Time Stepping Algorithms for Classical Molecular Dynamics

Colin John Cotter* and Sebastian Reich†

Department of Mathematics, Imperial College

180 Queen's Gate, London, SW7 2AZ

(Dated: March 27, 2004)

*Electronic address: colin.cotter@ic.ac.uk

†Electronic address: s.reich@ic.ac.uk

Contents

x1 Introduction

This survey is intended to provide an introduction to the special time integration techniques which are used in *classical molecular dynamics* (MD). Conservation of important physical quantities such as energy and momentum are crucial for long time simulations; these conservation laws

are represented by a rich symmetry structure in the numerical algorithms which is both mathematically elegant and physically intuitive.

The survey is primarily devoted to time stepping algorithms for systems of N particles which interact through both long and short range forces through Newton's 2nd law:

$$\dot{\mathbf{r}}_i = \mathbf{p}_i/m_i, \tag{1}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i^{\mathrm{c}} + \mathbf{F}_i^{\mathrm{s}}, \qquad i = 1,\ldots,N \tag{2}$$

where m_i is the mass of particle i with position vector $\mathbf{r}_i = (x_i, y_i, z_i)^T \in \mathbb{R}^3$ and momentum $\mathbf{p}_i = m_i\dot{\mathbf{r}}_i \in \mathbb{R}^3$. It is also assumed that the force acting on the i^{th} particle can be split into a conservative contribution $\mathbf{F}_i^{\mathrm{c}}$ and a non-conservative contribution $\mathbf{F}_i^{\mathrm{s}}$.

The survey begins by recalling a number of specific deterministic and stochastic MD formulations that fit into the general form (1)-(2). The rest of the survey is devoted to numerical algorithms for these formulations. Special focus will be put on methods that can be used for long time simulations. For a more complete perspective on molecular simulations, the reader is referred to a text on the subject such that of SCHLICK [1], ALLEN & TILDESLEY [2], RAPPAPORT [3], or FRENKEL & SMIT [4]. Finally an application of the particle algorithms to continuum modelling is discussed.

x1.1 Constant Energy Molecular Dynamics

For a particle system with forces $\mathbf{F}_i^{\mathrm{s}} = \mathbf{0}$ and

$$\mathbf{F}_i^{\mathrm{c}} = -\nabla_{\mathbf{r}_i} V(\mathbf{r}_1,\ldots,\mathbf{r}_N)$$

(where $V(\mathbf{r}_1,\ldots,\mathbf{r}_N)$ is some potential energy function) the equations of motion (1)-(2) conserve total energy

$$E = \frac{1}{2}\sum_{i=1}^{N} \|\mathbf{p}_i\|^2/m_i + V(\mathbf{r}_1,\ldots,\mathbf{r}_N). \tag{3}$$

Here $\nabla_{\mathbf{r}_i} V(\mathbf{r}) \in \mathbb{R}^3$ represents the column vector of partial derivatives given by

$$\nabla_{\mathbf{r}_i} V(\mathbf{r}) = (\partial_{x_i} V(\mathbf{r}), \partial_{y_i} V(\mathbf{r}), \partial_{z_i} V(\mathbf{r}))^T.$$

Hence *conservative MD* (*i.e.* $\mathbf{F}_i^{\mathrm{s}} = \mathbf{0}$) allows one to model a molecular system under the assumption of constant particle number N, volume V, and energy E. The resulting ensemble is commonly termed the *microcanonical* or constant NVE ensemble.

The mathematical structure of this conservative case has important implications for the design of practical numerical methods for MD. It is easily verified that equations (1)-(2) with $\mathbf{F}_i^s = \mathbf{0}$ are equivalent to the more abstract formulation

$$\dot{\mathbf{r}}_i = +\nabla_{\mathbf{p}_i}\mathcal{H}, \qquad \dot{\mathbf{p}}_i = -\nabla_{\mathbf{r}_i}\mathcal{H},$$

$i = 1,\ldots,N$, where the *Hamiltonian function* $\mathcal{H}$ is identical to the total energy (3)

$$\mathcal{H} = \frac{1}{2}\sum_{i=1}^{N} \|\mathbf{p}_i\|^2/m_i + V(\mathbf{r}_1,\ldots,\mathbf{r}_N). \tag{4}$$

Hence classical molecular dynamics falls into the category of *Hamiltonian mechanics*. This observation has important implications for the numerical treatment relating to the *symplectic structure* of Hamiltonian mechanics [5, 6]. To see this symplectic structure, take a solution $(\mathbf{r}_i(t), \mathbf{p}_i(t))$, $i = 1,\ldots,N$, of the equations (1)-(2) and linearize (1)-(2) along that solution to obtain the following time-dependent linear equations

$$\dot{\mathbf{R}}_i = \mathbf{P}_i/m_i, \tag{5}$$

$$\dot{\mathbf{P}}_i = -\sum_{j=1}^{N} \mathbf{A}_{ij}(t)\,\mathbf{R}_j \tag{6}$$

in the variables $\mathbf{R}_i \in \mathbb{R}^3$, $\mathbf{P}_i \in \mathbb{R}^3$, where

$$\mathbf{A}_{ij}(t) = D_{\mathbf{r}_i\mathbf{r}_j} V(\mathbf{r}_1(t),\ldots,\mathbf{r}_n(t))$$

is a symmetric, time-dependent 3×3 matrix of second order partial derivatives of V with respect to $\mathbf{r}_i$ and $\mathbf{r}_j$. Let $(\mathbf{R}_i^{(1)}(t), \mathbf{P}_i^{(1)}(t))$ and $(\mathbf{R}_i^{(2)}(t), \mathbf{P}_i^{(2)}(t))$ denote any two solutions of (5)-(6). Then it is easy to check that

$$\frac{d}{dt}\left[\sum_{i=1}^{N} \mathbf{R}_i^{(1)}(t)\cdot\mathbf{P}_i^{(2)}(t) - \mathbf{R}_i^{(2)}(t)\cdot\mathbf{P}_i^{(1)}(t)\right] = 0.$$

The expression in the bracket is linear in each of the arguments and defines a two-form Ω called the *symplectic form*. Hence the more abstract property of *conservation of symplecticity* $\dot{\Omega} = 0$ is derived. An immediate consequence of conservation of symplecticity is conservation of volume in phase space. See Appendix A for a precise definition and more details.

Additional conservation properties may apply. For example, consider the force $\mathbf{F}_{ij}^c$ acting of particle i due to particle j, where

$$\mathbf{F}_i^c = \sum_{j=1}^{N} \mathbf{F}_{ij}^c.$$

A wide range of the potential energy functions commonly used in MD lead to particle interactions that satisfy Newton's third law:

$$\mathbf{F}_{ij}^{\mathrm{c}} = -\mathbf{F}_{ji}^{\mathrm{c}}$$

and the force $\mathbf{F}_{ij}$ acts into the direction of $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. The immediate consequence is that such systems conserve total linear and angular momentum:

$$\mathbf{P} = \sum_{i=1}^{N} \mathbf{p}_i, \qquad \mathbf{L} = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{p}_i.$$

x1.2 Stochastic Molecular Dynamics

Often one wishes to perform a simulation at constant temperature rather than constant NVE. The associated ensemble is then called the constant NVT or *macrocanonical ensemble*. There are several techniques available to perform a constant NVT ensemble simulation. In this survey we focus on stochastic methods and consider the extended dynamics models of NOSÉ [7], HOOVER [8], and BOND, LAIRD & LEIMKUHLER [9] only briefly in §4. Throughout the text, we will also point out a few relevant details concerning *hybrid Monte-Carlo* (HMC) simulations [10].

The most widely used stochastic MD model fitting into the framework (1)-(2) is provided by the *Langevin dynamics* (LD) equations

$$\dot{\mathbf{r}}_i = \mathbf{p}_i/m_i, \tag{7}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i^{\mathrm{c}} - \gamma\dot{\mathbf{r}}_i + \sigma\dot{\mathbf{W}}_i(t), \tag{8}$$

where $\mathbf{W}_i(t) \in \mathbb{R}^3$ is a vector of independent Wiener (Brownian) processes [11, 12], γ is a friction coefficient, $\sigma = \sqrt{2\gamma k_B T}$, k_B is Boltzmann's constant, and T is the temperature. It is useful to define a velocity reaction time $\tau_i^* = m_i/\gamma$. τ_i^* is the time it takes the i^{th} particle velocity to be reduced by a factor of $1/e$ in the absence of any forcing.

Another stochastic model is provided by conventional Brownian dynamics (see ALLEN & TILDESLEY [2] and SCHLICK [1]). Algorithms for Brownian dynamics are, for example, discussed in [1, 2, 13] and will not be covered by this survey.

Mathematically, one has to be somewhat careful with the interpretation of (7)-(8) as a *stochastic differential equation* (SDE). In particular, the time derivative $\dot{\mathbf{W}}(t)$ of a Wiener process is not rigorously defined and equation (8) should be replaced by the formulation

$$d\mathbf{p}_i = \mathbf{F}_i^{\mathrm{c}}\,dt - \gamma\dot{\mathbf{r}}_i\,dt + \sigma\,d\mathbf{W}_i(t).$$

See the text [11, 12] and Appendix C for more details. Throughout this survey we will work for simplicity with the less stringent formulation (8).

The Langevin model takes the form of equations (1)-(2) with

$$\mathbf{F}_i^{\mathrm{s}} = -\gamma\dot{\mathbf{r}}_i + \sigma\dot{\mathbf{W}}_i(t).$$

Under fairly general assumptions one can show that the solutions of (7)-(8) indeed sample from a constant NVT ensemble.

There are two important generalizations of (7)-(8) available in the literature. The first one is the *generalized Langevin dynamics* (GLD) model:

$$\dot{\mathbf{r}}_i(t) \;=\; \mathbf{p}_i(t)/m_i, \tag{9}$$

$$\dot{\mathbf{p}}_i(t) \;=\; \mathbf{F}_i^{\mathrm{c}}(t) - \int_0^t \mathcal{K}(t-s)\,\dot{\mathbf{r}}_i(s)\,ds + \mathbf{U}_i(t), \tag{10}$$

where $\mathcal{K}(\tau)$ is a memory kernel and $\mathbf{U}_i(t) = (U_{i,1}, U_{i,2}, U_{i,3})^T \in \mathbb{R}^3$ is a vector of independent and stationary zero-mean Gaussian processes. See the text by ALLEN & TILDESLEY [2] and the original papers by MORI [14, 15] for more details. The auto-covariance function has to satisfy the fluctuation-dissipation relation

$$\mathbb{E}[U_{i,k}(t)U_{i,k}(s)] = k_B T\,\mathcal{K}(t-s),$$

for all $k = 1,2,3$, $i = 1,\ldots,N$. Here $\mathbb{E}[z]$ denotes the expectation value of a stochastic variable z.

The memory term in (10) consists of an integral over the history of the particle paths and makes the generalized Langevin equations very adaptable to various model situations. However it is also difficult to implement numerically. It is therefore convenient to approximate the generalized Langevin equations by higher-dimensional SDEs as suggested by MORI [15].

In this survey we only consider the special but very relevant case of an exponentially decaying kernel

$$\mathcal{K}(\tau) = \lambda e^{-\alpha|\tau|}, \tag{11}$$

where α is the decay rate and λ is a scaling (coupling) parameter. The associated SDE will be stated below in a slightly more general setting. See also ERMAK & BUCKHOLTZ [16].

The second generalization of (7)-(8) allows for a more specific coupling of the noise terms to the conservative part of the model and is called *dissipative particle dynamics* (DPD). The basic

equations due to HOOGERBRUGGE & KOELMAN [17] are

$$\dot{\mathbf{r}}_i = \mathbf{p}_i/m_i, \tag{12}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i^{\mathrm{c}} - \gamma \sum_{j\neq i} \omega(r_{ij})^2 (\mathbf{e}_{ij}\cdot\dot{\mathbf{r}}_{ij})\mathbf{e}_{ij} + \sigma \sum_{j\neq i} \omega(r_{ij})\mathbf{e}_{ij}\dot{W}_{ij}, \tag{13}$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, $\mathbf{e}_{ij} = \mathbf{r}_{ij}/r_{ij}$, $\dot{\mathbf{r}}_{ij} = \dot{\mathbf{r}}_i - \dot{\mathbf{r}}_j$. There is a freedom in the choice of the dimensionless weight function $\omega(r)$ which depends on the particular model in question. However, as pointed out by ESPAÑOL & WARREN [18], to reproduce the constant temperature macro-canonical ensemble the friction coefficient γ and the noise amplitude σ have to satisfy the fluctuation-dissipation relation $\sigma = \sqrt{2k_BT\gamma}$ as before. Finally, $W_{ij}(t) = W_{ji}(t)$ are independent Wiener processes.

An important aspect of DPD is that the non-conservative forcing term

$$\mathbf{F}_i^{\mathrm{s}} = \sum_{j=1}^{N} \mathbf{F}_{ij}^{\mathrm{s}}, \qquad \mathbf{F}_{ij}^{\mathrm{s}} = -\gamma\omega(r_{ij})^2(\mathbf{e}_{ij}\cdot\dot{\mathbf{r}}_{ij})\mathbf{e}_{ij} + \sigma\sum_{j\neq i}\omega(r_{ij})\mathbf{e}_{ij}\dot{W}_{ij},$$

satisfies $\mathbf{F}_{ij}^{\mathrm{s}} = -\mathbf{F}_{ji}^{\mathrm{s}}$ and the force acts in the direction of $\mathbf{r}_{ij}$. Hence, Newton's third law applies and total angular and linear momenta are conserved. Note, however, that there is no longer a conserved energy.

For later reference, write the equations (12)-(13) in a generalized form [19]:

$$\dot{\mathbf{r}} = \mathbf{M}^{-1}\mathbf{p}, \tag{14}$$

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{r}} V(\mathbf{r}) - \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}) \left[\gamma \dot{h}_k(\mathbf{r}) + \sigma \dot{W}_k\right], \tag{15}$$

where $\mathbf{r}$ is the collection of the N particle position vectors $\mathbf{r}_i$, $\mathbf{p} = \mathbf{M}\dot{\mathbf{r}}$ is the associated momentum vector, $\mathbf{M}$ is the diagonal mass matrix, $V(\mathbf{r})$ is the potential energy,

$$\dot{h}_k(\mathbf{r}) = \nabla_{\mathbf{r}} h_k(\mathbf{r})\cdot\dot{\mathbf{r}},$$

and the functions $h_k(\mathbf{r})$, $k = 1,\ldots,K$, can be chosen quite arbitrarily. The particular choice

$$h_k(\mathbf{r}) = \phi(r_{ij}), \qquad \omega(r) = \phi'(r),$$

$k = 1,\ldots,(N-1)N/2$, in (15) leads back to the standard DPD model (12)-(14). However, one can also set $K = 3N$ and

$$h_i(\mathbf{r}) = x_i, \qquad h_{i+N}(\mathbf{r}) = y_i, \qquad h_{i+2N}(\mathbf{r}) = z_i,$$

$i = 1, \ldots, N$, in (15), which leads to the standard Langevin model (7)-(8).

