Universiteit
Leiden
The Netherlands

Nuclear quantum effects in solid water: new insights from computational modeling

Rasti, S.

Citation

Rasti, S. (2022, October 25). *Nuclear quantum effects in solid water: new insights from computational modeling*. Retrieved from https://hdl.handle.net/1887/3484763

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3484763

Note: To cite this publication please use the final published version (if applicable).

CHAPTER 2

Theory and Methods

Atomistic modelling of water in its condensed form can be done at very different levels of theory. Here, those which are relevant to this thesis are briefly described, including some numerical and computational details where appropriate. First, the stage is set in Section 2.1 by the Born-Oppenheimer approximation, which is most fundamental for the description of interactions and dynamics at the atomic scale. After that, Section 2.2 introduces density functional theory (DFT) as one of the most popular and successful first-principles methods for inter- and intramolecular interactions between individual water molecules. Given the importance of van-der-Waals interactions for the latter, this chapter also pays special attention on how these interactions can be accounted for in the context of DFT. Coarse-graining away the electronic structure, force fields are another commonly-used approach for modelling water-water interactions as presented in Section 2.3. Those which are used in this thesis can be grouped into two categories, fixed-charged and polarizable force fields of different complexity. Since the main goal of this thesis is studying properties of crystalline forms of water ice which are heavily affected by the lattice vibrations of the latter, this chapter ends with concise descriptions of phonons in Section 2.4. In particular, their relation to the Helmholtz free energy is described and the so-called quasi-harmonic approximation is presented in order to account for thermal expansion.

2.1 Born-Oppenheimer Approximation

The description of a chemical system from first principles starts from the quantum mechanical many-body problem that is given in its time-independent form by the electron-nuclear Schrödinger equation

$$\hat{H}\Psi = E\Psi \quad . \tag{2.1}$$

E is the ground state (total) energy of the system (at zero temperature), Ψ is the electron-nuclear wave function that describes the spatial coordinates $\{\mathbf{R}_A\}_{A=1}^{M}$ of M nuclei and the spatial and spin coordinates $\{\mathbf{r}_i, \sigma_i\}_{i=1}^{N}$ of N electrons. The Hamilton operator is given by

$$\hat{H} = -\sum_{i=1}^{N}\frac{\nabla_i^2}{2} - \sum_{A=1}^{M}\frac{\nabla_A^2}{2M_A} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} - \sum_{i=1}^{N}\sum_{\mathrm{A}=1}^{M}\frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \tag{2.2}$$

in Hartree atomic units. M_A and Z_A are the mass and charge of a nucleus A, respectively, ∇_A (∇_i) denotes the Laplacian operator with respect to nuclear (electronic) coordinates. The first and third term are the kinetic energy of and Coulomb repulsion between the electrons, and the second and fourth term describe the same for the nuclei. The last term accounts for the Coulomb attraction between electrons and nuclei.

Exact analytical solutions of Eq. (2.2) are possible only for very few chemically relevant systems. Also numerically it is enormously difficult to calculate the full electron-nuclear wave function. Instead, electrons and nuclei are usually decoupled from each other according to the Born-Oppenheimer approximation [1], which is based on the large mass difference between the former: even for the lightest atom M_A is already about 2000 times larger than the electron mass. The electronic part of $\hat{H}$ is given by

$$\hat{H}_{\mathrm{elec}} = -\sum_{i=1}^{N}\frac{\nabla_i^2}{2} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \quad , \tag{2.3}$$

which yields an electronic Schrödinger equation analogous to Eq. (2.2)

$$\hat{H}_{\mathrm{elec}}\Psi_{\mathrm{elec}} = E_{\mathrm{elec}}\Psi_{\mathrm{elec}} \tag{2.4}$$

with energy E_{elec} for the electronic ground state. This equation is easier to solve than Eq. (2.2), but still approximations are needed for chemically interesting systems. One of the most popular approximations is described in Section 2.2.

The Born-Oppenheimer approximation allows to express the total (potential) energy of a chemical system for a set of nuclear positions ($\{\mathbf{R}_A\}_{A=1}^{M}$) as

$$E_{\text{total}} = E_{\text{elec}} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \quad . \tag{2.5}$$

The real-valued function $E_{\text{total}}(\{\mathbf{R}_A\}_{A=1}^{M})$ of $3M$ variables constitutes the system's potential energy surface (PES). Instead of solving Eq. (2.4), a PES can also be constructed (semi-)empirically based on physical-chemical properties of the system and thus without explicitly accounting for the electrons – which is commonly referred to as a force field. There are many such force fields for water, and those which are relevant in the context of this thesis are described in Section 2.3.

2.2 Density Functional Theory

One of the most popular and successful approaches for solving the electronic Schrödinger equation (Eq. (2.4)) is density functional theory. According to the Hohenberg-Kohm theorem it is fundamentally possible to write E_{elec} as a functional of the electron density $n(\mathbf{r})$ [2]. Unfortunately, a universal (Hohenberg-Kohn) density functional is not known. However, Kohn and Sham have suggested to use an auxiliary system of non-interacting electrons as a starting point for to approximate this functional. This system has a well-defined kinetic energy

$$T_{\text{s}}(\{\phi_i\}) = \int d\mathbf{r}\, \tau_{\text{s}}(\mathbf{r}) = -\frac{1}{2} \sum_i \int d\mathbf{r}\, \phi_i^*(\mathbf{r})\, \nabla^2\, \phi_i(\mathbf{r}) \quad , \tag{2.6}$$

where ϕ_i are the single-particle (so-called Kohn-Sham) orbitals and $\tau_{\text{s}}(\mathbf{r})$ is the concomitant kinetic energy density. The Kohn-Sham orbitals yield the electron density [1]

[1] The sums in Eqs. (2.6) and (2.7) extend over an appropriate amount of orbitals depending on the treatment of electron spin.

as

$$n(\mathbf{r}) = \sum_i |\phi_i|^2 \quad . \tag{2.7}$$

T_s constitutes a large contribution to E_elec, and the same holds for the classical electrostatic repulsion

$$E_\mathrm{Hartree}[n(\mathbf{r})] = \int d\mathbf{r} \int d\mathbf{r}' \, \frac{n(\mathbf{r})\, n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.8}$$

and electron-nuclear attraction

$$E_\mathrm{ext}[n(\mathbf{r})] = -\sum_{A=1}^{M} \int d\mathbf{r} \, \frac{n(\mathbf{r})\, Z_A}{|\mathbf{r} - \mathbf{R}_A|} \tag{2.9}$$

energies. This results in the decomposition

$$E_\mathrm{elec}[n] \;=\; T_\mathrm{s}\{\phi_i\} \;+\; E_\mathrm{Hartree}[n(\mathbf{r})] \;+\; E_\mathrm{ext}[n(\mathbf{r})] \;+\; E_\mathrm{xc}[n(\mathbf{r})] \quad , \tag{2.10}$$

which is employed by the vast majority of practical DFT calculations nowadays. The so-called exchange-correlation functional $E_\mathrm{xc}[n(\mathbf{r})]$ is supposed to capture all the quantum-mechanical contributions of the actual system of interacting electrons that are missed by the other terms. Again, the exchange-correlation functional is not known exactly, but since it is a much smaller contribution to E_elec than the Hohenberg-Kohn functional, much simpler approximations can thus yield very accurate results. Minimizing Eq. (2.10) with respect to the electron density yields the Kohn-Sham equations, which are Schrödinger-like equations for the Kohn-Sham orbitals. This equations need to be solved self-consistently due to the relation between the latter and the electron density given by Eq. (2.7).

2.2.1 Exchange-Correlation Functionals

Nowadays there are many different exchange-correlation functionals available. They are usually categorized according to their complexity by the scheme originally suggest by John Perdew [3]. This Jacob's ladder (see Fig. 2.1) is comprised of several rungs starting from the Hartree approximation, i.e., only accounting for exchange ("earth") and ascending to the exact exchange-correlation functional ("heaven"). The simplest and therefore most common approximations for E_xc (first three rungs) are used in

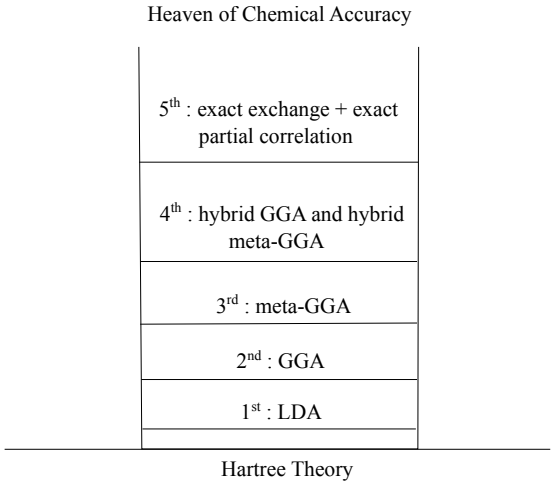

Figure 2.1: Jacob's ladder for categorizing different approximations for E_{xc} according to Perdew *et al.* [3]

the scope of this thesis (see Chapter 3 in particular) and therefore described in the following subsections.

2.2.1.1 Local Density Approximation

In the local density approximation (LDFA) for E_{xc} the electron density at each point ($\mathbf{r}$) is taken to define a homogenous electron gas (HEG), so that the corresponding exchange-correlation functional is obtained as the superposition

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int d\mathbf{r}\, n(\mathbf{r}) \left(\epsilon_{\mathrm{x}}^{\mathrm{HEG}}[n(\mathbf{r})] + \epsilon_{\mathrm{c}}^{\mathrm{HEG}}[n(\mathbf{r})]\right) \quad . \tag{2.11}$$

The exchange energy density of the HEG is known analytically

$$\epsilon_{\mathrm{x}}^{\mathrm{HEG}}[n] = -\frac{3}{4}\sqrt[3]{\frac{3\,n(\mathbf{r})}{\pi}} \quad . \tag{2.12}$$

For the correlation energy density of the HEG $\epsilon_{\mathrm{c}}^{\mathrm{HEG}}[n]$ accurate Quantum Monte Carlo energy are available for a large range of electron densities [4, 5]. Historically, these results have been interpolated differently by, e.g. Vosko-Wilk-Nusair (VWN) [6], Perdew and Zunger (PZ) [7] as well as Perdew and Wang (PW) [8], which has led to numeri-

cally (but not conceptually) different LDAs. The differences are very small so that one usually still refers to "the LDA".

In general, for inhomogeneous systems, the LDA underestimates the correlation and overestimates the exchange energy so that the total balance, $E_{\mathrm{xc}}^{\mathrm{LDA}}$, benefits from error cancellation. Since LDA satisfies several so-called sum rules [9, 10], this error cancellation is not accidental but systematical. Still, the cohesive energies of ice and water clusters are overestimated by more than 50 % [11–13]. In addition, the LDA completely misses long-range van der Waals interactions due to its strictly local dependence on the electron density.

2.2.1.2 Generalized Gradient Approximation

Generalized gradient approximations (GGAs) for the exchange-correlation energy are based on the general form

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n] = \int d\mathbf{r}\, n(\mathbf{r})\, \epsilon_{\mathrm{xc}}^{\mathrm{HEG}}[n(\mathbf{r})]\, F_{\mathrm{xc}}^{\mathrm{GGA}}[n(\mathbf{r}), \nabla n(\mathbf{r})] \quad . \tag{2.13}$$

In the spirit of a Taylor expansion the so-called gradient enhancement factor $F_{\mathrm{xc}}^{\mathrm{GGA}}$ depends not only on the density but also on its gradient at a point $\mathbf{r}$, which is why GGAs are also called semi-local functionals. There is no unique way how to construct $F_{\mathrm{xc}}^{\mathrm{GGA}}$, which has resulted in a variety of different GGA functionals. Some examples are the GGAs developed by Lee, Yang and Parr (LYP) [14], Perdew and Wang (PW91) [15] and Perdew, Burke and Ernzerhof (PBE) [16]. In general, they are more accurate for chemical systems because the density varies strongly when covalent bonds are formed, and they come at essentially the same computational cost as the LDA.

2.2.1.3 Meta-generalized Gradient Approximation

Meta-generalized gradient approximations (MGGAs) for E_{xc} sit on the third rung of Jacob's ladder. This class of exchange-correlation functionals take the second derivative of the electron density ($\nabla^2 n(\mathbf{r})$) and/or the kinetic energy density (see Eq. (2.6)) into account. Previous studies indicate that meta-GGAs such as TPSS [17] and SCAN [18] can improve the accuracy for numerous systems, including (small) water clusters and

ice [19–22] The computational cost for meta-GGAs is slightly higher than for GGAs, which is mainly because it is more difficult to achieve self-consistency when solving the Kohn-Sham equations.

2.2.2 Van der Waals Interactions

Van der Waals (vdW) interactions can appear (with different meaning) in the context of electrostatic, induction, and dispersion interactions. In the following, the focus is on the latter, i.e. the attractive interaction caused by fluctuating dipoles which are constituted by the correlated motion of electrons in individual atoms. This kind of interaction primarily contributes at the long range and therefore none of the exchange correlation functionals described in the previous section can properly account for them. Nowadays there are many different approaches available that can account for van der Waals interactions in DFT calculations [23–26]. Those which have been employed in this thesis are described very briefly in the following subsections.

2.2.2.1 C_6 Correction and Many-Body Dispersion

Dispersion interactions were introduced by London [27] as so-called London forces. Using the second-order quantum mechanical perturbation theory, the concomitant energy can be expanded as

$$E_{\mathrm{vdW}}^{AB} = -\sum_{n\geq 6}^{\infty} \frac{C_n^{AB}}{R_{AB}^n} \quad , \tag{2.14}$$

where C_n^{AB} are dispersion coefficients for the interaction between a pair of atoms A and B at distance R_{AB} from each other. Focusing on the first non-zero term, an accurate determination of the C_6^{AB} coefficient becomes the primary target for all DFT-based vdW methods described in the remainder of this section. It is calculated according to the Casimir-Polder equation [28]

$$C_6^{AB} = \frac{3}{\pi} \int_0^{\infty} d\omega \, \alpha_A(i\omega) \, \alpha_B(i\omega) \quad , \tag{2.15}$$

where $\alpha_{A,B}$ are the dipole polarizabilities of the respective atoms due to an external electric field. Accurate calculations of $\alpha(i\omega)$ are challenging because a large amount of

accurately described electronically excited states are required. Dipole polarizabilities can also be obtained experimentally in the condensed or gas phase from measurements of the dielectric function [29], Rayleigh scattering [30] and molecular beam [31] experiments.

