

B 3 Molecular Dynamics Simulations

R. G. Winkler

Institut für Festkörperforschung

Forschungszentrum Jülich GmbH

Contents

1	Introduction	2
2	Equations of motion	2
3	Interaction Potentials, Force Field	3
3.1	Atomic systems	3
3.2	Molecular systems	4
4	Periodic Boundary Conditions, Truncated Potentials	5
4.1	Periodic boundary conditions	5
4.2	Potential truncation; <i>minimum</i> image convention, neighbor list	6
5	Algorithms for Molecular Dynamics Simulations	7
5.1	Liouville formulation of time-reversible algorithms	8
5.2	Accuracy check, equilibration	10
6	Molecular Dynamics at Various Ensembles	13
6.1	Constant temperature molecular dynamics simulations	13
6.2	Isothermal-isobaric molecular dynamics simulations	22
7	Long-Range Interactions	26
A	Appendices	30
A.1	Extended phase-space partition function	30
A.2	Derivation of the internal pressure	30

1 Introduction

Molecular dynamics (MD) simulation is a technique to compute the equilibrium and transport properties of classical many-body systems. *Classical* means that the basis of the approach are Newton's equations of motion, i.e., quantum mechanical effects are not taken into account explicitly. Atoms are treated within the Born-Oppenheimer approximation, where the nucleus moves according to the laws of classical mechanics and the electrons follow instantaneously. Hence, there is no need to model them explicitly and it is sufficient to account for their average effect by an analytical interaction potential among nuclei.

Molecular dynamics simulations generate information on the microscopic level, in particular, atomic positions and velocities as a function of time, i.e., the complete description of the system in the sense of classical mechanics. This microscopic information has to be translated into macroscopic observables like pressure, heat capacity, diffusion coefficient, etc., by means of statistical mechanics.

Molecular dynamics simulations are exploited in a wide range of applications in fundamental and applied science. The MD method was introduced by Alder and Wainwright in the late 1950's [1, 2] to study the interactions of hard spheres. Many important insights concerning the behavior of simple liquids emerged from their studies. The next major advance was in 1964, when Rahman carried out the first simulation using a realistic potential for liquid argon [3]. The first molecular dynamics simulation of a realistic system was done by Rahman and Stillinger in their simulation of liquid water in 1974 [4]. The first protein simulations appeared in 1977 with the simulation of the bovine pancreatic trypsin inhibitor (BPTI) [5]. Today the literature is full of molecular dynamics simulation results ranging from atomic and solid state physics to soft matter applications [6, 7].

Since then, the number of simulation techniques has greatly expanded. Specialized techniques for particular problems, including mixed quantum mechanical - classical simulations, have been developed [8]. To cover the length- and time-scale gap of complex fluids, hybrid simulation techniques are exploited, where MD simulations are one of the components.

In this contribution the important aspects of molecular dynamics simulations are outlined. It is not a cookbook in the sense that formulas are provided only. There are a number of excellent books available like, e.g., [9–15], which describe the various aspect of simulations in great detail. The current article summarizes the state of the art aspects of simulations, in particular with respect to integration algorithms and isothermal simulations. Of course, the necessary formulas and algorithms are also provided.

2 Equations of motion

The basic dynamical equations of classical mechanics are Newton's equations of motion. For a system of N point particles of mass m_i ($i = 1, \dots, N$) at positions \mathbf{r}_i and velocities $\dot{\mathbf{r}}_i$, they are given by

$$m_i \frac{d^2}{dt^2} \mathbf{r}_i = \mathbf{F}_i. \quad (1)$$

The forces \mathbf{F}_i on particle i are obtained from the potential energy $U(\{\mathbf{r}\})$ via $\mathbf{F}_i = -\nabla_{\mathbf{r}_i} U(\{\mathbf{r}\})$. The solution of these equations provides the complete information of a system for particular initial conditions, say at $t = 0$, $\mathbf{r}_i(0)$ and $\dot{\mathbf{r}}_i(0)$.

Alternatively, Hamilton's equations of motion for the (generalized) momenta \mathbf{p}_i and positions \mathbf{r}_i follow from the Hamiltonian

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\{\mathbf{r}\}) \quad (2)$$

via the canonical equations

$$\dot{\mathbf{r}}_i = \nabla_{\mathbf{p}_i} H, \quad \dot{\mathbf{p}}_i = -\nabla_{\mathbf{r}_i} H. \quad (3)$$

For Cartesian coordinates Hamilton's equations become

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i}, \quad (4)$$

$$\dot{\mathbf{p}}_i = -\nabla_{\mathbf{r}_i} U(\{\mathbf{r}\}) = \mathbf{F}_i. \quad (5)$$

The following considerations are restricted to conservative systems, i.e., there is a potential energy which is independent of time. As a consequence, the total energy of the system $E = H = E_k + U$, where $E_k = \sum_i m_i \dot{\mathbf{r}}_i^2 / 2$ is the kinetic energy, is conserved. Moreover, the systems under consideration are invariant with respect to translations which implies the conservation of the linear momentum, i.e., $\sum_i m_i \ddot{\mathbf{r}}_i = 0$. Typically three dimensional periodic boundary conditions are considered (cf. Sec. 4.1). For such systems the total angular momentum is not a conserved quantity. It is important to ensure that the numerical solution of the equations of motion obeys the conservation laws.

An important aspect of the equations of motion (1) is that they are reversible in time. By changing the signs of all velocities, the particles retrace their trajectories. If the equations of motion are solved correctly, the computer-generated trajectories will also exhibit this property. Moreover, the equations of motion are canonical. This implies the conservation of phase-space volume [10, 16, 17]. This is another aspect to be incorporated into the integration algorithm.

3 Interaction Potentials, Force Field

All macroscopic properties of materials are tightly related to the forces among their elementary building blocks. The spectrum of properties is broad. It ranges from the spacial structure of solids to the secondary and ternary structure of biological supramolecular systems. Thus, it is desirable to achieve a representation of the actual interactions in terms of the classical potential energy $U(\{\mathbf{r}\})$ (*force field*) as accurate as possible. The potential energy is an empirical quantity. Hence, there is no 'correct' functional form. Its functional form is rather a compromise between accuracy and efficiency.

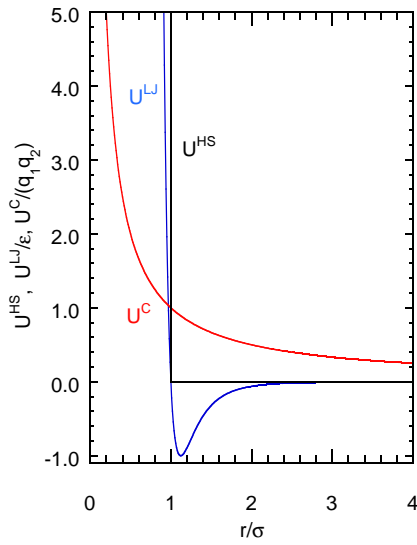
3.1 Atomic systems

The potential energy of N interacting atoms can be divided into terms depending on the coordinates of individual atoms, pairs, triplets etc.:

$$U(\{\mathbf{r}\}) = \sum_i u_1(\mathbf{r}_i) + \sum_i \sum_{j>i} u_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \sum_{j>i} \sum_{k>j>i} u_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k). \quad (6)$$

The first term (u_1) represents the effect of an external field on the system, e.g., gravitational force, electric fields, container walls, etc.. The remaining terms represent particle interactions. The pair potential term u_2 is the most important. The u_3 term, involving triples of atoms, is undoubtedly significant at liquid densities. Higher order terms are expected to be small. Usually, triplet terms are not included in computer simulations because their calculation is very time consuming. Typically the pair potential is chosen such that it includes triplet and higher order interactions. Thus, it has to be regarded as an effective pair potential, representing all the many body effects. A consequence of this approximation is that the effective pair potential needed to reproduce experimental data may turn out to depend on the density, temperature, etc., while the true two-body potential does not.

Typical atomic pair potentials are:



- Hard-sphere potential

$$U^{HS}(r) = \begin{cases} \infty, & r < \sigma \\ 0, & r > \sigma \end{cases} \quad (7)$$

- Lennard-Jones potential

$$U^{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (8)$$

- Coulomb potential

$$U^C(r) = \frac{q_1 q_2}{r}, \quad (9)$$

where r denotes the distance among the two atoms and the q_i s are the charges in cgs units.

The Lennard-Jones potential has an attractive tail at large separations ($\sim r^{-6}$), essentially due to correlations between the electron clouds surrounding the atoms ('van der Waals' or 'London' dispersion). The steep rising repulsive wall at short distances ($\sim r^{-12}$) captures the non-bonded overlap between the electron clouds.

3.2 Molecular systems

Molecules are composed of atoms and can also be described by potential energy terms. However, to account for all aspects of chemical bonding, including the reactions which form and break bonds would require a quantum mechanical description. Typically this problem is solved by accounting for the internal molecular degrees of freedom – bond length, bond bending, torsion – in terms of appropriate potential functions. A simple force field for a molecule is

$$U(\{\mathbf{r}\}) = \sum_{bonds} \frac{k_i^b}{2} (r_i - l_i)^2 + \sum_{angles} \frac{k_i^a}{2} (\theta_i - \theta_{eq,i})^2 + \sum_{torsion} \frac{k_i^t}{2} [1 + \cos(n_i \phi_i - \gamma_i)] + U^{LJ} + U^C. \quad (10)$$

The first term describes the covalent bonds among atoms with the equilibrium bond lengths l_i . The second term accounts for restrictions in the bond angle bending due to overlapping

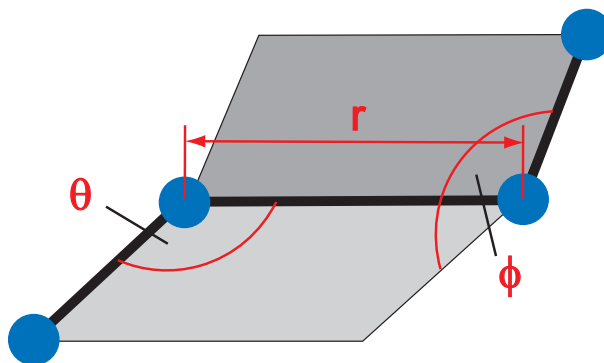


Fig. 1: Internal coordinates of a molecule: bond length r , bond angle θ , and torsion angle ϕ .

electronic clouds of bound atoms. Similarly, the third term describes the interactions of the electronic clouds of atoms separated by three bonds (torsional potential). Further apart atoms in a molecule interacted via the non-bonded Lennard-Jones and Coulomb potentials, similar to unbound atoms. Figure 1 illustrates the internal variables.

Instead by harmonic potentials, bond length and bond bending is often accounted for by constraints, i.e, the bond lengths l_i and/or the bond angles θ_i are assumed to possess fixed values. The constraints are taken into account in the equations of motion by Lagrangian multipliers and specific algorithms are used to solve them [7, 10].

In simulations often convenient units of energy and length are chosen and all other quantities are expressed in terms of these units. In a system with Lennard-Jones interactions typically the unit of length is set equal to σ and the unit of energy to ϵ . As a consequence, the unit of time is $\sigma\sqrt{m/\epsilon}$, where m is the mass of one sort of atoms, and the unit of temperature (T) is ϵ/k_B , where k_B is the Boltzmann factor.

4 Periodic Boundary Conditions, Truncated Potentials

Computer simulations are usually performed with a small number of particles $10 \leq N < 10^7$. The system size is essentially limited by the speed of execution of the program. The time required to evaluate the intermolecular forces, e.g., Lennard-Jones and Coulomb forces, and/or the potential energy is proportional to N^2 . Special techniques will be outlined below, which reduce this dependence to $\mathcal{O}(N)$, for large systems, but the force/energy evaluation almost inevitably dictates the overall speed, and, clearly, smaller systems will always be less expensive. In a small system, however, a large fraction of particles is located at the surface of the confining container. These particles experience quite different forces from those in the bulk.

4.1 Periodic boundary conditions

To overcome surface effects, periodic boundary conditions are implemented [18]. The primary – in the simplest case – cubic box containing the N particles is replicated throughout space to form an infinite lattice (see also contribution B2). The pair potential energy of such an infinite

system reads

$$U(\{\mathbf{r}\}) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' U_{ij}(\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}), \quad (11)$$

where $\mathbf{R}_{\mathbf{n}} = (n_x L, n_y L, n_z L)^T = L \mathbf{n} = \sqrt[3]{V} \mathbf{n}$ is the lattice vector, $n_x, n_y, n_z \in \mathbf{Z}$, and V is the volume of the primary simulation box. The prime indicates that for $\mathbf{n} = 0$ $i \neq j$, i.e., no self-interaction of a particle is taken into account.

