



universität
wien

MASTERARBEIT

Titel der Masterarbeit

„Melting and Freezing of Water Clusters“

Verfasser

Raimund Hirmer, B.Sc.

angestrebter akademischer Grad

Master of Science (M.Sc.)

Wien, 2013

Studienkennzahl lt. Studienblatt: A 066876

Studienrichtung lt. Studienblatt: Physik

Betreuerin / Betreuer: Univ.-Prof. Mag. Dr. Christoph Dellago

Contents

1. Introduction	1
2. Nucleation Theory	4
2.1. Classical Nucleation Theory	5
2.2. Capillarity Approximation	8
2.3. Nucleation Rate in Condensed Systems	10
3. Computational Methods	12
3.1. Molecular Dynamics	12
3.1.1. Potential	12
3.1.2. Integration of the equations of motion	13
3.1.3. Molecular Dynamics in the Canonical Ensemble	15
3.1.4. Constraints	15
3.2. Monte Carlo Methods	16
3.2.1. Metropolis Monte Carlo	16
3.2.2. Hybrid Monte Carlo	18
3.3. Umbrella Sampling	18
3.4. Replica Exchange	21
3.5. Weighted Histogram Analysis Method	22
4. Model Definition	25
5. Ice Configurations	29
5.1. Ice Ih	29
5.2. Ice Ic	30
5.3. Algorithm for Ice Generation	31
6. Determination of the Order Parameter	32
6.1. Size of the Largest Cluster	32
6.2. Structural Order Parameters	33
6.3. The Solid Cluster Criterion	35
7. Identifying of the Melting Point	42
7.1. Methods	42
7.2. Caloric Curve and Largest Crystalline Cluster	44

8. Calculation of the Free Energy Barrier	52
8.1. Details of Umbrella Sampling with 216 molecules	52
8.2. Results	52
8.3. Problems and Proposals for Solution	54
8.4. Details of Umbrella Sampling with 432 molecules	59
8.5. Results and Discussion	60
9. Conclusion	65
A. Free Energy Barrier:	
Derivation via the Partition Function	66

List of Figures

2.1. Minimal Work of nucleus formation	10
3.1. Evaluation of free energy in umbrella sampling	21
3.2. Successful swapping attempts with replica exchange algorithm	23
4.1. Phase diagram of the TIP4P ₂₀₀₅ model	28
4.2. Phase diagram of the TIP3P model	28
5.1. Unit cell of ice Ih	30
5.2. Unit cell of ice Ic	31
6.1. Probability distributions of $a_3(k, j)$, bulk	36
6.2. Probability distributions of $a_3(k, j)$, finite system	38
6.3. Composition of a 216 molecules cluster	39
6.4. Snapshot of an ice Ic droplet	40
6.5. Time evolution of the order parameter at 130 K and 200 K	41
7.1. Melting temperatures in dependence of system size	43
7.2. Time evolution of the order parameter. Ice Ic.	46
7.3. Time evolution of the order parameter. Ice Ih.	46
7.4. Potential energy per molecule and largest crystalline cluster of an original ice Ic droplet as function of temperature	47
7.5. Potential energy per molecule and largest crystalline cluster of an original ice Ih droplet as function of temperature	48
7.6. Time evolution of the order parameter at selected temperatures for a starting ice Ih configuration with 432 molecules	50
7.7. Potential energy per molecule and largest crystalline cluster of an original ice Ih droplet with 432 molecules as function of temperature	51
8.1. Free energy curves for four high temperatures.	53
8.2. Free energy curves for four low temperatures.	54
8.3. Probability distribution of the order parameter at 180 K.	55
8.4. Replica exchange moves at different temperatures.	57
8.5. Possible free energy landscape.	58
8.6. Free energy in each window at all studied temperatures.	58
8.7. Biased probability distribution and free energy at 150 K.	61
8.8. Energy landscape for different temperatures with 432 molecules.	61
8.9. Determination of the phase equilibrium.	64

List of Tables

- 4.1. Parameters of SPC/E, TIP4P and TIP4P₂₀₀₅ model 27
- 4.2. Dipole and quadrupole moment of SPC/E, TIP4P and TIP4P₂₀₀₅ 27
- 8.1. Selfdiffusion coefficients at different temperatures. 56
- 8.2. Location n^* and height ΔG of the free energy barrier. Probability P for the solid state. 60

1. Introduction

Water plays a central role in our life. A lot of organisms consist largely of water and the earth itself exhibits immense water reservoirs. Therefore it is reasonable that water is a object of research. But many properties of water are still not fully understood, although this molecule might seem at first sight very simple. Water has one of the most complex phase diagrams: In comparison to other liquids water exhibits several anomalies in the liquid state. It can crystallize into 15 different phases and even amorphous solid phases are known. Even more difficult situations arise, if one studies small and smallest molecule accumulations instead of bulk systems. While bulk water can be studied relatively easily and many theories focus on that, more attention must be payed to the analysis of such micro- and nanosystems. Chemical reactions at the interface of ocean-spray aerosols underline the relevance of water nanoclusters.¹ It is particularly necessary, too, because it seems that small waterclusters play an important role in our atmosphere and in interstellar media.^{2,3} At mesospheric conditions (temperatures about 120 K) crystalline phases of water such as hexagonal ice Ih or cubic ice Ic can form beside amorphous water without crystalline order.² It is an interesting and timely question, which phase is the stable one for nanoclusters and knowledge about mechanisms of nucleation and growth of nanoparticles is of importance, especially what influence the hydrogen-bonds of water have in the core of the cluster as well as on its surface. Ostwald's step rule is often assumed to be valid for waterclusters under these conditions.^{2,4} It states, that the thermodynamically stable phase does not necessarily form directly from an unstable state. Instead initial formation of metastable phases takes place, because these offer a lower free energy barrier to the original phase and resemble it more closely.

The direct approach via experiments often turns out to be challenging. Especially the determination of the actual size of the generated clusters is a non-trivial problem. Nevertheless substantial efforts are made both in theory and in experiment, covering research on the whole range of sizes of nanoclusters. In this respect structural evolution as a function of cluster size is of particular interest.⁵ There is a series of theoretical and experimental studies, which focus on cluster sizes of 2 to 10 water molecules. They look for global minima in the potential energy and analyze melting behavior employing electronic structure techniques and a variety of analytical potentials.^{6,7} For instance, for water clusters consisting of 6 molecules several two- or three-dimensional structures are predicted to have almost the same low energy. In fact, in experiments a cage-form⁷, a ring-form⁸ and a book-form⁹ have been detected until today. Considering only such a small number of molecules one is not able to speak about liquid

or crystalline phases. A phase is crystalline if there are periodic structures of atoms in the three-dimensional space and for that a significantly larger number of molecules is needed. For example, water ices Ih and Ic consist of an almost perfect tetrahedral network, in which each oxygen is coordinated by four hydrogen-bonds. But even if clusters consist of tens to hundreds molecules, it emerges that these clusters do not necessarily crystallize at low temperatures, but transform into amorphous ice phases.⁵ Reason for that are the surface effects of clusters: in case of a crystalline cluster (supposed it has a spherical-like shape rather than a perfect crystal geometry), water molecules lying on the surface have unbonded hydrogens, which strongly contribute to the energy because of electrostatic repulsion.⁵ In turn they arrange themselves to build more hydrogen-bonds, which is at the cost of the crystal geometry. Below a critical molecule number hence a crystalline water cluster transforms completely into an amorphous structure. But even at larger system sizes the amorphous character on the surface of the clusters can still be observed.⁵

Experimental studies with electron diffraction showed that the onset of crystallization is at cluster sizes of about 200-1000 molecules.¹⁰ This quite inaccurate statement underlines again the difficulty of generating clusters of certain sizes in experiment and to measure them. Two distinct procedures have proved themselves to generate different cluster sizes: generation by expansion beams¹¹ or by collisional cooling in cluster cells^{12,5}. With different experimental set-ups (pressure, temperature,...) the cluster size distribution can be controlled. The actual size of one cluster can then be estimated reactively by, for example, the following methods: an additional sodium atom can be attached, which allows to photoionize the cluster and to separate its mass subsequently with mass spectroscopy.¹³ Another technique is to cover the cluster with a very thin layer of CF_4 and to measure the infrared band intensities for that adsorbed monolayer.¹² IR-spectroscopy can then be applied to measure the OH-stretch of the cluster and in this way the quality of hydrogen-bonds: crystalline water exhibits a lower lying absorption peak in the spectrum than amorphous phases.¹³ Using this technique, Pradzynski and co-workers were able to identify crystallization at the core of clusters starting from sizes of about 275 particles. The degree of crystallization grows with increasing cluster sizes. Furthermore, electron diffraction patterns give rise to the assumption that cubic ice Ic is present within the cluster rather than hexagonal ice Ih, the thermodynamically stable state for bulk water at these conditions.¹⁰

In this work, I study some aspects of ice nucleation using computer simulations. The advantage of computer simulations is the easy feasibility of certain cluster sizes as well as the analy-

sis of molecular structures. The disadvantage is the oversimplification of real water molecules and their interactions and, of course, the limited simulation time. In the first chapter the theoretical background of nucleation is discussed, which is known as Classical Nucleation Theory. Originally designed for the description of droplet condensation from supercooled vapors, it can be applied with some extensions for liquid-solid interfaces, too. Because nucleation is an activated process, there is a free energy barrier separating the liquid and solid phases. Usually this barrier can be derived by the so-called Capillarity Approximation of the nucleation theory. After explaining the computational methods and some statistical analysis methods, a proper water model is chosen. This model called TIP4P₂₀₀₅, describes the water molecule as rigid body with four interaction sites. Some aspects of ice Ih and ice Ic are discussed in the subsequent section. The practical part of this work firstly consists of elaboration of proper order parameters, which can distinguish between solid and liquid molecules. The chosen order parameter goes back to ten Wolde and Frenkel, who use spherical harmonics to determine the largest crystalline cluster within the droplet. To study the stability of ice phases in water clusters consisting of 216 and 432 molecules, I run constant temperature molecular dynamic simulations in order to estimate the temperature range, at which crystalline structures begin to melt. Furthermore, I use the umbrella sampling technique to investigate the free energy landscape as function of the order parameter, especially to determine the free energy barrier, which is then compared to the barrier predicted by the nucleation theory.

2. Nucleation Theory

Most liquids and vapors can be undercooled and superheated, respectively. If that happens, the liquid or vapor will become metastable. If the temperature is decreased, or in the case of superheating increased, furthermore, the system will become unstable. Metastable and unstable liquids relax into the stable equilibrium by different mechanisms of new phase formation. On the one hand unstable liquids reach equilibrium by the so-called spinodal decomposition.¹⁴ In this regime there is no free energy barrier the system must overcome, the transition occurs spontaneously. This happens by density fluctuations of small amplitude, but at least of a critical wavelength. Since this work does not deal with unstable liquids spinodal decomposition is not discussed here in detail. So on the other hand the transition of metastable liquids towards stability is of more interest. Here the mechanism is called nucleation. Small embryos of the stable phase are formed within the system. While smaller embryos dissolve back into the metastable phase, larger ones, i.e. those which reach a critical size and are therefore called critical nuclei or critical clusters, grow spontaneously towards the stable phase. So also this happens by density fluctuations, but different to spinodal decomposition, here the system relax by the activated growth of localized fluctuations.¹⁴ This is an activated process and the system must overcome a free energy barrier in order to form embryos of sufficient size. For this reason superheating and undercooling can occur, because in order to form a critical embryo work must be performed. One can differentiate furthermore: If the system has suspended and dissolved impurities, or imperfectly wetted solid boundaries, heterogeneous nucleation can happen.¹⁴ This means, that these impurities and surfaces provide preferential sites for nucleation and embryos will be formed there. If no such nucleation agents exist, homogeneous nucleation will become the fundamental mechanism.

The goal of nucleation theories is to quantify the net rate at which embryos grow to a critical size. This knowledge is important for any investigation of metastable liquids, because such systems can only be studied in time scales, which are small compared to the typical time scale of nucleation.¹⁴

In Sections 2.1 and 2.2, I will first introduce the Classical Nucleation Theory (CNT) in combination with the Capillarity Approximation. In order to calculate the nucleation rate, here it is necessary to compute the free energy of embryo formation, which is written as a function of the embryo radius. This theory makes some assumptions, especially the embryo of the new phase having the same properties as a bulk system. These estimations cause in some experimental cases the breakdown of CNT.¹⁵ Critical reconsideration leads to more accurate approaches like the Kinetic Nucleation Theory.¹⁶ Originally CNT was designed for the de-

scription of droplet condensation from supercooled vapors. With some extensions it can also be applied for liquid-solid interfaces.¹⁷ In Sec. 2.3 I will discuss some changes in that case. A more microscopic consideration would be the Classical Density Functional Theory. Instead of the radius, the free energy in this theory is written as a functional of the density profile, whose minima determine the thermodynamically stable states. For this approach I refer to the literature.¹⁸

2.1. Classical Nucleation Theory

First, I consider condensation of a liquid droplet from the metastable vapor. As mentioned before, classical nucleation theory deals with the determination of the net rate of forming a critical cluster. To do this, one considers the rate $J(n)$, at which a $(n - 1)$ -cluster becomes a n -cluster. This rate is given by the difference between the rate, at which a cluster containing $n - 1$ molecules gets a new single molecule by condensation and the rate, at which this n -cluster is destroyed by losing a single molecule to become a $(n - 1)$ -cluster,

$$J(n) = f(n - 1)k_{n-1}^+ A(n - 1) - f(n)k_n^- A(n). \quad (2.1)$$

Here $f(n - 1)$ and $f(n)$ are the concentrations of $(n - 1)$ - and n -clusters in a system of a certain volume V and $A(n)$ and $A(n - 1)$ are their respective surface areas. k_{n-1}^+ is the flux per unit time and area of single molecules arriving at a $(n - 1)$ -cluster, while k_n^- is the flux of single molecules leaving a n -cluster. Though it is easy to get k_{n-1}^+ from the kinetic theory of gases, it is hard to find an independent way to calculate k_n^- . To circumvent this problem, CNT makes an important assumption: the so-called Constrained Equilibrium Hypothesis says that an equilibrium distribution of clusters f_{eq} can be established and because of microscopic reversibility one can write

$$f_{eq}(n - 1)k^+ A(n - 1) = f_{eq}(n)k^- A(n), \quad (2.2)$$

where it is assumed, that k^- and k^+ are independent of the embryo size. Because now one can solve Equ. (2.2) for k^- and substitute it into Equ. (2.1), one gets

$$J(n) = f_{eq}(n - 1)k^+ A(n - 1) \left[\frac{f(n - 1)}{f_{eq}(n - 1)} - \frac{f(n)}{f_{eq}(n)} \right]. \quad (2.3)$$

The constrained equilibrium hypothesis, which affects the calculation of k^- , transforms the CNT into a thermodynamic theory, though it should be a kinetic one. This approximation is of course questionable, because nucleation is a non-equilibrium process inherently. As mentioned before this is a point, which is reflected by kinetic theories.

Equation (2.1) leads to an expression for the time dependence of the embryo distribution:

$$\frac{\partial f(n,t)}{\partial t} = J(n) - J(n+1). \quad (2.4)$$

The next assumption CNT makes is that the population of n -clusters does not depend on time, i.e. that “steady states” exist. In other words this means that the lag time, the required time for reaching the steady state, is of a smaller magnitude than the time of measurement. Analytical expressions for the lag time were determined for example by Shneidman and Weinberg.¹⁹ So if $\partial f/\partial t = 0$, J becomes independent of n and after rearranging Equ. (2.3) one gets

$$\frac{J}{k^+ f_{eq}(n-1)A(n-1)} = \frac{f(n-1)}{f_{eq}(n-1)} - \frac{f(n)}{f_{eq}(n)}. \quad (2.5)$$

At this point the flux J is termed as the nucleation rate. Now summing from $n = 2$ to $n = \Omega$, where Ω is a large but undefined number, yields the expression

$$J = \frac{\frac{f(1)}{f_{eq}(1)} - \frac{f(\Omega+1)}{f_{eq}(\Omega+1)}}{\sum_{n=1}^{\Omega} \frac{1}{k^+ f_{eq}(n)A(n)}} = \frac{1}{\sum_{n=1}^{\Omega} \frac{1}{k^+ f_{eq}(n)A(n)}}. \quad (2.6)$$

To derive the second equality it is used that the metastable phase consists mostly of single monomers, so the actual embryo distribution $f(1)$ and the equilibrium distribution $f_{eq}(1)$ should be equal. Besides, for a large Ω the actual distribution $f(\Omega)$ should vanish, because otherwise the new phase already would be present. Equ. (2.6) shows that the nucleation rate depends on the equilibrium embryo distribution. It follows from the thermodynamics of fluctuations that this quantity is proportional to the exponential of the free energy:²⁰

$$f_{eq}(n) \propto \exp\left[-\frac{\Delta G(n)}{k_B T}\right]. \quad (2.7)$$

In the literature the quantity N_n is often referred as the total number of embryos consisting of n particles, instead of the concentration $f_{eq}(n)$, i.e. $N_n/V = f_{eq}(n)$. Since there is no work for forming a single molecule embryo, the proportional constant should be the total number N of

particles in the metastable system¹⁴:

$$N_n = N \exp \left[-\frac{\Delta G(n)}{k_B T} \right]. \quad (2.8)$$

A fundamental rather than heuristic derivation of this equation is given in the appendix A of this work. There N_n is derived starting with the partition function of the canonical ensemble. Once we have determined ΔG (see the following section), we can substitute the equilibrium embryo distribution in Equ. 2.6. Besides it is common to exchange the sum for an integral, where the lower boundary is far below the critical size n^* of the nucleus and the upper boundary is far above it.

$$J = k^+ \frac{N}{V} \left[\int_{n \ll n^*}^{n \gg n^*} \exp \left(\frac{\Delta G(n)}{k_B T} \right) \frac{1}{A(n)} dn \right]^{-1} \quad (2.9)$$

Since the exponential in this equation is sharply peaked at n^* as we will also see in the next section, we can expand ΔG up to the second order, in which the first order vanishes because of the maximum condition, i.e. $\Delta G'(n^*) = 0$.

$$\Delta G(n) \approx \Delta G(n^*) + \Delta G'(n^*)(n - n^*) + \frac{1}{2} \Delta G''(n^*)(n - n^*)^2. \quad (2.10)$$

In addition, the limits can be extended because of the sharpness of the integrand. The homogeneous nucleation rate then becomes

$$J = k^+ \frac{N}{V} A(n^*) \exp \left[\frac{-\Delta G(n^*)}{k_B T} \right] \times \left[\int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2k_B T} [-\Delta G''(n^*)] (n - n^*)^2 \right\} dn \right]. \quad (2.11)$$

Finally the integral can be evaluated. Its value is named *Zeldovich factor*²¹ and it is denoted by Z ,

$$J = \left\{ \frac{k^+ A(n^*)}{V} \right\} \left\{ \sqrt{\frac{-\Delta G''(n^*)}{2\pi k_B T}} \right\} \left\{ N \exp \left[-\frac{\Delta G(n^*)}{k_B T} \right] \right\} = j(n^*) \times Z \times N_{n^*}. \quad (2.12)$$

Beside the already mentioned Zeldovich factor Z , the final result consists of two other terms. The first one, $j(n^*)$, can be seen as the rate, at which single particles arrive the critical cluster. The second one, N_{n^*} , is the equilibrium number of embryos with critical size. So the Zeldovich

factor represents a correction in relation to the fact, that the number of critical nuclei differs from the equilibrium value.¹⁴

2.2. Capillarity Approximation

The aim of Capillarity Approximation is to get an expression for the free energy ΔG , which is identified with the minimal work of forming a cluster. The approximation consists of assuming small clusters to have the same properties as the bulk material. One can show¹⁴ that the minimal work ΔG_{min} to form a cluster consisting of n particles by homogeneous nucleation is determined by

$$\Delta G_{min} = \sigma A + (P - P')V' + n[\mu'(T, P') - \mu(T, P)], \quad (2.13)$$

where σ is the surface tension and A is the respective surface of the cluster. P and μ are the pressure and the chemical potential of the bulk system, while P' and μ' are the corresponding quantities of the cluster, which has a volume V' . Beside on the pressures the chemical potentials μ and μ' depend also on the temperature T .

In the case of heterogeneous nucleation one has to add further terms to Equ. (2.13). If for example nucleation occurs through the presence of a solid wall denoted by the subscript w , the equation for the minimal work will read

$$\Delta G_{min} = \sigma_{\alpha\beta}A_{\alpha\beta} + (\sigma_{w\beta} - \sigma_{\alpha w})A_{\beta w} + (P - P')V' + n[\mu'(T, P') - \mu(T, P)], \quad (2.14)$$

where α and β are the particular phases of the transition $\alpha \rightarrow \beta$ and $\sigma_{\alpha\beta}, \sigma_{w\beta}, \sigma_{\alpha w}, A_{\alpha\beta}$ and $A_{\beta w}$ are the appropriate surface tensions and interfaces.

