
Tomkins, R. P. T.: 400
Tommasini, M.: 32
Topsakal, M.: 399
Torre, R.: 293
Torricelli, F.: 32
Townes, C. H.: 15
Toyota, K.: 58
Träger, J.: MB33
Trautmann, S.: 300
Triolo, A.: MB21, MB46, MB50
Tripathi, D.: MB44
Trucks, G. W.: 58
Truhlar, D. G.: 242
Trunschke, A.: 78
Tsujino, H.: 24
Tuckerman, M. E.: 121, 123, 126, 184, 188, 190, 192, 207
Tukey, J. W.: 393
Turpin, P. Y.: 430
Turq, P.: 217

Uhlemann, U.: 73
Uhlig, F.: 371
Umari, P.: 137
Uno, T.: 24

Valero, R.: 242
Valiullin, R.: MB6
van Gisbergen, S. J. A.: 287
van Craen, J.: 273
van der Spoel, D.: 269, 270
van der Vorst, H. A.: 310, 312, 313, 314
van Duijnen, P. T.: 237
van Gastel, M.: 80, 82
van Gunsteren, W. F.: 434
van Hemert, M. C.: 77
van Wüllen, C.: 289
Van-Oanh, N.-T.: 388
Vanderbilt, D.: 136, 193, 203
VandeVondele, J.: 213, 214, 372, 402, 403, 404, 405, 414
Vass, E.: 319
Verevkin, S. P.: 148
Verlet, L.: 182
Villani, C.: 32
Vöhringer, P.: 10, 83, MB14
Voigt, T.: MB4
von Domaros, M.: 352, MB11, MB30
von Helden, G.: 285, 286
Voora, V. K.: 289
Voronoi, G.: 230
Vorontsov-Velyaminov, P. N.: 174
Voth, G. A.: 170, 171, 172, 189
Vrabec, J.: 181
Vrbka, L.: 141
Vreven, T.: 58
Vuilleumier, R.: 65, 66, 124, 127, 151, 217, 276, 337, 338, 386

Wagner, A.: MB24
Wagner, M. R.: 305

Wagner, T.: 322
Wainwright, T. E.: 179
Walker, D. S.: 157
Wan, Q.: 138
Wang, H.: 117
Wanner, G.: 183
Wannier, G. H.: 202
Wassermann, T. N.: 56
Watermann, T.: MB27
Waterston, J. J.: 277, 280
Watkins, M.: 214
Weber, H.: MB8, MB13, MB17, MB21, MB22
Weber, K.: 5, 7
Weber, V.: 214, 404
Weckhuysen, B. M.: 302
Wei, X.: 355
Weigend, F.: 288, 289
Weiß, M.: MB43, MB47
Welton, T.: MB30
Wendler, K.: 266, MB9
Weng, T.-C.: 28
Wennmohs, F.: 92, 93, 94, 292
Werner, H.-J.: 96
Weyhermüller, T.: 80
Weymuth, T.: 49
Wheeler, D. J.: 254
Wheeler, R. A.: 383, 384
White, G. M.: 275
Widdra, W.: MB37
Wiebeler, H.: 359
Wiener, N.: 281
Wild, M.: MB12
Wilde, D. J.: 408
Wilhelm, J.: 214
Williams, T.: 226
Williams-Young, D.: 58
Wilson, K. R.: 275
Wilson, M.: 397
Windiks, R.: 191
Wise, A. M.: 28
Witsch, D.: 104
Witt, A.: 390
Wittler, K.: 145
Wodyński, A.: 289
Wolff, S. K.: 64
Wondrousch, D.: MB24
Woody, R. W.: 30
Wu, F.: 399, MB49
Wu, T.: 344, 345
Wulf, A.: 433
Wurpel, G. W. H.: 349

Xantheas, S. S.: 235
Xu, Y.: 301, 344, 345, 419, 420, 424, MB51

Yamada, K. M. T.: 104
Yamaguchi, S.: 350, 353
Yamashita, T.: 24
Yang, S.: 336
Yang, W.: 410, 445
Yang, Y.: MB51
Yanovska, L. A.: 20
Yatsiv, S.: 21
Ye, S.: 80
York, D. M.: 87
Young, G.: 248
Yu, J. M.: 289
Yuan, K.-J.: 167
Yurchenko, S. N.: 46