It is important to assure that the properties of the small, infinitely periodic system and those of the macroscopic system, which it represents, are the same. This will depend both on the range of the intermolecular potential and the phenomenon under investigation. For a fluid of Lennard-Jones atoms, it should be possible to perform a simulation in a cubic box of side length $L \approx 6\sigma$, without a particle being able to 'sense' the symmetry of the periodic lattice. If the potential is long range, there will be a substantial interaction between a particle and its own images in neighboring boxes, and consequently the symmetry of the cell structure is imposed on a fluid which is in reality isotropic [10].

The use of periodic boundary conditions inhibits the occurrence of long-wavelength fluctuations. This crucially affects the simulation of phase transitions, where such fluctuations are of fundamental importance. If external potentials are present, such as surfaces, they have to be of the same periodicity.

4.2 Potential truncation; *minimum image convention*, neighbor list

The potential energy $U(\{\mathbf{r}\})$ (11) contains infinitely many terms and, hence, can not be calculated accurately. It is often not even desirable to take all periodic images into account. The attractive tail of the Lennard-Jones potential decays like r^{-6} and the potential energy of two particles at a distance of 3σ is only half a percent of the minimal value $-\epsilon$. To determine the remaining part with an immense effort is useless. In addition, the assumed periodicity is artificial. For long-range interactions, however, the summation of the image interactions is unavoidable. For short-range potentials, e.g., the Lennard-Jones potential, it is sufficient to calculate pair interactions up to a certain cutoff radius only. At larger distances, the potential energy is zero. Typical values for the cutoff radius are $r_c = 2.5\sigma$ or $r_c = 3.0\sigma$. To determine the principle properties of fluids, the Lennard-Jones potential is often truncated at its minimum, $r_c = \sqrt[6]{2}\sigma$, i.e., the reflecting part of the potential is taken into account only. To avoid jumps of the potential at the point of truncation, the potential is shifted by its value at $r = r_c$ and it finally reads

$$U = \begin{cases} U^{LJ}(r) - U^{LJ}(r_c), & r < r_c \\ 0, & r \geq r_c \end{cases}. \quad (12)$$

Naturally, the cutoff radius has to be smaller than half of the length of the simulation box. In this case, the minimum image convention can be applied. By the minimum image convention the interaction among closest neighboring particles is calculated only. Since only the equations of motion of the real particles are solved, these are either distances between real particles or a real and an image particle [10, 15].

Advantage of the cutoff is taken only if the interactions of particles with distances larger than the cutoff are left out in the force calculations. Verlet suggested a technique for improving the speed of a program by maintaining a list of the neighbors of a particular particle, which

is updated at intervals [10, 15]. Between updates, the program does not check the distances of all particles, but just those appearing in the list. The number of pair separations explicitly considered is dramatically reduced and the algorithm is $\mathcal{O}(N)$. For large systems ($N \geq 1000$) a cell list is more appropriate. Details can be found in Refs. [10, 15].

5 Algorithms for Molecular Dynamics Simulations

A number of algorithms have been suggested to integrate Newton's equations of motion. Some of them are more suitable than others. A good integration algorithm should satisfy the following criteria:

- The number of force evaluations per integration step should be small.

The evaluation of the forces is typically the most time consuming part of molecular dynamics simulations. Thus, the number of force evaluations should be as small as possible. In comparison to this, the speed of the integration algorithm is of less importance and the time spent for integration is small.

- Accuracy for a large time step.

For a larger time step a smaller number of force evaluations per unit of simulation time is required and hence a larger time interval can be covered.

- Energy and momentum have to be conserved.

In that respect, the long time conservation of energy and momentum is important, i.e., there should be no drift for long times.

- The algorithm should reproduce the classical trajectory as closely as possible.

In fact, no such algorithm exists [15]. The trajectories of all systems through phase space depend sensitively on the initial conditions. This means that two trajectories which are initially close will diverge exponentially as time progresses. The integration error caused by the algorithm can be considered as the source of the initial small difference between the 'true' trajectory of the system and the trajectory generated in the simulation. Any small integration error, no matter how small, will always cause the simulated trajectory to diverge exponentially from the true trajectory compatible with the initial conditions.¹ This so-called Lyapunov instability does not render molecular dynamics simulations useless. First of all, the aim of a molecular dynamics simulation is not to predict precisely a particular trajectory, rather the statistical properties are of interest. We wish to determine the average behavior of a system, which is achieved when the trajectory covers an appropriate fraction of the energy surface in phase space. Still, this would not justify the use of inaccurate trajectories unless the trajectories obtained numerically, in some sense, are close to true trajectories.

- The algorithm should be time reversible and phase-space conserving.

Newton's equations of motion are time reversible, a fact which should be captured in the integration algorithm. Moreover, the equations are symplectic, i.e., the phase-space volume is, among others, a conserved quantity [16, 17]. After sufficiently long times, we

¹This applies to chaotic systems only.

expect that the non-phase-space conserving algorithms will have greatly expanded the volume of the system in phase space, which is not compatible with energy conservation. Reversible, phase-space conserving algorithms will not change the magnitude of the volume in phase space. This property is not sufficient to guarantee the absence of a long-term energy drift, but it is at least compatible with it.

It should be noted, that no integration algorithm on a digital computer is really time reversible, because the computer has finite machine precision using floating point arithmetic that result in rounding errors.

A simply but very good algorithm, which satisfies the most important criteria is the **Verlet algorithm** [19]. It can be derived by a Taylor expansion of the coordinates of a particle at time t :

$$\mathbf{r}(t+h) = \mathbf{r}(t) + \dot{\mathbf{r}}(t)h + \frac{h^2}{2m}\mathbf{F}(t) + \frac{h^3}{3!}\frac{d^3}{dt^3}\mathbf{r} + \mathcal{O}(h^4), \quad (13)$$

similarly,

$$\mathbf{r}(t-h) = \mathbf{r}(t) - \dot{\mathbf{r}}(t)h + \frac{h^2}{2m}\mathbf{F}(t) - \frac{h^3}{3!}\frac{d^3}{dt^3}\mathbf{r} + \mathcal{O}(h^4), \quad (14)$$

where h is the time step. Summing of the two equations yields

$$\mathbf{r}(t+h) \approx 2\mathbf{r}(t) - \mathbf{r}(t-h) + \frac{h^2}{m}\mathbf{F}(t). \quad (15)$$

The error of the estimate of the new position is of the order of h^4 . Note that the Verlet algorithm does not require the velocity to compute the new position. The velocity follows by subtraction of Eqs. (13) and (14):

$$\dot{\mathbf{r}}(t) = \frac{\mathbf{r}(t+h) - \mathbf{r}(t-h)}{2h} + \mathcal{O}(h^2). \quad (16)$$

5.1 Liouville formulation of time-reversible algorithms

For most molecular dynamics simulations, the Verlet algorithm is perfectly adequate. Alternative formulations of the Verlet scheme can systematically be derived by a method introduced by Tuckerman, Berne, and Martyna [15,20]. The method is based on a formal solution of Newton's equations of motion.

We again consider a system of N mass points in 3 dimensions. The equation of motion for the phase-space vector $\mathbf{\Gamma} = (\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)^T$ is given by the Liouville equation

$$\dot{\mathbf{\Gamma}}(t) = i\mathcal{L}\mathbf{\Gamma}(t), \quad (17)$$

with

$$i\mathcal{L} = \sum_{i=1}^N \left(\mathbf{F}_i \nabla_{\mathbf{p}_i} + \frac{\mathbf{p}_i}{m_i} \nabla_{\mathbf{r}_i} \right). \quad (18)$$

The formal solution is

$$\mathbf{\Gamma}(t) = \mathbf{U}(t)\mathbf{\Gamma}(0), \quad (19)$$

with

$$\mathbf{U}(t) = \exp(i\mathcal{L}t). \quad (20)$$

As is obvious, \mathcal{L} is a hermitian operator, i.e., $\mathcal{L} = \mathcal{L}^\dagger$. Hence, \mathbf{U} is unitary, i.e., $\mathbf{U}^{-1}(t) = \mathbf{U}^\dagger(t) = \mathbf{U}(-t)$, and the solution of the equation of motion is time reversible. Unfortunately,

$\mathbf{U}(t)\Gamma(0)$ cannot be evaluated in general. However, an integration algorithm can be derived when \mathbf{U} is replaced by a product of operators which approximate it to a certain order and which can be evaluated individually. For a Liouville operator \mathcal{L} composed of two parts ($i\mathcal{L} = i\mathcal{L}_1 + i\mathcal{L}_2$) a factorization is obtained by the Trotter theorem [21] or equivalently by the Campbell-Baker-Hausdorff formula

$$\mathbf{U}(t) = \mathbf{U}_1(t/2)\mathbf{U}_2(t)\mathbf{U}_1(t/2) + \mathcal{O}(t^3). \quad (21)$$

Using the factorization $i\mathcal{L} = i\mathcal{L}_x + i\mathcal{L}_p$ for the Liouville operator (18) with the operators

$$\mathbf{U}_x(t) = e^{i\mathcal{L}_x t} = \exp\left(t \sum_{i=1}^N \frac{\mathbf{p}_i}{m_i} \nabla_{\mathbf{r}_i}\right), \quad (22)$$

$$\mathbf{U}_p(t) = e^{i\mathcal{L}_p t} = \exp\left(t \sum_{i=1}^N \mathbf{F}_i \nabla_{\mathbf{p}_i}\right), \quad (23)$$

we find the following formal solution of the equations of motion:

$$\Gamma(h) = \mathbf{U}_p(h/2)\mathbf{U}_x(h)\mathbf{U}_p(h/2)\Gamma(0) = \mathbf{G}(h)\Gamma(0). \quad (24)$$

With the relation

$$e^{c\partial/\partial q} f(q) = f(q + c) \quad (25)$$

the operators can be evaluated and we obtain

$$\mathbf{r}_i(h) = \mathbf{r}_i(0) + h\dot{\mathbf{r}}_i(0) + \frac{h^2}{2m_i} \mathbf{F}_i(0), \quad (26)$$

$$\dot{\mathbf{r}}_i(h) = \dot{\mathbf{r}}_i(0) + \frac{h}{2m_i} (\mathbf{F}_i(0) + \mathbf{F}_i(h)). \quad (27)$$

This integration scheme is called **velocity Verlet algorithm**. Compared to the Verlet algorithm, the velocity appears explicitly in the calculation of the position. By combining two steps, this integration scheme is identical to the Verlet algorithm. However, the above scheme yields the positions and velocities at the same time.

Since the operators (22) and (23) are unitary operators, $\mathbf{G}(t)$ is unitary too, i.e., $\mathbf{G}^{-1}(t) = \mathbf{G}(-t)$, and, hence, the integration scheme is time reversibel.

The integration of Newton's equations of motion is as follows by the velocity Verlet algorithm:

Initial conditions: positions $\{\mathbf{r}(t)\}$, velocities $\{\dot{\mathbf{r}}(t)\}$, forces $\{\mathbf{F}(t)\}$ ²

1. Calculation of new positions according to Eq. (26)

$$\mathbf{r}_i(t + h).$$

Calculation of velocities

$$\dot{\mathbf{r}}_i(t + h/2) = \dot{\mathbf{r}}_i(t) + \frac{h}{2m_i} \mathbf{F}_i(t).$$

²Note that we consider conservative forces, i.e., $\mathbf{F}_i = \mathbf{F}_i(\{\mathbf{r}\})$, and, hence, the forces can be calculated via the positions.

2. Calculation of forces using positions $\mathbf{r}(t+h)$

$$\mathbf{F}_i(t+h) = \mathbf{F}_i(\{\mathbf{r}(t+h)\}).$$

3. Calculation of velocities

$$\dot{\mathbf{r}}_i(t+h) = \dot{\mathbf{r}}_i(t+h/2) + \frac{h}{2m_i} \mathbf{F}_i(t+h).$$

We will come back to the Liouville procedure when we determine the integration scheme for isothermal molecular dynamics simulations in Sec. 6.1.

5.2 Accuracy check, equilibration

A newly written molecular dynamics simulation program will not work right away in the desired manner. To remove errors and to control the program, various physical quantities can be considered. The first check must be that the conservation laws are properly obeyed, in particular the energy and the momentum should be 'constant'. In fact small changes in the energy will occur due to the finite precision of the integration scheme. For a simple Lennard-Jones system, fluctuations of the order of 0.1% are generally considered to be acceptable. Energy fluctuations may be reduced by decreasing the time step.