If crystalline clusters grow from the liquid, certain polyhedra would arise depending on what substances are studied. Because there are several surface areas on such a polyhedron and their respective surface tension can differ, one actually has to sum over these terms:

$$\Delta G_{min} = \sum_i \sigma_i A_i + (P - P')V' + n[\mu'(T, P') - \mu(T, P)]. \quad (2.15)$$

Here, the index i numbers the facets of the polyhedron with their corresponding surface tensions σ_i and their surface areas A_i . However, in the Capillarity Approximation the idea is to assume, that the embryo has the same physical properties as the bulk metastable phase. In particular, this means that the surface tension of the critical nucleus equals the surface tension of a flat interface. Besides it is assumed that the nucleus has a spherical shape, so one does not have to account the various facets of a crystal. If we further treat the embryo as incompressible

we can write

$$\mu'(T, P') = \mu'(T, P) - v'(P - P'). \quad (2.16)$$

Here, v' is the volume per particle of the cluster. Substituting this expression into Equ. (2.13) yields

$$\Delta G_{min} = \sigma A + n\Delta\mu, \quad (2.17)$$

where $\Delta\mu$ is the difference in the chemical potentials of the metastable and stable phase at bulk conditions T and P . Besides $\Delta\mu$ is a negative quantity. So at this point one can see that the work is determined by two terms, which compete against each other. On the one hand there is the positive surface term $\Delta G_A = \sigma A$. The larger it becomes, the more work must be performed against it. So large interfaces are more unfavorable than smaller ones. This way the assumption of a spherical shape of the nucleus is justified, because a sphere has a minimal surface. On the other hand the volume term $\Delta G_{V'} = n\Delta\mu$ (the particle number n is proportional to the volume V') becomes more negative with growing embryo. At the transition of a metastable phase towards a stable one there should be a maximum of the resulting formation work. Exactly this maximum is the barrier the system must overcome and the size of the nucleus at the maximum is identified with the critical size. Embryos which are smaller than the critical size will shrink spontaneously to dissolve in the metastable phase, while embryos which are larger will further grow to transform the whole system into the stable phase. To illustrate this further let us write Equ. (2.17) in terms of the particle number n of the embryo and its radius r , respectively. The number of particles on the surface of the embryo are proportional to $n^{2/3}$ and the surface itself is proportional to r^2 . The volume of the embryo is proportional to n and to r^3 . With the proportional constants a , b , c and d one can write

$$\Delta G_{min} = an^{2/3} - bn = cr^2 - dr^3. \quad (2.18)$$

In Fig. 2.1 ΔG_{min} is shown in dependence of r . This function has its maximum at

$$n^* = \left(\frac{2a}{3b}\right)^3 = \frac{32\pi}{3} \left[\frac{(v')^{2/3}\sigma}{-\Delta\mu} \right]^3, \quad (2.19)$$

$$r^* = \frac{2c}{3d} = \frac{2\sigma v'}{-\Delta\mu}. \quad (2.20)$$

The height of the maximum is then given by

$$\Delta G^* = \frac{4a^3}{27b^2} = \frac{16\pi}{3} \left[\frac{v'\sigma^{3/2}}{-\Delta\mu} \right]^2. \quad (2.21)$$

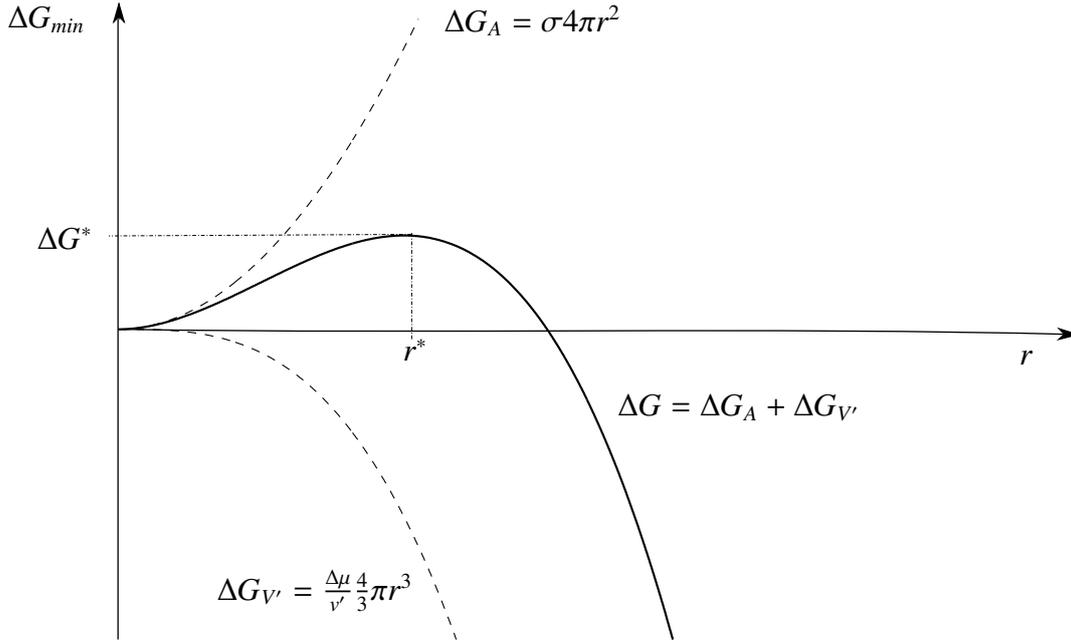


Figure 2.1: Work ΔG_{min} of nucleus formation as a function of nucleus size

2.3. Nucleation Rate in Condensed Systems

In the above sections it was assumed that the particles of the metastable phase can diffuse freely to the surface of the critical nucleus. This fact changes for an undercooled liquid, because the diffusion across the solid-liquid interface is affiliated with a second activation energy barrier. The solid and liquid phase differ not only in the density, but also in the degree of order. This affects the calculation of the forward and backward rate constants k_n^+ and k_n^- , for which Reaction Rate Theory can be used.¹⁷ The nucleation rate can then be written in the following form²²:

$$J = \mathcal{D} \times Z \times \frac{N}{V} \exp\left[-\frac{\Delta G(n^*)}{k_B T}\right]. \quad (2.22)$$

\mathcal{D} is the size diffusion of nuclei, for which Turnbull and Fisher²³ derived the expression

$$\mathcal{D} = i^* \frac{k_B T}{h} \exp\left[-\frac{E}{k_B T}\right], \quad (2.23)$$

where $i^* = 4n_*^{2/3}$ is the number of molecules at the surface of the cluster, and E is the aforementioned activation energy of crossing the liquid-solid interface. $\frac{k_B T}{h}$ with the Plank constant h is the molecular vibration frequency at the crystal surface. It is often assumed, that several rate processes such as viscous flow or self diffusion have the same activation mechanism, i.e. that the jump frequency at the interface is essentially the same as the jump frequency for bulk

diffusion or viscous flow. That way \mathcal{D} can also be written as

$$\mathcal{D} = \frac{24Dn_*^{2/3}}{l^2} = \frac{4n_*^{2/3}k_B T}{3\pi d_0^3 \eta}, \quad (2.24)$$

where D is the self diffusion constant, l is the characteristic diffusion length, d_0 is the effective molecule diameter and η is the liquid viscosity. Furthermore, if one considers an undercooled liquid at temperature T , where the equilibrium temperature (melting temperature) is $T_m > T$, one can expand in Equ. (2.17) the chemical potentials in the solid and liquid phases about their equilibrium values at the melting temperature T_m and the pressure P . Differentiation and solving for the radius r gives

$$r^* = \frac{2\sigma v^s T_m}{\Delta T \Delta h}, \quad (2.25)$$

where Δh is the heat of fusion, v^s is the solid volume per molecule and ΔT is $T_m - T$. The minimal work associated with forming of such an embryo of radius r^* is then

$$\frac{\Delta G^*}{k_B T} = \frac{16\pi \Delta s}{3 k_B} \left[\frac{\sigma(v^s)^{2/3}}{\Delta h} \right]^3 \frac{T_m^3}{T(T_m - T)^2}. \quad (2.26)$$

If one substitutes this into Equ. (2.22) and differentiates with respect to T , one will get an expression for maximizing the nucleation rate:

$$\frac{[1 - \theta(J_{max})]^3}{5 - 3\theta(J_{max})} = \frac{16\pi \Delta h}{3 E} \left[\frac{\sigma(v^s)^{2/3}}{\Delta h} \right]^3, \quad (2.27)$$

where $\theta(J_{max}) = T(J_{max})/T_m$. This maximum of the nucleation rate can be interpreted as follows: with increasing undercooling the critical size of embryos becomes smaller and the free energy barrier becomes lower. Thus the nucleation rate has a sharp rise. But with further undercooling the mobility of the particles strongly decreases and so does D . Since the nucleation rate in condensed systems is related to D via Equ. (2.22) and Equ. (2.24), J decreases, too.

3. Computational Methods

Computer simulations become important whenever a problem is analytically unsolvable, because it is too complicated or the solution of the equations, which describe the problem, cannot be expressed in a closed form. Unfortunately this is the case in most problems of nature. They also are a functional instrument if experimental conditions are hard to realize, for example very high temperatures or pressures, or just like in this work, a nano-sized system. A common application of computer simulations is the prediction of properties of a material.²⁴ In this context on the one hand their results can be compared with experiments to test if the underlying model is appropriate. On the other hand comparison with the properties predicted by the underlying theory can show us if the theory is possibly flawed. In this chapter I will give an overview of some computational methods in simulating molecular systems. All these methods are applied in this work. Beside basic algorithms like molecular dynamics (MD) or Monte Carlo methods, some advanced techniques are required like umbrella sampling. This is due to the fact that I study a phase transition, which is an activated process. With straightforward MD it would take too much time to observe even only one of such a transition.

Already here I want to refer what programs and program packages I use:

- to calculate the dynamics of the system I use the package GROMACS.²⁵ In this package a MD program with very high performance is offered, which can be run on single processors as well as on parallel machines.
- the umbrella sampling code is a slightly modified version of the program by Andreas Singraber. In this code a replica exchange method is implemented, too. The windows can be simulated in parallel on different cores.
- to analyze the data from the umbrella sampling, I use the WHAM routine by Alan Grossfield.²⁶

3.1. Molecular Dynamics

3.1.1. Potential

To simulate the motion of atoms or molecules for a certain time t one can use Molecular Dynamics (MD). From knowledge of the positions and velocities of the particles, Newton's

equations of motion (3.1) and (3.2) can be integrated simultaneously for each particle:

$$\dot{\mathbf{r}}_i(t) = \mathbf{v}_i(t), \quad (3.1)$$

$$\dot{\mathbf{v}}_i(t) = \frac{\mathbf{F}_i(t)}{m_i} = -\frac{1}{m_i} \nabla U(t), \quad (3.2)$$

where \mathbf{r}_i is the position of particle i , \mathbf{v}_i is its velocity and m_i is its mass. For determination of the forces \mathbf{F}_i a potential $U(\mathbf{r}^N)$ is required, which is dependent of the particle coordinates. In many cases it is sufficient to use classical empirical potentials.²⁴ One distinguishes between bonded forces, which correspond to chemical bonds, and non-bonded forces, the forces due to van der Waals and electrostatic interaction. Most empirical models take advantage of pair potentials, i.e. each pair of particles contributes to the total energy. Furthermore an empirical potential model usually has free parameters such as partial charges or bond lengths. Through adjustment to experimental data one determines these parameters and gets an appropriate potential model for the particles in the studied system. In this work I use such an empirical potential for water, which is named TIP4P₂₀₀₅²⁷ and is further described in Chapter 4. Sometimes it is possible to gather several particles into groups and to treat them as one pseudo-atom with an own potential. Such a coarse-graining approach for water is used in the mW model.²⁸ If one is interested in a more accurate potential, which is rather based on theory than on empirical fitting, one has to use quantum instead of classical mechanics. The Schrödinger equation has to be solved in order to get the electronic structure of the system. Such methods are called ab-initio methods. Such calculations can be done with Density Functional Theory (DFT)²⁹, in which the ground state of the potential is assumed to be fully determined by the electron density. Although this method also uses approximations, it gives much more information about the system, for instance the density of electronic states. Unfortunately, the computational cost is much higher than in classical molecular dynamics.

3.1.2. Integration of the equations of motion

To solve Newton's equations of motion (3.1) and (3.2) the time must be discretized into small time steps Δt . With knowledge of the positions and velocities at a certain time t , one wants to compute these quantities at a time $t + \Delta t$. To do so, one formally integrates both sides of Equ.

(3.1) and (3.2):

$$\begin{aligned}\int_t^{t+\Delta t} dt' \dot{\mathbf{r}}_i &= \int_t^{t+\Delta t} dt' \mathbf{v}_i(t'), \\ \int_t^{t+\Delta t} dt' \dot{\mathbf{v}}_i &= \int_t^{t+\Delta t} dt' \frac{\mathbf{F}_i(t')}{m_i}.\end{aligned}\quad (3.3)$$

While the integral before the equal sign can easily be calculated, the integrand after the equal sign must be approximated by a Taylor expansion, because it is time dependent itself:

$$\begin{aligned}\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t) &= \int_t^{t+\Delta t} dt' \sum_{k=0}^{\infty} \frac{(t' - t)^k}{k!} \left(\frac{d^k}{dt^k} \mathbf{v}_i \right)_t, \\ \mathbf{v}_i(t + \Delta t) - \mathbf{v}_i(t) &= \int_t^{t+\Delta t} dt' \sum_{k=0}^{\infty} \frac{(t' - t)^k}{k!} \left(\frac{d^k}{dt^k} \mathbf{F}_i/m_i \right)_t.\end{aligned}\quad (3.4)$$

After solving the remaining integral and rearrangement one gets

$$\begin{aligned}\mathbf{r}_i(t + \Delta t) &= \mathbf{r}_i(t) + \sum_{k=1}^{\infty} \frac{\Delta t^k}{k!} \left(\frac{d^{k-1}}{dt^{k-1}} \mathbf{v}_i \right)_t, \\ \mathbf{v}_i(t + \Delta t) &= \mathbf{v}_i(t) + \sum_{k=1}^{\infty} \frac{\Delta t^k}{k!} \left(\frac{d^{k-1}}{dt^{k-1}} \mathbf{F}_i/m_i \right)_t.\end{aligned}\quad (3.5)$$

The so-called Velocity Verlet algorithm³⁰, which I use in my MD simulations, solves these equations up to the second order:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \Delta t^2 \frac{\mathbf{F}_i(t)}{2m_i}, \quad (3.6)$$

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \Delta t \frac{1}{2m_i} [\mathbf{F}_i(t) + \mathbf{F}_i(t + \Delta t)]. \quad (3.7)$$

The choice of the time step plays an important role: if it is too small, the simulation of long time scales will take too much time, because many time steps are necessary. If it is too large, conservation of the energy will degrade, because rounding errors will build up. Another important feature of the Velocity Verlet algorithm is the symmetry of time. This means, that if one chooses a negative time step, one will get older configurations of the system. Thus the trajectory is reversible.

3.1.3. Molecular Dynamics in the Canonical Ensemble

As mentioned before, the original Velocity Verlet algorithm conserves the total energy with high accuracy. This situation is given in the (N, V, E) -ensemble, also called microcanonical ensemble, in which the temperature fluctuates. In usual experiments it is the other way around, the total energy fluctuates and the temperature is constant. Thus the (N, V, T) - or canonical ensemble is the correct description. If a computer simulation should be at constant temperature, a thermostat must be added. The idea of the Nosé - Hoover thermostat³¹, which is used in my constant temperature simulations, is the following: because on average the velocities are related to the temperature via the kinetic energy,

$$E_{kin} = \sum_i^N \frac{1}{2} m_i v_i^2 = \frac{3}{2} N k_B T, \quad (3.8)$$

where the sum goes over all N particles, an additional term with a free parameter ζ appears in the differential equation for the velocities. A third differential equation evolves this parameter in time so that the positions and velocities are canonically distributed and the kinetic energy conforms to the reference value $(3/2)Nk_B T$,

$$\dot{\mathbf{r}}_i = \mathbf{v}_i \quad (3.9)$$

$$\dot{\mathbf{v}}_i = \mathbf{F}_i/m_i - \zeta \mathbf{v}_i \quad (3.10)$$

$$\dot{\zeta} = \frac{1}{\tau^2} \left[\frac{\sum_i^N m_i v_i^2}{3Nk_B T} - 1 \right]. \quad (3.11)$$

In the last equation a new parameter τ with the dimension of time appears, which is termed relaxation time of the thermostat. It should be chosen carefully and in any case smaller than the simulation time to ensure the thermostat to be effective. On the other hand it should be larger than the typical relaxation times of the system to avoid a too large perturbation of the equations of motion.

3.1.4. Constraints

Because in the studied system there are fixed bond lengths and bond angles like the O-H distance and the H-O-H angle, constraints must be introduced. The constraints, which must be fulfilled by the equations of motion, can be expressed as

$$\sigma_k(\mathbf{r}_1 \dots \mathbf{r}_N) = 0, \quad (3.12)$$

with $k = 1 \dots K$, where K is the total number of constraints. Each constraint can depend on the coordinates of all particles in the system. This way, the constraint $\sigma = \frac{1}{2}[(\mathbf{r}_\alpha - \mathbf{r}_\beta)^2 - d^2] = 0$ defines the distance d between two atoms α and β . By using the method of the Lagrange multipliers the equation of motion for a certain particle i becomes

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i - \sum_{k=1}^K \lambda_k \frac{\partial \sigma_k}{\partial \mathbf{r}_i} \quad (3.13)$$

$$= \mathbf{F}_i + \mathbf{G}_i, \quad (3.14)$$

where λ_k is the Lagrange multiplier of the k -th constraint and the sum in Equ. (3.13) can be interpreted as constraint force \mathbf{G}_i . Constraint forces do not change the total energy, because they perform no work. After integrating both sides of Equ. (3.13) twice in time, the position $\mathbf{r}_i(t + \Delta t)$ can be written as

$$\mathbf{r}_i(t + \Delta t) = \hat{\mathbf{r}}_i(t + \Delta t) - \frac{\Delta t^2}{m_i} \sum_{k=1}^K \lambda_k \frac{\partial \sigma_k}{\partial \mathbf{r}_i}. \quad (3.15)$$

Here the unconstrained position is expressed by $\hat{\mathbf{r}}_i(t + \Delta t)$. To enforce the constraint one claims that $\sigma_k(t + \Delta t) = 0$ and therefore one has to solve a system of non-linear equations for the λ_k . For the special case of three constraints, which is given at a water molecule, this system can be solved in an efficient way and analytically by the SETTLE algorithm.³² This algorithm is also implemented in the GROMACS package and I use it in my MD simulations.

3.2. Monte Carlo Methods

3.2.1. Metropolis Monte Carlo

These techniques are called Monte Carlo methods (MC), which solve problems by the use of random numbers. One of the most important features is the sampling of probability distributions. The Metropolis Monte Carlo algorithm achieves this by so-called random walks.³³ Based on a starting point x a trial move is performed into the neighborhood $x + \Delta x$. Here the direction and amount of the displacement Δx is chosen randomly, but cannot exceed a certain maximum amount. If the probability $f(x + \Delta x)$ of the new point is higher than the one of the old point $f(x)$, the trial move is accepted. $x + \Delta x$ then becomes the new starting point and the procedure continues. Otherwise, the trial move is accepted only with a probability $\frac{f(x + \Delta x)}{f(x)}$.

The acceptance rule can be simply expressed by the formula

$$P_{acc}(x \rightarrow x + \Delta x) = \min\left(1, \frac{f(x + \Delta x)}{f(x)}\right). \quad (3.16)$$

The minimum function $\min(x, y)$ returns the smaller one of its arguments. The acceptance probability can be controlled by the maximum amount of displacement Δx . It should be about 20% - 70%. It is important that rejected moves also count as moves, i.e. the number of occurrences of the old point is increased. If the number of the random walk steps is large enough, the Metropolis Monte Carlo algorithm will sample the probability distribution $f(x)$ in a satisfactory way.

The algorithm can be transferred to the problem of molecular systems. For example, the canonical probability distribution is given by

$$f(\mathbf{r}^N, \mathbf{p}^N) = \frac{\exp[-\beta H(\mathbf{r}^N, \mathbf{p}^N)]}{\iint d\mathbf{r}^N d\mathbf{p}^N \exp[-\beta H(\mathbf{r}^N, \mathbf{p}^N)]}, \quad (3.17)$$

where \mathbf{r}^N and \mathbf{p}^N are the positions and momenta of all N particles and H is the Hamiltonian of the system, which corresponds to the total energy. A trial move consists of the random choice of one particle i and its displacement in space. Since the momenta of the particles are not considered, the acceptance probability of such a move is given by

$$P_{acc}(\mathbf{r}_i^{old} \rightarrow \mathbf{r}_i^{new}) = \min(1, \exp[-\beta \Delta U]), \quad (3.18)$$

where $\Delta U = \Delta U(\mathbf{r}^{new}) - \Delta U(\mathbf{r}^{old})$ is the difference of the potential energies of the old and new configuration. In many cases this difference can easily be calculated, because the potential energy depends only on the particle positions and consists of pair contributions, of which most stay unchanged. After N trial single particle moves a MC step is finished and can be repeated. In the canonical ensemble the expectation value of a quantity A , which is only dependent on the particle positions, is given by

$$\langle A(\mathbf{r}^N) \rangle = \frac{1}{Z} \int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)] A(\mathbf{r}^N), \quad (3.19)$$

where $Z = \int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)]$ is the partition function, which normalizes the distribution. For a correctly sampled canonical distribution this expectation value can be written as average

value over all MC steps M :

$$\langle A(\mathbf{r}^N) \rangle \approx \bar{A} = \frac{1}{M} \sum_{i=1}^M A(\mathbf{r}_1^{(i)}, \dots, \mathbf{r}_N^{(i)}). \quad (3.20)$$

The error of such a mean value would be proportional to the square root of the number of measurements M , if the correlation between successive states vanished. But since new states are created by old ones through only small displacements this is not true. Thus one alternative is to wedge the sample, i.e. to evaluate for example only every tenth or twentieth step of the sampling. Although then the sample is smaller, the correlation within the sample is reduced.

3.2.2. Hybrid Monte Carlo

Another alternative to reduce the correlation within the sample is to use the Hybrid Monte Carlo algorithm (HMC).³⁴ Instead of randomly choosing a particle and displacing it, the whole system is evolved according to MD. So at each HMC step the following happens: at first all particles are assigned a velocity at random according to a Gauß distribution corresponding to the simulated temperature. Then a trajectory is generated for several timesteps. After finishing the MD simulation the difference of total energy is determined. Now the same acceptance rule as above (Equ. (3.18)) is applied. The acceptance probability can be controlled by the size of the timestep, where large timesteps imply smaller acceptance probabilities. This way new, independent configurations are created effectively.

3.3. Umbrella Sampling

As described in Chapter 2 the transition of a metastable to a stable phase occurs very quickly and most of the time the system resides in one of the two states. This means, that the probability of a transition state is very low. Even if a conventional MC or MD simulation show such a transition, the statistics will be too poor to evaluate thermodynamic properties. Unfortunately it is the transition state, which is of interest, and we want to study especially this region of phase space. One way out is the use of Umbrella Sampling.³⁵ With the aid of this method the system is forced to move and stay in the phase space region of interest by introducing a biasing potential. This bias potential can be realized by a harmonic function $V(\Phi)$:

$$V(\Phi(\mathbf{r}^N)) = \frac{1}{2} k_{\Phi} (\Phi(\mathbf{r}^N) - \Phi_0)^2, \quad (3.21)$$

where Φ is the order parameter, which describes the phase transition, Φ_0 is the umbrella centre, around which Φ should be sampled, and k_Φ is the umbrella spring constant. By an appropriate choice of k_Φ one controls the narrowness around the umbrella centre. In practice one undertakes several simulations with different umbrella windows. In each window one tries to sample over a certain width in order to cover the whole order parameter range without using too much windows. Nevertheless the distributions of the order parameter of neighboring windows should overlap, so that good statistics of each value of Φ are given.