As suggested by COTTER & REICH [19], one can combine the idea of DPD and that of the generalized Langevin model (9)-(10). Under the assumption of an exponentially decaying memory kernel (11), the following equivalent SDE can be derived:

$$\dot{\mathbf{r}} = \mathbf{M}^{-1}\mathbf{p}, \tag{16}$$

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{r}} V(\mathbf{r}) - \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}) s_k, \tag{17}$$

$$\dot{s}_k = -\alpha s_k + \lambda \dot{h}_k(\mathbf{r}) + \sqrt{2k_B T \lambda \alpha}\, \dot{W}_k, \qquad k = 1, \ldots, K, \tag{18}$$

where $W_k(t)$ is a Wiener process for each k. We call (16)-(17) the *extended DPD* (EDPD) model [19]. Note that the variable s_k constitutes an Ornstein-Uhlenbeck process for $\dot{h}_k = 0$ [11].

The standard DPD model (14)-(15) can be recovered in the limit of $\alpha \gg 1$ subject to $\lambda/\alpha = const.$, in which case equation (18) reduces to

$$0 \approx -s_k + \frac{\lambda}{\alpha}\dot{h}_k(\mathbf{r}) + \sqrt{\frac{2k_B T \lambda}{\alpha}}\dot{W}_k$$

and we identify $\gamma = \lambda/\alpha$ and $\sigma = \sqrt{2k_B T \lambda/\alpha}$. We shall return to this limiting behaviour in the section on numerical methods.

The appropriate choice of a stochastic model should be determined by the physical system which it is to model. Most often the fluctuation-dissipation terms are introduced to represent unresolved degrees of freedom (see, for example, [14, 15, 20–22]). While it is desirable to derive such a stochastic parametrization on strict mathematical grounds, more often heuristic arguments need to be applied due to the complexity of the underlying physical processes. An example will be briefly discussed in §5.

x2 Numerical Methods for Constant Energy Molecular Dynamics

x2.1 The Störmer-Verlet Algorithm

The most widely used numerical method for MD is the *Störmer-Verlet* method, which is written here in the velocity/momentum formulation:

$$\mathbf{p}^{n+1/2} = \mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n), \tag{19}$$

$$\mathbf{r}^{n+1} = \mathbf{r}^n + \Delta t \mathbf{M}^{-1}\mathbf{p}^{n+1/2}, \tag{20}$$

$$\mathbf{p}^{n+1} = \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}), \tag{21}$$

where Δt is the step size. The method was first used in the context of MD by VERLET [23] and has been very popular with the MD community since. However there is indication that the method was already known the Newton and used in his *Principia* from 1687 to prove Kepler's second law (see [24]).

There exist a number of essentially equivalent formulations of the Störmer-Verlet method. See, for example, SCHLICK [1]. Introduce the abbreviation *kick* for the velocity updates (19) and (21), and the abbreviation *drift* for the position update (20). Hence the velocity Störmer-Verlet method may be characterized as *kick-drift-kick*.

Why is the Störmer-Verlet method so successful for constant energy MD simulations? Several reasons can be given. The method is easy to implement, it exactly conserves total linear and angular momentum, it is time-reversible, and the total energy (3) is very well conserved over long simulation times even for large and complex systems.

In fact, taking a slightly different perspective, these conservation properties become less of a surprise. Recall that the conservative MD equations can be derived from a Hamiltonian function (4). Write the Hamiltonian function $\mathcal{H}$ as the sum of three parts

$$\mathcal{H}_1 = \frac{1}{2}V(\mathbf{r}), \qquad \mathcal{H}_2 = \frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p}, \qquad \mathcal{H}_3 = \frac{1}{2}V(\mathbf{r}).$$

Each of these terms gives rise to Hamiltonian equations of motion. For example, take the Hamiltonian $\mathcal{H}_1$. The associated equations are

$$\dot{\mathbf{p}} = -\frac{1}{2}\nabla_{\mathbf{r}}V(\mathbf{r}), \qquad \dot{\mathbf{r}} = \mathbf{0}.$$

These equations can be solved analytically, producing the momentum update (*kick*):

$$\mathbf{p}(t) = \mathbf{p}(0) - \frac{t}{2}\nabla_{\mathbf{r}}V(\mathbf{r}), \qquad \mathbf{r}(t) = \mathbf{r}(0).$$

Denote the associated map from time $t = 0$ to $t = \tau$ by $\Psi_{\tau,\mathcal{H}_1} : (\mathbf{r}(0),\mathbf{p}(0)) \to (\mathbf{r}(\tau),\mathbf{p}(\tau))$. Of course $\Psi_{\tau,\mathcal{H}_3} = \Psi_{\tau,\mathcal{H}_1}$ and it remains to find $\Psi_{\tau,\mathcal{H}_2}$. A simple calculation yields the position update (*drift*):

$$\mathbf{p}(\tau) = \mathbf{p}(0), \qquad \mathbf{r}(\tau) = \mathbf{r}(0) + \tau\mathbf{M}^{-1}\mathbf{p}(0).$$

Not much has been gained yet. However, if one applies the map $\Psi_{\tau,\mathcal{H}_1}$ with $\tau = \Delta t$ to the numerical approximation $(\mathbf{r}^n,\mathbf{p}^n)$ and calls the result $(\mathbf{r}^n,\mathbf{p}^{n+1/2})$, and then applies $\Psi_{\Delta t,\mathcal{H}_2}$ to that approximation followed by an application of $\Psi_{\Delta t,\mathcal{H}_3}$, then the result $(\mathbf{r}^{n+1},\mathbf{p}^{n+1})$ is exactly equivalent to the outcome from the Störmer-Verlet method (19)-(21). Hence the Störmer-Verlet method

may be written as a map $\Phi_{\Delta t} : (\mathbf{r}^n, \mathbf{p}^n) \to (\mathbf{r}^{n+1}, \mathbf{p}^{n+1})$ which itself is a concatenation of three maps:

$$\Phi_{\Delta t} = \Psi_{\Delta t, \mathcal{H}_3} \circ \Psi_{\Delta t, \mathcal{H}_2} \circ \Psi_{\Delta t, \mathcal{H}_1}.$$

In other words, each of the three maps corresponds to exactly one step in the *kick-drift-kick* sequence of the Störmer-Verlet method (19)-(21). A number of very useful conclusions can be drawn from this abstract result.

(i) Each map $\Psi_{\Delta t, \mathcal{H}_i}$ conserves total linear and angular momentum. Hence the composition of these maps (*i.e.* the Störmer-Verlet method) conserves total linear and angular momentum.

(ii) Each map $\Psi_{\Delta t, \mathcal{H}_i}$ is the exact solution to a Hamiltonian differential equation and so it conserves the symplectic two-form Ω introduced in §1.1. Hence the Störmer-Verlet method preserves the symplectic two-form Ω from time step to time step, *i.e.* $\Omega^{n+1} = \Omega^n$. A numerical integrator with this property is called a *symplectic method* [6, 25, 26].

(iii) Time-reversibility of the Störmer-Verlet method follows from the obvious symmetry property

$$\Psi_{\Delta t, \mathcal{H}_3} \circ \Psi_{\Delta t, \mathcal{H}_2} \circ \Psi_{\Delta t, \mathcal{H}_1} = \Psi_{\Delta t, \mathcal{H}_1} \circ \Psi_{\Delta t, \mathcal{H}_2} \circ \Psi_{\Delta t, \mathcal{H}_3}$$

in the composition of the three flow maps. This symmetry also implies that the Störmer-Verlet method is second order.

We next describe a truly remarkable observation for the Störmer-Verlet method which relies on its conservation of the symplectic two-form Ω. Namely, one can find a time-dependent Hamiltonian function $\tilde{\mathcal{H}}(\mathbf{r}, \mathbf{p}, 2\pi t/\Delta t)$ which is 2π-periodic in its third argument such that the solution of

$$\begin{aligned} \dot{\mathbf{r}} &= +\nabla_{\mathbf{p}} \tilde{\mathcal{H}}(\mathbf{r}, \mathbf{p}, 2\pi t/\Delta t), \\ \dot{\mathbf{p}} &= -\nabla_{\mathbf{r}} \tilde{\mathcal{H}}(\mathbf{r}, \mathbf{p}, 2\pi t/\Delta t) \end{aligned}$$

with initial conditions $\mathbf{r}(0) = \mathbf{r}^n$ and $\mathbf{p}(0) = \mathbf{p}^n$ is exactly equivalent to $(\mathbf{r}^{n+1}, \mathbf{p}^{n+1})$ at $t = \Delta t$. (See the paper by KUKSIN & PÖSCHEL [27] and MOAN [28] for the mathematical details.) Using a more abstract notation, we can state

$$\Phi_{\Delta t} = \Psi_{\Delta t, \tilde{\mathcal{H}}}.$$

This result tells us that the solution behaviour of the Störmer-Verlet method is completely characterized by a *time-dependent* Hamiltonian problem with Hamiltonian function $\tilde{\mathcal{H}}$.

This statement is not entirely satisfactory as it is well known that energy is *not* conserved for time-dependent Hamiltonian problems. However, as first pointed out by NEISHTADT [29], the time-dependence in $\tilde{\mathcal{H}}$ averages itself out up to negligible terms of size $O(e^{-c/\Delta t})$ for sufficiently small step-sizes Δt (here $c > 0$ is a constant which depends on the particular problem). Hence we can claim that the Störmer-Verlet method is the nearly exact solution of a Hamiltonian problem with *time-independent* Hamiltonian $\hat{\mathcal{H}}_{\Delta t}(\mathbf{r},\mathbf{p})$. This time-independent Hamiltonian possesses an asymptotic expansion in the step-size Δt of the form

$$\hat{\mathcal{H}}_{\Delta t} = \mathcal{H} + \Delta t^2 \delta\mathcal{H}_2 + \Delta t^4 \delta\mathcal{H}_4 + \Delta t^6 \mathcal{H}_6 + \ldots,$$

with

$$\mathcal{H}_2 = \frac{1}{12}\mathbf{p}^T\mathbf{M}^{-1}[D_{\mathbf{rr}}V(\mathbf{r})]\mathbf{M}^{-1}\mathbf{p} + \frac{1}{24}[\nabla_{\mathbf{r}}V(\mathbf{r})]^T\mathbf{M}^{-1}\nabla_{\mathbf{r}}V(\mathbf{r}),$$

where $D_{\mathbf{rr}}V(\mathbf{r})$ denotes the Hessian matrix of the potential energy V. Expressions for the higher order correction terms $\delta\mathcal{H}_i$, $i = 4,6,\ldots$, can be found using the Baker-Campbell-Hausdorff (BCH) formula (see [25], [26], or [6]).

A practical algorithm for assessing energy conservation with respect to a modified Hamiltonian has been proposed by SKEEL & HARDY [30]. See also [31].

The fact that the modified energy $\tilde{\mathcal{H}}_{\Delta t}$ is essentially preserved exactly under the Störmer-Verlet method and the fact that

$$|\mathcal{H} - \hat{\mathcal{H}}_{\Delta t}| = O(\Delta t^2)$$

explains the observed approximate conservation of energy under the Störmer-Verlet method. See BENETTIN & GIORGILLI [32], HAIRER & LUBICH [33] and REICH [34] for rigorous results. On a less rigorous mathematical basis one can even claim that the Störmer-Verlet method samples from a constant $NV\hat{E}$ ensemble with modified energy $\hat{E} = \hat{\mathcal{H}}_{\Delta t}$.

Let us summarize this section. The Störmer-Verlet method can be written as the concatenation of exact solution maps to a sequence of Hamiltonian functions $\mathcal{H}_i$, $i = 1,2,3$. This implies that the Störmer-Verlet method is symplectic and hence is equivalent to the exact solution of a time-dependent modified Hamiltonian problem. For sufficiently small step-sizes this time-dependence averages itself out and an excellent long time conservation of energy is thus observed numerically. Time averages computed along numerical trajectories will follow a constant $NV\hat{E}$ ensemble under the assumption of ergodicity (see [34] for a precise result).

This result has also implications for *hybrid Monte-Carlo* (HMC) simulations [10, 35] (see also the text [1]). Detailed balance requires that the underlying MD algorithm is time-reversible and

volume conserving (see MEHLIG, HERMANN & FOREST [36]). These properties are satisfied by the Störmer-Verlet method. Furthermore, if one accepts the point of view that it is sufficient to sample from the macrocanonical ensemble associated with the modified Hamiltonian $\hat{\mathcal{H}}_{\Delta t}$, then the quasi-exact conservation of $\hat{\mathcal{H}}_{\Delta t}$ under the Störmer-Verlet method allows one to accept virtually all candidate moves (see HAMPTON & IZAGUIRRE [37]).

The above arguments apply to all numerical methods that can be derived by a splitting of a Hamiltonian function $\mathcal{H}$ into exactly solvable parts. The most interesting recent application of this approach to MD is provided by *rigid body dynamics* and we refer the reader to the publications TOUMA & WISDOM [38], REICH [39], and DULLWEBER, MCLACHLAN & LEIMKUHLER [40]. See also Appendix B for a summary of the basic ideas. Another application of the splitting idea will be discussed below in the context of *multiple-time-stepping* (MTS) methods, which will also point to some pitfalls of symplectic integration methods.

For a more complete perspective on symplectic integration methods and classical mechanics, the reader is referred to a text on the subject such as that of SANZ-SERNA & CALVO [25], HAIRER, LUBICH & WANNER [26], and LEIMKUHLER & REICH [6].

x2.2 Constraints and the SHAKE/RATTLE Algorithm

Along with the benefits of the Störmer-Verlet method there is one crucial problem with the method. The step-size Δt has to be smaller than the period of the fastest oscillations present in the molecular system [41]. Since these oscillations can be of the order of femtoseconds, this step-size is very short compared to the timescale showing dynamical behaviour of interest such as protein folding and conformational changes. Since bond vibrations are the source of the fastest vibrational modes, it was recognized early on that an explicit removal of these vibrations *via* rigid length constraints

$$\|\mathbf{r}_i - \mathbf{r}_j\| = L \tag{22}$$

would allow for an increase in step-size. Mathematically such an approach leads to constrained Hamiltonian dynamics of the form

$$\dot{\mathbf{r}} = \mathbf{M}^{-1}\mathbf{p}, \tag{23}$$

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{r}} V(\mathbf{r}) - \nabla_{\mathbf{r}} g(\mathbf{r})\,\lambda, \tag{24}$$

$$0 = g(\mathbf{r}), \tag{25}$$

where the function $g(\mathbf{r}) = \|\mathbf{r}_i - \mathbf{r}_j\| - L$ represents the length constraint (22) and where λ is the associated Lagrange multiplier. For simplicity, the discussion here is restricted to a single constraint but it is easy to generalize the results to a set of constraints $\mathbf{g}(\mathbf{q}) = \mathbf{0}$.

Differentiate (25) twice with respect to time to obtain first

$$0 = \frac{d}{dt} g(\mathbf{r}) = \nabla_{\mathbf{r}} g(\mathbf{r}) \cdot \mathbf{M}^{-1} \mathbf{p} \tag{26}$$

and then

$$0 = \frac{d^2}{dt^2} g(\mathbf{r}) = -\nabla_{\mathbf{r}} g(\mathbf{r}) \cdot \mathbf{M}^{-1} \left[\nabla_{\mathbf{q}} V(\mathbf{q}) + \nabla_{\mathbf{r}} g(\mathbf{r}) \lambda\right] + \mathbf{p}^T \mathbf{M}^{-1} D_{\mathbf{rr}} g(\mathbf{r}) \mathbf{M}^{-1} \mathbf{p}.$$

Here $D_{\mathbf{rr}} g(\mathbf{r}) \in \mathbb{R}^{3N \times 3N}$ denotes the Hessian matrix of g. The last equation can be solved for the Lagrange multiplier λ which is then substituted back into equation (25) to yield an explicit ODE system (23)-(24) with (25) as an invariant. The hidden constraint (26) places a restriction on the momenta $\mathbf{p}$.