A slow upwards drift of the energy may be due to a time step that is too long or might indicate a program error. Drifts due to an inadequate time step can be removed by decreasing it.

In general, it is recommended to study small systems initially with well defined initial conditions.

When the program appears to be running correctly, the user should check that the monitored quantities are in fact evolving in time. A time step which is too small will be very wasteful of computer time, and the extent to which the time step can be increased without prejudicing the stability of the simulation should be investigated.

So far, only aspects concerning the correct solution of the equations of motion are discussed [10]. A separate question is the proper thermodynamic equilibration of a system.

The particles of liquid-like systems are initially typically placed on lattice sites. Similarly, for polymer molecules stretched conformations are selected often with particles on lattice sites. To reach the real state point requires to run the program for a certain time. At the end of this equilibration period, all memory of the initial configuration should have been lost. Various quantities can be monitored to control the equilibration process. If a certain system temperature is required, the kinetic energy of the particles has to be monitored and adjusted to the desired value by an appropriate procedure (isothermal MD simulations are described in Sec. 6.1). A simple way to monitor system equilibration is to record the instantaneous values of the potential energy and pressure during that period. In the case of a lattice start, the potential energy rises. The equilibration period should be extended at least until these quantities have ceased to show a systematic drift and have started to oscillate about steady state values.

Among other quantities, which can be monitored while the initial lattice melts is the translational order parameter $\rho(\mathbf{k}) = \sum_{i=1}^N \cos(\mathbf{k}\mathbf{r}_i)/N$, where \mathbf{k} is a reciprocal lattice vector of the initial lattice [10]. For example $\mathbf{k} = (2\pi\sqrt[3]{N}/L)(1, 1, 1)^T$. For a solid, the magnitude of $\rho(\mathbf{k})$ is of order unity, whereas for a liquid it oscillates about zero with amplitude $\mathcal{O}(N^{-1/2})$.

For molecular systems other quantities can be used like the mean square end-to-end distance or the radius of gyration. In addition to the translational order parameter, rotational order parameters can be monitored [10].

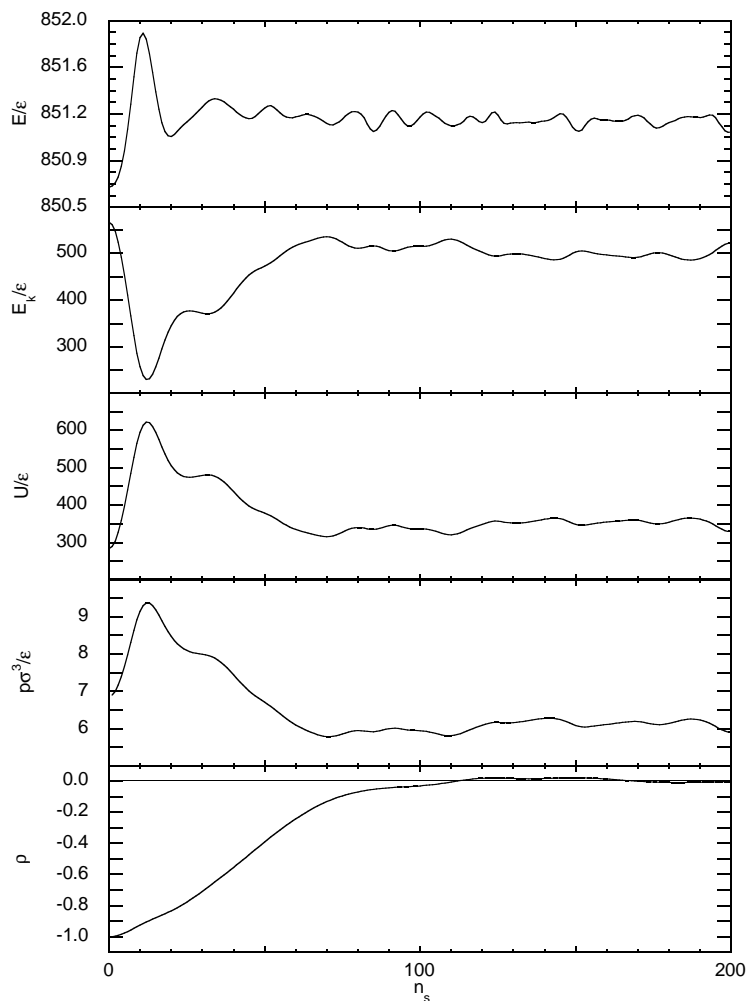


Fig. 2: *Equilibration of a Lennard-Jones fluid in a molecular dynamics simulation. The starting configuration is a cubic lattice. From top to bottom, the total energy, the kinetic energy, the potential energy, the pressure, and the translational order parameter are shown.*

Moreover, dynamical quantities, like the mean square displacement of particles or molecules from their initial lattice positions can be calculated. This function increases in the course of time in a liquid, but oscillates around a mean value in solids. A useful rule of thumb is that the root-mean-square displacement per particle exceeds 0.5σ and is clearly increasing, then the system has 'melted' and the equilibration can be terminated.

One useful trick that may be used to speed up equilibration from a lattice, is to raise the kinetic temperature to a higher value for a number of simulation steps.

It is difficult to say how long a run has to be to guarantee equilibration. The above quantities have to be examined carefully and at the end of the equilibration period, they should have clearly reached the expected stationary state. At the end of the equilibration period, the accumulators for the ensemble averages are reset to zero, and the production phase of the simulation begins.

As an example, the time dependence of certain quantities during the equilibration process of a simple Lennard-Jones fluid are shown in Fig. 2. The system consists of 512 particles, which

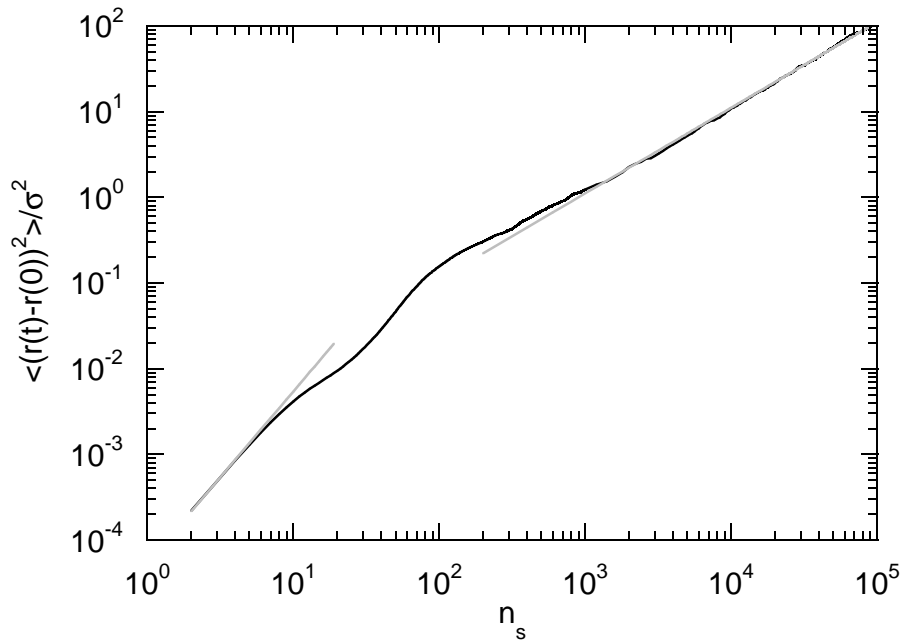


Fig. 3: Mean square displacement of a Lennard-Jones particle as a function of the number of simulation steps. The gray lines indicate ballistic ($n_s < 10$) and diffusive motion ($n_s > 10^3$), respectively.

interact with each other via the Lennard-Jones potential (12) with the cutoff radius $r_c = \sqrt[6]{2}\sigma$. The side length of the cubic simulation box is $L = 8.46\sigma$, which gives a number density of $\rho = N/V = 0.8456/\sigma^3$. The time step is $h = 5 \times 10^{-3} \sqrt{m\sigma^2/\epsilon}$ and the initial temperature is $k_B T/\epsilon = 0.722$. As initial configuration, a cubic lattice was chosen. After about 70 simulation steps (n_s), the potential energy, the kinetic energy, and the pressure are close to their equilibrium values. The total energy varies at the fourth digit, as expected. The mean value of the kinetic energy $\langle E_k \rangle \approx 500\epsilon$ is about 10 % smaller than the thermal value $3k_B T/2 = 554.5\epsilon$. The adjustment to the desired temperature requires thermalization of the system. Various available techniques are described in Sec. 6.1.

The translational order parameter requires somewhat more simulation steps before it reaches its limiting behavior. Hence, it is a more sensitive quantity with respect to the equilibration process.

In addition, dynamical quantities can be monitored. An example is the mean square displacement of the particles from their initial (lattice) positions. Figure 3 displays the mean square displacement for the above Lennard-Jones system. For short times ($n_s < 10$), the particles hardly interact with each other and therefore move ballistically. Correspondingly, the mean square displacement is given by $\langle (\mathbf{r}_i(t) - \mathbf{r}_i(0))^2 \rangle = \langle \mathbf{v}_i^2 \rangle t^2 = 3k_B T t^2 / m$ as indicated by the gray line. For $n_s > 10^3$, the mean square displacement exhibits the expected linear increase with time. At $n_s \approx 200$, the root-mean-square displacement per particle is on the order of $\sigma/2$ and is clearly increasing. Thus, the system has 'melted' [10]. As can be seen in Fig. 3, a linear mean square displacement is reached after approximately 500 simulation steps. Hence, this dynamical quantity requires a more extended equilibration period than the thermodynamic quantities of Fig. 2.

6 Molecular Dynamics at Various Ensembles

The simulations described so far apply to a closed system with a given number of particles N in a fixed volume V at constant energy E . If we assume that time averages are equivalent to ensemble averages, then the averages obtained in such a simulation are equivalent to ensemble averages in the microcanonical ensemble (NVE ensemble). However, various physical situations require the simulation of other ensembles by inclusion of the environment into the simulation. Examples are:

- heat bath (NVT ensemble)
- particle reservoir (μ VE ensemble)
- heat and particle reservoir (μ VT ensemble)
- temperature and/or pressure reservoir (NpH_e ensemble, NpT ensemble)³
- external fields, which induce non-equilibrium states by gradients, e.g., shear flow

There are various ways to couple the Newtonian system to the desired variables. In this contribution, I will focus on constant temperature and constant pressure simulations.

6.1 Constant temperature molecular dynamics simulations

In canonical ensemble, the number of particles, the temperature, and the volume are constant. The energy is no longer a conserved quantity, only the mean of the energy is the same for all members of the ensemble. The relevant thermodynamic variable is the temperature (T) of the (background) heat bath. In a molecular dynamics simulation, the system under consideration has to be coupled to the heat bath. There are various method to achieve such a coupling.

Stochastic methods

A canonical ensemble is obtained by coupling the particles of the systems to a background friction and a stochastic force. The equations of motion of the particles are the given by the Klein-Kramers equation (Langevin equation) [22]

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i - \gamma m_i \dot{\mathbf{r}}_i + \mathbf{\Lambda}_i, \quad (28)$$

where γ is the friction constant and the $\mathbf{\Lambda}_i$ s are stochastic forces. A canonical ensemble follows for a white noise stochastic force, i.e., a stationary, Gaussian, and Markovian process, with the moments

$$\langle \mathbf{\Lambda}_i(t) \rangle = 0, \quad (29)$$

$$\langle \Lambda_{i\alpha}(t) \Lambda_{j\beta}(t') \rangle = 2\gamma k_B T \delta_{ij} \delta_{\alpha\beta} \delta(t - t') \quad \alpha, \beta \in x, y, z. \quad (30)$$

By transformation into a Fokker-Planck equation, the stationary state canonical distribution function is obtained [22].

Aside for the fact that the simulation of a canonical ensemble is guaranteed, the stochastic force stabilizes the integration algorithm. The disadvantage of the algorithm is that the dynamics is

³ $H_e = H + pV$ is the enthalpy

modified. The friction force causes, e.g., a slower diffusion than in the case of a bare molecular dynamics simulation.

The Brownian dynamics simulation technique presented in the contribution B4 is a special limit of the above equation for very large friction.

A further method was proposed by Andersen [15, 23]. Here, the coupling to the heat bath is achieved by occasional changes of the velocity of a randomly selected particle. The velocity is taken from the Maxwellian distribution. Between the collisions, the particles move according to Newton's equations of motion. The mixing of Newtonian dynamics with stochastic collisions turns the molecular dynamics simulation into a Markov process.