The bias potential can be added to the normal potential U of the system, $U' = U + V$. This way a kind of weight function $W = e^{-\beta V}$ appears when calculating the expectation value of $A(\mathbf{r}^N)$ in the biased ensemble (denoted with the subscript W) :

$$\begin{aligned}\langle A(\mathbf{r}^N) \rangle_W &= \frac{\int d\mathbf{r}^N \exp[-\beta U'(\mathbf{r}^N)] A(\mathbf{r}^N)}{\int d\mathbf{r}^N \exp[-\beta U'(\mathbf{r}^N)]} \\ &= \frac{\int d\mathbf{r}^N \exp[-\beta U] W(\mathbf{r}^N) A(\mathbf{r}^N)}{\int d\mathbf{r}^N \exp[-\beta U] W(\mathbf{r}^N)}.\end{aligned}\quad (3.22)$$

The expectation value of the unbiased canonical ensemble is related to the biased one via

$$\begin{aligned}\langle A(\mathbf{r}^N) \rangle_{NTV} &= \frac{\int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)] A(\mathbf{r}^N)}{\int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)]} \\ &= \frac{\int d\mathbf{r}^N \exp[-\beta U] W(\mathbf{r}^N) A W(\mathbf{r}^N)^{-1}}{\int d\mathbf{r}^N \exp[-\beta U] W(\mathbf{r}^N) W(\mathbf{r}^N)^{-1}} \\ &= \frac{\langle A/W(\mathbf{r}^N) \rangle_W}{\langle 1/W(\mathbf{r}^N) \rangle_W}.\end{aligned}\quad (3.23)$$

An important application of umbrella sampling is the determination of the free energy as function of the order parameter. It can be shown - I demonstrated it in Chapter 2 in the special case of the cluster size n as order parameter - that the free energy ΔF is related to the order parameter via

$$\beta \Delta F(\Phi) = -\ln[P(\Phi)].\quad (3.24)$$

The probability $P(\Phi)$, i.e. the probability, that the system assumes a configuration \mathbf{r}^N with

order parameter Φ can be expressed by

$$\begin{aligned}
P(\Phi) &= \langle \delta(\Phi - \Phi(\mathbf{r}^N)) \rangle \\
&= \frac{\langle e^{\beta V(\Phi)} \delta(\Phi - \Phi(\mathbf{r}^N)) \rangle_W}{\langle e^{\beta V(\Phi)} \rangle_W} \\
&= \frac{e^{\beta V(\Phi)} \langle \delta(\Phi - \Phi(\mathbf{r}^N)) \rangle_W}{\langle e^{\beta V(\Phi)} \rangle_W}.
\end{aligned} \tag{3.25}$$

Here $\delta(\Phi - \Phi(\mathbf{r}^N))$ is the Dirac delta-distribution and Equ. (3.23) is used. It is interesting to see that, while the numerator can easily be determined from the histogram of Φ obtained in the simulation, the denominator does not have to be calculated. If one considers two neighboring windows with umbrella centers Φ_1 and Φ_2 but overlapping Φ -distributions, one can write

$$P(\Phi) = \frac{e^{\beta V_1(\Phi)} \langle \delta(\Phi - \Phi(\mathbf{r}^N)) \rangle_{W_1}}{\langle e^{\beta V_1(\Phi)} \rangle_{W_1}} = \frac{e^{\beta V_2(\Phi)} \langle \delta(\Phi - \Phi(\mathbf{r}^N)) \rangle_{W_2}}{\langle e^{\beta V_2(\Phi)} \rangle_{W_2}}. \tag{3.26}$$

After taking the logarithm at both sides in rearranging one gets

$$\begin{aligned}
\beta V_1(\Phi) + \ln(\langle \delta(\Phi - \Phi(\mathbf{r}^N)) \rangle_{W_1}) - \beta V_2(\Phi) - \ln(\langle \delta(\Phi - \Phi(\mathbf{r}^N)) \rangle_{W_2}) = \\
\ln(\langle e^{\beta V_1(\Phi)} \rangle_{W_1}) - \ln(\langle e^{\beta V_2(\Phi)} \rangle_{W_2}).
\end{aligned} \tag{3.27}$$

The right hand side of this equation is constant with respect to Φ . Thus measuring $F(\Phi)$ in the two different windows, i.e. with different umbrella potentials, leads to constantly shifted results (see also in Fig. 3.1). The easiest way to make the free energy function continuously is simply by shifting these parts one upon the other. So the free energy in window i can be calculated by

$$\beta \Delta F_i(\Phi) = -\beta V_i - \ln(\langle \delta(\Phi - \Phi(\mathbf{r}^N)) \rangle_{W_i}) + b_i, \tag{3.28}$$

where b_i is the shifting constant. The histogram of the sampled Φ is close to a Gauß-distribution. Its center usually is slightly right or left of the umbrella centre of the biasing potential. After taking the logarithm, as prescribed by Equ. (3.28), the quadratic term roughly cancels with the quadratic term of the bias potential. So in each window the remaining contribution to the free energy is roughly linear in Φ . Whether the free energy increases or decreases with Φ depends on the relative position of the quadratic curves. Generally it is possible to compose each small linear segment of the free energy by adjusting b_i in order to get the whole free energy landscape as function of the entire order parameter range. In Fig. 3.1 one can see the schematic

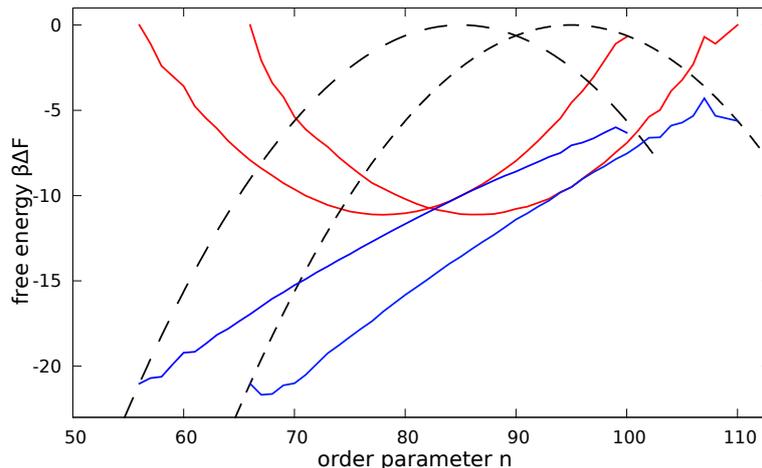


Figure 3.1: Schematic evaluation of the free energy in two neighboring windows. As order parameter the largest clustersize n is used. The black dashed lines represent the umbrella potentials around the values $n = 85$ and $n = 95$. The red solid lines are the logarithms of the Φ -histograms in these windows. Adding the red and black curves leads to the linear blue lines, which finally represent the free energies. One can see, that the free energies differ approximately by a constant in the overlapping region.

evaluation of the free energy in two neighboring windows as described here.

3.4. Replica Exchange

Replica exchange³⁶ is also known as parallel tempering. The main idea is to simulate several copies of the same system at different temperatures in parallel and to allow the exchange of configurations between them in order to enhance the sampling. If the energy landscape of a system has multiple minima, it can be trapped in one of them for a longer time. This way the phase space is not covered completely and the sampling is poor. One way out is to allow the system once in a while to swap its configuration with a system of higher temperature in order to overcome energy barriers. The different temperatures do not need to correspond with real physical temperatures. In the present case the different windows of the order parameter Φ , as described in the previous section, represent the replicas. Replica exchange only takes place between neighboring windows with overlapping Φ -distributions. The frequency of swapping attempts is chosen in such a way that the system can equilibrate between two attempts. The acceptance rule for such swapping attempts follows the Metropolis algorithm. Once a swapping attempt between the configurations i and j is performed, the energy before (V_{old}) and after the change (V_{new}) is computed. Because I use the configurations of the umbrella sampling

windows as replicas, the energy V is based on the bias potential of the particular window:

$$V_{old} = \frac{1}{2}k_i(\Phi_i - \Phi_{0,i})^2 + \frac{1}{2}k_j(\Phi_j - \Phi_{0,j})^2 \quad (3.29)$$

$$V_{new} = \frac{1}{2}k_i(\Phi_j - \Phi_{0,i})^2 + \frac{1}{2}k_j(\Phi_i - \Phi_{0,j})^2 \quad (3.30)$$

The acceptance probability of the exchange is then given by

$$P_{acc}(i \leftrightarrow j) = \min(1, \exp[-\beta(V_{new} - V_{old})]). \quad (3.31)$$

In Fig. 3.2 an example of replica exchange paths is shown. One can see that in principle it is possible to have, for instance, an original window 0 configuration after some swapping moves in window 10 and vice versa. Hence more configurations can be sampled in the phase space.

3.5. Weighted Histogram Analysis Method

We have seen that the free energy in each umbrella sampling window is determined up to a constant b_i . As said before, it is generally possible to adjust every segment to make the free energy continuous. But this method can cause some practical problems if the free energies in overlapping regions of neighboring windows are not exactly parallel. This is why I use the weighted histogram analysis method (WHAM).³⁷ To explain this technique in a clear way let us rename some quantities of Equ.(3.25).

$$\begin{aligned} \langle \delta(\Phi - \Phi_j) \rangle_{W_i} &= \frac{n_{ij}}{N_i}, \\ e^{-\beta V_i(\Phi_j)} &= c_{ij}, \\ \langle e^{\beta V_i} \rangle &= e^{\beta b_i} = f_i. \end{aligned} \quad (3.32)$$

In the first line the order parameter Φ has been discretized into j bins. In the i -th simulation, n_{ij} is the number of counts in the histogram bin associated with Φ_j . N_i is the total number of samples in the i -th simulation. This way the fraction n_{ij}/N_i is the biased probability of bin j . In the last line one can see again how the denominator of Equ. (3.25) is related to the shifting constant b_i . With this abbreviated notation, an estimation for the unbiased probability P_{ij} can be written as

$$P_{ij} = \frac{n_{ij}}{N_i c_{ij} f_i}. \quad (3.33)$$

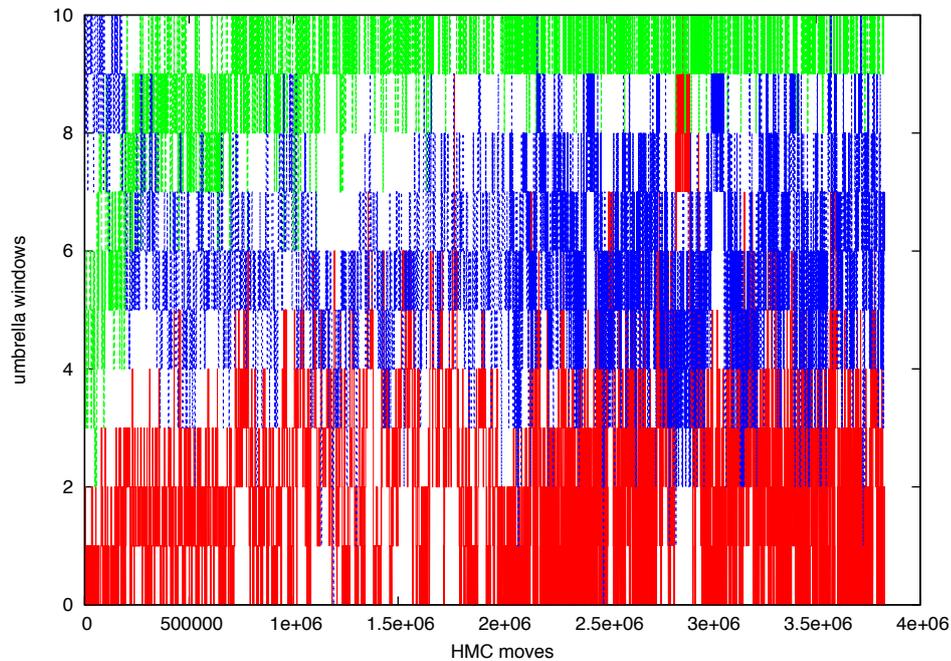


Figure 3.2: Process of successful swapping attempts in an umbrella sampling simulation with replica exchange. As order parameter the size n of the largest crystalline cluster within the system is used. The umbrella centers are located at $n = 5, 10, 15, 25, 35, 45, 55, 65, 75, 85, 95, 105$. Different colors denote configurations from different original windows (red line: window 0 with $n_0 = 5$, green line: window 5 with $n_0 = 55$, blue line: window 10 with $n_0 = 105$). This diagram was generated by an umbrella sampling simulation with 216 water molecules at 180 K and $3.8 \cdot 10^6$ HMC steps.

The real unbiased probability for bin j is defined as the sum over these estimates of each simulation with certain weight factors w_i :

$$P^0(\Phi_j) = \sum_i^{sim} w_i P_{ij}. \quad (3.34)$$

The idea is now to choose the weights in such a way that the variance of $P^0(\Phi_j)$ gets minimized. Considering the number of counts in bin j to be Poisson distributed, this variance can be written as

$$\text{var} [P^0(\Phi_j)] = \sum_i^{sim} w_i^2 \text{var} [P_{ij}] = \sum_i^{sim} \frac{w_i^2 \text{var} [n_{ij}]}{N_i^2 c_{ij}^2 f_i^2} = \sum_i^{sim} \frac{w_i^2 P^0(\Phi_j)}{N_i c_{ij} f_i}. \quad (3.35)$$

Minimizing this expression with respect to w_i and with the constraint $\sum_i^{sim} w_i = 1$ one gets the following condition for the optimal w_i 's, which can be substituted in Equ.(3.34):

$$w_i = \frac{N_i c_{ij} f_i}{\sum_k^{sim} N_k c_{kj} f_k}. \quad (3.36)$$

At last the unbiased probability for the order parameter Φ_j is given by

$$P^0(\Phi_j) = \frac{\sum_i^{sim} n_{ij}}{\sum_i^{sim} N_i c_{ij} f_i}, \quad (3.37)$$

where $P^0(\Phi_j)$ is still dependent of f_i , which is unknown yet. But in Equ. 3.33 one recognizes, that f_i must fulfill the function of a normalization constant, since $\sum_j^{bins} n_{ij}/N_i = 1$. This can be achieved by setting $f_i^{-1} = \sum_j^{bins} c_{ij} P_{ij}$. Because P_{ij} itself is the estimation of $P^0(\Phi_j)$ one can finally write

$$f_i^{-1} = \sum_j^{bins} c_{ij} P^0(\Phi_j). \quad (3.38)$$

The last two equations (3.37) and (3.38) are known as the WHAM equations. They are a system of $sim + bins$ nonlinear equations, which must be solved iteratively until self-consistency is reached. Getting the right unbiased probabilities and shifting factors one can calculate the free energy via Equ. (3.24).

4. Model Definition

There are several different potential models, which were developed to reproduce the properties of water. Results of the simulation with each particular model should be in accord with experimental reports, for example the phase diagram, the pair distribution function or the dielectric constant. In addition, the model should be as simple as possible for reasons of computational cost.

I decided on the TIP4P₂₀₀₅ model, because it has several advantages and it reproduces those properties very well, which are important in this work^{27,38}. The TIP4P₂₀₀₅ model is a reparametrization of the TIP4P model, i.e., it uses four sites of interaction. The water molecule is considered as rigid. Both the oxygen-hydrogen distances and the angle spanned by the two oxygen-hydrogen bonds, are fixed. They have a value of $d_{OH} = 0.9572 \text{ \AA}$ and $\angle(HOH) = 104.52^\circ$ respectively, and are geared to the experimental results of an isolated molecule.³⁹ The fourth site M is coplanar with the molecule and lies on the bisector of the (HOH)-angle with distance $d_{OM} = 0.1546 \text{ \AA}$ from the oxygen. Molecules interact via two kinds of pair potentials. On the one hand there is a Lennard-Jones potential, whose center is the oxygen. It has the form

$$u_{LJ} = 4\epsilon[(\sigma/r_{OO})^{12} - (\sigma/r_{OO})^6], \quad (4.1)$$

where r_{OO} is the distance between the two oxygens of the interacting molecules, $\sigma = 3.1589 \text{ \AA}$ is the location of the zero-value of the potential and $\epsilon/k_B = 93.2 \text{ K}$ is its depth (k_B is the Boltzmann constant). On the other hand, two molecules interact via a Coulomb-potential, too. It has the form

$$u_{CI} = \frac{e^2}{4\pi\epsilon_0} \sum_{a,b} \frac{q_a q_b}{r_{ab}}, \quad (4.2)$$

where e is the charge of an electron, ϵ_0 is the permittivity of vacuum and a, b are the indices of the charged centers of the molecule. They are located on the hydrogens with a magnitude of $q_H = 0.5564 e$ and on the M-site with an amount of $q_M = -2q_H$. The parameters σ , ϵ , q_H and d_{OM} were determined by nonlinear fits to the following properties:

- temperature at maximal density T_{md}
- enthalpy of vaporization $\Delta_v H$
- densities of liquid water at ambient conditions
- density of ice II at 123 K and 0 MPa
- density of ice V at 223 K and 530 MPa

- range of temperatures at which ice III is the thermodynamic stable phase at 300 MPa

Here is a brief summary, why the authors of the TIP4P₂₀₀₅-model think that these properties are important:²⁷ While 3-site models like TIP3P or SPC/E describe the phase diagram only in an unsatisfactory way because they overstabilize ice II (this is probably a consequence of the negative charge located at the oxygen site), 4-site models achieve better results. To improve this situation with ice II further, it is convenient to adjust the fit parameters especially to the experimental results of the densities of ice II and ice V, which is nearby in the phase diagram. For this reason the authors also show that it is advisable to take account of the area of stable ice III, too.⁴⁰ In this way not only the area of stable ice II shrinks, but also ice Ih becomes the most stable phase at ambient conditions, which is in agreement with experimental observation. By the way, the temperature of maximal density, T_{md} , plays an important role, because it is closely related to the melting temperature T_m and the enthalpy of vaporization $\Delta_v H$. Because it is impossible to find a set of fit parameters that give simultaneously good results for both of these quantities, the proper approximation of T_{md} of a model will provide a good balance between the deviation of T_m , which is always too low, and $\Delta_v H$, which is too large in TIP4P-like models. The results of the adjustment of the potential parameters are collected in Table 4.1. For comparison, the parameters of the original TIP4P and of the SPC/E model are included, too. Table 4.2 shows the dipole and quadrupole moments of these models. The dipole moments are always too large, because it is necessary to implement the induced polarization in an averaged way into the model. This is the price of using a rigid geometry and pair interaction. If one compares gas phases with condensed phases (which is not the case in this work), one has to add a correction term for the so-called self energy.⁴¹ The correction term takes account of the fact that in the gas phase the molecules can move freely and are not electrically induced by others. If μ_i is the dipole moment of the model molecule, μ_g is the dipole moment of a gas molecule and α is the polarizability of water, the correction term is

$$\Delta u_{pol} = (\mu_i - \mu_g)^2 / 2\alpha. \quad (4.3)$$

The quadrupole moments of these rigid, unpolarizable models are lower than of isolated molecules. But an improvement of these values could lead to an improvement of the properties as several studies with polarizable models had shown.^{42,43}

In comparison with other rigid models the results of TIP4P₂₀₀₅ are very good. For example the phase diagram, whose reproduction is an essential test of every model, is in good agreement with experiment (except of a shift in the $T - P$ plane). This can be seen in Fig. 4.1. In contrast the phase diagram of TIP3P is shown in Fig. 4.2. As mentioned before, there is a large area of

stable ice II while other ice phases seem not to be stable.

	d_{OH} [Å]	\angle (HOH) [deg]	σ [Å]	ϵ/k_B [K]	q_h [e]	d_{OM} [Å]
SPC/E	1.0	109.47	3.1656	78.20	0.4238	0
TIP4P	0.9572	104.52	3.1540	78.02	0.52	0.15
TIP4P₂₀₀₅	0.9572	104.52	3.1589	93.20	0.5564	0.1546

Table 4.1: Parameters of SPC/E (Ref.(41)), TIP4P (Ref. (44)) and TIP4P₂₀₀₅ (Ref. (27)) . d_{OH} and \angle (HOH) are given parameters. The other parameters are adjusted by nonlinear fits on the experimental results. The charges are in units of the elementary charge.

	μ [D]	Q_{xx} [DÅ]	Q_{yy} [DÅ]	Q_{zz} [DÅ]
SPC/E	2.350	2.19	-1.88	-0.30
TIP4P	2.177	2.20	-2.09	-0.11
TIP4P₂₀₀₅	2.305	2.36	-2.23	-0.13
Gas (expt.)	1.850	2.63	-2.50	-0.13

Table 4.2: Dipole moment μ and quadrupole moment Q of SPC/E, TIP4P and TIP4P₂₀₀₅ in comparison with the experimental values. When Q is calculated, the x - and z -axes are in the plane of the molecule, the z -axis lies on the C_2 -symmetry axis. The origin is the center of mass of the molecule.

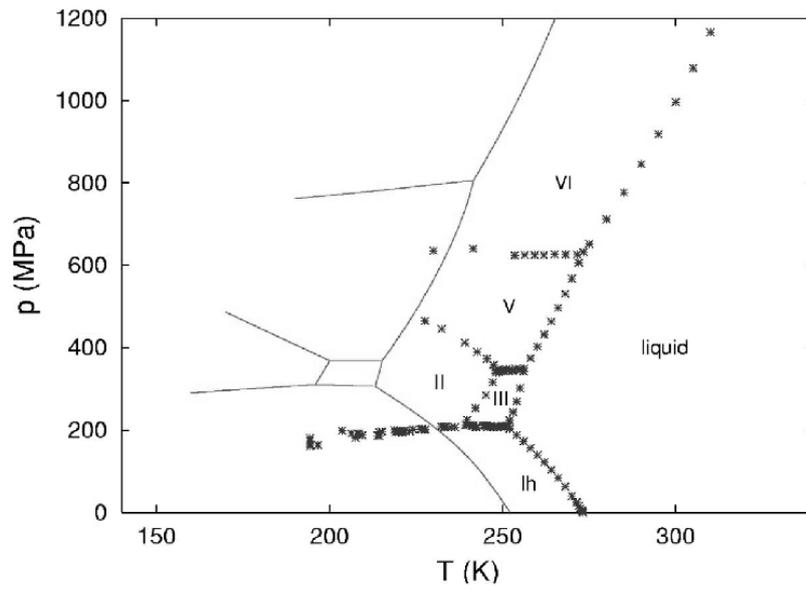


Figure 4.1: Phase diagram of the TIP4P₂₀₀₅ model. Stars represent experimental values while solid lines are simulation results. The graphic is taken from Ref. (27)

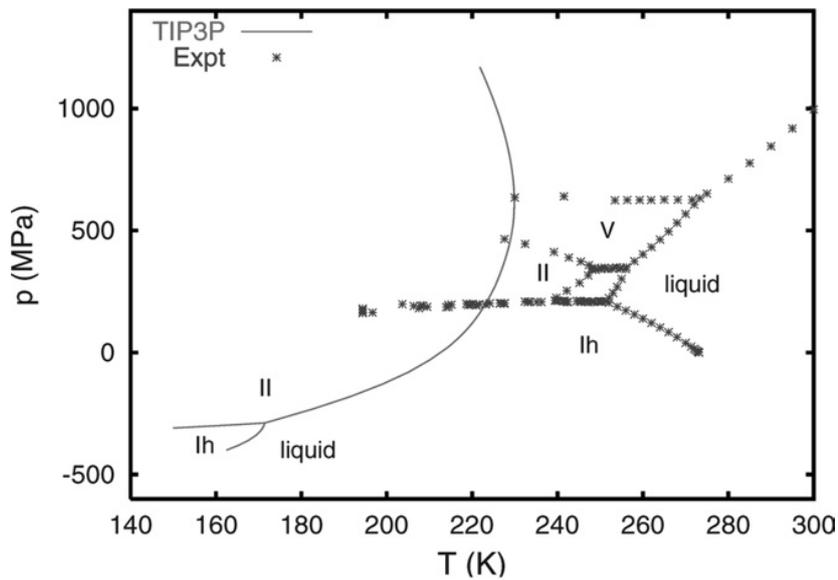


Figure 4.2: Phase diagram of the TIP3P model. Stars represent experimental values while solid lines are simulation results. The graphic is taken from Ref. (38)

5. Ice Configurations

In this chapter I want to give some information about the structure of hexagonal ice Ih and cubic ice Ic and the routine, which I use to build a spherical cluster of these two ice configurations. Both ice configurations consists of tetrahedral networks, meaning that each molecule is at the center of a tetrahedron and hydrogen bonded to its four nearest neighbors.⁴⁵

5.1. Ice Ih

The unit cell of hexagonal ice consists of four molecules and has symmetry $P6_3/mmc$ having lattice parameters $a = 4.5181 \text{ \AA}$, $c = 7.3560 \text{ \AA}$ ($90^\circ, 90^\circ, 120^\circ$, see Fig. 5.1).⁴⁶ So, the cell vectors can be written as follows, where \mathbf{X} , \mathbf{Y} and \mathbf{Z} are the cartesian unit vectors:

$$\begin{aligned} \mathbf{A}_1 &= \frac{1}{2}a\mathbf{X} - \frac{\sqrt{3}}{2}a\mathbf{Y} \\ \mathbf{A}_2 &= \frac{1}{2}a\mathbf{X} + \frac{\sqrt{3}}{2}a\mathbf{Y} \\ \mathbf{A}_3 &= c\mathbf{Z} \end{aligned} \quad (5.1)$$

Along the c -axis the crystal can be seen as sheets lying on top of each other. This sheets contain hexagonal rings in the chair-form, consisting of six H-bonded water molecules. There are also hexagonal rings parallel to the c -axis, but these ones have the boat-form.⁴⁵ This way each molecule has three bonds in a staggered arrangement to its neighbors in the plane of the chair-rings and one eclipsed bond to a molecule of a neighboring sheet along the c -axis. The exact positions of the four oxygen atoms can be written in terms of the cell vectors,

$$\begin{aligned} \mathbf{B}_1 &= \frac{1}{3}\mathbf{A}_1 + \frac{2}{3}\mathbf{A}_2 & \mathbf{B}_2 &= \frac{2}{3}\mathbf{A}_1 + \frac{1}{3}\mathbf{A}_2 + \frac{1}{2}\mathbf{A}_3 \\ \mathbf{B}_3 &= \frac{1}{3}\mathbf{A}_1 + \frac{2}{3}\mathbf{A}_2 + u\mathbf{A}_3 & \mathbf{B}_4 &= \frac{2}{3}\mathbf{A}_1 + \frac{1}{3}\mathbf{A}_2 + \left(\frac{1}{2} + u\right)\mathbf{A}_3, \end{aligned} \quad (5.2)$$

where u is defined by the condition that each O-atom has the same distance to its neighbors. Hence one gets an expression for this parameter, $u = \frac{1}{4} + \frac{a^2}{3c^2}$. With these definitions the nearest neighbor distance is about 2.8 \AA . Fig. 5.1 shows the unit cell of ice Ih. It should be noted that, whereas only proton ordered molecules are shown in the picture, in reality there is proton disorder and the H-bonds are distributed at random.