While differentiation of (25) is a useful analytical approach it is not well suited to numerical integration. Instead a direct discretization of the constrained formulation (23)-(25) is required. The following modification of the Störmer-Verlet method (19)-(21), called RATTLE, has proven successful in practice:

1.

$$\mathbf{p}^{n+1/2} = \mathbf{p}^n - \frac{\Delta t}{2} \nabla_{\mathbf{r}} V(\mathbf{r}^n) - \nabla_{\mathbf{r}} g(\mathbf{r}^n) \lambda^n, \tag{27}$$

$$\mathbf{r}^{n+1} = \mathbf{r}^n + \Delta t \mathbf{M}^{-1} \mathbf{p}^{n+1/2}, \tag{28}$$

$$0 = g(\mathbf{r}^{n+1}). \tag{29}$$

2.

$$\mathbf{p}^{n+1} = \mathbf{p}^{n+1/2} - \frac{\Delta t}{2} \nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}) - \nabla_{\mathbf{r}} g(\mathbf{r}^{n+1}) \mu^{n+1}, \tag{30}$$

$$0 = \nabla_{\mathbf{r}} g(\mathbf{r}^{n+1}) \cdot \mathbf{M}^{-1} \mathbf{p}^{n+1}. \tag{31}$$

RATTLE was introduced by ANDERSEN [42] based on an earlier method called SHAKE [43]. Both methods are essentially equivalent and have been shown to be symplectic by LEIMKUHLER & SKEEL [44]. In fact, RATTLE can be understood as yet another application of the splitting idea described in the previous section [45]. Hence the method also conserves linear and angular momentum and is time-reversible.

Let us briefly discuss the implementation of RATTLE. Step 1 requires the solution of a non-linear equation in λ^n to satisfy the position constraint (29). A fast way to achieve this is to apply Newton's method to

$$g\left(\bar{\mathbf{r}} - \Delta t \mathbf{M}^{-1} \nabla_{\mathbf{r}} g(\mathbf{r}^n)\, \lambda^n\right) = 0, \qquad \bar{\mathbf{r}} = \mathbf{r}^n + \Delta t \mathbf{M}^{-1} \left(\mathbf{p}^n - \frac{\Delta t}{2} \nabla_{\mathbf{r}} V(\mathbf{r}^n)\right).$$

Step 2 requires the computation of another Lagrange multiplier μ^{n+1}, which has to be chosen such that the linear momentum constraint (31) holds. This leads to a linear equation which may be solved explicitly to give

$$\mu^{n+1} = \frac{1}{\nabla_{\mathbf{r}} g(\mathbf{r}^{n+1}) \cdot \mathbf{M}^{-1} \nabla_{\mathbf{r}} g(\mathbf{r}^{n+1})} \nabla_{\mathbf{r}} g(\mathbf{r}^{n+1}) \cdot \mathbf{M}^{-1} \left[\mathbf{p}^{n+1/2} - \frac{\Delta t}{2} \nabla_{\mathbf{r}} V(\mathbf{r}^{n+1})\right].$$

Clearly these two steps easily generalize to larger sets of constraints and we point the reader to the publication [46] for an explicit discussion.

The elimination of fast bond stretching and possibly also bond bending modes certainly allows for larger time steps Δt but it is not without effect on the ensemble behaviour of the molecular system. This has first been pointed out by FIXMAN [47]. There are also other problems which relate to the fact that the molecular system might become too "rigid" [45]. Artifacts which are introduced by constraints have been demonstrated by VAN GUNSTEREN & KARPLUS through simulations of the protein BPTI in vacuum [48]. Nevertheless constraint dynamics formulations have become standard for water molecules and other small, essentially rigid molecular units. In fact, rigid body motion is best described in terms of rotation matrices and efficient numerical implementations can be given using RATTLE. This has been demonstrated by KOL, LAIRD & LEIMKUHLER [49] based on earlier work by MCLACHLAN & SCOVEL [50] and REICH [39].

To be able to treat polarization and other essentially quantum dynamical effects within classical MD, there have been attempts to develop algorithms for flexible or soft constraints. See the work by ZHOU, REICH & BROOKS [51] and HESS, SAINT-MARTIN & BERENDSEN [52]. These methods replace, for example, a rigid length constraint $\|\mathbf{r}_i - \mathbf{r}_j\| = L$ by a variable length constraint $\|\mathbf{r}_i - \mathbf{r}_j\| = L(t)$, where $L(t)$ is determined self-consistently by minimizing the total energy along the bond stretching mode $\mathbf{r}_i - \mathbf{r}_j$. This procedure is more expensive than rigid constraints, but it reintroduces flexibility into the molecular system without having to resolve the fast bond vibrations. In certain cases it provides also a more realistic model for chemical bonds.

A detailed account on constraint and rigid body dynamics and their numerical implementation can be found in the text by LEIMKUHLER & REICH [6].

MTS methods were introduced in 1970s by STREETT, TILDESLEY & SAVILLE [53] and FINNEY [54] in an effort to reduce computation costs. More recently symplectic MTS methods based on the idea of Hamiltonian splitting were suggested by TUCKERMAN, BERNE & MARTYNA [55] (r-RESPA) and GRUBMÜLLER, HELLER, WINDEMUTH & SCHULTEN [56] (Verlet-I).

MTS methods may be used for systems with Hamiltonian functions of the form

$$\mathcal{H} = T(\mathbf{p}) + V_{\text{slow}}(\mathbf{r}) + V_{\text{fast}}(\mathrm{r}),$$

where T represents the kinetic energy of the system (*i.e.* $T = \mathbf{p}^T\mathbf{M}^{-1}\mathbf{p}/2$), V_{slow} represents for slow and long-range interactions such as electrostatics, and V_{fast} collects all the short-range and stiff potential energy contributions such as bond stretching and bending. The aim is to minimize the number of computationally expensive evaluations of the long-range force field while simultaneously maximizing the step size Δt. The definition of V_{slow} and V_{fast} often involves the usage of cut-off functions [1, 2]. To maintain the good energy conservation property of the Störmer-Verlet method these cut-off functions should be sufficiently smooth (see [57] and [6])

The formulation again proceeds by splitting the Hamiltonian into exactly solvable problems. However this time it is also necessary to take into account the relative strength of the forces. For that reason an inner time step $\delta t = \Delta t/M$ is introduced, where $M \gg 1$ is an appropriate integer. Next write

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 + \mathcal{H}_3 + \cdots \mathcal{H}_{M+3}$$

with

$$\begin{aligned}
\mathcal{H}_1 &= \frac{1}{2}V_{\text{slow}}(\mathbf{r}),\\
\mathcal{H}_2 &= \frac{1}{2M}T(\mathbf{p}),\\
\mathcal{H}_{2i+1} &= \frac{1}{M}V_{\text{fast}}(\mathbf{r}), \quad (i = 1,\ldots,M),\\
\mathcal{H}_{2i+2} &= \frac{1}{M}T(\mathbf{p}), \quad (i = 1,\ldots,M-1),\\
\mathcal{H}_{2M+2} &= \frac{1}{2M}T(\mathbf{p}),\\
\mathcal{H}_{2M+3} &= \frac{1}{2}V_{\text{slow}}(\mathbf{r}).
\end{aligned}$$

Each of the Hamiltonian functions $\mathcal{H}_i$, $i = 1,\ldots,M+3$, gives rise to equations of motion that can be trivially solved along the lines discussed for the Störmer-Verlet method. Denote the corresponding

solution maps by $\Psi_{\Delta t,\mathcal{H}_i}$. A symplectic MTS method then can be written as a composition of these elementary maps:

$$\Phi_{\Delta t} = \Psi_{\Delta t,\mathcal{H}_{M+3}} \circ \Psi_{\Delta t,\mathcal{H}_{M+2}} \circ \cdots \circ \Psi_{\Delta t,\mathcal{H}_3} \circ \Psi_{\Delta t,\mathcal{H}_2} \circ \Psi_{\Delta t,\mathcal{H}_1}. \tag{32}$$

A numerical implementation would proceed by writing several subroutines performing updates according to $\Psi_{\Delta t,\mathcal{H}_i}$ which are called sequentially. Note that only three different subroutines are required and that the slow forces are evaluated only once per time step.

A list of applications of r-RESPA to various molecular systems with multiple time scales can be found in a survey by BERNE [58].

The method (32) conserves linear and angular momentum, is symplectic, and, because of symmetry, is also time-reversible. Hence it shares all these desirable properties with the Störmer-Verlet method. One might also expect the method to be able to take much larger outer time steps. However this is *not* in fact the case.

Where do things go wrong? Following the discussion from Section §2.1, there exists a time-dependent Hamiltonian $\tilde{\mathcal{H}}(\mathbf{r},\mathbf{p},2\pi t/\Delta t)$ such that

$$\Phi_{\Delta t} = \Psi_{\Delta t,\tilde{\mathcal{H}}}.$$

Now recall why we were able to remove the time-dependence in $\tilde{\mathcal{H}}$ in case of the Störmer-Verlet method. Since $\tilde{\mathcal{H}}$ is 2π-periodic in its third argument, the period of $\tilde{\mathcal{H}}$ with respect to time t is $T = \Delta t$. For $\Delta t \to 0$ this period becomes much shorter than any natural period of the molecular system and that allows the time-dependence in $\tilde{\mathcal{H}}$ to be transformed away up to terms exponentially small in Δt.

For the MTS method the situation is quite different. The aim is to use the method with an outer step-size Δt that is comparable or even larger than the period of the fastest molecular oscillations. Hence there is a very good chance that $T_{\text{step}} = \Delta t$ will be close to some natural period T_{osc} of the molecular system and that coincidence can give rise to numerically induced resonance instabilities. The situation is very similar to sitting in a swing. What one normally attempts to do there is to adjust one's own movements to the natural period of the swing. If that happens the swing will start swinging higher and higher and in case of an MTS method the numerical trajectory will start gaining energy and will eventually blow up. The danger of resonances for symplectic MTS methods such as Verlet-I/r-RESPA has been discussed in detail by BIESIADECKI & SKEEL [59].

To get a clearer picture of what is going on, consider the following linear problem.

Example. Consider a harmonic oscillator subject to a "perturbation" $\dot{p} = -r$, *e.g.*

$$\dot{r} = \omega p, \qquad \dot{p} = -\omega r - r.$$

The solutions in $\mathbf{z}(t) = (r(t), p(t))^T$ are oscillatory and the eigenvalues of the matrix-valued solution operator

$$\mathbf{S}(t) = e^{t\mathbf{A}}, \qquad \mathbf{A} = \begin{bmatrix} 0 & \omega \\ -\omega - 1 & 0 \end{bmatrix},$$

are on the unit circle for all $\omega \geq 0$, *i.e.*, the eigenvalues have modulus equal to one.

Assume now that $\omega \gg 1$. From a numerical point of view, one could be tempted to split the equations of motion into the highly oscillatory contribution

$$\dot{r} = \omega p, \qquad \dot{p} = -\omega r$$

and the perturbation

$$\dot{r} = 0, \qquad \dot{p} = -r.$$

Denoting the associated matrix-valued solution operators by $\mathbf{S}_1(t)$ and $\mathbf{S}_2(t)$ respectively, a second-order numerical propagator is obtained, for example, *via* the matrix product

$$\mathbf{M}_{\Delta t} = \mathbf{S}_2(\Delta t/2)\,\mathbf{S}_1(\Delta t)\,\mathbf{S}_2(\Delta t/2).$$

The modulus of the eigenvalues of $\mathbf{M}_{\Delta t}$ for $\omega = 50$ was computed as a function of the step-size $\Delta t \leq 0.2$. Regions of instabilities can be clearly seen in Fig. §1 for $\omega\Delta t \approx k\pi$, $k = 1, 2, 3$. A similar behaviour would be observed for the MTS splitting method (32) with $T = \omega p^2/2$, $V_{\text{fast}} = \omega r^2/2$, $V_{\text{slow}} = r^2/2$, and $\omega\,\delta t \ll 1$. □

The rather disappointing outcome from this linear model problem is that the symplectic MTS method (32) is limited to a step-size $\Delta t < \pi/\omega$, ω the fastest frequency of the system. Even worse, nonlinear 3:1 resonances between fast and slow modes potentially limit the step-size further to $\Delta t < 2\pi/(3\omega)$ and for flexible water this limit corresponds to 3.3 fs (see MA, IZAGUIRRE & SKEEL [60]). Compared to the Störmer-Verlet method, which has to satisfy the linear stability bound $\Delta t < 2/\omega$, the small improvement is hardly worth the effort. One should of course try to avoid the resonance zones by a careful choice of the step-size. However, such an approach seems impossible for the almost continuous frequency spectrum of densely packed molecular systems.

To explain the instabilities of the MTS method (32), the method can be usefully paraphrased as *kick-drift/oscillate-kick*. Clearly, if there is no slow force then the method reduces to *drift/oscillate* and as such is stable for δt small enough. On the other hand, if there is no fast force, then the MTS method reduces to the standard Störmer-Verlet sequence of *kick-drift-kick*. So the problem must arise from the interaction of the *kick* with the *oscillate* steps. Indeed the *kick*

$$\mathbf{p}^n + \frac{\Delta t}{2}\mathbf{F}_{\text{slow}}(\mathbf{r}^n) \to \mathbf{p}^n \tag{33}$$

should rather be replaced by an integral of the form

$$\mathbf{p}^n + \frac{1}{2}\int_{-\Delta t/2}^{+\Delta t/2} \mathbf{F}_{\text{slow}}(\mathbf{r}(t_n+\tau))\,d\tau \to \mathbf{p}^n, \tag{34}$$

where $\mathbf{r}(t_n+\tau)$ comes from the *drift/oscillate* step and $\mathbf{r}(t_n) = \mathbf{r}^n$. There are two fundamental obstacles to a numerical implementation of this idea:

1. The integral should be easy to evaluate and, in particular, only one evaluation of the long range force $\mathbf{F}_{\text{slow}}$ should be necessary.

2. The overall method should remain symplectic.

These requirements sound almost impossible to fulfill. But GARCÍA-ARCHILLA, SANZ-SERNA & SKEEL [61] showed that the impossible is possible for molecular systems. The basic idea is relatively simple. First note that the amplitude of the fast oscillations will be very small, allowing the integral in (34) to be replaced by a simple average

$$\overline{\mathbf{r}}^n = \int_{\Delta t/2}^{+\Delta t/2} \mathbf{r}(t_n+\tau)\,d\tau \tag{35}$$

which may be combined with the approximation

$$\int_{-\Delta t/2}^{\Delta t/2} \mathbf{F}_{\text{slow}}(\mathbf{r}(t_n+\tau))\,d\tau \approx \Delta t\mathbf{F}_{\text{slow}}(\overline{\mathbf{r}}^n).$$

That addresses item 1. The second observation is that (35) actually defines a map from the current value of $\mathbf{r}^n = \mathbf{r}(t_n)$ to its local time average $\overline{\mathbf{r}}^n$; denote this map by

$$\overline{\mathbf{r}}^n = \mathcal{A}(\mathbf{r}^n).$$

The averaged (or mollified) slow force is now defined as the gradient of a mollified potential

$$V_{\text{molly}}(\mathbf{r}) = V_{\text{slow}}(\mathcal{A}(\mathbf{r})).$$

Finally the *kick* (33) gets replaced by

$$\mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V_{\text{molly}}(\mathbf{r}^n) \to \mathbf{p}^n.$$

The resulting *mollified MTS* method is symplectic. See the paper by IZAGUIRRE, REICH & SKEEL [62] for implementation details and test results.