C_6 correction schemes, which are also commonly referred to as DFT+vdW or DFT-D schemes, add an additional term to the total energy of a DFT calculation that is of the form

$$E_{\mathrm{vdW}} = -\sum_{A}\sum_{B>A} f_{\mathrm{damp}}(R_{AB})\,\frac{C_6^{AB}}{R_{AB}^6} \quad . \tag{2.16}$$

Maintaining the short range interaction between two atoms A and B at the DFT level, the damping function $f_{\mathrm{damp}}(R_{AB})$ smoothly switches off the dispersion interactions at small distances. Obviously, Eq. (2.16) can be easily evaluated after the actual DFT calculation and therefore comes at hardly any additional computational cost on top of the former.

In this thesis, the so-called TS correction scheme suggested by Tkatchenko and Scheffler [32] is used. It is based on accurate TD-DFT calculations of C_6 coefficients for the vdW interactions between individual atoms, $C_{6,0}$, by Chu and Dalgarno [33]. To account for changes for atoms in molecules (or in a solid), these coefficients are rescaled to

$$C_{6,\mathrm{eff}}^{AA} = v_{AA}[n]\; C_{6,0}^{AA} \quad . \tag{2.17}$$

Here the scaling factor $v_{AA}[n]$ describes the change of atomic volume based on Hirshfeld partitioning [34] of the electron density n that is obtained from the DFT calculation for the "non-free" atoms A. The corresponding C_6 coefficients for heteronuclear pairs of atoms are obtained from

$$C_{6,\mathrm{eff}}^{AB} = \frac{2\,C_{6,\mathrm{eff}}^{AA}\,C_{6,\mathrm{eff}}^{BB}}{\frac{\alpha_{0,B}}{\alpha_{0,A}}C_{6,\mathrm{eff}}^{AA} + \frac{\alpha_{0,A}}{\alpha_{0,B}}C_{6,\mathrm{eff}}^{BB}} \quad , \tag{2.18}$$

where the (static dipole) polarizabilities of the free atoms α_0 are also taken from the work of Chu and Dalgarno [33]. $f_{\mathrm{damp}}(R_{AB})$ is chosen as a Fermi-type damping function with parameters that are optimized to the S22 database from Jurečka *et al.* [35].

The TS scheme provides a surprisingly accurate description of van der Waals interactions in the DFT context in particular for molecular systems. However, it struggles with accounting for (long-range) screening in the condensed phase, which has been addressed by an extension to the TS scheme Tkatchenko *et al.* [36]. This so-called many-body dispersion scheme (MBD) replaces Eq. (2.16) by a quantum mechanical many-body expression for a system of coupled dipole oscillators with frequency-dependent polarizabilities $\alpha(i\omega)$, which for the "non-free" atoms are again obtained from Hirshfeld partitioning. MBD still comes at moderate computational costs compared to the underlying DFT calculation and therefore can easily be applied to condensed phases of water.

2.2.2.2 van der Waals Density Functionals

Van der Waals density functionals (vdW-DFs)[2] add a long-range contribution typically to semi-local exchange-correlation functionals

$$E_{\mathrm{xc}}[n] = E_{\mathrm{x}}^{\mathrm{GGA}}[n] \; + \; E_{\mathrm{c}}^{0}[n] \; + \; E_{\mathrm{c}}^{\mathrm{lr}}[n] \quad . \tag{2.19}$$

Usually, the exchange energy is taken from a particular GGA [38, 39], and the short-range correlation energy E_{c}^{0} is treated at the LDA level [26]. The simplest expression for the long-range correlation energy is

$$E_{\mathrm{c}}^{\mathrm{lr}} = \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}' \; n(\mathbf{r}) \, \Phi(\mathbf{r}, \mathbf{r}') \, n(\mathbf{r}') \quad . \tag{2.20}$$

$\Phi(\mathbf{r}, \mathbf{r}')$ is the non-local kernel, for which different expressions have been suggested [38, 40, 41] and some of them been applied to water [42] and ice [43] systems.

From a computational point of view, the straightforward numerical implementation of Eq. (2.20) by evaluating the six dimensional integral is computationally expensive. However, many DFT codes employ the technique suggested by Román-Pérez and Soler [44]. This technique significantly reduces the computational effort and thus enables self-consistent calculations which come at almost the same cost as calculations with GGA functionals.

[2] Acknowledging the work of two pioneers in their development, they also referred to as Langreth-Lunqvist functionals. [37–39]

2.2.3 Computational Aspects

In principal, all-electron calculations for condensed phases of water can nowadays be carried out very accurately and computationally conveniently based by using numerical atom-centered basis sets. However, in practice, not all exchange-correlation functionals have been available in the code implementing the latter, which is why it was necessary to resort to plane-waves basis sets for some of the DFT calculations in this thesis.

2.2.3.1 Numerical Atom-Centered Basis Sets

One type of basis sets employed in this thesis are all-electron numerical atom-centered orbitals (NAO) as implemented in the FHI-AIMS code [45], which can be written as

$$\psi_i(\mathbf{r}) \;=\; \psi_i(r,\theta,\varphi) \;=\; \frac{u_i(r)}{r}\, Y_{l_i m_i}(\theta,\varphi) \quad . \tag{2.21}$$

Here $Y_{l(i),m(i)}$ is a spherical harmonic function with (atomic angular momentum indexes) l and m that implicitly depend on the basis function index i. $u_i(r)$ is obtained as numerical solution of Schrödinger-like radial equation

$$\left[-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l_i(l_i+1)}{r^2} + v_i(r) + v_{\mathrm{cut}}(r)\right] u_i(r) = \epsilon_i\, u_i(r) \tag{2.22}$$

for a particular (radial) potential $v_i(r)$ and a reasonably chosen confining potential $v_{\mathrm{cut}}(r)$. $v_i(r)$ determines the shape of the concomitant basis function. To construct the minimal basis set it is taken from a DFT calculation for the free neutral (and non-spin-polarized) atom of a given chemical element using the exchange-correlation functional selected for the target system. This way, core electrons are described very accurately and efficiently because their orbitals hardly change due the formation of chemical bonds. Additional basis functions are added by using $v_i(r)$ from free hydrogen-like atoms or ions, analogous to polarization functions in Gaussian basis sets. In FHI-AIMS these basis functions are ranked and grouped into tiers for each element based on how much they reduce the total energy for dimers at different bond distances covering the binding curve for that element [45]. $v_{\mathrm{cut}}(r)$ is a steeply increasing potential that ensures $u_i(r)$ quickly goes to zero beyond a certain distance, which is usually chosen

to be between 4 to 5 Å for light elements. This allows to minimize the computational effort for numerical operations involving the basis functions (e.g. integrations), ideally resulting in linear scaling of the latter with respect to the amount of atoms. Predefined choices for basis set tiers and parameters determining the numerical accuracy of the basis set representation (including $v_{\mathrm{cut}}(r)$) are available in FHI-AIMS in form of the so-called *light*, *tight* and *very tight* settings [45]. In this thesis, adequate settings have been selected based on thorough convergence tests.

2.2.3.2 Plane-Wave Basis Sets

According to Bloch's theorem electronic states $\phi_{n,\mathbf{k}}(\mathbf{r})$ in a crystalline solid with band index n at point $\mathbf{k}$ in the first Brillouin zone can be described exactly by the Fourier expansion

$$\phi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_{n,\mathbf{G}+\mathbf{k}} \ \exp(i(\mathbf{G}+\mathbf{k})\mathbf{r}) \quad . \tag{2.23}$$

The discrete set of plane wave vectors $\mathbf{G}$ are given by the reciprocal lattice of the crystal. For numerical applications of Eq. (2.23) in a computer code, the plane-wave expansion needs to be cut off to keep the amount of basis set coefficients $C_{n,\mathbf{G}+\mathbf{k}}$ finite. In practice, this is achieved by choosing a cut-off energy E_{cut} such that

$$\frac{1}{2}|\mathbf{G}+\mathbf{k}|^2 \leq E_{\mathrm{cut}} \quad . \tag{2.24}$$

Obviously, the choice of the cut-off energy determines the numerical accuracy of a plane-wave DFT calculation. Core electrons require very large values for E_{cut} because they are strongly localised at a particular nucleus and cause the radial parts of the corresponding valence electrons to oscillate in this core redion in order to maintain orthogonality. Since the computational effort increases significantly with increasing E_{cut}, practical plane-wave calculations freeze the core electrons into the nuclei to avoid the two aforementioned problems. Each chemical element is thus represented by a so-called pseudopotential.

In this thesis, the VASP code is used for all plane-wave DFT calculations, which also provides a set pre-constructed pseudopotentials. The latter and concomitant E_{cut} have

been chosen carefully based on convergence tests and comparison with FHI-aims for each system under investigation.

2.2.3.3 Reciprocal Space Sampling

For a crystalline solid many quantities, like e.g. the electron density, are obtained by integration over the first Brillouin zone Since the electronic wave-functions for different **k**-points are very similar[3], it is possible to approximate the integrals by a summation over a discrete set of **k**-points. Throughout this thesis, uniform **k**-point grids as suggested by Monkhorst and Pack [46] are used, and convergence of the results with respect to the grid density has been carefully checked.

This method generates a uniform **k** points mesh along three lattice vectors in reciprocal space. It is worth noting that an inadequate **k** point sampling can lead to inaccurate results. Nevertheless, setting a denser set of **k** points can solve this problem. Therefore, it is crucial to check the number of **k** points for having accurate and converged results.

2.3 Force Field Methods for Modeling Ice

The potential energy surface that the nuclei experience after the electronic problem has been solved separately according to the Born-Oppenheimer approximation (see Section 2.1) can be describe by a so-called force field model, which is given by a particular (often physically motivated) functional form together with a set of parameters. The construction of such force fields for modelling water in both the liquid and solid phase has a long history. As a result, there are now several tens of such water models available, many of which have been demonstrated to provide an accurate description of many physical properties of water and ice. [49, 57–59].

Figure 2.2 provides an overview of the particular force fields used in this thesis, grouped according to their complexity. The simplest ones, i.e., the SPC/E [60] and TIPnP

[3]See Eq. (2.23), but this also hold for an appropriate expansion of Block states in atom-centered basis sets.

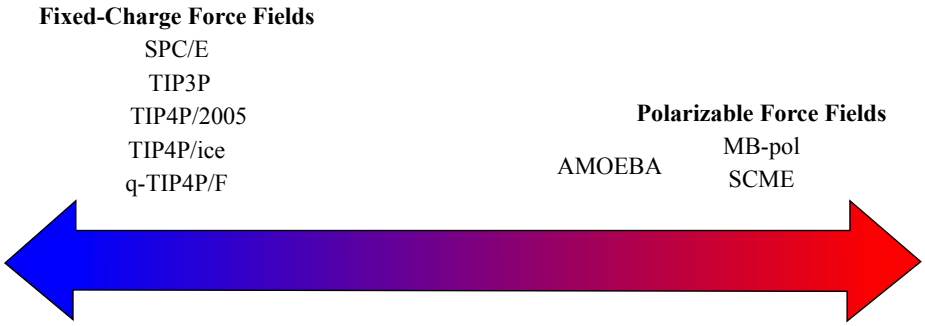

Figure 2.2: Spectrum of the complexity of water force fields. The force fields used in this thesis are SPC/E [47], TIP3P [48], TIP4P/2005 [49], TIP4P/ice [50], and q-TIP4P/F [51], AMOEBA [52, 53], MB-pol [54, 55], and SCME [56].

families of potentials [48], at the left side of this spectrum use fixed point charges for modelling electrostatic and Lennard-Jones potentials (for approximating van der Waals interactions, the latter of which have been discussed earlier in Section 2.2.2 in the context of DFT). The resulting electrostatic and vdW pair potentials are parametrized to reproduce the properties of liquid water or ice. Consequently, the so obtained empirical parameters of these two-body interactions also need to approximately account for any interactions involving three or more water molecules that are not included systematically. The majority of these force fields treat a single water molecule as a rigid monomer – assuming that the energy cost to deform individual water molecules is much larger than intermolecular interactions. SPC/fW [61] and q-TIP4P/F [51] are two examples where flexibility has been added. The explicit expressions and parameters of all the fixed-charged pair potentials used in this thesis are summarized in Section 2.3.1.

Force fields towards the right of Fig. 2.2 all explicitly include (higher) electrostatic multipole moments as well as polarization effects. The latter implicitly accounts for many-body interactions in the electrostatics. Both provide a well-defined classical limit for the interaction between water molecules at large distances, which is relevant

for modelling many delicate properties of water and ice correctly [52, 54, 55, 62, 63]. The common approach for including polarization effects is the Appliquès polarizable point dipole model [64]. This approach is further extended by Thole [65] to address the polarization catastrophe. Some of known Thole-type polarization force fields are TTM3-F [66], TTM4-F [62], AMOEBA [52, 53], SCME [56] and MB-pol [54, 55], the last three of which are being employed in this thesis. Other examples for polarizable force fields are DPP2 [67], CC-pol [68, 69] and WHBB [70]. They have not been considered here because they cannot be used for large systems as they are primarily targeted for water in the gas-phase [71] or because they are rigid water models (DPP2 and CC-pol). In its original formulation, the latter also partially holds for SCME [56], but lifting this limitation completely is part of this thesis (see Chapter 5). The main difference between AMOEBA, SCME and MB-pol is how the quantum mechanical many-body contributions to dispersion and repulsion between water molecules at the short range are being accounted for. As described in Section 2.3.2, AMOEBA is empirically parametrized similar to fixed-charge force fields, but it includes atomic multipole moments (up to quadrupoles) and polarizabilites (up to dipoles). In contrast, both SCME and MB-pol rely on highly accurate *ab initio* quantum chemistry calculations. The main difference is that SCME primarily tries to capture as much as possible by including high-order molecular electrostatic contributions (up to permanent molecular hexadecapoles and induced quadrupoles), which leads to a still fairly small amount of parameters (see Section 2.3.3). On the other hand, MB-pol utilizes a rather complex mathematical construct to fit the contributions to the intermolecular interactions at short distances involving (up to) three water molecules directly to quantum chemical data (see Section 2.3.4).