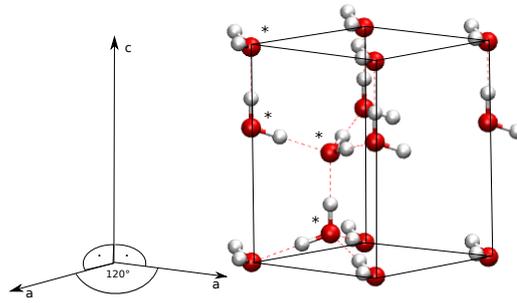


Figure 5.1: Unit cell of ice Ih. Molecules, which belong to the cell, are marked with a star. Note, that the structure is proton ordered, whereas ice Ih is proton disordered in reality.

5.2. Ice Ic

Ice Ic has a cubic unit cell with lattice constant $c = 6.358 \text{ \AA}$ ($90^\circ, 90^\circ, 90^\circ$) and the symmetry group $Fd\bar{3}m$.⁴⁶ There are eight oxygen atoms in the unit cell, which have the diamond structure:

$$\begin{aligned}
 \mathbf{B}_1 &= \frac{1}{4}\mathbf{X} + \frac{1}{4}\mathbf{Y} + \frac{1}{4}\mathbf{Z} & \mathbf{B}_2 &= \frac{3}{4}\mathbf{X} + \frac{3}{4}\mathbf{Y} + \frac{1}{4}\mathbf{Z} \\
 \mathbf{B}_3 &= \frac{1}{2}\mathbf{X} + \frac{1}{2}\mathbf{Z} & \mathbf{B}_4 &= \mathbf{X} + \frac{1}{2}\mathbf{Y} + \frac{1}{2}\mathbf{Z} \\
 \mathbf{B}_5 &= \frac{1}{4}\mathbf{X} + \frac{3}{4}\mathbf{Y} + \frac{3}{4}\mathbf{Z} & \mathbf{B}_6 &= \frac{3}{4}\mathbf{X} + \frac{1}{4}\mathbf{Y} + \frac{3}{4}\mathbf{Z} \\
 \mathbf{B}_7 &= \frac{1}{2}\mathbf{X} + \frac{1}{2}\mathbf{Y} + \mathbf{Z} & \mathbf{B}_8 &= \mathbf{X} + \mathbf{Z}
 \end{aligned} \tag{5.3}$$

In contrast to ice Ih all four hydrogen bonds are staggered, which leads to six-membered rings in chair form in any of the tetrahedrally arranged planes. The nearest neighbor distance is about 2.8 \AA like in ice Ih. In Fig. 5.2 the unit cell of ice Ic is shown, again in a proton ordered version, whereas in reality it is proton disordered.

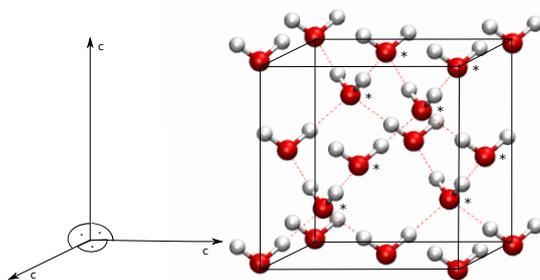


Figure 5.2: Unit cell of ice Ic. Molecules belonging to the cell are marked with a star. Note, that the structure is proton ordered, whereas ice Ic is proton disordered in reality.

5.3. Algorithm for Ice Generation

To generate disordered ice Ih and ice Ic for the subsequent simulations, I follow the guidelines of Ref. (47): A periodic lattice of the desired structure (Ic or Ih) of oxygen atoms is built. Even if there is no proton order in ice, each molecule must fulfill the so-called ice-rule. This means that each oxygen must have exactly two chemical bonded hydrogens, while there are other two hydrogens, which are hydrogen bonded. So a hydrogen atom is set between each two neighboring oxygens at one of the two allowed sites at random to build a chemical bond with the particular oxygen. Because in this stage an oxygen can have up to four chemical bonds, a Monte Carlo (MC) algorithm is applied to correct the hydrogen configuration according to the ice-rule: At each MC step an oxygen-oxygen axis is picked at random and the hydrogen sitting on that axis is shifted with a certain probability. Denoting the chemical coordination of the two oxygens with c_i and c_j , the hydrogen is shifted with probability 1 towards the other oxygen, if the difference of coordination, $|c_i - c_j|$, decrease. If there is no change, the probability is 0,5. Otherwise the shifting is rejected. This procedure repeats until each oxygen has only two chemical bonds, i.e. $c_i = 2$. To get a spherical cluster of the ice configuration, I simply cut a sphere of sufficient size out of this structure. However, the resulting ice droplet is not a realistic one, because dangling surface atoms can lead to a strong electrostatic interaction. These surface atoms should recombine in a subsequent MD simulation with energy minimization. Besides, the H-O-H angle of a water molecule, which is placed into the lattice by this algorithm, is 109.47° , the perfect tetrahedral angle. But the angle of the TIP4P₂₀₀₅ model of water is 104.52° . This discrepancy also should be corrected after a few steps by the algorithm of the MD program handling the constraints.

6. Determination of the Order Parameter

6.1. Size of the Largest Cluster

As mentioned above it is necessary to introduce an order parameter Φ to describe the states, between which a transition occurs. If this transition is an activated process, there is a free energy barrier, which can be expressed by

$$\Delta G(\Phi) = -k_B T \ln[P(\Phi)], \quad (6.1)$$

where $P(\Phi)$ is the probability to find the order parameter around the given value Φ . One has to distinguish between local and global order parameters. Global order parameters only measure the overall degree of crystallinity in the system, i.e. how much of the new phase is present in the system, but not how it is distributed in space. Local order parameters are able to identify regions of more or less structural order by inspection of the environment of each particle.

There are some disadvantages in using global order parameters in nucleation studies:

First of all one has to consider the question, whether one large cluster of the new phase is forming within the system or many smaller ones comprising the same number of particles. As usually assumed, it is more favorable with respect of the free energy to build only one cluster, because then the surface area is smaller. But in fact, as ten Wolde showed,⁴⁸ the entropy can play an important role as it increases with an increasing number of nuclei. This entropy contribution competes with the interfacial free energy and at an early stage of the nucleation process the entropic term is larger. However, ten Wolde also showed that the entropic break-up into smaller nuclei is dependent on the system size. Hence it is possible that a, with respect to CNT, post-critical cluster may still be precritical for entropic reasons. Global measure of crystallinity effectively induces the entropic break-up and is for this reason inappropriate when comparing the results to CNT.⁴⁹

Secondly, because of the underlying activated process, in computer simulations nucleation must be driven by use of the order parameter. Reinhardt and coworkers pointed out that when driving the nucleation process by global order parameters the pathways show an unphysical behavior.⁴⁹ Especially if the lowest free energy path is dynamically slow, a faster pathway, but perhaps one with a higher free energy, will be observed, because its barrier is negated by the rare event method. This is due to the fact that by using global order parameters a single particle can be influenced by solid clusters, which are far away, even though such a long-range interaction is unnatural.

In CNT the free energy barrier is a function of the cluster size distribution N_n . Hence using

a similar order parameter is an obvious advantage if one wants to compare results with CNT. I decided to use the size n of the largest crystalline cluster within the system as suggested by ten Wolde and Frenkel.⁵⁰ Because one is able to decide for each particle whether it belongs to a solid cluster or not, this order parameter is local in its nature. In an umbrella sampling scheme it is possible to control the size of one cluster directly and to sample all cluster sizes with equal accuracy.⁵⁰ To take into account only the largest cluster is in agreement with the cluster size distribution in CNT. It is justified by the fact that cluster building is a rare event and the cluster size distribution N_n can be approximated as the probability $p_n(1)$ to find only one n -cluster in the system (see also appendix A).

Of course, first we need a proper criterion to distinguish between solid- and liquid-like particles. So in the next section I give some ideas to identify the structural order of a system. While CNT provides an expression for the work of embryo formation based on the assumption that the embryo has the same properties as the bulk system, the following determination takes a microscopic and more fundamental view on the system.^{14,51,52} Furthermore this derivation does not need any estimation about the macroscopic properties of the embryo. This approach is very suitable for computer implementation due to the fact that in a computer simulation the exact positions and velocities of all particles are known at all times.

6.2. Structural Order Parameters

Several approaches exist to distinguish ordered crystal structures from disordered liquid structures or gas phases. Methods strongly depend on the studied substances, since for example their solid phases can have different crystal structures. It also depends whether the system is finite or bulk, because in the case of a finite system, particles on the surface maybe must be treated in a different way. For this reasons I discuss the cluster criterion already here in a very concrete manner, namely for a nano-sized water/ice system. Under ambient conditions water crystallizes in an ice Ih structure. But it is often found that also ice Ic is formed in finite systems. Both structures build tetrahedral networks, in which the oxygen atoms form with their four neighboring oxygens an angle of approximately 109° . The order parameter q_t takes advantage of this feature⁵³⁻⁵⁵:

$$q_t(k) = 1 - \frac{3}{8} \sum_{i=1}^3 \sum_{j=i+1}^4 \left(\cos \theta_{ikj} + \frac{1}{3} \right)^2. \quad (6.2)$$

Here k indicates the index of the central oxygen, i and j are the indices of two of the neighboring oxygens. This order parameter is normalized in such a way that it results in a value of $q_l(k) = 1$ for perfect tetrahedral configurations, while its expectation value $\langle q_l(k) \rangle$ vanishes for gas-like configurations. In bulk systems this order parameter can be used to distinguish ice and liquid water.^{55,56} However, it breaks down at differentiation of ice and LDA-glass, because of a too strong overlap of its distribution function.⁵⁶ Unfortunately cluster with LDA structure arise just when liquid water is cooled, i.e. angles of 109° build without forming an ice structure. A second disadvantage is the requirement of the positions of four neighbors. For the nanodroplet studied here approximately $2/3$ of the molecules are situated on the surface and hence do not have four neighboring molecules. Therefore one would have to omit the larger part in the calculation.

In ice, however, the tetrahedral structure is well defined, whereas in water or LDA structures it occurs randomly. This is used by the following method, which goes back to Steinhardt et al..⁵⁷ Local orientational bond order parameters $q_l(k)$ are defined for each molecule, whose $2l + 1$ complex components are given by

$$q_{lm}(k) = \frac{1}{N_{nb}(k)} \sum_{j=1}^{N_{nb}(k)} Y_{l,m}(\hat{r}_{kj}), \quad (6.3)$$

where k and j are molecule indices and N_{nb} is the number of neighboring molecules. $Y_{l,m}(\theta_{kj}\phi_{kj}) = Y_{l,m}(\hat{r}_{kj})$ is the spherical harmonic for the radial unit vector \hat{r}_{kj} between molecule k and j . The question whether a molecule is a neighbor of another is mostly answered by the specification of a cutoff radius. Many authors use this order parameter in its rotational invariant forms and with a specific variable l , also for studying water⁵⁵:

$$q_l(k) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^l |q_{lm}(k)|^2} \quad (6.4)$$

and

$$w_l(k) = \frac{\sum_{m_1+m_2+m_3=0} \begin{pmatrix} l & l & l \\ m_1 & m_2 & m_3 \end{pmatrix} q_{lm_1}(k) q_{lm_2}(k) q_{lm_3}(k)}{\left(\sum_{m=-l}^l |q_{lm}(k)|^2 \right)^{3/2}}, \quad (6.5)$$

where the term in the large parentheses is the Wigner-3j symbol. It was noted, that averaged local bond order parameters give an even more accurate determination of crystal structures.⁵⁸

Thereby one averages over the local orientational order vector $q_{lm}(k)$ of particle k and its surrounding particles.

$$\tilde{q}_{lm}(k) = \frac{1}{\tilde{N}_b(k)} \sum_{n=0}^{\tilde{N}_b(k)} q_{lm}(n). \quad (6.6)$$

The sum runs over all neighbors of particle k plus the particle k itself. The rotational invariant expressions are then calculated with this $\tilde{q}_{lm}(k)$. This way not only the first shell around particle k is considered, but also the second, and the accuracy of distinction of different structures is increased. On that score $\tilde{q}_{lm}(k)$ is a similar order parameter like the normalized dot product $a_l(k, j)$ between each pair of neighboring molecules. $a_l(k, j)$ is a measure for the similarity of the environment of the particular molecules and has defined values for special crystal structures. For disordered systems one obtains a wide distribution of values.⁵⁹

$$a_l(k, j) = \frac{q_l(k) \cdot q_l(j)}{|q_l(k)| |q_l(j)|} \quad (6.7)$$

$$= \frac{\sum_{m=-l}^l q_{lm}(k) q_{lm}^*(j)}{\left(\sum_{m=-l}^l q_{lm}(k) q_{lm}^*(k) \right)^{1/2} \left(\sum_{m=-l}^l q_{lm}(j) q_{lm}^*(j) \right)^{1/2}}. \quad (6.8)$$

Starred arguments indicate complex conjugate values.

6.3. The Solid Cluster Criterion

In the following I use a modified version of an algorithm by Reinhardt et al. to distinguish solid TIP4P₂₀₀₅ water particles from liquid ones.⁴⁹ This algorithm goes back to Moore et al., who applied it with slightly different parameters on an one-atom representation of water (mW model).⁵⁶ The algorithm uses Equ. (6.3) with a cutoff radius of 3.2 Å to identify neighboring molecules. This distance is the first minimum in the pair distribution function for the TIP4P₂₀₀₅ model, based on the oxygen-oxygen distance.^{27,60} (Reinhardt et al. use a cutoff radius of 3.5 Å.) Now one has to decide, which value of l is the best for applying Equ. (6.8). Moore and coworkers tested different values for l and found the best results for distinction of liquid and solid water at $l = 3$ and $l = 4$ ⁵⁶ (but also $l = 6$ can be found at the literature)⁶¹. Therefore, they chose $l = 3$ because it involves the least number of calculations. In Fig. 6.1 one can see the probability distribution of $a_3(k, j)$ for bulk systems. For these diagrams MD simulations in the canonical assemble were executed in a rectangular box with periodic boundary conditions. The electrostatic interaction was realized with Particle Mesh Ewald (PME) and the temperature was regulated with a Nosé-Hoover thermostat. For the distribution of ice Ic, I

simulated 10 ns with 216 molecules at 100 K and every 100 ps I extracted a configuration. For ice Ih, I used 432 molecules, apart from that I proceeded with the same parameters. Liquid water was first produced by melting an ice Ic configuration at 320 K. Furthermore 300 configurations contribute to the statistics here. Indeed the diagram shows a wide range of values for liquid water, while there are accurate peaks for ices Ic and Ih. Ice Ic has only one sharp peak at about $a_3 = -1$. This is because all oxygens in ice Ic have a staggered arrangement of intermolecular bonds with respect to their four neighbors. The difference to ice Ih is that there are only three staggered bonds and one, which is parallel to Ih c-axis, is eclipsed. That eclipsed bond corresponds to the peak at around $a_3 = -0.11$.

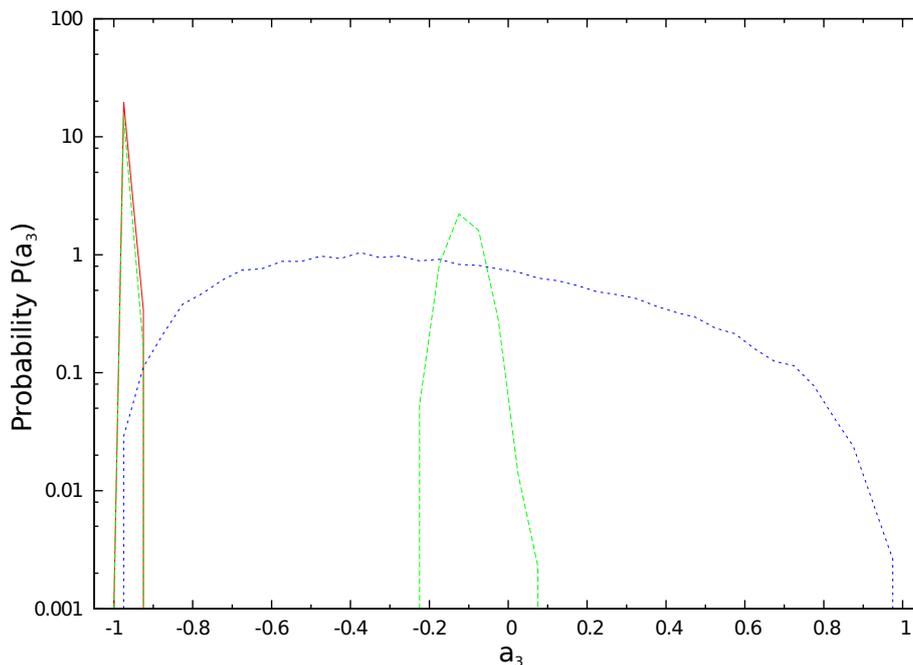


Figure 6.1: Probability distribution of $a_3(k, j)$ for liquid water (blue), ice Ic (red) and ice Ih (green). Note that the y-scale is logarithmic. Also note, that these are the probability distributions of bulk systems.

On this basis Reinhardt et al. classify each molecule according to its orientation with respect to its neighbors:⁴⁹ A molecule will belong to an ice phase, if it has at least three neighbors with a staggered or with an eclipsed bond. If two molecules have such a bond, they are called connected. To allow fluctuations, a value of $a_3 < -0.825$ is counted as a staggered bond. An

eclipsed bond is established if $-0.23 < a_3 < 0.01$:

$$n_{connections}(i) = \sum_{j=1}^{N_{neighs}(i)} \Gamma(a_3(i, j)) \geq 3 \quad (6.9)$$

where

$$\Gamma(x) = \begin{cases} 1 & \text{if } [(x < -0.825) \vee (-0.23 < x < 0.01)], \\ 0 & \text{otherwise.} \end{cases} \quad (6.10)$$

Again, the distribution of a_3 in Fig. 6.1 is for the bulk system. Because in a finite system the surface molecules are not fully bonded they are able to move and rotate more freely. This is the reason, why the distribution of a_3 for finite systems (see Fig. 6.2) does not have such sharp peaks. Here 216 molecules were arranged in a spherical shape and in the particular structure. MD simulations were executed for ice at 50 K and for water at 200 K. In each case altogether 5000 configurations were used for the statistics, extracted in 2 ps steps. One can see that there is still a much higher probability at around $a_3 = -1$ and, for ice Ih, a peak at $a_3 = -0.11$ is visible. So I decided to apply the same criteria as explained above to distinguish solid and liquid particles, with one exception: Because of having a finite system with many molecules located on the surface (or on the surface of a solid embryo), one needs a weaker criterion to count them as a solid molecule as well. So I decided to also count molecules as solid like, which have only two connections, but at least one neighbor with three connections. All molecules, which do not satisfy these criteria belong to a liquid or amorphous phase.

Once one has identified all molecules to be solid or liquid, it is easy to assign them to different clusters: If two solid-like molecules are neighbored, they belong to the same solid cluster.

With this choice I analyzed the above mentioned clusters consisting of 216 water molecules. For the statistics the same configurations were used as for Fig.6.2. In Fig. 6.3 one can see the average composition of molecules within the clusters. Altogether about 55% of the molecules of an ice cluster at 50 K are detected as belonging to ice. Even if that ratio does not seem to be very high, one has to keep in mind, that about $2/3$ of the molecules are located at the surface and can move more freely. Thus, these molecules have to be counted as liquid and form a liquid layer around the cluster. Without the weaker criterion (only two bonds, but at least one neighbor with minimum three bonds) only about 40% of an ice cluster is identified as ice. Each of the surface molecules must be treated as liquid and it would be impossible to figure out, if nucleation occurs on the surface. At 200 K almost all molecules are identified

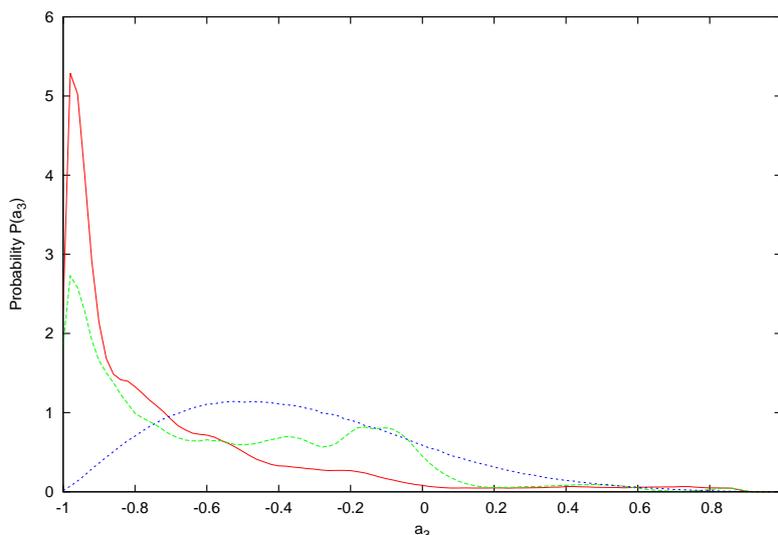


Figure 6.2: $a_3(k, j)$ - probability distribution of the 216 molecules containing nanodroplet for liquid water (blue), ice Ic (red) and ice Ih (green).

as amorphous, but about 9% are detected as 2- or 3-coordinated molecules. It seems, that the cluster is completely melted already and only random molecules are identified as solid (see also Fig. 6.5). This could be due to the fact that liquid water forms similar tetrahedral networks as ice.

In Fig. 6.4 a typical snapshot of an ice Ic cluster is shown. Indeed all inner molecules are identified as ice molecules, while some molecules at the very surface are liquid molecules. Nevertheless there exist some solid molecules at the surface, which is an evidence, that the weaker criterion works. For an ice Ih cluster there is a similar outcome.