There now follows a detailed description of a variant of the method that is particularly easy to implement in the context of bond stretching and bending modes. The method is called EQUILIBRIUM. For simplicity, consider a single bond stretch

$$g(\mathbf{r}) = \|\mathbf{r}_i - \mathbf{r}_j\|$$

with associated potential energy entry

$$V_{\text{fast}} = \frac{\kappa}{2}\left[g(\mathbf{r}) - L\right]^2,$$

where L is the equilibrium length of the bond and κ is a force constant. Assume also that this bond is the only fast vibrational mode and that all other entries in the potential energy function V go into V_{slow}. How is V_{molly} calculated? The fundamental observation is that the average $\bar{\mathbf{r}}$ will (to very good approximation) satisfy $g(\bar{\mathbf{r}}) = 0$ *i.e.*

$$\|\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j\| = L.$$

Hence $\mathcal{A}$ may defined as a simple projection step:

$$\bar{\mathbf{r}} = \mathbf{r} + \mathbf{M}^{-1}\nabla_{\mathbf{r}} g(\mathbf{r})\lambda, \tag{36}$$

$$L = g(\bar{\mathbf{r}}). \tag{37}$$

More explicitly, the single bond leads to an update

$$\bar{\mathbf{r}}_i = \mathbf{r}_i + \frac{1}{m_i}(\mathbf{r}_i - \mathbf{r}_j)\,\lambda, \tag{38}$$

$$\bar{\mathbf{r}}_j = \mathbf{r}_j + \frac{1}{m_j}(\mathbf{r}_j - \mathbf{r}_i)\,\lambda, \tag{39}$$

$$L^2 = \|\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j\|^2, \tag{40}$$

and $\bar{\mathbf{r}}_k = \mathbf{r}_k$ for all $k \neq i, j$, which can be solved for the Lagrange multiplier λ using Newton's method (compare Section §2.2). There is only one difficulty left. To be able to compute the gradient $\nabla_{\mathbf{q}} V_{\text{molly}}(\mathbf{r})$, the Jacobian matrix $\mathcal{A}_{\text{r}} = D_{\mathbf{r}}\bar{\mathbf{r}}$ of $\mathcal{A}$ needs to be calculated since

$$\nabla_{\mathbf{r}} V_{\text{molly}}(\mathbf{r}) = \mathcal{A}_{\mathbf{r}}^T \mathbf{F}_{\text{slow}}(\mathcal{A}(\mathbf{r})).$$

The Jacobian matrix $\mathcal{A}_{\mathbf{r}}$ is identical to the identity matrix except for the 3×3 block entries related to the position vectors $\mathbf{r}_i$, $\mathbf{r}_j$. Ignore the identity part for a moment and write

$$\mathcal{A}_{\text{r}} = \begin{bmatrix} D_{\mathbf{r}_i}\bar{\mathbf{r}}_i & D_{\mathbf{r}_i}\bar{\mathbf{r}}_j \\ D_{\mathbf{r}_j}\bar{\mathbf{r}}_i & D_{\mathbf{r}_j}\bar{\mathbf{r}}_j \end{bmatrix}.$$

The requested partial derivatives can be computed from (38)-(40) via implicit differentiation.

On a more abstract level equations (36)-(37) give:

$$\begin{aligned} \mathcal{A}_{\mathbf{r}} &= \mathbf{I} + \lambda \mathbf{M}^{-1} D_{\mathbf{rr}} g(\mathbf{r}) + \mathbf{M}^{-1} \nabla_{\mathbf{r}} g(\mathbf{r}) D_{\mathbf{r}} \lambda, \\ 0 &= D_{\mathbf{r}} g(\bar{\mathbf{r}}) \mathcal{A}_{\mathbf{r}} \end{aligned}$$

and, after elimination of $D_{\mathbf{r}}\lambda$,

$$\mathcal{A}_{\mathbf{r}} = \left[\mathbf{I} - \frac{1}{\nabla_{\text{r}} g(\mathbf{r}) \cdot \mathbf{M}^{-1} \nabla_{\mathbf{r}} g(\bar{\mathbf{r}})} \nabla_{\mathbf{r}} g(\mathbf{r}) D_{\mathbf{r}} g(\bar{\mathbf{r}}) \right] \left[\mathbf{I} + \lambda \mathbf{M}^{-1} D_{\mathbf{rr}} g(\mathbf{r}) \right].$$

A slightly more stable projection is to use

$$\bar{\mathbf{r}} = \mathbf{r} + \mathbf{M}^{-1} \nabla_{\mathbf{r}} g(\bar{\mathbf{r}}) \lambda$$

instead of (36). The resulting method is called EQUILIBRIUM* in [63] and is somewhat more expensive to implement than EQUILIBRIUM.

EQUILIBRIUM has been implemented in [62] and combined with a weak stochastic stabilization (see the paper by IZAGUIRRE, CATARELLO, WOZNIAK & SKEEL [63]) leading to significantly longer time steps of 14 fs up to 25 fs for explicitly modelled flexible water molecules (as compared to step sizes of 0.5-1 fs for the Störmer-Verlet method and 4 fs for Verlet-I/r-RESPA). Note that EQUILIBRIUM is conceptionally very different from RATTLE/SHAKE as it still resolves all the high frequency bond vibrations.

The described mollified MTS methods are ideally suited for hybrid Monte-Carlo simulations (see GROMOV & DE PABLO [64] for an implementation of r-RESPA into HMC). They are time-reversible and volume conserving. There is also a modified Hamiltonian $\hat{H}_{\Delta t}$ that in many circumstances is quasi-exactly conserved which implies a high acceptance rate of candidate moves.

Other related attempts on mollified/averaged long time step methods have been made by by LEIMKUHLER & REICH (see [65] and [6]) as well as HAIRER, HOCHBRUCK & LUBICH (see the monograph [26] for a summary).

x3 Numerical Methods for Stochastic Constant Temperature Simulations

x3.1 Langevin Dynamics

While there is common agreement that Störmer-Verlet provides the gold standard for constant energy MD simulations, the situation is less clear for Langevin dynamics. This survey will focus on two popular approaches. The first one is the well-known BBK method due to BRÜNGER, BROOKS & KARPLUS [66]. The second method is more recent and is based on a splitting idea originally suggested by SKEEL [57] and implemented in [63]. An elementary introduction to general numerical methods for SDEs can be found in [67].

A slightly modified BBK discretization for (7)-(8) is provided by

$$\mathbf{p}^{n+1/2} = \mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n) - \frac{\gamma}{2}\left(\mathbf{r}^{n+1} - \mathbf{r}^n\right) + \sigma\sqrt{\frac{\Delta t}{2}}\,\mathbf{R}^n, \tag{41}$$

$$\mathbf{r}^{n+1} = \mathbf{r}^n + \Delta t \mathbf{M}^{-1}\mathbf{p}^{n+1/2}, \tag{42}$$

$$\mathbf{p}^{n+1} = \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n) - \frac{\gamma}{2}\left(\mathbf{r}^{n+1} - \mathbf{r}^n\right) + \sigma\sqrt{\frac{\Delta t}{2}}\,\mathbf{R}^{n+1} \tag{43}$$

where $\mathbf{R}^n$ is a vector of independent Gaussian random variables with mean zero and variance one. The BBK scheme reduces to the Störmer-Verlet method for $\gamma = 0$ and is second order accurate in the deterministic part (*i.e.* for $\sigma = 0$). The stochastic scheme is however only weakly first order convergent. For a further discussion of the BBK scheme see [68].

It turns out that the BBK scheme becomes unstable in the Brownian dynamics limit of vanishing inertia, *i.e.* $\mathbf{M} \to 0$, under the assumption of a constant step size Δt. This instability and the limited accuracy of the BBK can be overcome by more advanced integration techniques such as the schemes suggested by VAN GUNSTEREN & BERENDSEN [69] and ALLEN [70]. A variant of these methods, that fits nicely into the framework of splitting, has been recently been suggested by SKEEL [57] and implemented by IZAGUIRRE, CATARELLO, WOZNIAK & SKEEL [63]. The basic splitting idea is simple and can be phrased as *kick-drift/fluctuate-kick*. The new ingredient is

the *drift/fluctuate* step which consists of the exact solution to the SDE

$$\begin{aligned}\dot{\mathbf{p}} &= -\gamma \mathbf{M}^{-1}\mathbf{p} + \sigma \dot{\mathbf{W}}(t),\\ \dot{\mathbf{r}} &= \mathbf{M}^{-1}\mathbf{p}.\end{aligned}$$

The essential steps to derive the exact solution of this linear SDE are provided in Appendix C.

The method *Langevin Impulse* (LI) [63] can be stated as follows:

1. (kick)
$$\overline{\mathbf{p}}^n = \mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n).$$

2. (drift/fluctuate)
$$\mathbf{r}^{n+1} = \mathbf{r}^n + \frac{1-e^{-\gamma \Delta t \mathbf{M}^{-1}}}{\gamma}\overline{\mathbf{p}}^n + \sqrt{\frac{2k_BT\Delta t}{\gamma}}\mathbf{R}_2^n, \tag{44}$$
$$\overline{\mathbf{p}}^{n+1} = e^{-\gamma \Delta t \mathbf{M}^{-1}}\overline{\mathbf{p}}^n + \sqrt{k_BT}\mathbf{M}^{1/2}\mathbf{R}_1^n, \tag{45}$$

   where $\mathbf{R}_1^n$ and $\mathbf{R}_2^n$ are bivariant Gaussian random variables of zero mean and covariance matrix $\mathbf{C}$ given in Appendix C.

3. (kick)
$$\mathbf{p}^{n+1} = \overline{\mathbf{p}}^{n+1} - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}).$$

If $m_i \to 0$ for finite step size Δt, then the LI scheme reduces to the Euler-Maruyama method

$$\mathbf{r}^{n+1} = \mathbf{r}^n - \frac{\Delta t}{\gamma}\nabla_{\mathbf{r}} V(\mathbf{r}^n) + \sqrt{\frac{2k_BT\Delta t}{\gamma}}\mathbf{R}_2^n,$$

which is a popular method for Brownian dynamics (see [1]).

As shown by SKEEL & IZAGUIRRE [71] the evaluation of two sets of random variables per time step can be reduced to one set if one reformulates the LI algorithm in position variables only.

A quite detailed comparison of various algorithms for Langevin dynamics has recently been provided by WANG & SKEEL [72].

x3.2 Extended Dissipative Particle Dynamics (EDPD)

It turns out that it is easier to derive numerical schemes for the extended DPD model (16)-(18) than for the original DPD equations (14)-(15). Furthermore, we will later see that one of the

popular DPD schemes can be derived as the limiting case of an extended DPD algorithm. This situation is similar to the Brownian dynamics limit of the LI scheme for Langevin dynamics. For that reason we first discuss the extended DPD equations before coming back to the standard DPD model.

The following modification (see COTTER & REICH [19]) of the Störmer-Verlet method suggests itself naturally for the extended DPD equations (16)-(18):

$$\mathbf{p}^{n+1/2} = \mathbf{p}^n - \frac{\Delta t}{2}\left[\nabla_{\mathbf{r}} V(\mathbf{r}^n) + \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}^n) s_k^n\right], \tag{46}$$

$$\mathbf{r}^{n+1} = \mathbf{r}^n + \Delta t \mathbf{M}^{-1}\mathbf{p}^{n+1/2}, \tag{47}$$

$$(1+\alpha\Delta t) s_k^{n+1} = s_k^n + \lambda\left[h_k(\mathbf{r}^{n+1}) - h_k(\mathbf{r}^n)\right] + \sqrt{2k_B T\lambda\alpha\Delta t}\, R_k^{n+1}, \tag{48}$$

$$\mathbf{p}^{n+1} = \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\left[\nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}) + \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}^{n+1}) s_k^{n+1}\right], \tag{49}$$

where R_k^{n+1} are independent Gaussian random variables with zero mean and variance one.

The scheme (46)-(49) requires only one force field evaluation per time step, conserves linear and angular momentum within the standard DPD setting, but it not symmetric (or self-consistent in the sense of PAGONABARRAGA, HAGEN & FRENKEL [73]). Hence, replace (48) by the implicit midpoint approximation

$$\left(1+\frac{\alpha\Delta t}{2}\right) s_k^{n+1} = \left(1-\frac{\alpha\Delta t}{2}\right) s_k^n + \lambda\left[h_k(\mathbf{r}^{n+1}) - h_k(\mathbf{r}^n)\right] + \sqrt{2k_B T\lambda\alpha\Delta t}\, R_k^{n+1/2}, \tag{50}$$

where $R_k^{n+1/2}$ are again independent Gaussian random variables with zero mean and variance one. The resulting scheme can be viewed as a generalization of the BBK scheme to the extended DPD model (16)-(18). In fact the scheme shares the advantages and disadvantages of the BBK scheme.

It is possible to derive a more accurate discretization based on a generalization of the LI scheme described in the previous section. The following implementation suggests itself:

1.

$$\begin{aligned}
\bar{\mathbf{p}}^n &= \mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n),\\
\mathbf{r}^{n+1/2} &= \mathbf{r}^n + \frac{\Delta t}{2}\mathbf{M}^{-1}\bar{\mathbf{p}}^n,\\
\bar{s}_k^n &= s_k^n + \lambda\left(h_k(\mathbf{r}^{n+1/2}) - h_k(\mathbf{r}^n)\right).
\end{aligned}$$

2.

$$\begin{aligned}\bar{\mathbf{p}}^{n+1} &= \bar{\mathbf{p}}^{n} + \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}^{n+1/2}) \left[\frac{1-e^{-\alpha\Delta t}}{\alpha} \bar{s}_k^{n} + \sqrt{\frac{2k_BT\lambda\Delta t}{\alpha}} \mathbf{R}_2^n \right], \\ \bar{s}_k^{n+1} &= e^{-\alpha\Delta t} \bar{s}_k^{n} + \sqrt{k_BT\lambda}\,\mathbf{R}_1^n,\end{aligned}$$

where $\mathbf{R}_1^n$ and $\mathbf{R}_2^n$ are joint Gaussian random variables of zero mean and covariance matrix $\mathbf{C}$ given in Appendix C.