2.3.1 Fixed-Charge Pair Potentials

The first fixed-charge water model was proposed by Bernal and Fowler [72] in 1933. Nowadays there are now several tens of these models available, among which the extended single charge model (SPC/E [47]) and some of the members of the family of transferable intermolecular potential (TIP) models are employed in this thesis. These

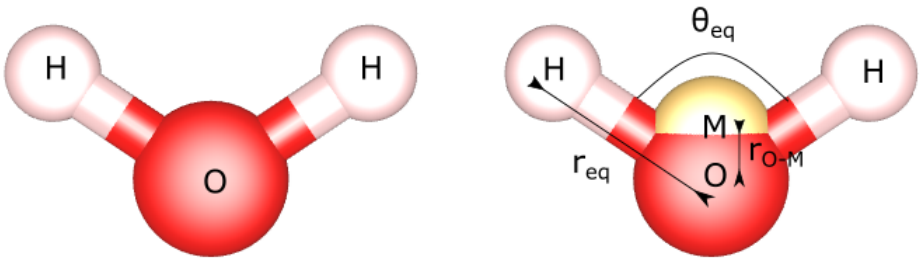

Figure 2.3: Three-site (left) and four-site (right) geometries of fixed-charge water force field. The oxygen atom is shown in red, hydrogen in white, and the virtual site in yellow. r_{eq} represents the equilibrium bond distance between oxygen and hydrogen, θ_{eq} is the equilibrium H−O−H bond angle, and $r_{\mathrm{O\text{-}M}}$ shows the distance between oxygen and virtual site in case of four-site force fields.

models are either based on three (TIP3P [48]) or four (TIP4P/2005 [49]) atomic sites per water molecule as show in Fig. 2.3. They have been modified specifically for ice (TIP4P/ice [50]) or to effectively account for quantum effects while introducing molecular flexibility (q-TIP4P/F [51]).

All of these force fields models describe the total energy based on four energy terms:

$$E_{\mathrm{total}} = \underbrace{E_{\mathrm{elec}} + E_{\mathrm{vdW}}}_{\text{intermolecular}} + \underbrace{E_{\mathrm{bond}} + E_{\mathrm{angle}}}_{\text{intramolecular}} \tag{2.25}$$

Here, the electrostatic (E_{elec}) and van der Waals (E_{vdW}) contributions constitute the intermolecular, and the O−H bond stretching (E_{bond}) and H−O−H angle bending (E_{angle}) constitute the intramolecular contributions.

The electrostatic interactions are described by pairwise Coulomb potentials between the sites i and j that carry the electric charges q_i and q_j:

$$E_{\mathrm{elec}} = \frac{1}{4\pi\epsilon_0} \sum_i \sum_{j>i} \frac{q_i\, q_j}{|\mathbf{r}_i - \mathbf{r}_j|} \quad \text{if} \quad |\mathbf{r}_i - \mathbf{r}_j| < r_{\mathrm{c,elec}} \tag{2.26}$$

ϵ_0 is the electric constant. In the simplest case, to decrease computational cost, only interactions between sites at distances $|\mathbf{r}_i - \mathbf{r}_j|$ smaller than the the cut-off radius

Table 2.1: Electrostatic and Lennard-Jones parameters for SPC/E, TIP3P, TIP4P/2005, TIP4P/ice, and q-TIP4P/F force fields. Maintaining charge neutrality for each water molecule, the oxygen and hydrogen partial charges for the three-site models is $+q$ and $-q/2$, respectively. In four-site models (TIP4P/2005, TIP4P/ice, and q-TIP4P/F), $+q$ is located at the virtual site (see Fig. 2.3). The Lennard-Jones potentials only act between oxygen atoms and thus only $\epsilon_{\mathrm{O-O}}$ and $\sigma_{\mathrm{O-O}}$ in Eq. (2.27) are non-zero. The only exception to that is TIP3P, which is why the corresponding parameters are listed separately in Table 2.2.

	$q(e)$	$\epsilon_{\mathrm{O-O}}(\mathrm{kcal\,mol^{-1}})$	$\sigma_{\mathrm{O-O}}(\text{Å})$
SPC/E	0.8200	0.1553	3.1660
TIP3P	0.8300	[see Table 2.2]	
TIP4P/2005	1.1128	0.1852	3.1589
TIP4P/ice	1.1794	0.2108	3.1668
q-TIP4P/F	1.1128	0.1852	3.1589

Table 2.2: Lennard-Jones parameters for the intermolecular interactions between O−O, O−H and H−H pairs in the TIP3P model.

	O−O	O−H	H−H
ϵ ($\mathrm{kcal\,mol^{-1}}$)	0.1521	0.0836	0.0460
σ (Å)	3.1507	1.7753	0.4000

$r_{\mathrm{c,elec}}$ are taken into account. Contributions from sites beyond the latter can be efficiently accounted for by reciprocal space techniques like for example Ewald summation [73].

E_{vdW} accounts for Pauli repulsion at small and van der Waals attraction at large distances (see Section 2.2.2 for a discussion of the latter in the DFT context). It is given by:

$$E_{\mathrm{vdW}} = \sum_i \sum_{j>i} 4\epsilon_{ij} \left[\left(\frac{\sigma}{\mathbf{r}_i - \mathbf{r}_j} \right)^{12} - \left(\frac{\sigma}{\mathbf{r}_i - \mathbf{r}_j} \right)^{6} \right] \quad \text{if} \quad |\mathbf{r}_i - \mathbf{r}_j| < r_{\mathrm{c,vdW}} \qquad (2.27)$$

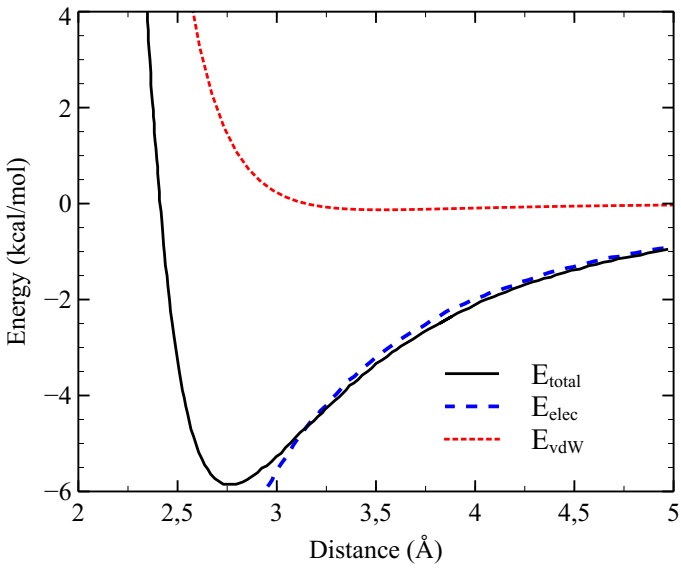

Figure 2.4: SPC/E electrostatic, van der Waals and total energy without intramolecular contribution in kcal mol^{-1} unit when two single water molecules oriented on their hydrogen-bond network.

Here $\sqrt[6]{2}\sigma$ is the equilibrium distance that yields $-\epsilon_{ij}$, i.e. the depth of the potential well. Again, for computational convenience, pairs at distances larger than the cut-off distance $r_{\mathrm{c,vdW}}$ are usually neglected. Unlike for the Coulomb potential, the use of reciprocal space techniques is usually not necessary due the much faster decay of this 6-12 Lennard-Jones potential as a function of distance. The relevance of E_{vdW} for the bonding in different ice phases is analyzed in Chapter 3.

Table 2.1 compiles the parameters related to intermolecular interactions (q, ϵ, and σ) as introduced in Eqs. (2.26) and (2.27) above for each force field. **SPC/E** force field is a result of the re-parametrization of the original SPC [74] potential. Fig. 2.4 shows the SPC/E electrostatic energy (Eq. (2.26)), Lennard-Jones energy (Eq. (2.27)), and total energy without intra-molecular contribution when two single molecule are interacting on their hydrogen-bond side. **TIP3P** is very similar to the SPC/E model but with some minor modifications. One of the main changes is adding additional Lennard-Jones

Table 2.3: Geometry parameters for five fixed-charge water models used in this work. The r_{eq}, θ_{eq}, $r_{\text{O-M}}$ are illustrated in Fig. 2.3.

Water models	r_{eq}(Å)	$r_{\text{O-M}}$(Å)	θ_{eq}(degree)
SPC/E	1.000	n/a	109.5
TIP3P	0.957	n/a	104.5
TIP4P/2005	0.957	0.155	104.5
TIP4P/ice	0.957	0.158	104.5
q − TIP4P/F	0.942	0.155	107.4

parameters for the interactions between hydrogen-hydrogen and oxygen-hydrogen pairs, which are compiled in Table 2.2. In the four-site models **TIP4P/2005**, **TIP4P/ice**, **q-TIP4P/F**, the negative charge of the oxygen atom is located on an additional virtual site on the bisector of the H−O−H angle. Figure 2.3 illustrates this construction, and the model-specific distances from the oxygen atom $r_{\text{O-M}}$ are listed in Table 2.3. This way, the (continuous) electron density as obtained by a quantum mechanical description is mimicked better than with a three-site model. As evidenced by Table 2.1, the parameters related to intermolecular interactions are the same for TIP4P/2005 and q-TIP4P/F. They were obtained in a way to reproduce the phase-diagram of water and ice [49, 51]. TIP4P/ice results from a subsequent reparametrization to better match thermodynamic properties of ice. For example, TIP4P/ice can accurately describe the melting temperature of ice [50].

Intramolecular energy contributions in Eq. (2.25) have not been included in the original **SPC/E, TIP3P, TIP4P/2005, TIP4P/ice** models, i.e. they all describe rigid water molecules. In order to account for intramolecular vibrations, in this thesis I have adapted the common approach to add flexibility by employing simple harmonic potentials. Bond stretching is described by the harmonic potential

$$E_{\mathrm{bond}} = \sum_i K_{\mathrm{r}} \left[(r_{i,1} - r_{\mathrm{eq}})^2 + (r_{i,2} - r_{\mathrm{eq}})^2 \right] \tag{2.28}$$

where $r_{i,1}$ and $r_{i,2}$ are the two O−H bond lengths of a single water molecule and r_{eq} is the model-specific equilibrium bond distance provided in Table 2.3. The same bond

force constant $K_\mathrm{r} = 450.372\,\mathrm{kcal\,mol^{-1}\,\AA^{-2}}$ obtained from the measured Raman spectrum of liquid water [75] is employed for all of the aforementioned models. Equivalently, bending is described by

$$E_\mathrm{angle} = \sum_i K_\theta(\theta_i - \theta_\mathrm{eq})^2 \tag{2.29}$$

where θ_i is the H−O−H angle of each water molecule and θ_eq is the model-specific equilibrium bond angle given in Table 2.3. Again, the same bending force constant $K_\theta = 55\,\mathrm{kcal\,mol^{-1}\,rad^{-2}}$ obtained from experimental data is used for all the models mentioned above.

q-TIP4P/F is a flexible water model "by design" [51]. The bond stretching is accounted for based on a Morse potential

$$E_\mathrm{bond} = \sum_i E_\mathrm{Morse}(r_{i,1}) + E_\mathrm{Morse}(r_{i,2}) \tag{2.30}$$

which is approximated according to

$$E_\mathrm{Morse}(r) \approx D_\mathrm{r}\left[\alpha_r^2(r - r_\mathrm{eq})^2 - \alpha_r^3(r - r_\mathrm{eq})^3 + \frac{7}{12}\alpha_\mathrm{r}^4(r - r_\mathrm{eq})^4\right] \tag{2.31}$$

Thanks to this anharmonic potential, q-TIP4P/F is able to capture the anharmonic and anticorrelation effects between the O−H bond and the hydrogen bonding. [51] The bending contribution to the intramolecular interactions is a harmonic potential like in Eq. (2.29) above. The corresponding parameters, $D_\mathrm{r} = 116.09\,\mathrm{kcal\,mol^{-1}}$ and $\alpha_\mathrm{r} = 2.287\,\mathrm{\AA^{-1}}$ as well as $K_\theta = 43.925\,\mathrm{kcal\,mol^{-1}\,rad^{-2}}$, respectively, are optimized such that path integral molecular dynamics simulations yield best possible agreement with the experimental absorption spectrum, diffusion constant, and vibrational properties of liquid water [51].

In this thesis, all of the fixed-charge pair potentials described in this section are employed via the LAMMPS MD package [76].

2.3.2 AMOEBA

Ren and Ponder [52] have developed the so-called *Atomic Multipole Optimized Energies for Biomolecular Applications* (AMOEBA), which includes polarization effects as well

Table 2.4: Parameters for the electrostatic interactions in the AMOEBA model. For the dipole and quadrupole moments of oxygen atoms, the H−O−H bisector is chosen as z-axis and the x-axis is in the plane defined by the H_2O molecule. For hydrogen, the z-axis is aligned along the O−H bond direction, and the x-axis lies in the molecular pointing towards the other H atom. In both cases, the y-axis is added such that right-handed coordinate systems are completed. Only non-zero components in these reference frames of both multipole moments are listed here.

