universität
wien

MASTERARBEIT

Titel der Masterarbeit

„A simulation study on the thermo-polarisation effect in water“

Verfasser

Peter Wirnsberger, BSc MPhil

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2014

Studienkennzahl lt. Studienblatt:	A 066 876
Studienrichtung lt. Studienblatt:	Masterstudium Physik
Betreuerin / Betreuer:	Univ.-Prof. Mag. Dr. Christoph Dellago

Zusammenfassung

Starke Temperaturgradienten können in einem polaren Medium zu molekularer Reorientierung führen und es dadurch polarisieren. Dieses Phänomen wird als „Thermo-Polarisationseffekt“ bezeichnet. Kürzlich ist es Bresme und Mitarbeitern gelungen, das in Wasser dabei entstehende Feld mit Hilfe von Computersimulationen nachzuweisen. Es konnte eine lineare Abhängigkeit des Feldes vom Temperaturgradienten demonstriert werden, was im Einklang mit den theoretischen Vorhersagen der Nichtgleichgewichtsthermodynamik steht. In der vorliegenden Arbeit wurde die Abhängigkeit des auftretenden elektrostatischen Feldes von der Behandlung langreichweitiger Wechselwirkungen untersucht. Dazu wurden Computersimulationen von Wasser mit zwei unterschiedlichen Summationstechniken, nämlich Wolf-Summation und Ewald-Summation, durchgeführt. Dabei ist aufgefallen, dass ein oftmals zuvor verwendeter Algorithmus die Gesamtenergie im System nicht erhält und dadurch ein erheblicher Energieverlust entsteht. Der Grund dafür konnte identifiziert und behoben werden, was zu einem neuen, verbesserten Algorithmus führt. Des Weiteren konnte gezeigt werden, dass die Behandlung langreichweitiger Wechselwirkungen nur geringe Auswirkungen auf das Feld hat, wenn die in der Simulation verwendeten Randbedingungen bei der Berechnung berücksichtigt werden. Diese Ergebnisse stehen damit im Gegensatz zu früheren Berichten.

Abstract

In a polar fluid, strong temperature gradients can lead to molecular reorientation and polarisation, a phenomenon called 'thermo-polarisation effect'. Recently Bresme and co-workers reported the establishment of an electrostatic field in simulations of liquid water in the presence of strong temperature gradients. They were able to demonstrate that the fields so obtained are proportional to the temperature gradients as predicted by non-equilibrium thermodynamics. Motivated by the work of Bresme and co-workers, we investigated the dependence of the electrostatic field on the underlying treatment of electrostatic interactions for Wolf summation and Ewald summation. We noticed that the algorithm employed previously causes an unphysical energy drift. Based on analytical work, we were able to identify the cause of this problem and suggest an improved algorithm that conserves the total energy in the system. Furthermore, in contrast to previous reports, we show that the polarisation effect is largely independent of the treatment of long-range electrostatic interactions, if the field is calculated in accordance with the employed boundary conditions.

Acknowledgements

I would first like to thank my supervisor, Professor Christoph Dellago, for his invaluable guidance and support throughout the entire project. He always helped me immediately when I was looking for advice which I highly appreciate. I am also very grateful to Professor Daan Frenkel for co-supervising the project and for providing me with the opportunity of carrying out part of my research in his group in Cambridge. I enjoyed many useful discussions with members of both research groups. I would especially like to thank Dr Aleks Reinhardt for carefully reading my manuscript and for his helpful comments.

I would also like to thank Professor Martin Neumann for his advice and many insightful ideas. His lecture series on Computational Physics, including the practical work, helped me to understand how molecular dynamics simulations are implemented from first principles and were therefore invaluable for the success of this project. The sections on electrostatic interactions owe almost everything to his elegant derivation of the Ewald sum from a Green's function perspective.

Financial support from the Federation of Austrian Industry (IV) Carinthia and the LLP Erasmus Programme is acknowledged with gratitude.

Finally, I would like to thank my friends and my family for their support. They have always believed in me and have encouraged me no matter what I was pursuing.

Contents

1 Introduction **1**

1.1 Motivation and previous work 1

1.2 Non-equilibrium thermodynamics 3

1.2.1 Entropy production and local equilibrium 3

1.2.2 Phenomenological equations 4

1.2.3 Stationary states 5

1.2.4 Thermal polarisation 6

1.2.5 Some critical comments 7

1.3 Molecular simulations 8

1.3.1 Non-equilibrium molecular dynamics 9

1.4 Outline 11

2 Governing equations of motion **13**

2.1 Short-range interaction 13

2.2 Long-range interaction 14

2.2.1 Ewald summation 15

2.2.2 Wolf summation 16

2.3 Numerical integration 17

3 Thermal gradient **19**

3.1 Heat exchange algorithm 20

3.1.1 Formulation 20

3.1.2 Energy conservation 22

3.2 Conservative heat exchange algorithm 23

3.2.1 Modified equation analysis 23

3.2.2 Formulation 27

3.2.3 Energy conservation 29

4 Thermal polarisation in water **31**

4.1 Simulation setup 32

4.1.1 The SPC/E model 32

4.1.2 Box setup 33

4.2 Simulation protocol 33

4.3 Validation 34

4.3.1 Radial distribution function and VACF 35

4.3.2 Kirkwood g-factor and dielectric constant 35

4.3.3 Energy conservation 39

4.3.4 Centre of mass velocity 40

4.3.5 Coordinate constraints 40

4.4 Calculation of the electrostatic field 43

4.4.1 Ewald summation 44

4.4.2 Wolf summation 45

4.5 Results 46

4.5.1 Temperature 46

4.5.2 Number densities and charge density 48

4.5.3 Electrostatic potential 50

4.5.4 Electrostatic field 50
4.5.5 Comparison with literature data 52
4.6 Discussion 54

5 Conclusion and future work 57

References 59

Appendix A Long-range interaction 63

A.1 Ewald summation 63
A.2 Wolf summation 64

Appendix B Heat exchange algorithm 65

B.1 Formulation 65
B.2 Lennard-Jones NEMD simulation 66

Appendix C Electrostatic field 70

C.1 Wolf summation 70

Chapter 1

Introduction

In this work we investigate the behaviour of water subjected to a strong thermal gradient. In particular, we carry out a series of non-equilibrium molecular dynamics (NEMD) simulations to study the effect of thermal polarisation due to a strong gradient in temperature. We first explain our motivation for carrying out this work in §1.1. The underlying theoretical framework is given by non-equilibrium thermodynamics (NET) and a brief summary of the relevant predictions of the theory is given in §1.2. Subsequently, the basic idea of NEMD is introduced in §1.3.1 in order to illustrate how the technique can be utilised for studying the system under consideration. Finally, we give an outline of the remainder of this report in §1.4.

1.1 Motivation and previous work

Non-equilibrium phenomena are ubiquitous in nature and of importance in chemistry, physics, biology and materials science [1], to name but a few. Several of these effects are related to strong thermal gradients which may, for example, be caused by chemical reactions at surfaces [1] or local hot spots arising from ultrasonic insonation [2]. The Peltier effect, as an example of a thermoelectric effect, as well as the Soret effect, or thermophoresis, both fall in this category [1]. From a practical standpoint, this could be of significance for reusing dissipated heat by converting it into other forms of energy [3]. As water is undoubtedly a very important substance in all these fields [4], understanding its microscopic behaviour under such conditions is highly relevant.

Recently, Bresme and co-workers expended considerable effort investigating the behaviour of water in the vicinity of strong thermal gradients [1, 3–8]. Motivated by the predictions of NET, they carried out NEMD simulations and were able to confirm the hypothesised rearrangement of molecules in response to a temperature gradient, consequently leading to a net polarisation. Furthermore, they could demonstrate that the induced electrostatic field does indeed scale linearly with the temperature gradient. The reported fields were as high as 10^8 V/m for gradients of 3 K/Å. In subsequent work [3], Armstrong and co-workers determined the thermodynamic states that maximise the polarisation effect, thus providing a useful guideline for future investigations.

We find these results very interesting and believe that the findings of Bresme and co-workers are potentially relevant for a phenomenon called *sonocrystallisation*. This phenomenon comprises the enhancement of crystallisation processes by ultrasonic radiation. Many researchers agree that cavitation is a crucial factor for sonocrystallisation. Cavitation is the process of nucleation in a liquid when the pressure falls below the vapor pressure. If sound waves lead to the formation of cavitation bubbles, it is referred to as acoustic cavitation. The bubbles so formed can either oscillate around a stable equilibrium radius (stable cavitation) or collapse violently after an initial period of rapid growth (transient cavitation). The latter can trigger a shock wave and very high pressures and temperatures in the vicinity of the collapsing bubble [9, 10]. The resulting temperature gradients were estimated to be comparable to the ones considered in previous investigations on the thermo-polarisation effect [11]. We therefore think that a better understanding of molecular reorientation as a response to strong thermal gradients could be of direct importance for sonocrystallisation.

Another potential application could be the effective interaction of nanoparticles in a water bath. Experimentally, it is nowadays possible to heat up nanoparticles individually, for example by electromagnetic irradiation [12, 13]. Furthermore, it is known that the temperature decays as r^{-1} around a nanoparticle [12]. For strong enough gradients, the thermo-polarisation effect might lead to a non-negligible effective Coulomb interaction. This effect could be probed in NEMD simulations and experiments.

As a starting point, we decided to follow the approach of Bresme and co-workers in order to confirm their results. Surprisingly, we found that the treatment of the long-range electrostatic interactions has a strong influence on the resulting field if the field is calculated as suggested by Bresme and co-workers [1, 3, 8]. In particular, we could only reproduce their results with the Wolf method [14], but not with the commonly used Ewald summation technique. In a very recent paper [3], Armstrong and co-workers claim to have shown that both implementations, Ewald summation and Wolf summation, produce the same results referring to one of their earlier papers [8]. However, to the best of our knowledge, they did not present the results for the electrostatic field obtained with Ewald summation [8]. This is intriguing, as the authors undermine the validity of the Wolf method for their setup by comparing various simulation quantities to their Ewald summation counterparts. What is more, we found that the heat exchange (HEX) algorithm [15, 16], previously used to establish the temperature gradient, leads to a considerable loss of energy over the relevant simulation time scales.

In this work we closely follow the simulation protocol of Armstrong and co-workers [8] and reproduce some of their results for validation. Subsequently, we argue that the way they calculated the electrostatic field in several recent publications is incorrect. This mistake partly explains the strong dependence of the electrostatic field on the underlying treatment of the electrostatic interactions. We will show that both methods lead to very similar results if the fields are calculated consistently. Additionally, we identify the major cause for the energy loss in the HEX algorithm and suggest an easy fix based on some analytical work. To the best of our knowledge, neither of these issues has successfully been addressed in previous work.

1.2 Non-equilibrium thermodynamics

In order to present the basic concepts of non-equilibrium thermodynamics, we will closely follow de Groot and Mazur [17]. Unless explicitly stated otherwise, all the derivations shown in the sections §1.2.1–1.2.3 are taken from that source. A full derivation of equation (1.23) relevant for the later chapters would be beyond the scope of this work and may be found elsewhere [1, 17].

1.2.1 Entropy production and local equilibrium

As a starting point, we consider the variation of the entropy

$$\mathrm{d}S = \mathrm{d_i}S + \mathrm{d_e}S, \tag{1.1}$$

where $\mathrm{d_i}S$ arises from the entropy production inside the system and $\mathrm{d_e}S$ is the contribution to the entropy supplied by the surroundings. The second law of thermodynamics may then be written as

$$\mathrm{d_i}S \geq 0, \tag{1.2}$$

where the equality holds for reversible transformations [17]. The aim now is to relate the internal entropy production, $\mathrm{d_i}S$, to the irreversible processes of interest that occur inside the system. The contributions to the entropy in Eqn (1.1) can be expressed as

$$S = \int\limits_V \rho_\mathrm{m} s \; \mathrm{d}V, \tag{1.3}$$

$$\frac{\mathrm{d}S_\mathrm{i}}{\mathrm{d}t} = \int\limits_V \sigma \; \mathrm{d}V, \tag{1.4}$$

$$\frac{\mathrm{d}S_\mathrm{e}}{\mathrm{d}t} = -\int\limits_\Omega \mathbf{J}_\mathrm{s,tot} \cdot \mathrm{d}\boldsymbol{\Omega}, \tag{1.5}$$

where Ω is the enclosing surface of the volume V, ρ_m the mass density, s the entropy per unit mass, σ the entropy production per unit volume and unit time and $\mathbf{J}_\mathrm{s,tot}$ the total entropy flow per unit area and unit time [17]. As the equations (1.3)–(1.5) must hold for an arbitrary volume, it follows from Eqn (1.1) that

$$\frac{\partial \rho_\mathrm{m} s}{\partial t} = -\nabla \cdot \mathbf{J}_\mathrm{s,tot} + \sigma, \tag{1.6}$$

$$\sigma \geq 0. \tag{1.7}$$

By writing Eqns (1.1)–(1.2) in the form of Eqns (1.6)–(1.7), it is implicitly assumed that the same macroscopic laws also apply for infinitesimally small parts of the system. Additionally, it is assumed that the system is in *local equilibrium*. In order to explain what this means, let us consider a system whose equilibrium state can be defined by the specific internal energy u and the specific volume v, i.e. $s = s(u, v)$. The assumption of local equilibrium then implies that for a small volume element there exists a local entropy function $s_\mathrm{l}(u, v)$ which is the same function as s, although the overall system is not in equilibrium. This is a very important assumption, as it implies that the entropy variation is still given by

$$T\mathrm{d}s = \mathrm{d}u + p\mathrm{d}v, \tag{1.8}$$

$$T\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\mathrm{d}u}{\mathrm{d}t} + p\frac{\mathrm{d}v}{\mathrm{d}t}, \tag{1.9}$$

even though the overall system is out of equilibrium [17]. Equation (1.9) is relevant for the entropy balance equation, Eqn (1.6), and based on the assumption that Eqn (1.8) is valid along the centre of gravity motion of a small mass element [17].

1.2.2 Phenomenological equations

From the preceding section, we know that the entropy production σ vanishes when the system is in thermodynamic equilibrium. The next step is to relate the entropy production to the independent fluxes and thermodynamic forces. Several phenomenological laws, such as Fick's law or Fourier's law, Eqn (1.13), suggest that there is a linear relation between the flow and the corresponding thermodynamic force. Allowing also for cross-effects, we can therefore write

$$J_i = \sum_k L_{ik} X_k, \tag{1.10}$$

$$\sigma = \sum_i J_i X_i, \tag{1.11}$$

where J_i is any Cartesian component of the independent flux and X_i is a component of the thermodynamic force [17]. Moreover, the *phenomenological coefficients* L_{ik} were introduced and the equations (1.10) are called *phenomenological equations*. The phenomenological coefficients satisfy certain conditions based on spatial symmetry properties (Curie principle) and the time reversal symmetry of the microscopic equations of motion (Onsager reciprocal relations) [17].

1.2.3 Stationary states

If the state variables are independent of time, the thermodynamic state is called a *stationary state*. For a one-component system at constant volume, a stationary state would, for example, be defined by $\frac{\partial u}{\partial t} = 0$. For a wide range of phenomena including thermal conduction, diffusion, chemical reactions and cross-effects, it can be shown that stationary states minimise the entropy production [17]. One could therefore in theory quantitatively determine the stationary state by minimising the entropy production.

For simplicity we will only present the introductory example of Ref. [17] *(Chapter V, §3)* in a stepwise manner and show the proof of this assertion for a solid one-component isotropic system in the absence of viscous phenomena. Let the substance be contained in a vessel and the temperature at the walls be independent of time with non-uniform values. The entropy production for this case is given by

$$\sigma = \mathbf{J}_q \cdot \nabla \frac{1}{T}, \tag{1.12}$$

$$\mathbf{J}_q = L_{qq} \nabla \frac{1}{T}, \tag{1.13}$$

with $\mathbf{J}_q$ being the heat flow, T the temperature and $L_{qq} = \lambda T^2$ the corresponding phenomenological coefficient, where λ denotes the thermal conductivity. For this proof we will assume that L_{qq} is constant throughout the system and depends only on the overall equilibrium temperature [17]. As the system is solid by construction, we are justified in neglecting the thermal expansion and we can write the energy equation as

$$\rho_{\mathrm{m}} \frac{\partial u}{\partial t} = \rho_{\mathrm{m}} c_v \frac{\partial T}{\partial t} = -\nabla \cdot \mathbf{J}_q, \tag{1.14}$$

with c_v being the constant volume specific heat. Considering the total entropy production P given by

$$P = \int\limits_V \sigma \mathrm{d}V = \int\limits_V \mathbf{J}_q \cdot \nabla \frac{1}{T} \, \mathrm{d}V = \int\limits_V L_{qq} \left(\nabla \frac{1}{T} \right)^2 \mathrm{d}V, \tag{1.15}$$

we can then try to solve the variational problem

$$\delta P = \delta \left(\int\limits_V L_{qq} \left(\nabla \frac{1}{T} \right)^2 \mathrm{d}V \right) = 0 \tag{1.16}$$

in order to minimise the entropy production. Together with Eqn (1.13) and the assumption of a constant temperature at the boundary and constant L_{qq}, the solution to this problem is given by [17]

$$\nabla \cdot \mathbf{J}_q = 0. \tag{1.17}$$

From Eqn (1.14) it then follows that

$$\frac{\partial u}{\partial t} = c_v \frac{\partial T}{\partial t} = 0, \tag{1.18}$$

and it can be concluded that the stationary state compatible with the boundary conditions minimises the entropy production. Under the same assumptions it can be proven that this state is also stable with respect to local perturbations in T [17].

From this example we learn that if we are able to relate the irreversible phenomenon of heat conduction to the entropy production σ, it follows from a variational principle that the state of minimum entropy production is a stationary state. This state can be either an equilibrium or a non-equilibrium stationary state depending on the imposed boundary conditions. Although we have only presented the proof of this theorem for a simple example, it can be shown in a more general way under the following assumptions [17]: local equilibrium must hold and the phenomenological coefficients are constants. Additionally, the Onsager reciprocal relations have the form $L_{ik} = L_{ki}$. Furthermore, the phenomenological equations (1.10) hold, i.e. the fluxes depend only linearly on the thermodynamic forces. However, even if not all of the above conditions are satisfied, it is possible to derive a more general theorem [17].

1.2.4 Thermal polarisation

Using the same theoretical framework, we now consider the reorientation of an isotropic liquid in the vicinity of a thermal gradient [1]. If the system is polarisable, the entropy production is given by [1, 17]

$$\sigma = -\frac{1}{T} \frac{\partial \mathbf{P}}{\partial t} \cdot (\mathbf{E}_{\mathrm{eq}} - \mathbf{E}) - \frac{1}{T^2} \mathbf{J}_q \cdot \nabla T, \tag{1.19}$$

where $\mathbf{P}$ denotes the polarisation, $\mathbf{E}_{\text{eq}}$ the equilibrium electrostatic field and $\mathbf{E}$ the electrostatic field [1]. The polarisation is related to the equilibrium field according to

$$\mathbf{P} = \chi\epsilon_0\mathbf{E}_{\text{eq}}, \tag{1.20}$$

where $\chi = \epsilon_{\text{r}} - 1$ is the electric susceptibility, ϵ_{r} the relative dielectric constant and ϵ_0 the vacuum permittivity. Assuming that the phenomenological equations hold, the displacement current and the heat flux are given by [1]

$$\frac{\partial \mathbf{P}}{\partial t} = -\frac{L_{pp}}{T}(\mathbf{E}_{\text{eq}} - \mathbf{E}) - \frac{L_{pq}}{T^2}\nabla T, \tag{1.21}$$

$$\mathbf{J}_q = -\frac{L_{qp}}{T}(\mathbf{E}_{\text{eq}} - \mathbf{E}) - \frac{L_{qq}}{T^2}\nabla T. \tag{1.22}$$

If we write the fluxes in this form, the cross-phenomena become immediately apparent. The temperature gradient affects the displacement current and the electrostatic field affects the heat flow in turn. The phenomenological constants are again assumed to obey the Onsager reciprocal relations, i.e. $L_{pq} = L_{qp}$. At the stationary state and in the absence of free charges, i.e. $\frac{\partial \mathbf{P}}{\partial t} = 0$ and $\mathbf{P} = -\epsilon_0\mathbf{E}$, it follows from Eqn (1.21) that [1]

$$\mathbf{E} = \left(1 - \frac{1}{\epsilon_{\text{r}}}\right)\frac{L_{pq}}{L_{pp}}\frac{\nabla T}{T}, \tag{1.23}$$

which is the central result of this section. According to this equation, the electrostatic field arising from thermal reorientation is proportional to the imposed temperature gradient.

1.2.5 Some critical comments

The theory established by de Groot and Mazur has been heavily criticised by Jaynes in Ref. [18]. In particular, Jaynes questions the validity of the assumption that the phenomenological coefficients are constants. For the example of thermal conductivity, this assumption implies that $\lambda \propto T^{-2}$, and there is no known substance for which this scaling holds [18]. What is more, he argues that in the more general case of heat conduction, diffusion and chemical reactions, the condition of minimum entropy production simply reduces to the conservation laws, although very restrictive additional assumptions are necessary. Therefore, the necessity of the theory is questioned altogether. A detailed discussion can be found in Ref. [18].

We note that even without using the Onsager reciprocal relations or the assumption that the phenomenological coefficients are constants, a more general theorem can be derived; however, in that case, the state of minimum entropy production does not necessarily correspond to a stationary state [17]. The system investigated in this report is certainly not isotropic, which will be shown in Chapter 4. However, a theory without the restrictive assumptions and simplifications that were necessary to derive Eqn (1.23) is yet to be developed.

1.3 Molecular simulations

Computer simulations have become a powerful instrument for studying molecular systems on a microscopic basis. On the one hand, they complement experiments by providing valuable experimental parameters. On the other hand, some physical phenomena, such as crystallisation or chemical reactions, happen on time and length scales which are difficult to study experimentally. What is more, in many cases carrying out experiments would be either too dangerous or simply too expensive; one such example would be high-pressure fluid dynamics experiments. Molecular dynamics (MD) is a technique that allows us to study the time evolution of a microscopic system by solving Newton's equation of motion,

$$m_i \ddot{\mathbf{r}}_i = -\nabla_{\mathbf{r}_i} U(\mathbf{r}^N) = \mathbf{f}_i, \tag{1.24}$$

where $U(\mathbf{r}^N)$ is the many-body potential depending on the positions of all N atoms, $\mathbf{r}^N = \{\mathbf{r}_1, ..., \mathbf{r}_N\}$, $\nabla_{\mathbf{r}_i}$ denotes the gradient with respect to atom i and $\mathbf{f}_i$ is the force acting on atom i. Often $U(\mathbf{r}^N)$ is a sum of pairwise, spherically symmetric interactions of the form

$$U(\mathbf{r}^N) = \sum_{i=1}^{N} \sum_{j<i} u(r_{ij}), \tag{1.25}$$

where $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$. The system is evolved in time by a certain timestep and a vast amount of information is available. Thus, many thermodynamic quantities can readily be calculated with little or moderate overhead. However, computational power is a limiting factor and only comparatively small systems of $\mathcal{O}(10^3 - 10^6)$ molecules can be studied at present [19], depending on, for example, the type of interaction. Although a million molecules is several orders of magnitudes more than what was possible a few decades ago, we are still far from being able to simulate realistic system sizes comprising of the order of 10^{23} particles. Thus, special care has to be taken in order to avoid finite size effects. Let us consider the three-dimensional Coulomb potential to illustrate the problem. The potential decays very slowly as it is proportional to r^{-1}, with r being the distance between two charged atoms. The integral over the tail of the potential from a cutoff radius to infinity therefore diverges. Although for certain systems, truncating the long-range electrostatic interaction might be a reasonable approximation [14], it is not justified in general.