In another variant, the velocities of all particle are changed at the same time (massive stochastic collisions). By the changes of the velocities, the average is calculated over various initial conditions, where the initial state is determined by the canonical distribution. In fact, an ensemble average is calculated.

Velocity scaling

A method which is very well suited to control temperature during thermalization is velocity scaling. Here, the velocities of all particles are scaled such that at any integration step the desired kinetic energy and hence temperature is obtained. If E_k the kinetic energy at a certain time we have

$$E_k = \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{r}}_i^2 = \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{r}}_i'^2 c^2 = c^2 \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{r}}_i'^2 = c^2 \frac{3}{2} N k_B T, \quad (31)$$

where the $\dot{\mathbf{r}}_i'$ s are the scaled velocities. The scaling factor is then given by $c = \sqrt{2E_k/3Nk_B T}$ and the new velocities by $\dot{\mathbf{r}}_i' = \dot{\mathbf{r}}_i/c$.

Such a scaling of velocities does not provide a canonical ensemble and may lead to temperature gradients in the system.

Extended phase-space isothermal molecular dynamics simulations

The above methods rely on stochastic interactions of the system with the heat bath. Over the last decades, deterministic methods to simulate canonical ensembles have been introduced and are nowadays routinely used [15, 17, 24–27]. The novel idea of these methods is that the phase space of the physical system under consideration is extended by a single degree of freedom. This is in contrast to traditional ideas, where the heat bath possesses many more degrees of freedom than the system itself. In order to yield the desired canonical distribution function, the trajectory of the extended system has to cover the phase space sufficiently dense. For a small number of degrees of freedom, this requires that the system exhibits chaotic motion.

For the derivation of equations of motion, we consider two different systems. A physical system with coordinates $\{\mathbf{r}\}$ and momenta $\{\mathbf{p}\}$. In addition, a virtual system with primed variables $\{\mathbf{r}'\}$ and $\{\mathbf{p}'\}$. The coordinate of the extended phase space is denoted by s and its canonically conjugated momentum by p_s . The corresponding physical variables are η and p_η .

Virtual system: $\{\mathbf{r}', \mathbf{p}', s, p_s\}$

Physical system: $\{\mathbf{r}, \mathbf{p}, \eta, p_\eta\}$

The Hamiltonian of the extended system is **chosen** as [24, 25]

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i'^2 s^{2n}}{2m_i} + U(\{\mathbf{r}'\}) + \frac{p_s^2}{2Q} + gk_B T \ln s = H_0(\{\mathbf{p}'s^n\}, \{\mathbf{r}'\}) + \frac{p_s^2}{2Q} + gk_B T \ln s. \quad (32)$$

The parameter g is determined such that averages of the physical system correspond to canonical averages. Q is the mass associated with the dynamics of s . In its original derivation, Nosé used $n = -1$. For certain systems, however, it is desirable to use a different exponent [17]. I will come back to this point later on. Since H is a Hamiltonian, it is a conserved quantity (energy conservation in the extended phase space).

The equations of motion of the extended virtual system are given by Hamilton's equations:

$$\frac{d\mathbf{r}_i'}{dt'} = \frac{\mathbf{p}_i' s^{2n}}{m_i}, \quad \frac{d\mathbf{p}_i'}{dt'} = \mathbf{F}_i, \quad (33)$$

$$\frac{ds}{dt'} = \frac{p_s}{Q}, \quad \frac{dp_s}{dt'} = -\frac{1}{s} \left[\sum_{i=1}^N n \frac{\mathbf{p}_i'^2 s^{2n}}{m_i} + gk_B T \right]. \quad (34)$$

t' denotes the time of the virtual system. Using the relations between the virtual variables and the physical variables:

$$\mathbf{r}_i = \mathbf{r}_i', \quad \mathbf{p}_i = \mathbf{p}_i' s^n, \quad (35)$$

$$\eta = \ln s, \quad p_\eta = p_s s^{-(n+1)}, \quad (36)$$

as well as the scaling relation for the time t'

$$t = \int^{t'} s(t'')^n dt'', \quad (37)$$

which means that $dt = s(t')^n dt'$, the following equations of motion for the physical variables are obtained

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i}, \quad \frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i + \frac{n}{Q} p_\eta \mathbf{p}_i, \quad (38)$$

$$\frac{d\eta}{dt} = \frac{p_\eta}{Q}, \quad \frac{dp_\eta}{dt} = -e^{-2(n+1)\eta} \left[\sum_{i=1}^N n \frac{\mathbf{p}_i^2}{m_i} + gk_B T \right] - \frac{(n+1)}{Q} p_\eta^2. \quad (39)$$

The time scaling (37) is required to obtain the proper velocity momentum relation Eq. (38).

The **Nosé-Hoover equations of motion** follow for $n = -1$

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i}, \quad \frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i - \frac{p_\eta}{Q} \mathbf{p}_i, \quad (40)$$

$$\frac{d\eta}{dt} = \frac{p_\eta}{Q}, \quad \frac{dp_\eta}{dt} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - gk_B T. \quad (41)$$

Since the equations of motion for $\{\mathbf{r}\}$, $\{\mathbf{p}\}$, and p_η are independent of η in this set of equations, it is not necessary to integrate the equation for η . However, I prefer to solve the whole set of equations, because

$$H^{NH} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\{\mathbf{r}\}) + \frac{p_\eta^2}{2Q} + gk_B T \eta \quad (42)$$

is a conserved quantity. Although H^{NH} is not a Hamiltonian, it can be used to measure how well the integration algorithm conserves energy.

Before an integration algorithm is derived, it will be shown that an isothermal system is simulated by the extended phase-space method. Moreover, the parameter g will be determined.

Since the extended system corresponds to a microcanonical ensemble, its partition function is given by

$$Z = \int \delta(H - E) s^n d^N \mathbf{r}' d^N \mathbf{p}' ds dp_s, \quad (43)$$

with H defined in Eq. (32) and s^n the weighting of the distribution function due to the time scaling (37) [28]. Transformation of the variables according to Eqs. (35) (this variable transformation can be used to define the relation between the virtual and physical coordinates and momenta) and integration over s straightforwardly yields (cf. Appendix A1)

$$Z = C \int \exp \left(-\frac{-3nN + n + 1}{gk_B T} H_0 \right) d^N \mathbf{r} d^N \mathbf{p}. \quad (44)$$

The constant C includes factors which cancel in the calculation of averages of observables depending on $\{\mathbf{r}, \mathbf{p}\}$ only. $H_0 = H_0(\{\mathbf{r}\}, \{\mathbf{p}\})$ is defined in Eq. (32). Comparison of Eq. (44) with the Boltzmann distribution implies that $g = n + 1 - 3nN$, or for the Nosé-Hoover thermostat with $n = -1$, $g = 3N$.

As it turns out, the Nosé-Hoover equations of motion (40) and (41) do not preserve the phase-space volume. The Jacobian is a function of time and not a constant (for details see Ref. [17]). The factor n is another parameter which can be adjusted such that the equations of motion for the physical variables conserve phase-space volume. As discussed in [17], this is easily achieved for systems with a small number of degrees of freedom. For a system with one degree of freedom, e.g., a one-dimensional harmonic oscillator, the equations are symplectic for $n = -2$ [29]. However, due to numerical problems this is no longer possible for systems with a large number of degrees of freedom and we recommend to either use the Nosé-Hoover equations or Eqs. (38)-(39) with $n = -2$.

It can easily be shown that the microcanonical ensemble average of a function $A(\{\mathbf{p}' s^n\}, \{\mathbf{r}'\}) = A(\{\mathbf{p}\}, \{\mathbf{r}\})$ in the extended phase space is equal to the canonical ensemble average in the physical phase space $(\{\mathbf{p}\}, \{\mathbf{q}\})$. With the distribution function $\rho = s^n \delta(H - E)/Z$, a calculation similar to that for the partition function (44) (cf. Appendix A1) yields

$$\langle A \rangle = \frac{1}{Z} \int A(\{\mathbf{p}' s^n\}, \{\mathbf{r}'\}) \delta(H - E) s^n d^N \mathbf{r}' d^N \mathbf{p}' ds dp_s \quad (45)$$

$$= \frac{1}{\tilde{Z}} \int A(\{\mathbf{p}\}, \{\mathbf{r}\}) \exp(-\beta H_0) d^N \mathbf{r} d^N \mathbf{p}, \quad (46)$$

$$\tilde{Z} = \int \exp(-\beta H_0) d^N \mathbf{r} d^N \mathbf{p}, \quad (47)$$

with $\beta = 1/(k_B T)$. For the time average of the function $A(\{\mathbf{p}\}, \{\mathbf{r}\}) = A(t)$ follows

$$\langle A \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau A(\{\mathbf{p}\}, \{\mathbf{r}\}) dt = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau A(\{\mathbf{p}' s^n\}, \{\mathbf{r}'\}) dt \quad (48)$$

$$= \lim_{\tau' \rightarrow \infty} \frac{\frac{1}{\tau'} \int_0^{\tau'} A(\{\mathbf{p}' s^n\}, \{\mathbf{r}'\}) s^n dt'}{\frac{1}{\tau'} \int_0^{\tau'} s^n dt'} \quad (49)$$

$$\stackrel{\text{ergodic}}{=} \frac{\frac{1}{Z} \int A(\{\mathbf{p}' s^n\}, \{\mathbf{r}'\}) \delta(H - E) s^n d^N \mathbf{r}' d^N \mathbf{p}' ds dp_s}{\frac{1}{Z} \int \delta(H - E) s^n d^N \mathbf{r}' d^N \mathbf{p}' ds dp_s} \quad (50)$$

$$= \frac{1}{Z} \int A(\{\mathbf{p}\}, \{\mathbf{r}\}) \exp(-\beta H_0) d^N \mathbf{r} d^N \mathbf{p}, \quad (51)$$

which means that

$$\langle A(\{\mathbf{p}' s^n\}, \{\mathbf{r}'\}) \rangle_m = \langle A(\{\mathbf{p}\}, \{\mathbf{r}\}) \rangle_c. \quad (52)$$

An important assumption is the validity of the ergodicity hypothesis.

Again, the formal solution of the Liouville equation $\Gamma(t) = e^{i\mathcal{L}t}\Gamma(0)$ is exploited to derive an integration algorithm for the Nosé-Hoover equations of motion. The Liouville operator now reads

$$i\mathcal{L} = \sum_{i=1}^N \left[\frac{\mathbf{p}_i}{m_i} \nabla_{\mathbf{r}_i} + \left(\mathbf{F}_i - \frac{p_\eta}{Q} \mathbf{p}_i \right) \nabla_{\mathbf{p}_i} \right] + \frac{p_\eta}{Q} \frac{\partial}{\partial \eta} + f_\eta \frac{\partial}{\partial p_\eta}, \quad (53)$$

with

$$f_\eta = f_\eta(\{\mathbf{p}\}) = \sum_{i=1}^N \mathbf{p}_i^2 / m_i - g k_B T. \quad (54)$$

A factorization of \mathcal{L} , which yields an integration scheme equivalent to the velocity Verlet algorithm is

$$i\mathcal{L} = i\mathcal{L}_x + i\mathcal{L}_p + i\mathcal{L}_\eta + i\mathcal{L}_{p_\eta} + i\mathcal{L}_{pp}. \quad (55)$$

The operators \mathbf{U}_x and \mathbf{U}_p are defined in Eqs. (22) and (23). The others are defined as

$$\mathbf{U}_\eta = e^{i\mathcal{L}_\eta t} = \exp \left(t \frac{p_\eta}{Q} \frac{\partial}{\partial \eta} \right), \quad (56)$$

$$\mathbf{U}_{p_\eta} = e^{i\mathcal{L}_{p_\eta} t} = \exp \left(t f_\eta \frac{\partial}{\partial p_\eta} \right), \quad (57)$$

$$\mathbf{U}_{pp} = e^{i\mathcal{L}_{pp} t} = \exp \left(-t \sum_{i=1}^N \frac{p_\eta}{Q} \mathbf{p}_i \nabla_{\mathbf{p}_i} \right). \quad (58)$$

The solution of the equation is then $\Gamma(t) = \mathbf{G}(t)\Gamma(0) + \mathcal{O}(t^3)$ with

$$\mathbf{G}(t) = \mathbf{U}_{p_\eta}(t/2) \mathbf{U}_\eta(t/2) \mathbf{U}_{pp}(t/2) \mathbf{U}_p(t/2) \mathbf{U}_x(t) \mathbf{U}_p(t/2) \mathbf{U}_{pp}(t/2) \mathbf{U}_\eta(t/2) \mathbf{U}_{p_\eta}(t/2). \quad (59)$$