	O	H
charge	$q_{\mathrm{O}} = -0.426\,16\,e$	$q_{\mathrm{H}} = +0.213\,08\,e$
dipole moment	$\mu_{\mathrm{O},z} = +0.062\,51\,e\,a_0$	$\mu_{\mathrm{H},x} = -0.101\,17\,e\,a_0$
		$\mu_{\mathrm{H},y} = -0.271\,71\,e\,a_0$
quadrupole moment	$\Theta_{\mathrm{O},xx} = +0.175\,76\,e\,{a_0}^2$	$\Theta_{\mathrm{H},xx} = +0.122\,83\,e\,{a_0}^2$
	$\Theta_{\mathrm{O},yy} = -0.231\,60\,e\,{a_0}^2$	$\Theta_{\mathrm{H},yy} = +0.089\,50\,e\,{a_0}^2$
	$\Theta_{\mathrm{O},zz} = +0.055\,84\,e\,{a_0}^2$	$\Theta_{\mathrm{H},zz} = -0.212\,33\,e\,{a_0}^2$
		$\Theta_{\mathrm{H},xy} = -0.069\,89\,e\,{a_0}^2$
polarizability	$\alpha_{\mathrm{O}} = +0.920\,\text{Å}^3$	$\alpha_{\mathrm{H}} = +0.539\,\text{Å}^3$
	$= +6.209 \cdot 4\pi\epsilon_0 a_0^3$	$= +3.637 \cdot 4\pi\epsilon_0 a_0^3$

as intramolecular flexibility. In this thesis, the most recent parameterization for water given in 2014 by Laury *et al.* [77] (AMOEBA14) is used. The total energy expression of AMOEBA consists of the following terms:

$$E_{\mathrm{total}} = \underbrace{E_{\mathrm{elec}}^{\mathrm{perm}} + E_{\mathrm{elec}}^{\mathrm{ind}} + E_{\mathrm{vdW}}}_{\text{intermolecular}} + \underbrace{E_{\mathrm{bond}} + E_{\mathrm{angle}} + E_{\mathrm{UB}}}_{\text{intramolecular}} \tag{2.32}$$

In contrast to the fixed-charge models described in the previous section, the electrostatic contribution takes both permanent ($E_{\mathrm{elec}}^{\mathrm{perm}}$) and induced ($E_{\mathrm{elec}}^{\mathrm{ind}}$) atomic multipole moments into account. The permanent electrostatic moments include monopoles, dipoles, and quadrupoles obtained from high-level electronic structure calculations for the hydrogen and oxygen atomics sites (see Table 2.4) enumerated by i, which are written conveniently in poly-tensorial notation $M_i = [q_i, \mu_{i,x}, \mu_{i,y}, \mu_{i,z}, \Theta_{i,xx}, \Theta_{i,xy}, \ldots, \Theta_{i,zz}]^T$.

This allows to write

$$\begin{aligned} E_{\text{elec}}^{\text{perm}} &= \sum_{\substack{i,j \\ j>i}} q_i \left[T^{ij}, T_x^{ij}, T_y^{ij}, T_z^{ij}, T_{xx}^{ij}, T_{xy}^{ij}, \ldots, T_{zz}^{ij}\right] M_j \\ &+ \sum_{\substack{i,j \\ j>i}} \sum_{\alpha} \mu_{i,\alpha} \left[T_\alpha^{ij}, T_{\alpha x}^{ij}, T_{\alpha y}^{ij}, T_{\alpha z}^{ij}, T_{\alpha xx}^{ij}, T_{\alpha xy}^{ij}, \ldots, T_{\alpha zz}^{ij}\right] M_j \\ &+ \sum_{\substack{i,j \\ j>i}} \sum_{\alpha,\beta} \Theta_{i,\alpha\beta} \left[T_{\alpha\beta}^{ij}, T_{\alpha\beta x}^{ij}, T_{\alpha\beta y}^{ij}, T_{\alpha\beta z}^{ij}, T_{\alpha\beta xx}^{ij}, T_{\alpha\beta xy}^{ij}, \ldots, T_{\alpha\beta zz}^{ij}\right] M_j \ , \end{aligned} \tag{2.33}$$

where the left parts of the "scalar-products" with M_j are the multipole interaction tensor elements in Stone's [78] notation

$$T = \frac{1}{R} \tag{2.34}$$

$$T_\alpha = -\frac{R_\alpha}{R^3} \tag{2.35}$$

$$T_{\alpha\beta} = \frac{1}{R^5}\left[3R_\alpha R_\beta - R^2\delta_{\alpha\beta}\right] \tag{2.36}$$

$$T_{\alpha\beta\gamma} = \frac{1}{R^7}\left[15R_\alpha R_\beta R_\gamma - 3R^2\left(R_\alpha\delta_{\beta\gamma} + R_\beta\delta_{\alpha\gamma} + R_\gamma\delta_{\alpha\beta}\right)\right] \tag{2.37}$$

$$\begin{aligned} T_{\alpha\beta\gamma\delta} = \frac{1}{R^9}\big[&105R_\alpha R_\beta R_\gamma R_\delta \\ &- 15R^2\left(R_\alpha R_\beta\delta_{\gamma\delta} + R_\alpha R_\gamma\delta_{\beta\delta} + R_\alpha R_\delta\delta_{\beta\gamma} + R_\beta R_\gamma\delta_{\alpha\delta} + R_\gamma R_\delta\delta_{\alpha\beta}\right) \\ &+3R^4\left(\delta_{\alpha\beta}\delta_{\gamma\delta} + \delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma}\right)\big] \end{aligned} \tag{2.38}$$

Here R denotes the distances between two sites and the site indicators i and j have been omitted for brevity.

Since AMOEBA includes polarizabilities of hydrogen and oxygen atoms as given in Table 2.4, a smeared charge distribution

$$\rho = \frac{3a}{4\pi}\exp(-au^3) \tag{2.39}$$

with dimensionless width parameter $a = 0.39$ is used in practice, where $u = R_{ij}(\alpha_i\alpha_j)^{-\frac{1}{6}}$. This has originally been suggested by Thole [65, 79] in order to avoid a polarization catastrophy at short-distances. Consequently, the multipole interaction tensor elements $T_{\ldots}$ need to be replaced by their damped counterparts $T_{\ldots}^{\text{D}}$, which include additional weights for the individual terms resulting from Eq. (2.39) but are otherwise of equivalent form [52].

Table 2.5: Parameters for the intermolecular vdW interactions between O−O, O−H and H−H pairs in the AMOEBA model.

	O−O	O−H	H−H
ϵ (kcal mol^{-1})	0.1512	0.0263	0.0105
R^0 (Å)	3.5791	3.2001	2.1176

Induced dipoles are calculated via the product of the polarizabilities and the electric field $\mathbf{F}^i = (F^i_x, F^i_y, F^i_z)$ at atomic site i

$$\begin{aligned}\mu^{\text{ind}}_{i,\alpha} = \alpha_i \, F^i_\alpha = \alpha_i \sum_{\{j\}} [T^{ij}_\alpha, T^{ij}_{\alpha x}, T^{ij}_{\alpha y}, T^{ij}_{\alpha z}, T^{ij}_{\alpha xx}, T^{ij}_{\alpha xy}, \ldots, T^{ij}_{\alpha zz}] \, M_j \\ + \sum_{\{j'\}} [T^{ij'}_{\alpha x}, T^{ij'}_{\alpha y}, T^{ij'}_{\alpha z}] \cdot [\mu^{\text{ind}}_{j',x}, \mu^{\text{ind}}_{j',y}, \mu^{\text{ind}}_{j',z}]^T \quad .\end{aligned} \tag{2.40}$$

$\mathbf{F}^i$ has two contributions. The first one results from the permanent multipole moments of all the water molecules surrounding the molecule containing the atomic site i (summation over $\{j\}$). The second one is from the induced dipoles of all other atomic sites except site i (summation over $\{j'\}$) Equation (2.40) is solved iteratively to yield a self-consistent solution for all the $\mu^{\text{ind}}_{i,\alpha}$ [52]. Substituting $\mu_{i,\alpha} + \mu^{\text{ind}}_{i,\alpha}$ for $\mu_{i,\alpha}$ in Eq. (2.33) yields additional terms which define the induced electrostatic energy contribution $E^{\text{ind}}_{\text{elec}}$.

Also in contrast to the fixed-charge models, the van der Waals intermolecular interactions (E_{vdW}) are based on the buffered 14-7 potential originally suggested by [80]. Dudek and Ponder [81] showed that when combined with quadrupole moments, this potential allows to better reproduce reference energetics obtained from *ab initio* calculations – in particular in the repulsive regime when two interacting sites approach each other. In the AMOEBA14 parametrization the buffered 14-7 potential is given by

$$E^{ij}_{\text{vdw}} = \epsilon_{ij} \left(\frac{1.07}{\rho_{ij} + 0.07} \right)^7 \left(\frac{1.12}{\rho^7_{ij} + 0.12} - 2 \right) \quad , \tag{2.41}$$

where ϵ_{ij} is the minimum energy at distance R^0_{ij} for the interacting sites i and j. $\rho_{ij} = \frac{R_{ij}}{R^0_{ij}}$ describes the rescaled distance between this pair. Like in TIP3P (see Section 2.3.1)

all oxygen and hydrogen atomic sites contribute, and the parameters are tabulated in Table 2.5.

The intramolecular interactions in Eq. (2.32) consist of bonding ($E_{\text{bond}} = \sum_{i,o} E^{i,o}_{\text{bond}}$), H−O−H bending ($E_{\text{angle}} = \sum_i E^i_{\text{angle}}$), and the Urey-Bradley ($E_{\text{UB}} = \sum_i E^i_{\text{UB}}$) contributions.

E_{bond} is based on an anharmonic potential, which includes deviations of the O−H bond length r from its equilibrium value $r_0 = 0.9565\,\text{Å}$ up to fourth order

$$E^{i,o}_{\text{bond}} = K_{\text{r}}(r_{i,o} - r_0)^2 \left[1 - 2.55(r_{i,o} - r_0) - 3.793125(r_{i,o} - r_0)^2\right] \tag{2.42}$$

where $K_{\text{r}} = 556.82\,\text{kcal}\,\text{mol}^{-1}\,\text{Å}^{-2}$ is the stretching force constant i enumerates the water molecules and $o = 1, 2$ the O−H bonds in each molecule.

E_{angle} is described by

$$E^i_{\text{angle}} = K_\theta\,(\theta_i - \theta_0)^2\,\eta(\theta_i, \theta_0) \tag{2.43}$$

where $K_\theta = 48.98\,\text{kcal}\,\text{mol}^{-1}\,\text{rad}^{-2}$ is the bending force constant and

$$\begin{aligned}\eta(\theta, \theta_0) = 1 - 0.014(\theta - \theta_0) \;+\; 5.6 \cdot 10^{-5}(\theta - \theta_0)^2 \\ + 7.0 \cdot 10^{-7}(\theta - \theta_0)^3 \;+\; 2.2 \cdot 10^{-8}(\theta - \theta_0)^4\end{aligned} \tag{2.44}$$

accounts for anharmonicity up to second order deviations from the equilibrium bond angle $\theta_0 = 107.91°$.

E_{UB} is based on the simple harmonic expression

$$E^i_{\text{UB}} = K_l(l_i - l_0)^2 \tag{2.45}$$

where l_0 is the ideal length of virtual bond between the two hydrogens inside a single water molecule and $K_l = -7.6\,\text{kcal}\,\text{mol}^{-1}\,\text{Å}^{-2}$ is the force constant associated with this virtual bond.

The implementation of AMOEBA14 available in the TINKER package [53, 82] is used in this thesis.

Table 2.6: Parameters for the electrostatic interactions in the SCME model. The multipole moments are with respect to the center of mass of a H_2O molecule. The H−O−H bisector is chosen as the negative z-axis and the x-axis is in the plane defined by the H_2O molecule. The y-axis is added such that right-handed coordinate system is completed. Only non-zero components in this reference frames are listed here.

dipole moment ($e\,a_0$)		
$\mu_z = -0.729\,81$		
quadrupole moment ($e\,a_0{}^2$)		
$\Theta_{xx} = +1.955\,32$	$\Theta_{yy} = -1.858\,67$	$\Theta_{zz} = -0.096\,65$
octupole moment ($e\,a_0{}^3$)		
$\Omega_{xxy} = -3.271\,90$	$\Omega_{yyz} = +1.366\,06$	$\Omega_{zzz} = +1.905\,85$
hexadecapole moment ($e\,a_0{}^4$)		
$\Phi_{xxxx} = -0.949\,03$	$\Phi_{xxyy} = -3.384\,90$	$\Phi_{xxzz} = +4.333\,93$
$\Phi_{yyyy} = +4.098\,35$	$\Phi_{yyzz} = -0.713\,45$	$\Phi_{zzzz} = -3.620\,48$

2.3.3 SCME

The *Single Center Multipole Expansion* (SCME) is a water model that in its original formulation by Wikfeldt *et al.* [56] is based on rigid water molecules. Chapter 5 of this thesis describes an extension that adds flexibility and thus intramolecular energy contributions. Unlike AMOEBA, SCME employs permanent molecular multipole moments which are defined with respect to the center of mass of a H_2O molecule up to and including hexadecapoles. Like in AMOEBA, polarizability is accounted for, but also including induced molecular quadrupoles instead of only atomic dipoles. The total energy used by SCME

$$E_{\text{total}} = \underbrace{E_{\text{elec}}^{\text{perm+ind}} + E_{\text{disp}} + E_{\text{rep}}}_{\text{intermolecular}} \tag{2.46}$$

consists of permanent and induced electrostatic ($E_{\text{elec}}^{\text{perm+ind}}$), dispersion ($E_{\text{disp}}$) as well as short-range repulsion (E_{rep}) contributions.