At next I analyzed whether the order parameter, i.e. the largest crystalline cluster size within the droplet, can distinguish between systems at clearly separated temperatures. For this purpose I undertook a MD simulation at 130 K and another one at 200 K. Both simulations started with an ice Ic configuration and ran for 10 ns. The time evolution is shown at Fig. (6.5). On the one hand at 200 K the crystalline cluster shrinks within a time of 500 ps. Then the largest crystalline cluster size fluctuates about a value of 6. A snapshot of such a droplet shows, that only few molecules are identified as solid, but no structural order is visible. On the other side at 130 K, though the crystalline cluster also melts, the inner region of the droplet stays

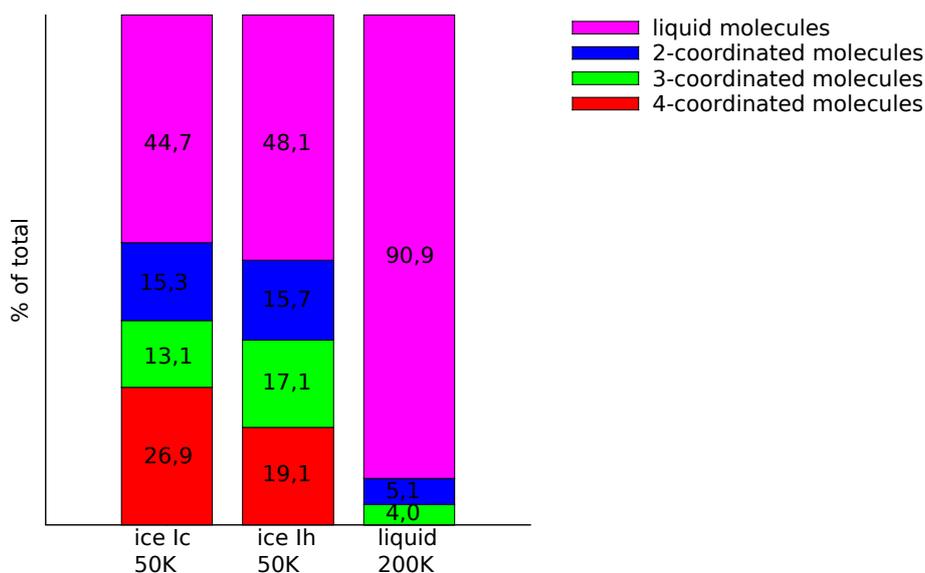


Figure 6.3: Statistics of the composition of a cluster consisting of 216 molecules. 4-coordinated molecules build four staggered bonds or three staggered bonds and one eclipsed, respectively. They can be seen as true ice Ic or ice Ih molecules. 3-coordinated molecules only have three bonds and 2-coordinated ones have two, but at least one neighbor with minimum three bonds. All others are liquid or amorphous molecules, respectively.

crystalline and the order parameter fluctuates around a value of 80. There is no overlap of the order parameter at these two temperatures and it seems that the largest crystalline cluster size is a proper way to distinguish these systems.

To conclude this section I give a brief summary of the solid cluster criterion:

- Two molecules are neighbored, if their distance $r_{O-O} < 3.2 \text{ \AA}$.
- If two neighbored molecules have a staggered or an eclipsed bond, they are connected. This can be measured using the normalized dot product $a_3(i, j)$.
- If a molecule has at least three connections, it is solid.
- If it has only two connections, it might be solid as well if it has at least one neighbor with three connections. This way also surface molecules can be identified as solid like.
- All other molecules are liquid.

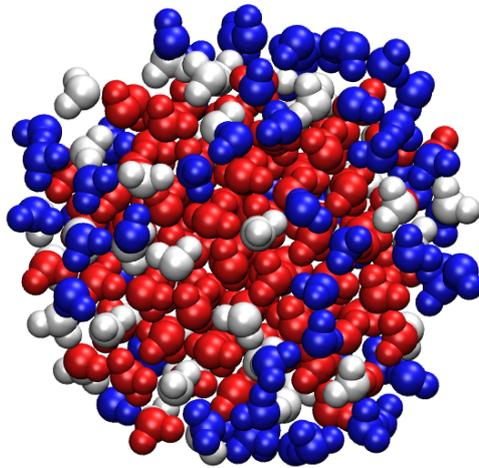


Figure 6.4: Snapshot of an ice Ic droplet. Red molecules are identified as solid with aid of the criterion provided by Reinhardt et al., i.e. these molecules have at least three connections. White molecules are defined as solid surface molecules, having only two connections, but at least one neighbor with three connections. Blue molecules must be treated as liquid.

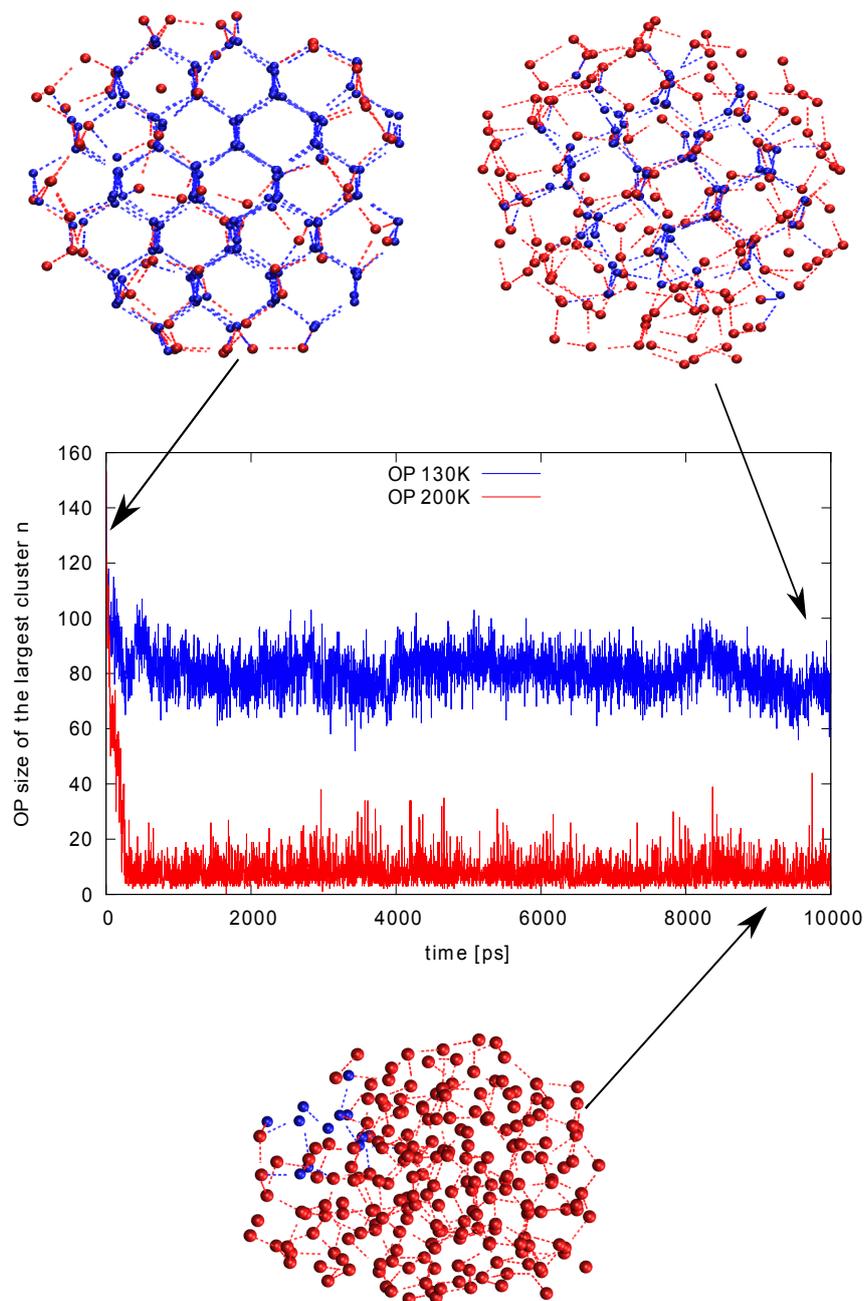


Figure 6.5: Time evolution of a 216-molecule cluster at 130 K (blue line) and 200 K (red line). Red balls in the snapshots represent the liquid oxygens, while blue ones represent the ice like oxygens. Dotted lines represent the hydrogen bonds.

7. Identifying of the Melting Point

7.1. Methods

In the thermodynamic limit and at normal pressure ice melts at about 273 K.⁶² One can see in Fig. 4.1 that the phase diagram of the TIP4P₂₀₀₅ model differs from experiment in a shift to lower temperatures and higher pressures. So also the melting temperature T_m of ice Ih (which is the thermodynamic most stable phase) is lower, at about 252.1 K²⁷.

Furthermore the melting point depends on the system size. In finite systems the curved surface is leading to the Gibbs-Thomson effect⁶³. The positive interfacial energy σ between liquid and solid phase is responsible for small particles having an increased vapor pressure and thus also a decreased melting temperature. For spherical particles this can be derived from Equ. 2.25 after rearranging. Note, that the names of the variables change here: T_m must identified with the bulk melting temperature T_m^{bulk} and T with the melting temperature $T_m(R)$ of the spherical cluster. It leads to the Gibbs-Thomson relation:

$$\Delta T_m(R) = T_m^{bulk} - T_m(R) = T_m^{bulk} \frac{2\sigma}{\Delta h \rho_s R}. \quad (7.1)$$

R is the radius of the droplet, σ is the solid-liquid interface energy, Δh is the bulk enthalpy of fusion and ρ_s is the density of the solid. The Gibbs-Thomson equation can also be written in a simplified version:

$$T_m(R) = T_m^{bulk} - \frac{K_{GT}}{R - d}, \quad (7.2)$$

where K_{GT} is the Gibbs-Thomson constant, which has different values for different liquids and different geometries. In this equation the parameter d is used to include the fact that a pre-melted layer with thickness d is formed around the crystal, which is in accordance with the impossibility to identify the outer molecules as solid as mentioned in Sec. 6.3, because they re-arrange themselves in order to build more hydrogen bonds.

In a recent work C. Johnston and V. Molinero studied the system size dependence of the melting point for the mW water model⁴. They melted spherical ice cluster consisting of 13824, 4235, 2149, 915, 417 and 159 molecules and fitted the resulting melting temperatures, which they identified as the inflection point in the fraction of ice versus temperature, with Equ. 7.2 (see Fig. 7.1). Thus the Gibbs-Thomson constant was calculated as $K_{GT} = (82 \pm 5)$ K nm and the thickness of the pre-melted layer as $d = (0.26 \pm 0.05)$ nm. Adopting the Gibbs-Thomson relation with these parameters to a droplet with $R = 1.1$ nm, which is approximately the radius for 216 molecules, one gets a melting temperature of $T_m = 176$ K. Of course this value is only

a rough estimation, because already the melting temperature of the bulk system in mW model is different (274 K) than in TIP4P₂₀₀₅ model, but it can be used for a first orientation.

Another way to determine the melting temperature of nano-sized water droplets was used by D. Pan and coworkers⁶⁴. They melted nano-crystals of different sizes with TIP4P model, which is very similar to TIP4P₂₀₀₅, and observed a jump in the potential energy once the crystal was melted. This jump reflects a first-order phase transition, because according to the Ehrenfest classification, such a transition exhibits a discontinuity in a first derivative of the free energy, as for example the potential energy. The temperature at which this jump occurred was identified as the melting temperature. Also they confirmed the Gibbs-Thomson relation with slightly different constants. It should be noted, that the jump also occurs when a liquid cluster is frozen. Because the crystallization is a first-order phase transition a hysteresis can appear, which means that there are metastable phases which cause the freezing temperature to be lower than the melting temperature.

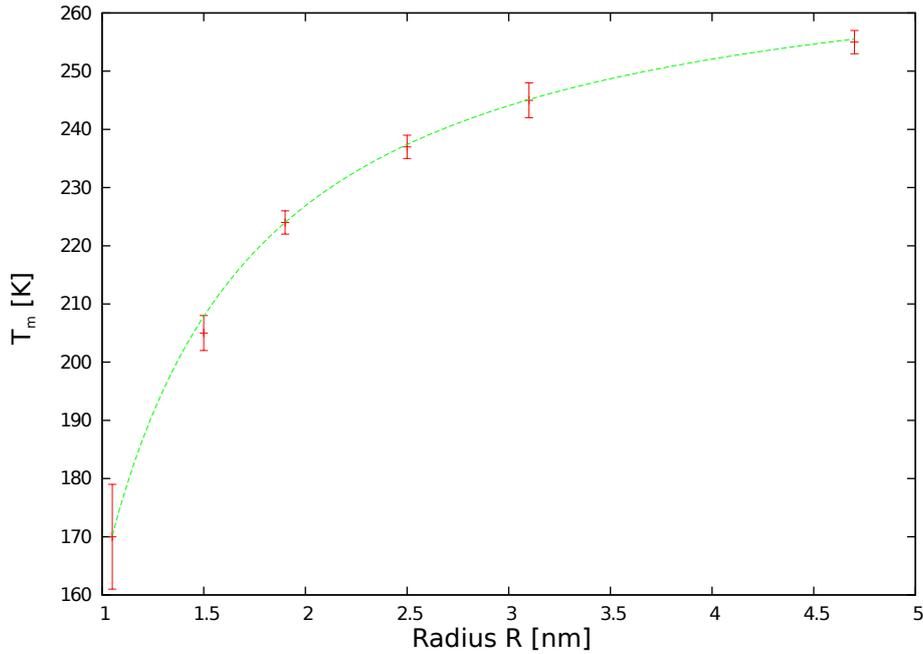


Figure 7.1: Melting temperatures in dependence of system size. Data (red) are from Ref. (4). $K_{GT} = (82 \pm 5) \text{ K nm}$ and $d = (0.26 \pm 0.05) \text{ nm}$ are calculated from least squared fit (green) via Gibbs-Thomson Equ. (7.2). $T_m^{bulk} = 274 \text{ K}$ for the mW model.

7.2. Caloric Curve and Largest Crystalline Cluster

To determine the melting point of a water droplets consisting of 216 and 432 molecules with the TIP4P₂₀₀₅ model I follow the above ideas. I monitor the potential energy while simulating an ice crystal at different temperatures and calculate the fraction of ice within the cluster. While the fraction of ice-like molecules could be a proper indication of melting, but instead I use the above mentioned order parameter, namely the largest crystalline cluster within the droplet. In fact, this is a very similar quantity.

For this purpose I undertook several MD simulations at constant temperature, starting both with ice Ic and ice Ih, and equilibrated the systems at temperatures from 100 K up to 200 K in steps of 10 K. Depending on the temperature, it can last differently long until the system is equilibrated. Indications for an imperfect equilibration can be fluctuations of the potential energy as well as of the order parameter. As time step I used $dt = 2$ fs, the time constant for the Nosé-Hoover thermostat was $\tau = 20$ ps. Furthermore the cut-off radius for both electrostatic and van-der-Waals forces was $r_{cut} = \infty$. To circumvent the problem of flying ice cubes (the energy of high-frequency modes is drained into low-frequency modes, which means that the thermal motion within a system is frozen while the center of mass gets translational and rotational energy, which is an artifact of numerical integration and violates the principle of equipartition of energy),⁶⁵ I removed the translational and rotational energy of the center of mass every 100 time steps. This should be rarely enough to let the physical movement of the system undisturbed. I calculated the potential energy every 100 timesteps (0.2 ps) and saved the actual configuration, i.e. positions and velocities of the molecules, every 1000 time steps (2 ps). With these data it was possible to determine the average values of the potential energy and of the largest crystalline cluster within the whole time range and to give an approximation of error via the standard deviation.

In Figs. 7.2 and 7.3 the time evolution of the order parameter at selected temperatures is shown, starting both with ice Ic and ice Ih configurations. The common ground of both diagrams is that at 130 K the crystalline core of the cluster stays stable at about 60-70 molecules. Besides, at temperatures above 150 K the crystalline part shrinks rapidly and the whole droplet can be seen as amorphous. On the one hand the crystallinity of ice Ic at 140 K-150 K conserves for about 150 ns, but finally solid clusters dissolve, too. On the other hand at ice Ih these solid clusters fluctuate, but stay alive at about 20-40 molecules, at least in the observed time range of 250 ns. It seems that solid and liquid phases are not well separated and the transition in respect to the temperature is quite undefined. Actually, this is a well-known fact for finite systems. The melting transition does not occur at one certain temperature, but within a temperature range and generally, the smaller the system is, the broader is this range.⁶

The results of the order parameter and potential energy calculations for an original ice Ic cluster are collected in Fig.7.4. The potential energy is quite linear from 100 K to 140 K with a slowly increasing slope. At about 150 K a discontinuity shows up and the potential energy becomes linear with temperature again in the range of 160 K to 200 K, but with an enhanced slope. This linearity of potential energy around the melting point and its abrupt increase at the melting temperature can be confirmed by other, both experimental and computational works about nano-clusters of water.^{6,64,66} There is a small discrepancy on the fact that at 150 K, where the jump in the potential energy occurs, the order parameter already has reached its minimum value. Actually one would expect that the jump in the potential energy occurs at the same time as the crystallinity within the cluster vanishes. Perhaps it can be explained thereby that, though no larger crystalline cluster can be formed at 140 K, there are still solid-like bonds in the system, which will be destroyed at higher temperatures. This in turn leads to the observed discontinuity of the potential energy. Even if it is not possible to define the melting temperature precisely because of the above mentioned finite size effects, the observed data, especially the developing of the order parameter, suggest that the melting of the ice Ic cluster occurs at about (140 ± 10) K.

In Fig.7.5 the same dependence on temperature is shown for an original ice Ih cluster. While the caloric curve is almost identical to that of the ice Ic cluster, the transition from larger crystalline clusters to an amorphous structure is smeared out. At 150 K there is still a crystalline core, even if it consists of only about 25 molecules. Anyway, to estimate the melting temperature I fitted the data with the following function, which describes a two-state transition:

$$g(T) = \frac{A}{1 + \exp(-B \cdot (T - C))} + D. \quad (7.3)$$

In this equation T is the temperature, A , B , C and D are fit parameters and especially C is the inflection point of the function. The temperature at this inflection point can be interpreted as melting temperature and it shows that it is very close to 140 K, too. So I am able to give the same approximation for the melting temperature as for ice Ic. It is remarkable that one can clearly see in this diagram the stronger increase of the potential energy once the order parameter has reached its minimum value. This is also an evidence of observing the melting transition.

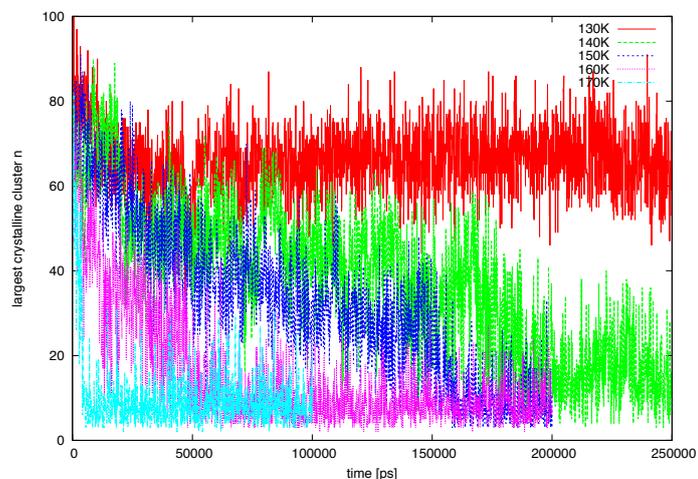


Figure 7.2: Time evolution of the order parameter at selected temperatures for a starting ice Ic configuration. At 170 K the system is evolved until 100 ns, because it already can be seen as equilibrated. Due to the same reason, I evolved the system at 160 K and 150 K for 200 ns

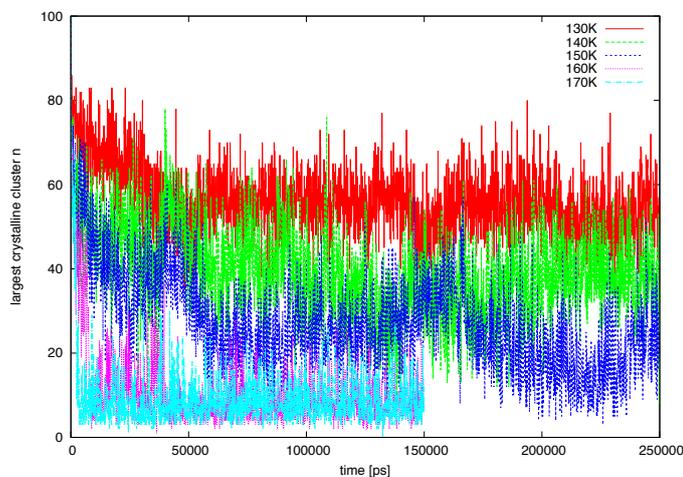


Figure 7.3: Time evolution of the order parameter at selected temperatures for a starting ice Ih configuration. At 160 K and 170 K I evolved the system for 150 ns, because then it already can be seen as equilibrated.

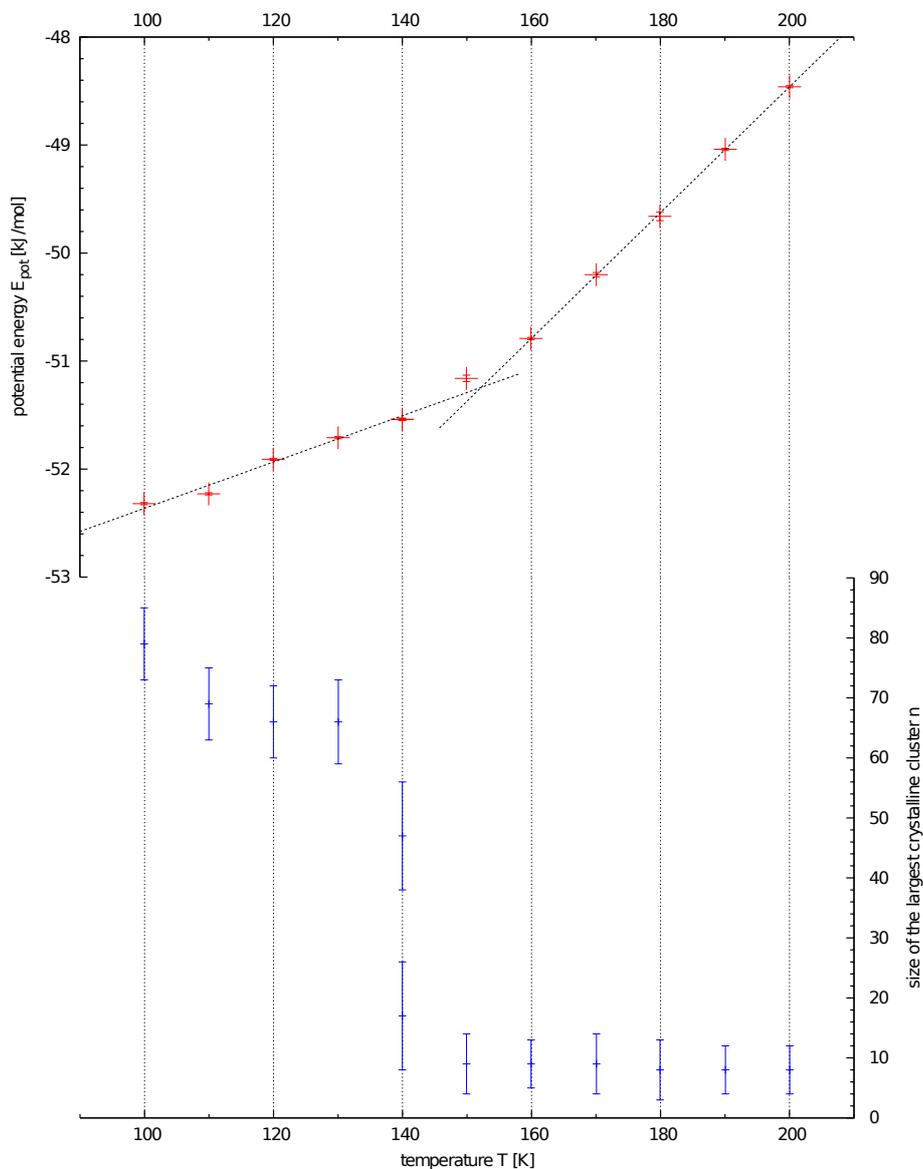


Figure 7.4: Potential energy per molecule and largest crystalline cluster of an original ice Ic droplet as function of temperature. Note that the error bars of the potential energy are mostly smaller than the data-symbols. The lines are just guides for the eyes to demonstrate the linearity. At 140 K I plotted two data points of the order parameter, because there was a crystalline core in the system as long as about 150 ns before it finally shrank to about 15 molecules.

