2 Molecular and stochastic dynamics

In this chapter we discuss the most commonly used ensembles in molecular simulations and the different equations of motion that can be used to generate each ensemble. For generating a canonical ensemble one can use stochastic dynamics with or without inertia. The temperature in a simulation can also be maintained by coupling to a heat bath. We compare the different methods for generating constant temperature trajectories using several test systems.

2.1 Equations of motion

We will briefly introduce two ways to describe classical dynamics. For a more thorough treatment see Goldstein [7] or Landau and Lifshitz [8].

Lagrangian dynamics

Classical dynamics can be described in the most general way with the Lagrangian approach. The equations of motion are expressed using generalized coordinates $\mathbf{q}(t)$. These coordinates do not have to be coordinates in the usual sense, which span a high dimensional space, like Cartesian coordinates, they should be seen as 'just' variables. Given the kinetic energy K and the potential energy V, the Lagrangian L of the system is defined as:

$$L(\dot{\mathbf{q}}(t), \mathbf{q}(t), t) = K - V \tag{2.1}$$

where $\dot{\mathbf{q}}(t)$ is the derivative of $\mathbf{q}$ with respect to time:

$$\dot{\mathbf{q}}(t) = \frac{\mathrm{d}\mathbf{q}(t)}{\mathrm{d}t} \tag{2.2}$$

The kinetic and potential can both be a function of $\dot{\mathbf{q}}$, $\mathbf{q}$ and t explicitly. Lagrange's equations of motion are:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial \dot{q}_i} = 0 \ , \quad i = 1, \ldots, N \tag{2.3}$$

The elegance of this formulation is that it is completely independent of the choice of coordinates $\mathbf{q}$. This is convenient when the potential energy is known in some appropriate coordinate system. For Cartesian coordinates $\mathbf{q} = \mathbf{x}$, with the kinetic energy given by:

$$K(\dot{\mathbf{x}}) = \sum_i \frac{m_i}{2}(\dot{x}_i)^2 \tag{2.4}$$

and a potential energy which only depends on the coordinates, Lagrange equations of motion reduce to Newton's equations of motion:

$$m_i \frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} = f_i \tag{2.5}$$

where the forces are defined as minus the gradient of the potential:

$$f_i = -\frac{\partial}{\partial x_i} V(\mathbf{q}) \tag{2.6}$$

The Lagrange formulation is most useful for solving the equations of motion, but for statistical mechanics the Hamiltonian formulation is more convenient.

Hamiltonian dynamics

The main conceptual difference between the Lagrangian and the Hamiltonian formulation is that in the latter the generalized velocities q_i are replaced by the generalized momenta p_i:

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.7}$$

The generalized momenta are said to be conjugate to the generalized coordinates. The Hamiltonian H is a function of the generalized momenta and coordinates and it can depend on time explicitly:

$$H(\mathbf{p}(t), \mathbf{q}(t), t) = \sum_i p_i \dot{q}_i - L(\dot{\mathbf{q}}(t), \mathbf{q}(t), t) \tag{2.8}$$

The equations of motion can be derived from this definition [7]:

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{2.9}$$

$$-\dot{p}_i = \frac{\partial H}{\partial q_i} \tag{2.10}$$

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} \tag{2.11}$$

Equations (2.9) and (2.10) are called the canonical equations of Hamilton.

If the generalized coordinates can be expressed as a function of Cartesian coordinates, the kinetic energy will be a quadratic function of the generalized velocities, with coefficients which depend on the generalized coordinates, but not explicitly on time:

$$K = \frac{1}{2} \sum_{ij} a_{ij}(\mathbf{q})\, \dot{q}_i \dot{q}_j \tag{2.12}$$

$$a_{ij} = \sum_k m_k \frac{\partial x_k}{\partial q_i} \frac{\partial x_k}{\partial q_j} \tag{2.13}$$

When the potential is not a function of the generalized velocities, the momenta are given by:

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial K}{\partial \dot{q}_i} = \sum_j a_{ij}(\mathbf{q})\, \dot{q}_i \tag{2.14}$$

Using (2.8) it can be shown that under these conditions the Hamiltonian is equal to the total energy:

$$H = K + V \tag{2.15}$$

From (2.11) we can see that when the potential does not depend explicitly on time the total energy is conserved. Such systems are called conservative.

2.2 Molecular dynamics

In molecular dynamics (MD) the classical equations of motion are solved in time for a set of atoms. The interactions between the atoms, which are in principle given by the Schrödinger equation, are approximated by mainly pairwise potentials. Some bonded interactions, like angles and dihedrals, involve more-body terms. The accuracy of a molecular dynamics simulation is mainly determined by the quality of the inter-atomic potentials, which together are called the forcefield.

Simulations which include solvent are almost always performed in Cartesian coordinates. When no constraints are present, the system obeys Newton's equations of motion. When one or more inter-atomic distances are constrained one can remove these coordinates and use generalized coordinates, as described for Hamilton's equation of motion. For analytical derivations this can be useful. But when simulating the system on a computer it is often more convenient to simulate in Cartesian space and correct the distances between the constrained particles using a constraint algorithm like SHAKE [9] or LINCS [10]. A simulation of a single molecule with fixed bond lengths and bond angles might be performed more efficiently in so called internal coordinates, which are usually dihedral angles.

2.3 Ensembles used in MD

Pure conservative Hamiltonian simulations of many-particle systems yield equilibrium distributions in $(\mathbf{p}, \mathbf{q})$ space (called phase space) with constant energy. By proper modification of the equations of motion, other ensembles can be generated. There are four ensembles which are commonly used in molecular simulations. These ensembles are indicated by the three thermodynamic quantities which are kept constant. The ensembles are: the constant-NVE or microcanonical ensemble, the constant-NVT or canonical ensemble, the constant-NPT or isothermal-isobaric ensemble and the constant-μVT or grand canonical ensemble. Another possible ensemble is the constant-μPT ensemble.

When MD is performed using Newton's equations of motion, the total energy is conserved, assuming we integrate accurately with a stable integration scheme. This results in a microcanonical ensemble. When we want to obtain a different ensemble, the equations of motion need to be modified. A way of doing this is the extended system approach. To obtain a canonical ensemble, the temperature can be added as a generalized coordinate to the Lagrangian. This approach is called the Nosé-Hoover scheme, or coupling with the Nosé-Hoover thermostat [11]. Although a canonical ensemble is obtained, the temperature will oscillate around the average, where the period of the oscillations depends on the effective "mass" associated with the temperature coordinate in the Lagrangian. This can result in unwanted kinetic effects. A method which leaves the fast kinetics unaffected is weak coupling to a bath, also called the Berendsen thermostat [12]. The

weak coupling is achieved with a first other differential equation for the temperature. This causes any deviations from the reference temperature to decay exponentially. The main disadvantage of the Berendsen thermostat is that it is unknown if it generates an ensemble. Any temperature coupling method will affect the long-term velocity correlations in the system. For the Berendsen thermostat a large part of these correlations can be retrieved after performing the simulation [13].