With the relation (25) and

$$e^{cq\partial/\partial q}q = qe^c \quad (60)$$

evaluation of the operators yields

$$\mathbf{r}_i(h) = \mathbf{r}_i(0) + h\dot{\mathbf{r}}(0)e^{-\frac{h}{2}[\dot{\eta}(0) + \frac{h}{2Q}f_\eta(0)]} + \frac{h^2}{2m_i}\mathbf{F}_i(0), \quad (61)$$

$$\dot{\mathbf{r}}_i(h) = \left(\dot{\mathbf{r}}_i(0)e^{-\frac{h}{2}[\dot{\eta}(0) + \frac{h}{2Q}f_\eta(0)]} + \frac{h}{2m_i}(\mathbf{F}_i(0) + \mathbf{F}_i(h)) \right) e^{-\frac{h}{2}[\dot{\eta}(0) + \frac{h}{2Q}f_\eta(0)]}, \quad (62)$$

$$\eta(h) = \eta(0) + h\dot{\eta}(0) + \frac{h^2}{2Q}f_\eta(0), \quad (63)$$

$$\dot{\eta}(h) = \dot{\eta}(0) + \frac{h}{2Q}(f_\eta(0) + f_\eta(h)). \quad (64)$$

The exponential terms include contributions higher than $\mathcal{O}(h^3)$. This seems unnecessary for an $\mathcal{O}(h^3)$ integration algorithm. They ensure, however, that the integration scheme is time reversible and of $\mathcal{O}(h^3)$. A number of other reversible integration algorithms are presented in literature [17, 31, 32], but, the velocities are obtained typically within $\mathcal{O}(h^2)$ only. Some of them are based on the standard Verlet integration scheme [32]. In Ref. [33], an extension to Nosé-Hoover chains is presented.

Applications

To illustrate the features of the extended phase-space methods, in particular the Nosé-Hoover thermostat, three examples are presented.

Harmonic oscillator

Since the equations of motion of the harmonic oscillator can be solved analytically, it is often used to test algorithms.

The Hamiltonian of the one-dimensional harmonic oscillator is

$$H_0 = \frac{p^2}{2m} + \frac{1}{2}Kx^2. \quad (65)$$

In the evaluation of the equations of motion we use $m = 1$ and $K = 1$, i.e., we introduce reduce variables by choosing $K^{-1/2}$ as the unit of length and $(m/K)^{1/2}$ as the unit of time.

As it turns out, the Nosé-Hoover thermostat yields non-ergodic trajectories, i.e., the system trajectory of the extended system is not covering the energy hypersurface densely. Thus, the Nosé-Hoover thermostat does not yield a canonical distribution in phase space. The reason is that the dynamics of the oscillator is not sufficiently chaotic to sample all of the phase space. As a consequence, time averages are not equivalent to ensemble averages.

The extended phase-space equations of motion obtained for $n = -2$ (Eqs. (38)-(39)) contain additional non-linear terms and factors. As shown in figure 4, these non-linearities are sufficient to generate chaotic trajectories. Naturally, the influence of the non-linearities of the thermostat on the dynamics of the harmonic oscillator depends on the coupling parameter Q . In the limit $Q \rightarrow \infty$, the oscillator decouples from the variable η . Correspondingly, we expect a strong coupling for small Q values. Figure 4 shows that for larger Q values regular trajectories exist,

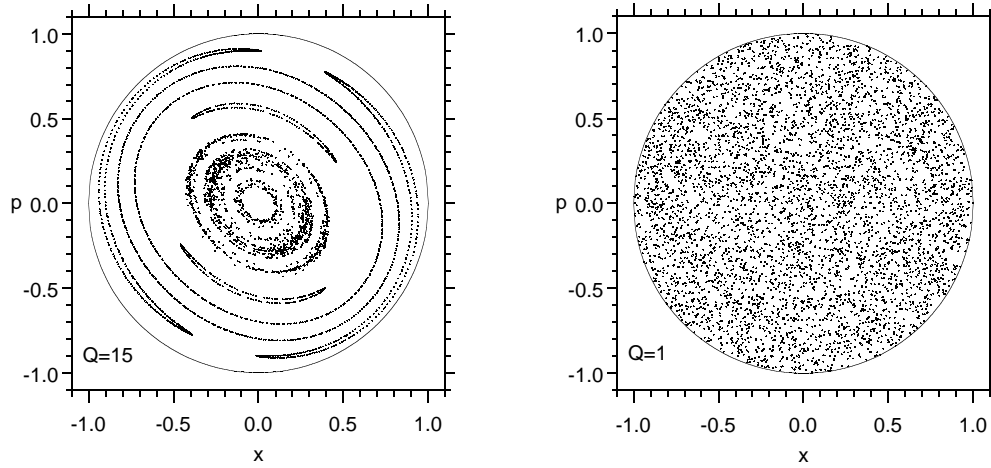


Fig. 4: Poincaré surfaces of section for various initial conditions and the parameters $Q = 15$ (left) and $Q = 1$ (right). The energy is set to $H_0 = 1/2$ and $k_B T = 1/2$.

whereas for small Q ($Q \lesssim 1$) the phase space is covered densely by a single trajectory. Hence, the extended phase-space thermostat with $n = -2$ yields canonical time averages. More details are presented in Ref. [29].

Double-well potential

The investigation of the dynamics of a particle in the double-well potential

$$U(x) = x^2 + \frac{1}{0.25 + x^2} \quad (66)$$

shows that the failure of the Nosé-Hoover thermostat to produce a canonical distribution is not related to the linear forces of the harmonic oscillator [17]. By solving Eqs. (38)-(39) for $n = -2$ with the reversible integration scheme described in Ref. [17] as well as those for the Nosé-Hoover thermostat, we determined the probability distribution for the position x . The initial conditions are: $x(0) = 0$, $\dot{x}(0) = 0.5$, $\eta(0) = -\dot{x}(0)^2/2k_B T$, and $\dot{\eta}(0) = 0$, which gives the initial energy $H(0) = 4$ ($m = 1$). The temperature is set to $k_B T = 4$, the thermostat mass to $Q = 0.09$, and the time step to $h = 10^{-2}$.

The probability distribution $P(x)$ is presented in Fig. 5(a). The dots give the distribution after 10^8 integration steps. The solid line is the Boltzmann distribution $\exp(-U(x)/k_B T)/Z$ and the dashed line is found from the Nosé-Hoover equations. The probability distributions are normalized such that the areas under the curves are equal to one. As is obvious from the figure, the probability distribution obtained by the equations with $n = -2$ is hardly distinguishable from the Boltzmann distribution, whereas the lack of full chaotic behavior of the Nosé-Hoover thermalization is apparent in the deviation from the Boltzmann factor. This is even more obvious, when calculating the deviations of the numerical results from the Boltzmann distribution. Figure 5(b) shows the integrated square of the deviation from the Boltzmann factor as a function of the number of integration steps. Obviously, the distribution for the Nosé-Hoover thermostat is not significantly improved after 10^4 integration steps. However, the distribution obtained by the equations with $n = -2$ converges monotonically towards the Boltzmann distribution. Hence, by the latter choice, canonical averages are obtained.

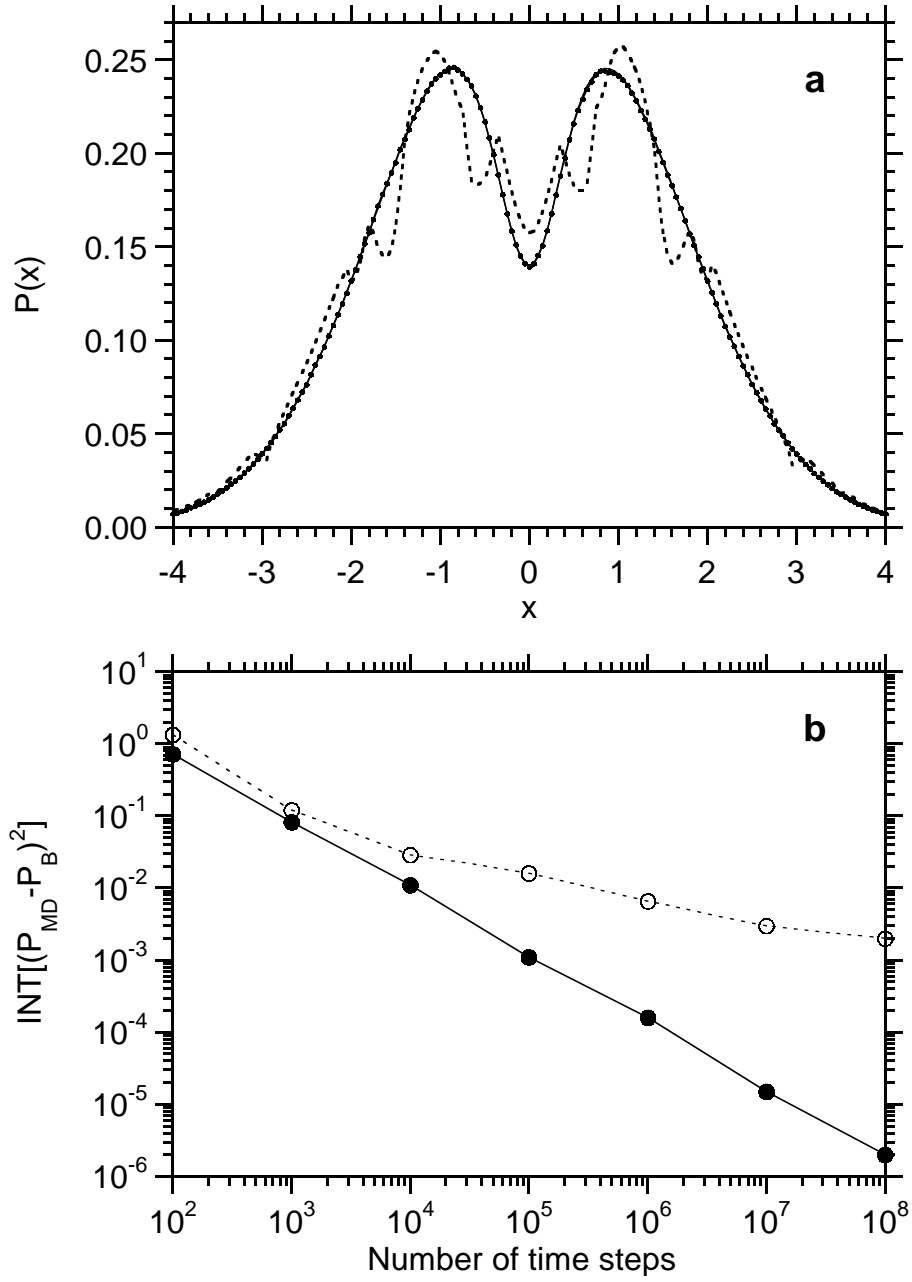


Fig. 5: (a) Probability distribution for the position of a particle in a double-well potential. The dots are obtained by the algorithm described in Ref. [17]. The dashed line is the result of the Nosé-Hoover thermostat, and the solid line is the Boltzmann distribution. The parameters are $h = 10^{-2}$, $Q = 0.09$, and $H = k_B T = 4$. (b) Integral of the square of the deviation from the Boltzmann distribution. The open symbols correspond to the Nosé-Hoover thermostat.

Various other methods have been suggested to remove the lack of fully chaotic motion of the Nosé-Hoover thermostat. References to various articles are given in Ref. [17]. In practice, so-called Nosé-Hoover chains are typically used. In this scheme the Nosé-Hoover thermostat is coupled to another thermostat, or, if necessary, to a whole chain of thermostats [15, 30].

Lennard-Jones liquid

Typically, physical systems possess a large number of degrees of freedom $\mathcal{O}(10^{23})$. For such systems, ensemble averages of a system with $3N$ degrees of freedom are essentially identical to those of a system with $3N - 1$ degrees of freedom, i.e, it does not matter much whether the system is confined to a $6N$ dimensional hypersurface in phase space or a $6N - 1$ dimensional one. As we know, microcanonical ensemble averages become identical to canonical ensemble averages in the limit $N \rightarrow \infty$ (the number of conserved quantities is assumed to be much smaller than N). Hence, there is no need to generate a chaotic trajectory on the $6N$ -dimensional hypersurface anymore and many of the isothermal simulation methods are equally well suited for the simulation of systems with many degrees of freedom.