However, what one can do to reduce finite size effects considerably is to employ *periodic boundary conditions* and actually simulate an infinite system. This system

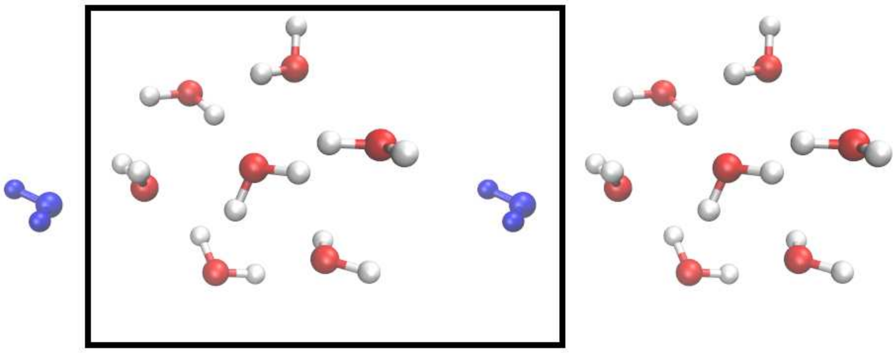

Figure 1.1: Periodic boundary conditions: a molecule (blue) that leaves the reference box (black solid line) will enter it again at the same relative position where the original molecule entered the neighbouring box.

can be thought of as an infinite number of identical copies of a reference box which are aligned in such a way that they fill up all space. Additionally, the molecules in different boxes can interact with each other. It is still impossible to simulate an infinite system of truly independent molecules, but we are capable of simulating an infinite system of periodically repeated units. The entire system can be constructed from the knowledge of one box, i.e. the reference box, and it suffices to keep track of the molecules in that box only. Once a molecule leaves the reference box, it enters a neighbouring box, as illustrated in Fig. 1.1. At the same time a periodic copy of this molecule will enter the reference box at the same relative position where the original molecule entered the neighbouring box. All this can be incorporated by two modifications: when calculating distances between molecules, say i and j, we take the closest distance between i and any image of j (*minimum image convention*). Additionally, the long-range potential contribution requires modification, as it should also take into account all the interactions within the infinite, periodic system.

1.3.1 Non-equilibrium molecular dynamics

Commonly, MD is used to study systems that are in equilibrium. If the system is Hamiltonian with no explicit time dependence, it is straightforward to carry out microcanonical, or NVE, simulations, as the system evolves on a constant energy surface in phase space. However, it is also possible to use a thermostat to adjust the temperature in a MD simulation. A Nosé–Hoover thermostat [20, 21], for example, can be coupled to the equations of motion in such a way that the resulting trajectory samples the canonical distribution. This particular case would correspond to an NVT simulation, but other ensembles are also possible.

In contrast to equilibrium MD, the system can be far from equilibrium in a non-equilibrium MD simulation. For instance, this would be the case when an external force couples to the system and causes strong dissipative fluxes and transport phenomena. Such a scenario can be studied with NEMD simulations and one usually forces the system to remain in a non-equilibrium stationary state [22].

Let us further illustrate the difference between equilibrium MD and NEMD using the calculation of the thermal conductivity as an example. For simplicity, we assume that the centre of mass velocity of our system is zero. We can calculate the thermal conductivity λ with either method, but the approaches are completely different. For the MD case, it is straightforward to employ the Green–Kubo relation [23],

$$\lambda = \frac{V}{k_\mathrm{B}T^2}\int\limits_0^\infty \mathrm{d}\tau \, \langle J_{q,\alpha}(\tau) J_{q,\alpha}(0)\rangle \ , \tag{1.26}$$

$$\mathbf{J}_q = \frac{1}{V}\sum_{i=1}^{N}\left[\mathbf{v}_i\left(\frac{m_i v_i^2}{2} + \frac{1}{2}\sum_{j\ (\neq i)} u_{ij}\right) - \frac{1}{2}\sum_{j\ (\neq i)} \mathbf{r}_{ij}\left(\mathbf{v}_i\cdot\mathbf{f}_{ij}\right)\right] , \tag{1.27}$$

where k_B is Boltzmann's constant, T the temperature, V the volume, $u_{ij} = u(r_{ij})$, $\mathbf{f}_i$ the force acting on atom i and $\mathbf{f}_{ij} = \mathbf{f}_j - \mathbf{f}_i$. The quantity $J_{q,\alpha}$ denotes the α component of $\mathbf{J}_q$, $\alpha \in \{x, y, z\}$, and defines the heat flux through the surface perpendicular to α averaged over the volume V containing the N particles in the summation. The notation $\sum_{j\ (\neq i)}$ refers to a single summation over j, where the term corresponding to $j = i$ is skipped. (By contrast, $\sum_{j\neq i}$ is a double summation without the term corresponding to $i = j$.) Equation (1.26) relates the thermal conductivity to the integral over a time autocorrelation function (ACF). In an equilibrium simulation of length τ_0, the time ACF of a dynamical quantity $C(t)$ can be estimated as [24]

$$\langle C(\tau)C(0)\rangle \approx \frac{1}{\tau_0}\int\limits_0^{\tau_0} \mathrm{d}t \; C(t+\tau)C(t). \tag{1.28}$$

It is important to note that with the help of Eqn (1.26), it is possible to estimate the reaction of the system to an external perturbation–in this case a temperature gradient along the coordinate α–just by analysing equilibrium properties.

An alternative approach would be to measure the thermal conductivity directly in a NEMD simulation by actually establishing a temperature gradient [23]. Let us assume that the gradient along the coordinate α arises due to an external field $\mathcal{A}$ that couples to the equations of motion in such a way that the dissipation can be written as

$$\frac{\mathrm{d}E}{\mathrm{d}t} = V J_{q,\alpha}\mathcal{A}. \tag{1.29}$$

The thermal conductivity can then be directly calculated with

$$\lambda = \frac{1}{T} \lim_{\mathcal{A}\to 0} \lim_{t\to\infty} \frac{\langle J_{q,\alpha}(t)\rangle}{\mathcal{A}}. \tag{1.30}$$

This expression should give exactly the same result as Eqn (1.26). However, the key difference really is that in the latter approach, we do actually simulate a temperature gradient, whereas the former only requires an equilibrium simulation in the absence of a gradient.

Heat conduction is one of many examples where NEMD has proved to be a powerful technique. The method was also successfully applied to studies of shear-flow in liquids [22], shock-wave propagation in solids [25] and very recently to studies of thermal polarisation effects in liquid water [1].

1.4 Outline

The remainder of this thesis is organised as follows. The governing equations of motion are discussed in Chapter 2. We first explain how the short-range interactions are modelled and then we present two summation techniques for the treatment of long-range interactions. We begin Chapter 3 with the derivation of an algorithm which is often used in NEMD simulations for imposing a stationary temperature gradient onto the system. Based on simple test cases, we identify a major drawback of this method and suggest a novel algorithm. In Chapter 4, we first depict the simulation setup and afterwards we discuss the calculation of the electrostatic field. Finally, the results are presented followed by a critical discussion. A summary of the relevant findings is given in Chapter 5, where we also suggest possible improvements for future work. Some of the derivations are discussed in more detail in the Appendices A–C.

Chapter 2

Governing equations of motion

In the previous chapter, we briefly introduced the tools and theoretical framework we will need in order to study the thermo-polarisation effect in water. This chapter deals in some detail with the two different types of interaction relevant to many common water models. Let us assume that N_m water molecules are contained in a rectangular box defined by

$$\Omega = \prod_{\alpha \in \{x,y,z\}} \left[-\frac{L_\alpha}{2}, \frac{L_\alpha}{2} \right], \tag{2.1}$$

where L_α is the box length along the coordinate α and the box of volume $V = |\Omega|$ is centred about the origin. Anticipating the detailed description of the water model in §4.1.1, let the molecules be rigid and each consist of $n = N/N_\mathrm{m} = 3$ sites. These interact with the sites of other molecules through two types of interaction. Short-range interactions only consider neighbours within a certain cutoff distance, for example a spherical region with cutoff radius r_sr. The second type is long-range electrostatic interaction, which requires special treatment, as mentioned above. The interaction potential can be split up into two contributions according to

$$U(\mathbf{r}^N) = U^\mathrm{sr}(\mathbf{r}^N) + U^\mathrm{lr}(\mathbf{r}^N), \tag{2.2}$$

where the superscripts indicate ‘short-range’ and ‘long-range’, respectively.

2.1 Short-range interaction

The interaction between uncharged atoms separated by distance r is often described by the Lennard-Jones potential,

$$u^\mathrm{lj}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]. \tag{2.3}$$

At large distances, $r \gg \sigma$, this potential models the van-der-Waals–London dispersion forces, whereas at short distances, $r < \sigma$, it models the strong repulsion due to the Pauli exclusion principle. The energy ϵ is the well depth of the potential and σ is the effective atomic diameter. The potential decays rapidly and at a distance of 2.5σ its value is only about $1/60$th of the well depth [24]. Typically the simulation box is much larger than a sphere with a similar radius. However, simply truncating the potential can have significant effects [24] even if the tail only contributes little to the potential energy. What one could do is limit the interaction to a sphere with radius $r_{\mathrm{sr}} \in [2.5\sigma, \min_\alpha(L_\alpha/2)]$ and correct for ignoring the tail contribution of the potential. An alternative approach would be to truncate and shift the potential instead, such that it is continuous. The truncated and shifted counterpart of Eqn (2.3) is given by [24]

$$\tilde{u}^{\mathrm{lj}}(r) = \begin{cases} u^{\mathrm{lj}}(r) - u^{\mathrm{lj}}(r_{\mathrm{sr}}) & \text{if } r \le r_{\mathrm{sr}}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Additional correction terms for the virial and the total energy, which arise from this modification, should be taken into account as well. For the energy conservation studies in Chapter 3, we used the truncated and shifted potential given by Eqn (2.4). For all water simulations in Chapter 4, we used the unshifted potential given by Eqn (2.3) with a spherical cutoff and standard correction terms as described in Ref. [24].

2.2 Long-range interaction

In addition to the short-range interaction we need to take into account the Coulomb interactions of all particles in the box with all other particles in all periodic images. This ambitious task can be achieved by a technique called Ewald summation [26]. Since this method is computationally expensive, several alternatives have recently been proposed [14, 27]. With regard to previous work on the thermo-polarisation effect and a comparison to literature data [3, 8], we are particularly interested in the Wolf summation method [14].

Much of the complexity of the problem lies in finding a generalised Green's function for the underlying geometry and boundary conditions. Once the Green's function G for the particular problem is known, the long-range contribution to the potential energy can be calculated according to [28]

$$\begin{aligned} U^{\mathrm{lr}}(\mathbf{r}^N) = {} & \frac{1}{2}\sum_{j \neq l}\sum_{A,B} q_{jA} q_{lB} G(\mathbf{r}_{jAlB}) + \frac{1}{2}\sum_{j}\sum_{A \neq B} q_{jA} q_{jB}\left[G(\mathbf{r}_{jAjB}) - \frac{1}{r_{jAjB}}\right] \\ & + \frac{1}{2}\sum_{j}\sum_{A} q_{jA}^2 \lim_{r \to 0}\left[G(\mathbf{r}) - \frac{1}{r}\right], \end{aligned} \tag{2.5}$$

where $\mathbf{r}_{jA}$ is the position vector of the site $A \in \{1, ..., n\}$ on molecule $j \in \{1, ..., N_\mathrm{m}\}$, q_{jA} the charge of that site and $\mathbf{r}_{jAlB} = \mathbf{r}_{lB} - \mathbf{r}_{jA}$. Therefore, we will first show how one can calculate the three-dimensional generalised Green's function for Ewald summation, G_E, starting from Poisson's equation. Subsequently, we present the corresponding Green's function for Wolf summation, G_W, and briefly comment on the differences.

2.2.1 Ewald summation

In this section, we follow Ref. [28] unless explicitly stated otherwise. Before we start with the derivation of G_E, let us first recall Poisson's law in Gaussian units,

$$\nabla^2 \Phi(\mathbf{r}) = -4\pi\rho_q(\mathbf{r}), \tag{2.6}$$

where Φ denotes the electrostatic potential and ρ_q the charge density. If we write

$$\Phi(\mathbf{r}) = \int \mathrm{d}^3 r' \frac{\rho_q(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \int \mathrm{d}^3 r' G_{\mathbb{R}^3}(\mathbf{r} - \mathbf{r}')\rho_q(\mathbf{r}'), \tag{2.7}$$

with the Green's function

$$G_{\mathbb{R}^3}(\mathbf{r}) = \frac{1}{r}, \tag{2.8}$$

then it follows from Eqn (2.6) that

$$\nabla^2 G_{\mathbb{R}^3}(\mathbf{r}) = -4\pi\delta(\mathbf{r}), \tag{2.9}$$

where $\delta(\mathbf{r})$ is the three-dimensional Dirac delta function. Note that the integration is carried out over $\mathbb{R}^3$ and thus there are no boundary effects. In the simulation, however, we use periodic boundary conditions and we need to reformulate Eqn (2.9) to account for that. In particular, we are looking for a Green's function G that satisfies [28]

$$\text{i)} \quad \nabla^2 G(\mathbf{r}) = -4\pi \left[\delta(\mathbf{r}) - \frac{1}{V}\right], \tag{2.10}$$

$$\text{ii)} \quad G(\mathbf{r}) \text{ is periodic with } \tilde{G}(\mathbf{0}) = 0, \tag{2.11}$$

where $\tilde{G}$ denotes the Fourier transform of G. The condition $\tilde{G}(\mathbf{0}) = 0$ is required for uniqueness and also leads to the additional term $4\pi/V$ on the RHS of Eqn (2.10). By Fourier transforming Eqn (2.10), one can easily verify that this condition is satisfied. Sometimes this term is also referred to as a neutralising background charge [29]. As the periodicity of G is naturally incorporated in its Fourier transform, it is advantageous to formulate the problem in k-space. The solution of Eqns (2.10)–(2.11) is then formally given by

$$G(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{k} \neq \mathbf{0}} \frac{4\pi}{k^2} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.12}$$

where the sum includes all $\mathbf{k}$-vectors compatible with Ω except for $\mathbf{k}=\mathbf{0}$. The sum in Eqn (2.12) is only conditionally convergent. We therefore introduce a convergence or damping factor η and change the order of summation by splitting the expression into

$$G(\mathbf{r}) = G_{\mathrm{I}}(\mathbf{r}) + G_{\mathrm{II}}(\mathbf{r}), \tag{2.13}$$

$$G_{\mathrm{I}}(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi}{k^2}\left[1-\mathrm{e}^{-\frac{k^2}{4\eta^2}}\right]\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.14}$$

$$G_{\mathrm{II}}(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi}{k^2}\mathrm{e}^{-\frac{k^2}{4\eta^2}}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}. \tag{2.15}$$

An important role of η is to weight the contributions of G_{I} and G_{II}. The expression of G_{I} can be further simplified by noticing that it only differs from the Fourier transform of $\operatorname{erfc}(\eta\mathbf{r})/r$ over $\mathbb{R}^3$ for $\mathbf{k}=\mathbf{0}$. Correcting for the missing term $\mathbf{k}=\mathbf{0}$ by $-\pi(\eta^2 V)^{-1}$, one can show that the Fourier transform of

$$G(\mathbf{r}) = \sum_{\mathbf{n}} \frac{\operatorname{erfc}(\eta|\mathbf{r}+\mathbf{n}L|)}{|\mathbf{r}+\mathbf{n}L|} - \frac{\pi}{\eta^2 V} + \frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi}{k^2}\mathrm{e}^{-\frac{k^2}{4\eta^2}}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.16}$$

is indeed given by Eqn (2.13), where $\mathbf{n}$ is a shift vector between a molecule and its periodic image and the summation runs over all periodic images [28]. Choosing η carefully, one can achieve that the first term converges fast and the contributions for $\mathbf{n}\neq\mathbf{0}$ are small. If we ignore these terms, Eqn (2.16) finally reduces to

$$G_{\mathrm{E}}(\mathbf{r}) = \frac{\operatorname{erfc}(\eta r)}{r} - \frac{\pi}{\eta^2 V} + \frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi}{k^2}\mathrm{e}^{-\frac{k^2}{4\eta^2}}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}. \tag{2.17}$$

As demonstrated in Appendix A, inserting this expression back into Eqn (2.5) yields the standard Ewald summation expression as presented in textbooks, e.g. in Ref. [24]. For performance reasons, one chooses a real space cutoff, r_{c}, for the first term in Eqn (2.17) and a k-space cutoff, k_{c}, for the last term. Nevertheless, the evaluation of the energy still remains expensive. With an optimal choice of η, r_{c} and k_{c} a scaling of $\mathcal{O}(N^{3/2})$ can be achieved [24]. However, there are extensions to Ewald summation, such as Particle Mesh Ewald, that evaluate the k-space contribution efficiently using a fast Fourier transformation and scale as $\mathcal{O}(N\log N)$ [30].

2.2.2 Wolf summation

Recently, Wolf and co-workers showed that in a condensed ionic system the net Coulomb potential is effectively short-ranged [14]. Based on this insight, they devised a novel summation method that avoids the expensive k-space term in Eqn (2.16). Instead, the potential is shifted in such a way that charge neutrality is enforced [27]. Comparison with other methods reveals that this technique performs well [8, 27]. For a full derivation and detailed discussion of the Wolf method, we refer to Ref. [14]. Here, we will only present the Green's function which, when inserted into Eqn (2.5), gives the final result for the energy of a charge neutral system (Eqn (5.13) in Ref. [14]). It can be easily verified (see Appendix A) that the corresponding Green's function is given by [28]

$$G_{\mathrm{W}}(\mathbf{r}) = \begin{cases} \frac{\operatorname{erfc}(\zeta r)}{r} - \frac{\operatorname{erfc}(\zeta r_{\mathrm{c}})}{r_{\mathrm{c}}} & \text{if } r \leq r_{\mathrm{c}}, \\ 0 & \text{otherwise}, \end{cases} \tag{2.18}$$

where r_{c} is again a cutoff radius and ζ a damping parameter. It is obvious from the functional form of G_{W} that the function is continuous at the cutoff distance. We can also note that the first term in G_{W} is identical to the one in G_{E}, although the optimal choice of the damping parameter is not necessarily the same. In sharp contrast to standard Ewald summation, the Wolf method can achieve linear, i.e. $\mathcal{O}(N)$, scaling [27].

2.3 Numerical integration

The dynamics of the system are governed by Eqn (1.24), which we can discretise to evolve all coordinates in time by a fixed timestep Δt. The velocity Verlet algorithm, given by

$$\mathbf{v}_i^{n+\frac{1}{2}} = \mathbf{v}_i^n + \frac{\Delta t}{2m_i}\mathbf{f}_i^n, \tag{2.19}$$

$$\mathbf{r}_i^{n+1} = \mathbf{r}_i^n + \Delta t \mathbf{v}_i^{n+\frac{1}{2}}, \tag{2.20}$$

$$\mathbf{v}_i^{n+1} = \mathbf{v}_i^{n+\frac{1}{2}} + \frac{\Delta t}{2m_i}\mathbf{f}_i^{n+1}, \tag{2.21}$$

is a simple but good choice for solving the equations of motion, because it exhibits only little long-term energy drift [24]. Carrying out the update procedure (Eqn (2.19)–(2.21)) evolves the coordinates and velocities, $\mathbf{r}_i^n$ and $\mathbf{v}_i^n$, from time $t^n = n\Delta t$ to a time $t^{n+1} = t^n + \Delta t$. We will use this notation throughout the report and sometimes refer to the time t^n simply as time n. Since a typical trajectory of interest contains more than 10^6 steps, long-term energy conservation is a desirable feature. The same update procedure is applied to each site on any molecule individually. Therefore, there is no connectivity between the sites of a rigid molecule unless we explicitly impose one. If we fail to do so, the molecule would be strongly distorted even after only a few timesteps. A very elegant way of freezing certain degrees of freedom is using constraining forces with Lagrange multipliers. Just focussing on a single water molecule and employing this method, one can reformulate the equations of motion as [31]

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i - \sum_\alpha \lambda_\alpha \frac{\partial \sigma_\alpha}{\partial \mathbf{r}_i} = \mathbf{f}_i + \mathbf{g}_i, \tag{2.22}$$

where $i \in \{1, 2, 3\}$ is the index running over the sites of the reference molecule. The additional force term $\mathbf{g}_i$ consists of Lagrange multipliers, λ_α, and holonomic constraints of the form

$$\sigma_\alpha(\mathbf{r}_j) \equiv 0 \quad \forall \quad \alpha \in \{12, 13, 23\}, \tag{2.23}$$

$$\sigma_{12} = \frac{1}{2}\left[(\mathbf{r}_2 - \mathbf{r}_1)^2 - d_{12}^2)\right], \tag{2.24}$$

$$\sigma_{23} = \frac{1}{2}\left[(\mathbf{r}_3 - \mathbf{r}_2)^2 - d_{23}^2)\right], \tag{2.25}$$

$$\sigma_{13} = \frac{1}{2}\left[(\mathbf{r}_3 - \mathbf{r}_1)^2 - d_{13}^2)\right], \tag{2.26}$$

where d_α is the oxygen-hydrogen or hydrogen-hydrogen distance we want to impose. Next, we need to determine the Lagrange multipliers in order to solve Eqn (2.22). Let us assume that the distance between sites 1 and 2 should be correct after updating the coordinates with the velocity Verlet algorithm at time $n+1$. Using Eqns (2.20), (2.22) and (2.24), this condition can be formulated as

$$\begin{aligned}\left(r_{12}^2\right)^{n+1} = \Bigg[\bar{\mathbf{r}}_{12}^{n+1} &- \frac{\Delta t^2}{2}\left(\frac{1}{m_2} + \frac{1}{m_1}\right)\lambda_{12}\mathbf{r}_{12}^n \\ &+ \frac{\Delta t^2}{2m_2}\lambda_{23}\mathbf{r}_{23}^n - \frac{\Delta t^2}{2m_1}\lambda_{13}\mathbf{r}_{13}^n\Bigg]^2 \overset{!}{=} d_{12}^2,\end{aligned} \tag{2.27}$$

with

$$\bar{\mathbf{r}}_i^{n+1} = \mathbf{r}_i^n + \Delta t \mathbf{v}_i^n + \frac{\Delta t^2}{2m_i}\mathbf{f}_i^n, \tag{2.28}$$

$$\mathbf{r}_i^{n+1} = \bar{\mathbf{r}}_i^{n+1} - \frac{\Delta t^2}{2m_i}\sum_\alpha \lambda_\alpha \frac{\partial \sigma_\alpha}{\partial \mathbf{r}_i}, \tag{2.29}$$

where $\bar{\mathbf{r}}_i^{n+1}$ is the updated (uncorrected) position the site i would take if it was simply integrated with Eqn (2.20). Equation (2.27) imposes one condition on the three Lagrange multipliers and proceeding analogously for d_{23} and d_{13} leads to a quadratic system of equations in λ_α, which can be solved with an arbitrary method of choice, e.g. iteratively.

This procedure describes how rigidity is imposed in LAMMPS, the parallel simulation package we used for all our simulations [32]. Although there are similarities, we note that this is neither the original SHAKE algorithm [33] for Verlet integration nor the RATTLE algorithm [31] for velocity Verlet integration. We discuss some issues involved with the LAMMPS implementation in §4.3.5.