Methods for keeping the pressure constant in an MD simulation are analogous to the temperature coupling method. The volume of the periodic unit-cell, or all three vectors of the unit-cell can be added as an extra generalized coordinate in the Lagrangian. This method is called Parrinello-Rahman pressure coupling [14]. In combination with the Nosé-Hoover thermostat this generates an isothermal-isobaric ensemble. Just like the Nosé-Hoover thermostat, Parrinello-Rahman coupling gives oscillations of the unit-cell vectors. Berendsen coupling can also be used for the pressure. Again, this has the advantage that the fast kinetics are unaffected and the disadvantage that it is unknown if an ensemble is generated.

2.4 Monte Carlo

When one is not interested in the dynamics Monte Carlo simulations can be used to generate a canonical ensemble. The most simple algorithm is Metropolis Monte Carlo [15]. Starting from a configuration $\mathbf{r}^n$ a trial move $\mathbf{r}'$ is generated by giving one or more particles a random displacement. The trial move is accepted, $\mathbf{r}^{n+1} = \mathbf{r}'$, with probability:

$$\min\left(1, \exp\left(-\frac{1}{k_B T}\left[V\left(\mathbf{r}^{n+1}\right) - V\left(\mathbf{r}^n\right)\right]\right)\right) \tag{2.16}$$

If the trial move is not accepted then $\mathbf{r}^{n+1} = \mathbf{r}^n$. It is essential that the random displacement should not introduce a bias into the sampling. For some types of biases it is possible apply a correction to the probability 2.16, an example is given in section 2.7. The efficiency of the algorithm depends on the choice of the random displacement. When the displacement is very small almost every move will be accepted, but it will take a large number of moves to cover the whole available phase space. When the displacement is very large the acceptance ratio is very small. One can also think of complicated non-Cartesian moves. For instance, (combined) dihedral moves, which are especially efficient for proteins. An advantage of Monte Carlo over MD is that only the energy needs to be calculated, not the forces. For small systems Monte Carlo is more efficient in sampling than MD. But for large systems, such as proteins in explicit solvent, Monte Carlo is less efficient, since there will be large, collective motions present.

2.5 Langevin dynamics

Another method to obtain a canonical ensemble is Langevin dynamics (LD). In LD a frictional force, proportional to the velocity, is added to the conservative force. The friction removes kinetic energy from the system. A random force adds kinetic energy to the system. To generate a canonical ensemble, the friction and random force have to obey the fluctuation-dissipation theorem [6].

In general a Langevin dynamics system arises from a classical system by removing degrees of freedom. The degrees of freedom which are removed exert conservative and frictional forces on the rest of the system. All other forces are assumed to add up to a random force. The textbook example is a colloidal particle in a solvent. When we consider only the degrees of freedom of the colloidal particle, we can describe the system with Langevin dynamics. The frictional and random forces are caused by collisions of solvent molecules with the colloidal particle.

A formal way of deriving Langevin dynamics is the projection operator formalism of Zwanzig[16, 17, 18] and Mori[19, 20]. In this approach the phase space is divided into two parts, which we will call interesting and uninteresting degrees of freedom (dof), For the approach to be useful, the uninteresting dof should be rapidly varying with respect to the interesting dof. Mori introduced two projection operators, which project the whole phase space into the sets of interesting and uninteresting dof respectively. These projection operators are used to project the full equations of motion into the set of interesting dof. The result is a differential equation with three force terms: a mean force between the interesting dof, a dissipative or frictional force exerted by the uninteresting dof onto the interesting coordinates and a rest term which consists of forces which are neither correlated with the positions nor with the velocities. The frictional and uncorrelated force have the same memory function, the shape of which is given by the projection operator formalism. When the uncorrelated force is approximated by a random force the interesting dof are independent of the uninteresting dof, so the number of dof can be reduced. While this method is formally completely correct, it is of no practical use, since it does not give any hint on how to choose the division in interesting and uninteresting dof. Even if one already has such a division the method is of little use, since obtaining the potential of mean force and the memory function is often computationally more expensive than a simulation with all dof.

We will not treat Langevin dynamics with memory in the noise term. The formal equations of motion for LD for a one-dimensional system with a δ-function as friction kernel are:

$$m\,\mathrm{d}v = -m\,\xi\,v\,\mathrm{d}t + f(r)\,\mathrm{d}t + \mathrm{d}W \tag{2.17}$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = v \tag{2.18}$$

where ξ is the friction constant and W is a Wiener process [21]. The conditions for the

stochastic process $[W(t), t \geq 0]$ to be a Wiener process are:

1. $[W(t), t \geq 0]$ has stationary independent increments
2. $W(t)$ is normally distributed for $t \geq 0$
3. $\langle\, W(t)\,\rangle = 0$ for $t \geq 0$
4. $W(0) = 0$

The only parameter with has not been specified is the spread of the distribution of $W(t)$, this is given by:

$$\langle\, W(t)\, W(t+\tau)\,\rangle = 2\, m\, k_B T\, \xi\, \tau \tag{2.19}$$

With this spread a canonical ensemble is generated for temperature T. Langevin equations are often written in differential form, which can be obtained by dividing both sides of equation (2.17) by $\mathrm{d}t$. Formally this is incorrect, since $W(t)$ is not differentiable. The differential form for the Langevin equation in Cartesian coordinates is:

$$m_i \frac{\mathrm{d}^2 \mathbf{r}_i}{\mathrm{d}t^2} = -m_i \sum_j \mathbf{\Xi}_{ij} \frac{\mathrm{d}\mathbf{r}_j}{\mathrm{d}t} + \mathbf{f}_i(\mathbf{r}) + \boldsymbol{\phi}_i \tag{2.20}$$

where the subscript denotes the particle number, so $\mathbf{r}_i$ are the coordinates of particle i, $\mathbf{\Xi}_{ij}$ is a 3×3 friction matrix, $\mathbf{f}(\mathbf{r}) = -\nabla V(\mathbf{r})$ and $\boldsymbol{\phi}_i$ is a random force. Without memory the noise is Gaussian distributed and δ-correlated in time:

$$\left\langle\, \boldsymbol{\phi}_i(t)\, \boldsymbol{\phi}_j^T(t+\tau)\,\right\rangle = 2\sqrt{m_i m_j} k_B T\, \mathbf{\Xi}_{ij} \delta(\tau) \tag{2.21}$$

where k_B is Boltzmann's constant, T is the temperature and $\delta(\tau)$ is the Dirac delta function. In practice the friction matrix is often (chosen) diagonal and isotropic, which simplifies the equations of motion to:

$$m_i \frac{\mathrm{d}^2 \mathbf{r}_i}{\mathrm{d}t^2} = -m_i \xi_i \frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} + \mathbf{f}_i(\mathbf{r}) + \boldsymbol{\phi}_i \tag{2.22}$$

where ξ_i is the friction constant of particle i. The noise correlation is given by:

$$\langle\, \phi_{ik}(t)\, \phi_{jl}(t+\tau)\,\rangle = 2 m_i k_B T\, \xi_i\, \delta_{ij} \delta_{kl} \delta(\tau) \tag{2.23}$$

where k and l indicate the Cartesian vector components and δ_{ij} is the Kronecker delta.