Zahn, S.: MB17
Zaitsau, D. H.: 148
Zakrzewski, V. G.: 58
Zedler, L.: 26
Zeng, H. J.: 147
Zentel, T.: 146
Zhang, Y.: 410
Zhao, L. L.: 368
Zhao, W.: MB3
Zheng, G.: 58
Zhu, J.: 23
Zhu, X. D.: 37
Zhu, Z.: 121
Ziegler, B.: 102
Ziegler, T.: 90, 287
Zindel, D.: 322
Zora, J. A.: 432
Zukowski, S. R.: 149
Zysk, F.: 359

List of Figures

1.3.1 Snapshot of atomistic simulation 9

2.1.1 Two-dimensional illustration of Voronoi tessellation 18
2.1.2 Separation plane between Voronoi sites with radii 20
2.1.3 Radical Voronoi tessellation of benzene 21
2.1.4 Voronoi and Wannier IR and Raman spectra 24
2.1.5 Frame times for BOMD with Wannier localization 25
2.1.6 Wannier-based IR spectrum of benzene with artificial peaks 26
2.1.7 Wannier localization and normal mode for benzene 27
2.1.8 Wannier-based noisy Raman spectrum of benzene 28
2.2.1 Compressing volumetric data 29
2.2.2 Illustration of different extrapolation polynomials 31
2.2.3 Illustration of grid points in the volumetric extrapolation 32
2.2.4 Illustration of Hilbert curve in 2D and 3D 35
2.2.5 Example of the Burrows–Wheeler transformation 37
2.2.6 Example of the Move-to-Front transformation 38
2.2.7 Logo of the bqb file format 43
2.2.8 Comparing compression ratio/timing for a volumetric trajectory . . . 45
2.2.9 Comparing compression ratio/timing for a position trajectory 47
2.3.1 Power spectrum of 20 harmonic oscillators 50
2.6.1 Schematic example of classical approach 59
2.8.1 Schematic approach to ROA electromagnetic moments 64
2.10.1 Schematic simulation protocol for resonance Raman 70
2.10.2 Fourier transform of dipole moment step response 70
2.10.3 Dynamic polarizability at different times 71
2.13.1 Logo of TRAVIS 81
2.13.2 Command line interface of TRAVIS 81
2.13.3 Example of Sankey diagram created with TRAVIS 82

3.1.1 Infrared spectra of methanol in gas and liquid 84
3.2.1 Molecular structure of *(R)*-propylene oxide 86
3.2.2 Depolarized Raman spectrum of propylene oxide 86
3.2.3 Depolarization ratio of propylene oxide 87
3.2.4 Snapshot of propylene oxide simulation cell 88
3.2.5 Comparison of Raman spectra for $MgCl_2$/KCl 89
3.2.6 Snapshot of $MgCl_2$ simulation cell 90
3.3.1 Predicted VCD spectrum of liquid propylene oxide 91
3.4.1 Predicted ROA spectrum of liquid propylene oxide 93
3.4.2 Molecular structure of NALC 95

3.4.3 Predicted and experimental ROA spectra of NALC 95
3.4.4 Snapshot of NALC in water cell 96
3.4.5 Low-energy conformations of NALC 97
3.4.6 Restrained dihedral angles in NALC 97
3.4.7 Overlay of NALC structures along trajectory 98
3.4.8 ROA spectra of six NALC conformers 99
3.5.1 Uracil integrates into hydrogen bond network 100
3.5.2 Predicted resonance Raman spectrum of uracil in water 101
3.5.3 Predicted excitation profile of uracil in water 102
3.5.4 Snapshot of uracil in water simulation 103
3.5.5 Predicted resonance Raman spectrum of *o*-nitrophenol 105
3.5.6 Predicted excitation profile of *o*-nitrophenol 106
3.6.1 Molecular structure of [EMIm][OAc] 107
3.6.2 Normal mode spectra of C–H stretching in [EMIm][OAc] 108
3.6.3 Reference structures for $[\mathrm{EMIm}]^+$ 109
3.6.4 Snapshot of [EMIm][OAc]/water simulation cell 110