3.

$$\begin{aligned}\mathbf{r}^{n+1} &= \mathbf{r}^{n+1/2} + \frac{\Delta t}{2}\mathbf{M}^{-1}\bar{\mathbf{p}}^{n+1}, \\ \mathbf{p}^{n+1} &= \bar{\mathbf{p}}^{n+1} - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}), \\ s_k^{n+1} &= \bar{s}_k^{n+1} + \lambda\left(h_k(\mathbf{r}^{n+1}) - h_k(\mathbf{r}^{n+1/2})\right).\end{aligned}$$

x3.3 Dissipative Particle Dynamics (DPD)

In this section a few popular algorithms for DPD are outlined, starting with an algorithm that can be obtained as the limit of the extended DPD algorithm (46)-(49). More precisely, one can formally consider the limit $\alpha \to \infty$ subject to $\lambda/\alpha = \gamma = const.$ and $\Delta t = const.$ A straightforward calculation yields the limiting scheme

$$\mathbf{p}^{n+1/2} = \mathbf{p}^{n} - \frac{\Delta t}{2}\left[\nabla_{\mathbf{r}} V(\mathbf{r}^n) + \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}^n) s_k^n\right], \tag{51}$$

$$\mathbf{r}^{n+1} = \mathbf{r}^{n} + \Delta t\,\mathbf{M}^{-1}\mathbf{p}^{n+1/2}, \tag{52}$$

$$s_k^{n+1} = \frac{\gamma}{\Delta t}\left[h_k(\mathbf{r}^{n+1}) - h_k(\mathbf{r}^n)\right] + \sqrt{2k_BT\gamma/\Delta t}\, w_k^{n+1}, \tag{53}$$

$$\mathbf{p}^{n+1} = \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\left[\nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}) + \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}^{n+1}) s_k^{n+1}\right], \tag{54}$$

which becomes identical to a scheme suggested for DPD by GROOT & WARREN [74] once $\left[h_k(\mathbf{r}^{n+1}) - h_k(\mathbf{r}^n)\right]/\Delta t$ has been replaced by $\nabla_{\mathbf{r}} h_k(\mathbf{r}^{n+1}) \cdot \mathbf{M}^{-1}\mathbf{p}^{n+1/2}$.

The situation changes if (48) is replaced by the implicit midpoint approximation (50). Let us again investigate the limit $\alpha_k \Delta t \gg 1$. After a few straightforward calculations the limiting equations (51)-(54) with (53) replaced by

$$\frac{s_k^{n+1} - s_k^{n-1}}{2\Delta t} = \gamma\frac{h_k(\mathbf{r}^{n+1}) - 2h_k(\mathbf{r}^n) + h_k(\mathbf{r}^{n-1})}{\Delta t^2} + \sqrt{2k_BT\gamma/\Delta t}\,\frac{w_k^{n+1/2} - w_k^{n-1/2}}{\Delta t}. \tag{55}$$

are derived. Ignoring the noise term for a moment, (55) corresponds to an explicit midpoint discretization in s_k of

$$\frac{d}{dt}s_k = \gamma\frac{d^2}{dt^2}h_k(\mathbf{r}).$$

The explicit midpoint method is known to be unconditionally unstable. This has the implication that the symmetric variant of (46)-(49) cannot be used with a step-size $\Delta t \gg 1/\alpha$ and does not lead to a applicable standard DPD scheme.

Return now to the DPD equations (14)-(15). PAGONABARRAGA, HAGEN & FRENKEL [73] suggested a symmetric (or self-consistent) discretization. Here is the variant of the scheme introduced by BESOLD, VATTULAINEN, KARTTUNEN & POLSON [75]:

$$\begin{aligned}
\mathbf{p}^{n+1/2} &= \mathbf{p}^n - \frac{\Delta t}{2}\left\{\nabla_{\mathbf{r}}V(\mathbf{r}^n) + \sum_{k=1}^{K}\nabla_{\mathbf{r}}h_k(\mathbf{r}^n)\times \right. \\
&\qquad \left.\left[\gamma\nabla_{\mathbf{r}}h_k(\mathbf{r}^n)\cdot\mathbf{M}^{-1}\mathbf{p}^n - \frac{\sigma}{\sqrt{\Delta t}}R_k^n\right]\right\}, \qquad (56)\\
\mathbf{r}^{n+1} &= \mathbf{r}^n + \Delta t\mathbf{M}^{-1}\mathbf{p}^{n+1/2}, \qquad (57)\\
\mathbf{p}^{n+1} &= \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\left\{\nabla_{\mathbf{r}}V(\mathbf{r}^{n+1}) + \sum_{k=1}^{K}\nabla_{\mathbf{r}}h_k(\mathbf{r}^{n+1})\times \right. \\
&\qquad \left.\left[\gamma\nabla_{\mathbf{r}}h_k(\mathbf{r}^{n+1})\cdot\mathbf{M}^{-1}\mathbf{p}^{n+1} - \frac{\sigma}{\sqrt{\Delta t}}R_k^{n+1}\right]\right\}, \qquad (58)
\end{aligned}$$

where R_k^n are independent standard Gaussian random variables with mean zero and variance one.

The self-consistent scheme is essentially a generalization of the BBK scheme [66]. However it is significantly more expensive to implement since the linear system of equations in $\mathbf{p}^{n+1}$ arising from (58) is not diagonal. For a comparison of several numerical algorithms for DPD see BESOLD, VATTULAINEN, KARTTUNEN & POLSON [75].

We finally mention that the complexity of the self-consistent DPD scheme (56)-(58) can be reduced by applying the idea of splitting to the sum of the K dissipation-fluctuation terms. See SHARDLOW [76] for details.

x3.4 A Numerical Example

In this section the stiff spring pendulum equations

$$\begin{aligned}
\dot{\mathbf{r}} &= \mathbf{p},\\
\dot{\mathbf{p}} &= \mathbf{g} - \kappa\frac{\mathbf{r}}{\|\mathbf{r}\|}(\|\mathbf{r}\| - L)
\end{aligned}$$

are used as a model to demonstrate the difference between standard Langevin dynamics, the DPD model, and the extended version of DPD. Here $\mathbf{r},\mathbf{p} \in \mathbb{R}^3$, $\mathbf{g} = (0,0,1)^T$, $\kappa = 1000$ is the spring constant, and $L = 1$ is the equilibrium length of the pendulum. The stiff pendulum equations are Hamiltonian with conserved energy

$$E = \frac{1}{2}\mathbf{p}^T\mathbf{p} + \frac{\kappa}{2}(r-L)^2 - \mathbf{r}^T\mathbf{g}, \qquad r = \|\mathbf{r}\|.$$

Furthermore, because of the relatively large spring constant $\kappa = 1000$, the (oscillatory) energy in the spring is an adiabatic invariant, *i.e.*, J given by

$$J = \frac{1}{2}p_r^2 + \frac{\kappa}{2}(r-L)^2, \qquad p_r = \mathbf{r}^T\mathbf{p}/r,$$

is approximately conserved over long intervals of time [5, 6, 29]. This is demonstrated by using initial conditions $\mathbf{r} = (1,0,0)^T$ and $\mathbf{p} = (1,1,1)^T$ and simulating the system with the Störmer-Verlet method over a time period $t \in [0,100]$ with a step-size of $\Delta t = 0.01$. In Fig. §2, plots of $E(t)$, $J(t)$, and the "slow" rotational energy

$$E_s(t) = E(t) - J(t)$$

are given.

Next, replace the Hamiltonian equations of motion by the Langevin model

$$\begin{aligned} \dot{\mathbf{r}} &= \mathbf{p}, \\ \dot{\mathbf{p}} &= \mathbf{g} - \kappa\frac{\mathbf{r}}{r}(r-L) - \gamma\mathbf{p} + \sigma\dot{\mathbf{W}}(t), \end{aligned}$$

with $\gamma = 0.1$ and $\sigma = 0.2$. It is apparent from Fig. §3 that none of the energy contributions are (even approximately) conserved any longer. Note that, by definition, $J \geq 0$, $E \geq -r \approx -1$, and $E_s \geq -r \approx -1$.

The DPD model, in contrast, leads to a stochastic coupling only in the direction of the stiff spring force and the equations of motion become

$$\begin{aligned} \dot{\mathbf{r}} &= \mathbf{p}, \\ \dot{\mathbf{p}} &= \mathbf{g} - \frac{\mathbf{r}}{r}\left[\kappa(r-L) + \gamma p_r - \sigma\dot{W}(t)\right], \end{aligned}$$

with parameters values as before. Indeed, the simulation results, displayed in Fig. §4, reveal that the slow energy is now approximately conserved while both the total energy and the oscillatory

energy drift in time. This demonstrates the capability of the DPD model to thermostat degrees of freedom in a rather targeted manner. An application of this property will be encountered in the subsequent section.

Finally, replace the DPD model by the extended DPD model

$$\begin{aligned}\dot{\mathbf{r}} &= \mathbf{p},\\ \dot{\mathbf{p}} &= \mathbf{g} - \frac{\mathbf{r}}{r}\left[\kappa(r-L)+s\right],\\ \varepsilon\dot{s} &= -s+\gamma p_r - \sigma\dot{W}(t)\end{aligned}$$

for various values of the parameter $\varepsilon = 1/\alpha$. Note that the extended model reduces to the standard DPD model for $\varepsilon \to 0$. The rapid change in energy behaviour shown in Fig. §5 for parameter values from an interval $1 \geq \varepsilon \geq 0.01$ is quite striking and indicates that the model's response to added fluctuation-dissipation can depend crucially on the decay rate α of the auto-covariance matrix.

x3.5 Stochastic Multiple-Time-Stepping Methods

An obvious way to further improve the performance of the stochastic algorithms presented so far is to combine them with the idea of multiple time stepping (MTS). This combination has been investigated in detail by BARTH & SCHLICK [77] and SANDU & SCHLICK [78]. They use an MTS procedure that is different from Verlet-I/r-RESPA and relies instead on the idea of *extrapolation*. Extrapolation based MTS methods were first considered by STREETT, TILDESLEY & SAVILLE [53]. These MTS methods exhibit a systematic energy drift, a result of not being symplectic. However, this is not such a severe problem if used in combination with friction and stochastic forcing as present in the Langevin or DPD equations. In fact, the LN method of BARTH & SCHLICK [77] alleviates resonance effects of standard MTS methods and outer time steps of 50 fs are reported in [63] for explicit water simulations.

Stochastic formulations alter the dynamics of molecular systems. If the dissipation-fluctuation terms are included to model unresolved degrees of freedom, or to sample from a constant temperature ensemble, that effect is clearly desired. On the other hand, if Langevin dynamics is used to stabilize numerical methods then the impact of the dissipation-fluctuation terms on the otherwise Newtonian description should be minimized. This issue has been addressed by IZAGUIRRE, CATARELLO, WOZNIAK & SKEEL [63]. The essential idea is to combine the mollified MTS method with Langevin dynamics. The resulting methods can be summarized as *kick-*

drift/fluctuate/oscillate-kick. The *kick* step uses a mollified slow energy function as described in Section §2.4 with an outer time step Δt. The *drift/fluctuate/oscillate* step integrates

$$\begin{aligned}\dot{\mathbf{r}} &= \mathbf{M}^{-1}\mathbf{p},\\ \dot{\mathbf{p}} &= -\nabla_{\mathbf{r}} V_{\text{fast}}(\mathbf{r}) - \gamma\dot{\mathbf{r}} + \sigma\dot{\mathbf{W}}(t),\end{aligned}$$

over M time steps with an inner step size $\delta t = \Delta t/M$. The resulting method is called Langevin-Molly (LM).

The LM method applies fluctuation-dissipation terms to all degrees of freedom. As noted by MA & IZAGUIRRE [79] this is not necessary and the Langevin part of the dynamics can be replaced by DPD. To explain the basic procedure, again consider the simple case of a single bond stretching mode

$$V_{\text{fast}} = \frac{\kappa}{2}\left(\|\mathbf{r}_i - \mathbf{r}_j\| - L\right)^2$$

and an otherwise arbitrary slow potential energy function V_{slow}. Denote the mollified slow potential energy by V_{molly} (see Section §2.4). The method *Targeted MOLLY* (TM) [79] consists now of a *kick*

$$\mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V_{\text{molly}}(\mathbf{r}^n) \to \mathbf{p}^n.$$

followed by M integration steps of

$$\dot{\mathbf{r}} = \mathbf{M}^{-1}\mathbf{p}, \tag{59}$$

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{r}} g(\mathbf{r})\left[\kappa g(\mathbf{r}) + \gamma\dot{g}(\mathbf{r}) - \sigma\dot{W}(t)\right], \qquad g(\mathbf{r}) = \|\mathbf{r}_i - \mathbf{r}_j\|, \tag{60}$$

with step size $\delta t = \Delta t/M$. This step replaces the *drift/fluctuate/oscillate* step of the method LM. Finally another *kick* is performed. The method TM is stable with an outer step size of 16 fs and an inner step size of 2 fs when simulating flexible TIP3P water [79].

The function g in (60) can be replaced by a bond angle ϕ, i.e. $g(\mathbf{r}) = \phi(\mathbf{r})$ or even a torsion angle θ, i.e. $g(\mathbf{r}) = \theta(\mathbf{r})$. For a tri-atomic molecule such as water, one could also set

$$g(\mathbf{r}) = \|\mathbf{r}_O - \mathbf{r}_{H1}\| + \|\mathbf{r}_O - \mathbf{r}_{H2}\| + \|\mathbf{r}_{H1} - \mathbf{r}_{H2}\|$$

to simultaneously stabalize/thermostate all three vibrational modes with one stochastic variable $W(t)$.

Another option is to replace the DPD equations (59)-(60) by the corresponding extended DPD equations. However this method has not yet been implemented for flexible water.

We finally mention the idea of replacing bond stretching and bending modes by constraints as outlines in Section §2.2. In the context of stochastic MD algorithms one might want to substitute the symplectic SHAKE/RATTLE algorithm by a cheaper (non-symplectic) method. A wide variety of suitable candidate methods can be found in the multi-body dynamics literature. See the text by EICH-SOELLNER & FÜHRER [80] and the survey by ASCHER, CHIN, PETZOLD & REICH [81] on constraint stabilization techniques. The method LINCS by HESS, BEKKER, BERENDSEN & FRAAIJE [82] represents one such stabilized constraint algorithm applied to molecular dynamics.

x4 Numerical Methods for Nosé-Hoover Constant Temperature Dynamics

The NOSÉ-HOOVER [7, 8] thermostat is a *deterministic* molecular dynamics simulation method that allows one to sample from the *canonical distribution* for a given temperature T. The derivation of the thermostat model starts with the virtual variable formulation due to NOSÉ [7]. Postulate a Hamiltonian

$$\mathcal{H}_{\text{Nosé}} = \frac{1}{2}\sum_{i=1}^{N} \|\tilde{\mathbf{p}}_i\|^2/(m_i s^2) + V(\mathbf{r}) + p_s^2/(2Q) + \tilde{g}k_B T \ln s, \tag{61}$$

where p_s is the variable conjugate to s, Q is a fixed constant of dimension energy·(times)^2, and $\tilde{g} = 3N+1$. The equations of motion derived from the Hamiltonian are

$$\frac{d}{d\tau}\mathbf{r} = s^{-2}\mathbf{M}^{-1}\tilde{\mathbf{p}}, \tag{62}$$

$$\frac{d}{d\tau}\tilde{\mathbf{p}} = -\nabla_{\mathbf{r}} V(\mathbf{r}), \tag{63}$$

$$\frac{d}{d\tau}s = \frac{p_s}{Q}, \tag{64}$$

$$\frac{d}{d\tau}p_s = s^{-3}\tilde{\mathbf{p}}^T\mathbf{M}^{-1}\tilde{\mathbf{p}} - s^{-1}\tilde{g}k_B T. \tag{65}$$

NOSÉ [7] has shown that the dynamics associated with this system generates canonically distributed configurations in the variables $\mathbf{p} = \tilde{\mathbf{p}}/s$, $\mathbf{r}$ provided the equations of motion are ergodic. However, time is sampled at a virtual time scale τ with real time t related by the differential transformation

$$dt = s^{-1}d\tau.$$

The most popular reformulation of (62)-(65) in terms of real time t and real momenta $\mathbf{p}$ is due to HOOVER [8] and has the appealing form

$$\frac{d}{dt}\mathbf{r} = \mathbf{M}^{-1}\mathbf{p}, \tag{66}$$

$$\frac{d}{dt}\mathbf{p} = -\nabla_{\mathbf{r}}V(\mathbf{r}) - \xi\mathbf{p}, \tag{67}$$

$$\frac{d}{dt}\xi = \left[\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} - gk_BT\right]/Q, \tag{68}$$

with $g = 3N$ and $\xi = p_s/Q$. The Nosé-Hoover equations (66)-(68) resemble Langevin dynamics except that there is no noise term and the damping factor has become a dynamic variable.