The probability density $\rho(\mathbf{r}, \mathbf{p}, t)$ of trajectories generated by equation (2.22), also called the phase space density, obeys the Fokker-Planck equation:

$$\begin{aligned} \frac{\partial}{\partial t}\rho(\mathbf{r}, \mathbf{p}, t) + \sum_i \left(\frac{\mathbf{p}_i}{m_i} \cdot \nabla_{\mathbf{r}_i} \rho(\mathbf{r}, \mathbf{p}, t) + \mathbf{f}_i \cdot \nabla_{\mathbf{p}_i} \rho(\mathbf{r}, \mathbf{p}, t) \right) = \\ \sum_i \xi_i \nabla_{\mathbf{p}_i} \cdot (\mathbf{p}_i \rho(\mathbf{r}, \mathbf{p}, t) + m_i k_B T \nabla_{\mathbf{p}_i} \rho(\mathbf{r}, \mathbf{p}, t)) \end{aligned} \tag{2.24}$$

For a proof of the equivalence of (2.22) and (2.24) see Stratonovich [22]. The stationary solution of the Fokker-Planck equation is the canonical ensemble:

$$\rho(\mathbf{r},\mathbf{p}) \propto \exp\left(-\frac{1}{k_B T}\left(V(\mathbf{r}) + \sum_i \frac{1}{2m_i}\mathbf{p}_i \cdot \mathbf{p}_i\right)\right) \tag{2.25}$$

We have left out the normalization constant, which is, by definition, one over the integral of the term on the right-hand side of (2.25) over the whole phase space ($\mathbf{r}$ and $\mathbf{p}$).

To numerically integrate equation (2.22), we use the complicated third order scheme of Van Gunsteren and Berendsen [23]. In the limit of zero friction this scheme converges to the leap-frog integration scheme for Molecular Dynamics.

The harmonic oscillator

The stochastic differential equation can be integrated analytically when the force depends linearly on the position. For a one dimensional system with a harmonic potential equation (2.22) becomes:

$$m\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -m\xi\frac{\mathrm{d}r}{\mathrm{d}t} - kr + \phi(t) \tag{2.26}$$

The solution is:

$$\begin{aligned} r(t) &= \frac{1}{2\sqrt{d}}\left(e^{-\beta_0 t}\left(\beta_1 r(0) + v_0 + \frac{1}{m}\int_0^t e^{\beta_0 s}\phi(s)\mathrm{d}s\right)\right. \\ &\qquad \left. -e^{-\beta_1 t}\left(\beta_0 r(0) + v_0 + \frac{1}{m}\int_0^t e^{\beta_1 s}\phi(s)\mathrm{d}s\right)\right) \qquad (2.27) \\ d &= \frac{\xi^2}{4} - \frac{k}{m} \qquad (2.28) \\ \beta_0 &= \frac{\xi}{2} - \sqrt{d} \qquad (2.29) \\ \beta_1 &= \frac{\xi}{2} + \sqrt{d} \qquad (2.30) \end{aligned}$$

The autocorrelation function of the velocity ($v = \mathrm{d}r/\mathrm{d}t$) can be calculated directly from this:

$$\begin{aligned} \langle\, v(0)v(t)\,\rangle &= \frac{1}{2\sqrt{d}}\left(-\beta_0 e^{-\beta_0 t} + \beta_1 e^{-\beta_1 t}\right)\left\langle\, v_0^2\,\right\rangle \qquad (2.31) \\ &= \frac{1}{2\sqrt{d}}\frac{k_B T}{m}\left(-\beta_0 e^{-\beta_0 t} + \beta_1 e^{-\beta_1 t}\right) \qquad (2.32) \end{aligned}$$

where we have used that the variance of the velocity is $k_B T/m$.

The over-damped limit

For small ξ the dynamics is close to that of a harmonic oscillator. A large difference with deterministic dynamics ($\xi = 0$) is that the amplitude of the oscillations varies in time. As ξ increases the period of the oscillations decreases. When $\alpha = 4k/m\xi^2 < 1$, all oscillatory behavior is lost. In the latter case we can make an expansion for small α:

$$\alpha < 1 \implies \sqrt{d} = \frac{\xi}{2} - \frac{k}{m\xi} + O\left(\alpha^2\right) \implies \begin{cases} \beta_0 = \dfrac{k}{m\xi} + O\left(\alpha^2\right) \\ \beta_1 = \xi - \dfrac{k}{m\xi} + O\left(\alpha^2\right) \end{cases} \tag{2.33}$$

Looking at the velocity autocorrelation (2.32) we can see that it is a sum of two exponentials. First there is a decay with time constant $1/\xi$, on long time scales there is a decay with time constant $m\xi/k$. The pre-factor of the slow decay is negative and a factor α smaller than the one of the fast decay. Thus inertia effects become negligible on time-scales longer than $1/\xi$. When $k > 0$, the integral over the velocity autocorrelation is zero, which is correct, since the mass can not escape from the harmonic potential. When $k = 0$, the system is purely diffusive. In this case the integral over the velocity autocorrelation is D, which corresponds to the Green-Kubo formula.

When no force is present, the position is given by:

$$r(t) = r(0) + \frac{1}{\xi}\left(\left(1 - e^{-\xi t}\right) v_0 + \frac{1}{m}\int_0^t \phi(s)\mathrm{d}s + \frac{1}{m}e^{-\xi t}\int_0^t e^{\xi s}\phi(s)\mathrm{d}s\right) \tag{2.34}$$