A.1 Power spectra of one-dimensional oscillators 130

A.2 The "AK Brehm" at the STC 2019 in Rostock 199

List of Abbreviations

$\mathbb{R}$	Set of real numbers
$\mathbb{Z}$	Set of integer numbers
δ	Dirac delta distribution
$\overline{f}$	Complex conjugate of f
$\mathbf{x}$	A vector *(denoted by lowercase bold letter)*
$\mathbf{A}$	A matrix *(denoted by uppercase bold letter)*
$\dot{x}$	Time derivative of x
$\mathcal{F}$	Fourier transform
$\cdot$	Vector dot product
$\times$	Vector cross product
$\langle\cdot\rangle$	Average of some quantity
∇	Gradient operator
Δ	Laplace operator, $\Delta = \nabla^2$
$\sphericalangle$	Angle between two vectors
$\mathrm{Re}(\cdot)$	Real part of complex value
$\mathrm{Im}(\cdot)$	Imaginary part of complex value
a. u.	Atomic units
AAT	Atomic axial tensor
ACF	Autocorrelation function
ADF	Amsterdam density functional *(software)*
AIMD	*Ab initio* molecular dynamics
APT	Atomic polar tensor
ASPC	Always stable predictor–corrector
BLYP	Becke–Lee–Yang–Parr *(DFT functional)*
BOMD	Born–Oppenheimer molecular dynamics
bqb	The bqb lossless compression format
BSD	Berkeley software distribution

CA	Complete adiabatic correction
CCF	Cross-correlation function
CG	Conjugate gradient method
CMD	Centroid molecular dynamics *(a variant of PIMD)*
COM	Center of mass
CPMD	Car–Parrinello molecular dynamics
CPU	Central processing unit
CRC-32	Cyclic redundancy check
DCPin	In-phase dual circularly polarized radiation
DCPout	Out-of-phase dual circularly polarized radiation
DFPT	Density functional perturbation theory
DFT	Density functional theory
DIIS	Direct inversion in iterative subspace
EMD	Ehrenfest molecular dynamics
$[\mathrm{EMIm}]^+$	1-Ethyl-3-methylimidazolium cation
eV	Electron volt *(energy unit)*
FFT	Fast Fourier transform
FP32	Single precision floating point format (32 bits)
FT	Fourier transform
GGA	Generalized gradient approximation
GNU GPL	GNU general public license
GPW	Gaussian plane waves approach
GTH	Goedecker–Teter–Hutter pseudopotentials
HDF5	Hierarchical data format 5
HF	Hartree–Fock
ICP	Incident circularly polarized radiation
IL	Ionic Liquid
INMA	Instantaneous normal mode analysis
IR	Infrared
JPEG	Joint photographic experts group

LAMMPS	Large-scale atomic/molecular massively parallel simulator *(software)*
LJ	Lennard-Jones *(interatomic potential)*
LR-TDDFT	Linear response time-dependent density functional theory
MC	Monte Carlo
MD	Molecular dynamics
MFPT	Magnetic field perturbation theory
MLWCs	Maximally localized Wannier centers
MLWFs	Maximally localized Wannier functions
MO	Molecular orbital
MP2	Second-order Møller–Plesset perturbation theory
MP3	MPEG audio layer III, an audio file format
MTS	Multiple time step approach
NALC	*N*-Acetyl-L-cysteine
NHC	Nosé–Hoover chain thermostat
NpT	The isothermal–isobaric ensemble
NVE	The microcanonical ensemble
NVPT	Nuclear velocity perturbation theory
NVT	The canonical ensemble
$[\mathrm{OAc}]^-$	Acetate anion
ODE	Ordinary differential equation
OPLS–AA	Optimized potentials for liquid simulations – All-atom *(force field)*
OT	Orbital transformation *(algorithm in CP2k)*
PBC	Periodic boundary conditions
PBE	Perdew–Burke–Ernzerhof *(DFT functional)*
PDE	Partial differential equation
PDF	Portable document format
PIMD	Path integral molecular dynamics
PIM	Polarizable ion model
PMA	Principal mode analysis

PPPM	Particle-particle particle-mesh method
QS	Quickstep *(algorithm in CP2k)*
revPBE	Revised Perdew–Burke–Ernzerhof *(DFT functional)*
RMSD	Root mean square deviation
RNA	Ribonucleic acid
ROA	Raman optical activity
RPMD	Ring polymer molecular dynamics *(a variant of PIMD)*
RT-TDDFT	Real-time time-dependent density functional theory
RTP	Real-time propagation
Ry	Rydberg *(energy unit)*
SCF	Self-consistent field method
SCP	Scattered circularly polarized radiation
SERS	Surface-enhanced Raman spectroscopy
SFG	Sum frequency generation
sinc	Sinus cardinalis
SVD	Singular value decomposition
TCF	Time-correlation function formalism
TDDFT	Time-dependent density functional theory
TERS	Tip-enhanced Raman spectroscopy
TIP4P–EW	A water force field
TRAVIS	Trajectory analyzer and visualizer *(software)*
UV	Ultraviolet light
VCD	Vibrational circular dichroism
VCI	Vibrational configuration interaction
vDOS	Vibrational density of states
vdW	Van der Waals *(atom radii)*
VMD	Visual molecular dynamics *(software)*
VPT2	2nd order vibrational perturbation theory
VSCF	Vibrational self consistent field theory