Table 2.7: Same as Table 2.6, but for polarizabilities.

dipole–dipole polarizability ($4\pi\epsilon_0 a_0^3$)		
$\alpha_{xx} = +10.31146$	$\alpha_{yy} = +9.54890$	$\alpha_{zz} = +9.90656$
dipole–quadrupole polarizability ($4\pi\epsilon_0 a_0^4$)		
$A_{x,xz} = -8.42037$	$A_{y,yz} = -1.33400$	$A_{z,xx} = -2.91254$
$A_{z,yy} = +4.72407$	$A_{z,zz} = -1.81153$	
quadrupole–quadrupole polarizability ($4\pi\epsilon_0 a_0^5$)		
$C_{xx,xx} = +12.11907$	$C_{xx,yy} = -6.95326$	$C_{xx,zz} = -5.16582$
$C_{xy,xy} = +7.86225$	$C_{xz,xz} = +11.98862$	$C_{yy,yy} = +11.24741$
$C_{yy,zz} = -4.29415$	$C_{yz,yz} = -6.77226$	$C_{zz,zz} = +9.45997$

The electrostatic energy component is given by

$$E_{\text{elec}}^{\text{perm+ind}} = -\frac{1}{2}\sum_i \left(\mu_\alpha^i \tilde{F}_\alpha^i + \frac{1}{3}\Theta_{\alpha\beta}^i \tilde{F}_{\alpha\beta}^i + \frac{1}{15}\Omega_{\alpha\beta\gamma}^i \tilde{F}_{\alpha\beta\gamma}^i + \frac{1}{105}\Phi_{\alpha\beta\gamma\delta}^i \tilde{F}_{\alpha\beta\gamma\delta}^i \right) , \tag{2.47}$$

where the prefactor of $\frac{1}{2}$ avoids double counting. For the remainder of this section, the Einstein convention is being employed, i.e. doubly occurring Cartesian components (x, y, z), which are denoted by Greek letters, are being summed over. The summation over different water molecules on the other hand is written out explicitly. Dipole, quadrupole, octopole and hexadecapole moments of a particular water molecule i are denoted by μ^i, Θ^i, Ω^i and Φ^i, respectively. The dipoles and quadrupoles are based on measured values [83, 84], whereas the higher multipoles have been obtained from quantum chemical calculations at the MP2 level [85]. They are all defined with respect to the center of mass of a H_2O molecule and compiled in Table 2.6. $\tilde{F}_\alpha^i$ is the α-component of the electric field, caused by all the other molecules except for molecule i and scaled by a switching function f^{sw} that depends on the center-of-mass distance r_{ij}^{com} between molecules i and j:

$$\tilde{F}_\alpha^i = \sum_{\substack{j \\ j\neq i}} f^{\text{sw}}(r_{ij}^{\text{com}})\, F_\alpha^{ij} \tag{2.48}$$

F_α^{ij} is the multipole expansion of the electric field caused by water molecule j at the

center of mass of molecule i

$$\begin{aligned} F_{\alpha}^{ij} = T_{\alpha\beta}^{ij}\left(\mu_{\beta}^{j} + \Delta\mu_{\beta}^{j}\right) - \frac{1}{3}T_{\alpha\beta\gamma}^{ij}\left(\Theta_{\beta\gamma}^{j} + \Delta\Theta_{\beta\gamma}^{j}\right) \\ + \frac{1}{15}T_{\alpha\beta\gamma\delta}^{ij}\,\Omega_{\beta\gamma\delta}^{j} - \frac{1}{105}T_{\alpha\beta\gamma\delta\epsilon}^{ij}\,\Phi_{\beta\gamma\delta\epsilon}^{j} \quad , \end{aligned} \tag{2.49}$$

where again Stone's T notation is used for the multipole interaction tensors as in the previous Section 2.3.2 (see Eqs. (2.34) to (2.38))[4]. The derivatives of this multipole expansion (with respect to Cartesian directions) are denoted by adding additional (Greek) indices, i.e. n additional indices for the n-th derivative:

$$F_{\alpha\ldots\delta}^{ij} = \frac{\partial}{\partial r_{\beta}} \cdots \frac{\partial}{\partial r_{\delta}} F_{\alpha}^{ij} \tag{2.50}$$

These derivatives also define $\tilde{F}_{\alpha\ldots\delta}^{i}$ by substituting F_{α}^{ij} in Eq. (2.48) accordingly. The induced dipole ($\Delta\mu_{\alpha}^{i}$) and quadrupole ($\Delta\Theta_{\alpha\beta}^{i}$) moments of water molecule i are obtained self-consistently from

$$\Delta\mu_{\alpha}^{i} = \alpha_{\alpha\beta}^{i}\,\tilde{F}_{\beta}^{i} \;+\; \frac{1}{3}A_{\alpha,\beta\gamma}^{i}\,\tilde{F}_{\beta\gamma}^{i} \tag{2.51}$$

$$\Delta\Theta_{\alpha\beta}^{i} = A_{\gamma,\alpha\beta}^{i}\,\tilde{F}_{\gamma}^{i} \;+\; C_{\gamma\delta,\alpha\beta}^{i}\,\tilde{F}_{\gamma\delta}^{i} \quad . \tag{2.52}$$

Here $\alpha_{\alpha\beta}^{i}$, $A_{\alpha,\beta\gamma}^{i}$ and $C_{\alpha\beta,\gamma\delta}^{i}$ are the corresponding dipole-dipole, dipole-quadrupole and quadrupole-quadrupole polarizabilities, respectively. Again, the values compiled in Table 2.7 are based on experiments ($\alpha_{\alpha\beta}$) and quantum chemical calculations ($A_{\alpha,\beta\gamma}$ and $C_{\alpha\beta,\gamma\delta}$) [56]. Equations (2.51) and (2.52) together with Eqs. (2.48) and (2.49) (and their multi-subscripts counterparts) are solved iteratively until all components of the electric field and its gradients as well as the induced multipole moments are converged up to certain threshold.[5] f^{sw} smoothly switches off any electrostatic interactions at short ($r_{ij}^{\mathrm{com}} < 5\,\text{Å}$) and long ($r_{ij}^{\mathrm{com}} > 11\,\text{Å}$) distances.[6] The overlap of electron densities associated with different water molecules at short distances leads to strong repulsion due to quantum mechanical effects (Pauli repulsion), which is captured by the E_{rep} term in Eq. (2.46) (*vide infra*). Like for the simplest water force fields described in Section 2.3.1, the long-range cut-off included in f^{sw} is an approximation that makes

[4] $T_{\alpha\beta\gamma\delta\epsilon} = \nabla_{\epsilon} T_{\alpha\beta\gamma\delta}$

[5] 10^{-7} a.u. in the original work of Wikfeldt *et al.* [56].

[6] See Wikfeldt *et al.* [56] for details on the definition of f^{sw}.

Table 2.8: Parameters for the dispersion and repulsion interactions in the SCME model. All values are giving in atomic units matching to Eqs. (2.53) to (2.57)

damping		
$\tau_\mathrm{d} = 2.32837906$		
dispersion		
$C_6 = 46.44309964$	$C_8 = 1141.70326668$	$C_{10} = 33441.11892923$
repulsion		
$A = 1857.45898793$	$C = 1.68708507 \cdot 10^{6}$	
$b = 1.44350000$	$c = 1.83402715$	$d = 0.35278471$
$a_0 = +1.02508535 \cdot 10^{-1}$	$a_1 = -1.72461186 \cdot 10^{-4}$	$a_2 = +1.02195556 \cdot 10^{-7}$
$a_3 = -2.60877107 \cdot 10^{-11}$	$a_4 = +3.06054306 \cdot 10^{-15}$	$a_5 = -1.32901339 \cdot 10^{-19}$

the evaluation of the SCME computationally convenient, in particular when dealing with infinitely extended (bulk) ice structures.

The dispersion and repulsion energies are obtained by only considering the oxygen-oxygen distances r_{ij}^{OO} between different water molecules. The dispersion energy is given by

$$E_\mathrm{disp} = -\sum_{\substack{i,j \\ j>i}} \left[f_6^{\mathrm{TT}}(r_{ij}^{\mathrm{OO}}) \frac{C_6}{(r_{ij}^{\mathrm{OO}})^6} + f_8^{\mathrm{TT}}(r_{ij}^{\mathrm{OO}}) \frac{C_8}{(r_{ij}^{\mathrm{OO}})^8} + f_{10}^{\mathrm{TT}}(r_{ij}^{\mathrm{OO}}) \frac{C_{10}}{(r_{ij}^{\mathrm{OO}})^{10}} \right] \quad , \tag{2.53}$$

with coefficients C_n ($n \in \{6, 8, 10\}$) compiled in Table 2.8 and the Tang-Toennies [86] damping function

$$f_{2n}^{\mathrm{TT}}(r_{ij}^{\mathrm{OO}}) = 1 \; - \; \exp(-\tau_\mathrm{d}\, r_{ij}^{\mathrm{OO}}) \sum_{k=0}^{2n} \frac{(\tau_\mathrm{d}\, r_{ij}^{\mathrm{OO}})^k}{k!} \quad , \tag{2.54}$$

The damping factor τ_d is physically motivated by the inverse decay length of the charge density in a water monomer (see Table 2.8).

The original formulation of SCME [56] employs a modified Born-Mayer potential for

the repulsion energy

$$E_{\text{rep}} = A \sum_{\substack{i,j \\ j>i}} [1 + B(\rho_j) + B(\rho_i)] \ (r_{ij}^{\text{OO}})^{-b} \ \exp(-c \cdot r_{ij}^{\text{OO}}) \quad , \tag{2.55}$$

where

$$B(\rho_i) = \begin{cases} 0 & \text{if } \rho_i \leq 1600 \\ \sum_{n=0}^{5} a_n \rho_i^n & \text{if } 1600 < \rho_i < 8000 \\ 0.0875 & \text{if } 8000 \leq \rho_i \quad . \end{cases} \tag{2.56}$$

ρ_i is the superposition of the electron densities of all water molecules surrounding a given molecule i

$$\rho_i = \sum_{\substack{j \\ j \neq i}} \exp(-d \cdot r_{ij}^{\text{OO}}) \frac{C}{(r_{ij}^{\text{OO}})^3} \tag{2.57}$$

The eleven parameters A, C, a_n $(n = 0, 1, \ldots, 5)$, b, c and d compiled in Table 2.8 have been obtained by Wikfeldt *et al.* [56] via fitting to quantum chemical calculations for small water clusters (consisting of up to six H_2O molecules, at the MP2 level). Finally, in the extension of SCME for flexible water molecules that is presented in Chapter 5 of this thesis, the Born-Mayer term needed to be substituted by a different expression.

2.3.4 MB-pol

The MB-pol potential has been developed by Paesani and coworkers [54, 55, 87]. Compared to the AMOEBA and SCME models described in the two previous subsections, MB-pol relies on a much larger number of parameters, which are consequently obtained by fitting to accurate quantum chemical calculations at the CCSD(T) level [88, 89]. Therefore, in the interest of conciseness and unlike in the previous sections, not all analytical expressions are provided here. Instead, only the essential concepts are explained and compared to AMOEBA and SCME wherever relevant.

Like all water force fields described in the preceding sections, MB-pol is based on the notion that interactions between water molecules can be separated into different contributions of varying relevance at different separation distances [78]. The electron

densities of individual molecules hardly overlap in the long-range part, which is therefore dominated by dispersion and electrostatic interactions. MB-pol (only) considers pair-wise contributions to the dispersion interactions – analogous to the fixed-charge force fields, AMOEBA and SCME. For the electrostatic interactions, MB-pol employs a Thole-type model (TTM) [65], which is based on (smeared) atomic charges and includes polarization and thus many-body effects at the long range similar to AMOEBA but rather different from SCME. Building on earlier work resulting in the so-called HBB2-pol force field [88, 89], MB-pol explicitly accounts for quantum mechanical many-body effects involving up to three water molecules at the short range. This is a unique feature compared to all other force fields considered so far. Going beyond three-body terms has not been found to be necessary for achieving chemical accuracy [90–93]. Since no "simple" analytical form is known for the concomitant short-range two- and three-body potentials, these are heavily parametrised based on a large dataset generated by the aforementioned quantum chemical calculations.

Consequently, the total energy of the MB-pol model

$$E_{\text{total}} = \underbrace{E_{\text{elec}} + E^{\text{2B}}_{\text{short}} + E^{\text{3B}}_{\text{short}}}_{\text{intermolecular + intramolecular}} + \underbrace{E_{\text{disp}}}_{\text{intermolecular}} + \underbrace{E^{\text{1B}}}_{\text{intramolecular}} \tag{2.58}$$

is decomposed into permanent and induced electrostatic (E_{elec}), dispersion (E_{disp}), short-range two-body ($E^{\text{2B}}_{\text{short}}$) and three-body ($E^{\text{3B}}_{\text{short}}$) intermolecular contributions.

The intramolecular contribution that accounts for the flexibility of a single water molecule (one-body term E^{1B}) is taken "as is" from Partridge and Schwenke [94].

The electrostatic energy in MB-pol is based on the so-called TTM4-F model as originally constructed by Burnham *et al.* [62], Burnham *et al.* [95], and Burnham and Xantheas [96]. This bares similarities to the AMOEBA model presented in 2.3.2, i.e., it accounts for permanent and induced electrostatic intermolecular interactions

$$E_{\text{elec}} = E_{\text{TTM4-F}} = E^{\text{perm}}_{\text{TTM4-F}} + E^{\text{ind}}_{\text{TTM4-F}} \quad , \tag{2.59}$$

with smeared charges and induced dipoles located on the H and shifted slightly away from O *atomic sites* of a H_2O molecule. Only charge-charge, charge-dipole and dipole-

dipole interactions (the latter also intramolecular) are included in Eq. (2.59), i.e., unlike in SCME (higher order) molecular multipole moments are not considered explicitly. The dependence of the atomic charges on the geometry of a single water molecule is obtained from the dipole moment surface (DMS) by Partridge and Schwenke [94]. Babin *et al.* [54] use the same values for the polarizabilities and charge (Thole) smearing parameters as given by Burnham *et al.* [62], except for a reduced dipole-dipole interaction related smearing width.[7]

The expression for the dispersion energy

$$E_{\mathrm{disp}} = \sum_{a>b} V_{6,8}(\mathbf{x}_a, \mathbf{x}_b) = -\sum_{a>b} \sum_{\substack{i\in a \\ j\in b}} \left[f_6^{\mathrm{TT}}(\delta_6^{ij}\, r_{ij})\, \frac{C_6^{ij}}{r_{ij}^6} + f_8^{\mathrm{TT}}(\delta_8^{ij}\, r_{ij})\, \frac{C_8^{ij}}{r_{ij}^8} \right] \tag{2.60}$$

is similar to SCME. $\mathbf{x}_a$ and $\mathbf{x}_b$ denote all the atomic positions belonging to water molecules a and b. $V_{6,8}$ obviously contributes to two-body interactions at the long range as it includes summation over all atomic sites i and j in two different H_2O molecules a and b, respectively, with concomitant distances r_{ij} only considered up to r_{ij}^{-8}. f_6^{TT} is the Tang-Toennies damping function already given by Eq. (2.54). The parameters $\delta_{6,8}^{\mathrm{HH}}$, $\delta_{6,8}^{\mathrm{OH}}$, $\delta_{6,8}^{\mathrm{OO}}$ and $C_{6,8}^{\mathrm{HH}}$, $C_{6,8}^{\mathrm{OH}}$, $C_{6,8}^{\mathrm{OO}}$ for Eq. (2.60) are obtained from fitting to the results from the aforementioned quantum chemical calculations.