However, even though equations (62)-(65) were derived from a Hamiltonian $\mathcal{H}_{\text{Nosé}}$, the Nosé-Hoover equations (66)-(68) are no longer Hamiltonian due to the non-symplectic nature of the intermediate transformations. This problem was rectified by BOND, LEIMKUHLER & LAIRD [9] who suggested applying a Poincaré time transformation. The transformation leads to the new extended Hamiltonian

$$\mathcal{H}_{NVT} = s\left(\mathcal{H}_{\text{Nosé}} - H_0\right),$$

where H_0 is the initial value of $\mathcal{H}_{\text{Nosé}}$ at time zero. The associated Hamiltonian equations of motion in real time t are

$$\frac{d}{dt}\mathbf{r} = s^{-1}\mathbf{M}^{-1}\tilde{\mathbf{p}}, \tag{69}$$

$$\frac{d}{dt}\tilde{\mathbf{p}} = -s\nabla_{\mathrm{r}}V(\mathbf{r}), \tag{70}$$

$$\frac{d}{dt}s = s\frac{p_s}{Q}, \tag{71}$$

$$\frac{d}{dt}p_s = \frac{1}{2s^2}\tilde{\mathbf{p}}^T\mathbf{M}^{-1}\tilde{\mathbf{p}} - V(\mathbf{r}) - \frac{p_s^2}{2Q} - gk_BT(1+\ln s) + H_0. \tag{72}$$

Symplectic integration in time implies an excellent long-time stability of the simulation and that has been demonstrated by BOND, LEIMKUHLER & LAIRD in [9]. A particularly elegant implementation of such a symplectic method has been proposed by NOSÉ [83]. His idea is based on a splitting of the Hamiltonian $\mathcal{H}_{NVT}$ into integrable sub-problems, *i.e.*

$$\mathcal{H}_{\text{NVT}} = \mathcal{H}_1 + \mathcal{H}_2 + \mathcal{H}_3 + \mathcal{H}_4 + \mathcal{H}_5$$

with $\mathcal{H}_1 = \mathcal{H}_5 = sp_s^2/(4Q)$, $\mathcal{H}_2 = \mathcal{H}_4 = sV(\mathbf{r})/2$ and

$$\mathcal{H}_3 = s\left[\frac{1}{2s^2}\tilde{\mathbf{p}}^T\mathbf{M}^{-1}\tilde{\mathbf{p}} + gk_BT\ln s - H_0\right].$$

The numerical method is then given by the concatenation of flow-maps

$$\Phi_{\Delta t} = \Psi_{\Delta t,\mathcal{H}_1} \circ \Psi_{\Delta t,\mathcal{H}_2} \circ \Psi_{\Delta t,\mathcal{H}_3} \circ \Psi_{\Delta t,\mathcal{H}_4} \circ \Psi_{\Delta t,\mathcal{H}_5}.$$

The exact solution operator $\Psi_{\Delta t,\mathcal{H}_2} = \Psi_{\Delta t,\mathcal{H}_4}$ is easily found. It also turns out that the system of equations

$$\begin{aligned}
\frac{d}{dt}\mathbf{r} &= s^{-1}\mathbf{M}^{-1}\tilde{\mathbf{p}}, \\
\frac{d}{dt}\tilde{\mathbf{p}} &= \mathbf{0}, \\
\frac{d}{dt}s &= 0, \\
\frac{d}{dt}p_s &= \frac{1}{2s^2}\tilde{\mathbf{p}}^T\mathbf{M}^{-1}\tilde{\mathbf{p}} - gk_BT(1+\ln s) + H_0.
\end{aligned}$$

corresponding to $\mathcal{H}_3$ is straightforward to integrate. It remains to solve the coupled system

$$\frac{d}{dt}s = s\frac{p_s}{2Q}, \qquad \frac{d}{dt}p_s = -\frac{p_s^2}{4Q}$$

arising from the Hamiltonian $\mathcal{H}_1 = \mathcal{H}_5$. The analytic solutions are given by

$$\frac{1}{p_s(t)} = \frac{1}{p_s(0)} + \frac{t}{4Q}$$

and

$$s(t) = s(0) + \frac{s(0)p_s(0)}{2Q}t + \frac{s(0)p_s(0)^2}{16Q}t^2 = s(0)\left(1 + \frac{p_s(0)}{4Q}t\right)^2.$$

After a few manipulations, the following numerical scheme in terms of the real momentum

approximations $\mathbf{p}^n = \tilde{\mathbf{p}}^n/s^n$, $\mathbf{p}^{n+1/2} = \tilde{\mathbf{p}}^{n+1/2}/s^{n+1/2}$, and $\mathbf{p}^{n+1} = \tilde{\mathbf{p}}^{n+1}/s^{n+1}$ is obtained:

$$
\begin{aligned}
s^{n+1/2} &= s^n \left(1 + \frac{p_s^n}{2Q}\frac{\Delta t}{2}\right)^2, \\
p_s^* &= p_s^n / \left(1 + \frac{p_s^n}{2Q}\frac{\Delta t}{2}\right), \\
\mathbf{p}^{n+1/2} &= \left(1 - \frac{p_s^*}{2Q}\frac{\Delta t}{2}\right)^2 \mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n), \\
\mathbf{r}^{n+1} &= \mathbf{r}^n + \Delta t \mathbf{M}^{-1}\mathbf{p}^{n+1/2}, \\
p_s^{**} &= p_s^* + \Delta t \left[\frac{1}{2}(\mathbf{p}^{n+1/2})^T \mathbf{M}^{-1}\mathbf{p}^{n+1/2} - \right. \\
&\qquad \left. - \frac{1}{2}(V(\mathbf{r}^{n+1}) + V(\mathbf{r}^n)) - g k_B T (1 + \ln s^{n+1/2}) + H_0\right], \\
s^{n+1} &= s^{n+1/2}\left(1 + \frac{p_s^{**}}{2Q}\frac{\Delta t}{2}\right)^2, \\
p_s^{n+1} &= p_s^{**} / \left(1 + \frac{p_s^{**}}{2Q}\frac{\Delta t}{2}\right), \\
\mathbf{p}^{n+1} &= \left(1 - \frac{p_s^{n+1}}{2Q}\frac{\Delta t}{2}\right)^2 \left(\mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n)\right).
\end{aligned}
$$

Time averages for an observable $\mathcal{A}$ computed along a numerical trajectory will converge to the macro-canonical ensemble average value of $\mathcal{A}$ under the assumption of micro-canonical ergodicity of the equations of motion (which we assume to carry over to the numerical scheme). However, one has to be careful when giving a dynamic interpretation of trajectories since artificial forces were introduced to generate a constant temperature ensemble simulation. On the other hand, Nosé-Hoover constant molecular dynamics algorithms lead to less expensive implementations than constant temperature stochastic algorithms. This is due to the fact that the generation of a large set of random numbers at each time step can take up a significant amount of computer time. There is also the danger that the introduction of stochastic terms could perturb the classical molecular dynamics picture in a more significant manner than NOSÉ's reformulation.

x5 Another Application. Particle Methods for Ideal Fluids

Throughout this survey we have so far considered algorithms suitable for constant energy or constant temperature molecular dynamics (MD) simulations. On the other hand, dissipative particle dynamics was introduced in an attempt to model coarse-grained molecular dynamics with an essentially hydrodynamic behavior [84]. More recently the original DPD algorithm has been

extended by several groups (see, for example, TROFIMOV, NIES & MICHELS [85], ESPAÑOL & REVENGA, [86] and the survey by ESPAÑOL [87]) to allow for a more accurate description of the thermodynamics of real systems. These results have led to an interesting link between DPD and the *Smoothed Particle Hydrodynamics* (SPH) method of LUCY [88], GINGOLD & MONAGHAN [89], and SALMON [90]. As proposed originally SPH is a Lagrangian particle method for ideal hydrodynamics that leads to a large conservative system of interacting particles.

Newtonian equations of motion describing ideal fluid dynamics can also be derived within the classical particle-mesh methodology [91–93] and that is the approach focused on in this section. Below we will outline how dissipative particle dynamics naturally enters into the particle-mesh model equations.

First the basic ideas for the two-dimensional shallow-water equations (SWEs) describing a shallow layer of fluid subject to gravity [94] are explained. The Lagrangian formulation of the SWEs is

$$\ddot{\mathbf{X}} = -c_0^2 \nabla_{\mathbf{X}} h(\mathbf{X}),$$

where $\mathbf{X} = (X,Y)^T$ are the particle positions (a continuous function of both space and time), $c_0 = \sqrt{gH}$, g is the gravitational constant, and H is the mean layer-depth. The normalized layer-depth h is given by the convolution

$$h(\mathbf{x},t) = \int h_0(\mathbf{a})\, \delta(\mathbf{x} - \mathbf{X}(\mathbf{a},t))\, d\mathbf{a},$$

where $\mathbf{a} = \mathbf{X}(\mathbf{a},0)$ are the initial particle positions and $h_0(\mathbf{a})$ is the initial layer-depth at position $\mathbf{a}$. See [95, 96] for more details.

FRANK, GOTTWALD & REICH [95, 96] suggested a particle-mesh method, called the Hamiltonian particle-mesh (HPM) method, for the solution of the two-dimensional shallow-water equations. The HPM method may be viewed as an accurate numerical discretisation of the regularised fluid equations:

$$\ddot{\mathbf{X}} = -c_0^2 \nabla_{\mathbf{X}} [\mathscr{A} * h(\mathbf{X})]$$

where $\mathscr{A}$ is a smoothing operator with some smoothing length Λ. This smoothing length is of course much larger than any molecular length scale but shorter than the desired resolution of the discrete model. Hence one can think of the HPM method as a Newtonian particle system representing mesoscale fluid dynamics. For later reference, denote the numerically unresolved part of the layer-depth by

$$\eta = h - \mathscr{A} * h. \tag{73}$$

The HPM method uses a regular grid $\mathbf{x}_{kl} = (k\Delta x, l\Delta y)^T$, particles $\mathbf{X}_i = (X_i, Y_i)^T$, grid-centred basis functions $\psi_{kl}(\mathbf{X})$, and the layer-depth approximation

$$\tilde{h}_{kl}(t) = \sum_{i=1}^{N} m_i \psi_{kl}(\mathbf{X}_i)$$

at $\mathbf{x}_{kl}$. The basis functions form a partition of unity, *i.e.* $\sum_{k,l} \psi_{kl}(\mathbf{X}) = 1$. The smoothing operator $\mathscr{A}$ is now defined as the discretization of the inverse modified Helmholtz operator with smoothing length $\Lambda = 4\Delta x$ over the grid $\mathbf{x}_{kl}$. The discrete approximation is denoted by $\{a_{kl}^{mn}\}$. Consequently, the finite-dimensional Hamiltonian equations of motion are given by

$$\dot{\mathbf{X}}_i = \mathbf{V}_i, \tag{74}$$

$$\dot{\mathbf{V}}_i = -c_0^2 \sum_{k,l} \nabla_{\mathbf{X}_i} \psi_{kl}(\mathbf{X}_i) \hat{h}_{kl}, \qquad \hat{h}_{kl} = \sum_{m,n} a_{kl}^{mn} \tilde{h}_{mn}. \tag{75}$$

For further implementation details see again [95, 96].

Within this setting the numerical issues of propagating (74)-(75) in time are now very similar to what has been discussed for constant energy molecular dynamics. However, there is also a close link to stochastic molecular dynamics algorithms. Namely, the numerically unresolved gravity waves in the layer-depth (73) can be modelled by a generalized Langevin process. This idea can be mathematically motivated by representing η as the solution of a linear wave equation coupled to the particle system and subsequent reduction following the Kac-Zwanzig approach (see ESPAÑOL [20], REY-BELLET & THOMAS [21], and KUPFERMAN, STUART, TERRY & TUPPER [22] for related approaches). The assumption of exponentially decaying kernel (which can be obtained by assuming that the waves are 'localised') leads to the EDPD extension of the HPM equations given by COTTER & REICH [19] as

$$\dot{\mathbf{X}}_i = \mathbf{V}_i, \tag{76}$$

$$\dot{\mathbf{V}}_i = -c_0^2 \sum_{k,l} \left[\hat{h}_{kl} + \eta_{kl}\right] \nabla_{\mathbf{X}_i} \psi_{kl}(\mathbf{X}_i), \tag{77}$$

$$\dot{\eta}_{kl} = -\alpha \eta_{kl} + \lambda d_{kl} + c_0^{-1} (2k_B T \alpha \lambda)^{1/2} \dot{W}_{kl}, \tag{78}$$

where d_{kl} is defined by

$$d_{kl} = \frac{d}{dt} \hat{h}_{kl} = \sum_i m_i \nabla_{\mathbf{X}_i} \psi_{kl}(\mathbf{X}_i) \cdot \mathbf{V}_i.$$

The only necessary modification is a scaling of the term multiplying $\dot{W}_{kl}$ by c_0^{-1}. The scaling is necessary because c_0^2 multiplies the force term in (77). It should be noted that $h_{kl} = \hat{h}_{kl} + \eta_{kl}$ is

ment to be a better approximation to the "true" layer-depth h than $\hat{h}_{kl}$ alone. See the PhD thesis of COTTER [97] for further details.

The extended model still preserves circulation (see FRANK AND REICH [96]).

Lattice Boltzmann hydrodynamics simulations provide another very interesting link between molecular statistical mechanics and macroscopic continuum descriptions. See SUCCI, KARLIN & CHEN [98] for a survey. However the numerical issues involved are quite different from those discussed in this survey.

As demonstrated by FRANK & REICH [99], particle methods can also be used for the *shallow water equations on a rotating sphere*. These equations then provide a simple model for *large scale atmospheric circulation* and can easily be generalized to multi-layer *primitive equation models* [94, 100]. Hence it transpires from these discussions that some of the basic algorithmic challenges addressed in this survey not only concern molecular and nano-scale modelling but carry over to the modelling of large scale atmospheric circulation patterns and all the way to climate research.

Acknowledgments. We would like to thank David Heyes, Jesus Izaguirre, Mark Ma and Andrew Stuart for carefully reading the paper and for providing us with plenty of feedback.