To provide an example, the energy conservation is studied as a function of time step h for a run of n_s steps of a system of Lennard-Jones particles. The three-dimensional periodic system consists of $N = 512$ particles which interact with each other via the potential (12), where the cutoff radius is $r_c = 3\sigma$. The temperature of the system is $k_B T/\epsilon = 1$, the density $\rho\sigma^3 = 0.8$, and the mass $Q/(m\sigma^2) = 2$. The deviation in energy is defined as

$$\Delta H(h) = \frac{1}{n_s} \sum_{n=1}^{n_s} \left| \frac{H(nh) - H(0)}{H(0)} \right|. \quad (67)$$

A total time of $t = n_s h = 1/\sqrt{\epsilon/m\sigma^2}$ is considered. The center of mass motion of the whole system is eliminated initially. The center of mass velocity then remains zero in the course of integration except for numerical errors. The energy is given by Eq. (32) (in physical variables) in general and by Eq. (42) for the Nosé-Hoover equations of motion, where $g = 6(N - 1) - 1$ for $n = -2$ and $g = 3(N - 1)$ for the Nosé-Hoover thermostat.

Figure 6 shows the dependence of ΔH on the time step h for the extended phase-space equations of motion with $n = -2$ (solid line) and the Nosé-Hoover thermostat (dotted line). In addition, results of Newton's equations of motion are presented, which are integrated by the velocity Verlet algorithm (dashed line). In this case the kinetic energy is adjusted such that its mean value corresponds to the thermal energy at the same temperature as that of the other systems. There are only minor difference among the results of the various equations of motion. Hence, each integration algorithm is suitable for molecular dynamics simulations as far as stability is concerned. In Ref. [17], additional results for other integration schemes are presented.

The results of Fig. 6 provide no conclusion on the long time stability of the individual integration schemes. However, a test shows that all of them are stable for long times.

The various curves of Fig. 6 have a slope of approximately two. This is related to fact that they are accurate to $\mathcal{O}(h^3)$. There are higher order integration schemes (even reversible ones), which yield larger slopes and, hence, smaller energy deviations at (very) short times. Thus, such integration schemes are more precise on short time scales, but often lack of long time stability [10].

In summary, I highly recommend to use the velocity Verlet algorithm to solve Newton's equations of motion. In the case of isothermal simulations, integration algorithms derived via the Liouville procedure work very well.

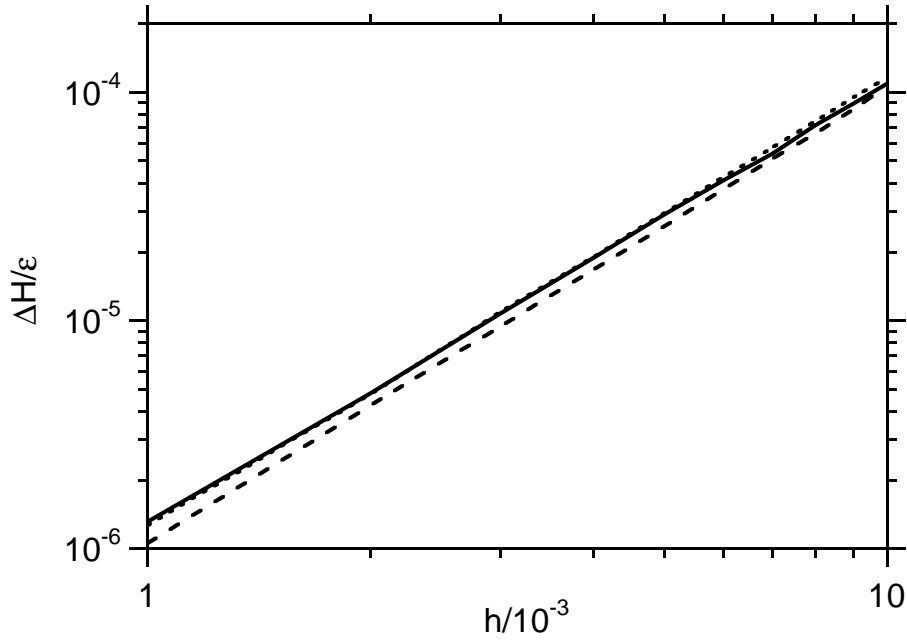


Fig. 6: Conservation of energy as a function of the integration time step for the extended phase-space method with $n = -2$ (solid line), the the Nosé-Hoover thermostat (dotted line), and the velocity Verlet algorithm (dashed line) [17]. The total time is $t = n_s h = 1/\sqrt{\epsilon/m\sigma^2}$.

6.2 Isothermal-isobaric molecular dynamics simulations

Experiments are more often performed under constant pressure conditions than under constant volume conditions. Thus, it is often desirable to perform simulations applying a constant external pressure rather than a constant volume. This particularly applies to systems which exhibit phase transitions.

For a system confined in a container, the meaning of pressure is obvious: It is the external force exerted on the system divided by the area ($p = F/A$). In the case of a periodic system, the meaning of pressure is less obvious. There is a large amount of literature on how to calculate pressure in periodic systems [10, 23, 34–38] and also on how to perform constant pressure simulations in particular when solid-solid structural changes appear [10, 23, 24, 30, 33–40].

There are various ways to derive expressions for the pressure in periodic systems. As shown in Ref. [35], the pressure (p) of a periodic system – corresponding to the mechanical definition $p = F/A$ – is given by

$$p = -\frac{1}{6V} \left\langle \sum_{i=1}^N \sum_{j=1}^N \sum_n' \mathbf{F}_{ij}^n \mathbf{R}_n \right\rangle, \quad (68)$$

where $\mathbf{F}_{ij}^n = \mathbf{F}_{ij}(\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_n)$ is obtained from the pair potential (11) and $\langle \dots \rangle$ is either the time or ensemble average. To make the expression plausible, we assume that all N particles of the system are in the primary box. For short range interactions, there will be only contributions then for $n_\alpha = \pm 1$, where the contributions for $n_\alpha = -1$ are identical to those for $n_\alpha = 1$. In the force calculation, one particle is always an image particle, thus, the forces always act across the formal boundaries of the periodic system. For a cubic box of length L , p can be written as $p = -\sum_{ij} \sum_\alpha F_{ij\alpha} L/3V$. Since the sum gives the total force for $n_\alpha = 1$, the pressure is

indeed equal to the force per area. The factor 3 accounts for the average over the three spacial directions.

Using Newton's equations of motion, we can derive the virial theorem and the virial pressure for a system of periodic boundary conditions. Multiplication of Newton's equations of motion (1), with the forces derived from the potential (11), by \mathbf{r}_i and summation over all particles yields

$$\frac{d}{dt} \sum_{i=1}^N m_i \dot{\mathbf{r}}_i \mathbf{r}_i = \sum_{i=1}^N m_i \dot{\mathbf{r}}_i^2 + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \mathbf{F}_{ij}^{\mathbf{n}} [\mathbf{r}_i - \mathbf{r}_j]. \quad (69)$$

As can easily be shown, the average of the left-hand side of this equation is zero for any standard ensemble. Similarly the time average is zero when the mean square displacement increases with time by a power smaller than two. Thus, we find Clausius' virial theory [41,42]

$$\sum_{i=1}^N \langle m_i \dot{\mathbf{r}}_i^2 \rangle + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \langle \mathbf{F}_{ij}^{\mathbf{n}} [\mathbf{r}_i - \mathbf{r}_j] \rangle = 0, \quad (70)$$

where the term with the forces is denoted as virial. Multiplying this equation with $1/3V$ and adding Eq. (68) at both sides yields the following expression for the pressure

$$p = \frac{1}{3V} \sum_{i=1}^N \langle m_i \dot{\mathbf{r}}_i^2 \rangle + \frac{1}{6V} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \langle \mathbf{F}_{ij}^{\mathbf{n}} [\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}] \rangle. \quad (71)$$

In appendix A.2 the same expression is derived on a statistical and thermodynamical basis. Consequently, we can define the following two expressions for the instantaneous pressure of a periodic system

$$p^{ex} = -\frac{1}{6V} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \mathbf{F}_{ij}^{\mathbf{n}} \mathbf{R}_{\mathbf{n}}, \quad (72)$$

$$p^{in} = \frac{1}{3V} \sum_{i=1}^N m_i \dot{\mathbf{r}}_i^2 + \frac{1}{6V} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \mathbf{F}_{ij}^{\mathbf{n}} [\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}], \quad (73)$$

where we denote p^{ex} as external and p^{in} as internal pressure. The above calculations show that the averages of these two expressions are identical.

To derive equations of motion for a constant pressure simulation, again an extended phase-space is considered. As will be discussed below, such an extension is not necessary for an isobaric simulation, but the integration of the equations of motion of the extended system is less time consuming [35]. Now, the volume V is considered an additional degree of freedom with a mass W and the moment p_V . The Hamiltonian then reads

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\{\mathbf{r}\}, V) + \frac{p_V^2}{2W} + pV. \quad (74)$$

The equations of motion are given by

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i, \quad (75)$$

$$W \ddot{V} = p^{ex} - p, \quad (76)$$

where $p^{ex} = -\partial U/\partial V = -\frac{1}{2}\sum'_{ijn}(\partial U_{ij}/\partial[\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_n])(-\partial \mathbf{R}_n/\partial V) = -\frac{1}{6V}\sum'_{ijn}\mathbf{F}_{ij}^n \mathbf{R}_n$ (72). These equations of motion include the external pressure and are of particular simple form. The equations of motion for the positions are unchanged. The change of the volume leads to a displacement of the image particles, which affects the dynamics of the real particles.

In the limit $W \rightarrow 0$, the external pressure has to be identical to the applied pressure p ($p_V \equiv 0$). Thus, the actual volume – consistent with p and the positions of the particles – has to be determined from the equation $p^{ex}(\{\mathbf{r}\}, V) = p$, i.e., the volume has to be adjusted instantaneously. The conserved quantity is then the enthalpy $H_e = H_0 + pV(\{\mathbf{r}\})$. Strictly speaking, only this ensemble is a constant pressure ensemble. In the case of the extended phase-space (75), (76) the systems are in contact with a pressure reservoir.

The equations of motion (75) and (76) can be integrated using the velocity Verlet algorithm, because the forces and the pressure are independent of velocities.

The isobaric-isoenthalpic equations of motion (75), (76) are easily extended to isobaric-isothermal equations by combining them with a Nosé-Hoover thermostat. The Hamiltonian for the extended (virtual) system is defined as

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i'^2}{2m_i s^2} + U(\{\mathbf{r}'\}, V) + \frac{p_s^2}{2Q} + \frac{p_V'^2}{2W s^2} + pV + gk_B T \ln s. \quad (77)$$

The coupling of the momentum p_V to the variable s ensures that also the volume degree of freedom is thermalized. The canonical equations of motion of the virtual system are given by

$$\frac{d\mathbf{r}_i'}{dt'} = \frac{\mathbf{p}_i'}{m_i s^2}, \quad \frac{d\mathbf{p}_i'}{dt'} = \mathbf{F}_i, \quad (78)$$

$$\frac{ds}{dt'} = \frac{p_s}{Q}, \quad \frac{dp_s}{dt'} = \frac{1}{s} \left(\sum_{i=1}^N \frac{\mathbf{p}_i'^2}{m_i s^2} - gk_B T \right) + \frac{p_V'^2}{W s^3}, \quad (79)$$

$$\frac{dV}{dt'} = \frac{p_V'}{W s^2}, \quad \frac{dp_V'}{dt'} = p^{ex} - p. \quad (80)$$

Transformation to physical variables according to Eqs. (35), (36), and (37) ($n = -1$) as well as $p_V'/s = p_V$ yields

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i}, \quad \frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i - \frac{1}{Q} p_\eta \mathbf{p}_i, \quad (81)$$

$$\frac{d\eta}{dt} = \frac{p_\eta}{Q}, \quad \frac{dp_\eta}{dt} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} + \frac{p_V^2}{W} - gk_B T, \quad (82)$$

$$\frac{dV}{dt} = \frac{p_V}{W}, \quad \frac{dp_V}{dt} = p^{ex} - p - \frac{1}{Q} p_\eta p_V. \quad (83)$$

The equations of motion of the particles (81) are obviously not directly affected by the extension of phase space by the variable V . They are identical to those of the isothermal ensemble. The equation of motion for the thermostat momentum p_η couples now to the momentum p_V and vice versa. Correspondingly, the calculation of the partition function – in analogy to the calculation presented in appendix A.1 – yields now $g = 3N + 1$. If the conservation of the linear momentum of a periodic system is taken into account, $g = 3N - 2$.