Index

AAT, *see* Atomic axial tensor
ab initio molecular dynamics, 7
Absolute configuration, 3
Absorption, 71
Absorption coefficient, 52
Acetate, 9, 45, 46, 82, 107
AIMD, *see ab initio* molecular dynamics
Always stable predictor–corrector, 112
Anharmonic effects, 7, 52, 85, 103, 130
Anisotropy, 55, 72
APT, *see* Atomic polar tensor
Aromaticity, 26, 100
Artificial bands, 26
ASPC, *see* Always stable predictor–corrector
Atomic axial tensor, 60
Atomic partial charges, 23
Atomic polar tensor, 60
Atomic radii, 23
Autocorrelation, 49, 52, 53, 72, 77, 79, 118

Backbone, 97
Backscattering, 95
Barostat
  Nosé–Hoover, 111
Base-2 bijective numeration, 39
Becke–Johnson damping, 112
Benzene, 26
Berendsen thermostat, 111
Berry phase, 13, 53, 89
BiCGstab, 58
Biconjugate gradient stabilized method, 58
Bijective numeration, 39
Binary tree, 40
BLYP functional, 112
Born–Oppenheimer approximation, 57
Boys–Foster localization, 13
bqb format, 30, 43, 70, 104, 112
Breathing mode, 87
Burrows–Wheeler transformation, 37

C_{2v} surface, 68
Canonical Huffmann coding, 41
Car–Parrinello molecular dynamics, 7, 52
CCl_4, 84
Cellulose, 82
Center of mass, 22, 61, 63
CH_4, 87
Charge transfer, 24
Charge variance, 23
Combination bands, 7, 52, 130
Commutator trick, 61, 65, 79
Complete adiabatic correction, 60
Complex correlation, 72
Compression of volumetric data, 29
Compression ratio, 46, 47
Conformers, 96, 99
Conjugate gradient, 90
Conservative field, 57
Conservative system, 116
Conserved quantity, 12
Continuity equation, 57
Continuum solvation, 6
Convolution theorem, 120
Coordinate origin, 22, 61, 63
COSMO, 6
Coulomb solver, 111
CP2k, 16, 23, 25, 88, 90, 96, 103, 104, 112
CPMD, *see* Car–Parrinello molecular dynamics
Crazy angle algorithm, 14, 25
CRC-32 code, 43
Cross-correlation, 13, 53, 55, 60, 65, 72, 74, 76, 77, 79, 120
Cross-correlation theorem, 77, 120
Cube file, 29, 45, 104, 112
Cutoff radius, 111
Cyclohexatriene, 27

Density, 90, 103, 111
Depolarization ratio, 56, 87
Depolarized band, 87
Depolarized Raman spectrum, 56, 73, 86, 89
Diagonalization, 14, 25, 75
Dihedral angle, 97
DIIS, *see* Direct inversion in iterative subspace
Dipole moment, 52, 60, 62, 68, 70, 89
Dipole–dipole interaction tensor, 15
Direct inversion in iterative subspace, 112
Dispersion, 71
Dispersion correction, 90, 112
Dynamic polarizability, 69, 70

Eckart frame, 75
Eigenvalue problem, 5, 74
Electric current, 57, 61
Electronic excitation, 2, 69, 100
[EMIm][OAc], 9, 45, 46, 82, 107
Enantiomeric excess, 3
Energy conservation, 10
Ensemble
  NpT, 90, 103, 111
  NVE, 11, 111
  NVT, 11, 88, 90, 111, 112
Ensemble average, 11, 79, 115, 117
Entropy, 7
  information entropy, 30, 31
Entropy encoding, 36
Equation of motion, 4, 10–12, 121
Equilibration, 111
Equipartition theorem, 50, 74
Ergodic theorem, 117
Ergodicity, 11, 115
Euclidean space, 17
Ewald summation, 15
Exact factorization, 60
Excitation profile, 73, 102, 106
Expectation value, 13
Exponential window function, 78
Extrapolation, 30, 31