Appendix A (Hamiltonian Mechanics)

Let $\mathbf{r}$ denote an n-vector of particle positions and $\mathbf{p}$ the associated vector of conjugate momenta. Given a Hamiltonian $\mathcal{H}$ the associated canonical Hamiltonian equations of motion are

$$\dot{\mathbf{r}} = +\nabla_{\mathbf{p}}\mathcal{H}(\mathbf{r},\mathbf{p}), \qquad \dot{\mathbf{p}} = -\nabla_{\mathbf{r}}\mathcal{H}(\mathbf{r},\mathbf{p}).$$

Upon concatenating the positions $\mathbf{r}$ and the momenta $\mathbf{p}$ into one vector $\mathbf{z} = (\mathbf{r}^T, \mathbf{p}^T)^T \in \mathbb{R}^{2n}$, the Hamiltonian equations can be condensed into the compact form

$$\dot{\mathbf{z}} = \mathbf{J}\nabla_{\mathbf{z}}\mathcal{H}(\mathbf{z}), \quad \mathbf{J} = \begin{bmatrix} \mathbf{0}_n & \mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{0}_n \end{bmatrix}.$$

The linearized equations along $\mathbf{z}(t)$ are given by

$$\dot{\mathbf{Z}} = \mathbf{J}\mathbf{A}(t)\,\mathbf{Z}, \qquad \mathbf{A}(t) = D_{\mathbf{zz}}H(\mathbf{z}(t)).$$

Let $\mathbf{V}(t) \in \mathbb{R}^{2n}$ and $\mathbf{U}(t) \in \mathbb{R}^{2n}$ denote any two solutions of the linearized equations. Note that the matrix $\mathbf{J}$ is skew-symmetric while $\mathbf{A}(t)$ is symmetric. Then one can easily verify that

$$\frac{d}{dt}\left[\mathbf{V}(t)^T \mathbf{J}^{-1} \mathbf{U}(t)\right] = 0.$$

The *symplectic two-form* Ω is now defined by

$$\Omega(\mathbf{U}, \mathbf{V}) = \mathbf{U}^T \mathbf{J}^{-1} \mathbf{V}.$$

Hence conservation of symplecticness (*i.e.* $\dot{\Omega} = 0$) can be concluded along solutions of the linearized equations.

A similar procedure applies to maps. Given a map Φ the linearization (Jacobian matrix) can be used to propagate two vectors $\mathbf{U}^n$ and $\mathbf{V}^n$. The map Φ is called *symplectic* if

$$\Omega(\mathbf{U}^{n+1}, \mathbf{V}^{n+1}) = \Omega(\mathbf{U}^n, \mathbf{V}^n).$$

For planar maps conservation of sympleticness is equivalent to conservation of area. For higher dimensional maps this analogy becomes slightly more complex but a consequence is conservation of volume in phase space.

The idea that a numerical method for classical mechanics should preserve the symplectic two-form can first be found in a technical report by VOGELAERE [101].

Appendix B (Rigid Body Dynamics)

In this appendix a symplectic splitting method for rigid body motion is briefly outlined. The algorithm can be used to simulated (small) molecular units that may be modelled as a rigid body (such as water). Only a single rigid body consisting of N rigidly linked particles is considered for simplicity. Each particle has a mass m_i, coordinates $\mathbf{r}_i$, and an applied force $\mathbf{F}_i$ acting on. The total mass of the rigid body is then $M = \sum_i m_i$ and the center of mass is given by

$$\mathbf{r}_c = \frac{1}{M} \sum_{i=1}^{N} m_i \mathbf{r}_i.$$

Also define the relative position vectors $\xi_i = \mathbf{r}_i - \mathbf{r}_c$.

Next, introduce a body coordinate system with its three orthogonal basis vectors denoted by $\hat{\mathbf{e}}_\alpha$, $\alpha = 1, 2, 3$. This coordinate system is chosen such that its origin coincides with the center of mass

$\mathbf{r}_c$ of the rigid body and the three basis vectors point in the direction of the principle moments of intertia, *i.e.*, the intertia tensor $\mathbf{I} \in \mathbb{R}^{3\times 3}$ with entries

$$I_{\alpha\beta} = \sum_{i=1}^{N} m_i(\hat{\mathbf{e}}_\alpha \times \xi_i) \cdot (\hat{\mathbf{e}}_\beta \times \xi_i)$$

is diagonal.

Assume also a fixed spatial coordinate system with its basis vectors denoted by $\mathbf{e}_\alpha$, $\alpha = 1,2,3$. There is an orthogonal matrix $\mathbf{Q}$ that link the two sets of basis vectors, *i.e.*

$$\mathbf{e}_\alpha = \mathbf{Q}\hat{\mathbf{e}}_\alpha, \qquad \alpha = 1,2,3.$$

As the rigid body changes position as a function of time the body coordinate system also moves and the following time-dependent relation is derived:

$$\mathbf{e}_\alpha = \mathbf{Q}(t)\hat{\mathbf{e}}_\alpha(t), \qquad \alpha = 1,2,3.$$

Differention of this relation with respect to time yields the equation

$$\begin{aligned} \mathbf{0} &= \left(\frac{d}{dt}\mathbf{Q}\right)\hat{\mathbf{e}}_\alpha + \mathbf{Q}\left(\frac{d}{dt}\hat{\mathbf{e}}_\alpha\right) \\ &= \left(\frac{d}{dt}\mathbf{Q}\right)\hat{\mathbf{e}}_\alpha - \mathbf{Q}(\omega \times \hat{\mathbf{e}}_\alpha) \\ &= \left[\frac{d}{dt}\mathbf{Q} - \mathbf{Q}\Omega\right]\hat{\mathbf{e}}_\alpha, \end{aligned} \tag{79}$$

where ω is the body angular velocity vector and

$$\Omega = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}, \quad \omega = (\omega_1, \omega_2, \omega_3)^T$$

is the associated skew-symmetric matrix which satisfies

$$\frac{d}{dt}\hat{\mathbf{e}}_\alpha = -\omega \times \hat{\mathbf{e}}_\alpha = -\Omega\hat{\mathbf{e}}_\alpha.$$

Equation (79) provides the differential equation used to update the rotation matrix $\mathbf{Q}$:

$$\frac{d}{dt}\mathbf{Q} = \mathbf{Q}\Omega.$$

The *Euler equation* (see ARNOLD [5] or LEIMKUHLER & REICH [6]) for the body angular momentum vector $\pi = \mathbf{I}\omega$ is

$$\frac{d}{dt}\pi = \pi \times \mathbf{I}^{-1}\pi + \mathbf{Q}^T \sum_{i=1}^{N} (\mathbf{F}_i \times \xi_i).$$

Finally, the spatial position of the i^{th} particle is given by

$$\mathbf{r}_i(t) = \mathbf{r}_c(t) + \mathbf{Q}(t)\mathbf{Q}^T(0)(\mathbf{r}_i(0) - \mathbf{r}_c(0)),$$

where the center of mass moves according to the system of equations

$$\begin{aligned} \frac{d}{dt}\mathbf{r}_c &= \mathbf{p}_c/M, \\ \frac{d}{dt}\mathbf{p}_c &= \sum_{i=1}^{N} \mathbf{F}_i. \end{aligned}$$

This closes the equations of motion for the rigid body.

The rigid body equations of motion are Hamiltonian with a non-canonical Lie-Poisson structure [5, 6]. The Hamiltonian is given by

$$\mathcal{H} = \frac{1}{2M}\mathbf{p}_c^T\mathbf{p}_c + \frac{1}{2}\pi^T\mathbf{I}^{-1}\pi + V(\mathbf{r}_1, \ldots, \mathbf{r}_N),$$

which can be split into kinetic and potential energy. Hence a symplectic splitting scheme can follow the *kick-drift-kick* methodology of the Störmer-Verlet method with the important difference that the *drift* step gets replaced by *drift/rotate*. The *drift/rotate* step itself gets decomposed into three sub-step. More specifically, decompose the inverse of the (diagonal) inertia tensor as follows:

$$\begin{aligned} \mathbf{I}^{-1} &= \begin{bmatrix} 1/I_1 & 0 & 0 \\ 0 & 1/I_2 & 0 \\ 0 & 0 & 1/I_3 \end{bmatrix} \\ &= \frac{1}{I_1}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{I_1-I_2}{I_1I_2} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{I_1-I_3}{I_1I_3} \end{bmatrix} \\ &= \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3. \end{aligned}$$

It is easy to verify that each associated system of equations

$$\begin{aligned} \frac{d}{dt}\pi &= \pi \times \mathbf{T}_i\pi, \\ \frac{d}{dt}\mathbf{Q} &= \mathbf{Q}\Omega, \qquad \omega = \mathbf{T}_i\pi \ (= const.), \end{aligned}$$

$i = 1, 2, 3$, can be integrated exactly. Hence, a splitting method can be implemented as follows:

$$\Phi_{\Delta t} = \Psi_{\Delta t, \mathcal{H}_1} \circ \Psi_{\Delta t, \mathcal{H}_2} \circ \Psi_{\Delta t, \mathcal{H}_3} \circ \Psi_{\Delta t, \mathcal{H}_4} \circ \Psi_{\Delta t, \mathcal{H}_5} \circ \Psi_{\Delta t, \mathcal{H}_6}$$

with

$$\begin{aligned}
\mathcal{H}_1 &= \mathcal{H}_6 = \frac{1}{2}V(\mathbf{r}_1,\ldots,\mathbf{r}_N),\\
\mathcal{H}_2 &= \frac{1}{2M}\mathbf{p}_c^T\mathbf{p}_c + \frac{1}{2I_1}\pi^T\pi,\\
\mathcal{H}_3 &= \mathcal{H}_5 = \frac{1}{4}\pi^T\mathbf{T}_2\pi,\\
\mathcal{H}_4 &= \frac{1}{4}\pi^T\mathbf{T}_3\pi.
\end{aligned}$$

The equations of motion for $\mathcal{H}_1 = \mathcal{H}_6$ are given by

$$\begin{aligned}
\frac{d}{dt}\pi &= \frac{1}{2}\mathbf{Q}^T\sum_{i=1}^{N}(\mathbf{F}_i\times\xi_i),\\
\frac{d}{dt}\mathbf{Q} &= \mathbf{0},\\
\frac{d}{dt}\mathbf{p}_c &= \frac{1}{2}\sum_{i=1}^{N}\mathbf{F}_i,\\
\frac{d}{dt}\mathbf{q}_c &= \mathbf{0}.
\end{aligned}$$

It turns out that the splitting scheme is also time-reversible because some of the flow maps commute.

Appendix C (Stochastic Differential Equations)

First review the explicit solution of a scalar (for simplicity) stochastic differential equation

$$dy = -ay\,dt + b\,dW(t),$$

where $W(t)$ is a Wiener process, a and b are constants. The Wiener process can be defined by

$$W(t+\tau) - W(t) = \sqrt{\tau}Z(t),$$

where $W(0) = 0$ and $Z(t)$ are independent standard Gaussian random variables (with mean zero and variance one).

The integrating factor formulation

$$d\left[e^{at}y\right] = e^{at}b\,dW$$

gives

$$y(\tau) = e^{-a\tau}y(0) + \int_0^\tau be^{-a(\tau-t)}dW(t).$$

The last integral represents a Gaussian process $r_1(\tau)$ with zero mean and variance

$$E\left[r_1(\tau)^2\right] = \int_0^\tau b^2 e^{-2a(\tau-t)}dt = \frac{b^2}{2a}\left(1-e^{-2a\tau}\right)$$

Often the integral of the Ornstein-Uhlenbeck process $y(\tau)$ is also required:

$$\Delta I(\tau) = \int_0^\tau y(t)dt = \frac{1-e^{-a\tau}}{a}y(0) + \int_0^\tau \int_0^s be^{-a(s-t)}dW(t)ds$$

Using integration by parts, the double integral again defines a Gaussian process

$$r_2(\tau) = \frac{b}{a}\int_0^\tau \left[1-e^{-a(\tau-t)}\right]dW(t)$$

with zero mean and variance

$$E\left[r_2(\tau)^2\right] = \frac{b^2}{a^2}\int_0^\tau \left[1-e^{-a(\tau-t)}\right]^2 dt = \frac{\tau b^2}{a^2}\left(1+\frac{4e^{-a\tau}-3-e^{-2a\tau}}{2a\tau}\right).$$

The two processes $r_1(\tau)$ and $r_2(\tau)$ are also cross correlated according to

$$E\left[r_1(\tau)r_2(\tau)\right] = \frac{b^2}{a}\int_0^\tau e^{-a(\tau-t)}\left[1-e^{-a(\tau-t)}\right]dt = \frac{b^2}{2a^2}\left(1-2e^{-a\tau}+e^{-2a\tau}\right).$$

Hence the covariance matrix

$$\begin{aligned}\mathbf{C}(\tau) &= \begin{bmatrix} E\left[r_1(\tau)r_1(\tau)\right] & E\left[r_1(\tau)r_2(\tau)\right] \\ E\left[r_1(\tau)r_2(\tau)\right] & E\left[r_2(\tau)r_2(\tau)\right]\end{bmatrix} \\ &= \begin{bmatrix} b/\sqrt{2a} & 0 \\ 0 & \sqrt{\tau}b/a\end{bmatrix}\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22}\end{bmatrix}\begin{bmatrix} b/\sqrt{2a} & 0 \\ 0 & \sqrt{\tau}b/a\end{bmatrix},\end{aligned}$$

is obtained where

$$c_{11} = 1-e^{-2a\tau}, \quad c_{12} = c_{21} = \frac{1-2e^{-a\tau}+e^{-2a\tau}}{\sqrt{2a\tau}}, \quad c_{22} = 1+\frac{4e^{-a\tau}-3-e^{-2a\tau}}{2a\tau}.$$

Applying a Cholesky factorization to the covariance matrix, a numerical scheme now reads as follows

$$\begin{aligned}y^{n+1} &= e^{-a\Delta t}y^n + \frac{b}{\sqrt{2a}}R_1^n, \\ \Delta I^n &= \frac{1-e^{-a\tau}}{a}y^n + \frac{\sqrt{\Delta t}b}{a}R_2^n,\end{aligned}$$

where

$$\begin{bmatrix} R_1^n \\ R_2^n\end{bmatrix} = \begin{bmatrix} \sqrt{c_{11}} & 0 \\ \frac{c_{21}}{\sqrt{c_{11}}} & \sqrt{c_{22}-\frac{c_{21}^2}{c_{11}}}\end{bmatrix}\begin{bmatrix} Z_1^n \\ Z_2^n\end{bmatrix}$$

and Z_i^{n+1}, $i = 1,2$, are independent Gaussian random numbers with mean zero and variance one.

As an application, consider the Langevin equation for a single particle with mass m and friction coefficient γ. One finds that $a = \gamma/m$, $b = \sqrt{2k_BT\gamma}$ and the numerical scheme becomes

$$\begin{aligned}
\mathbf{p}^{n+1} &= e^{-\gamma\Delta t/m}\mathbf{p}^n + \sqrt{\frac{k_BT}{m}}\mathbf{R}_1^n, \\
\mathbf{r}^{n+1} &= \mathbf{r}^n + \frac{1-e^{-\gamma\Delta t/m}}{\gamma}\mathbf{p}^n + \sqrt{\frac{2k_BT\Delta t}{\gamma}}\mathbf{R}_2^n
\end{aligned}$$

where

$$\begin{bmatrix} \mathbf{R}_1^n \\ \mathbf{R}_2^n \end{bmatrix} = \begin{bmatrix} \sqrt{c_{11}}\,\mathbf{I} & \mathbf{0} \\ \frac{c_{21}}{\sqrt{c_{11}}}\,\mathbf{I} & \sqrt{c_{22} - \frac{c_{21}^2}{c_{11}}}\,\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{Z}_1^n \\ \mathbf{Z}_2^n \end{bmatrix}.$$

The result easily generalizes to systems of particles. For further details see also [57, 63].