Chapter 3

Thermal gradient

In an equilibrium simulation, the average temperature distribution should be homogeneous. However, we are interested in the physical behaviour of a system that remains in a non-equilibrium stationary state. NEMD is a useful tool for studying the underlying physics. With this method, a temperature gradient can be imposed, which is independent of time after an initial transient period, i.e. $\frac{\partial T(t,z)}{\partial t} = 0$ for $t \geq t_0$. In the past, many algorithms have been developed for this purpose [15, 34–37]. In previous investigations on the thermo-polarisation effect, the algorithm of Ikeshoji and Hafskjold [15] proved to be particularly useful [1]. We explain this algorithm in detail in §3.1.

Furthermore, we want to clearly distinguish between the method employed to create the gradient itself and the combination of this method with the underlying integration algorithm. Ikeshoji and Hafskjold originally used the Leap-Frog algorithm for integrating the equations of motion in Ref. [15]. In Ref. [16], the method of Ikeshoji and Hafskjold was called the heat exchange (HEX) algorithm; however, the original algorithm was combined with velocity Verlet time integration. In this report, we will call the method for establishing the gradient the Ikeshoji and Hafskjold algorithm. We understand the HEX algorithm to be the combination of velocity Verlet integration and the Ikeshoji and Hafskjold algorithm.

In this chapter, we first introduce the HEX algorithm in §3.1.1. Then, it is demonstrated that the algorithm does not conserve the total energy in §3.1.2. We analyse the integration scheme in §3.2.1 by considering the limit $\Delta t \to 0$, which is called *modified equation analysis*. Based on that analysis, we identify the major cause of the energy conservation problem and propose an improved algorithm in §3.2.2. Finally, we demonstrate that the new algorithm conserves the energy very well for simulation times of interest in §3.2.3.

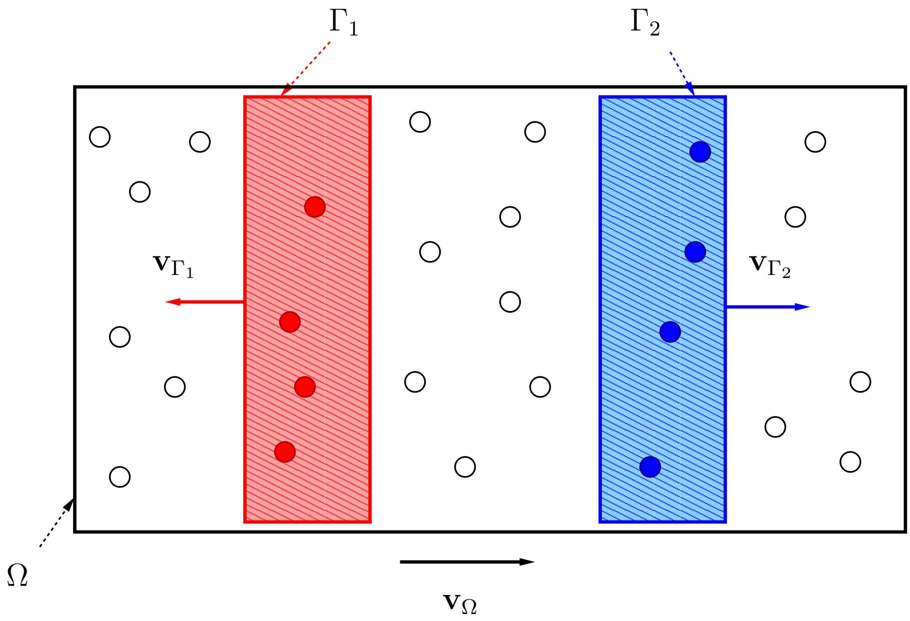

Figure 3.1: Illustration of the simulation box, Ω, with a red shaded hot region, Γ_1, and a blue shaded cold region, Γ_2. The centre of mass velocities are $\mathbf{v}_\Omega$, $\mathbf{v}_{\Gamma_1}$ and $\mathbf{v}_{\Gamma_2}$, respectively. Atoms are represented by red/blue circles, if they are contained in the hot/cold region and by empty circles otherwise.

3.1 Heat exchange algorithm

The original formulation of the Ikeshoji and Hafskjold algorithm has been simplified and presented in a cleaner way [38]. In what follows, the reformulated version is considered, but using a different notation. The algorithm has previously been integrated into LAMMPS exactly as presented in Ref. [38]. Throughout the report we will use two different versions of LAMMPS: the original version 1Feb14 and a slightly modified version 1Feb14/mod, which contains the new algorithm proposed in §3.2.2.

3.1.1 Formulation

Suppose we want to add a heat amount ΔQ to a region $\Gamma \subset \Omega$. For example, this could be either the region Γ_1 or Γ_2 in Fig. 3.1. We also assume that the algorithm is to be applied sequentially to different, disjoint regions. Let us define the index set $\gamma = \{j \mid \mathbf{r}_j \in \Gamma\}$, where $\mathbf{r}_i$ denotes the position vector of atom i. Furthermore, let $\mathbf{v}_i$ be the velocity of that atom and $\mathbf{v}_\Gamma$ and $\mathbf{v}_\Omega$ the centre of mass velocities of the region Γ and the box Ω, respectively.

The basic idea of the algorithm is to scale all the velocities in Γ by the same factor η and shift them by a constant, such that $\mathbf{v}_\Gamma$ is unaffected and the non-translational kinetic energy of region Γ changes by ΔQ. To formulate this idea in an equation, we first define what we mean by non-translational kinetic energy. The total kinetic energy contained in the region Γ, $\mathcal{K}_\Gamma$, is given by

$$\mathcal{K}_\Gamma = \sum_{i\in\gamma} \frac{m_i v_i^2}{2}, \tag{3.1}$$

and it can be further split into translational and non-translational contributions,

$$\mathcal{K}_\Gamma = \mathcal{K}_\Gamma^{\mathrm{t}} + \mathcal{K}_\Gamma^{\mathrm{nt}}. \tag{3.2}$$

The translational kinetic energy, $\mathcal{K}_\Gamma^{\mathrm{t}}$, is calculated as

$$\mathcal{K}_\Gamma^{\mathrm{t}} = \frac{m_\Gamma v_\Gamma^2}{2}, \quad \text{where} \tag{3.3}$$

$$\mathbf{v}_\Gamma = \frac{1}{m_\Gamma} \sum_{i\in\gamma} m_i \mathbf{v}_i, \tag{3.4}$$

and where m_Γ is the total mass contained in region Γ. The non-translational kinetic energy, $\mathcal{K}_\Gamma^{\mathrm{nt}}$, is then simply defined as

$$\mathcal{K}_\Gamma^{\mathrm{nt}} = \mathcal{K}_\Gamma - \mathcal{K}_\Gamma^{\mathrm{t}}. \tag{3.5}$$

If an updated quantity has an overbar in addition to the standard symbol, the application of the algorithm should satisfy

$$\bar{\mathcal{K}}_\Gamma^{\mathrm{nt}} = \mathcal{K}_\Gamma^{\mathrm{nt}} + \Delta Q = \eta\, \mathcal{K}_\Gamma^{\mathrm{nt}}, \tag{3.6}$$

which implicitly defines the scaling factor η. Solving Eqn (3.6) for η gives

$$\eta = \frac{\mathcal{K}_\Gamma^{\mathrm{nt}} + \Delta Q}{\mathcal{K}_\Gamma^{\mathrm{nt}}}. \tag{3.7}$$

With the definition $\xi = \sqrt{\eta}$ it can easily be shown (see Appendix B) that the update step

$$\mathbf{v}_i \mapsto \bar{\mathbf{v}}_i = \xi \mathbf{v}_i + (1-\xi)\mathbf{v}_\Gamma \quad \forall i \in \gamma, \tag{3.8}$$

satisfies Eqn (3.6) and conserves the centre of mass velocity of Γ, namely

$$\bar{\mathbf{v}}_\Gamma = \mathbf{v}_\Gamma. \tag{3.9}$$

If this procedure is applied to regions Γ_1 and Γ_2 in Fig. 3.1 such that ΔQ is taken from Γ_2 and added to Γ_1, the total energy in the system as well as $\mathbf{v}_\Omega$ should be unaffected. The velocity Verlet integration scheme together with Eqn (3.8) define the HEX algorithm. The full set of equations is summarised in §3.2.1.

3.1.2 Energy conservation

Intrigued by the observation of a slight energy drift in a simulation of rigid water, we decided to investigate this problem further using a much simpler system and this algorithm. To illustrate the energy conservation problem, we performed NEMD simulations of various different temperature gradients along the z-axis in a Lennard-Jones system comprising 1728 particles. Reduced units (see Appendix B) apply to all results presented in this chapter. The individual quantities should therefore be labelled separately, e.g. with an asterisk (*). However, we drop this label for readability. The rectangular box is defined by the dimensions $L_x = L_y = L_z/2$. The full input script and input parameters for a typical simulation are given in the Appendix B.

We performed two sets of simulations for a similar but not identical setup as depicted in Fig. 3.1. In particular, the centres of the two reservoirs are separated by a distance $L_z/2$ and the hot reservoir is aligned with $z = 0$. The same binary restart file was used for all gradient simulations in order to guarantee identical initial conditions. First, we simulated the same kinetic energy gradient using various timesteps Δt (see Fig. 3.2). We then simulated different gradients for a fixed timestep (see Fig. 3.3). This can be achieved by adjusting the heat production rate,

$$\kappa = \frac{\Delta Q}{\Delta t} \geq 0, \tag{3.10}$$

which defines the heat $\Delta Q = \kappa \Delta t$ being added to the hot reservoir and ΔQ being removed from the cold reservoir during every timestep. Periodicity and the fact that the two regions are separated exactly by $L_z/2$ require half the heat supplied to the hot reservoir to dissipate to each neighbouring cold reservoir (periodic boundary conditions). The resulting heat fluxes are therefore given by

$$\mathbf{J}_q = \pm \frac{\kappa}{2 L_x L_y} \mathbf{e}_z, \tag{3.11}$$

where $\mathbf{e}_z$ is the unit vector in z-direction. We note that the energies were well conserved in an *NVE* simulation at the highest average kinetic energy of 4.5 (see Fig. 3.2) using the largest timestep of $\Delta t = 0.0012$ (black symbols in Fig. 3.4). This suggests that the loss in energy cannot simply be attributed to the choice of an unstable timestep.

In Fig. 3.2, we can see that the gradient remains unaffected by the choice of timestep for the timestep range considered. All simulations exhibit a fairly linear kinetic energy profile between the reservoirs. The maximum and minimum kinetic energies are obtained at the centres of the reservoirs, respectively. On the other hand, changing κ for a fixed timestep leads to different kinetic energy gradients, as demonstrated in Fig. 3.3. However, most importantly, in both cases the total energy per

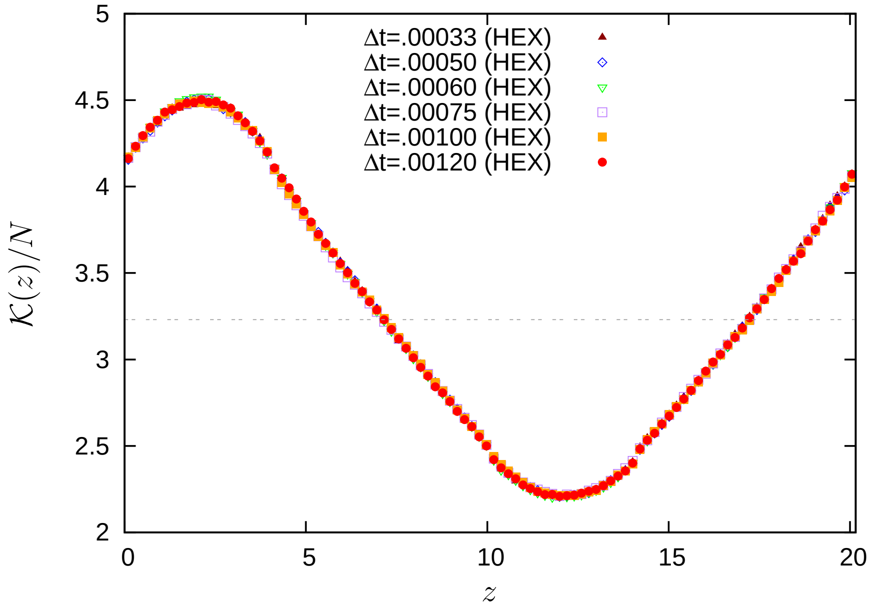

Figure 3.2: Kinetic energy profile along the z-axis for a fixed value of $\kappa = 300$ at varying timesteps Δt. The results were obtained using the original HEX algorithm as implemented in LAMMPS (1Feb14). The dashed line represents the average kinetic energy per particle.

particle decreases as a function of time, as illustrated in Figs 3.4–3.5, respectively. Plotting the loss of energy per particle at the final time $t = 1000$ as a function of the timestep (Fig. 3.6) and κ (Fig. 3.7), respectively, one can clearly see a quadratic dependence on both quantities. This observation is crucial for identifying the cause of the observed energy drift.

3.2 Conservative heat exchange algorithm

From the results presented in the previous section, we know that the energy loss scales with Δt^2. We are therefore particularly interested in identifying which terms in the integration scheme scale in this way. Moreover, it is important for understanding the HEX algorithm to know which modified equation is actually being integrated.

3.2.1 Modified equation analysis

The standard integration scheme in LAMMPS is the velocity Verlet algorithm. Evolving the momenta and coordinates in time according to Eqns (2.19)–(2.21) satisfies Newton's law, i.e. $\ddot{\mathbf{r}}_i(t) = \dot{\mathbf{v}}_i(t) = \mathbf{f}_i(t)/m_i$, up to terms of $\mathcal{O}(\Delta t^2)$. The HEX algorithm requires an additional velocity update step and the integration

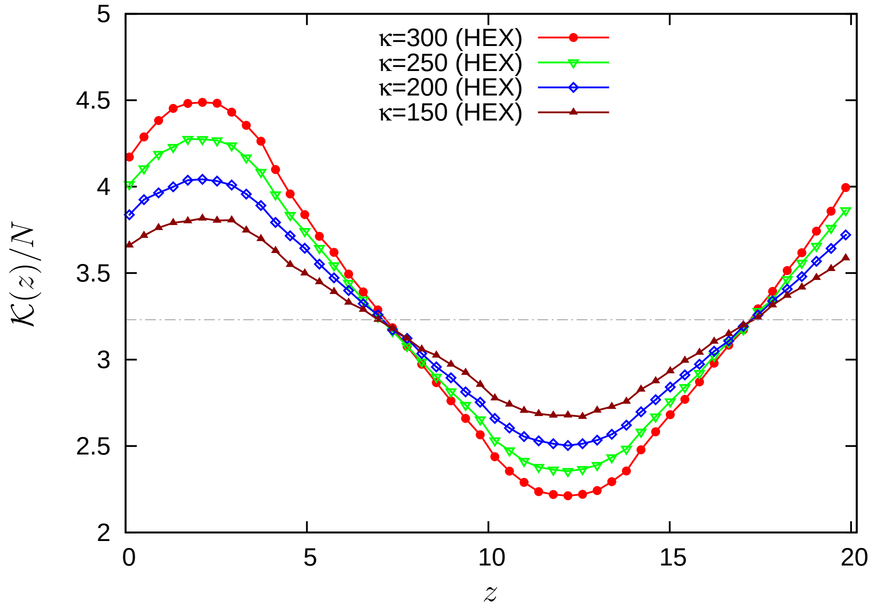

Figure 3.3: Kinetic energy profile along the z-axis for various values of κ at a fixed timestep $\Delta t = 0.001$. The results were obtained using the original HEX algorithm as implemented in LAMMPS (1Feb14). The dashed line represents the average kinetic energy per particle.

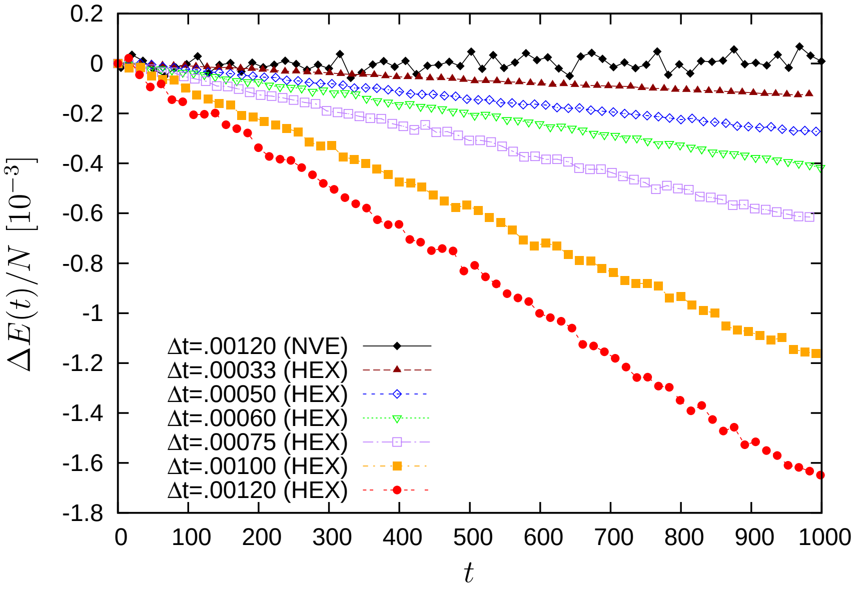

Figure 3.4: Energy loss per particle as a function of time for various timesteps and a fixed $\kappa = 300$. The results were obtained using the original HEX algorithm as implemented in LAMMPS (1Feb14). For comparison, we also added the results of an NVE simulation at the highest kinetic energy, $\mathcal{K}(z_{\max})/N = 4.5$, and the largest timestep $\Delta t = 0.0012$.

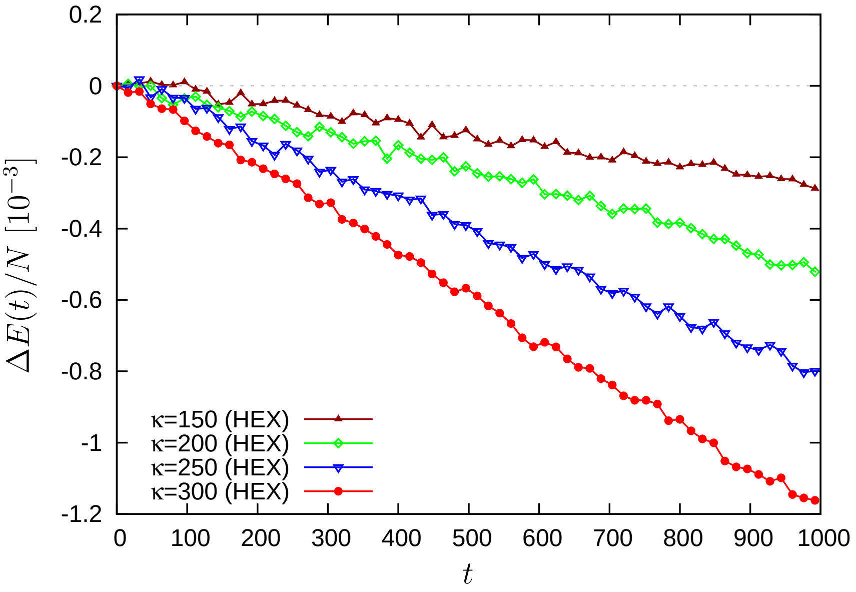

Figure 3.5: Energy loss per particle as a function of time for various values of κ and a fixed timestep $\Delta t = 0.001$. The results were obtained using the original HEX algorithm as implemented in LAMMPS (1Feb14). The dashed horizontal line is added for comparison.

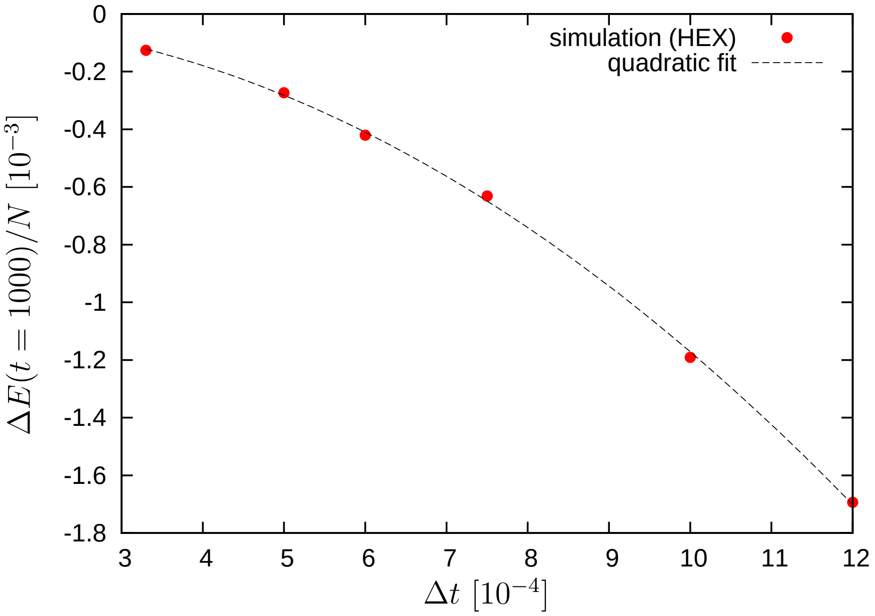

Figure 3.6: Energy loss per particle at time $t = 1000$ as a function of Δt for a fixed $\kappa = 300$. The results were obtained using the original HEX algorithm as implemented in LAMMPS (1Feb14).