The short-range two-body energy contribution

$$E_{\mathrm{short}}^{\mathrm{2B}} = \sum_{a>b} V_{\mathrm{short}}^{\mathrm{2B}}(\mathbf{x}_a, \mathbf{x}_b) \tag{2.61}$$

is based on the short-range two-body potential

$$V_{\mathrm{short}}^{\mathrm{2B}}(\mathbf{x}_a, \mathbf{x}_b) = s^{\mathrm{2B}}(r_{\mathrm{O}_a\mathrm{O}_b})\, V_{\mathrm{poly}}^{\mathrm{2B}}(\xi_1^{\mathrm{2B}}, \ldots, \xi_{31}^{\mathrm{2B}}) \quad , \tag{2.62}$$

which involves all the atomic positions $\mathbf{x}_a$ and $\mathbf{x}_b$ of two different water molecules plus two oxygen lone-pair sites each. These two additional sites L_1 and L_2 are schematically depicted for one water molecule in Fig. 2.5. Babin *et al.* [54] have represented $V_{\mathrm{poly}}^{\mathrm{2B}}(\xi_1^{\mathrm{2B}}, \ldots, \xi_{31}^{\mathrm{2B}})$ by permutationally invariant polynomials up to fourth degree in the variables ξ_i^{2B}, which are based on the six intramolecular distances ($\xi_i^{\mathrm{2B}} = \exp(-k_i^{\mathrm{2B}} d_i)$,

[7] It has been set identical to its intermolecular counterpart.

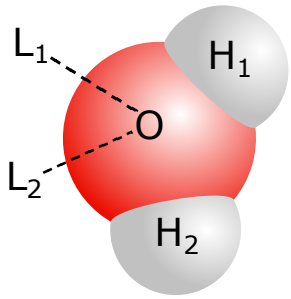

Figure 2.5: Oxygen lone-pair sites L_1 and L_2 used in the short-range two-body potential V^{2B}_{short} (see Eq. (2.62)).

$i = 1, \ldots, 6$), nine intermolecular Coulomb-like terms ($\xi_i^{2B} = \exp(-k_i^{2B} d_i)/d_i$, $i = 7, \ldots, 15$) and the sixteen intermolecular distances involving the aforementioned oxygen lone-pair sites ($\xi_i^{2B} = \exp(-k_i^{2B} d_i)$, $i = 16, \ldots, 31$). Here d_i denote actual distances between sites and k_i^{2B} ($i = 1, \ldots, 31$) are fit parameters, which are obtained from the aforementioned quantum chemical calculations after (long-range) two-body contributions resulting from E_{elec} and E_{disp} have been subtracted. s^{2B} smoothly switches on V^{2B}_{short} in a range $5.5\,\text{Å} < r_{O_a O_b} < 7.5\,\text{Å}$ based on the distance $r_{O_a O_b}$ between the two oxygen atoms of the molecules a and b.

The short-range three-body energy contribution

$$E^{short}_{3B} = \sum_{a>b>c} V^{3B}_{short}(\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c) \tag{2.63}$$

is based on the short-range three-body potential

$$\begin{aligned} V^{3B}_{short}(\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c) = \left[s^{3B}(t_{ab}) s^{3B}(t_{ac}) + s^{3B}(t_{ab}) s^{3B}(t_{bc}) + s^{3B}(t_{ac}) s^{3B}(t_{bc}) \right] \\ V^{3B}_{poly}(\xi^{3B}_1, \ldots, \xi^{3B}_{36}) \quad , \end{aligned} \tag{2.64}$$

involving the atomic positions of three different water molecules ($\mathbf{x}_a$,$\mathbf{x}_b$,$\mathbf{x}_c$). Babin *et al.* [55] have constructed V^{3B}_{poly} in similar to V^{2B}_{poly} by using permutationally invariant polynomials up to fourth degree in the variables ξ_i^{3B}. All monomials contributing to V^{3B}_{poly} are based on the 9 intramolecular distances ($\xi_i^{3B} = \exp(-k_i^{3B} d_i)$, $i = 1, \ldots, 9$

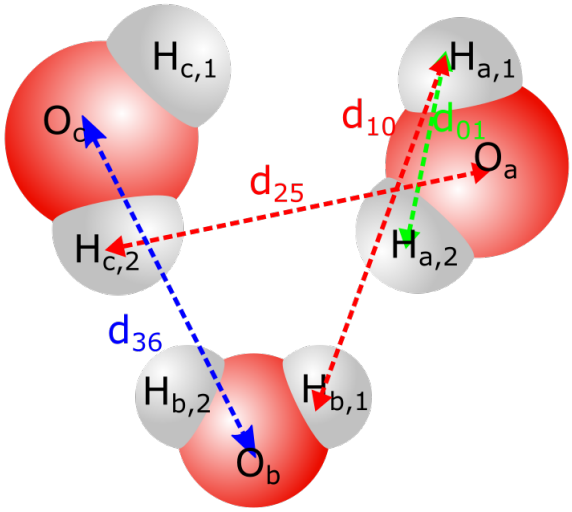

Figure 2.6: Schematic representation of four particular distances d_{10}, d_{25} (both red), d_{36} (blue), and d_1 (green) between atoms of the three depicted H_2O molecules a, b and c. The corresponding variables $\xi_i^{3B} = \exp(-k_i^{3B} d_i)$, $i \in \{1, 10, 25, 36\}$ yield monomials that contribute to V_{poly}^{3B} (see Eq. (2.64)), i.e., $V_{\text{poly}}^{3B} = \ldots + \xi_{10}\xi_{25} + \xi_{10}\xi_{25}\xi_{36} + \xi_{10}\xi_{25}\xi_{36}\xi_{01} + \ldots$

– 3 pairs each per H_2O) and the 27 intermolecular distances ($\xi_i^{3B} = \exp(-k_i^{3B} d_i)$, $i = 10, \ldots, 36$ – 3 pairs of molecules with 9 distance pairs each). This is further illustrated in Fig. 2.6. k_i^{3B} ($i = 1, \ldots, 36$) are fit parameters, which are obtained from the aforementioned quantum chemical calculations after (long-range) three-body contributions resulting from $E^{\text{ind}}_{\text{TTM4-F}}$ (see Eq. (2.59)) has been subtracted. $s^{3B}(t_{ab}) = s^{3B}(t_{ab}(r_{O_a O_b}))$ smoothly deactivates V^{3B}_{short} for oxygen-oxygen distances $r_{O_a O_b} > 4.5\,\text{Å}$ between the two H_2O molecules a and b.

Throughout this thesis, the original parameterisation and implementation of MB-pol developed by the Paesani group [54, 55, 87] is used.

2.4 Phonons

This section first introduces the harmonic approximation for the total energy of a crystalline solid. Solving the lattice dynamical equations then yields uncoupled vibrational modes (phonons) characterised by frequencies and displacement eigenvectors as a function of phonon wave vector. The (harmonic) phonon system is then conveniently characterised by its density of states and subject to statistical treatment in the canonical ensemble to describe non-zero temperatures. Afterwards, the quasi-harmonic approximation is introduced in order to account for thermal expansion. The section concludes with computational aspects focussing on the finite-displacement technique that is used throughout in this thesis.

2.4.1 Harmonic Approximation

Not considering any defects, the total internal energy E of a perfect crystal with equilibrium lattice vectors $\mathbf{R}^{\text{eq}} = (\mathbf{a}^{\text{eq}}, \mathbf{b}^{\text{eq}}, \mathbf{c}^{\text{eq}})$[8] can be expanded around the equilibrium positions of the atoms. Terminating this expansion after the second order is called the

[8] i.e., those which yield the lowest total energy of the crystal for atoms in their equilibrium positions

harmonic approximation

$$\begin{aligned} E^{\mathrm{harm}}(\mathbf{u};\mathbf{R}^{\mathrm{eq}}) &= E_0(\mathbf{R}^{\mathrm{eq}}) + E_1(\mathbf{u};\mathbf{R}^{\mathrm{eq}}) + E_2(\mathbf{u};\mathbf{R}^{\mathrm{eq}}) \\ &= E(\mathbf{u}=\mathbf{0};\mathbf{R}^{\mathrm{eq}}) + \sum_{n,i,\alpha} \underbrace{\left.\frac{\partial E}{\partial u_{n,i,\alpha}}\right|_{\mathbf{u}=\mathbf{0}}}_{=-F_{n,i,\alpha}(\mathbf{u}=\mathbf{0};\mathbf{R}^{\mathrm{eq}})=0} u_{n,i,\alpha} \\ &\quad + \frac{1}{2}\sum_{n,i,\alpha}\sum_{n',i',\alpha'} \left.\frac{\partial^2 E}{\partial u_{n,i,\alpha}\partial u_{n',i',\alpha'}}\right|_{\mathbf{u}=\mathbf{0}} u_{n,i,\alpha}\, u_{n',i',\alpha'} \\ &= E^{\mathrm{eq}} + \frac{1}{2}\sum_{n,i,\alpha}\sum_{n',i',\alpha'} \Phi_{n,i,\alpha}^{n',i',\alpha'}(\mathbf{u}=\mathbf{0};\mathbf{R}^{\mathrm{eq}})\, u_{n,i,\alpha}\, u_{n',i',\alpha'} \quad . \end{aligned} \tag{2.65}$$

Here the summations extend over the infinitely many unit cells of the crystal (n'), the N_{cell} atoms within a unit cell ($i' = 1, \ldots, N_{\mathrm{cell}}$) and their Cartesian coordinates ($\alpha' = 1, 2, 3$). $u_{n,i,\alpha}$ denotes the displacement along Cartesian direction α of atom i in a replica n. When no atoms are displaced ($\mathbf{u} = \mathbf{0}$), only the zeroth order term E_0 contributes and thus defines the (constant) equilibrium energy $E^{\mathrm{eq}} = E(\mathbf{u} = \mathbf{0}; \mathbf{R}^{\mathrm{eq}})$ of the crystal. In the following, E^{eq} is chosen as energy zero such that it does not need to be explicitly denoted. The first order term E_1 vanishes because all force components $F_{n,i,\alpha}(\mathbf{u} = \mathbf{0}; \mathbf{R}^{\mathrm{eq}})$ acting on atom i in a cell n vanish by definition when the atoms are sitting at their equilibrium positions. This only leaves the second order (or so-called harmonic) term E_2, which is fully defined by the force constants $\Phi_{n,i,\alpha}^{n',i',\alpha'}(\mathbf{u} = \mathbf{0}; \mathbf{R}^{\mathrm{eq}})$ of the material. The equilibrium lattice vectors are explicitly denoted in Eq. (2.65) in preparation of the so-called quasi-harmonic approximation, which is introduced below in Section 2.4.4.

2.4.2 Lattice Dynamics

According to Eq. (2.65), the forces in the harmonic approximation are given by

$$F_{n,i,\alpha}^{\mathrm{harm}} = -\frac{\partial E^{\mathrm{harm}}}{\partial u_{n,i,\alpha}} = -\sum_{n',i',\alpha'} \Phi_{n,i,\alpha}^{n',i',\alpha'} u_{n',i',\alpha'} \quad , \tag{2.66}$$

resulting in the following Newtonian equations of motion

$$-\sum_{n',i',\alpha'} \Phi_{n,i,\alpha}^{n',i',\alpha'} u_{n',i',\alpha'} = M_i \frac{\partial^2 u_{n,i,\alpha}}{\partial t^2} \tag{2.67}$$

for each atom with mass M_i in the crystal. In principal, the force constant matrix $\Phi^{n',i',\alpha'}_{n,i,\alpha}$ couples the motion of all atoms in the crystal, with the strength of the couplings typically decreasing with increasing distance from a reference atom. Equation (2.67) can be solved by the following plane-wave-like ansatz for the displacement patterns

$$u_{n,i,\alpha}(\mathbf{q},t) = \frac{A}{\sqrt{M_i}}\,\tilde{u}_{i,\alpha}(\mathbf{q})\,\exp(i(\mathbf{q}\cdot\mathbf{R}_n - \omega(\mathbf{q})t)) \quad , \tag{2.68}$$

where the linear combination of lattice vectors $\mathbf{R}_n = n_a(n)\mathbf{a}^{\mathrm{eq}} + n_b(n)\mathbf{b}^{\mathrm{eq}} + n_c(n)\mathbf{c}^{\mathrm{eq}}$ describes the offset of replica cell n from the (reference) unit cell. The wave vector $\mathbf{q}$ characterizes the periodicity of such a displacement pattern with (angular) frequency $\omega(\mathbf{q})$. All possible displacement patterns can be described by wave vectors that are contained in the (first) Brillouin zone of the crystal (Wigner-Seitz cell of the reciprocal lattice). The unit vector $\tilde{u}_{i,\alpha}(\mathbf{q})$ describes the displacements of each atom in a unit cell and A is the amplitude of the resulting displacement pattern. Substitution of this ansatz into Eq. (2.67)

$$\omega^2\tilde{u}_{i,\alpha}(\mathbf{q}) = \sum_{i',\alpha'}\sum_{n}\frac{1}{\sqrt{M_i M_{i'}}}\Phi^{n',i',\alpha'}_{n,i,\alpha}\,\tilde{u}_{i,\alpha}(\mathbf{q})\,\exp(i\mathbf{q}\cdot\mathbf{R}_n) \tag{2.69}$$

then allows to determine the frequencies $\omega(\mathbf{q})$ and directions of atomic displacements $\tilde{u}_{i,\alpha}(\mathbf{q})$ that are described by Eq. (2.68). Introducing the so-called dynamical matrix

$$D^{i',\alpha'}_{i,\alpha}(\mathbf{q}) = \sum_{n}\frac{1}{\sqrt{M_i M_{i'}}}\Phi^{n',i',\alpha'}_{n,i,\alpha}\exp(i\mathbf{q}\cdot\mathbf{R}_n) \quad , \tag{2.70}$$

Eq. (2.69) constitutes an eigenvalue problem for this $3j \times 3j$ dimensional matrix

$$\omega^2\tilde{u}_{i,\alpha}(\mathbf{q}) = \sum_{i',\alpha'} D^{i',\alpha'}_{i,\alpha}(\mathbf{q})\,\tilde{u}_{i,\alpha}(\mathbf{q}) \quad . \tag{2.71}$$

Since $D(\mathbf{q})$ is Hermitian it can be diagonalized and the $3N_{\mathrm{cell}}$ different eigenvectors and concomitant eigenvalues provide solutions for the displacements $\nu_b(\mathbf{q})$ and their concomitant frequencies $\omega_b(\mathbf{q})$, respectively, where $b = 1,\ldots,3N_{\mathrm{cell}}$. Each solution it describes is a particular phonon state. At each $\mathbf{q}$-point these states are enumerated by the band index b, and the total amount of bands is determined by the amount of atoms in the primitive unit cell. In practice, solutions for a discrete set of $\mathbf{q}$-points sampling the Brillouin zone are being calculated.