The mean square displacement can be calculated directly from (2.34):

$$\begin{aligned} \left\langle \left(r(t) - r(0)\right)^2 \right\rangle &= \frac{1}{\xi^2}\left(1 - e^{-\xi t}\right)^2 \left\langle v_0^2 \right\rangle \\ &\quad + \frac{2k_BT}{m\xi}\left(t + \frac{1}{\xi}\left(\frac{3}{2} - e^{-\xi t} - \frac{1}{2}e^{-2\xi t}\right)\right) && (2.35) \\ &= 2D\left(t + \frac{2}{\xi}\left(1 - e^{-\xi t}\right)\right) && (2.36) \end{aligned}$$

where $D = k_BT/m\xi$ is the diffusion constant. Note that the mean square displacement converges to the correct value of $2Dt$. In the derivation of expression (2.35) we used the following two expectations of integrals over stochastic functions:

$$\begin{aligned} \left\langle \left(\int_0^t \phi(s)\mathrm{d}s\right)^2 \right\rangle &= \int_0^t\int_0^t \left\langle \phi(s)\phi(u) \right\rangle \mathrm{d}u\mathrm{d}s && (2.37) \\ &= \int_0^t 2mk_BT\xi\mathrm{d}s && (2.38) \\ &= 2mk_BT\xi t && (2.39) \end{aligned}$$

$$\left\langle e^{-at}\int_0^t e^{as}\phi(s)\mathrm{d}s\; e^{-bt}\int_0^t e^{bs}\phi(s)\mathrm{d}s\right\rangle \tag{2.40}$$

$$= e^{-(a+b)t}\int_0^t\int_0^t e^{as+bu}\left\langle\,\phi(s)\phi(u)\,\right\rangle \mathrm{d}u\mathrm{d}s \tag{2.41}$$

$$= e^{-(a+b)t}\int_0^t e^{(a+b)s}2mk_BT\xi\mathrm{d}s \tag{2.42}$$

$$= \frac{2mk_BT\xi}{a+b}\left(1-e^{-(a+b)t}\right) \tag{2.43}$$

2.6 Brownian dynamics

When the friction is high, correlations in the velocity will decay in a period over which changes in the force are negligible. Such a system is called overdamped. In this case the left-hand side of equation (2.22) can be neglected, after averaging over short times. The result is Brownian dynamics, which is described by the position Langevin equation:

$$\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = \frac{1}{\gamma_i}\mathbf{f}_i(\mathbf{r}) + \boldsymbol{\upsilon}_i(t) \tag{2.44}$$

where $\gamma_i = m_i\xi_i$ is the friction coefficient of particle i. The random velocities are δ-correlated with variance:

$$\left\langle\,\upsilon_{ik}(t)\upsilon_{jl}(t+\tau)\,\right\rangle = \frac{2k_BT}{\gamma_i}\delta_{ij}\delta_{kl}\delta(\tau) \tag{2.45}$$

In Brownian dynamics the friction coefficient γ scales the time. This means that one only needs to perform a simulation at one value of γ. Afterwards the time of the results can be scaled to obtain results for any value of γ.

For the harmonic oscillator we can perform a more rigorous derivation. When $\alpha = 4k/m\xi^2 \ll 1$, correlations in the velocity decay exponentially with time constant $1/\xi$. Changes in the position are negligible on this time scale. The separation of time scales makes it possible to exchange the second order stochastic differential equation (2.26) for a first order stochastic differential equation without affecting the dynamics on time scales longer than $1/\xi$. To prove this we will look at displacements over a time τ for which the following inequalities hold:

$$1 \ll \xi\tau \ll \frac{1}{\alpha} \tag{2.46}$$

So over time scales of length τ the velocity has lost its correlation, but the effects of the potential are still negligible compared to the random forces. The position at time τ is given by (2.27), in which all terms with $-\beta_1\tau$ in the exponent can be neglected, except

for the integral $e^{-\beta_1\tau}\int e^{\beta_1 s}\phi \mathrm{d}s$, which we will call Q:

$$\begin{aligned}
r(\tau) &= \frac{1}{2\sqrt{d}}e^{-\beta_0\tau}\left(\beta_1 r(0)+v_0+\frac{1}{m}\int_0^\tau e^{\beta_0 s}\phi(s)\mathrm{d}s\right)+\frac{Q}{2m\sqrt{d}} \\
&\quad + O\left(e^{-\xi\tau}\right) && (2.47)\\
&= \frac{1}{\xi}e^{-\beta_0\tau}\left(\xi r(0)+v_0+\frac{1}{m}\int_0^\tau e^{\beta_0 s}\phi(s)\mathrm{d}s\right)+\frac{Q}{m\xi}+O\left(\alpha,e^{-\xi\tau}\right) && (2.48)\\
&= e^{-\beta_0\tau}\left(r(0)+\frac{1}{m\xi}\int_0^\tau e^{\beta_0 s}\phi(s)\mathrm{d}s\right)+\frac{Q}{m\xi}+O\left(\sqrt{\alpha},e^{-\xi\tau}\right) && (2.49)\\
&= \left(1-\frac{\alpha\xi\tau}{4}\right)r(0)+\frac{1}{m\xi}e^{-\beta_0\tau}\int_0^\tau e^{\beta_0 s}\phi(s)\mathrm{d}s+\frac{Q}{m\xi} \\
&\quad + O\left(\sqrt{\alpha},(\alpha\xi\tau)^2,e^{-\xi\tau}\right) && (2.50)\\
&= \left(1-\frac{\alpha\xi\tau}{4}\right)r(0)+\frac{1}{m\xi}e^{-\beta_0\tau}\left(\int_0^\tau \phi(s)\mathrm{d}s+\int_0^\tau\left(e^{\beta_0 s}-1\right)\phi(s)\mathrm{d}s\right) \\
&\quad + \frac{Q}{m\xi}+O\left(\sqrt{\alpha},(\alpha\xi\tau)^2,e^{-\xi\tau}\right) && (2.51)\\
&= r(0)-\frac{k}{m\xi}\tau r(0)+\frac{1+O(\alpha\xi\tau)}{m\xi}\int_0^\tau\phi(s)\mathrm{d}s \\
&\quad + \frac{1}{m\xi}e^{-\beta_0\tau}\int_0^\tau\left(1-e^{\beta_0 s}\right)\phi(s)\mathrm{d}s+\frac{Q}{m\xi}+O\left(\sqrt{\alpha},(\alpha\xi\tau)^2,e^{-\xi\tau}\right) && (2.52)
\end{aligned}$$

where in going from (2.48) to (2.49) we have used that v_0 is on average $\sqrt{a}$ smaller than $\xi r(0)$, since in the canonical distribution $\langle\, v_0 v_0\,\rangle = k_B T/m$ and $\langle\, r(0)\, r(0)\,\rangle = k_B T/k$. The variances of the three stochastic integrals in (2.52) are:

$$\left\langle\left(\int_0^\tau\phi(s)\mathrm{d}s\right)^2\right\rangle = 2mk_BT\xi\tau \qquad (2.53)$$

$$\left\langle\left(e^{-\beta_0\tau}\int_0^\tau\left(e^{\beta_0 s}-1\right)\phi(s)\mathrm{d}s\right)^2\right\rangle = \left(\frac{1}{3}(\alpha\xi\tau)^2+O\left((\alpha\xi\tau)^3\right)\right)2mk_BT\xi\tau \qquad (2.54)$$

$$\left\langle\, Q^2\,\right\rangle = \left\langle\left(e^{-\beta_1\tau}\int_0^\tau e^{\beta_1 s}\phi(s)\mathrm{d}s\right)^2\right\rangle = 2mk_BT+O\left(\alpha,e^{-\xi\tau}\right) \qquad (2.55)$$