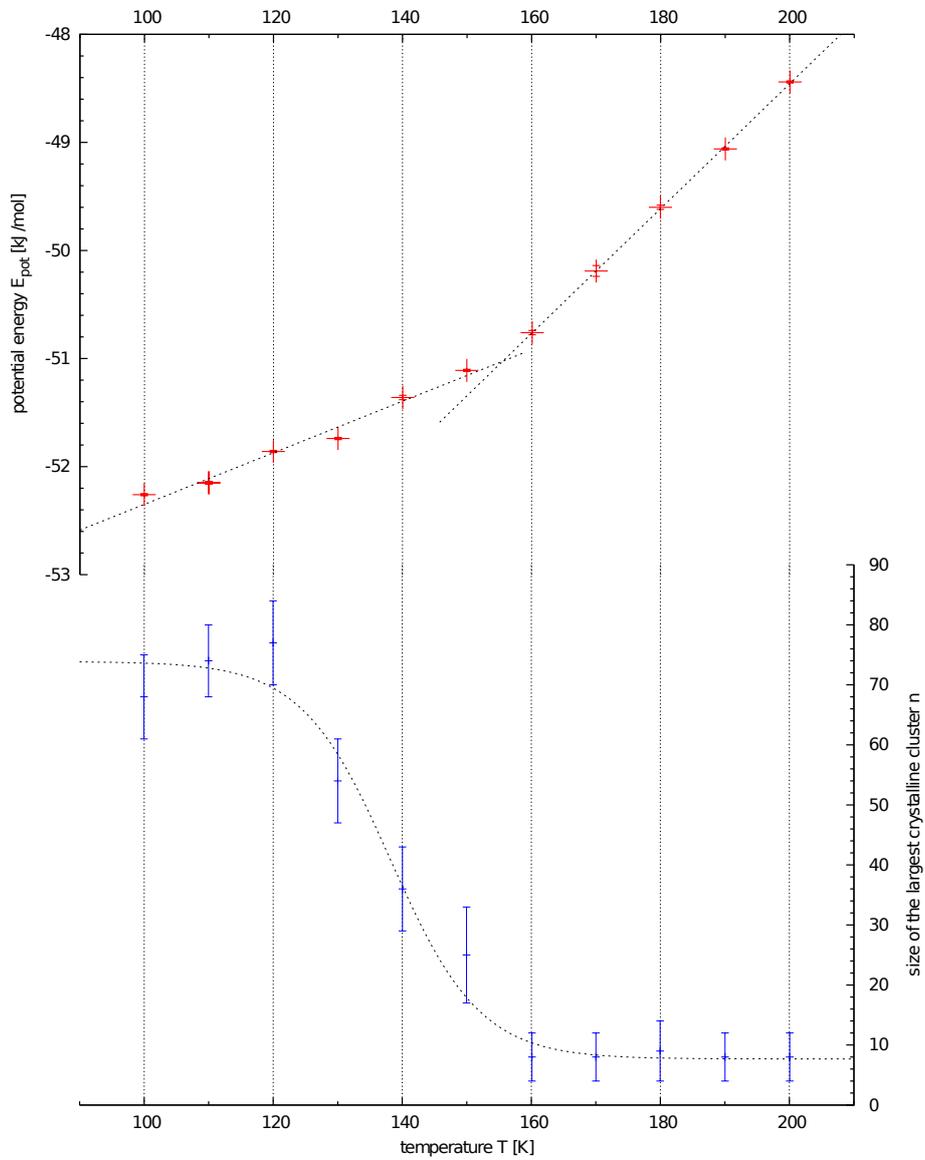


Figure 7.5: Potential energy per molecule and largest crystalline cluster of an original ice Ih droplet as function of temperature. Note that the error bars of the potential energy are mostly smaller than the data-symbols. The lines are just guides for the eyes to demonstrate the linearity. The dotted line in the plot of the order parameter is calculated using Equ. 7.3 to get an inflection point.

In a third set of MD-simulations I analyzed the same quantities for an larger water cluster with 432 molecules, starting with ice Ih configurations. In Fig.7.6 the time evolution of the order parameter at selected temperatures is shown. This time one can see a clear separation of stable and unstable arrangements at a given temperature. While from 120 K-160 K the system equilibrates with a stable crystalline core of over 150 molecules, this crystalline core completely vanishes at higher temperatures within 20 ns - 120 ns. This confirms again the above statement that finite size effects become more and more irrelevant with larger systems and the melting point can be defined more precisely. It also underlines that there must be a minimal size of the crystalline core, here about 150 molecules, to establish a stable equilibrium of the solid and the liquid phase. Fig.7.7 shows the potential energy and the order parameter as function of temperature for such larger systems. For comparison the potential energy of a 216-molecules system is also plotted. Obviously the potential energy per molecule of the larger system lies lower than that of the smaller one. That is again because of the finite size effects and the ratio of surface and volume, which decrease with increasing system size with inverse radius. In other words, inner lying molecules of large systems feel more attractive forces because they are surrounded by more molecules than in smaller systems. Apart from that, there is again a linear relation between potential energy and temperature and the slopes are quite parallel to them of the 216-molecules system. Unfortunately it is not visible that the jump of the potential energy occurs at higher temperatures, what we would expect actually. But on the basis of the order parameter one is able to identify the melting transition at about (165 ± 5) K. So the trend of the Gibbs-Thomson equation 7.2 can be confirmed, i.e. the larger the systems are the higher is the melting temperature and it finally converges to the bulk melting temperature at about 273 K.

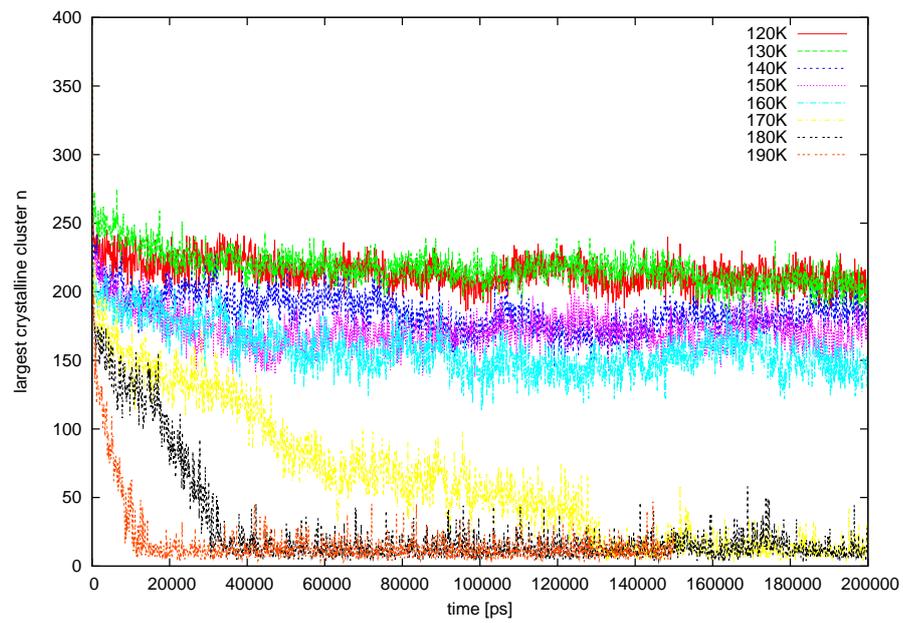


Figure 7.6: Time evolution of the order parameter at selected temperatures for a starting ice Ih configuration with 432 molecules.

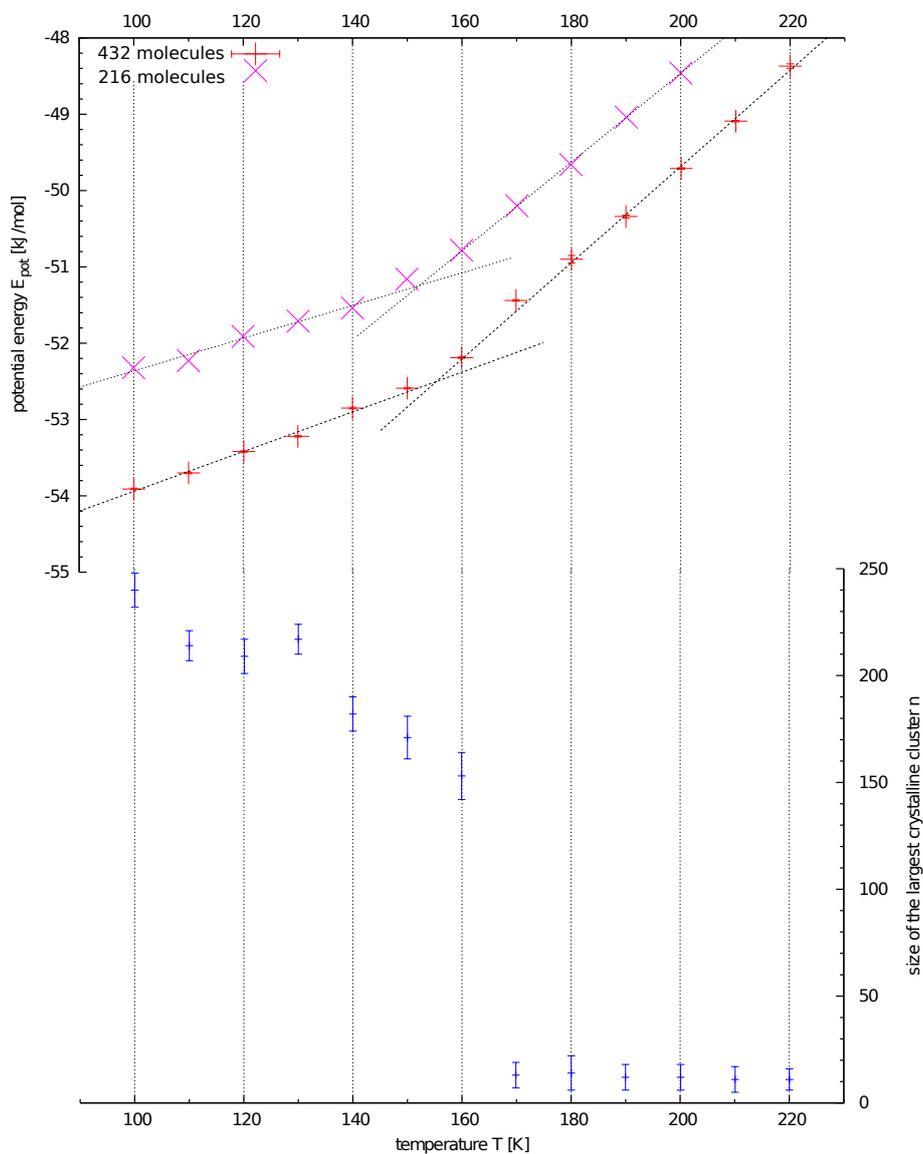


Figure 7.7: Potential energy per molecule and largest crystalline cluster of an original ice Ih droplet with 432 molecules as function of temperature. Note that the error bars of the potential energy are mostly smaller than the data-symbols. The lines are just guides for the eyes to demonstrate the linearity. For comparison the caloric curve of the 216-molecules system is added. The melting temperature seems to be defined more accurately and the transition occurs at higher temperatures than in Figs. 7.4 and 7.5.

8. Calculation of the Free Energy Barrier

8.1. Details of Umbrella Sampling with 216 molecules

For the umbrella sampling simulations with 216 molecules I studied the order parameter range from $n = 0$ to about $n = 110$. I divided this range into 11 windows, where the centers of the windows are located at $n = 5, 15, 25, \dots, 95, 105$. The starting configuration for each window were selected by melting an ice cluster and picking a configuration with the appropriate solid cluster size out of this trajectory. The spring constant of the bias potential was chosen to be $\kappa = 0.05$ kJ/mol, so the overlaps of the biased probability distributions were large enough. For the dynamics I used the MD package GROMACS²⁵. At each call of the MD program a small trajectory consisting of 10 steps with a time step of 3 fs to 4 fs was calculated, depending on temperature and number of window. The particular choice is to ensure that the acceptance probability of the HMC moves is between 20% and 70%. For high temperatures and window numbers this cannot be achieved always and the probability falls down to about 10%. Replica exchange was enabled on average every 100 HMC moves. Mostly the acceptance of replica exchange attempts reaches a value between 10% and 20%, again depending on temperature and window. Altogether umbrella sampling simulations for 8 different temperatures were executed, starting at 110 K, increasing in 10 K steps until reaching 180 K. Each simulation consists of at least 3.5×10^6 HMC moves. After every HMC move the resulting information, especially the order parameter of the actual configuration was written down in a log-file to be able to build the biased histogram of the largest crystalline cluster size. With the help of these histograms and Equ.(3.28) the free energy in each window can then be computed up to a constant. These shifting constants were determined with WHAM²⁶, getting the free energy landscape with respect of the order parameter.

8.2. Results

In Fig. (8.1) and Fig. (8.2) the results of the free energy calculations are shown. At the temperatures 180 K to 160 K there exists only one minimum at $n \approx 6$, which can be interpreted as liquid state. This is consistent with the results of the molecular dynamic simulations, which already showed that at these temperatures a crystalline cluster melts and that the value of the order parameter fluctuates around $n = 6$. In Fig.(8.3) one can see that the distribution of the order parameter in the MD simulation at 180 K agrees very well with the unbiased distribution of US simulations at the same temperature. This can be interpreted as evidence that the free energy in this range of temperatures is described correctly, because both simulation methods

give the same results. The minimum of the free energy is the more accurate, the higher the temperatures are and the free energy increases faster for larger n . At 150 K one can see a second minimum arising. The free energy barrier between these minima is still very small and the liquid state represents the global minimum. Nevertheless, this new minimum could originate from a crystalline core conserved at 150 K for a long time, which is confirmed by the MD simulation.

The free energy curves for the temperatures 140 K to 110 K offer the picture that there is one solid stable state. The crystalline core consists of about $n = 60$ molecules at 110 K and decreases slowly at higher temperatures, reaching a size of $n = 40$ at 140 K. This crystalline state is separated from the liquid state by a energy barrier at about $n = 20$. The free energy difference between both states is about $5k_B T$ at 110 K whereas it shrinks to about $1k_B T$ at 140 K.

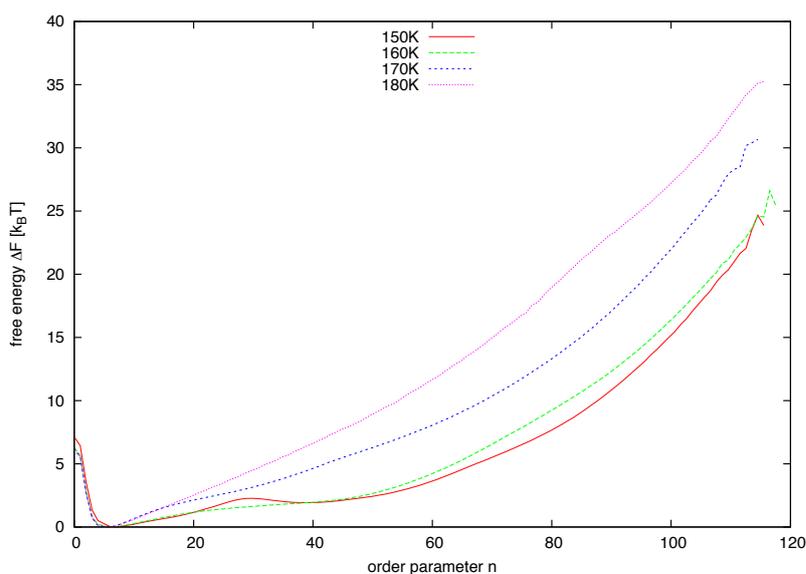


Figure 8.1: Free energy curves for four high temperatures (180 K - 150 K). As above n is the largest crystalline cluster size.

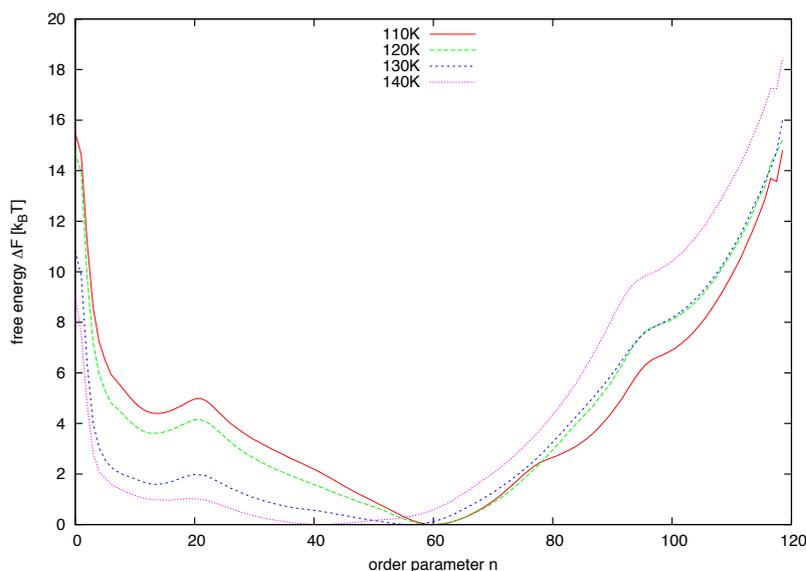


Figure 8.2: Free energy curves for four low temperatures (140 K - 110 K). As above n is the largest crystalline cluster size.

8.3. Problems and Proposals for Solution

Though the result qualitatively fits with the results of the MD simulation - until 140 K a crystalline core is conserved - they must be critically questioned. Especially the height of the energy barrier seems to be too small. Because if the height is correct, then one should see a phase transition relatively often. Particularly the transition from the liquid to the solid state should be easy to realize, because the energy barrier on this site only amounts maximal $1k_B T$, which corresponds to the size of thermal fluctuations. However, I never observed such a transition in a MD simulation starting at an amorphous phase. Even in the US simulations, where a quasi-transition is possible via the replica exchange mechanism, the transition is not observed. Instead of this, configurations of low and high lying windows mainly remain among themselves (see also Fig.(8.4)). For this sampling problem there could be several reasons: First of all, the dynamics at low temperatures are very slow. At low temperatures, molecules hardly can rearrange themselves and are so to speak frozen, even if they do not appear crystalline order. A measure for that is the self-diffusion constant. It can be determined from

$$\lim_{t \rightarrow \infty} \langle \|r_i(t) - r_i(0)\|^2 \rangle = 6Dt, \quad (8.1)$$

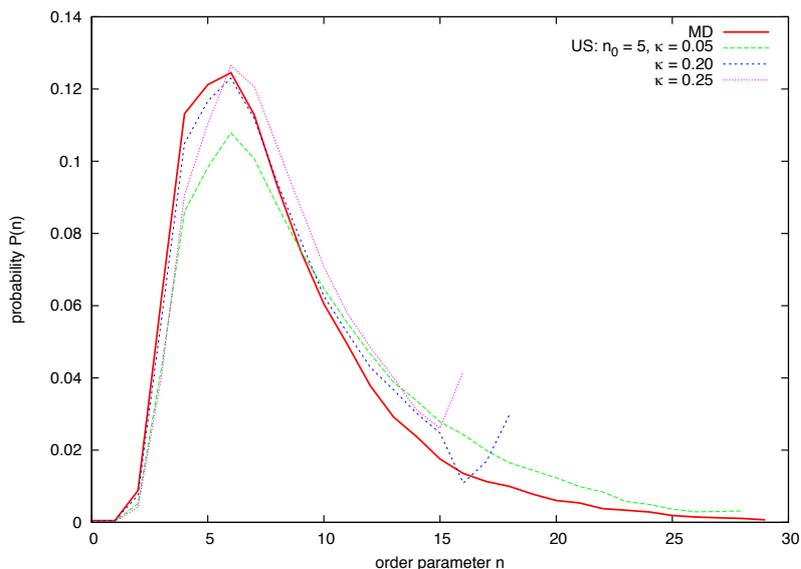


Figure 8.3: Probability distribution of the order parameter at 180 K for both MD and US simulations. Several US simulations with umbrella center at $n_0 = 5$ and different umbrella constants κ were performed to show that they give the same results as the MD simulation. For the statistics 25000 configurations of a 100 ns lasting MD simulation are used. The umbrella simulations consisted of 1.5×10^6 HMC moves and the biased histograms were recalculated via Equ. (3.25). All histograms are normalized.

where $\langle \|r_i(t) - r_i(0)\|^2 \rangle$ is the mean square displacement of particle i at time t . This expression is proportional to t and the self-diffusion constant D can be determined from its slope. For its calculation I cooled down liquid configurations to the particular temperature and measured the mean square displacement over 10 ns. The configurations did not crystallize. The results are shown in Tab.(8.1). In the considered temperature range the self-diffusion constants already differ by a factor 10^2 . In comparison to the bulk value at 278 K, thus slightly above the melting temperature of bulk water, the value at 180 K lies lower by a further factor 10^3 . So the movement of the molecules at low temperatures indeed is very reduced. Liquid configurations cannot completely rearrange themselves to build crystalline networks, at least at the considered timescale. This means that the slow dynamics prevent the crystallization, even if the energy barrier is low. Accordingly, the kinetic pre-factor in the equation for the nucleation rate must be very small. Looking at the replica exchange moves of the US simulation this assumption is confirmed. In Fig.(8.4) the replica exchange moves of window 0 ($n_0 = 5$) and window 10 ($n_0 = 105$) are shown. On the one hand there are accepted exchange moves also for low temperatures. But on the other hand a for example original window 0 configuration stays at lower window numbers and never climbs up to higher order parameters. This behavior improves upon increasing temperature and at 180 K there is quite good intermixing, at least after a sufficiently large number of HMC moves. Unfortunately at this temperature a solid phase is not stable anymore.

120 K	140 K	160 K	180 K	bulk 278 K
$(3 \pm 2) \times 10^{-5}$	$(4 \pm 1) \times 10^{-4}$	$(4 \pm 1) \times 10^{-4}$	$(5 \pm 2) \times 10^{-3}$	1.27

Table 8.1: Selfdiffusion coefficients (in $10^9 D \text{ m}^2/s$) at different temperatures. The bulk value for the TIP4P₂₀₀₅ model at 278 K is cited by Ref.(27). It is very close to the experimental value of $1.313 \times 10^{-9} \text{ m}^2/s$ at the same temperature. The values of the finite system are calculated by a linear fit over 10 ns of a MD simulation.