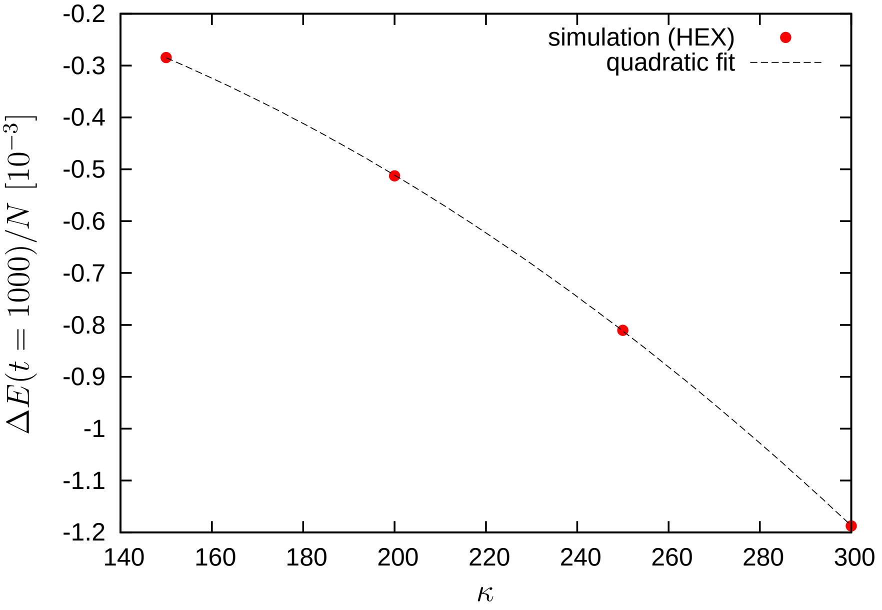

Figure 3.7: Energy loss per particle at time $t = 1000$ as a function of κ for a fixed $\Delta t = 0.001$. The results were obtained using the original HEX algorithm as implemented in LAMMPS (1Feb14).

scheme may be written as

$$\begin{aligned} \mathbf{v}_i^{n+\frac{1}{2}} &= \mathbf{v}_i^n + \frac{\Delta t}{2m_i}\mathbf{f}_i^n, && (3.12)\\ \mathbf{r}_i^{n+1} &= \mathbf{r}_i^n + \Delta t \mathbf{v}_i^{n+\frac{1}{2}}, && (3.13)\\ \mathbf{f}_i^{n+1} &= -\nabla_{\mathbf{r}_i} U(\mathbf{r}_j^{n+1}) + \mathbf{g}_i(\mathbf{r}_j^{n+1}), && (3.14)\\ \mathbf{v}_i^{n+1} &= \mathbf{v}_i^{n+\frac{1}{2}} + \frac{\Delta t}{2m_i}\mathbf{f}_i^{n+1}, && (3.15)\\ \bar{\mathbf{v}}_i^{n+1} &= \xi^{n+1}\left(\mathbf{v}_i^n + \frac{\Delta t}{2m_i}\left[\mathbf{f}_i^n + \mathbf{f}_i^{n+1}\right]\right) && (3.16)\\ &\quad + \left[1-\xi^{n+1}\right]\frac{1}{m_\Gamma}\sum_{j\in\gamma} m_j\left(\mathbf{v}_j^n + \frac{\Delta t}{2m_j}\left[\mathbf{f}_j^n + \mathbf{f}_j^{n+1}\right]\right), \end{aligned}$$

where $\mathbf{g}_i$ is a constraining force for rigid molecules and $\mathbf{g}_i \equiv 0$ applies to atomistic systems. The forces need only be calculated once per timestep and the same scaling factor ξ^{n+1} applies to all velocities in Γ. Next, we consider the discrete velocities $\bar{\mathbf{v}}_i^n$ and $\bar{\mathbf{v}}_i^{n+1}$ in the limit $\Delta t \to 0$. It is straightforward to show that the velocity satisfies the modified equation

$$\begin{aligned} \bar{\mathbf{v}}_i(t) &= \xi(t)\mathbf{v}_i(t) + [1-\xi(t)]\,\mathbf{v}_\Gamma(t), && (3.17)\\ \dot{\mathbf{v}}_i(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\xi(t)\mathbf{v}_i(t) + (1-\xi)\mathbf{v}_\Gamma(t)\right] =: \frac{\bar{\mathbf{f}}_i(t)}{m_i}, && (3.18) \end{aligned}$$

which defines the modified force $\bar{\mathbf{f}}_i(t)$. Strictly speaking, ξ is also a piece-wise constant function of $\mathbf{r}_j$, but we ignore the discontinuities for simplicity. The modified force satisfies $\lim_{\xi\to 1}\bar{\mathbf{f}}_i(t) = \mathbf{f}_i(t)$, as one would expect since applying Eqn (3.17) with $\xi = 1$ is the identity operation. From Eqns (3.12)–(3.16), it can easily be seen that the coordinates $\mathbf{r}_i^{n+1}$ are not affected by the rescaling step, given by Eqn (3.16), at time n. With other words, the coordinates do not feel the acceleration

$$\frac{1}{m_i}\left[\bar{\mathbf{f}}_i - \mathbf{f}_i\right](t) = \frac{\mathrm{d}}{\mathrm{d}t}\left[\xi\mathbf{v}_i + (1-\xi)\mathbf{v}_\Gamma\right](t) - \dot{\mathbf{v}}_i(t) \tag{3.19}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t}\left[(\xi - 1)(\mathbf{v}_i - \mathbf{v}_\Gamma)\right](t), \tag{3.20}$$

acting on the particles in the time interval $[t^n, t^{n+1}]$ in order for the velocities to satisfy Eqn (3.16).

3.2.2 Formulation

Having noticed that the coordinates are ‘lagging behind’, the fix is actually trivial. Consider the Taylor expansion

$$\frac{\mathbf{r}_i(t+1) - \mathbf{r}_i(t)}{\Delta t} = \mathbf{v}_i(t) + \frac{\Delta t}{2m_i}\mathbf{f}_i(t) + \frac{\Delta t^2}{6m_i}\dot{\mathbf{f}}_i(t) + \mathcal{O}(\Delta t^3). \tag{3.21}$$

The update step given by Eqn (3.13) already takes into account the first two terms in Eqn (3.21). However, the change of the force within the time interval $[t^n, t^{n+1}]$ caused by scaling and shifting the velocities requires a consideration of the third term as well. We can approximate the time integral of that term by

$$\frac{\Delta t^2}{6m_i}\int_t^{t+\Delta t} \mathrm{d}t\ \dot{\mathbf{f}}_i(t) \approx \frac{\Delta t^2}{6}\frac{\mathrm{d}}{\mathrm{d}t}\Big((\xi - 1)(\mathbf{v}_i - \mathbf{v}_\Gamma)\Big)(t), \tag{3.22}$$

and subsequently correct the coordinates in an additional update step. This update step will move the particles in such a way that they are up-to-date with the velocities at the end of the timestep, whilst leaving the velocities unaffected. Based on this insight, we propose the **cHEX** (conservative HEX) **algorithm** given by Eqns (3.23)–(3.30).

$$\mathbf{v}_i^{n+\frac{1}{2}} = \bar{\mathbf{v}}_i^n + \frac{\Delta t}{2m_i}\bar{\mathbf{f}}_i^n, \tag{3.23}$$

$$\mathbf{r}_i^{n+1} = \bar{\mathbf{r}}_i^n + \Delta t \mathbf{v}_i^{n+\frac{1}{2}}, \tag{3.24}$$

$$\mathbf{f}_i^{n+1} = -\nabla_{\mathbf{r}_i}\ U(\mathbf{r}_j^{n+1}) + \mathbf{g}_i(\mathbf{r}_j^{n+1}), \tag{3.25}$$

$$\mathbf{v}_i^{n+1} = \mathbf{v}_i^{n+\frac{1}{2}} + \frac{\Delta t}{2m_i}\mathbf{f}_i^{n+1}, \tag{3.26}$$

$$\xi_i^{n+1} = \begin{cases} \left[\frac{\left(\mathcal{K}_\Gamma^{\mathrm{nt}}\right)^n + \Delta Q}{\left(\mathcal{K}_\Gamma^{\mathrm{nt}}\right)^n}\right]^{\frac{1}{2}} & \text{if } i \in \gamma^{n+1}, \\ 1 & \text{otherwise,} \end{cases} \tag{3.27}$$

$$\bar{\mathbf{v}}_i^{n+1} = \begin{cases} \xi_i^{n+1}\mathbf{v}_i^{n+1} + \left[1 - \xi_i^{n+1}\right]\mathbf{v}_\Gamma^{n+1} & \text{if } i \in \gamma^{n+1}, \\ \mathbf{v}_i^{n+1} & \text{otherwise,} \end{cases} \tag{3.28}$$

$$\bar{\mathbf{r}}_i^{n+1} = \begin{cases} \mathbf{r}_i^{n+1} + \frac{\Delta t^2}{6}\left[(\xi_i^{n+1} - 1)\left(\frac{\mathbf{f}_i^{n+1}}{m_i} - \frac{\mathbf{v}_\Gamma^{n+1} - \mathbf{v}_\Gamma^n}{\Delta t}\right)\right. & \\ \left. + \frac{\xi_i^{n+1} - \xi_i^n}{\Delta t}\left(\mathbf{v}_i^{n+1} - \mathbf{v}_\Gamma^{n+1}\right)\right] & \text{if } i \in \gamma^{n+1}, \\ \mathbf{r}_i^{n+1} & \text{otherwise,} \end{cases} \tag{3.29}$$

$$\bar{\mathbf{f}}_i^{n+1} = \begin{cases} \mathbf{f}_i^{n+1} + \mathbf{g}_i(\bar{\mathbf{r}}_j^{n+1}) & \text{if } i \in \gamma^{n+1}, \\ \mathbf{f}_i^{n+1} & \text{otherwise.} \end{cases} \tag{3.30}$$

There are a few differences compared to the HEX algorithm, the most important of which is the additional coordinate integration step (Eqn (3.29)). For rigid molecules, this step requires calculating additional constraining forces, as stated in Eqn (3.30). There is a very important detail we would like to point out: the coordinate integration (Eqn (3.29)) is carried out with the constraining forces from Eqn (3.25). Therefore, this update step will not preserve the geometry exactly. However, the constraints will be satisfied again after the subsequent integration step (Eqn (3.24)). Therefore, it is important that any sampling occurs after the coordinate integration given by Eqn (3.24). However, even after the second coordinate integration (Eqn (3.29)), the geometry of water, for example, remains preserved up to a precision of 10^{-4} for a timestep of 1 fs. Depending on the desired tolerance, it might not even be necessary to update the constraining force at all (Eqn (3.30)).

What is more, each atom i now has an individual ξ_i assigned to it, because this is required in order to approximate the time derivative of ξ in the coordinate integration (Eqn (3.29)). With regard to parallel efficiency, this is a slight drawback, as it requires additional communication every time the mapping of atoms to processors changes. However, this does not necessarily happen in every timestep. Therefore, it adds only little overhead to the basic communication that is necessary in any case in order to calculate $\mathbf{v}_\Gamma$, m_Γ and $\mathcal{K}_\Gamma^{\mathrm{nt}}$.

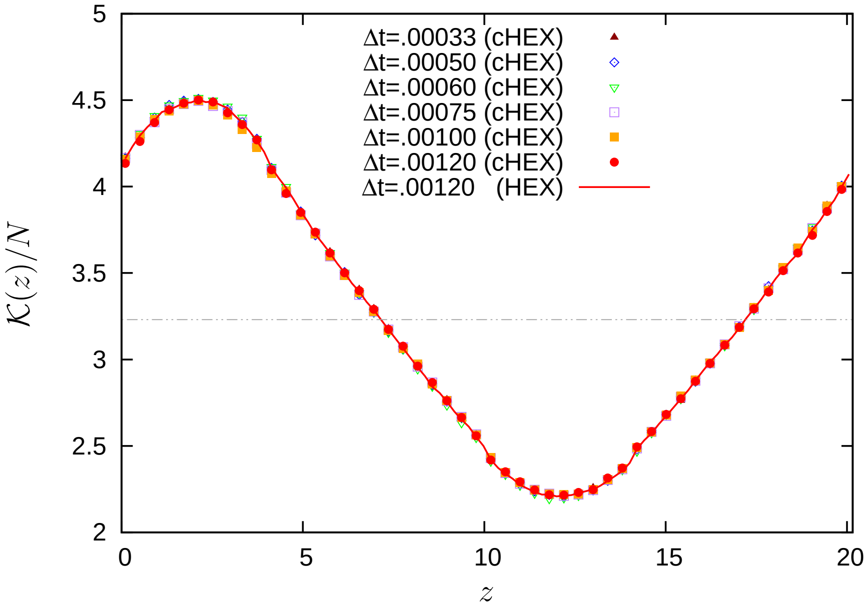

Figure 3.8: Kinetic energy profile along the z-axis for a fixed value of $\kappa = 300$ at varying timesteps Δt. The results represented by symbols were obtained using the cHEX algorithm in LAMMPS (1Feb14/mod). The $\Delta t = 0.0012$ result for the HEX algorithm (red solid line) is also shown for comparison.

It should also be mentioned that although the velocities and coordinates are integrated in every timestep, LAMMPS does allow heat to be added at an arbitrary frequency. With other words, the integration steps given by Eqns (3.27)–(3.30) do not necessarily have to be carried out in every timestep. Reformulating the equations to take this into account is straightforward. Finally, we note that for particles outside the region Γ, the update steps given by Eqns (3.27)–(3.30) are simply identity operations and the combined scheme reduces to velocity Verlet.

3.2.3 Energy conservation

In order to analyse the energy conservation properties of the cHEX algorithm, we reran all previous LJ simulations with LAMMPS (1Feb14/mod), which uses the new cHEX algorithm instead of the built-in HEX algorithm. All HEX and cHEX gradient simulations were started from the same binary restart file. The spatial kinetic energy profiles are presented in Fig. 3.8 and Fig. 3.9, respectively. As can be seen, there is no substantial difference and the new curves are overlapping with the old ones. When it comes to the error in energy, however, the improvement becomes immediately apparent. As compared to the original algorithm, there is no noticeable energy drift any more after 10^6 timesteps. This clearly demonstrates that we were able to identify and resolve the underlying problem.

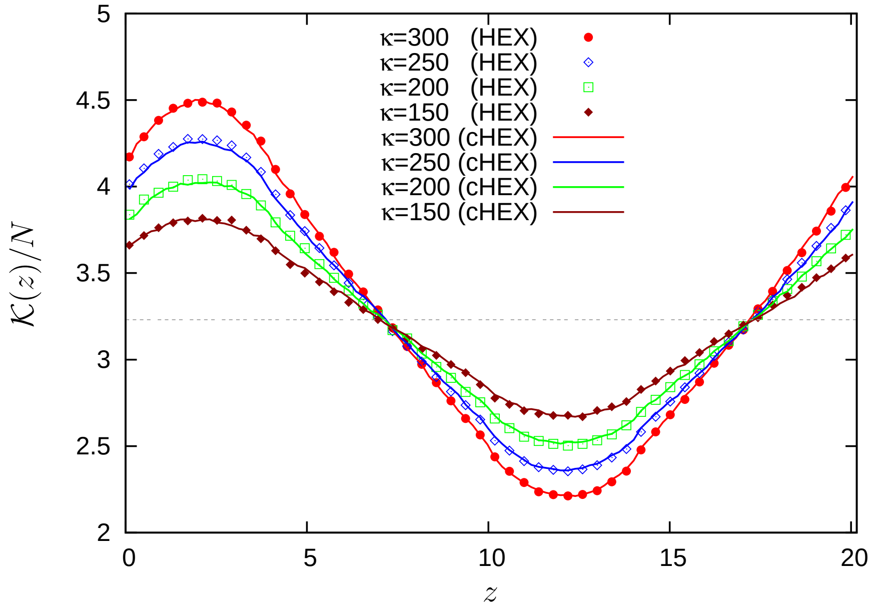

Figure 3.9: Kinetic energy profile along the z-axis for various values of κ at a fixed timestep $\Delta t = 0.001$. The solid curves represent the results for the cHEX algorithm (LAMMPS (1Feb14/mod)) and the symbols the old results for the HEX algorithm.

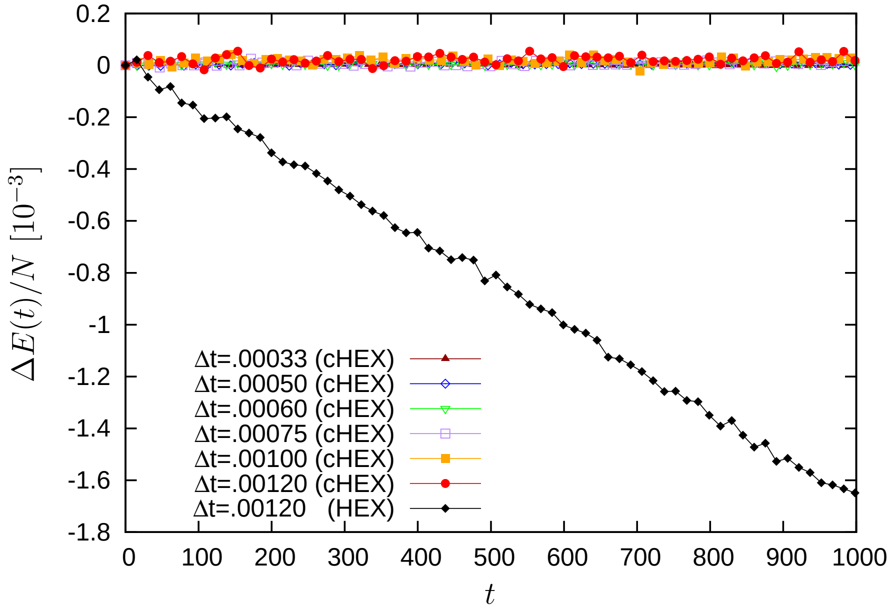

Figure 3.10: Energy loss per particle as a function of time for various timesteps and a fixed $\kappa = 300$. The results were obtained using the cHEX algorithm (LAMMPS (1Feb14/mod)). For comparison, we also show the result for $\kappa = 300$ and $\Delta t = 0.0012$ obtained with the HEX algorithm.

Chapter 4

Thermal polarisation in water

In the preceding chapters, we introduced the tools and algorithms which can be utilised to study the thermo-polarisation effect in water. Earlier, Bresme and co-workers expended considerable effort investigating the electrostatic field which arises due to molecular reorientation [1, 3, 6, 8]. Wolf summation was used for the treatment of electrostatic interactions in their work. Bresme and co-workers argue that Ewald summation leads to artificial effects which are caused by the interaction of the dipoles in the simulation box and their replicas [1]. However, it had previously been shown that several simulation quantities obtained with the Wolf method compare well to their Ewald summation counterparts [8, 14], and we therefore find it surprising that the electrostatic field should be affected by the underlying summation technique to such an extent. To the best of our knowledge, the thermo-polarisation effect has not been reported in simulations with Ewald summation yet. Since Ewald summation is a widely used technique, we therefore decided to follow the simulation protocol in Ref. [8] and carry out a comparison of the electrostatic field for both summation techniques.

We begin this central chapter of the current investigation by explaining in detail the simulation setup and the simulation protocol in §4.1–4.2. Subsequently, some results are compared to the ones obtained with an in-house code and to the literature data in §4.3, which confirms that our LAMMPS input scripts are correct. Based on the theoretical work of Neumann [28], we then discuss the calculation of the electrostatic field for a system with periodic boundary conditions in §4.4. All simulation results are presented in §4.5 and our main findings are discussed in §4.6.

m_O	m_H	d_OH	θ	q_H	q_O	ϵ	σ
15.9994	1.00794	1.0	109.47	0.4238	−0.8476	0.1553	3.166

Table 4.1: Parameters for the SPC/E model. The following units apply: $[m_\mathrm{x}] = \mathrm{u}$, $[d_\mathrm{x}] = [\sigma] = \text{Å}$, $[\theta] = \mathrm{deg}$ and $[\epsilon] = \mathrm{kcal/mol}$. The charge is specified with respect to the charge of a proton and u stands for atomic mass unit.

4.1 Simulation setup

For water is such an important substance, many models have been devised to investigate its properties. Some of these compare well to the experimental data. However, no single model outperforms all others across the board [39, 40]. For instance, the phase diagram of the TIP4P/2005 model compares qualitatively well to that of actual water [41]. However, the recently proposed coarse-grained, monatomic model (mW model) reproduces the melting temperature better than any other current model despite neglecting electrostatic interactions [40]. In turn, the mW model suffers from an unrealistically high mobility due to the reduced degrees of freedom [40]. Therefore, it becomes necessary to focus on specific aspects and choose the water model accordingly.

4.1.1 The SPC/E model

We used the extended simple point charge (SPC/E) model for the current investigation [42]. This choice is based on several considerations. Firstly, the results should be comparable to the literature data in order to verify the reported thermopolarisation effect in water [8]. The SPC/E model reproduces the dielectric constant of water at ambient conditions quite well ($\epsilon_\mathrm{SPC/E} = 68$ vs. $\epsilon_\mathrm{exp} = 78.5$) [39]. Moreover, the Kirkwood g-factor shows the correct temperature dependence in the temperature range of interest [43].

A schematic representation of the SPC/E model is shown in Fig. 4.1. The model has three sites and the oxygen atom acts as a Lennard-Jones centre. There is no hydrogen-hydrogen or oxygen-hydrogen short-range interaction, but all three sites carry partial charges and exhibit Coulomb interactions. All the relevant parameters are summarised in Tab. 4.1.

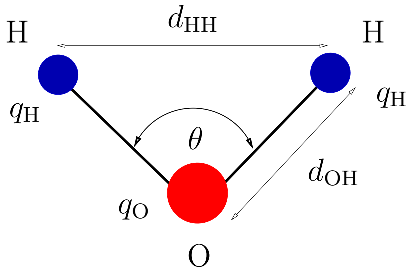

Figure 4.1: Schematic representation of the SPC/E model.

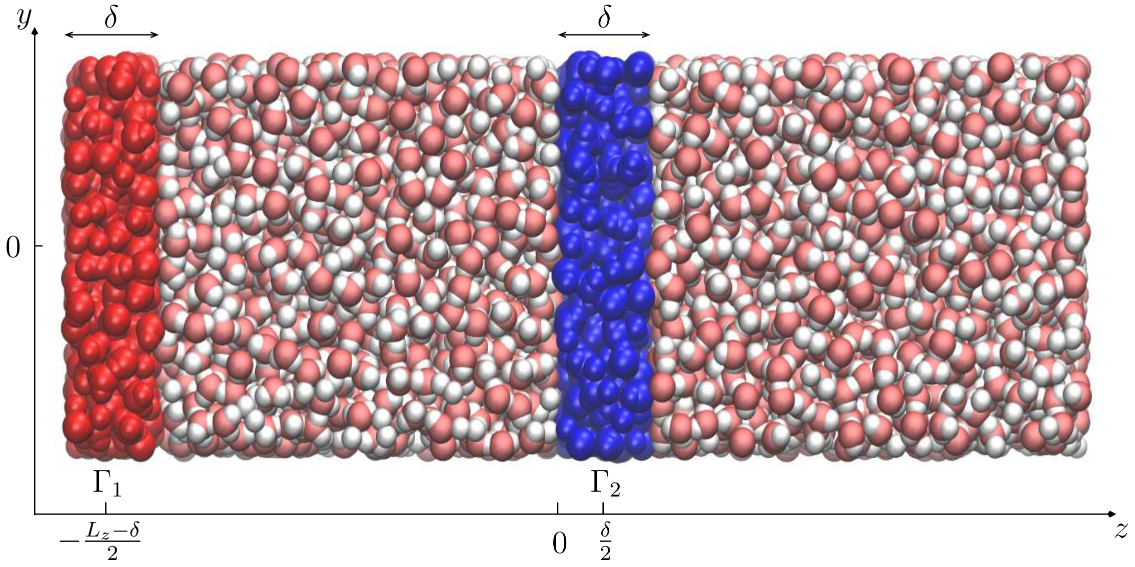

Figure 4.2: Projection of the simulation box onto the zy-plane. The box contains a hot reservoir (red region) and a cold reservoir (blue region).

4.1.2 Box setup

The general box layout was taken to be the same as in Ref. [8], but the systems considered in the current investigation comprise 4116 rather than 4500 molecules. The rectangular box illustrated in Fig. 4.2 is defined by Eqn (2.1) with the dimensions $L_x = L_y \approx L_z/3$. The exact values are given in Tab. 4.2. We remark that in Fig. 1 of Ref. [8], the hot reservoir aligns with $z = 0$, whereas our box is centred around the origin. Obviously, this should not have any effect on the results, because periodic boundary conditions are employed and the potential energy only depends on the separation vector between the individual atom pairs. The reservoirs are defined by the sets

$$\Gamma_1 = \left[-\frac{L_x}{2}, \frac{L_x}{2}\right] \times \left[-\frac{L_y}{2}, \frac{L_y}{2}\right] \times \left[-\frac{L_z}{2}, -\frac{L_z}{2} + \delta\right], \tag{4.1}$$

$$\Gamma_2 = \left[-\frac{L_x}{2}, \frac{L_x}{2}\right] \times \left[-\frac{L_y}{2}, \frac{L_y}{2}\right] \times \left[0, \delta\right], \tag{4.2}$$

where δ is the extent of the reservoir in z-direction. The centres of the reservoirs, $(-L_z + \delta)/2$ and $\delta/2$, are separated by a distance of $L_z/2$.