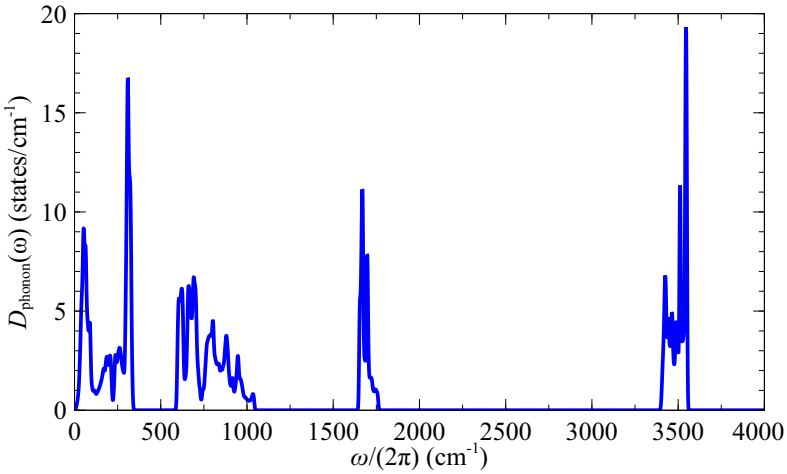

Figure 2.7: Phonon DOS of ice-Ih calculated with the q-TIP4P/F model. Five different sets of phonon modes are clearly separated by their respective frequency ranges, namely hydrogen-bond bending (0 to 150 cm^{-1}), hydrogen-bond stretching (150 to 500 cm^{-1}), librational (500 to 1500 cm^{-1}), H−O−H bending (1500 to 2000 cm^{-1}), and O−H stretching (2000 to 4000 cm^{-1}).

It is convenient to characterize the phonon spectrum by the amount of states per frequency and unit cell, which is expressed by the phonon density of states (DOS)

$$D_{\mathrm{phonon}}(\omega) = \frac{V}{(2\pi)^3} \sum_b \int_{\mathrm{BZ}} d\mathbf{q}\, \delta(\omega - \omega_b(\mathbf{q})) \quad . \tag{2.72}$$

V is the volume of the unit cell and the integration extends over the (first) Brillouin zone (BZ). Figure 2.7 shows D_{phonon} for the most common crystalline ice phase (ice Ih) as calculated with a particular fixed-charged force field (see Section 2.3.1).

2.4.3 Vibrational Free Energy

The canonical partition function of the phonon system Z_{phonon} at temperature T is obtained by summing over all phonon states, which are enumerated by their phonon wavevectors ($\mathbf{q}$ in the first Brillouin zone), band indexes (b) and the amount of vibrational quanta (m) that are "stored" in each state. According to the harmonic ap-

proximation there is no coupling between the different phonon modes. Consequently, the vibrational partition function factorizes into the contribution of individual phonon (harmonic) oscillators

$$Z_{\text{phonon}}(T) = \prod_{\mathbf{q},b} \sum_{m=0}^{\infty} \exp(-\frac{\hbar\omega_b(\mathbf{q})(m+\frac{1}{2})}{k_{\text{B}}T}) = \prod_{\mathbf{q},b} \frac{\exp(-\frac{\hbar\omega_b(\mathbf{q})}{2k_{\text{B}}T})}{1-\exp(-\frac{\hbar\omega_b(\mathbf{q})}{k_{\text{B}}T})} \quad , \tag{2.73}$$

where k_{B} is the Boltzmann constant. This yields the phonon contribution to the Helmholtz free energy

$$\begin{aligned} F_{\text{phonon}}(T) &= -k_{\text{B}}T \ln(Z_{\text{phonon}}) \\ &= \underbrace{\frac{1}{2}\sum_{\mathbf{q},b} \hbar\omega_b(\mathbf{q})}_{\text{ZPE}} + k_{\text{B}}T \sum_{\mathbf{q},b} \ln\left[1-\exp\left(\frac{-\hbar\omega_b(\mathbf{q})}{k_{\text{B}}T}\right)\right] \end{aligned} \tag{2.74}$$

The first term is the zero-point energy (ZPE), which will be the focus of Chapters 3 and 6. The partition function also yields the vibrational entropy of the phonon system

$$S_{\text{phonon}}(T) = \frac{1}{T}\sum_{\mathbf{q},b} \frac{\hbar\omega_b(\mathbf{q})}{\exp\left(\frac{-\hbar\omega_b(\mathbf{q})}{k_{\text{B}}T}\right)-1} - k_{\text{B}} \sum_{\mathbf{q},b} \ln\left[1-\exp\left(\frac{-\hbar\omega_b(\mathbf{q})}{k_{\text{B}}T}\right)\right] \quad . \tag{2.75}$$

The internal energy of the phonon system at finite temperature thus becomes

$$\begin{aligned} F_{\text{vib}}(T) &= F_{\text{phonon}}(T) + TS_{\text{phonon}}(T) \\ &= \frac{1}{2}\sum_{\mathbf{q},b} \hbar\omega_b(\mathbf{q}) + \sum_{\mathbf{q},b} \frac{\hbar\omega_b(\mathbf{q})}{\exp\left(\frac{-\hbar\omega_b(\mathbf{q})}{k_{\text{B}}T}\right)-1} \quad , \end{aligned} \tag{2.76}$$

which at $T = 0\,\text{K}$ is equal to the ZPE introduced in Eq. (2.74).

2.4.4 Quasi-Harmonic Approximation

Thermal expansion is not described by the harmonic approximation, which can be rationalized by the fact that the expectation value for the position of each phonon oscillator is zero. That means that even at higher temperatures the vibrational free energy given by Eq. (2.74) still describes a crystal where the atoms are (on average) at the same equilibrium positions as for zero temperature. Substituting the equilibrium

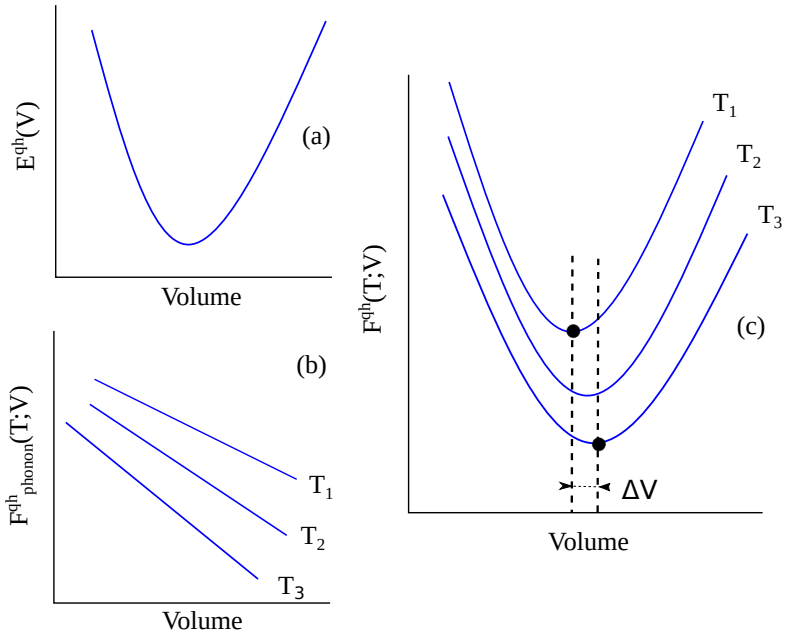

Figure 2.8: Schematic representation of (a) the internal energy $E^{\mathrm{qh}}(V)$ together with (b) the volume dependence of the phonon contribution to the Helmholtz free energy $F^{\mathrm{qh}}_{\mathrm{phonon}}(T;V)$ as well as (c) the total Helmholtz free energy $F^{\mathrm{qh}}(T;V) = E^{\mathrm{qh}}(V)+F^{\mathrm{qh}}_{\mathrm{phonon}}(T;V)$ according to the quasi-harmonic approximation. $F^{\mathrm{qh}}_{\mathrm{phonon}}(T;V)$ and $F^{\mathrm{qh}}(T;V)$ are shown for three different temperatures $T_3 > T_2 > T_1$. $\Delta V = V_0(T_3) - V_0(T_1)$ denotes the volume expansion as further detailed in the text.

lattice vectors $\mathbf{R}^{\mathrm{eq}}$ by general lattice vectors $\mathbf{R} = (\mathbf{a}, \mathbf{b}, \mathbf{c})$ in Eq. (2.65) leads to lattice-vector-dependent force constants $\Phi^{n',i',\alpha'}_{n,i,\alpha}(\mathbf{u} = \mathbf{0}; \mathbf{R})$. Consequently, the internal energy E^{harm} can be written as a function of the unit cell volume $V(\mathbf{R}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$

$$E^{\mathrm{qh}}(\mathbf{u}; V(\mathbf{R})) = E^{\mathrm{harm}}(\mathbf{u}; \mathbf{R}) \tag{2.77}$$

such that

$$\min_V E^{\mathrm{qh}}(\mathbf{u} = \mathbf{0}; V(\mathbf{R})) = E^{\mathrm{harm}}(\mathbf{u} = \mathbf{0}; V^{\mathrm{eq}}) = E^{\mathrm{eq}} \quad , \tag{2.78}$$

with $V^{\mathrm{eq}} = V(\mathbf{R}^{\mathrm{eq}})$. This is called quasi-harmonic approximation (QHA). $E^{\mathrm{qh}}(V)$ is schematically depicted in Fig. 2.8 (a).[9] Obviously, $E^{\mathrm{qh}}(V)$ alone still does not describe any thermal expansion.

Together with the force constants also the dynamical matrix and thus the solutions of the lattice dynamical equations (see Section 2.4.2) depend on V, i.e., in particular $\omega = \omega(\mathbf{q}; V)$. Substitution into the expressions introduced in Section 2.4.3 results in a volume dependent phonon contribution to the Helmholtz free energy $F^{\mathrm{qh}}_{\mathrm{phonon}}(T; V)$, which is shown schematically in Fig. 2.8 (b) for three different temperatures. Accounting for the internal and phonon contributions according to the QHA, the total Helmholtz free energy becomes

$$F^{\mathrm{qh}}(T; V) = E^{\mathrm{qh}}(V) + F^{\mathrm{qh}}_{\mathrm{phonon}}(T; V) \quad , \tag{2.79}$$

which is depicted in panel (c) of Fig. 2.8. $F^{\mathrm{qh}}_{\mathrm{phonon}}(T; V)$ leads to a temperature-dependent shift of the minimum of $E^{\mathrm{qh}}(V)$. The equilibrium volume is obtained from $\min_V F^{\mathrm{qh}}(T; V) = F(T; V_0(T))$ and changes according to $\Delta V^{\mathrm{eq}}(T) = V_0(T) - V^{\mathrm{eq}}$ compared to the $T = 0\,\mathrm{K}$. This means that the canonical ensemble given by the QHA can effectively describe volume expansion due to increasing temperature.

2.4.5 Computational Aspects

The essential computational challenge for phonon calculations is to obtain the force constants elements $\Phi^{n',i',\alpha'}_{n,i,\alpha}$ (see Eq. (2.65)), because this defines the computationally

[9] Assuming that for a given set of lattice vectors $\mathbf{R}$ only relaxed atomic coordinates are being considered (i.e., $\mathbf{u} = \mathbf{0}$), the dependence on the displacements and lattice vectors are no longer explicitly denoted here and in the following.

rather simple eigenvalue given in Eq. (2.71) the solution of which then yields phonon frequencies and displacements vectors and thus all thermodynamic properties describe in Section 2.4.3. Unfortunately, in particular for electronic structure calculations in periodic boundary conditions, second derivatives of the total internal energy with respect to atomic (displacement) coordinates are usually not as "easily" available as the corresponding first derivatives (forces). Density-functional perturbation theory (DFPT) [97] has been formulated for that reason, but (obviously) cannot be combined with force fields. To treat both DFT and force fields at equal footing, all phonon calculations in this thesis are based on the so-called finite displacement technique [98, 99]. The only systematic drawback of this technique is that it cannot account for the splitting of longitudinal and transversal optical phonon modes at the Γ-point (known as LO-TO splitting), which however is negligible for properties that are integrated over the entire Brillouin zone.

The basic idea behind the finite-displacement technique is to evaluate the elements of the force constants matrix by using finite differences of the respective force components

$$\Phi_{n,i,\alpha}^{n',i',\alpha'} \approx \frac{F_{n',i',\alpha'}(+u_{n,i,\alpha}) - F_{n',i',\alpha'},(-u_{n,i,\alpha})}{2\Delta} \quad . \tag{2.80}$$

Here $F_{n',i',\alpha'}(\pm u_{n,i,\alpha})$ is the force component α' acting on atom i' in unit cell n' of the crystal after only the single atom i in unit cell n has been displaced by $\pm u_{n,i,\alpha} = \pm\Delta$ along the Cartesian direction α while all other atoms remain in their equilibrium positions.[10] Of course, the latter is problematic when forces are calculated within periodic boundary conditions, because this always results in the displacement of infinitely many atoms. Using a supercell of the primitive unit cell allows to displace at least a certain amount of atom pairs $((n',i'),(n,i))$ with $(n',i') \neq (n,i)$ independently from each other. These pairs can be up to a certain maximum distance apart, which is inherent to the chosen supercell. Assuming that $\Phi_{n,i,\alpha}^{n',i',\alpha'} = 0$ for atom pairs at even larger distances then allows to obtain a reasonable approximation for the force constant matrix if the supercell is indeed large enough. This can and needs to be tested in practice by

[10] For a given set of lattice vectors.

checking the convergence of calculated phononic properties with respect to the supercell. The dynamical matrix resulting from the former is then exact for phonons with wavevectors $\mathbf{q}$ that are commensurate with the supercell, and Eq. (2.69) describes a Fourier interpolation for all other wave vectors. Consequently, for a given force constant matrix, phonons can be calculated with very modest computational effort throughout the entire Brillouin zone, and convergence of phononic properties with respect to the $\mathbf{q}$-grid used for the Brillouin zone sampling can be easily checked.