Secondly, the missing observation of the transition could also have the cause that the order parameter is not a good reaction coordinate. This means that there is a further coordinate being orthogonal to the original order parameter that must be considered, too. This situation is illustrated in Fig.(8.5). The order parameter n separates both minima of the free energy, but the reaction from state A to state B must be described with a second parameter q . On the one side the US simulation can force the system to move along the n -axis, but on the other side the actual transition also requires a shift on the q -axis, for which a further energy barrier must be overcome. Thus the overall barrier of the reaction between liquid and solid phase would be

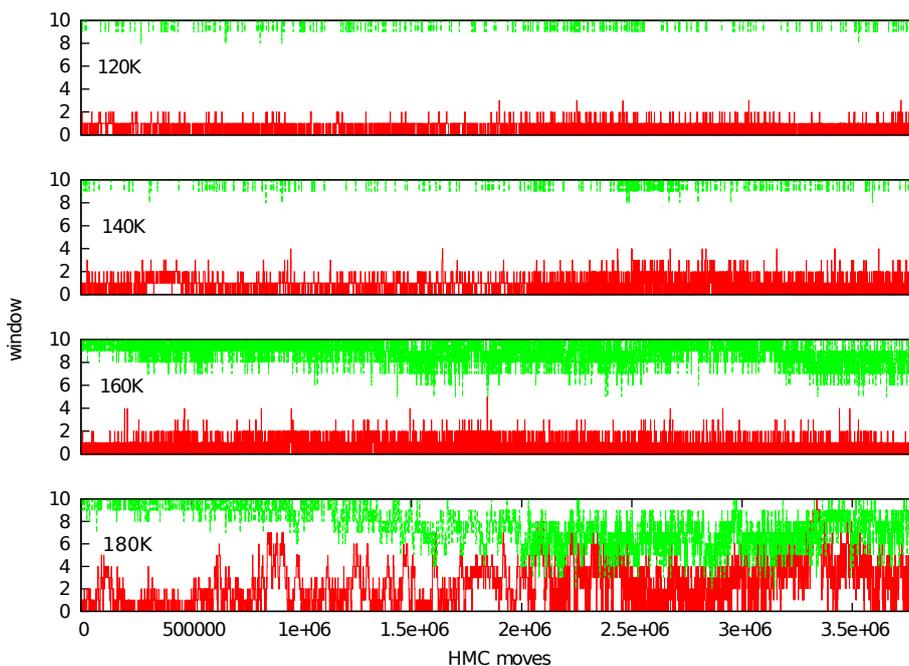


Figure 8.4: Replica exchange moves at different temperatures of original window 0 and window 10 configurations as function of Hybrid Monte Carlo steps(HMC).

larger.

But also if n is a suitable reaction coordinate in general, there could be difficulties in some cases as one can also see in Fig. 8.5. The position of the free energy minima with respect to the order parameter defines stable and metastable states in a unique way, but such states normally possess attractive basins. It means, if the system is close to a stable state, it will reach this state in the near future with a large probability. But if these basins do overlap, the energy barrier cannot be determined correctly. To test this, I evaluated the biased histograms for each window separately using Equ. (3.28), instead of calculating the global energy function with WHAM. As mentioned above the curves of successive windows should roughly be parallel, because $F(n)$ of different windows differs only by a constant. As one can see in Fig.(8.6) this is true in most of the windows. But indeed, at temperatures, at which a maximum in the free energy is formed (110 K-150 K), the free energy curves of the relevant windows are non-parallel. This is evidence of the above assumption and shows that the height of the free energy barrier has not been calculated correctly. Though both minima give information about the coexistence of liquid and solid state, nothing can be said about the height of the barrier and consequently also nothing about which of both states is the stable or metastable one, respectively.

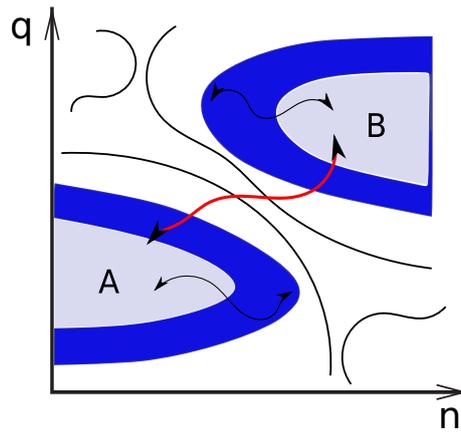


Figure 8.5: Possible free energy landscape as function of two order parameters n and q . A and B are two stable regions with attractive basins around them (dark blue area). Instead of moving along the right reaction path (red arrow), the US technique samples just along the n -axis, i.e. the largest crystalline cluster size (black arrows). Consider also the possible overlap of attractive basins with respect to n .

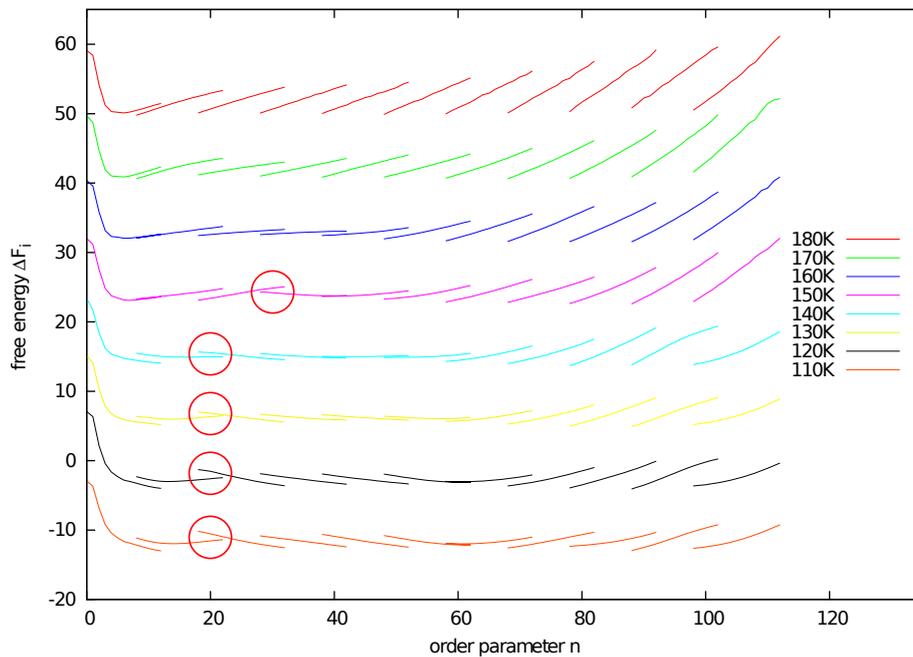


Figure 8.6: Free energy in each window at all studied temperatures. For a clear view, the free energy values of the particular temperatures are shifted on top of each other. The critical overlaps are marked with red circles.

Hence there are two separate problems in the calculation of the free energy. The first is related to the slow dynamics, which cause a bad intermixing of the windows in the replica exchange mechanism. The second is the improper choice of the order parameter, because the two stable phases are too close to each other which effects the overlap of their basins. In the first case one could improve the dynamics with parallel tempering not only in the different windows of the order parameter, but also in temperature as described in chapter 3. This is exactly what E. Mendez-Villuendas did for the free energy calculation in the freezing of gold nano-clusters.⁶⁷ Running the windows in parallel at different temperatures swapping attempts between them are accepted with probability

$$P_{acc}(i \leftrightarrow j) = \min\left(1, \exp[(\beta_i - \beta_j)(E_i - E_j)]\right). \quad (8.2)$$

E_i is the energy of the cluster simulated at temperature T_i whereas E_j is the energy of the cluster simulated at temperature T_j . One could also think of particular Monte Carlo moves to accelerate the sampling rate of the simulation.

The second problem could be solved either by choosing a better order parameter, which describes the reaction in better way, or by keeping the old order parameter and doing multidimensional umbrella sampling, i.e. also in an orthogonal direction as illustrated in Fig. 8.5. For example, Reinhardt et al. found out in their studies concerning crystallization of TIP4P₂₀₀₅ bulk water that molecules partially build chains, which increase the order parameter, but inhibit the crystallization of the entire system. In order to force the whole system to crystallize, they modified their order parameter in such a way that they did not count chain molecules as belonging to the solid cluster.⁴⁹ Unfortunately I do not monitor such chains, so this improvement of the order parameter is not suitable for my problem. But maybe it must be ensured that the crystalline cluster remains quite compact.

8.4. Details of Umbrella Sampling with 432 molecules

For the system with 432 molecules I studied the temperature range from 120 K-180 K. The setup of the US simulation consists of 32 windows, whose centers are located at $n = 5, 15, 25, 35, 45, 50, 55, 60, \dots, 165, 175, 185, 195$. Again, the start configurations originate from a MD melting trajectory of ice Ic. I decided to choose smaller windows in the middle of this range ($\Delta n_0 = 5$) because the barrier is expected to occur in this section and I wanted to produce a better sampling there. Because the centers of the first five and last four windows are wider apart ($\Delta n_0 = 10$), I gave these windows a weaker umbrella constant $\kappa = 0.1$ kJ/mol to ensure that there is a sufficient overlap of the distribution functions. The umbrella constants of the

other windows are $\kappa = 0.25$ kJ/mol. The dynamics are chosen with a timestep of 2 fs and each HMC move consists of 10 MD steps with subsequent acceptance test. This acceptance ratio is between 20% and 60%, whereby the lower boundary is reached, if the free energy curve has a large absolute slope. Replica exchange attempts take place every 300 HMC moves on average. Such attempts are accepted with a probability of 10-15%, but at 120 K and 130 K it can fall down under 10%. Again, this is due to the fact that at these low temperatures configurations cannot rearrange themselves sufficiently fast in order to climb up or down the umbrella windows. At each temperature an US run has 1.7×10^6 HMC moves, except for 150 K-170 K, where 2.7×10^6 HMC moves are used.

8.5. Results and Discussion

In Fig. 8.7 one sees the biased probability distribution at temperature 150 K and, resulting from that, the free energies for each window. In this larger system, the curves $F_i(n)$ are quite parallel for successive windows. Probably this might be because of the wider distance of the solid and the liquid state and because their basins do not overlap. To this end n is a good order parameter. There is a similar outcome for other temperatures.

The results of the free energy calculation with 432 molecules according to WHAM are shown in Fig. 8.8. The height and the location of the barrier in dependence on temperature is listed in Tab. 8.2. Concerning this matter, no trend is observable, the barrier is at about $n = 70$ and its height lies between $6 - 7k_B T$. For higher temperatures the crystalline phase is slightly shifted towards smaller n . For lower temperatures one can see that a further metastable state forms at about $n = 170$.

T [K]	n^*	$\Delta G(n^*) [k_B T]$	P(solid)
120	66	6.62	0.357
130	71	6.19	0.167
140	75	6.65	0.062
150	70	6.16	0.061
160	68	6.64	0.032
170	69	6.62	0.016
180	-	-	-

Table 8.2: Location n^* and height ΔG of the free energy barrier. Probability P for the solid state.

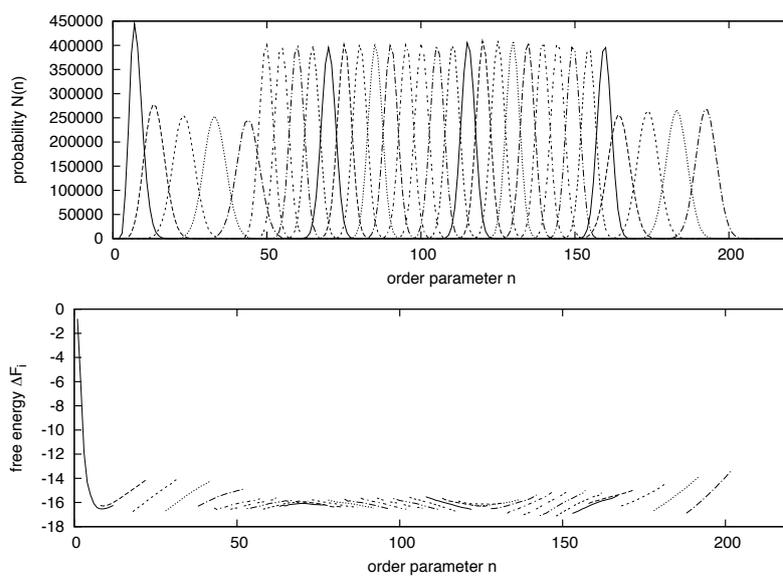


Figure 8.7: Biased probability distribution and free energy at 150 K. Note that the middle windows lie closer, but have a stronger umbrella constant. This time $F_i(n)$ is parallel in successive windows.

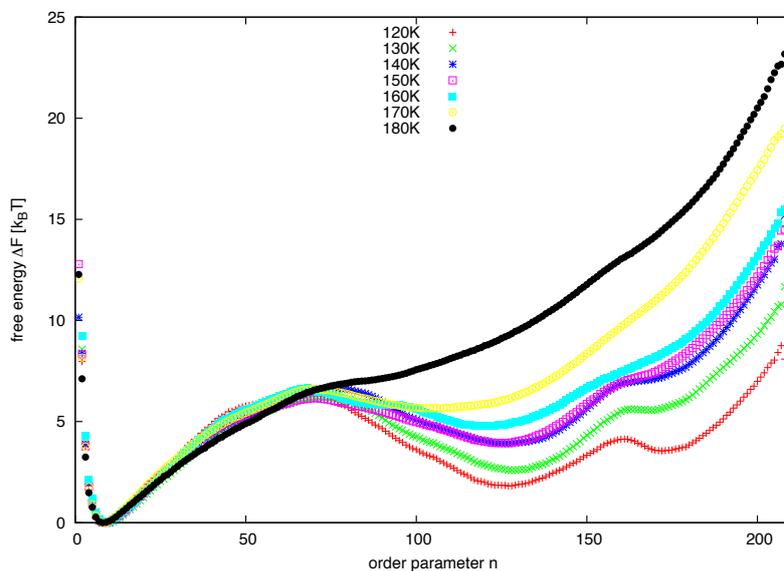


Figure 8.8: Energy landscape for different temperatures with 432 molecules.

It turns out that the liquid state, i.e. the state with an order parameter of $n \approx 9$ has the lowest lying free energy at all temperatures. At 180 K it even is the only minimum in the free energy landscape. However, this minimum is quite narrow. To identify the phase equilibrium it is necessary to integrate $P(n)$ over the stable regions in order to get the total probability of the solid and liquid state, respectively.

$$\begin{aligned} P(\text{liquid}) &= \int_0^{n^*} p(n) dq \\ P(\text{solid}) &= \int_{n^*}^N p(n) dq. \end{aligned} \quad (8.3)$$

The results are collected in Tab. 8.2 as well and it turns out that the equilibrium between the two phases, i.e. where $P(\text{solid}) = P(\text{liquid}) = 0,5$, is not even reached at 120 K. Accordingly, the melting temperature would be below 120 K. To determine the temperature, at which the equilibrium appears, one can make the following approach. At the melting point the free energy is approximated by expanding about T_m up to $O(\Delta T)$.⁶⁸:

$$\Delta F = \frac{\Delta H}{T_m}(T - T_m), \quad (8.4)$$

where ΔH is the enthalpy of fusion. Because $\frac{P(\text{solid})}{P(\text{liquid})} = \exp[-\beta\Delta F]$ and in addition $P(\text{solid}) + P(\text{liquid}) = 1$, the probability of the solid phase can be expressed as:

$$P(\text{solid}) = \frac{1}{1 + \exp\left(\frac{\Delta h}{k_B T_m} \frac{(T - T_m)}{T}\right)}. \quad (8.5)$$

In Fig. 8.9, $P(\text{solid})$ was fitted with Equ.(8.6). $P(\text{solid}) = 0,5$ is reached at a temperature of about 115 K. This temperature corresponds to the inflection point of the fit function. A further fit parameter is ΔH , the enthalpy of fusion. It turns out that $\Delta H/N = 0.03$ kJ/mol, which is more than a factor 10^2 smaller than the corresponding bulk value for the TIP4P₂₀₀₅ model ($(\Delta H/N)^{\text{bulk}} = 4.857$ kJ/mol).³⁸ Because of the finite size effects the transition is not sharp, but smeared out. Challa et al. give an approximation for the transition region, which

can be expressed as⁶⁸

$$\Delta T = \frac{4k_B T_m^2}{\Delta H}. \quad (8.6)$$

Using this approximation gives $\Delta T = 34$ K, which means that the melting point can be estimated to be (115 ± 17) K. This temperature seems to be very low considering that the melting is observed in the MD simulation not until 170 K, which is outside of the melting point uncertainty. Johnston et al. determined the melting point of a 417 molecule cluster with the mW-model to be (205 ± 3) K, which is higher than almost 80%. Nevertheless, one has to keep in mind that this result was not based on a free energy calculation and even the bulk melting temperature is higher in this model ($T_m^{bulk}(mW) = 274$ K). So one simple reason for seeing the crystal being stable up to 170 K in the MD simulation could be the relatively high energy barrier in combination with the wide minimum in the energy landscape. This leads to observing the melting process above and the crystallization process below the equilibrium transition temperature. Besides, it must be noted once more that at these low temperatures the molecules are quite immobile even without being in a crystalline structure. Because of the slow dynamics one probably has to wait a very long time until the crystal begins to melt. Probably this is also the reason why there again are problems with the replica exchange. Though the acceptance lies at 10-15%, a complete intermixing of the replicas does not take place. Parallel tempering in temperature could lead here to an improvement, as well.

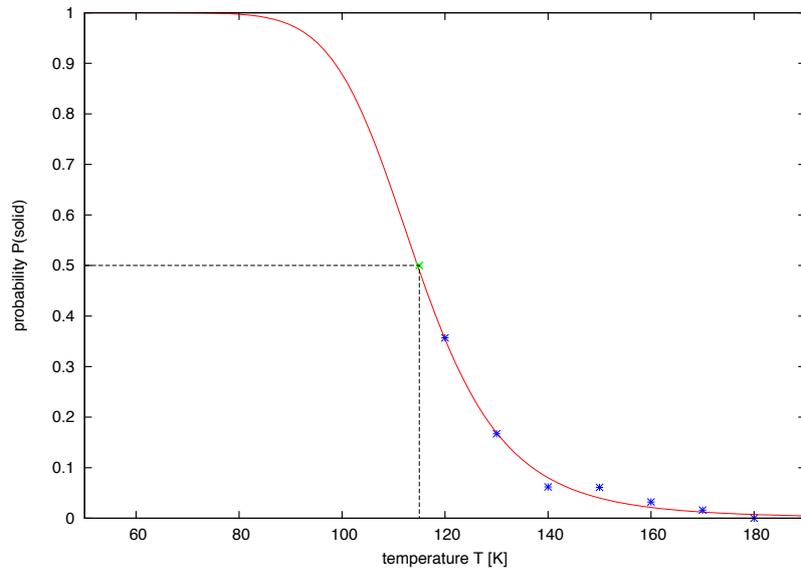


Figure 8.9: Determination of the phase equilibrium. Blue crosses are the simulation data. Red line: fit function according to Equ. (8.6). Green point: melting temperature T_m , $P(T_m) = 0.5$.

9. Conclusion

Based on local orientational bond order parameters it was possible to identify solid- and liquid-like molecules within water clusters. The largest crystalline cluster within the system corresponds with CNT, which provides a mechanism for nucleation in terms of the size of the nucleated embryo.

Analyzing ice clusters, it turns out that on the one hand a crystalline core is conserved, but on the other hand there is an amorphous surface, which is formed because of the rearranging of surface molecules in order to form more hydrogen bonds. This is at the cost of the crystal geometry.

Carrying out MD simulations at constant temperature, I was able to estimate the temperature, at which ice clusters, consisting of 216 and 432 molecules, melt. The larger cluster melts at higher temperatures, confirming qualitatively the Gibbs-Thomson relation.

Umbrella sampling simulations at different temperatures were used to determine the free energy barrier predicted by CNT and to analyze the stability of ice phases. Unfortunately this method breaks down for small water clusters because the order parameter cannot distinguish between the stable phases in a proper way. For a larger cluster the free energy barrier was able to be calculated. Hence it was possible to find the temperature of phase equilibrium.

The principal difficulty in the umbrella sampling simulation was the realization of the replica exchange moves. This behavior must be improved by parallel tempering in temperature in order to overcome local minima in the energy landscape.

A. Free Energy Barrier: Derivation via the Partition Function

To get an expression for the free energy, I start with the partition function of the (N, V, T) -ensemble, i.e. the canonical ensemble. Such a derivation can also be found in works of Reiss et al.⁵² and Mendez-Villuandes.⁵¹ Instead of the Gibbs free energy $\Delta G(n)$, which is the potential of the (N, P, T) -ensemble, this determines the Helmholtz free energy $\Delta F(n)$.

The canonical partition function for a system containing N particles in a volume V at temperature T is given by

$$Q(N, V, T) = \frac{1}{\Lambda^{3N} N!} \int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)], \quad (\text{A.1})$$

where $\beta = 1/k_B T$ and $\Lambda = h/(2\pi m k_B T)^{1/2}$ is the de Broglie wavelength, in which h is the Planck constant and m is the mass of a single particle. $U(\mathbf{r}^N)$ is the potential energy, which is dependent of the coordinates \mathbf{r} of all particles.