4.2 Simulation protocol

We used LAMMPS (1Feb14/mod) for all simulations in this chapter, apart form §4.3, where we also show results for the in-house code for validation. The initial configurations for the Ewald and Wolf production runs were prepared according to

Parameter	Ewald	Wolf	Equation(s)
r_{sr}	11	11	(2.3)
r_{c}	11	11	(2.17)–(2.18)
η	$9.42/L$	-	(2.17)
ζ	-	$7.2/L$	(2.18)
Δt	1	1	(3.23)–(3.30)
κ	0.1540	0.1415	(3.10)
δ	8	8	(4.1)–(4.2)
L_z	105.30	104.67	(2.1)
L_x	34.91	35.06	(2.1)
ρ_{m}	960	957	(1.3)
ρ	0.0321	0.0320	(4.3)

Table 4.2: Input parameters for the equilibration and production runs, where κ is only relevant for the latter. The numerical values for the damping factors η and ζ are given as multiples of $L = L_x = L_y$ for convenience. The following units apply: $[r_{\mathrm{sr/c}}] = \text{Å}$, $[\eta] = [\zeta] = \text{Å}^{-1}$, $[\Delta t] = \mathrm{fs}$, $[\kappa] = \mathrm{kcal/(mol\ fs)}$, $[\rho] = \text{Å}^{-3}$ and $[\rho_{\mathrm{m}}] = \mathrm{kg/m}^3$.

the same protocol. First, a lattice structure was created as a combination of three cubic fcc structures, each containing 343 unit cells with four molecules. The lattice was then melted in a 100 ps NVT run at 400 K using a Nosé–Hoover thermostat with a relaxation time of 1 ps. Subsequently, a 100 ps NpT run was carried out at the same temperature and a target pressure of 724.7 atm to get close to one of the thermodynamic states studied in Ref. [8]. A Nosé–Hoover barostat with a relaxation time of 2.5 ps was used for adjusting the pressure. Additional kinetic energy was then supplied to or removed from the system and the volume was rescaled such that the desired averages of T and p were achieved in a subsequent 2 ns NVE simulation (see Tab. 4.3). All the relevant input parameters are listed in Tab. 4.2.

4.3 Validation

Several tests were performed to validate the simulation setup and to confirm that the LAMMPS input scripts are correct. The in-house code, which was written by the author of this thesis, is referred to as MDSPCE and it uses Ewald summation for the treatment of electrostatic interactions. As MDSPCE was implemented to run in serial, the simulations are limited to system sizes much smaller than 4116 molecules. We therefore used a cubic box containing only 256 molecules. Apart from the box dimensions and the damping factor η, which was taken to be $5.5/L$ with respect to the small box, the settings were identical to the ones for the Ewald summation run listed in Tab. 4.2. We remark that finite size effects have to be taken into account when carrying out a comparison of the results for two systems which considerably differ in size [44]. Nevertheless, the comparison should at least allow us to identify any large-scale deviations which cannot be attributed to finite size effects alone. The LAMMPS (1Feb14/mod) results were sampled during the latter halves of the individual 2 ns *NVE* equilibration runs.

A comparison of the radial distribution functions and the velocity autocorrelation function is given in §4.3.1. The Kirkwood g-factor and the dielectric constant are discussed in §4.3.2. Energy conservation and conservation of total system momentum are analysed in §4.3.3–4.3.4. Finally, some complications with the coordinate constraining algorithm in LAMMPS are discussed in §4.3.5.

4.3.1 Radial distribution function and VACF

In contrast to a solid, a fluid has no long-range order. However, there is still some short-range order, which can be characterised by the radial distribution function. This quantity is a measure for the deviation of the actual distribution of atoms from a truly homogeneous one. Assuming a monatomic isotropic system, the radial distribution function can be expressed as [24]

$$g(r) = \frac{1}{2\pi r^2 \rho N} \left\langle \sum_{i<j} \delta(r - r_{ij}) \right\rangle, \tag{4.3}$$

where ρ is the number density. For example, $g(r) > 1$ means that there are on average more atoms in a thin spherical shell with radius r around any atom in the system than suggested by the homogeneous number density. For our system we can calculate the oxygen-oxygen (g_{OO}), oxygen-hydrogen (g_{OH}) and hydrogen-hydrogen (g_{HH}) radial distribution functions. As there is no long-range order, we expect the individual radial distribution function to approach unity for sufficiently large distances. Figure 4.3 compares the data sampled during the *NVE* equilibration runs to the results for the in-house code. As can be seen, there is perfect agreement.

Another important quantity for comparison is the velocity autocorrelation function, $\langle \mathbf{v}(\tau) \cdot \mathbf{v}(0) \rangle$, because it depends on the dynamics. It contains the information of how long the velocity of an atom or a molecule is, on average, correlated with itself at previous times. The results are shown in Fig. 4.4 and we notice very good agreement with the in-house code.

4.3.2 Kirkwood g-factor and dielectric constant

The dynamic dielectric constant $\epsilon(\omega)$, as a function of frequency ω, can be related to the dipole moment autocorrelation function $\langle \mathbf{M}(\tau) \cdot \mathbf{M}(0) \rangle$, where $\mathbf{M}$ is the total

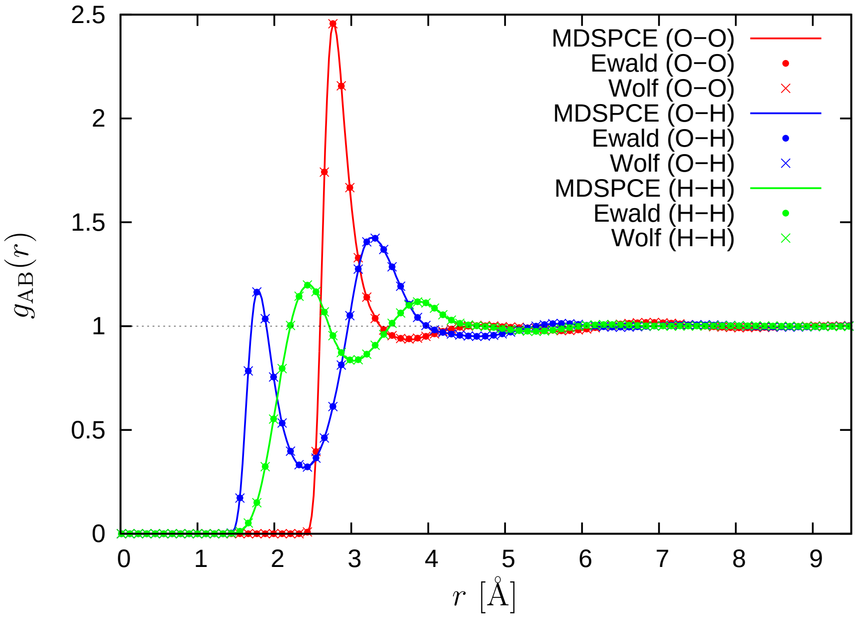

Figure 4.3: Radial distribution functions g_{OO}, g_{OH} and g_{HH} at $T = 400$ K for the MDSPCE in-house code (solid lines) and for the LAMMPS equilibration runs with Ewald summation (coloured circles) and Wolf summation (coloured crosses).

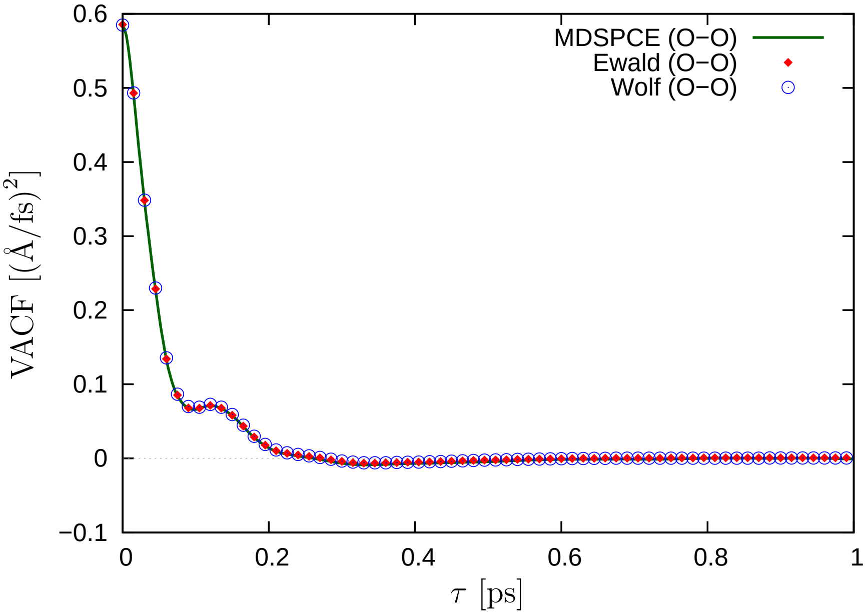

Figure 4.4: Oxygen-oxygen VACF at $T = 400$ K for the MDSPCE in-house code (solid green line) and the LAMMPS equilibration runs with Ewald summation (red diamonds) and Wolf summation (blue circles).

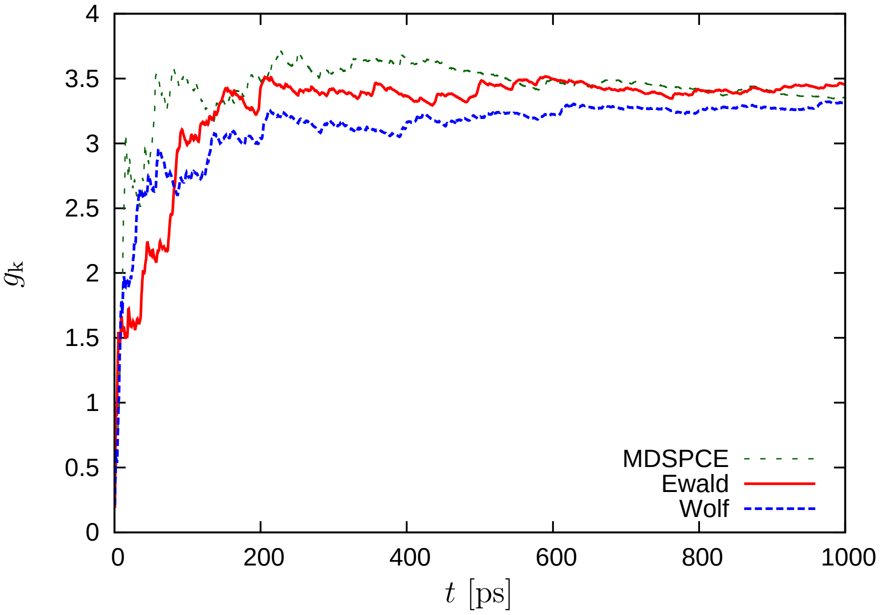

Figure 4.5: Time evolution of the Kirkwood g-factor over a time period of 1 ns for the MDSPCE in-house code (dashed green curve), LAMMPS with Ewald summation (solid red line) and Wolf summation (dashed blue curve).

dipole moment of the simulation box [45]. However, in order to determine the proportionality factor in Eqn (1.23), one only needs to know the static dielectric constant, $\epsilon = \epsilon(0)$. The calculation can therefore be simplified. We note that we used ϵ_r to denote the relative dielectric constant in §1.2.4, because SI units were used in that section. Here, we again use Gaussian units, for which $\epsilon \equiv \epsilon_\mathrm{r}$. For conducting boundary conditions, the static dielectric constant becomes [24]

$$\epsilon = 1 + \frac{4\pi\rho}{3k_\mathrm{B}T} g_\mathrm{k}\mu^2, \tag{4.4}$$

where $\boldsymbol{\mu}$ is the dipole moment of a water molecule and g_k is the Kirkwood factor, given by

$$g_\mathrm{k} = \frac{1}{N\mu^2}\left(\langle \mathbf{M}^2\rangle - \langle \mathbf{M}\rangle^2\right). \tag{4.5}$$

The simulation results are shown in Figs 4.5–4.6 and summarised in Tab. 4.3. The dielectric constants are in reasonable agreement with the values reported in Ref. [8]. The run with Wolf summation exhibits a slightly lower value, but longer simulation times would be necessary to confirm a deviation definitively. We also note that the density was slightly higher in our setup, which could partly account for the difference. In Ref. [43], ϵ was calculated for a system of 216 molecules at temperatures of 298 K and 373 K. Given these values, we can approximate the temperature derivative of ϵ and extrapolate the value for 373 K to 400 K, which yields $\epsilon \approx 44$. This indicates that the magnitude is correct and both comparisons therefore suggest a successful implementation.

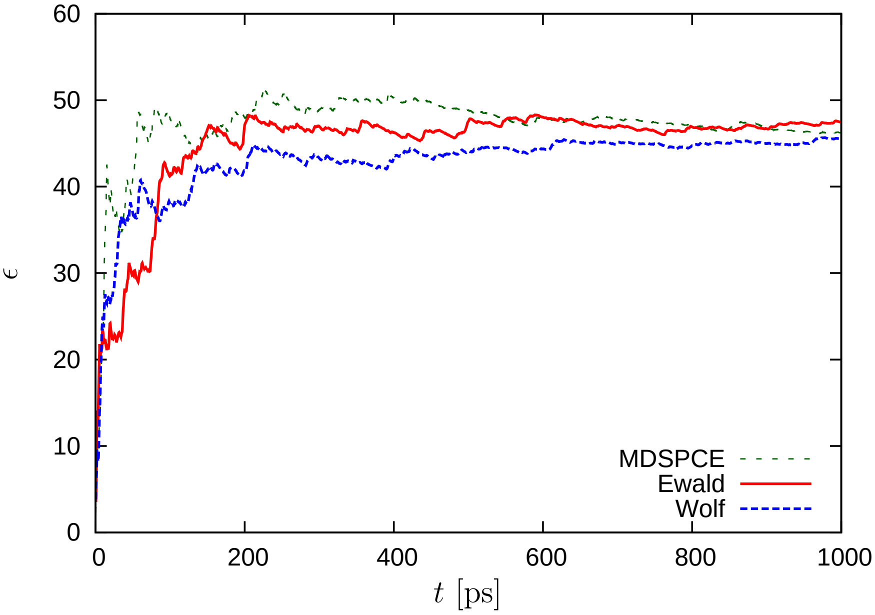

Figure 4.6: Time evolution of the static dielectric constant over a time period of 1 ns for the MDSPCE in-house code (dashed green curve), LAMMPS with Ewald summation (solid red line) and Wolf summation (dashed blue curve).

Run	T_{eq} [K]	P_{eq} [atm]	g_{k}	ϵ
MDSPCE	398.5 ± 0.3	640 ± 10	3.4	46.2
Ewald	400.54 ± 0.03	725 ± 3	3.5	47.5 (47.4)
Wolf	400.17 ± 0.03	721 ± 3	3.3	45.6 (46.2)

Table 4.3: Comparison of the Kirkwood g-factors and dielectric constants. The errors for temperatures and pressures were calculated using block average analysis. For g_{k} and ϵ, the value at 1 ns was taken, as both quantities are calculated from cumulative averages. Literature values for the dielectric constant were taken from Ref. [8] and are shown in brackets.

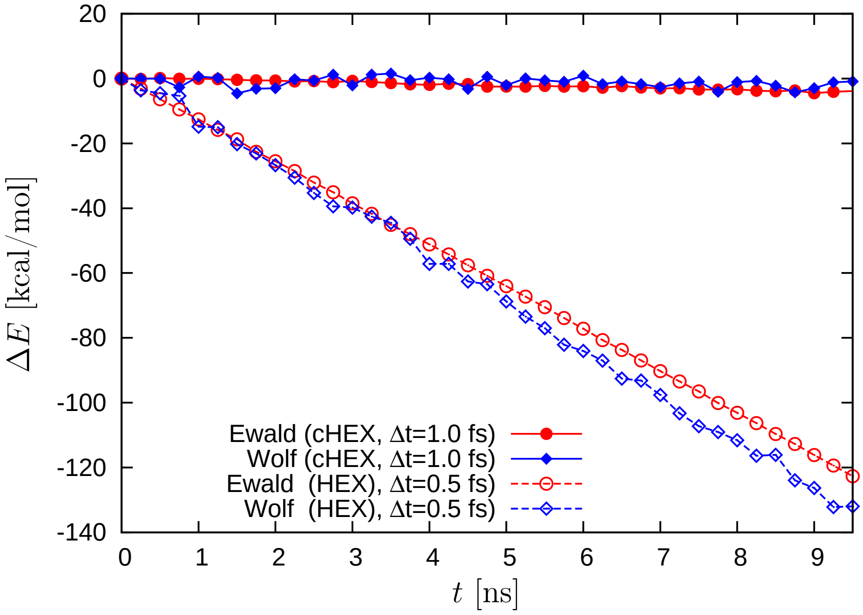

Figure 4.7: Time evolution of the total energy for gradient simulations with the cHEX algorithm and the HEX algorithm. The gradients in the cHEX simulations were 5.09 K/Å for Ewald summation (solid red line with filled circles) and 5.14 K/Å for Wolf summation (solid blue line with filled diamonds). The gradients in the simulations with the HEX algorithm were 5.22 K/Å for Ewald summation (dashed red line with circles) and 5.69 K/Å for Wolf summation (dashed blue line with diamonds). The average temperature was approximately 403 K in all cases.

4.3.3 Energy conservation

In §3.2.3, we demonstrated that the cHEX algorithm conserves the energy of a Lennard-Jones fluid very well over 10^6 steps. However, it has not been confirmed yet that this property also holds for rigid molecules, such as SPC/E water. In Fig. 4.7, the time evolution of the total energy during the first 9.5 ns of the production run is compared to previous results obtained with the HEX algorithm. The time range is limited to the first 9.5 ns of the production runs because the simulations using the HEX algorithm were not extended any further due to the high energy loss. In fact, the comparison was carried out in favour of the HEX algorithm, because a smaller timestep of 0.5 fs was employed in the HEX simulations.

The cHEX algorithm conserves the energy well on this scale, but a slight energy drift can be noticed on a finer scale. However, the drift is very small compared to the energy loss caused by the HEX algorithm. The gradient in the simulation using the HEX algorithm and Wolf summation was 5.69 K/Å and therefore slightly higher than the others ranging between 5.09 – 5.22 K/Å. For a quantitative comparison, we therefore consider the energy loss for Ewald summation which is about -120 kcal/mol after 9.5 ns. With an average temperature of $T = 400$ K, this corresponds to $-151\ k_{\mathrm{B}}T$. Even though a higher timestep was employed, the energy only changed by approximately 5 $k_{\mathrm{B}}T$ when the cHEX algorithm was employed. Assuming that the quadratic scaling of the error also holds for rigid molecules, we conclude that the HEX algorithm loses at least 100 times more energy than the cHEX algorithm.

4.3.4 Centre of mass velocity

In theory, the centre of mass velocity $\mathbf{v}_\Omega$ should not be affected by the HEX algorithm (see Eqn (3.9)). We did not explicitly verify that this was indeed the case for the runs shown in Fig. 4.7, but we tested the cHEX algorithm. Figure 4.8 shows the time evolution of v_Ω^2 on a log-scale for both production runs. Over the first few nanoseconds, we can observe an increase of this quantity by about three orders of magnitude. After 10 ns, the error keeps increasing sublinearly on this scale. The rapid initial growth might be caused by numerical round-off errors in the approximation of $\dot{\mathbf{v}}_\Gamma$ in Eqn (3.29), because two very small quantities are subtracted from one another, which can result in a loss of significance. This could be tested by setting the approximation of $\dot{\mathbf{v}}_\Gamma$ to zero if v_Γ^2 is below a certain threshold value.

A closer inspection of the time evolution of the individual components of $\mathbf{v}_\Omega$ in Fig. 4.9 reveals that the z-component is mainly responsible for the increase of v_Ω^2. It would be very interesting to see how the error scales with the timestep. Higher-order approximations of $\dot{\mathbf{v}}_\Gamma$ and $\dot{\xi}_i$ could possibly yield an improvement.

Previously, Lukes and Liang reported that the algorithm of Ikeshoji and Hafskjold does not strictly conserve the instantaneous total system momentum [46]. They found that this quantity was fluctuating rapidly around zero. Since we only collected 0.5 ps averages of $v_{\Omega,\alpha}$ during the simulation, the presence of fluctuations cannot directly be confirmed with the data shown in Fig. 4.9. We therefore launched very short simulations from a binary restart file at 10 ns with Ewald summation and exactly the same settings as for the production run. However, we could not observe any fluctuations of $v_{\Omega,\alpha}$ on a scale of 10^{-13} Å/fs. Whether this is an improvement arising from the additional coordinate integration step in the cHEX algorithm is subject to further investigation.

4.3.5 Coordinate constraints

The constraining algorithm used in LAMMPS suffers from a problem we wish to discuss in more detail. In §2.3, it was shown how rigidity of the molecules is imposed in LAMMPS in combination with the velocity Verlet integration scheme. Constrain-

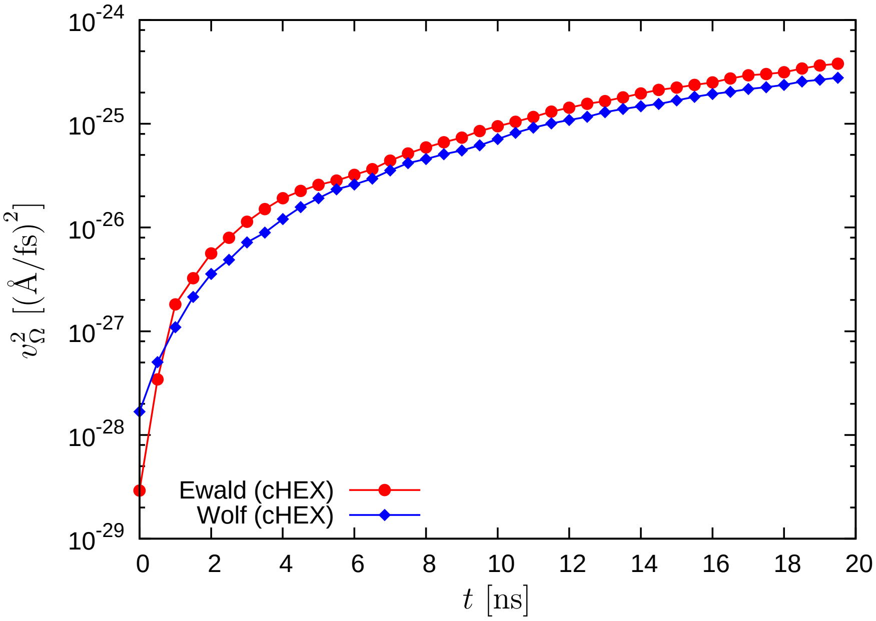

Figure 4.8: Time evolution of the magnitude of the centre of mass velocity for simulations with the cHEX algorithm. The datasets for Ewald summation (solid red line with circles) and Wolf summation (solid blue line with diamonds) were sampled during the 20 ns production runs. Each symbol represents a time average over 0.5 ps.