The second integral is $\alpha\xi\tau$ smaller than the first and Q is $\sqrt{\xi\tau}$ smaller than the first. From (2.52) we can calculate the displacement from time 0 to τ:

$$r(\tau)-r(0) = -\frac{k}{m\xi}\tau r(0)+\frac{1}{m\xi}\int_0^\tau\phi(s)\mathrm{d}s+O\left(\sqrt{\xi\tau},\sqrt{\alpha}\right) \qquad (2.56)$$

When we are not interested in the short time behavior, we can ignore the higher order terms in $\xi\tau$ and α. When we divide both sides of (2.56) by τ and take the limit of τ to 0,

we obtain a first order Langevin equation:

$$\frac{\mathrm{d}r}{\mathrm{d}t} = -\frac{k}{m\xi}r + \frac{1}{m\xi}\phi(t) \tag{2.57}$$

This equation will produce the same behavior as the Langevin dynamics equation (2.26) on time scales longer than $1/\xi$. The short time behavior will be completely different. The dynamics described by (2.57) is called position Langevin dynamics or Brownian dynamics. The derivation can be generalized to non-harmonic potentials by replacing k by the second derivative of the potential.

The equation for the probability density of trajectories generated by (2.44) is the Schmoluchowski equation:

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) + \sum_i \frac{1}{\gamma_i}\nabla_{\mathbf{r}_i}\cdot(\mathbf{f}_i\rho(\mathbf{r},t)) = \sum_i \frac{k_BT}{\gamma_i}\Delta_{\mathbf{r}_i}\rho(\mathbf{r},t) \tag{2.58}$$

The stationary solution is a canonical distribution for the positions:

$$\rho(\mathbf{r}) \propto \exp\left(-\frac{1}{k_BT}V(\mathbf{r})\right) \tag{2.59}$$

To numerically integrate equation (2.44), we use a simple first order Euler scheme:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \frac{\Delta t}{\gamma_i}\mathbf{f}_i(\mathbf{r}(t)) + \boldsymbol{\eta}_i(t) \tag{2.60}$$

where Δt is the time step. The random displacements $\boldsymbol{\eta}_i(t)$ are Gaussian distributed with mean zero and variance:

$$\left\langle \boldsymbol{\eta}_{ik}(t)\boldsymbol{\eta}_{jl}(t+n\Delta t)\right\rangle = \frac{2k_BT\Delta t}{\gamma_i}\delta_{ij}\delta_{kl}\delta_{0n} \tag{2.61}$$

Also in this discretized version of Brownian dynamics γ scales time, as Δt always occurs divided by γ.

2.7 Smart Monte Carlo

The integration scheme for Brownian Dynamics (equation (2.60)) looks like a Monte Carlo algorithm, except that there is no acceptance criterion. The Metropolis criterion (2.16) can not be used, since in addition to the random displacement there is a bias, which is proportional to the force. The correct acceptance probability was derived by Rossky, Doll and Friedman [24]:

$$\min\left(1, \exp\left(-\frac{1}{k_BT}\left[V' - V + \frac{1}{2}(\mathbf{f}'+\mathbf{f})\cdot(\mathbf{r}'-\mathbf{r}) + \frac{1}{4}\frac{\Delta t}{\gamma}(\mathbf{f}'+\mathbf{f})\cdot(\mathbf{f}'-\mathbf{f})\right]\right)\right) \tag{2.62}$$

where the primed quantities are at "time" $t + \Delta t$ and the unprimed quantities are at "time" t. When the new configuration is not accepted $\mathbf{r}(t + \Delta t)$ is identical to $\mathbf{r}(t)$. The resulting algorithm is called smart Monte Carlo. We quoted the word time because there is no real time in a Monte Carlo simulation. In practice however, one can chose Δt such that the acceptance ratio is almost 100%. Then t can still be interpreted as a physical time. Smart Monte Carlo makes it possible to increase the time step, and thus the sampling efficiency, without having to worry about the integration accuracy.

2.8 Simulations of a Lennard-Jones liquid

Using stochastic dynamics or Brownian dynamics for generating a canonical ensemble changes the dynamics compared to Newtonian dynamics. To analyze the changes in the dynamics we simulated one of the most simple systems of interacting particles: a Lennard-Jones liquid. This is a system of particles that interact via a pairwise, so-called, Lennard-Jones potential:

$$V_{ij}(r_{ij}) = 4\epsilon \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right) \tag{2.63}$$

where r_{ij} is the distance between particle i and j. We can make the system dimensionless using σ, ϵ, the mass of the particles m and the volume of the simulation box V. The quantities we will need are:

$$t^* = \frac{1}{\sigma}\sqrt{\frac{\epsilon}{m}}t \tag{2.64}$$

$$\xi^* = \sigma\sqrt{\frac{m}{\epsilon}}\xi \tag{2.65}$$

$$D^* = \frac{1}{\sigma}\sqrt{\frac{m}{\epsilon}}D \tag{2.66}$$

$$T^* = \frac{k_B T}{\epsilon} \tag{2.67}$$

$$\rho^* = \frac{N\sigma^3}{V} \tag{2.68}$$

where N is the number of particles. To enable simulation of the system with a computer, periodic boundary conditions are introduced, meaning that the simulation box is surrounded by an infinite number of identical copies, which are translated by an integer number of box lengths in each direction. We used a cubic box of volume V which contains 1000 particles. The reduced density and reduced temperature are both 1.

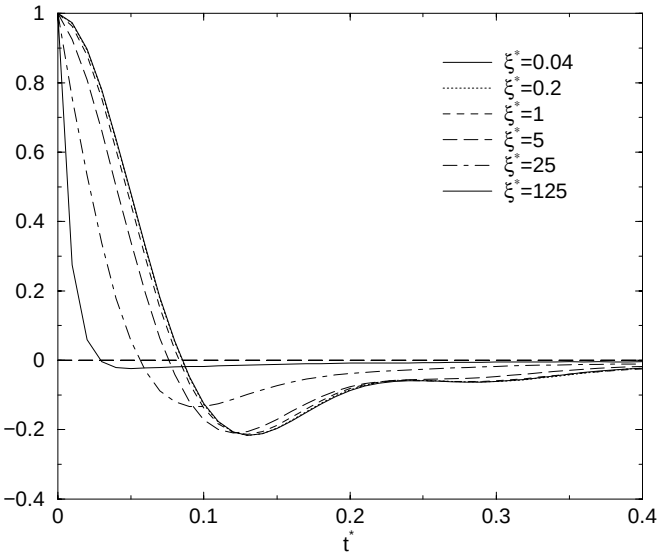

Figure 2.1: Velocity autocorrelation function for a Lennard-Jones liquid simulated with LD for 6 different values of the friction constant ξ^*.