Once a proper criterion to identify particles within the cluster is known and one has detected a cluster of n particles, one can define weight functions as follows:

$$\delta_n(\mathbf{r}^n) = \begin{cases} 1 & \text{if all } n \text{ particles belong to the embryo,} \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.2})$$

$$\delta_r(\mathbf{r}^n) = \begin{cases} 1 & \text{if no other particle belongs to the embryo,} \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.3})$$

These special delta functions are connected via the following relation:

$$\delta_r(\mathbf{r}^n) = \prod_{i=n+1}^N [1 - \delta_{n+1}(\mathbf{r}^n, \mathbf{r}_i)]. \quad (\text{A.4})$$

The partition function of a system containing at least one n -cluster is then given by

$$\begin{aligned} Q_n(N, V, T) &= \frac{1}{\Lambda^{3n} n!} \frac{1}{\Lambda^{3(N-n)} (N-n)!} \\ &\times \int d\mathbf{r}^n \delta_n \delta_r \exp[-\beta[U_n(\mathbf{r}^n) + U_\sigma(\mathbf{r}^n, \mathbf{r}^{N-n})]] \\ &\times \int d\mathbf{r}^{(N-n)} \exp[-\beta U_{N-n}(\mathbf{r}^{N-n})], \end{aligned} \quad (\text{A.5})$$

where U_n , U_σ and U_{N-n} are the potential energy within the n -cluster, the interaction energy

A FREE ENERGY BARRIER:
DERIVATION VIA THE PARTITION FUNCTION

between particles in the cluster and the remaining $N - n$ outer particles and the potential energy of the outer particles, respectively. There is one simplification in this formula, because the latter potential energy should be dependent on the location of the cluster within the system. This dependence can be ignored for macroscopic systems and short range forces.⁵² Now one can incorporate the weight function into U_n and U_σ , respectively:

$$\begin{aligned} U'_n &= U_n - k_B T \ln[\delta_n], \\ U'_\sigma &= U_\sigma - k_B T \ln[\delta_r]. \end{aligned} \quad (\text{A.6})$$

Furthermore one can substitute one integration variable \mathbf{r}_n for the center-of-mass \mathbf{R} of the cluster:

$$\begin{aligned} \mathbf{R} &= \frac{\sum_{i=1}^n \mathbf{r}_i}{n}, \\ d\mathbf{r}_1 \dots d\mathbf{r}_n &= n^3 d\mathbf{R} d\mathbf{r}'_1 \dots d\mathbf{r}'_{n-1}. \end{aligned} \quad (\text{A.7})$$

Primed coordinates are relative to the center-of-mass while unprimed coordinates denote a fixed laboratory system. So the partition function becomes

$$\begin{aligned} Q_n(N, V, T) &= \frac{n^3}{\Lambda^{3n} n!} \frac{1}{\Lambda^{3(N-n)} (N-n)!} \\ &\times \int d\mathbf{R} \int d\mathbf{r}'^{n-1} \exp\left[-\beta[U'_n(\mathbf{r}'^n) + U'_\sigma(\mathbf{r}'^{n-1}, \mathbf{r}'^{N-n})]\right] \\ &\times \int d\mathbf{r}'^{(N-n)} \exp\left[-\beta U_{N-n}(\mathbf{r}'^{N-n})\right]. \end{aligned} \quad (\text{A.8})$$

Considering a decoupled system, i.e. without interaction between n -cluster and remaining $N - n$ particles, the following two probability densities can be introduced:

$$\Phi_{N-n}^0 = \frac{\exp\left[-\beta U'_{N-n}\right]}{\int d\mathbf{r}'^{N-n} \exp\left[-\beta U'_{N-n}\right]}. \quad (\text{A.9})$$

$$\Phi_n^0 = \frac{\exp\left[-\beta U'_n\right]}{\int d\mathbf{R} \int d\mathbf{r}'^{n-1} \exp\left[-\beta U'_n\right]}. \quad (\text{A.10})$$

The two correspondent partition functions would be

$$Q(N-n, V, T) = \frac{1}{\Lambda^{3(N-n)} (N-n)!} \int d\mathbf{r}'^{(N-n)} \exp\left[-\beta U_{N-n}\right] \quad (\text{A.11})$$

for the remaining $N - n$ particles and

$$\begin{aligned} q_n^*(V, T) &= \frac{1}{\Lambda_n^3} \int d\mathbf{R} \int d\mathbf{r}'^{n-1} \frac{n^{3/2}}{\Lambda^{3(n-1)} n!} \exp[-\beta U'_n] \\ &= \frac{V}{\Lambda_n^3} \int d\mathbf{r}'^{n-1} \frac{n^{3/2}}{\Lambda^{3(n-1)} n!} \exp[-\beta U'_n] \end{aligned} \quad (\text{A.12})$$

for the decoupled n -cluster, where $\Lambda_n^3 = h/(2\pi nm k_B T)^{1/2}$. Using Equ. A.9 through Equ. A.12 in Equ. A.8 gives

$$\begin{aligned} Q_n(N, V, T) &= Q(N - n, V, T) q_n^*(V, T) \\ &\quad \times \int d\mathbf{R} \int d\mathbf{r}'^{n-1} \int d\mathbf{r}'^{N-n} \Phi_n^0 \exp[-\beta U_\sigma] \Phi_{N-n}^0 \\ &= Q(N - n, V, T) q_n^*(V, T) \langle e^{-\beta U_\sigma} \rangle_0 \\ &= Q(N - n, V, T) q_n(V, T), \end{aligned} \quad (\text{A.13})$$

where the average $\langle \cdot \rangle_0$ is taken in the decoupled system. The partition function $q_n(V, T)$ of the cluster now includes its interaction with the outer particles by the factor $\langle e^{-\beta U_\sigma} \rangle_0$. Furthermore it can be written as

$$\begin{aligned} q_n(V, T) &= q_n^*(V, T) \langle e^{-\beta U_\sigma} \rangle_0 \\ &= \frac{V}{\Lambda_n^3} \int d\mathbf{r}'^{n-1} \frac{n^{3/2}}{\Lambda^{3(n-1)} n!} \exp[-\beta U'_n] \langle e^{-\beta U_\sigma} \rangle_0 \\ &= N \frac{v}{\Lambda_n^3} q_n^{internal} \\ &= N q_n(v, T), \end{aligned} \quad (\text{A.14})$$

where the integration over \mathbf{R} gives V , $v = V/N$ is the volume per particle and $q_n^{internal}$ is the partition function of the cluster without regard of its possibility of translation within the system. The latter is expressed by the factor \tilde{v}/λ_n^3 . The probability to find at least one n -cluster in a canonical system can then be written as

$$\begin{aligned} \Pi_n &= \frac{Q_n(N, V, T)}{Q(N, V, T)} = \frac{Q_n(N - n, V, T) q_n(V, T)}{Q(N, V, T)} \\ &= q_n(V, T) e^{\beta n \mu}. \end{aligned} \quad (\text{A.15})$$

The factor $e^{\beta n \mu}$, where μ is the chemical potential of a particle in the remaining system, arises by the fact, that the statistical mechanics of the canonical ensemble is linked to thermodynamics by the relation $Q(N, V, T) = e^{-\beta F}$ and the derivation $dF/dN = \mu$. So $n\mu$ is the part of the

free energy the system would have, if the n particles were “solved” in the remaining system instead of arranged in the cluster.

Since the formation of a cluster is a rare event, the probability Π_n to find at least one cluster can be approximated with

$$\Pi_n = p_n(1) + p_n(2) + \dots \approx p_n(1), \quad (\text{A.16})$$

where $p_n(i)$ is the probability to find exactly i n -cluster. This approximation is based on the fact that the formation of the clusters is uncorrelated and hence $p_n(i) \approx [p_n(1)]^i$. Thus it is allowed to neglect higher order terms. By the same reason the average number N_n of n -clusters can be approximated:

$$N_n = p_n(1) + 2p_n(2) + 3p_n(3) + \dots \approx p_n(1) = \Pi_n. \quad (\text{A.17})$$

Finally, one gets Equ. 2.8 (with ΔF instead of ΔG), again by the relation $q_n(v, T) = e^{-\beta f_n(v)}$:

$$\begin{aligned} N_n &= q_n(V, T) e^{\beta n \mu} \\ &= N q_n(v, T) e^{\beta n \mu} \\ &= N \exp \left[-\beta [f_n(v) - n \mu] \right] \\ &= N \exp \left[-\beta \left(-\frac{\ln[v/\Lambda_n^3]}{\beta} + f_n^{\text{internal}} - n \mu \right) \right] \\ &= N \exp [-\beta \Delta F]. \end{aligned} \quad (\text{A.18})$$

It should be noted that the contribution from translational free energy is also included in ΔF . This contribution can be identified with the term $-\frac{\ln[v/\Lambda_n^3]}{\beta}$.⁵² Beside the translational degrees of freedom, also rotational degrees of freedom should appear, if they are regarded at the partition function. It was pointed out,⁶⁹⁻⁷¹ that such additional contributions to the free energy are neglected in CNT, although their contribution is considerable. Furthermore f_n^{internal} corresponds with the partition function q_n^{internal} and therefore denotes the free energy inside the cluster. So actually the quantity $N_n/N = P_n$, which is named the “intensive probability”⁵⁰ and which is directly accessible by computer simulations, is related to the free energy via

$$\begin{aligned} N_n/N = P_n &= \exp \left[-\beta \Delta F \right], \\ \Delta F &= -k_B T \ln \left[P_n \right]. \end{aligned} \quad (\text{A.19})$$

Software

Here is a list of the software, which was used for programming, simulation, data analysis and creating figures:

- GCC - version 4.4.7
The GNU Compiler Collection
<http://gcc.gnu.org>
- GROMACS - version 4.5.5
Package for molecular dynamic simulations and data analysis
<http://www.gromacs.org>
- WHAM - version 2.0.7
Implementation of the weighted histogram analysis method by A. Grossfield
<http://membrane.urmc.rochester.edu/content/wham>
- ReExUS
Umbrella sampling code with replica exchange by Andreas Singraber
- gnuplot - version 4.2
Command-line driven graphing utility
<http://www.gnuplot.info>
- Inkscape - version 0.48.4
Vector graphics editor
<http://www.inkscape.org>
- VMD - version 1.9.1
Molecular visualization program
<http://www.ks.uiuc.edu>

References

1. E. M. Knipping *et al.*, “Experiments and simulations of ion-enhanced interfacial chemistry on aqueous NaCl aerosols”, *Science* **288**, 301–306 (2000).
2. B. J. Murray, E. J. Jensen, “Homogeneous nucleation of amorphous solid water particles in the upper mesosphere”, *Journal of Atmospheric and Solar-Terrestrial Physics* **72**, 51–61 (2010).
3. M. Kulmala, “How particles nucleate and grow”, *Science* **302**, 1000–1001 (2003).
4. J. C. Johnston, V. Molinero, “Crystallization, Melting, and Structure of Water Nanoparticles at Atmospherically Relevant Temperatures”, *Journal of the American Chemical Society* **134**, 6650–6659 (2012).
5. V. Buch *, B. Sigurd, J. Paul Devlin, U. Buck, J. K. Kazimirski, “Solid water clusters in the size range of tens-thousands of H₂O: a combined computational, spectroscopic outlook”, *International Reviews in Physical Chemistry* **23**, 375–433 (2004).
6. T. Kaneko, T. Akimoto, K. Yasuoka, A. Mitsutake, X. C. Zeng, “Size-Dependent Phase Changes in Water Clusters”, *Journal of Chemical Theory and Computation* **7**, 3083–3087 (2011).
7. K. Liu *et al.*, “Characterization of a cage form of the water hexamer”, *Nature* **381**, 501–503 (1996).
8. K. Nauta, “Formation of Cyclic Water Hexamer in Liquid Helium: The Smallest Piece of Ice”, *Science* **287**, 293–295 (2000).
9. C. Steinbach *et al.*, “Detection of the book isomer from the OH-stretch spectroscopy of size selected water hexamers”, *Physical Chemistry Chemical Physics* **6**, 3320 (2004).
10. G. Torchet, P. Schwartz, J. Farges, M. F. de Feraudy, B. Raoult, “Structure of solid water clusters formed in a free jet expansion”, *The Journal of Chemical Physics* **79**, 6196 (1983).
11. J. Huang, L. S. Bartell, “Kinetics of Homogeneous Nucleation in the Freezing of Large Water Clusters”, *The Journal of Physical Chemistry* **99**, 3924–3931 (1995).
12. J. P. Devlin, C. Joyce, V. Buch, “Infrared Spectra and Structures of Large Water Clusters”, *The Journal of Physical Chemistry A* **104**, 1974–1977 (2000).
13. C. C. Pradzynski, R. M. Forck, T. Zeuch, P. Slavicek, U. Buck, “A fully size-resolved perspective on the crystallization of water clusters.” *Science* **337**, 1529–32 (2012).

14. P. G. DeBenedetti, *Metastable Liquids: Concepts and Principles* (Princeton University Press, New Jersey, 1996).
15. A. Laaksonen, I. Napari, “Breakdown of the Capillarity Approximation in Binary Nucleation: A Density Functional Study”, *The Journal of Physical Chemistry B* **105**, 11678–11682 (2001).
16. J. L. Katz, H. Wiedersich, “Nucleation theory without Maxwell Demons”, *Journal of Colloid and Interface Science* **61**, 351–355 (1977).
17. H. Ehrenreich, D. Turnbull, *Solid State Physics Vol. 45: Advances in Research & Applications* (Academic Press, 1991).
18. P. Harrowell, D. W. Oxtoby, “A molecular theory of crystal nucleation from the melt”, *The Journal of Chemical Physics* **80**, 1639 (1984).
19. V. A. Shneidman, M. C. Weinberg, “Induction time in transient nucleation theory”, *The Journal of Chemical Physics* **97**, 3621 (1992).
20. L. D. Landau, E. M. Lifshitz, *Statistical Physics* (Pergamon Press, 1958).
21. Y. B. Zeldovich, “On the theory of new phase formation, cavitation”, *Acta Physicochimica URSS* **18**, 1 (1943).
22. V. G. Baidakov, A. O. Tipsev, “Crystal nucleation and the solid-liquid interfacial free energy.” *The Journal of Chemical Physics* **136**, 074510 (2012).
23. D. Turnbull, J. C. Fisher, “Rate of Nucleation in Condensed Systems”, *The Journal of Chemical Physics* **17**, 71 (1949).
24. D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Elsevier Science, San Diego, London, 2001).
25. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, “GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation”, *Journal of Chemical Theory and Computation* **4**, 435–447 (2008).
26. A. Grossfield, “WHAM: the weighted histogram analysis method”, version 2.0.6, <http://membrane.urmc.rochester.edu/content/wham>
27. J. L. F. Abascal, C. Vega, “A general purpose model for the condensed phases of water: TIP4P/2005”, *The Journal of Chemical Physics* **123**, 234505 (2005).
28. V. Molinero, E. B. Moore, “Water modeled as an intermediate element between carbon and silicon.” *The journal of Physical Chemistry B* **113**, 4008–16 (2009).

29. N. Argaman, G. Makov, "Density functional theory: An introduction", *American Journal of Physics* **68**, 69 (2000).
30. W. C. Swope, "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters", *The Journal of Chemical Physics* **76**, 637 (1982).
31. W. Hoover, "Canonical dynamics: Equilibrium phase-space distributions", *Physical Review A* **31**, 1695–1697 (1985).
32. S. Miyamoto, P. A. Kollman, "Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models", *Journal of Computational Chemistry* **13**, 952–962 (1992).
33. W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications", *Biometrika* **57**, 97–109 (1970).
34. S. Duane, A. Kennedy, B. J. Pendleton, D. Roweth, "Hybrid Monte Carlo", *Physics Letters B* **195**, 216–222 (1987).
35. G. Torrie, J. Valleau, "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling", *Journal of Computational Physics* **23**, 187–199 (1977).
36. C. J. Geyer, E. A. Thompson, "Annealing Markov Applications Chain to With Ancestral Inference", *Journal of the American Statistical Association* **90**, 909–920 (1995).
37. S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, "THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method", *Journal of Computational Chemistry* **13**, 1011–1021 (1992).
38. C. Vega, J. L. F. Abascal, M. M. Conde, J. L. Aragones, "What ice can teach us about water interactions: a critical comparison of the performance of different water models", *Faraday Discussions* **141**, 251 (2009).
39. W. S. Benedict, N. Gailar, E. K. Plyler, "Rotation-Vibration Spectra of Deuterated Water Vapor", *The Journal of Chemical Physics* **24**, 1139 (1956).
40. J. L. F. Abascal, E. Sanz, R. Garcia Fernandez, C. Vega, "A potential model for the study of ices and amorphous water: TIP4P/Ice", *The Journal of chemical physics* **122**, 234511 (2005).
41. H. Berendsen, J. Grigera, "The missing term in effective pair potentials", *Journal of Physical Chemistry* **91**, 6269–6271 (1987).

42. H. Yu, W. F. van Gunsteren, “Charge-on-spring polarizable water models revisited: from water clusters to liquid water to ice”, *The Journal of Chemical Physics* **121**, 9549–64 (2004).
43. P. Paricaud, M. Predota, A. a. Chialvo, P. T. Cummings, “From dimer to condensed phases at extreme conditions: accurate predictions of the properties of water by a Gaussian charge polarizable model”, *The Journal of Chemical Physics* **122**, 244511 (2005).
44. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, “Comparison of simple potential functions for simulating liquid water”, *The Journal of Chemical Physics* **79**, 926 (1983).
45. D. Eisenberg, W. Kauzmann, *The Structure and Properties of Water (Oxford Classic Texts in the Physical Sciences)* (Oxford University Press, USA, 1997), p. 318.
46. M. Chaplin, “Water Structure and Science”, <http://www.lsbu.ac.uk/water/index.html>
47. V. Buch, P. Sandler, J. Sadlej, “Simulations of H₂O Solid, Liquid, and Clusters, with an Emphasis on Ferroelectric Ordering Transition in Hexagonal Ice”, *The Journal of Physical Chemistry B* **102**, 8641–8653 (1998).
48. P. R. T. Wolde, “Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations”, *Science* **277**, 1975–1978 (1997).
49. A. Reinhardt, J. P. K. Doye, E. G. Noya, C. Vega, “Local order parameters for use in driving homogeneous ice nucleation with all-atom models of water.” *The Journal of Chemical Physics* **137**, 194504 (2012).
50. P. R. ten Wolde, D. Frenkel, “Computer simulation study of gas-liquid nucleation in a Lennard-Jones system”, *The Journal of Chemical Physics* **109**, 9901 (1998).
51. E. Mendez-Villuendas, *Nucleation in Gold Nanoclusters* (University of Saskatchewan, Master’s Thesis, 2007).
52. H. Reiss, R. K. Bowles, “Some fundamental statistical mechanical relations concerning physical clusters of interest to nucleation theory”, *The Journal of Chemical Physics* **111**, 7501 (1999).
53. P.-L. Chau, A. J. Hardwick, “A new order parameter for tetrahedral configurations”, *Molecular Physics* **93**, 511–518 (1998).
54. J. R. Errington, P. G. Debenedetti, “Relationship between structural order and the anomalies of liquid water”, *Nature* **409**, 318–21 (2001).

55. R. Radhakrishnan, B. L. Trout, "Nucleation of hexagonal ice (Ih) in liquid water", *Journal of the American Chemical Society* **125**, 7743–7 (2003).
56. E. B. Moore, E. de la Llave, K. Welke, D. A. Scherlis, V. Molinero, "Freezing, melting and structure of ice in a hydrophilic nanopore", *Physical Chemistry Chemical Physics : PCCP* **12**, 4124–34 (2010).
57. P. Steinhardt, D. Nelson, M. Ronchetti, "Bond-orientational order in liquids and glasses", *Physical Review B* **28**, 784–805 (1983).
58. W. Lechner, C. Dellago, "Accurate determination of crystal structures based on averaged local bond order parameters", *The Journal of Chemical Physics* **129**, 114707 (2008).
59. P. Rein ten Wolde, M. J. Ruiz-Montero, D. Frenkel, "Numerical calculation of the rate of crystal nucleation in a Lennard-Jones system at moderate undercooling", *The Journal of Chemical Physics* **104**, 9932 (1996).
60. C. Vega, C. McBride, E. Sanz, J. L. F. Abascal, "Radial distribution functions and densities for the SPC/E, TIP4P and TIP5P models for liquid water and ices Ih, Ic, II, III, IV, V, VI, VII, VIII, IX, XI and XII", *Physical Chemistry Chemical Physics* **7**, 1450 (2005).
61. T. Li, D. Donadio, G. Russo, G. Galli, "Homogeneous ice nucleation from supercooled water", *Physical Chemistry Chemical Physics : PCCP* **13**, 19807–13 (2011).
62. R. Feistel, W. Wagner, "A New Equation of State for H₂O Ice Ih", *Journal of Physical and Chemical Reference Data* **35**, 1021 (2006).
63. S. J. Gregg, K. S. W. Sing, *Adsorption, surface area, and porosity* (Academic Press, 1991).
64. D. Pan, M. Liu, B. Slater, A. Michaelides, E. Wang, "Melting the ice: on the relation between melting temperature and size for nanoscale ice crystals", *ACS Nano* **5**, 4562–9 (2011).
65. S. C. Harvey, R. K. Z. Tan, T. E. Cheatham, "The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition", *Journal of Computational Chemistry* **19**, 726–740 (1998).
66. C. Hock, M. Schmidt, R. Kuhnen, C. Bartels, "Calorimetric observation of the melting of free water nanoparticles at cryogenic temperatures", *Physical Review* **103**, 073401 (2009).
67. E. Mendez-Villuendas, R. Bowles, "Surface Nucleation in the Freezing of Gold Nanoparticles", *Physical Review Letters* **98**, 185503 (2007).

68. M. Challa, D. Landau, K. Binder, “Finite-size effects at temperature-driven first-order transitions”, *Physical Review B* **34**, 1841–1852 (1986).
69. J. Lothe, G. M. Pound, “Reconsiderations of Nucleation Theory”, *The Journal of Chemical Physics* **36**, 2080 (1962).
70. V. Ruth, J. P. Hirth, G. M. Pound, “On the theory of homogeneous nucleation and spinodal decomposition in condensation from the vapor phase”, *The Journal of Chemical Physics* **88**, 7079 (1988).
71. H. Reiss, “Translation-Rotation Paradox in the Theory of Nucleation”, *The Journal of Chemical Physics* **48**, 5553 (1968).

Zusammenfassung

Nano-Wassercluster bestehend aus einigen hundert bis tausend Molekülen spielen eine wichtige Rolle in unsere Atmosphäre und im interstellaren Raum. Unter mesosphärischen Bedingungen können sich neben amorphen Phasen, die keine kristalline Ordnung aufweisen, auch kristalline Phasen, wie hexagonales oder kubisches Eis, ausbilden. Es ist eine interessante und aktuelle Frage, welche der Phasen unter diesen Bedingungen die stabilste ist. Das Wissen über den Mechanismus der Eis-Nukleation und über das Wachstum von Nanoteilchen hat große Bedeutung in der Forschung. Vor allem die Rolle, die die Wasserstoffbindungen sowohl im Inneren, als auch an der Oberfläche des Clusters dabei spielen, muss näher beleuchtet werden. Ab welcher Clustergröße sich überhaupt kristalline Strukturen ausbilden können ist ein weiterer Forschungsgegenstand.

In dieser Arbeit beschäftige ich mich mit einigen Aspekten der Eis-Nukleation mit Hilfe von Computer-Simulationen. Im ersten Kapitel wird der theoretische Hintergrund der Nukleation diskutiert, welcher als Klassische Nukleationstheorie bekannt ist. Ursprünglich wurde diese Theorie entworfen, um die Kondensation im unterkühlten Dampf zu beschreiben. Mit einigen Erweiterungen kann sie aber auch auf den flüssig-fest Übergang angewendet werden. Da Nukleation ein Aktivierungsprozess ist, gibt es eine Barriere in der freien Energie, die überwunden werden muss, damit der Phasenübergang stattfindet. Ein mathematischer Ausdruck für diese Barriere kann mit Hilfe der sogenannten Kapillaritätsapproximation abgeleitet werden. Nachdem die verwendeten Simulations- und Analysemethoden erklärt werden, wird ein geeignetes Wassermolekül gewählt. Das TIP4P₂₀₀₅ Modell beschreibt ein Wassermolekül als starren Körper mit vier Wechselwirkungszentren. Einige Aspekte von hexagonalem beziehungsweise kubischem Eis werden im nachfolgenden Kapitel diskutiert. Der praktische Teil dieser Arbeit besteht zunächst in der Ausarbeitung eines geeigneten Ordnungsparameters, der in der Lage ist zwischen festen und flüssigen Molekülen zu unterscheiden. Der gewählte Ordnungsparameter geht zurück auf ten Wolde und Frenkel, die zur Bestimmung des größten kristallinen Kerns innerhalb des Clusters Kugelflächenfunktionen verwenden. Um die Stabilität von Eisphasen in einem Cluster bestehend aus 216 und 432 Molekülen zu untersuchen, werden Molekular-Dynamik Simulationen bei konstanter Temperatur durchgeführt. Dadurch ist es möglich den Temperaturbereich zu schätzen, bei dem kristalline Strukturen zu schmelzen beginnen. Außerdem verwende ich die Umbrella-Sampling-Methode, um die Landschaft der freien Energie in Abhängigkeit des Ordnungsparameters zu ermitteln und vor allem die Barriere der freien Energie zu bestimmen, die von der Nukleationstheorie vorhergesagt wird.

Raimund Hirmer

Ausbildung

- 10/2010–09/2013 **Physik**, *Universität Wien*,
M.Sc..
- 03/2009–09/2010 **Physik**, *Universität Wien*,
B.Sc..
- 03/2008–02/2009 **Physik**, *Humboldt Universität zu Berlin*.
- 10/2006–02/2008 **Biotechnologie**, *Technischen Universität Berlin*.
- 09/1996–07/2005 **Abitur**, *Albrecht Altdorfer Gymnasium Regensburg*.

Wien, 29. Juli 2013