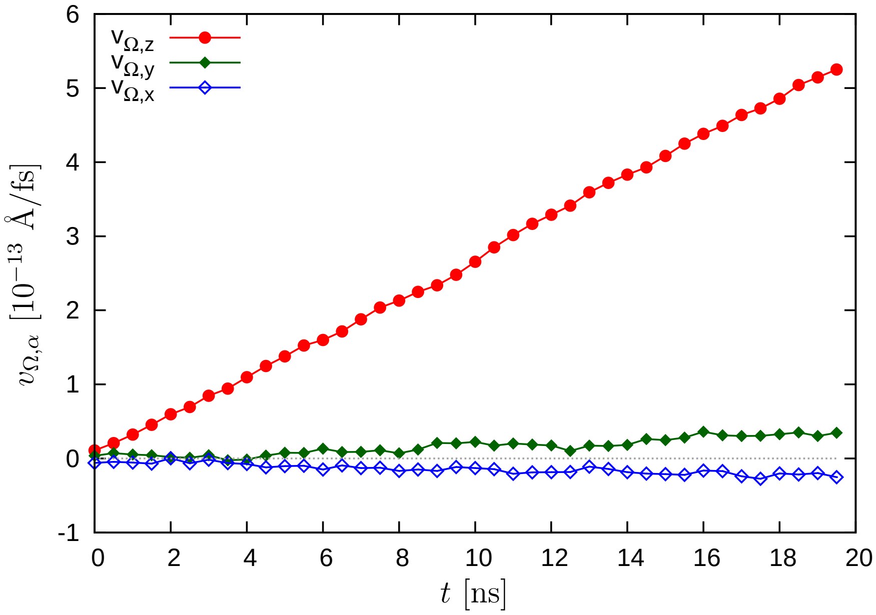

Figure 4.9: Time evolution of the individual components of $\mathbf{v}_\Omega$ for the production run with Ewald summation. Each symbol represents a time average over 0.5 ps. The dotted line represents zero and is shown for reference.

ing forces are calculated in such a way that the imposed bond lengths are satisfied after the next coordinate integration step. This procedure is very similar to the original SHAKE algorithm proposed by Ryckaert and co-workers in Ref. [33]. With SHAKE and Verlet integration, the time evolution of the coordinates is determined by

$$\mathbf{r}_i(t+\Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t-\Delta t) + \frac{\Delta t^2}{m_i}\Big[\mathbf{f}_i(t) + \mathbf{g}_{\mathrm{RR},i}(t)\Big], \tag{4.6}$$

where $\mathbf{g}_{\mathrm{RR},i}$ is a constraining force [31]. However, as pointed out by Andersen in Ref. [31], the velocity Verlet version of SHAKE requires an additional velocity correction step, or else the time derivatives of the constraints will not be satisfied exactly. Andersen called this modified algorithm RATTLE and the combined integration scheme is given by

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \Delta t\dot{\mathbf{r}}_i(t) + \frac{\Delta t^2}{2m_i}\Big[\mathbf{f}_i(t) + \mathbf{g}_{\mathrm{RR},i}(t)\Big], \tag{4.7}$$

$$\dot{\mathbf{r}}_i(t+\Delta t) = \dot{\mathbf{r}}_i(t) + \frac{\Delta t}{2m_i}\Big[\mathbf{f}_i(t) + \mathbf{g}_{\mathrm{RR},i}(t) + \mathbf{f}_i(t+\Delta t) + \mathbf{g}_{\mathrm{RV},i}(t)\Big], \tag{4.8}$$

where $\mathbf{g}_{\mathrm{RV},i}$ comprises the additional velocity correction [31]. Applying the SHAKE algorithm in combination with velocity Verlet integration, as described in §2.3, only guarantees that the constraints on the coordinates are satisfied at time $t+\Delta t$,

$$r_{ij}^2(t+\Delta t) = d_{ij}^2, \tag{4.9}$$

where $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$ and d_{ij} is the distance to be imposed between the sites i and j. The additional velocity correction in the RATTLE algorithm is such that the time derivative of Eqn (4.9) is satisfied as well,

$$2\mathbf{r}_{ij}(t+\Delta t)\cdot\mathbf{v}_{ij}(t+\Delta t) = 0. \tag{4.10}$$

In other words, there should be no velocity component along a bond, because this would correspond to a vibration, which cannot happen with rigid bonds. In order to test whether the current SHAKE implementation in LAMMPS satisfies Eqn (4.10), we ran very short simulations of a system comprising 108 SPC/E water molecules. The initial time was $t=0$ and we monitored the quantity

$$\phi = \max_{\substack{(i,j)\text{ fixed}\\ n}} \Big| 2\mathbf{r}_{ij}(n\Delta t)\cdot\mathbf{v}_{ij}(n\Delta t)\Big|, \tag{4.11}$$

which gives the largest absolute deviation from the condition (4.10). Of course the maximum over i,j in Eqn (4.11) is restricted to pairs (i,j) with fixed bonds. The coordinate constraints in Eqn (4.9) were perfectly satisfied up to a tolerance of less than 10^{-12} Å throughout the run. The dependence of ϕ on the timestep is shown in Fig. 4.10. We can clearly see that the maximum error ϕ scales as $\mathcal{O}(\Delta t^2)$.

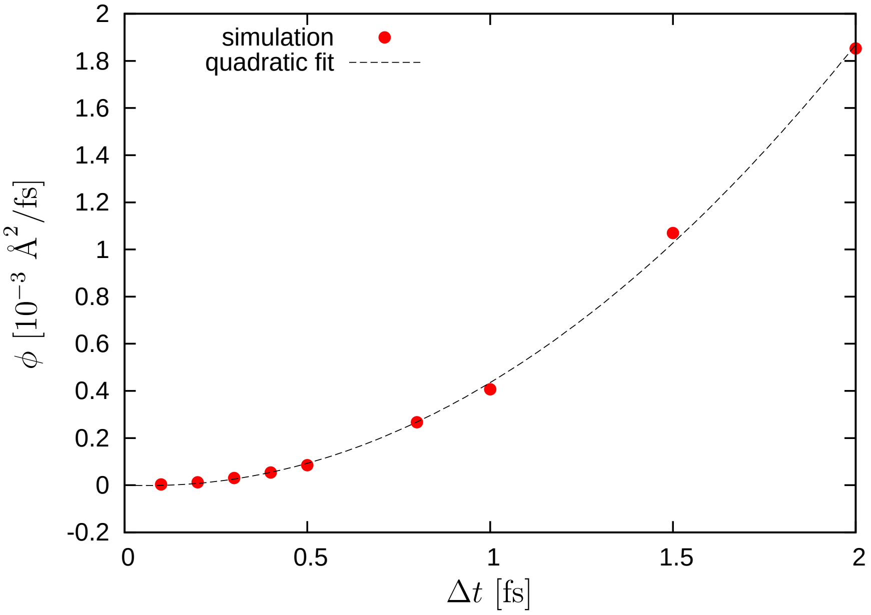

Figure 4.10: Dependence of ϕ on the timestep. Each value (red circles) was calculated in a simulation of 1000 timesteps. The temperature was approximately 298 K in all cases. The runs started from the same, previously equilibrated binary restart file. In addition, a quadratic fit (dashed black line) is shown.

This is exactly what one would expect, because the constraining forces $\mathbf{g}_{\mathrm{RV},i}$ are of $\mathcal{O}(\Delta t)$ [31].

How this inconsistency affects the dynamics of the system is hard to assess. Our validation tests in §4.3.1–4.3.2 show good agreement with the in-house code which uses SHAKE and Verlet integration. Therefore, we believe that the effect caused by this error is small and the validity of the results in §4.4 is not in jeopardy.

4.4 Calculation of the electrostatic field

In this section we discuss the calculation of the electrostatic field. Our considerations are based on the assumption that the underlying three-dimensional problem can be reduced to a one-dimensional problem, as suggested by Bresme and co-workers [1, 3, 6, 8]. However, the exact expression we use for calculating the field from the charge density differs from the approach in Refs [1, 3, 6, 8].

We first divided the simulation box into N_{bins} regions which we call *bins*. The xy-cross section of each bin is orthogonal to the z-direction and they are all of equal volume. The width of a bin is therefore given by $\Delta z = L_z/N_{\text{bins}}$. During the production runs we sampled the average number densities of oxygen, ρ_{O}, and hydrogen, ρ_{H}, for every bin in order to capture the spatial variation. Let us denote the average number densities for bin i by $\rho_{\mathrm{O}i}$ and $\rho_{\mathrm{H}i}$, respectively, and assign these values to the centre of the bin, i.e. $z_i = -0.5L_z + (i - 0.5)\Delta z$ for $i \in \{1, 2, .., N_{\mathrm{bins}}\}$. The average charge density for bin i is then given by [1]

$$\rho_{\mathrm{q}i} = q_{\mathrm{O}}\rho_{\mathrm{O}i} + q_{\mathrm{H}}\rho_{\mathrm{H}i}. \tag{4.12}$$

This quantity can be used to solve Poisson's equation numerically. In order to derive the relevant relations, we assume that the charge density is continuous. As a starting point, we consider the previously introduced expression for the electrostatic potential,

$$\Phi(\mathbf{r}) = \int_V \mathrm{d}^3r'\ G(\mathbf{r} - \mathbf{r}')\rho_q(\mathbf{r}'). \tag{4.13}$$

If we assume that the charge density is only a function of z, we can rewrite the above equation as [28]

$$\Phi(z) = \int_{-L_z/2}^{L_z/2} \mathrm{d}z\ \bar{G}(z - z')\rho_q(z'), \tag{4.14}$$

where we introduced the modified Green's function

$$\bar{G}(z) = \int_{-L_x/2}^{L_x/2} \mathrm{d}x \int_{-L_y/2}^{L_y/2} \mathrm{d}y\ G(x - x', y - y', z). \tag{4.15}$$

Using this definition the electrostatic field in the z-direction is formally given by

$$E_z(z) = -\frac{\mathrm{d}\Phi}{\mathrm{d}z}(z) = -\int_{-L_z/2}^{L_z/2} \mathrm{d}z'\ \frac{\mathrm{d}\bar{G}}{\mathrm{d}z}(z - z')\rho_q(z'). \tag{4.16}$$

The modified Green's function $\bar{G}$ depends on the three-dimensional Green's function used in the simulation (see §2.2.1–2.2.2). Following Neumann [28], we will show individually for both summation techniques how the modified Green's functions can be calculated from G_{E} and G_{W}, respectively.

4.4.1 Ewald summation

The Green's function G_{E} approximates the expression

$$G(\mathbf{r}) = \frac{1}{L_xL_yL_z}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi}{k^2}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}, \tag{4.17}$$

where

$$\mathbf{k} = (k_x, k_y, k_z) = 2\pi(m_x/L_x, m_y/L_y, m_z/L_z), \tag{4.18}$$

and m_x, m_y and m_z are integers [28]. If we integrate this expression, we get the modified Green's function

$$\begin{aligned}\bar{G}_{\mathrm{E}}(z) &= \int\limits_{-L_x/2}^{L_x/2} \mathrm{d}x \int\limits_{-L_y/2}^{L_y/2} \mathrm{d}y \; G_{\mathrm{E}}(x, y, z) \\ &= \int\limits_{-L_x/2}^{L_x/2} \mathrm{d}x \int\limits_{-L_y/2}^{L_y/2} \mathrm{d}y \frac{1}{L_x L_y L_z} \sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi}{k^2} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} \\ &= \frac{1}{L_z} \sum_{k_z\neq 0} \frac{4\pi}{k_z^2} \mathrm{e}^{ik_z z}. \end{aligned} \tag{4.19}$$

In the last step, we make use of the fact that the integration eliminates all terms in the summation for which $k_x \neq 0$ or $k_y \neq 0$. The inverse Fourier transform in Eqn (4.19) is given by [28]

$$\tilde{G}_{\mathrm{E}}(z) = 2\pi \left(-|z| + \frac{L_z}{6} + \frac{z^2}{L_z} \right). \tag{4.20}$$

We note that this function is not periodic unlike the RHS of Eqn (4.19). However, this can easily be corrected by modifying the argument of $\tilde{G}_{\mathrm{E}}$ by nesting the function

$$d_{\mathrm{pbc}}(z) = z - L_z \;\mathrm{nint}\left(\frac{z}{L_z}\right), \tag{4.21}$$

where $\mathrm{nint}(x)$ rounds its argument to the nearest integral number. Then we finally find that

$$\bar{G}_{\mathrm{E}}(z) = \tilde{G}_{\mathrm{E}}(d_{\mathrm{pbc}}(z)), \tag{4.22}$$

which is equivalent to Eqn (4.19).

4.4.2 Wolf summation

We can use exactly the same procedure to find the corresponding averaged Green's function for Wolf summation. The integration can be simplified by switching to polar coordinates.

With the definition of $r^2 = x^2 + y^2 + z^2 = s^2 + z^2$ and Eqn (2.18) we get the modified

Green's function

$$\tilde{G}_{\mathrm{W}}(z) = \int\limits_{-L_x/2}^{L_x/2} \mathrm{d}x \int\limits_{-L_y/2}^{L_y/2} \mathrm{d}y\ G_{\mathrm{W}}(x,y,z) \tag{4.23}$$

$$= 2\pi \int\limits_{0}^{r_{\mathrm{c}}} s\mathrm{d}s \left[\frac{\mathrm{erfc}(\zeta\sqrt{s^2+z^2})}{\sqrt{s^2+z^2}} - \frac{\mathrm{erfc}(\zeta r_{\mathrm{c}})}{r_{\mathrm{c}}} \right]. \tag{4.24}$$

In Appendix C it is shown that the integral is given by

$$\tilde{G}_{\mathrm{W}}(z) = 2\pi \Bigg[-|z|\mathrm{erfc}(|z|\zeta) + \sqrt{z^2+r_{\mathrm{c}}^2}\ \mathrm{erfc}(\zeta\sqrt{z^2+r_{\mathrm{c}}^2}) + \frac{\mathrm{e}^{-z^2\zeta^2}}{\sqrt{\pi}\zeta}\left(1-\mathrm{e}^{-r_{\mathrm{c}}^2\zeta^2}\right) - \frac{1}{2} r_{\mathrm{c}}\mathrm{erfc}(r_{\mathrm{c}}\zeta) \Bigg]. \tag{4.25}$$

If we modify the distance calculation using the function d_{pbc}, we finally obtain

$$\bar{G}_{\mathrm{W}}(z) = \tilde{G}_{\mathrm{W}}(d_{\mathrm{pbc}}(z)). \tag{4.26}$$

4.5 Results

In this section we present the results for the 20 ns production runs with Ewald and Wolf summation. In order to capture the spatial dependence, we used a fine resolution of $\Delta z = 0.554$ Å to calculate each quantity. However, sometimes we show the results at a coarser resolution $\Delta z \mapsto \Delta z' = m\Delta z$, with $m = 5$ or 10, where we average over the fine bins. All resolutions are compatible with $N_{\mathrm{bins}} = 190$.

4.5.1 Temperature

The spatial variation in temperature along the z-direction is shown in Fig. 4.11. There is no noticeable difference between the results with Wolf and Ewald summation. The peak temperature at the centre of the hot reservoir is about 542 K and the lowest temperature at the centre of the cold reservoir is about 288 K. The gradient is linear in the intervals $[-\frac{L_z}{2}+\delta, 0]$ and $[\delta, \frac{L_z}{2}-\delta]$, i.e. the regions between the reservoirs. Furthermore, the temperature profile is symmetric with respect to $z = 4$ Å, which is in accordance with the setup. The results for linear least square fits and heat fluxes are given in Tab. 4.4. The heat fluxes are calculated from Eqn (3.11) and κ is an input parameter which was adjusted by trial and error. It was reported previously that the truncation of long-range electrostatic interactions leads to a lower thermal conductivity [5], which explains why the lower heat flux for Wolf summation leads to a temperature gradient similar to that obtained for Ewald summation.

Run	$\lvert\nabla T\rvert$ [K/Å]	$\lvert J_{q,z}\rvert$ [10^{10} W/m^2]	T_{avg} [K]
Ewald	5.09 ± 0.02	4.39	403.61 ± 0.01
Wolf	5.14 ± 0.02	4.00	403.33 ± 0.01

Table 4.4: Temperature gradients and heat fluxes for the production runs. All errors were calculated using block average analysis. The values for $J_{q,z}$ do not have error estimates, since they are input parameters (Eqn (3.11)).

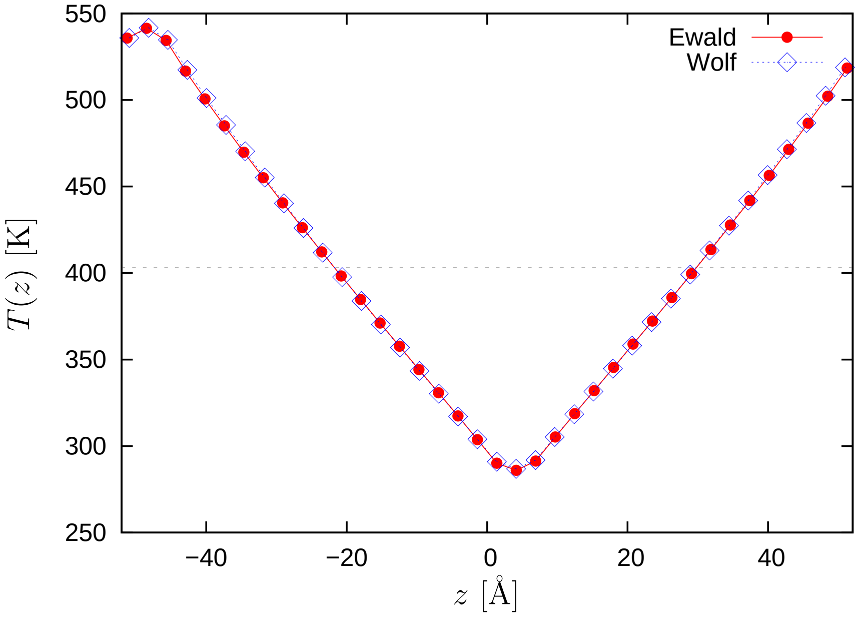

Figure 4.11: Spatial variation of the temperature along the z-axis for Ewald summation (red solid line with circles) and Wolf summation (blue dotted line with diamonds). The dashed line represents the average temperature of about 403 K for both runs. The spatial resolution is 2.77 Å (coarse-graining with $m = 5$).

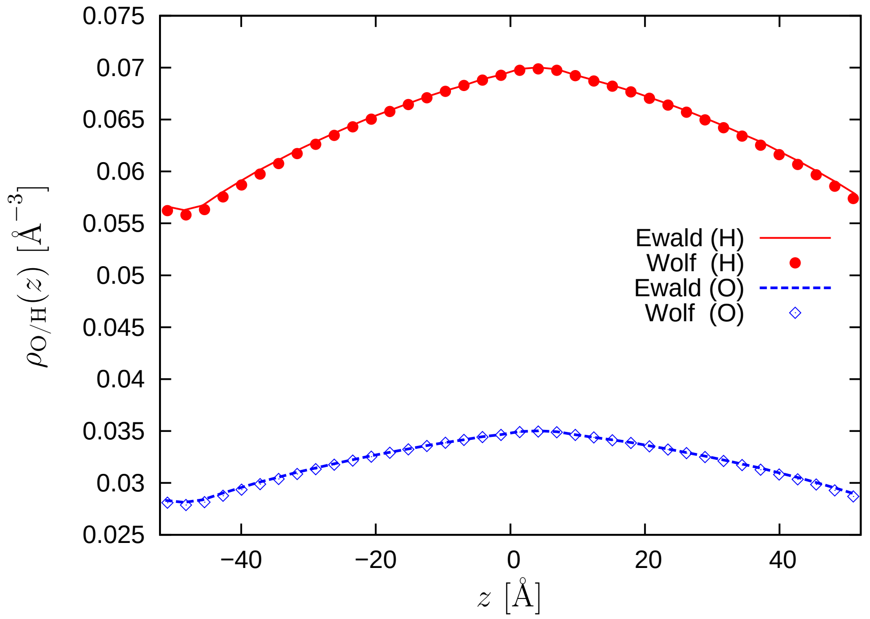

Figure 4.12: Spatial variation of the average oxygen and hydrogen number densities along the z-axis for Ewald summation (solid red line and dashed blue line) and Wolf summation (red circles and blue diamonds). The spatial resolution is 2.77 Å (coarse-graining with $m = 5$).

4.5.2 Number densities and charge density

The measured average number densities are shown in Fig. 4.12. All curves are symmetric with respect to $z = 4$ Å. The treatment of long-range electrostatic interactions does not seem to have a large effect on the individual average number densities of oxygen and hydrogen. However, there are some noticeable differences in the vicinity of the hot reservoir.

Figures 4.13–4.14 show the average charge densities for two different resolutions. The quantity fluctuates around zero for both methods and the fluctuations are strongest near the cold reservoir (see Fig. 4.13). Coarsening the resolution results in a decrease of the average charge density, as can be seen in Fig. 4.14. This is reasonable, because the molecules are charge neutral and therefore only the molecules which are not fully contained within an individual bin lead to a non-zero contribution. Since the net charge in a bin is divided by a larger volume, the average charge density necessarily has to decrease. In the limit of only one bin, the charge density would be exactly zero, because the system is charge neutral.

The average charge densities for Ewald summation and Wolf summation seem to follow a similar trend (Fig. 4.13), but longer simulation times would be necessary to improve the statistics.

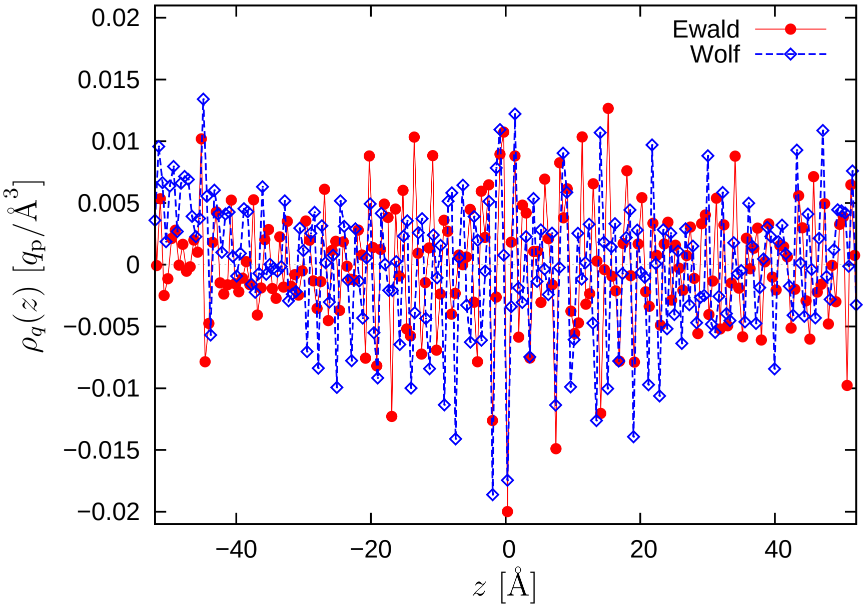

Figure 4.13: Spatial variation of the average charge density along the z-axis for Ewald summation (solid red line with circles) and Wolf summation (dashed blue line with diamonds). The spatial resolution is 0.554 Å. The quantity q_p denotes the charge of a proton.

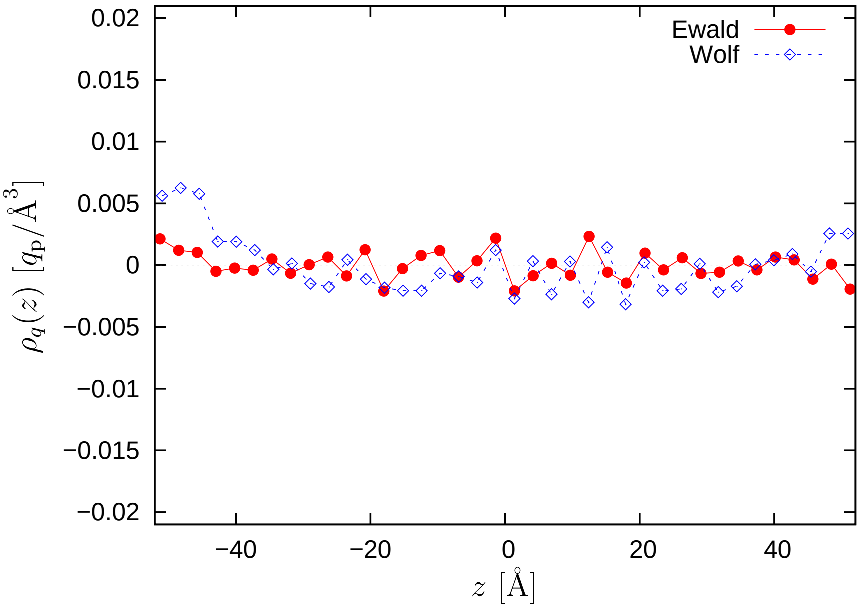

Figure 4.14: Spatial variation of the average charge density along the z-axis for Ewald summation (solid red line with circles) and Wolf summation (dashed blue line with diamonds). The spatial resolution is 2.77 Å (coarse-graining with $m = 5$). The quantity q_p denotes the charge of a proton.