Fast Fourier transform, 77, 118
FFT, *see* Fast Fourier transform
Finite differences, 15, 27, 58, 63, 78, 88, 96
Force field MD, 90
Fourier transform, 49, 52, 55, 60, 65, 70, 74, 77, 118
Fresnel factor, 67
Functional
  BLYP, 112
  PBE, 90
  PBE0, 59
  revPBE, 90

Gaussian Cube format, 45, 104, 112
Gaussian window function, 71, 78
Generalized gradient approximation, 101
GGA, *see* Generalized gradient approximation
GTH pseudopotentials, 90, 112

H_2O, 2, 7, 9, 15, 52, 82, 95, 100, 107, 111, 117, 130
Hamiltonian system, 10
Hann window function, 78
Harmonic oscillator, 50, 80, 122, 130
Harmonic potential, 122
HDF5, 43
Hessian matrix, 5, 74
Hilbert curve, 36
Huffmann coding, 40
Hybrid functional, 59
Hydrogen bond, 8, 82, 84, 100
Hyperpolarizability, 67

IL, *see* Ionic liquid
Imidazolium, 9, 45, 46, 82, 107
Implicit solvent, 6
Index frame, 43
Information entropy, 30, 31
Instantaneous normal mode analysis, 74
Interface, 3, 67
Interpolation, 23, 78
  trigonometric, 78
Ionic liquid, 9, 45, 46, 82, 107
Isotropic polarizability, 54, 72

Jacobi diagonalization, 14, 25, 75

KCl, 89
Key frame, 44
Kohn–Sham orbitals, 13
Kronecker delta, 22, 63

LAMMPS, 16, 111
Langevin thermostat, 111
Laplace operator, 57
Laser, 2, 67, 69, 100, 103, 106
Least-squares solution, 33
Lennard-Jones potential, 111
Linear response TDDFT, 69
Liouville theorem, 116
Locality, 36
Long-range Coulomb solver, 111
Lossless compression, 29, 30
LR-TDDFT, *see* Linear response TDDFT
LU decomposition, 58
LU factorization, 58

Magnetic field perturbation theory, 60
Magnetic moment, 57, 60, 61, 63
Mass weighting, 49, 52, 54
Mass-weighted coordinates, 5
Massive thermostat, 112
Matrix form, 33, 58
MC, *see* Monte Carlo
Measure space
  finite, 115
Measure-preserving transformation, 116
Melt, 89
Methane, 87
Methanol, 24, 84
Method B, 18, 19
Methyl oxirane, 86, 91, 93
MFPT, *see* Magnetic field perturbation theory
$MgCl_2$, 89
Microsolvation, 6
Mixture, 13
Modern theory of polarization, 13, 53
Molecular orbitals, 13
MOLOPT basis set, 90, 112
Monolayer, 67
Monte Carlo, 9
Moore–Penrose inverse, 33
Morse potential, 130
Move-to-front transformation, 38
MP2, 59
MTS, *see* Multiple time step approach
Multi-quantum transition, 130
Multiple time step approach, 12
Multipoles
  electric, 28

N-Acetyl-L-cysteine, 95
NALC, *see* *N*-Acetyl-L-cysteine
Near-resonant susceptibility, 67
Newton's equation, 4, 10
Newton–Raphson method, 12
NHC, *see* Nosé–Hoover chain thermostat
o-Nitrophenol, 105
Non-linear regression, 35
Normal mode, 5, 13, 27, 52, 54, 56, 60, 62, 74, 85
Nosé–Hoover barostat, 111
Nosé–Hoover chain thermostat, 11, 90, 111, 112, 128
NpT ensemble, 90, 103, 111
Nuclear velocity perturbation theory, 8, 57, 59, 60
NVE ensemble, 11, 111
NVPT, *see* Nuclear velocity perturbation theory
NVT ensemble, 11, 88, 90, 111, 112

o-Nitrophenol, 105
[OAc], 9, 45, 46, 82, 107
ODE, *see* Ordinary differential equation, 122
Optical activity tensor, 62
Optical field, 68
Optimal code, 40
Orbital transformation, 90, 112
Ordinary differential equation, 10, 12
Orientational averaging, 54
Orthogonal Raman spectrum, 56, 73, 87
Orthonormal transformation, 75
Orthorhombic cell, 21
OT, *see* Orbital transformation
Overlay, 98
Overtones, 7, 52, 130