Since η and V are bound, the following conclusions can be drawn from these equations of motion. Finite η and V values imply $\dot{\eta} = \dot{V} = 0$. Hence, $\langle \sum_{i=1}^N \mathbf{p}_i^2/m_i \rangle + \langle p_V^2/W \rangle = g$ and⁴

⁴ $\langle p_V p_\eta \rangle = 0$

$\langle p^{ex} \rangle = p$, in accordance with expectations.

So far we derived and discussed equations of motion involving the external pressure. By an additional scaling of the momentum and position variables, equations of motion are obtained with the internal pressure. The Hamiltonian is now defined as

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i'^2}{2m_i s^2 V^{2/3}} + U(\{V^{1/3} \mathbf{r}'\}) + \frac{p_s^2}{2Q} + \frac{p_V'^2}{2W s^2} + pV + gk_B T \ln s. \quad (84)$$

The equations of motion are given by

$$\frac{d\mathbf{r}_i'}{dt'} = \frac{\mathbf{p}_i'}{m_i s^2 V^{2/3}}, \quad \frac{d\mathbf{p}_i'}{dt'} = V^{-1/3} \mathbf{F}_i, \quad (85)$$

$$\frac{ds}{dt'} = \frac{p_s}{Q}, \quad \frac{dp_s}{dt'} = \frac{1}{s} \left(\sum_{i=1}^N \frac{\mathbf{p}_i'^2}{m_i s^2 V^{2/3}} - gk_B T \right) + \frac{p_V'^2}{W s^3}, \quad (86)$$

$$\frac{dV}{dt'} = \frac{p_V'}{W s^2}, \quad \frac{dp_V'}{dt'} = \frac{1}{3V} \left(\sum_{i=1}^N \frac{\mathbf{p}_i'^2}{m_i s^2 V^{2/3}} + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' F_{ij}^{\mathbf{n}} V^{1/3} [\mathbf{r}'_i - \mathbf{r}'_j - \mathbf{n}] \right) - p. \quad (87)$$

The transformations $\mathbf{r}_i = V^{1/3} \mathbf{r}'_i$, $\mathbf{p}_i = \mathbf{p}_i' / (s V^{1/3})$, $\eta = \ln s$, $p_\eta = p_s$, $p_V = p_V' / s$ yield the following equations of motion for the physical variables

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i} + \frac{\dot{V}}{3V} \mathbf{r}_i, \quad \frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i - \frac{1}{Q} p_\eta \mathbf{p}_i - \frac{1}{3VW} p_V \mathbf{p}_i, \quad (88)$$

$$\frac{d\eta}{dt} = \frac{p_\eta}{Q}, \quad \frac{dp_\eta}{dt} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} + \frac{p_V^2}{W} - gk_B T, \quad (89)$$

$$\frac{dV}{dt} = \frac{p_V}{W}, \quad \frac{dp_V}{dt} = p^{in} - p - \frac{1}{Q} p_\eta p_V. \quad (90)$$

The energy H (84), expressed in the physical variables, is an integral of motion. Similar to the above set of equations, $g = 3N + 1$, or, for conserved linear momentum, $g = 3N - 2$. The major difference between Eqs. (81) and (88) is the coupling of the volume to the positions and momenta of the particles. Now, a change of volume not only shifts the image particles, but also causes a displacement of the real particles and changes all momenta.

Integration algorithms for the various isothermal-isobaric equations of motion are easily obtained by the method described in section 6.1. The operator $G(t)$ (59) has to be extended by the volume operators only (see also Ref. [33]).

Aside from an isotropic volume change there is considerable interest in simulation algorithms which take into account full fluctuations of the cell size and shape. Equations of motion for such applications are presented, e.g., in Refs. [10,33,44–46]. Such simulations require a stress tensor rather than a pressure only. The concept of internal and external pressures is easily generalized to stress tensors [35,36].

In section 4.2, truncated potentials have been introduced. Strictly speaking, the pressure (and other quantities, like the energy) should be calculated with the full potential range, otherwise the numerical value is not correct. For a system of particles interacting via pair-potentials only, the mean energy and the pressure can be expressed by the pair distribution function $g(r) = g(|\mathbf{r}_i - \mathbf{r}_j|)$ according to

$$E = \frac{1}{2} \sum_{i=1}^N \langle m_i \dot{\mathbf{r}}_i^2 \rangle + 2\pi N \rho \int_0^\infty r^2 u(r) g(r) r^2 dr, \quad (91)$$

$$pV = \frac{2}{3} \sum_{i=1}^N \langle m_i \dot{\mathbf{r}}_i^2 \rangle - \frac{2}{3} \pi N \rho \int_0^\infty w(r) g(r) r^2 dr, \quad (92)$$

where u is the pair potential, e.g., the Lennard-Jones potential (8), w is the pair virial $w(r) = w(|\mathbf{r}_i - \mathbf{r}_j|) = -\mathbf{F}_{ij}^n[\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_n]$, and ρ is the density [10].

Corrections to the results for the truncated potentials are frequently estimated by assuming $g(r) \approx 1$ beyond the cutoff distance. Hence the full quantities are obtained as follows:

$$E_{\text{full}} = E_c + 2\pi N \rho \int_{r_c}^\infty r^2 u(r) g(r) r^2 dr, \quad (93)$$

$$(pV)_{\text{full}} = (pV)_c - \frac{2}{3} \pi N \rho \int_{r_c}^\infty w(r) g(r) r^2 dr. \quad (94)$$

E_c and $(pV)_c$ are the values determined from the simulation using the truncated potential [10].

7 Long-Range Interactions

For certain interaction it is inevitable to take the long-range part into account. A long-range force is defined as a force which falls off slower than r^{-d} in space, where d is the dimensionality of the system. Important interactions of this kind are the Coulomb interaction between charges ($\sim r^{-1}$) and dipole-dipole interactions ($\sim r^{-3}$). For such potentials, the interactions of a particle with all other particles real as well as image particles, even its own images, have to be taken into account.

An efficient method to sum the interactions between an ion and all its periodic images is the Ewald sum [10, 15, 47, 48]. The Coulomb potential for the periodic system is

$$U = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_n|}. \quad (95)$$

The q_i are the particle charges in *cgs* units.

In the Ewald method, each point charge is surrounded by a charge distribution of equal magnitude but opposite sign, which spreads out radially from the charge (see Fig. 7). Typically this distribution is taken to be Gaussian

$$\rho_i^c(\mathbf{r}) = -q_i \frac{\kappa}{\pi^{3/2}} \exp(-\kappa^2 \mathbf{r}^2), \quad \int \rho_i^c(\mathbf{r}) d\mathbf{r} = -q_i. \quad (96)$$

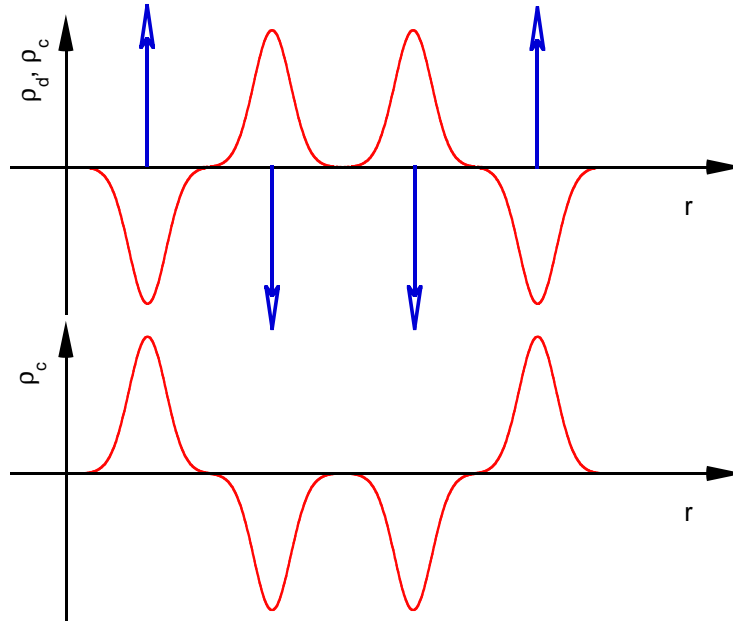


Fig. 7: Charge distribution in the Ewald sum. (top) point charges (ρ^d) plus screening distribution (ρ^c) and (bottom) smooth charge distribution. The arrows represent the δ functions of the point charges.

The parameter κ determines the width of the distribution and is chosen such that the appearing expression can be determined effectively. The total charge density is then of the form

$$\rho(\mathbf{r}) = \sum_{i=1}^N \sum_{\mathbf{n}} [q_i \delta(\mathbf{r} - \mathbf{r}_i - \mathbf{R}_{\mathbf{n}}) + \rho_i^c(\mathbf{r} - \mathbf{r}_i - \mathbf{R}_{\mathbf{n}}) - \rho_i^c(\mathbf{r} - \mathbf{r}_i - \mathbf{R}_{\mathbf{n}})] \quad (97)$$

$$= \sum_{i=1}^N \sum_{\mathbf{n}} q_i \left[\delta(\mathbf{r} - \mathbf{r}_i - \mathbf{R}_{\mathbf{n}}) - \frac{\kappa}{\pi^{3/2}} \exp(-\kappa^2(\mathbf{r} - \mathbf{r}_i - \mathbf{R}_{\mathbf{n}})^2) \right] \quad (98)$$

$$+ \sum_{i=1}^N \sum_{\mathbf{n}} q_i \frac{\kappa}{\pi^{3/2}} \exp(-\kappa^2(\mathbf{r} - \mathbf{r}_i - \mathbf{R}_{\mathbf{n}})^2) \quad (99)$$

$$= \rho_{\text{dir}} + \rho_{\text{rec}}. \quad (100)$$

The additional charge distribution ρ_i^c screens the point charges such that the total interaction is of short-range.

The potential energy of the charge density ρ_{dir} can be calculated without problems [49]. The basic expression

$$U = \frac{1}{2} \sum_{i=1}^N q_i \phi(\mathbf{r}_i) = \frac{1}{2} \sum_{i=1}^N q_i \int \frac{\rho(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|} d\mathbf{r} \quad (101)$$

yields

$$U_{\text{dir}} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' q_i q_j \frac{\text{erfc}(\kappa[\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}])}{|\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}|}. \quad (102)$$

Here, $\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-r^2} dr$ is the complimentary error function. Since U_{dir} is a short-range potential, the minimum image convention is applied in its calculation.

The evaluation of the energy U_{rec} for the charge density ρ_{rec} requires, aside from the calculation of the interaction of the particles in the primary box, the involved summation over all the particles of the other boxes. However, this potential is calculated effectively in Fourier space. Due to the smoothly varying charge density ρ_{rec} , a few Fourier coefficients have to be taken into account only. Following the procedure in Ref. [49], the following potential is obtained

$$U_{\text{rec}} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N q_i q_j \frac{\kappa^3}{\pi^{3/2}} \int \frac{\sum_{\mathbf{n}} \exp(-\kappa^2(\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}})^2)}{|\mathbf{r}_i - \mathbf{r}|} d\mathbf{r} \quad (103)$$

$$= \frac{2\pi}{V} \sum_{\mathbf{k} \neq 0} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{k^2} \exp\left(-\frac{k^2}{4\kappa^2}\right) \cos(\mathbf{k}[\mathbf{r}_i - \mathbf{r}_j]) - \frac{\kappa}{\sqrt{\pi}} \sum_{i=1}^N q_i^2 + \frac{2\pi}{3V} \left| \sum_{i=1}^N q_i \mathbf{r}_i \right|,$$

for a neutral system, i.e., $\sum_i q_i = 0$, with $\mathbf{k} = 2\pi\mathbf{n}/L$. The two last terms are denoted as self-interaction and surface term. The self-interaction is subtracted, because in the first term of Eq. (103) this self-term is contained.

The calculation of U_{rec} has to be performed carefully otherwise an result is obtained without surface term [49]. To overcome the problems caused by the conditionally convergent sums of the potential (95), a convergence factor can be introduced [49] in the charge density ρ_{rec} .

The difficulties in the calculation of the potential U_{rec} originate from the fact that the charge density does not vanish at infinity. Hence, the formula to calculate the potential $\phi = \int \rho(\mathbf{r})/|\mathbf{r} - \mathbf{r}'| d\mathbf{r}'$ is not applicable. Strictly speaking the Poisson equation $\Delta\phi = -4\pi\rho$ should be solved under the appropriate boundary conditions. Due to the convergence factor, the charge density disappears at infinity and the integral can be used to calculate the potential. Alternatively, the potential can be calculated for a finite crystal. At the end the limit to an infinite system has to be performed.