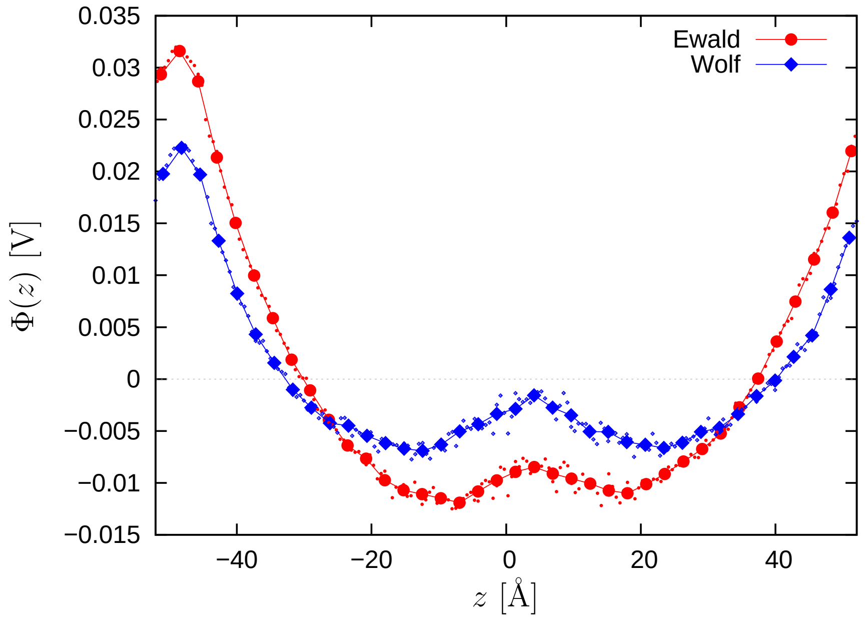

Figure 4.15: Spatial variation of the electrostatic potential along the z-axis. The coarse-grained results for Ewald summation (solid red line with big circles) and Wolf summation (solid blue line with big diamonds) have a resolution of 2.77 Å ($m = 5$). The corresponding high-resolution results (0.554 Å) are represented by small symbols using the same colours.

4.5.3 Electrostatic potential

Given a charge density, the electrostatic potential can be calculated by convolution with the appropriate modified Green's function (Eqn (4.14)). For Ewald summation we use $\bar{G}_\mathrm{E}$ (Eqn (4.22)) and for Wolf summation $\bar{G}_\mathrm{W}$ (Eqn (4.26)) to evaluate the integral in Eqn (4.14) numerically. The approximation was carried out using the midpoint rule and the average charge density with the highest resolution of 0.554 Å. A coarse-graining step with $m = 5$ followed afterwards and the results for both resolutions are presented in Fig. 4.15.

The potential is symmetric around $z = 4$ in accordance with our setup. There is also good qualitative agreement between the results for both summation techniques, but the gradients are slightly steeper for Ewald summation. Furthermore, we notice that the potential is continuous and the gradient vanishes at the centres of the reservoirs, i.e. at $z = -48.7$ Å and $z = 4$ Å. (It also vanishes at $z \approx 4 \pm 14$ Å.)

4.5.4 Electrostatic field

We only need to consider the z-component of the electrostatic field because of symmetry considerations. We did not directly compute the field from the potential. Instead, we calculated the convolution of the charge density with the derivative of

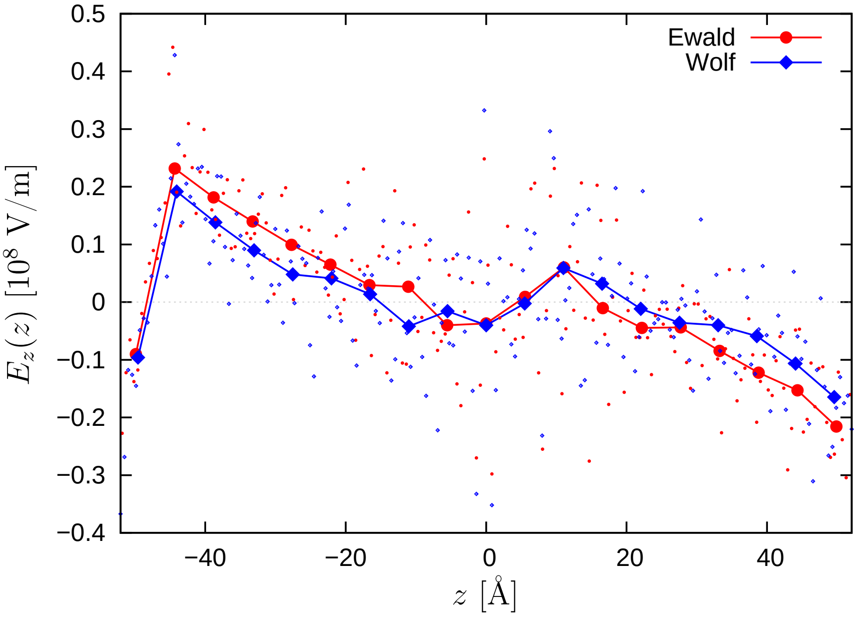

Figure 4.16: Spatial variation of the electrostatic field along the z-axis. The coarse-grained results for Ewald summation (solid red line with big circles) and Wolf summation (solid blue line with big diamonds) have a resolution of 5.54 Å ($m = 10$). The corresponding high-resolution results (0.554 Å) are represented by small symbols using the same colours.

the corresponding modified Green's function according to Eqn 4.16. The approximation was again carried out using the midpoint rule with the highest resolution and a coarse-graining step with $m = 10$ followed afterwards. The results for both resolutions are presented in Fig. 4.16. Due to the symmetry of the potential, both curves in Fig. 4.16 are necessarily antisymmetric around the centre of the cold reservoir at $z = 4$ Å. The fields are strongest in the vicinity of the hot reservoir; away from this point, the magnitude of the field decreases almost linearly. Interestingly, both curves exhibit local minima close to the cold reservoirs before they vanish at $z = 4$ Å. The absolute value of the field is almost everywhere higher for Ewald summation than for Wolf summation. However, the order of the magnitude of the field is the same in both cases. The high-resolution results for the field (Fig. 4.16) exhibit larger fluctuations as compared to those for the potential (Fig. 4.15). This is not surprising because the derivative with respect to z yields an additional factor ik_z in the summation in Eqn (4.19).

It ought to be borne in mind that the field strongly depends on the Green's function that was used for the calculation. To illustrate this point, we recalculated the Wolf summation result with the modified Green's function $\bar{G}_{\mathrm{E}}$ corresponding to Ewald summation. This result is shown in Fig. 4.17 and compared to the old one. We see a sudden increase by an order of magnitude compared to the previous results. Furthermore, the local extrema in the vicinity of the cold reservoir vanished and the

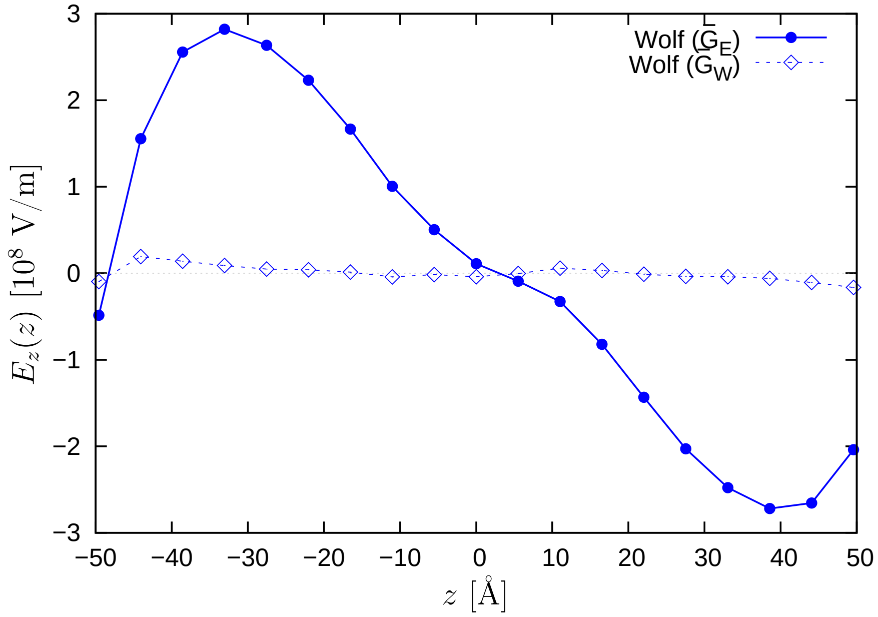

Figure 4.17: Spatial variation of the electrostatic field along the z-axis for Wolf summation calculated with $\bar{G}_\mathrm{E}$ (solid blue line with circles) and $\bar{G}_\mathrm{W}$ (dashed blue line with diamonds), respectively. The spatial resolution is 5.54 Å (coarse-graining with $m = 10$).

maximum and minimum are shifted towards the cold reservoir. Periodic boundary conditions are taken into account for both fields shown in Fig. 4.17. However, as the results demonstrate, the field also depends significantly on the Green's function employed during the simulation. This makes sense because the molecules move according to the short-range and long-range forces and the long-range contribution depends on the particular Green's function.

4.5.5 Comparison with literature data

In order to be able to compare our results more fully with the results reported by Armstrong and co-workers (Ref. [8]), we carried out an additional simulation. Similar settings to the ones in Ref. [8] were chosen in order to get a comparable stationary state (see Tab. 4.5). The heat flux we imposed was slightly higher than theirs, which implies an analogous increase in the temperature gradient. As can be seen from Tab. 4.5, the pressure reported in the literature data was almost 80 atm higher than the one we measured in our simulation. This discrepancy might also be caused by the different heat fluxes used in the simulation.

A considerable difference between the two approaches arises from the way we cal-

Parameter	Wolf	Ref. [8]
$\lvert\nabla T\rvert$	5.22 ± 0.02	5.15 ± 0.02
$J_{q,z}$	3.90	3.81
T_{eq}	400.60 ± 0.03	400
T	404.05 ± 0.01	
P	647 ± 10	724.7 ± 24.2
L_z	105.9	110
δ	8	4
Δz	0.56	0.54
ρ_{m}	0.934	0.934
ζ	$7.2/L_z$	$7.2/L_z$
r_{c}	11	11
r_{sr}	11	11

Table 4.5: Simulation settings and average values for the comparison with Ref. [8]. The following units apply: $[\nabla T] = [\mathrm{K/Å}]$, $[J_{q,z}] = [10^{-10}\ \mathrm{W/m^2}]$, $[T] = [\mathrm{K}]$, $[P] = [\mathrm{atm}]$, $[L_z] = [\delta] = [\Delta z] = [r_{\mathrm{c/sr}}] = [\mathrm{Å}]$, $[\rho_{\mathrm{m}}] = [\mathrm{g/cm^3}]$ and $[\zeta] = [\mathrm{Å}^{-1}]$.

culate the electrostatic field. In Ref. [8], the field is given by

$$E_z(z) = \frac{1}{\epsilon_0} \int\limits_{-\infty}^{z} \mathrm{d}z' \rho_q(z'), \tag{4.27}$$

as opposed to

$$E_z(z) = -\frac{1}{4\pi\epsilon_0} \int\limits_{-L_z/2}^{L_z/2} \mathrm{d}z' \, \frac{\mathrm{d}\bar{G}_{\mathrm{W}}}{\mathrm{d}z}(z - z')\rho_q(z'), \tag{4.28}$$

which is the expression we used. The division by $4\pi\epsilon_0$ is necessary because we used Gaussian units to derive $\bar{G}_{\mathrm{W}}$. It is not immediately obvious what the lower boundary $-\infty$ of the integral in Eqn (4.27) should be for a system with periodic boundary conditions. We interpreted the lower boundary $-\infty$ as the centre of the hot reservoir, where the field should vanish due to the symmetrical setup. In Ref. [8], the mirror symmetry of the field around the centre of the cold reservoir was then employed in order to improve the statistics. Their production run was 10 ns long, whilst ours was 20 ns long, and therefore we expect the statistics to be comparable even without the folding step.

The results are shown in Fig. 4.18. As can be seen, the fields differ significantly, although they were calculated from the same charge density but with two different formulas (Eqns (4.27)–(4.28)). We think that calculating the field with Eqn (4.27) is inconsistent with the underlying treatment of electrostatic interactions, whereas Eqn (4.28) is in accordance with the dynamics of the simulation.

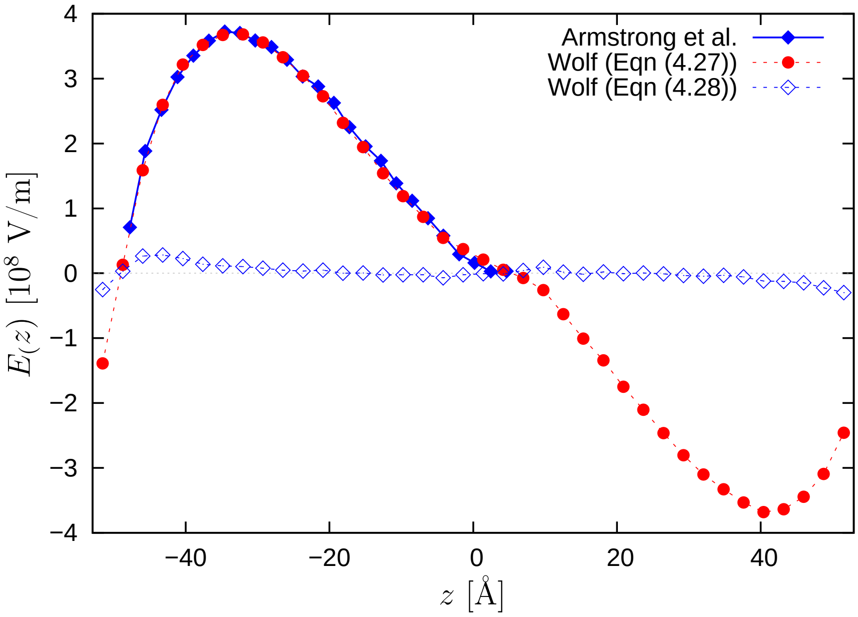

Figure 4.18: Comparison of the electrostatic field obtained for the new Wolf summation run with the literature data. The results of Armstrong and co-workers (solid blue line with filled diamonds) were taken from Ref. [8] and shifted such that the centres of the hot reservoirs coincide. The field as calculated with Eqn (4.27) (dashed red line with circles) is compared to the one obtained with Eqn (4.28) (dashed blue line with diamonds).

4.6 Discussion

The results in the previous section suggest that the electrostatic field caused by the thermo-polarisation effect does not depend significantly on the underlying treatment of long-range forces. In particular, we demonstrated that the field obtained with the truncated Wolf summation approach is comparable to that obtained with the widely used Ewald summation (Fig. 4.16). However, this is only the case if the calculation takes into account the correct modified Green's function. This function is derived directly from the three-dimensional Green's function used in the simulation (Eqn (4.15)) and is therefore compatible with the boundary conditions.

We can reproduce the results reported by Armstrong and co-workers (Fig. 4.18) if we calculate the field as suggested in Ref. [8] (Eqn (4.27)). However, we believe that this expression is compatible with neither the dynamics of the simulation nor the imposed boundary conditions. Taking both considerations into account results in weaker fields of the order of $\pm 10^7$ V/m rather than $\pm 10^8$ V/m (Fig. 4.18).

It is worth emphasising that the results reported both here and in Ref. [8] rely on the assumption that the three-dimensional problem can be reduced to one-dimensional considerations. This is achieved by first averaging the charge density over the bins and then calculating the field (see Eqn (4.16) and Fig. 4.13). However, we lose some

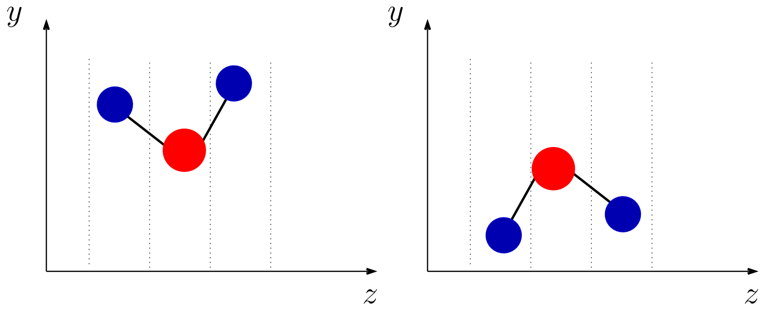

Figure 4.19: Illustration of a scenario where two molecules with different orientations yield the same one-dimensional charge density. The bins are represented by dotted lines.

information about the orientation of a molecule if we only consider the total charge contained in a bin. We illustrate this point in Fig. 4.19 with a simple example. In this scenario, two different orientations give exactly the same charge distribution. This might be a problem because the charge density is the only quantity that enters the calculation of the field.

In order to take the orientation into account, we could carry out additional simulations and calculate the three-dimensional field $\mathbf{E}(\mathbf{r}_{ijk}, \mathbf{r}^N(t))$ on a lattice indexed by (i, j, k), depending on the instantaneous positions $\mathbf{r}^N$ of all atoms. This can be done by convolution with the three-dimensional Green's function G_E or G_W, as explained in §2.2.1. The average field at each lattice site would then be given by

$$\mathbf{E}(\mathbf{r}_{ijk}) = \frac{1}{t} \int_0^t \mathrm{d}t' \; \mathbf{E}(\mathbf{r}_{ijk}, \mathbf{r}^N(t')). \tag{4.29}$$

This approach would probably require longer simulation times, but the validity of the reduction to one dimension could be directly probed by comparing to the previous results.

Chapter 5

Conclusion and future work

In this work, we carried out NEMD simulations of rigid water molecules in order to confirm the theoretically hypothesised and previously reported creation of an electric field arising from a temperature gradient. We found that the heat exchange algorithm, which was used in previous investigations, leads to a considerable decrease in the total system energy for the simulation times of interest. Based on analytical considerations, we were able to identify the major cause of the energy drift and proposed a new method, which we have termed the cHEX algorithm.

In contrast to previous reports, we found that the treatment of long-range electrostatic interactions does not strongly influence the electrostatic field if it is calculated in accordance with the boundary conditions employed. Our results suggest that the field was overestimated by about one order of magnitude in previous work. This was demonstrated to be a direct consequence of the incorrect expression used in the calculation of the field in previous work.

The validity of this approach is based on the assumption that the problem can be reduced to one-dimensional considerations, as was done in previous work. This simplification could potentially be problematic, because different orientations of the molecules can lead to the same one-dimensional charge density. We therefore suggest the field should first be calculated on a three-dimensional grid for each individual configuration and time averages calculated afterwards. This is a necessary test for consistency and we will address this point in future work.

In this work, we could establish the appropriate protocols which will allow us to better understand the thermo-polarisation effect on a molecular basis. This is of direct importance for studying the physical background of non-equilibrium processes involving strong temperature gradients and cross-phenomena.

Bibliography

[1] F Bresme, A Lervik, D Bedeaux and S Kjelstrup, 'Water polarization under thermal gradients', *Physical Review Letters*, **101**, 020602, 2008.

[2] S J Doktycz and K S Suslick, 'Interparticle collisions driven by ultrasound', *Science*, **247**, 1067–1069, 1990.

[3] J Armstrong, A Lervik and F Bresme, 'Enhancement of the thermal polarization of water via heat flux and dipole moment dynamic correlations', *Journal of Physical Chemistry B*, **117**, 14817–14826, 2013.

[4] F Bresme, 'Equilibrium and nonequilibrium molecular-dynamics simulations of the central force model of water', *Journal of Chemical Physics*, **115**, 7564–7574, 2001.

[5] J Muscatello and F Bresme, 'A comparison of Coulombic interaction methods in non-equilibrium studies of heat transfer in water', *Journal of Chemical Physics*, **135**, 234111, 2011.

[6] J Muscatello, F Römer, J Sala and F Bresme, 'Water under temperature gradients: Polarization effects and microscopic mechanisms of heat transfer', *Physical Chemistry Chemical Physics*, **13**, 19970–19978, 2011.

[7] F Römer, F Bresme, J Muscatello, D Bedeaux and J M Rubí, 'Thermomolecular orientation of nonpolar fluids', *Physical Review Letters*, **108**, 105901, 2012.

[8] J Armstrong and F Bresme, 'Water polarization induced by thermal gradients: The extended simple point charge model (SPC/E)', *Journal of Chemical Physics*, **139**, 014504, 2013.

[9] C Brennen, *Cavitation and bubble dynamics.* Oxford University Press, New York, 1995.

[10] R Pecha and B Gompf, 'Microimplosions: Cavitation collapse and shock wave emission on a nanosecond time scale', *Physical Review Letters*, **84**, 1328–1330, 2000.

[11] C Virone, H J M Kramer, G M van Rosmalen, A H Stoop and T W Bakker, 'Primary nucleation induced by ultrasonic cavitation', *Journal of Crystal Growth*, **294**, 9–15, 2006.

[12] A O Govorov, W Zhang, T Skeini, H H Richardson, J Lee and N A Kotov, 'Gold nanoparticle ensembles as heaters and actuators: Melting and collective plasmon resonances', *Nanoscale Research Letters*, **1**, 84–90, 2006.

[13] A O Govorov and H H Richardson, 'Generating heat with metal nanoparticles', *Nano Today*, **2**, 30–38, 2007.

[14] D Wolf, P Keblinski, S R Phillpot and J Eggebrecht, 'Exact method for the simulation of Coulombic systems by spherically truncated, pairwise r^{-1} summation', *Journal of Chemical Physics*, **110**, 8254–8282, 1999.

[15] T Ikeshoji and B Hafskjold, 'Non-equilibrium molecular dynamics calculation of heat conduction in liquid and through liquid-gas interface', *Molecular Physics*, **81**, 251–261, 1994.

[16] F Bresme, B Hafskjold and I Wold, 'Nonequilibrium molecular dynamics study of heat conduction in ionic systems', *Journal of Physical Chemistry*, **100**, 1879–1888, 1996.

[17] S R de Groot and P Mazur, *Non-equilibrium thermodynamics.* North-Holland Publishing Company, Amsterdam, first edition, 1962.

[18] E T Jaynes, 'The minimum entropy production principle', *Annual Review of Physical Chemistry*, **31**, 579–601, 1980.

[19] S Auer and D Frenkel, 'Numerical simulation of crystal nucleation in colloids', *Advances in Polymer Science*, **173**, 149–208, 2005.

[20] S Nosé, 'A unified formulation of the constant temperature molecular dynamics methods', *Journal of Chemical Physics*, **81**, 511–519, 1984.

[21] W G Hoover, 'Canonical dynamics: equilibrium phase-space distributions', *Physical Review A*, **31**, 1695–1697, 1985.

[22] H A Posch and W G Hoover, 'Nonequilibrium molecular dynamics of classical fluids', *Molecular Liquids: New Perspectives in Physics and Chemistry*, **379**, 527–547, 1992.

[23] D Macgowan and D J Evans, 'A comparison of NEMD algorithms for thermal conductivity', *Physics Letters A*, **117**, 414–416, 1986.

[24] D Frenkel and B Smit, *Understanding molecular simulation: From algorithms to applications.* Academic Press, San Diego, London, second edition, 2002.

[25] B L Holian and P S Lomdahl, 'Plasticity induced by shock waves in nonequilibrium molecular-dynamics simulations', *Science*, **280**, 2085–2088, 1998.

[26] P P Ewald, 'Die Berechnung optischer und elektrostatischer Gitterpotentiale', *Annalen der Physik*, **369**, 253–287, 1921.

[27] C J Fennell and J D Gezelter, 'Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics', *Journal of Chemical Physics*, **124**, 234104, 2006.