LD simulations

We performed LD simulations for 6 values of ξ^*, going from 0.04 to 125 in steps of a factor 5. The time step was chosen as $\min(0.005, 0.0625\xi^*)$ to insure accurate integration for both the low and high friction regime. A neighbor list was used, which was regenerated every 20 steps with a cut-off of 4σ. Two simulations were performed for each value of ξ^*, a short one of $20\tau^*$ for determining the short time correlations and a longer one of $100\tau^*$ for determining the diffusion constant. All simulations were started from the same equilibrated conformation, using different random seeds. The velocity autocorrelation for all values of ξ^* is shown in Figure 2.1. The plots for the two smallest friction constants coincide, because on this time scale the friction is negligible. The autocorrelation initially decays slowly, followed by a steep decrease. As the friction constant increases this initial "plateau" is lost and the decrease is purely exponential. To get an estimate of the decay time we plotted the time at which the velocity autocorrelation is $1/e$ (Figure 2.2). This time converges to $1/\xi^*$ in the limit of high friction. The diffusion constant was determined by fitting a straight line to the mean square displacement, using only the points from t^*=20 to 60. The results are shown in Figure 2.2. In the high friction limit the diffusion constant converges to a value which is approximately a factor 10 smaller than for the diffusion of non-interacting particles. This means that for large friction constants the interactions decrease the diffusion by a factor of 10. Although the dynamics is dominated by the random forces, the interactions reduce the diffusion, because particles can not move through each other.

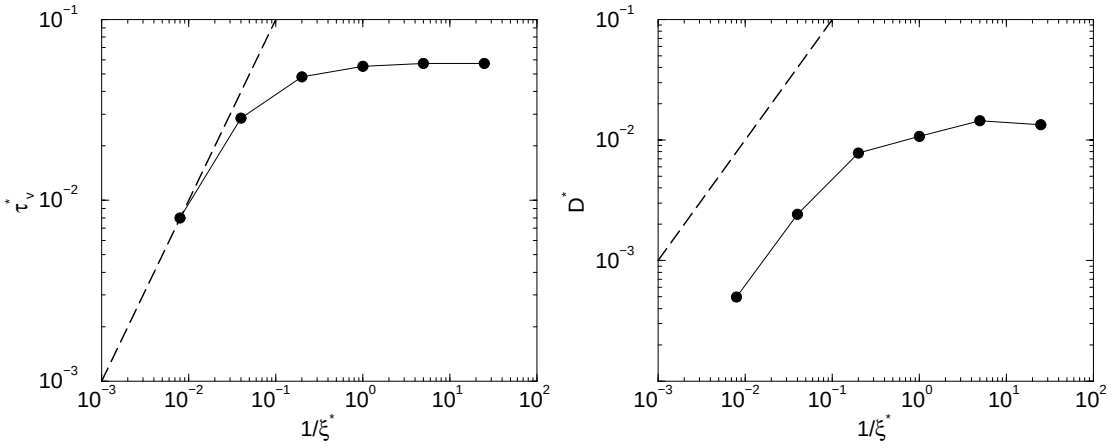

Figure 2.2: Correlation times in LD simulations of a Lennard-Jones liquid. The time at which the velocity autocorrelation function reaches $1/e$ (left-hand graph) and the diffusion constant (right-hand graph) as a function the inverse friction constant. The dashed lines are the curves for a system of non-interacting particles.

BD and SMC simulations

To compare Brownian dynamics and smart Monte Carlo we first tried the most simple system: a harmonic oscillator. Since in BD time scales with the friction coefficient, all times are reported in units of γ^*. We performed simulations of one million steps for both integrators with time steps $0.02\,\gamma^*$, $0.1\,\gamma^*$, $0.5\,\gamma^*$. and γ^*. The acceptance ratios and potential energies are shown in Table 2.1. For a single harmonic oscillator SMC is a clear improvement over BD, as for a time step of $0.5\,\gamma^*$ BD samples too many high energy states. The acceptance ratio gives a good indication of the quality of the sampling in BD.

We also performed BD and SMC simulations of the LJ system. The simulations were of length γ^* with six different time steps, ranging from $10^{-5}\,\gamma^*$ to $38 \cdot 10^{-5}\,\gamma^*$. BD simulations with time step $40 \cdot 10^{-5}\,\gamma^*$ crash and SMC simulations with time step $35 \cdot 10^{-5}\,\gamma^*$ tend to get stuck in a conformation where all steps are rejected. The acceptance ratios and potential energies are shown in Table 2.1. When the time step for both integrator is increased from $5 \cdot 10^{-5}\,\gamma^*$ to $25 \cdot 10^{-5}\,\gamma^*$ the potential energy increases slightly. The cause off this can be seen in the radial distribution functions (rdf's) (see Figure 2.3). With larger time steps the foot of the peak becomes a bit wider and the left flank of the peak shifts in slightly. This is because the second derivative of the potential is very non-linear at the steep wall of the LJ potential. For larger time steps some particles will move too far into the extremely steep wall of the LJ potential, which leads to a large decrease in acceptance, as can be seen in Table 2.1. Under these conditions even SMC does not produce an exact canonical ensemble. This is because the LJ forces are highly

system	time step $\Delta t^*/\gamma^*$	Brownian dynamics "acceptance ratio"	Brownian dynamics V/N [k_BT]	Smart Monte Carlo acceptance ratio	Smart Monte Carlo V/N [k_BT]
harmonic oscillator	0.02	0.999	0.51	0.999	0.50
	0.1	0.992	0.52	0.992	0.50
	0.5	0.925	0.68	0.920	0.51
	1	0.817	1.00	0.780	0.52
LJ liquid	$1 \cdot 10^{-5}$	0.979	-6.35	0.980	-6.35
	$5 \cdot 10^{-5}$	0.815	-6.34	0.823	-6.34
	$25 \cdot 10^{-5}$	0.298	-6.29	0.576	-6.27
	$30 \cdot 10^{-5}$	0.206	-6.26	0.549	-6.27
	$33 \cdot 10^{-5}$	0.139	-6.24	0.535	-6.26
	$38 \cdot 10^{-5}$	0.064	-6.24	-	-

Table 2.1: The acceptance ratios and potential energies for BD and SMC simulations of a harmonic oscillator and a LJ liquid. All acceptance ratios were calculated using the SMC criterion (2.62). Note that the real acceptance ratio for BD is 1.