The complete potential energy is

$$U = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum'_{\mathbf{n}} q_i q_j \frac{\text{erfc}(\kappa[\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}])}{|\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}|} \quad (104)$$

$$+ \frac{2\pi}{V} \sum_{\mathbf{k} \neq 0} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{k^2} \exp\left(-\frac{k^2}{4\kappa^2}\right) \cos(\mathbf{k}[\mathbf{r}_i - \mathbf{r}_j]) - \frac{\kappa}{\sqrt{\pi}} \sum_{i=1}^N q_i^2 + \frac{2\pi}{3V} \left| \sum_{i=1}^N q_i \mathbf{r}_i \right|.$$

The forces follow from the relation $\mathbf{F}_i = -\nabla_{\mathbf{r}_i} U$.

The potential (104) applies to an (infinite) system surrounded by vacuum ($\epsilon_r = 1$). It may appear strange that the form of the potential energy of an infinite periodic system of charges depends on the nature of the boundary condition at infinity. It is, however, a real effect with a physical interpretation [15]. For a system surrounded by a conducting medium ($\epsilon_r = \infty$, 'tin foil' boundary conditions), there is no surface term, i.e., the surface term in (104) has to be left out.

To calculate the potential energy (104), the parameter κ , the cutoff radius r_c , and the cutoff k_c in Fourier space have to be specified. It is common to set $k_c = 2\pi n_c/L$ with n_c a positive integer. The total number of Fourier components within this cutoff value is $(4\pi/3)n_c^3$ [15]. The values of these parameters depend on the desired accuracy ϵ_s , that is, the root mean-square difference

between the exact Coulombic energy and the result from the Ewald summation. Expressions for the cutoff errors have been derived in Refs. [50,51]. These calculations yield the expression $\epsilon_s = \exp(-s_\epsilon^2)/s_\epsilon^2$ for the error. The cutoffs are then obtained from the equations

$$r_c = \frac{s_\epsilon}{\kappa}, \quad n_c = \frac{s_\epsilon L \kappa}{\pi}. \quad (105)$$

In practice, the cutoff r_c is set to $L/2$. The other values follow then from the above equations. A reasonable value for κ is $\kappa = 5/L$ [10]. The required computer time is then $\mathcal{O}(N^{3/2})$ [52].

The above Ewald sum is appropriate for systems with a few hundred particles. For larger systems, so-call particle mesh methods are available, which are of $\mathcal{O}(N \ln N)$. The description of the particle mesh methods goes beyond the scope of this lecture. Details of the methods are presented in Refs. [10,54–57].

A Appendices

A.1 Extended phase-space partition function

The partition function for the system in extended phase space is given by

$$Z = \int \delta(H - E) s^n d^N \mathbf{r}' d^N \mathbf{p}' ds dp_s \quad (106)$$

$$= \int \delta \left(\sum_{i=1}^N \frac{\mathbf{p}_i'^2 s^{2n}}{2m_i} + U(\{\mathbf{r}'\}) + \frac{p_s^2}{2Q} + g k_B T \ln s \right) s^n d^N \mathbf{r}' d^N \mathbf{p}' ds dp_s. \quad (107)$$

With the substitution $\mathbf{r}'_i = \mathbf{r}_i$, $\mathbf{p}'_i s^n = \mathbf{p}_i$, it reads

$$Z = \int \delta \left(\sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\{\mathbf{r}\}) + \frac{p_s^2}{2Q} + g k_B T \ln s \right) s^n s^{-3nN} d^N \mathbf{r} d^N \mathbf{p} ds dp_s. \quad (108)$$

Using the rule $\delta(f(x)) = \delta(x - x_0)/|df(x)/dx|_{x=x_0}$, Z transforms into

$$Z = \int \delta \left(s - \exp \left(- \left[\sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\{\mathbf{r}\}) + \frac{p_s^2}{2Q} \right] / g k_B T \right) \right) \times \exp \left(- \left[\sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\{\mathbf{r}\}) + \frac{p_s^2}{2Q} \right] / g k_B T \right) \frac{s^{n(1-3N)}}{g k_B T} d^N \mathbf{r} d^N \mathbf{p} ds dp_s \quad (109)$$

Integration over p_s and s yields

$$Z = C \int \exp \left(- \frac{-3nN + n + 1}{g k_B T} H_0 \right) d^N \mathbf{r} d^N \mathbf{p}. \quad (110)$$

A.2 Derivation of the internal pressure

In an isochoric-isobaric ensemble, the pressure can be calculated by the well known relation

$$p = - \frac{\partial F}{\partial V} = k_B T \frac{\partial \ln Z}{\partial V}. \quad (111)$$

For a canonical ensemble, the partition function Z is given by $Z = \int \exp(-\beta H) d^N \mathbf{r} d^N \mathbf{p}$. Introducing the dimensionless variables $\mathbf{r}'_i = V^{-1/3} \mathbf{r}_i$ and $\mathbf{p}'_i = V^{1/3} \mathbf{p}_i$ leads to the Hamiltonian

$$H' = \frac{1}{V^{2/3}} \sum_{i=1}^N \frac{\mathbf{p}_i'^2}{2m_i} + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' U_{ij}(V^{1/3}[\mathbf{r}'_i - \mathbf{r}'_j - \mathbf{n}]), \quad (112)$$

with the new partition function

$$Z = \int \exp(-\beta H') d^N \mathbf{r}' d^N \mathbf{p}', \quad (113)$$

where the range of $\mathbf{r}'_{i\alpha}$ spans zero to unity [43]. Since the integral bounds are independent of the volume, the pressure is easily calculated by differentiation of the exponential function

$$p = - \frac{\beta}{Z} \int \frac{\partial H'}{\partial V} \exp(-\beta H') d^N \mathbf{r}' d^N \mathbf{p}'. \quad (114)$$

With

$$\begin{aligned}
 \frac{\partial H'}{\partial V} &= -\frac{1}{3V^{2/3}} \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \frac{\partial U_{ij}}{\partial (V^{1/3}[\mathbf{r}'_i - \mathbf{r}'_j - \mathbf{n}])} \frac{1}{3V^{2/3}} [\mathbf{r}'_i - \mathbf{r}'_j - \mathbf{n}] \\
 &= -\frac{1}{3V} \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - \frac{1}{6V} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \mathbf{F}_{ij}^{\mathbf{n}} [\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}],
 \end{aligned} \tag{115}$$

the internal pressure is given by

$$p = \frac{1}{3V} \left\langle \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} \right\rangle + \frac{1}{6V} \left\langle \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n}}' \mathbf{F}_{ij}^{\mathbf{n}} [\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_{\mathbf{n}}] \right\rangle. \tag{116}$$

References

- [1] B. J. Alder, T. E. Wainwright, *J. Chem. Phys.* **27**, 1208 (1957)
- [2] B. J. Alder, T. E. Wainwright, *J. Chem. Phys.* **31**, 459 (1959)
- [3] A. Rahman, *Phys. Rev. A* **136**, 405 (1964)
- [4] F. H. Stillinger, A. Rahman, *J. Chem. Phys.* **60**, 1545 (1974)
- [5] J. A. McCammon, B. R. Gelin, M. Karplus, *Nature* **267**, 585 (1977)
- [6] M. P. Allen, D. J. Tildesley, *Computer Simulation in Chemical Physics* (Kluwer Academic Publishers, Dordrecht, 1993)
- [7] M. Karttunen, A. Lukkari, I. Vattulainen, eds., *Novel Methods in Soft Matter Simulations*, *Lect. Notes Phys.* **640** (Springer-Verlag, Berlin, 2004)
- [8] R. Car, M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985)
- [9] R. W. Hockney, J. W. Eastwood, *Computer Simulations using Particles* (McGraw-Hill, New York, 1981)
- [10] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987)
- [11] D. J. Evans, G. P. Morriss, *Statistical Mechanics of Non-Equilibrium Liquids* (Academic Press, London, 1990)
- [12] D. W. Heermann, *Computer Simulations Methods in Theoretical Physics* (Springer, Berlin, 1990)
- [13] W. G. Hoover, *Molecular Dynamics* (Springer, Berlin, 1990)
- [14] J. M. Haile, *Molecular Dynamics Simulations* (Wiley, New York, 1992)
- [15] D. Frenkel, B. Smit, *Understanding Molecular Simulations* (Academic Press, San Diego, 2002)
- [16] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1978)
- [17] R. G. Winkler, V. Kraus, P. Reineker, *J. Chem. Phys.* **102**, 9018 (1995)
- [18] M. Born, Th. von Karman, *Physik Z.* **13**, 297 (1912)
- [19] L. Verlet, *Phys. Rev.* **159**, 98 (1967)
- [20] M. Tuckerman, B. J. Berne, G. J. Martyna, *J. Chem. Phys.* **97**, 1990 (1992)
- [21] H. F. Trotter, *Proc. Am. Math. Soc.* **10**, 545 (1959)
- [22] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1989)
- [23] H. C. Andersen, *J. Chem. Phys.* **72**, 2384 (1980)

- [24] S. Nosé, J. Chem. Phys. **81**, 511 (1984)
- [25] S. Nosé, Mol. Phys. **52**, 255 (1984)
- [26] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985)
- [27] S. Nosé, J. Phys. Soc. Jpn. **70**, 75 (2001)
- [28] J. Jellinek, J. Phys. Chem. **92**, 3163 (1988)
- [29] R. G. Winkler, Phys. Rev. A **45**, 2250 (1992)
- [30] G. J. Martyna, M. L. Klein, M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992)
- [31] S. Toxvaerd, Mol. Phys. **72**, 159 (1991)
- [32] B. L. Holian, A. J. de Groot, W. G. Hoover, C. G. Hoover, Phys. Rev. A **41**, 4522 (1990)
- [33] G. J. Martyna, M. T. Tuckerman, D. J. Tobias, M. L. Klein, Mol. Phys. **87**, 1117 (1996)
- [34] K. G. Honnell, C. K. Hall, R. Dickman, J. Chem. Phys. **87**, 664 (1987)
- [35] R. G. Winkler, H. Morawitz, D. Y. Yoon, Mol. Phys. **75**, 669 (1992)
- [36] R. G. Winkler, R. Hentschke, J. Chem. Phys. **99**, 5405 (1993)
- [37] R. G. Winkler, J. Chem. Phys. **117**, 2449 (2002)
- [38] B. Oliva, P. H. Hünenberger, J. Chem. Phys. **116**, 6898 (2002)
- [39] G. R. Kneller, T. Mülders, Phys. Rev. E **54**, 6825 (1996)
- [40] W. G. Hoover, Phys. Rev. A **34**, 2499 (1986)
- [41] R. Becker, *Theory of Heat* (Springer, Berlin, 1967)
- [42] R. J. Swenson, Am. J. Phys. **51**, 940 (1983)
- [43] H. S. Green, Proc. R. Soc. (London) Ser. A **189**, 103 (1947)
- [44] M. Parrinello, A. Rahman, Phys. Rev. Lett. **45**, 1196 (1980)
- [45] M. Parrinello, A. Rahman, J. Appl. Phys. **52**, 7182 (1981)
- [46] S. Nosé, M. L. Klein, Mol. Phys. **50**, 1055 (1983)
- [47] P. Ewald, Ann. Phys. **64**, 253 (1921)
- [48] E. Madelung, Phys. Z. **19**, 524 (1918)
- [49] S. W. de Leeuw, J. W. Perram, E. R. Smith, Proc. R. Soc. London A **373**, 27 (1980)
- [50] J. A. Kolafa, J. W. Perram, Mol. Sim. **9**, 351 (1992)
- [51] H. G. Petersen, J. Chem. Phys. **103**, 3668 (1995)

-
- [52] J. Perram, H. G. Petersen, *Mol. Phys.* **65**, 875 (1988)
- [53] R. W. Hockney, J. W. Eastwood, *Computer Simulation Using Particles* (Institute of Physics, London, 1988)
- [54] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. Pedersen, *J. Chem. Phys.* **103**, 8577 (1995)
- [55] M. Deserno, Ch. Holm, *J. Chem. Phys.* **109**, 7678 (1998)
- [56] M. Deserno, Ch. Holm, *J. Chem. Phys.* **109**, 7694 (1998)
- [57] Ch. Holm, in *Computational Soft Matter: From Synthetic Polymers to Proteins*, John von Neumann Winter School 2004, eds. N. Attig, K. Binder, H. Grubmüller, K. Kremer, (NIC-Directors Jülich, 2004)