[28] M Neumann, *Computational Physics II: Simulation. Lecture notes.* University of Vienna, 2013.

[29] S Tyagi, 'Coulomb potentials in two and three dimensions under periodic boundary conditions', *Journal of Chemical Physics*, **122**, 14101, 2005.

[30] T Darden, D York and L Pedersen, 'Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems', *Journal of Chemical Physics*, **98**, 10089, 1993.

[31] H Andersen, 'RATTLE: A 'Velocity' version of the SHAKE algorithm for molecular dynamics calculations', *Journal of Computational Physics*, **52**, 24–34, 1983.

[32] S Plimpton, 'Fast parallel algorithms for short-range molecular dynamics', *Journal of Computational Physics*, **117**, 1–19, 1995.

[33] J-P Ryckaert, G Ciccotti and H J C Berendsen, 'Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes', *Journal of Computational Physics*, **23**, 327–341, 1977.

[34] D J Evans, 'Homogeneous NEMD algorithm for thermal conductivity–application of non-canonical linear response theory', *Physics Letters A*, **91**, 457–460, 1982.

[35] F Müller-Plathe, 'A simple nonequilibrium molecular dynamics method for calculating the thermal conductivity', *Journal of Chemical Physics*, **106**, 6082–6085, 1997.

[36] A Tenenbaum, G Ciccotti and R Gallico, 'Stationary nonequilibrium states by molecular dynamics. Fourier's law', *Physical Review A*, **25**, 2778–2787, 1982.

[37] S Kuang and J D Gezelter, 'A gentler approach to RNEMD: Nonisotropic velocity scaling for computing thermal conductivity and shear viscosity', *Journal of Chemical Physics*, **133**, 164101, 2010.

[38] S Aubry, D J Bammann, J J Hoyt, R E Jones, C J Kimmer, P A Klein, G J Wagner, E B Webb III and J A Zimmerman. 'A robust, coupled approach for atomistic-continuum simulation'. Technical report, Sandia National Laboratories, 2004.

[39] C Vega and J L F Abascal, 'Simulating water with rigid non-polarizable models: A general perspective', *Physical Chemistry Chemical Physics*, **13**, 19663–19688, 2011.

[40] V Molinero and E B Moore, 'Water modeled as an intermediate element between carbon and silicon', *Journal of Physical Chemistry B*, **113**, 4008–4016, 2009.

[41] J L F Abascal and C Vega, 'A general purpose model for the condensed phases of water: TIP4P/2005', *Journal of Chemical Physics*, **123**, 234505, 2005.

[42] H J C Berendsen, J R Grigera and T P Straatsma, 'The missing term in effective pair potentials', *Journal of Physical Chemistry*, **91**, 6269–6271, 1987.

[43] M R Reddy and M Berkowitz, 'The dielectric constant of SPC/E water', *Chemical Physics Letters*, **155**, 173–176, 1989.

[44] I-C Yeh and G Hummer, 'System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions', *Journal of Physical Chemistry B*, **108**, 15873–15879, 2004.

[45] M Neumann, 'The dielectric constant of water. Computer simulations with the MCY potential', *Journal of Chemical Physics*, **82**, 5663–5672, 1985.

[46] J R Lukes, X-G Liang, C-L Tien and D Y Li, 'Molecular dynamics study of solid thin-film thermal conductivity', *Journal of Heat Transfer*, **122**, 536–543, 2000.

Appendix A

Long-range interaction

A.1 Ewald summation

Let us for simplicity assume that we have an ionic system instead of rigid molecules with partial charges. Furthermore, we assume that the system is charge neutral. We then wish to show that

$$G_{\mathrm{E}}(\mathbf{r}) = \frac{\mathrm{erfc}(\eta r)}{r} - \frac{\pi}{\eta^2 V} + \frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi}{k^2} \mathrm{e}^{-\frac{k^2}{4\eta^2}} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}, \tag{A.1}$$

when inserted into

$$U^{\mathrm{lr}}(\mathbf{r}^N) = \frac{1}{2}\sum_{i\neq j} q_i q_j G_{\mathrm{E}}(\mathbf{r}_{ij}) + \frac{1}{2}\sum_i q_i^2 \lim_{r\to 0}\left[G_{\mathrm{E}}(\mathbf{r}) - \frac{1}{r}\right], \tag{A.2}$$

is equivalent to

$$\begin{aligned} U^{\mathrm{lr}}(\mathbf{r}^N) = {} & \frac{1}{2V}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{4\pi}{k^2} \left|\sum_i q_i \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}_i}\right|^2 \mathrm{e}^{-k^2/4\eta^2} - \frac{\eta}{\sqrt{\pi}}\sum_{i=1} q_i^2 \\ & + \frac{1}{2}\sum_{i\neq j} q_i q_j \frac{\mathrm{erfc}(\eta r_{ij})}{r_{ij}}, \end{aligned} \tag{A.3}$$

which is the expression given by Eqn (12.1.25) of Ref. [24] ($\eta \equiv \sqrt{\alpha}$).

For the entire proof, we will follow Ref. [28]. Let us first consider the limit $r \to 0$ in Eqn (A.2). The function $\mathrm{erfc}(x)$ is defined as

$$\mathrm{erfc}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{d}t\ \mathrm{e}^{-t^2}. \tag{A.4}$$

If we use a Taylor expansion, we find that

$$\lim_{x\to 0}\left(\frac{\operatorname{erfc}(\eta x)}{x}-\frac{1}{x}\right)=\lim_{x\to 0}\left(\frac{1-\frac{2}{\sqrt{\pi}}\eta x}{x}-\frac{1}{x}+\mathcal{O}(x^2)\right)=-\frac{2\eta}{\sqrt{\pi}} \tag{A.5}$$

and we can therefore evaluate the expression

$$\lim_{r\to 0}\left[G(\mathbf{r})-\frac{1}{r}\right]=-\frac{2\eta}{\sqrt{\pi}}-\frac{\pi}{\eta^2 V}+\frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{4\pi}{k^2}\mathrm{e}^{-\frac{k^2}{4\eta^2}}. \tag{A.6}$$

We then consider the summation over i and j in Eqn A.2 and first collect all k-space contributions, which gives

$$\begin{aligned}&\frac{1}{2}\sum_{i\neq j}q_i q_j\frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{4\pi}{k^2}\mathrm{e}^{-\frac{k^2}{4\eta^2}}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}_{ij}}+\frac{1}{2}\sum_i q_i^2\frac{1}{V}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{4\pi}{k^2}\mathrm{e}^{-\frac{k^2}{4\eta^2}}\\&=\frac{1}{2V}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{4\pi}{k^2}\mathrm{e}^{-\frac{k^2}{4\eta^2}}\left|\sum_i q_i\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}_i}\right|^2.\end{aligned} \tag{A.7}$$

Furthermore, the sums involving the term $\pi/\eta^2 V$ cancel out, since

$$-\frac{1}{2}\sum_{i\neq j}q_i q_j\frac{\pi}{\eta^2 V}-\frac{1}{2}\sum_i q_i^2\frac{\pi}{\eta^2 V}=-\frac{\pi}{2\eta^2 V}\left(\sum_i q_i\right)^2=0 \tag{A.8}$$

holds, because we assumed charge neutrality. The remaining terms finally yield

$$\begin{aligned}U^{\mathrm{lr}}(\mathbf{r}^N)=&\frac{1}{2V}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{4\pi}{k^2}\left|\sum_i q_i\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}_i}\right|^2\mathrm{e}^{-k^2/4\eta^2}-\frac{\eta}{\sqrt{\pi}}\sum_{i=1}q_i^2\\&+\frac{1}{2}\sum_{i\neq j}q_i q_j\frac{\operatorname{erfc}(\eta r_{ij})}{r_{ij}},\end{aligned} \tag{A.9}$$

which is what we wanted to show.

A.2 Wolf summation

With the derivations in the previous section it follows immediately that

$$\begin{aligned}U^{\mathrm{lr}}(\mathbf{r}^N)&=\frac{1}{2}\sum_{i\neq j}q_i q_j G_{\mathrm{W}}(\mathbf{r}_{ij})+\frac{1}{2}\sum_i q_i^2\lim_{r\to 0}\left[G_{\mathrm{W}}(\mathbf{r})-\frac{1}{r}\right]\\&=\frac{1}{2}\sum_{i\neq j}q_i q_j\left(\frac{\operatorname{erfc}(\zeta r_{ij})}{r_{ij}}-\frac{\operatorname{erfc}(\zeta r_{\mathrm{c}})}{r_{\mathrm{c}}}\right)\\&-\left(\frac{\operatorname{erfc}(\zeta r_{\mathrm{c}})}{2r_{\mathrm{c}}}+\frac{\zeta}{\sqrt{\pi}}\right)\sum_i q_i^2.\end{aligned} \tag{A.10}$$

This is exactly the expression given in Eqn (5.13) of Ref. [14], if we make the identification $\zeta\equiv\alpha$ and only consider those terms in the double summation for which $r_{ij}\leq r_{\mathrm{c}}$.

Appendix B

Heat exchange algorithm

B.1 Formulation

We wish to show that the velocity update

$$\mathbf{v}_i \mapsto \bar{\mathbf{v}}_i = \xi \mathbf{v}_i + (1-\xi)\mathbf{v}_\Gamma \quad \forall i \in \gamma, \tag{B.1}$$

where $\xi = \sqrt{\eta}$, satisfies

$$\bar{\mathcal{K}}_\Gamma^{\mathrm{nt}} = \mathcal{K}_\Gamma^{\mathrm{nt}} + \Delta Q = \eta\, \mathcal{K}_\Gamma^{\mathrm{nt}} \tag{B.2}$$

and

$$\bar{\mathbf{v}}_\Gamma = \mathbf{v}_\Gamma, \tag{B.3}$$

where all the terms have been defined in §3.1.1. First, we show that the centre of mass velocity is conserved. That is easy to see, since

$$\begin{aligned} \bar{\mathbf{v}}_\Gamma &= \frac{1}{m_\Gamma} \sum_{i\in\gamma} m_i \bar{\mathbf{v}}_i = \frac{1}{m_\Gamma} \sum_{i\in\gamma} m_i \Big(\xi \mathbf{v}_i + (1-\xi)\mathbf{v}_\Gamma \Big) && \text{(B.4)} \\ &= \xi \mathbf{v}_\Gamma + (1-\xi)\mathbf{v}_\Gamma = \mathbf{v}_\Gamma. && \text{(B.5)} \end{aligned}$$

Therefore, the translational kinetic energy remains unaffected, namely

$$\bar{\mathcal{K}}_\Gamma^{\mathrm{t}} = \mathcal{K}_\Gamma^{\mathrm{t}}. \tag{B.6}$$

Physical quantity	Conversion
energy (e)	$e^* = e\epsilon^{-1}$
distance (d)	$d^* = d\sigma^{-1}$
time (t)	$t^* = t\sqrt{\frac{m\sigma^2}{\epsilon}}$
density (ρ)	$\rho^* = \rho\sigma^3$
temperature (T)	$T^* = Tk_{\mathrm{B}}\epsilon^{-1}$

Table B.1: The 'lj' reduced units system in LAMMPS. All quantities are defined in terms of the basic units of mass (m), energy (ϵ) and distance (σ).

The total kinetic energy in the region Γ after the velocity update step is given by

$$\begin{aligned}
\bar{\mathcal{K}}_\Gamma &= \sum_{i\in\gamma} \frac{m_i \bar{v}_i^2}{2} = \sum_{i\in\gamma} \frac{m_i}{2}\Big(\xi\mathbf{v}_i + (1-\xi)\mathbf{v}_\Gamma\Big)^2 \\
&= \sum_{i\in\gamma} \frac{m_i}{2}\Big(\eta v_i^2 - 2\xi(\xi-1)\mathbf{v}_\Gamma\cdot\mathbf{v}_i + (\xi-1)^2 {v_\Gamma}^2\Big) \\
&= \eta\mathcal{K}_\Gamma - \xi(\xi-1)m_\Gamma\mathbf{v}_\Gamma\cdot\mathbf{v}_\Gamma + (\xi-1)^2\frac{m_\Gamma {v_\Gamma}^2}{2} \\
&= \eta\mathcal{K}_\Gamma + \mathcal{K}_\Gamma^{\mathrm{t}}\big((\xi-1)^2 - 2\xi(\xi-1)\big) \\
&= \eta\mathcal{K}_\Gamma + \mathcal{K}_\Gamma^{\mathrm{t}}\big(\eta - 2\xi + 1 - 2\eta + 2\xi\big) \\
&= \eta\mathcal{K}_\Gamma + \mathcal{K}_\Gamma^{\mathrm{t}}\big(1-\eta\big) \\
&= \eta\big(\mathcal{K}_\Gamma^{\mathrm{t}} + \mathcal{K}_\Gamma^{\mathrm{nt}}\big) + \mathcal{K}_\Gamma^{\mathrm{t}}\big(1-\eta\big) \\
&= \eta\mathcal{K}_\Gamma^{\mathrm{nt}} + \mathcal{K}_\Gamma^{\mathrm{t}}.
\end{aligned} \tag{B.7}$$

With the definition of η given by Eqn (3.7), we can therefore easily see that

$$\begin{aligned}
\bar{\mathcal{K}}_\Gamma^{\mathrm{nt}} = \eta\mathcal{K}_\Gamma^{\mathrm{nt}} &= \left(\frac{\mathcal{K}_\Gamma^{\mathrm{nt}} + \Delta Q}{\mathcal{K}_\Gamma^{\mathrm{nt}}}\right)\mathcal{K}_\Gamma^{\mathrm{nt}} \\
&= \mathcal{K}_\Gamma^{\mathrm{nt}} + \Delta Q,
\end{aligned} \tag{B.8}$$

which proves the assertion (Eqn (B.2)).

B.2 Lennard-Jones NEMD simulation

The simulation software LAMMPS can interpret user-defined input scripts. A script is a textfile containing all the instructions that should be carried out sequentially. The NEMD simulation with the settings $\kappa = 300$ and a timestep of $\Delta t = 0.001$ can be fully reproduced with the example input script provided below. The conversion to reduced units is given in Tab. B.1 and all the relevant input parameters are summarised in Tab. B.2.

Parameter	Value
N	1728
ρ	0.8442
T_{eq}	≈ 2.3
κ	300
Δt	0.001

Table B.2: Relevant input parameters for a typical Lennard-Jones NEMD simulation.

```
# Setup initial lattice structure
#################################

units           lj
atom_style      atomic

# density = 0.8842
lattice         fcc 0.8442
region          box block 0 6 0 6 0 12
create_box      1 box
create_atoms    1 box
mass            1 1.0
velocity        all create 3.0 87287

# Define interaction potential
# (LJ, shifted, r_sr = 4)
##############################

pair_style      lj/cut 4
pair_coeff      1 1 1.0 1.0 4
neighbor        0.3 bin
neigh_modify    every 20 delay 0 check no
pair_modify     shift yes

# Melt initial configuration
############################
variable        Nwarmup equal 10000
variable        dt      equal 0.001

# temperature = 2.3
variable        T       equal 2.3

# use Langevin thermostat
fix             fNVE_melt  all  nve
fix             fLangevin all langevin ${T} ${T} 10.0 1000101
timestep        ${dt}
thermo          1000
thermo_style    custom step temp ke pe etotal press
run             ${Nwarmup}
unfix           fLangevin
unfix           fNVE_melt

# Carry out NVE equilibration run
#################################
variable       Nequi equal 100000
reset_timestep 0
fix            fNVE_equi all nve
run            ${Nequi}
unfix          fNVE_equi
```

```
# NEMD simulation
#################

reset_timestep     0

# define variables
variable N         equal 1000000

# box dimension in z-direction and resolution
variable Lz        equal zhi-zlo
variable delta     equal ${Lz}*0.2
variable dz        equal ${Lz}/100

# reservoir extents in z-direction
variable zlo_Thi   equal zlo
variable zhi_Thi   equal ${zlo_Thi}+${delta}
variable zlo_Tlo   equal ${Lz}/2.
variable zhi_Tlo   equal ${zlo_Tlo}+${delta}

# sampling and averaging frequencies
variable Nsamp     equal 10
variable Nevery    equal 1000
variable Nrepeat   equal ${Nevery}/${Nsamp}

# use NVE integration
fix    NVE  all  nve

# define regions
region Thi_reg   block 0 INF 0 INF ${zlo_Thi} ${zhi_Thi}
region Tlo_reg   block 0 INF 0 INF ${zlo_Tlo} ${zhi_Tlo}
region Tgrad_reg block 0 INF 0 INF ${zlo_Thi} ${zlo_Tlo}

# compute individual temperature
compute    cTlo    all temp/region Tlo_reg
compute    cThi    all temp/region Thi_reg
compute    cTgrad  all temp/region Tgrad_reg

# define fixes that add/remove heat during every timestep
fix        fHi     all heat 1 +300. region Thi_reg
fix        fLo     all heat 1 -300. region Tlo_reg

# calculate energies for spatial and temporal averaging
variable   vN      equal count(all)
compute    ke      all ke/atom
compute    cT      all temp
compute    cP      all pressure thermo_temp
compute    cPe     all pe
compute    cKe     all ke
variable   vpe     equal c_cPe/v_vN
variable   vke     equal c_cKe/v_vN
variable   ve      equal v_vke+v_vpe

# write time averages to the file tavg.dat
fix  fTavg  all ave/time    ${Nsamp} ${Nrepeat} ${Nevery} c_cT c_cTlo c_cTgrad
    c_cThi c_cP v_ve v_vpe v_vke file tavg.dat

# write spatial averages to the file savg.dat
fix  fSavg  all ave/spatial ${Nsamp} ${Nrepeat} ${Nevery} z lower ${dz} c_ke file
    savg.dat units box

# define timestep
timestep           ${dt}
```

```
# write to console every 1000 steps
thermo_style        custom step c_cT c_cTlo c_cTgrad c_cThi c_cP v_ve v_vpe v_vke

# carry out N simulation steps
run                 ${N}
```

Listing B.1: Full LAMMPS (1Feb14) input script for a Lennard-Jones NEMD simulation.

Appendix C

Electrostatic field

C.1 Wolf summation

The modified Green's function for Wolf summation is given by

$$\tilde{G}_{\mathrm{W}}(z) = 2\pi \int\limits_0^{r_{\mathrm{c}}} s\mathrm{d}s \, \left[\frac{\operatorname{erfc}(\zeta\sqrt{s^2+z^2})}{\sqrt{s^2+z^2}} - \frac{\operatorname{erfc}(\zeta r_{\mathrm{c}})}{r_{\mathrm{c}}} \right], \tag{C.1}$$

where $r^2 = x^2 + y^2 + z^2 = s^2 + z^2$. We will only show how to evaluate the first integral, because the second one is trivial. Let us therefore consider the expression

$$I = \int\limits_0^{r_{\mathrm{c}}} \mathrm{d}s \; s \frac{\operatorname{erfc}(\zeta\sqrt{s^2+z^2})}{\sqrt{s^2+z^2}}. \tag{C.2}$$

With the definition $\tau(s) = \sqrt{s^2+z^2}$, the integral can be rewritten as

$$I = \int\limits_{|z|}^{\tau(r_{\mathrm{c}})} \mathrm{d}\tau \; \operatorname{erfc}(\zeta\tau). \tag{C.3}$$

The error function is defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int\limits_0^x \mathrm{d}t \; \mathrm{e}^{-t^2}, \tag{C.4}$$

and the function is related to the complementary error function through $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$. Furthermore, the identity

$$\int \operatorname{erf}(z)\mathrm{d}z = z \operatorname{erf}(z) + \frac{\mathrm{e}^{-z^2}}{\sqrt{\pi}} \tag{C.5}$$

holds. With these identities it is straightforward to show that

$$\begin{aligned}\int_c^d \operatorname{erfc}(ax)\mathrm{d}x &= \int_c^d (1-\operatorname{erf}(ax))\,\mathrm{d}x\\ &= (d-c) - \frac{1}{a}\left[x\operatorname{erf}(x) + \frac{\mathrm{e}^{-x^2}}{\sqrt{\pi}}\right]_{ac}^{ad}\\ &= d\ \operatorname{erfc}(ad) - c\ \operatorname{erfc}(ac) + \frac{1}{\sqrt{\pi}a}\left[\mathrm{e}^{-a^2c^2} - \mathrm{e}^{-a^2d^2}\right]. \end{aligned} \tag{C.6}$$

If we make the identifications $c = |z|$, $d = \tau(r_c)$ and $a = \zeta$, the solution of Eqn (C.2) is given by

$$I = \tau(r_\mathrm{c})\operatorname{erfc}(\zeta\tau(r_\mathrm{c})) - |z|\operatorname{erfc}(\zeta|z|) + \frac{1}{\sqrt{\pi}\zeta}\left[\mathrm{e}^{-\zeta^2z^2} - \mathrm{e}^{-\zeta^2\tau(r_\mathrm{c})^2}\right] \tag{C.7}$$

and the expression for $\bar{G}_\mathrm{W}$ in Eqn (4.25) follows, since the integration of the second term is trivial.

Peter Wirnsberger, BSc MPhil

Curriculum vitae

EDUCATION

since Oct 2012	**Master of Science (MSc) in Physics**, Computational Physics Group, Faculty of Physics, University of Vienna, Austria Master's thesis: A simulation study on the thermo-polarisation effect in water, Supervisors: Prof. C. Dellago (University of Vienna) and Prof. D. Frenkel (University of Cambridge)
2011-2012	**Master of Philosophy (MPhil) on Scientific Computing** *(distinction/high pass)*, Centre for Scientific Computing, Cavendish Laboratory, University of Cambridge, United Kingdom Master's thesis: The grid orientation effect in miscible displacement, Supervisors: Dr N. Nikiforakis (University of Cambridge) and Dr F. Monmont (SCR)
2008-2011	**Bachelor of Science (BSc) in Physics** *(distinction)*, Faculty of Physics, University of Vienna, Austria Bachelor's thesis: Modelling atmospheric flows in Cartesian and spherical geometry, Supervisors: Prof. C. Dellago (University of Vienna) and Dr N. Nikiforakis (University of Cambridge)
2002-2007	**HTBLuVA Villach (EDVO)**; secondary school specialised on software engineering *(distinction)*

INTERNSHIPS, WORKING EXPERIENCE AND ACADEMIES

Oct 2013-Apr 2014	ERASMUS internship in the group of Prof. D. Frenkel, University of Cambridge
Feb 2013	VICOM-Winterschool on phase transitions and phase stability
Sep 2011	High Performance Computing (HPC) Autumn Academy, University of Cambridge
Jul-Sep 2010	Internship in the group of Dr N. Nikiforakis, Cavendish Laboratory, University of Cambridge Development of numerical models for earth system modelling
Jul-Aug 2009	Internship in the group of Prof. C. Dellago, Faculty of Physics, University of Vienna Dielectric properties of one-dimensional water chains confined in nano-pores
Jul-Aug 2008	Summer job at Infineon Technologies Austria, Production and Manufacturing
Jul-Aug 2006	Project in the field of Electronic and Equipment Engineering, Fachhochschule Kärnten
Jul-Aug 2005	Internship at Deutsche Bank, IT Centre Eschborn

TEACHING EXPERIENCE

Mar-Jun 2013	Teaching assistant (Scientific Computing), University of Vienna
Oct-Jan 2012	Teaching assistant (Introduction to Mathematics for Physicists), University of Vienna
Oct-Jan 2010	Teaching assistant (Introduction to Mathematics for Physicists), University of Vienna

AWARDS AND FELLOWSHIPS

Doctoral Fellowship (DOC) of the Austrian Academy of Sciences, 2014
Excellence Grant of the Federation of Austrian Industry (IV) Carinthia, 2013
Schlumberger Departmental Award, 2011

LANGUAGES

German (native speaker), English, Spanish (intermediate), French (beginner)