Packmol, 16, 111

Parallel Raman spectrum, 56, 73, 87
Parallelepiped, 21
Partial differential equation, 58
Path integral molecular dynamics, 8
PBC, *see* Periodic boundary conditions
PBE functional, 90
PBE0 functional, 59
PCM, 6
PDE, *see* Partial differential equation
Periodic boundary conditions, 7, 8, 13–15, 17, 21, 25, 52, 53, 58, 62, 70, 104
Perturbation theory, 13, 15, 62
Phase space, 115
Phonons, 2
PIM, *see* Polarizable ion model
PIMD, *see* Path integral molecular dynamics
Pipek–Mezey localization, 13
Plane wave cutoff, 90, 112
Poincaré recurrence theorem, 116
Polarizability, 13, 15, 27, 54, 62, 63, 67, 68, 88, 96
  dynamic, 69
Polarizable ion model, 90
Polarized band, 87
Polyhedron
  Voronoi, 18
Polynomial, 31
Post-Hartree–Fock method, 59
Power diagram, 19
Power spectrum, 49, 74, 130
PPPM method, 111
Preconditioning, 58, 90, 112
Prefix code, 40
Principal mode analysis, 74
Propagator function, 116
Propylene oxide, 86, 91, 93

Quadrupole moment, 22, 28, 62
Quicksort, 38
Quickstep, 90, 112

Racemate, 3
Radical Voronoi tessellation, 19
Raman scattering, 2, 54, 69
Raman spectrum
  depolarization ratio, 56, 87
  depolarized, 56, 73, 86, 89
  orthogonal, 56, 73, 87
  parallel, 56, 73, 87
  unpolarized, 56, 73, 86, 89
Random access, 44
RATTLE algorithm, 111
Real-time propagation, 104
Real-time TDDFT, 69
Reciprocal lattice vector, 14
Rectangular window function, 77
Recursion, *see* Recursion
Refinement, 23
Regression
  non-linear, 35
Residual, 31
Resonance Raman effect, 2, 69, 102
Resonant susceptibility, 67
Restraint, 97, 98
revPBE functional, 90
Ribonucleic acid, 100
RMSD, *see* Root mean square deviation
RNA, *see* Ribonucleic acid
Root mean square deviation, 76, 109
Rotational strength, 60
RT-TDDFT, *see* Real-time TDDFT
RTP, *see* Real-time propagation
Run-length encoding, 39

Salt, 89
Sankey diagram, 82
SCF, *see* Self-consistent field
Self-consistent field, 15, 112
Shannon’s theorem, 40
Shock waves, 111
σ-Algebra, 115
sinc, *see* Sinus cardinalis
Singular value decomposition, 34
Sinus cardinalis, 78
Solvent effect, 7, 100, 102
Solvent influence, 103
Sparse matrix, 58
Spread functional, 14
Spring constant, 98
Static–harmonic approximation, 4, 52, 54, 60, 62, 69, 105
Step response, 70, 104
Surface selectivity, 3, 67

Surface symmetry, 67
Susceptibility, 67
SVD, *see* Singular value decomposition
Symplectic integrator, 10

Taylor expansion, 4
TCF, *see* Time-correlation function formalism
TDDFT, *see* Time-dependent DFT
Temporal average, 11, 53, 55, 115, 117
Tessellation, 17, 19
Tetrachloromethane, 84
Theorem
  convolution, 120
  cross-correlation, 77, 120
  equipartition, 50
  ergodic, 117
  Poincaré recurrence, 116
  Shannon, 40
  Wiener–Khinchin, 51, 77, 118
Thermostat
  Berendsen, 111
  massive, 88, 112
  Nosé–Hoover chain, 11, 90, 111, 112, 128
Time reversibility, 74, 79
Time-correlation function formalism, 7, 13, 25, 52, 55, 60, 62, 68, 74, 118
Time-dependent DFT, 69
TIP4P–EW, 111
Total charge, 19
Total electron density, 59
Trace, 22, 62
Transformation
  measure-preserving, 116
  unitary, 13
Traversal scheme, 36
TRAVIS, 16, 21, 51, 73, 81, 82, 96, 112
Trigonometric interpolation, 78