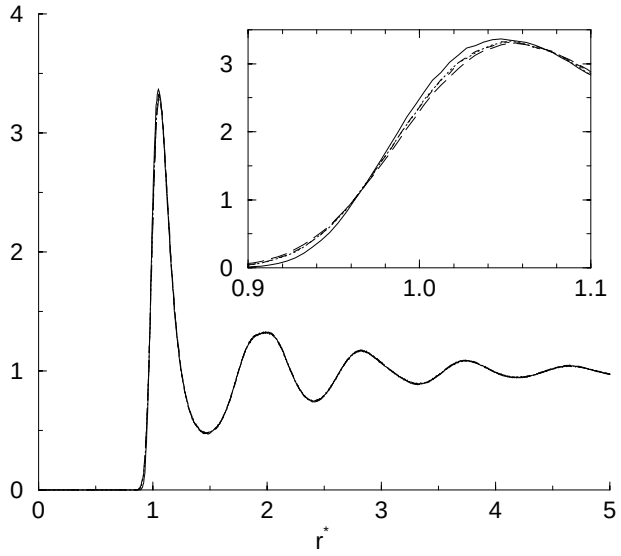

Figure 2.3: Radial distribution functions for a LJ liquid simulated with SMC and BD. The inset shows a magnification of the first peak. The lines in the peak are from top to bottom: SMC with time step $10^{-5}\,\gamma^*$ (solid line), SMC and BD with time step $30 \cdot 10^{-5}\,\gamma^*$ (dotted line and short dashed lines, these two are almost identical) and BD with time step $38 \cdot 10^{-5}\,\gamma^*$ (long dashed line).

non-linear at the steep wall, while acceptance criterion (2.62) uses a linear interpolation of the forces. A time step of $5 \cdot 10^{-5}\,\gamma^*$ is required to produce an accurate rdf. If one does not need very high accuracy, BD with the largest possible time step, $38 \cdot 10^{-5}\,\gamma^*$, will do fine. The acceptance ratio's for the SMC and BD simulations give no indication of the accuracy in this system, except for the details of the first peak in the rdf. But usually such accuracy is not required as the errors in most molecular force fields are much larger. SMC and BD produce exactly identical rdf's for equal time steps. This makes BD the preferable simulation method, as it retains some time information and probably also samples slightly faster, as no steps are rejected.

2.9 Simulations of SPC water

As a model system for more complex liquids we used the SPC water model [25]. This model consists of 3 masses in a fixed geometry, so there are 3 translational and 3 rotational degrees of freedom. The oxygen has a charge -0.82 (e) and Lennard-Jones interactions, the hydrogens have charge 0.41 (e) and no Lennard-Jones interactions. The simulation system consisted of 1728 water molecules in a cubic box of size 3.75 (nm), which gives a density of 980 (kg m^{-3}). We used the particle-mesh Ewald method [4] for the electrostatics, since a cut-off method would result in heating of the system. We performed 5 Langevin dynamics simulations, with friction constant ranging from 0.4 to 250 (ps^{-1}), 3 molecular dynamics simulation using the Berendsen thermostat, with coupling times of 2.5, 0.5 and 0.1 (ps). These simulations had a time step of 2 fs and a length of 80 ps. We also performed 2 Brownian dynamics simulations, one with $\xi_O = \xi_H$ and one with $\gamma_O = \gamma_H$ or $\xi_O = \xi_H/16$, both with a time step of $0.25/\xi_H$ and of length $1.25 \cdot 10^5/\xi_H$. All simulations were performed at a temperature of 300 (K).

The fast time-scale properties are shown in Figure 2.4. These are the decay time of the velocity autocorrelation of the centers of mass of the water molecules and the decay times (τ_r) of the 3 rotational degrees of freedom. The rotational decay times were determined by integrating the autocorrelation functions with an exponential fit of the tails. The diffusion constant is shown in Figure 2.5. In the high friction limit the interactions decrease the diffusion by a factor of 5. To examine the relation between short and long time scale behavior we plotted the diffusion constant against the dipole rotation correlation time Figure 2.6. This is also done for the Brownian dynamics, which results in straight lines, since time scales with γ, which linearly scales both the rotation correlation times and the diffusion constant. It can be seen that BD with constant γ is closer to the real behavior (LD with small ξ) of SPC water than BD with constant ξ. But for both BD simulations $D\tau_r$ deviates significantly from LD with low friction.

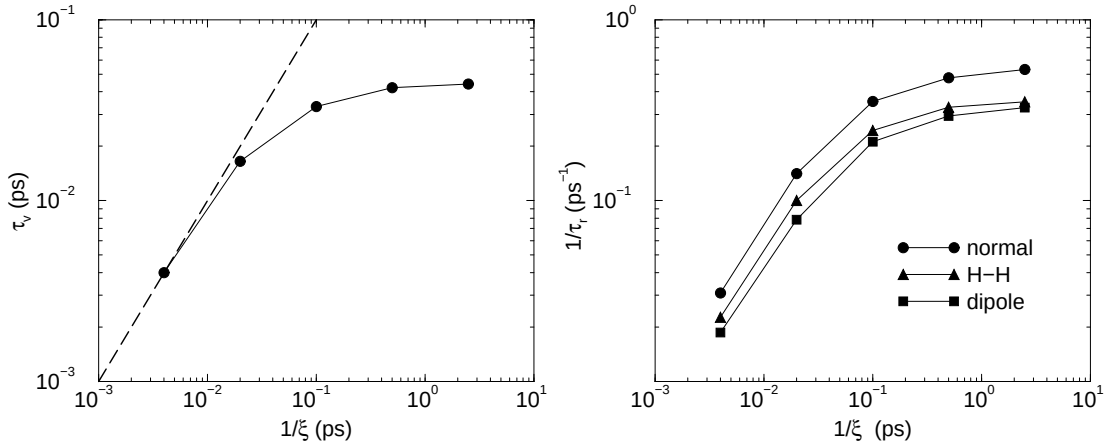

Figure 2.4: Correlation times in LD simulations of SPC water. The time at which the velocity autocorrelation function of the center of mass of the water molecule reaches $1/e$ (left-hand graph), the dashed line is the high friction limit $\tau_v = 1/\xi$. The decay time of the correlation function of the 3 rotational degrees of freedom (right-hand graph). The decay time is the integral over the rotational correlation function.