In principal, the central difference approximation used in Eq. (2.80) could also be replaced by forward or backward differencing, which both reduce the amount of force evaluations because only $F_{n',i',\alpha'}(+u_{n,i,\alpha})$ or $F_{n',i',\alpha'}(-u_{n,i,\alpha})$ would be required, respectively. At the same time, however, this results in a numerical truncation error for $\Phi_{n,i,\alpha}^{n',i',\alpha'}$ of order Δ, while it is of order Δ^2 with the central difference approximation. This is of particular importance since the displacement size Δ cannot be chosen arbitrarily small due to

1. limitations of floating point arithmetics when evaluating Eq. (2.80) and
2. numerical noise making it exceedingly difficult to accurately obtain "small" non-analytically defined forces (usually the case in electronic structure calculations).

The quasi-harmonic approximation (see Section 2.4.4) relies on solving lattice dynamics for harmonic Hamiltonians according to Eq. (2.71), but now for a different one for each different volume. Depending on the volume range to be covered, this comes at only moderately increased computational costs – in particular when considering alternatives as, e.g., a fully dynamical description. Furthermore, computer codes for lattice dynamics often already include the QHA or can be rather easily extended otherwise.

In this thesis, the PHONOPY package [100] is employed for all phonon calculations. Most important for this thesis is that it can be easily coupled to external codes that perform the force evaluations thanks to the well-structured and documented PYTHON code base. PHONOPY implements the finite displacement method originally introduced by Parlinski *et al.* [99], including an automatic recognition and exploitation of symme-

try to (potentially) significantly reduce the amount of $2 \times 3 \times N_{\text{cell}}$ force evaluations that are required when symmetry is not considered (or not present in the structure). Furthermore, it has been extended to enable calculations within the QHA. In the first step, a set of symmetry-reduced atomic displacements in the requested supercell are generated. After calculating forces for these structures with the external code (DFT or force fields in this work), PHONOPY is again invoked to generate the force constant matrix in Eq. (2.80) and solve the phonon eigenvalue problem (see Eq. (2.71)). In the QHA these steps need to be repeated for a set of different primitive cells each with a different cell volume. In the last step, e.g. phonon DOSs (Eq. (2.72)) or vibrational free energy (and entropy) contributions can then be calculated.

2.5 References

1. M. Born, R. Oppenheimer, *Annalen der physik* **389**, 457–484 (1927).
2. P. Hohenberg, W. Kohn, *Phys. Rev.* **136**, B864 (1964).
3. J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, *et al.*, *J. Chem. Phys* **123**, 062201 (2005).
4. D. M. Ceperley, B. J. Alder, *Phys. Rev. Lett.* **45**, 566–569 (1980).
5. M. Gell-Mann, K. A. Brueckner, *Phys. Rev.* **106**, 364 (1957).
6. S. H. Vosko, L. Wilk, M. Nusair, *Canadian J. Phys.* **58**, 1200–1211 (1980).
7. J. P. Perdew, A. Zunger, *Phys. Rev. B* **23**, 5048–5079 (1981).
8. J. P. Perdew, Y. Wang, *Phys. Rev. B* **45**, 13244–13249 (1992).
9. O. Gunnarsson, B. I. Lundqvist, *Phys. Rev. B* **13**, 4274 (1976).
10. K. Burke, J. P. Perdew, M. Ernzerhof, *J. Chem. Phys* **109**, 3760–3771 (1998).
11. C. Lee, D. Vanderbilt, K. Laasonen, R. Car, M Parrinello, *Phys. Rev. B* **47**, 4863 (1993).
12. D. Hamann, *Phys. Rev. B* **55**, R10157 (1997).
13. P. J. Feibelman, *Physical Chemistry Chemical Physics* **10**, 4688–4691 (2008).
14. C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **37**, 785–789 (1988).
15. J. P. Perdew, P Ziesche, H Eschrig, *Electronic structure of solids' 91*, 1991.
16. J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
17. J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
18. J. Sun, A. Ruzsinszky, J. P. Perdew, *Phys. Rev. Lett.* **115**, 036402 (2015).
19. V. N. Staroverov, G. E. Scuseria, J. Tao, J. P. Perdew, *J. Chem. Phys* **119**, 12129–12137 (2003).
20. Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **109**, 5656–5667 (2005).
21. J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, *et al.*, *Nature chemistry* **8**, 831 (2016).
22. Y. Zhang, J. Sun, J. P. Perdew, X. Wu, *Phys. Rev. B* **96**, 035143 (2017).
23. E. R. Johnson, I. D. Mackie, G. A. DiLabio, *J. Phys. Org. Chem.* **22**, 1127–1135 (2009).
24. S Grimme, L Goerigk, R. Fink, *Rev.: Comput. Mol. Sci* **1**, 211 (2011).
25. A. Tkatchenko, L. Romaner, O. T. Hofmann, E. Zojer, C. Ambrosch-Draxl, *et al.*, *MRS bulletin* **35**, 435–442 (2010).
26. J. Klimeš, A. Michaelides, *J. Chem. Phys* **137**, 120901 (2012).
27. F. London, *Transactions of the Faraday Society* **33**, 8b–26 (1937).
28. H. B. Casimir, D. Polder, *Phys. Rev.* **73**, 360 (1948).
29. H Sutter, R. Cole, *J. Chem. Phys* **52**, 132–139 (1970).
30. V. A. Maltsev, O. A. Nerushev, S. A. Novopashin, B. A. Selivanov, *Chemical physics letters* **212**, 480–482 (1993).
31. A. Salop, E. Pollack, B. Bederson, *Phys. Rev.* **124**, 1431 (1961).

32. A. Tkatchenko, M. Scheffler, *Phys. Rev. Lett.* **102**, 073005 (2009).
33. X. Chu, A. Dalgarno, *J. Chem. Phys.* **121**, 4083–4088 (2004).
34. F. L. Hirshfeld, *Theoretica chimica acta* **44**, 129–138 (1977).
35. P. Jurečka, J. Černý, P. Hobza, D. R. Salahub, *J. Comp. Chem.* **28**, 555–569 (2007).
36. A. Tkatchenko, R. A. DiStasio, R. Car, M. Scheffler, *Phys. Rev. Lett.* **108**, 236402 (2012).
37. D. C. Langreth, M. Mehl, *Phys. Rev. Lett.* **47**, 446 (1981).
38. M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, B. I. Lundqvist, *Phys. Rev. Lett.* **92**, 246401 (2004).
39. K. Lee, É. D. Murray, L. Kong, B. I. Lundqvist, D. C. Langreth, *Phys. Rev. B* **82**, 081101 (2010).
40. O. A. Vydrov, T. Van Voorhis, *Phys. Rev. Lett.* **103**, 063004 (2009).
41. O. A. Vydrov, T. Van Voorhis, *J. Chem. Phys.* **133**, 244103 (2010).
42. J. Wang, G. Román-Pérez, J. M. Soler, E. Artacho, M.-V. Fernández-Serra, *J. Chem. Phys* **134**, 024516 (2011).
43. É. D. Murray, G. Galli, *Phys. Rev. Lett.* **108**, 105502 (2012).
44. G. Román-Pérez, J. M. Soler, *Phys. Rev. Lett.* **103**, 096102 (2009).
45. V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, *et al.*, *Comp. Phys. Comm.* **180**, 2175–2196 (2009).
46. H. J. Monkhorst, J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
47. H. Berendsen, J. Grigera, T. Straatsma, *J. Phys. Chem.* **91**, 6269–6271 (1987).
48. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys* **79**, 926–935 (1983).
49. J. L. Abascal, C. Vega, *J. Chem. Phys* **123**, 234505 (2005).
50. J. L. F. Abascal, E. Sanz, R. García Fernández, C. Vega, *J. Chem. Phys.* **122**, 234511 (2005).
51. S. Habershon, T. E. Markland, D. E. Manolopoulos, *J. Chem. Phys* **131**, 024501 (2009).
52. P. Ren, J. W. Ponder, *J. Phys. Chem. B* **107**, 5933–5947 (2003).
53. J. W. Ponder *et al.*, *Washington University School of Medicine, Saint Louis, MO* **3** (2004).
54. V. Babin, C. Leforestier, F. Paesani, *J. Chem. Theory Comput.* **9**, 5395–5403 (2013).
55. V. Babin, G. R. Medders, F. Paesani, *J. Chem. Theory Comput.* **10**, 1599–1607 (2014).
56. K. T. Wikfeldt, E. Batista, F. Vila, H Jónsson, *Physical Chemistry Chemical Physics* **15**, 16542–16556 (2013).
57. E. Lambros, F. Paesani, *The Journal of Chemical Physics* **153**, 060901 (2020).
58. R. A. Gilmore, M. T. Dove, A. J. Misquitta, *Journal of chemical theory and computation* **16**, 224–242 (2019).

59. G. A. Cisneros, K. T. Wikfeldt, L. Ojam´ae, J. Lu, Y. Xu, *et al.*, *Chemical reviews* **116**, 7501–7528 (2016).
60. H. J. Berendsen, J. P. Postma, W. F. van Gunsteren, J. Hermans, in *Intermolecular forces* (Springer, 1981), pp. 331–342.
61. F. Paesani, W. Zhang, D. A. Case, T. E. Cheatham III, G. A. Voth, *J. Chem. Phys* **125**, 184507 (2006).
62. C. Burnham, D. Anick, P. Mankoo, G. Reiter, *J. Chem. Phys* **128**, 154519 (2008).
63. P. E. Lopes, B. Roux, A. D. MacKerell, *Theoretical Chemistry Accounts* **124**, 11–28 (2009).
64. J. Applequist, J. R. Carl, K.-K. Fung, *J. Am. Chem. Soc.* **94**, 2952–2960 (1972).
65. B. T. Thole, *Chemical Physics* **59**, 341–350 (1981).
66. G. S. Fanourgakis, S. S. Xantheas, *J. Chem. Phys* **128**, 074506 (2008).
67. R. Kumar, F.-F. Wang, G. R. Jenness, K. D. Jordan, *J. Chem. Phys* **132**, 014309 (2010).
68. R. Bukowski, K. Szalewicz, G. C. Groenenboom, A. van der Avoird, *J. Chem. Phys* **128**, 094314 (2008).
69. R. Bukowski, K. Szalewicz, G. C. Groenenboom, A. van der Avoird, *J. Chem. Phys* **128**, 094313 (2008).
70. Y. Wang, X. Huang, B. C. Shepler, B. J. Braams, J. M. Bowman, *J. Chem. Phys* **134**, 094509 (2011).
71. Y. Wang, V. Babin, J. M. Bowman, F. Paesani, *J. Am. Chem. Soc.* **134**, 11116–11119 (2012).
72. J. D. Bernal, R. H. Fowler, *J. Chem. Phys* **1**, 515–548 (1933).
73. D. Frenkel, B. Smit, *Understanding Molecular Simulation – From Algorithms to Applications* (Academic Press, San Diego, 2002), vol. 1.
74. K. Toukan, A. Rahman, *Phys. Rev. B* **31**, 2643 (1985).
75. D. M. Carey, G. M. Korenowski, *J. Chem. Phys.* **108**, 2669–2675 (1998).
76. S. Plimpton, *J. Comput. Phys.* **117**, 1–19 (1995).
77. M. L. Laury, L.-P. Wang, V. S. Pande, T. Head-Gordon, J. W. Ponder, *J. Phys. Chem. B* **119**, 9423–9437 (2015).
78. A. Stone, *The Theory of Intermolecular Forces* (OUP Oxford, 2013).
79. P. T. Van Duijnen, M. Swart, *J. Phys. Chem. A* **102**, 2399–2407 (1998).
80. T. A. Halgren, *J. Am. Chem. Soc.* **114**, 7827–7843 (1992).
81. M. J. Dudek, J. W. Ponder, *J. Comp. Chem.* **16**, 791–816 (1995).
82. J. W. Ponder *et al.*, *AMOEBA-WATER-2014 parameters.*
83. T. R. Dyke, J. Muenter, *J. Chem. Phys* **59**, 3125–3127 (1973).
84. J Verhoeven, A Dymanus, *J. Chem. Phys* **52**, 3222–3233 (1970).
85. E. R. Batista, S. S. Xantheas, H. Jónsson, *J. Chem. Phys* **109**, 4546–4551 (1998).
86. K. Tang, J. P. Toennies, *J. Chem. Phys* **80**, 3726–3741 (1984).
87. G. R. Medders, V. Babin, F. Paesani, *J. Chem. Theory Comput.* **10**, 2906–2910 (2014).

88. V. Babin, G. R. Medders, F. Paesani, *J. Phys. Chem. Lett.* **3**, 3765–3769 (2012).
89. G. R. Medders, V. Babin, F. Paesani, *J. Chem. Theory Comput.* **9**, 1103–1114 (2013).
90. S. S. Xantheas, *Chemical Physics* **258**, 225–231 (2000).
91. M. P. Hodges, A. J. Stone, S. S. Xantheas, *J. Phys. Chem. A* **101**, 9163–9168 (1997).
92. A. Hermann, R. P. Krawczyk, M. Lein, P. Schwerdtfeger, I. P. Hamilton, *et al.*, *Phys. Rev. A* **76**, 013202 (2007).
93. J. Cui, H. Liu, K. D. Jordan, *J. Phys. Chem. B* **110**, 18872–18878 (2006).
94. H. Partridge, D. W. Schwenke, *The Journal of Chemical Physics* **106**, 4618–4639 (1997).
95. C. J. Burnham, J. Li, S. S. Xantheas, M. Leslie, *J. Chem. Phys.* **110**, 4566–4581 (1999).
96. C. J. Burnham, S. S. Xantheas, *J. Chem. Phys.* **116**, 1479–1492 (2002).
97. S. Baroni, S. De Gironcoli, A. Dal Corso, P. Giannozzi, *Reviews of modern Physics* **73**, 515 (2001).
98. W Frank, C Elsässer, M Fähnle, *Phys. Rev. Lett.* **74**, 1791 (1995).
99. K Parlinski, Z. Li, Y Kawazoe, *Phys. Rev. Lett.* **78**, 4063 (1997).
100. A. Togo, I. Tanaka, *Scripta Mater.* **108**, 1–5 (2015).