Ultraviolet, 100
Unitary transformation, 13
Unpolarized Raman spectrum, 56, 73, 86, 89
Uracil, 100, 103
UV, *see* Ultraviolet

van der Waals radii, 20, 92, 94, 112
VCI, 7
vDOS, *see* Vibrational density of states
Velocity spectrum, 49
Verlet integrator, 10, 80, 121
Verlet propagator, 122
Vertex error, 19
Vibrational density of states, 49
Vibronic coupling, 2, 73, 102, 106
Vibronic theory, 69
Voronoi integration, 22, 53, 61, 63, 68, 70, 89, 104, 112
Voronoi tessellation, 17
VPT2, 7
VSCF, 7

Wannier localization, 13, 24
Wannier spread, 15
Water, 2, 7, 9, 15, 52, 82, 95, 100, 103, 107, 111, 117, 130
Wiener–Khinchin theorem, 51, 77, 118
Window function, 71, 77
  exponential, 78
  Gaussian, 71, 78
  Hann, 78
  rectangular, 77

Yoshida–Suzuki integrator, 12, 128, 129

Zero measure set, 116
Zero padding, 78

Acknowledgments

In first place, I would like to thank my mentor, colleague, and friend Daniel Sebastiani who supported me all the way in doing the research presented in this thesis.

Similarly important, I would like to thank all members of the small group which I have established in Halle *(see Figure A.2)*, namely Christoph, Moritz, Eliane, Christopher, and Thomas *(from left to right)* for the enjoyable time we have spent together.

Special thanks go to Martin Thomas for his contributions to many approaches which have been developed within the scope of this thesis.

I thank Christian Dreßler for the countless fruitful scientific *(and not-so-scientific)* discussions over the last years. I also thank Arne Scherrer and Sascha Jähnigen for the scientific exchange.

I would like to thank Barbara Kirchner, Jörg Kressler, Dariush Hinderberger, Alessandro Triolo, Vyacheslav Bryantsev, Marie-Pierre Gaigeot, and Yunjie Xu for the successfull collaborations.

My thanks go to Jürg Hutter for his and his group's good work on the CP2k code, and for giving me the chance to implement the Voronoi integration and bqb compression there during my 6-month guest visit in 2020 / 2021.

I am happy to thank Martin Thomas, Christian Dreßler, and Victoria Munsky for the proof-reading of this thesis.

I would like to express my gratitude to my family and in particular to Victoria Munsky for the continuous support during writing this thesis.

Finally, I would like to acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG) through the project grants BR 5494/1–1 and BR 5494/1–3.

Figure A.2.: The "AK Brehm" at the STC conference 2019 in Rostock.

Curriculum Vitae of the Author

1985	Born in Leipzig
2001 – 2004	Attended secondary school “Goethegymnasium Ilmenau”
10/2004 – 07/2007	Bachelor studies of chemistry at the University of Leipzig
02/2007 – 04/2007	Bachelor's thesis at the University of Toronto (Canada) Title: *“Neue Synthese von Tetrahydrofuran-Derivaten”*
10/2007 – 09/2009	Master studies of chemistry at the University of Leipzig
05/2009 – 09/2009	Master's thesis at the University of Leipzig Title: *“Entwicklung eines Programmpaketes zur Analyse und Visualisierung von MD-Trajektorien”*
Since 10/2009	Diploma studies of mathematics at the University of Leipzig
10/2009 – 12/2012	PhD student at the University of Leipzig in the group of Prof. Dr. Barbara Kirchner
01/2013 – 04/2014	PhD student at the University of Bonn in the group of Prof. Dr. Barbara Kirchner
04/2014	Obtained PhD *(“Dr. rer. nat.”)* at the University of Leipzig Title: *“Analyzing Trajectories from Molecular Simulation”*
01/2014 – 06/2016	Postdoc at the Helmholtz-Zentrum für Umweltforschung Leipzig (UFZ) in the group of Prof. Dr. Gerrit Schüürmann
Since 06/2016	Junior group leader at the Martin-Luther-Universität Halle–Wittenberg in the group of Prof. Dr. Daniel Sebastiani
11/2020 – 05/2021	Scientific guest visit at the University of Zurich (Switzerland) in the group of Prof. Dr. Jürg Hutter
09/2021	Received the Hellmann award 2021 of the “Arbeitsgemeinschaft Theoretische Chemie” (AGTC)

Statement of Authorship

Eidesstattliche Erklärung

Hiermit erkläre ich, dass die Habilitationsschrift selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht wurden.

Halle, 01.03.2023

Martin Brehm