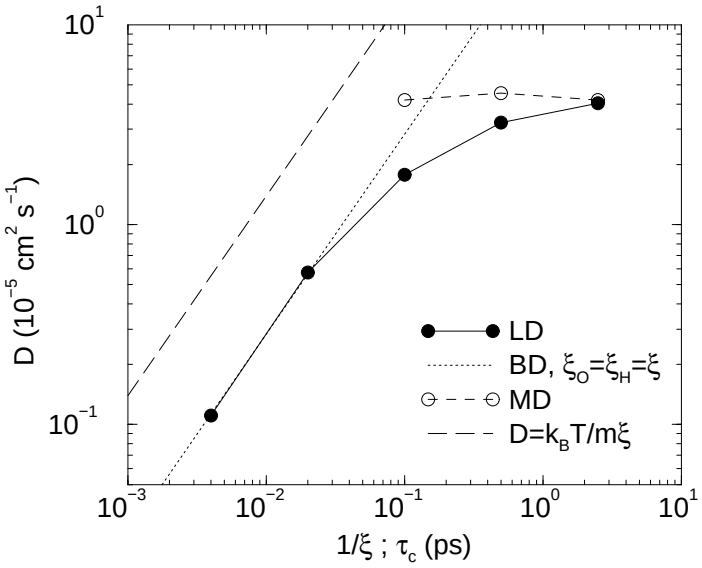

Figure 2.5: The diffusion constant of SPC water plotted as a function of the inverse friction constant for LD and BD and as a function of the coupling time of the Berendsen thermostat for MD. The straight line is the diffusion constant for non-interacting water molecules.

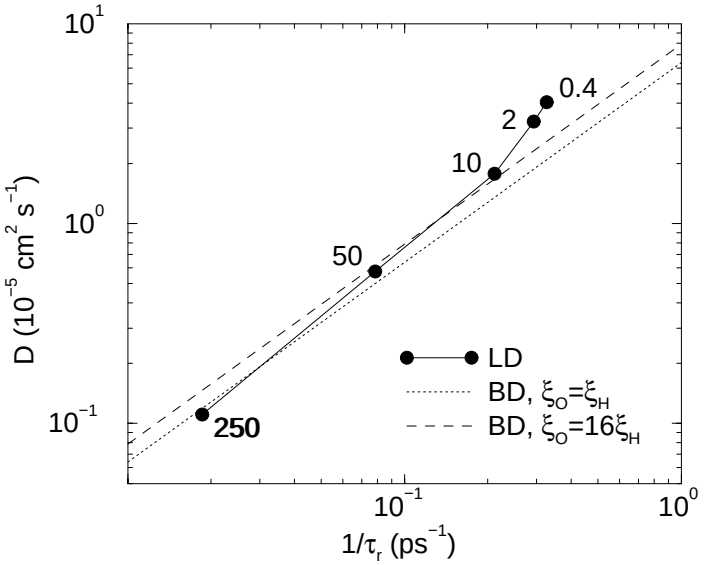

Figure 2.6: The diffusion constant of SPC water plotted as a function of the rotation correlation time of the water dipole. The dots are the Langevin dynamics simulation for the 5 ξ's between 0.4 to 250 (ps^{-1}). The 2 dashed lines are 2 Brownian dynamics simulations with different ratios of friction coefficients between the oxygens and the hydrogens.

2.10 Temperature jumps

To keep the temperature constant during an MD simulation one can use a Berendsen or Nosé-Hoover thermostat. As mentioned before the Berendsen thermostat has the disadvantage that it is unknown if an ensemble is generated. Also the equipartition is not reliable. Uncoupled or weakly coupled degrees of freedom tend to accumulate kinetic energy, for instance the center of mass of the system which leads to "flying ice-cubes". Another example is a periodic system of a water and alkane slab. The interaction between water and alkane is weak and the two slabs will start moving with respect to each other. For these examples the problem can be fixed by removing the center of mass motion of the groups. The Nosé-Hoover thermostat thermostat does not have this problem, but it can cause large fluctuations in the system Langevin dynamics generates a canonical ensemble, but suppresses the velocity correlations at times on the order of $1/\xi$.

To compare the three different temperature coupling methods we applied a temperature jump to the Lennard-Jones system described above. The system was made 8 times larger by stacking $2 \times 2 \times 2$ cubes. After equilibration with LD at $T = \epsilon/k_B$, we increased the temperature of the bath by a factor 1.2 and performed simulations of 5 ps with 3 different thermostats. One with LD with $\xi = 1$, one with MD and Berendsen

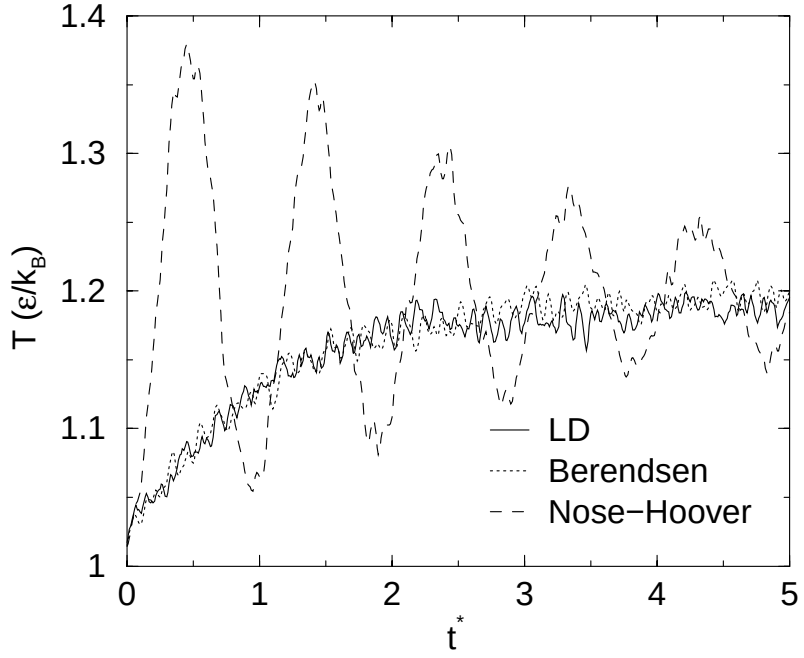

Figure 2.7: Response of three different thermostats on a temperature jump from 1 to 1.2 for a system of 8000 Lennard-Jones particles.

coupling with $\tau_c = 0.5$ and one with MD and a Nosé-Hoover thermostat with a period of oscillations of 1. The behavior of the temperature is shown in Figure 2.7. Figure 2.1. The temperature in the LD simulation and the MD simulation with Berendsen coupling converges exponentially with a time constant of 1. The temperature converges twice as slow as the actual coupling time in the differential equation for the temperature, because the energy fed into the system by the thermostat is distributed almost equally between kinetic and potential energy. This only holds when the coupling time is longer than the correlation times in the system. The Nosé-Hoover thermostat is not appropriate for enforcing a temperature jump. It causes oscillations with an initial amplitude of the size of the temperature difference. The amplitude decays slower than the period of the oscillations.

Note that when a system consisting of a large number of particles is in equilibrium and the integration method is accurate, the temperature will be practically constant and thus the Berendsen and Nosé-Hoover thermostats will have little influence on the system. In contrast, with LD all correlations on times scales longer than $1/\xi$ are suppressed. For instance LD of SPC water with $1/\xi = 0.1$ (ps) decreases the diffusion constant by a factor of 2 (see Figure 2.5).