The corresponding numerical method for a system of type

$$\begin{aligned}
\dot{\mathbf{r}} &= \mathbf{0}, \\
\dot{\mathbf{p}} &= -\nabla_{\mathbf{r}} h(\mathbf{r})s, \\
\dot{s} &= -\alpha s + \sqrt{2k_BT\alpha\lambda}\dot{W},
\end{aligned}$$

as needed for the EDPD splitting method, is now straightforward, using $a = \alpha$ and $b = \sqrt{2k_BT\lambda\alpha}$.

---

[1] T. Schlick, *Molecular Modeling and Simulation* (Springer-Verlag, New-York, 2002).

[2] M. Allen and D. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987).

[3] D. Rappaport, *The Art of Molecular Dynamics Simulation* (Cambridge University Press, New York, 1995).

[4] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, New York, 2001), 2nd ed.

[5] V. Arnold, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York, 1989), 2nd ed.

[6] B. Leimkuhler and S. Reich, *Geometric Numerical Methods for Hamiltonian Mechanics* (Cambridge University Press, Cambridge, 2004).

[7] S. Nosé, J. Chem. Phys. **81**, 511 (1984).

[8] W. Hoover, Phys. Rev. A **31**, 1695 (1985).

[9] S. Bond, B. Leimkuhler, and B. Laird, J. Comput. Phys. **151**, 114 (1999).

[10] S. Duane, A. Kennedy, B. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).

[11] Z. Breźniak and T. Zastawniak, *Basic Stochastic Processes* (Springer-Verlag, London, 1999).

[12] B. Oksendal, *Stochastic Differential Equations* (Springer-Verlag, Berlin-Heidelberg, 2000), 5th ed.

[13] D. Heyes and A. Brańka, Mol. Phys. **98**, 1949 (2000).

[14] H. Mori, Prog. Theor. Phys. **33**, 423 (1965).

[15] H. Mori, Prog. Theor. Phys. **34**, 399 (1965).

[16] D. Ermak and H. Buckholtz, J. Comput. Phys. **35**, 169 (1980).

[17] P. Hoogerbrugge and J. Koelman, Europhys. Lett **19**, 155 (1992).

[18] P. Español and P. Warren, Europhys. Lett. **30**, 191 (1995).

[19] C. Cotter and S. Reich, Europhys. Lett. **64**, 723 (2003).

[20] P. Español, Phys. Rev. E **53**, 1572 (1996).

[21] L. Rey-Bellet and L. Thomas, Commun. Math. Phys. **225**, 305 (2002).

[22] R. Kupferman, A. Stuart, J. Terry, and P. Tupper, Stoch. Dyn. **2**, 533 (2002).

[23] L. Verlet, Phys. Lett. **159**, 98 (1967).

[24] E. Hairer, C. Lubich, and G. Wanner, Acta Numerica **12**, 399 (2003).

[25] J. Sanz-Serna and M. Calvo, *Numerical Hamiltonian Problems* (Chapman & Hall, London, 1994).

[26] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration* (Springer-Verlag, Berlin Heidelberg, 2002).

[27] S. Kuksin and J. Pöschel, in *Seminar on Dynamical Systems (St. Petersburg, 1991)*, edited by S. Kuksin, V. Lazutkin, and J. Pöschel (Birkhäuser Verlag, Basel, 1994), vol. 12 of *Progr. Nonlinear Differential Equations Appl.*, pp. 96–116.

[28] P. Moan, Nonlinearity **17**, 67 (2004).

[29] A. Neishtadt, J. Appl. Math. Mech. **48**, 133 (1984).

[30] R. Skeel and D. Hardy, SIAM J. Sci. Comput. **23**, 1172 (2001).

[31] B. Moore and S. Reich, Numer. Math. **95**, 625 (2003).

[32] G. Benetin and A. Giorgilli, J. Stat. Phys. **74**, 1117 (1994).

[33] E. Hairer and C. Lubich, Numer. Math. **76**, 441 (1997).

[34] S. Reich, SIAM J. Numer. Anal. **36**, 475 (1999).

[35] B. Forrest and U. Suter, J. Chem. Phys. **101**, 2616 (1994).

[36] S. Duane, A. Kennedy, B. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).

[37] S. Hampton and J. Izaguirre, Tech. Rep., Department of Computer Science and Engineering, University of Notre Dame (2003).

[38] J. Touma and J. Wisdom, Astron. J. **107**, 1189 (1994).

[39] S. Reich, in *Integration Algorithms and Classical Mechanics*, edited by J. Mardsen, G. Patrick, and W. Shadwick (Amer. Math. Soc., 1996), vol. 10 of *Fields Inst. Com.*, pp. 181–192.

[40] A. Dullweber, B. Leimkuhler, and R. McLachlan, J. Chem. Phys. **107**, 5840 (1997).

[41] T. Schlick, M. Madziuk, R. Skeel, and K. Srinivas, J. Comput. Phys. **139**, 1 (1998).

[42] H. Andersen, J. Comput. Phys. **52**, 24 (1983).

[43] J. Ryckaert, G. Ciccotti, and H. Berendsen, J. Comput. Phys. **23**, 327 (1977).

[44] B. Leimkuhler and R. Skeel, J. Comput. Phys. **112**, 117 (1994).

[45] S. Reich, SIAM J. Numer. Anal. **33**, 475 (1996).

[46] E. Barth, K. Kuczera, B. Leimkuhler, and R. Skeel, J. Comput. Chem. **16**, 1192 (1995).

[47] M. Fixman, Proc. Nat. Acad. Sci. **71**, 3050 (1974).

[48] W. van Gunsteren and M. Karplus, Macromolecules **15**, 1528 (1982).

[49] A. Kol, B. Laird, and B. Leimkuhler, J. Chem. Phys. **107**, 2580 (1997).

[50] R. McLachlan and C. Scovel, J. Nonlinear Sci. **5**, 233 (1995).

[51] J. Zhou, S. Reich, and B. Brooks, J. Chem. Phys. **112**, 7919 (2000).

[52] B. Hess, H. Saint-Martin, and H. Berendsen, J. Chem. Phys. **116**, 9602 (2002).

[53] W. Streett, D. Tildesley, and G. Saville, Mol. Phys. **35**, 639 (1978).

[54] J. Finney, J. Comput. Phys. **28**, 92 (1978).

[55] M. Tuckerman, B. Berne, and G. Martyna, J. Chem. Phys. **97**, 1990 (1992).

[56] H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten, Mol. Sim. **6**, 121 (1991).

[57] R. Skeel, in *The Graduate Student's Guide to Numerical Analysis*, edited by M. Ainsworth, J. Levesley, and M. Marletta (Springer-Verlag, Berlin, 1999), vol. 4 of *SSCM*, pp. 119–176.

[58] B. Berne, in *Computational Molecular Dynamics*, edited by P. Deuflhard, J. Hermans, B. Leimkuhler, A. Mark, S. Reich, and R. Skeel (Springer-Verlag, Berlin, 1999), vol. 4 of *Lect. Notes Comput. Sci. Eng.*, pp. 281–296.

[59] J. Biesiadecki and R. Skeel, J. Comput. Phys. **109**, 318 (1993).

[60] Q. Ma, J. Izaguirre, and R. Skeel, SIAM J. Sci. Comput. **24**, 1951 (2003).

[61] B. García-Archilla, J. Sanz-Serna, and R. Skeel, SIAM J. Sci. Comput. **20**, 930 (1998).

[62] J. Izaguirre, S. Reich, and R. Skeel, J. Chem. Phys. **110**, 9853 (1999).

[63] J. Izaguirre, D. Catarello, J. Wozniak, and R. Skeel, J. Chem. Phys. **114**, 2090 (2001).

[64] D. Gromov and J. de Pablo, J. Chem. Phys. **103**, 8247 (1995).

[65] B. Leimkuhler and S. Reich, J. Comput. Phys. **171**, 95 (2001).

[66] A. Brünger, C. Brooks, and M. Karplus, Chem. Phys. Lett. **105** (1984).

[67] D. Higham, SIAM Review **43**, 525 (2001).

[68] R. Pastor, B. Brooks, and A. Szabo, Mol. Phys. **65**, 1409 (1988).

[69] W. V. Gunsteren and H. Berendsen, Mol. Phys. **45**, 637 (1982).

[70] M. Allen, Mol. Phys. **47**, 599 (1982).

[71] R. Skeel and J. Izaguirre, Mol. Phys. **100**, 3885 (2002).

[72] W. Wang and R. Skeel, Mol. Phys. **101**, 2149 (2003).

[73] I. Pagonabarraga, M. Hagen, and D. Frenkel, Europhys. Lett. **42**, 377 (1998).

[74] R. Groot and P. Warren, J. Chem. Phys. **107**, 4423 (1997).

[75] G. Besold, I. Vattulainen, M. Karttunen, and J. Polson, Phys. Rev. E **62**, R7611 (2000).

[76] T. Shardlow, SIAM J. Sci. Comput. **24**, 1267 (2003).

[77] E. Barth and T. Schlick, J. Chem. Phys. **109**, 1617 (1998).

[78] A. Sandu and T. Schlick, J. Comput. Phys. **151**, 74 (1999).

[79] Q. Ma and J. Izaguirre, Multiscale Modeling and Simulation **2**, 1 (2003).

[80] E. Eich-Soellner and C. Führer, *Numerical Methods in Multibody Dynamics* (B.G. Teuber, Stuttgart, 1998).

[81] U. Ascher, H. Chin, L. Petzold, and S. Reich, Mech. Struct. & Mach **23**, 135 (1995).

[82] B. Hess, H. Bekker, H. Berendsen, and J. Fraaije, J. Comput. Chem. **18**, 1463 (1997).

[83] S. Nosé, J. Phys. Soc. Japan **70**, 75 (2001).

[84] P. Español, Phys. Rev. E **52**, 1734 (1995).

[85] S. Trofimov, E. Nies, and M. Michels, J. Chem. Phys. **117**, 9383 (2002).

[86] P. Español and M. Revenga, Phys. Rev. E **67**, 026705 (2003).

[87] P. Español, in *Trends in nanoscale mechanics: Analysis of nanostructure materials and multi-scale modeling*, edited by V. Harik and M. Salas (Kluwer, 2003), pp. 1–23.

[88] L. Lucy, Astron. J. **82**, 1013 (1977).

[89] R. Gingold and J. Monaghan, Mon. Not. R. Astr. Soc. **181**, 375 (1977).

[90] R. Salmon, J. Fluid Mech. **132**, 431 (1983).

[91] R. Hockney and J. Eastwood, *Computer Simulations Using Particles* (Institute of Physics Publisher,

Bristol and Philadelphia, 1988).

[92] C. Birdsall and A. Langdon, *Plasma Physics via Computer Simulations* (McGraw-Hill, New York, 1981).

[93] A. Langdon, J. Comput. Phys. **12**, 247 (1973).

[94] R. Salmon, *Lectures on Geophysical Fluid Dynamics* (Oxford University Press, Oxford, 1999).

[95] J. Frank, G. Gottwald, and S. Reich, in *Meshfree Methods for Partial Differential Equations*, edited by M. Griebel and M. Schweitzer (Springer-Verlag, Berlin Heidelberg, 2002), vol. 26 of *Lect. Notes Comput. Sci. Eng.*, pp. 131–142.

[96] J. Frank and S. Reich, BIT **43**, 40 (2003).

[97] C. Cotter, Ph.D. thesis, Department of Mathematics, Imperial College London, London, SW7 2AZ (2004).

[98] S. Succi, I. Karlin, and H. Chen, Reviews of Modern Physics **74**, 1203 (2002).

[99] J. Frank and S. Reich, Atmospheric Science Letters (2004).

[100] D. Andrews, *An Introduction to Atmospheric Physics* (Cambridge University Press, Cambridge, 2000).

[101] R. de Vogelaere, Tech. Rep. 4, Dept. Math. Univ. of Notre Dame (1956).

Fig. x**1** (should go into Example of Section §2.3)

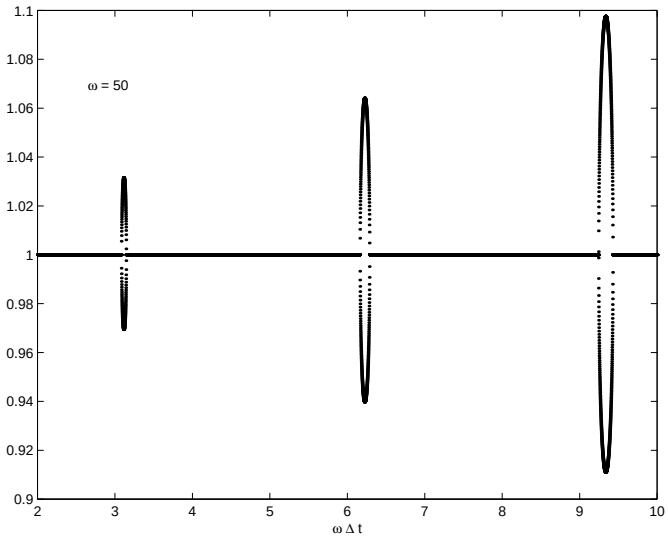

FIG. 1: Modulus of eigenvalues of numerical propagator $\mathbf{M}_{\Delta t}$ as a function of $\omega\Delta t$.

Fig. x**2** (should go into Section §3.4)

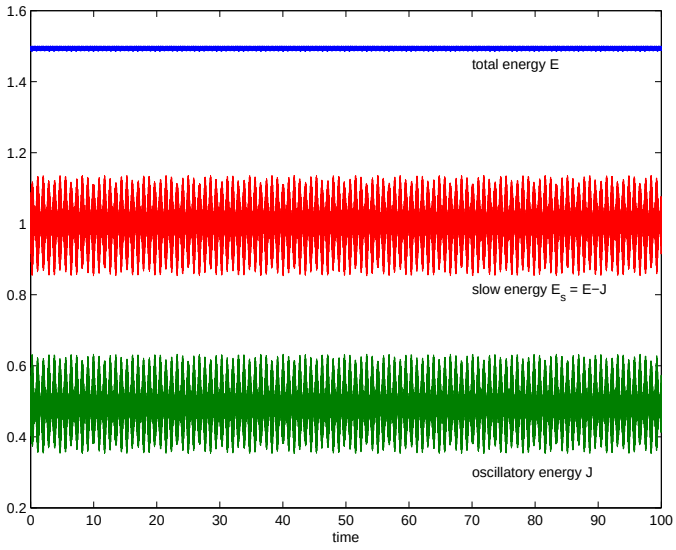

FIG. 2: Total energy $E(t)$, fast oscillatory energy $J(t)$, and slow rotational energy $E_s(t) = E(t) - J(t)$ as a function of time for the conserative stiff pendulum system.

Fig. x**3** (should go into Section §3.4)

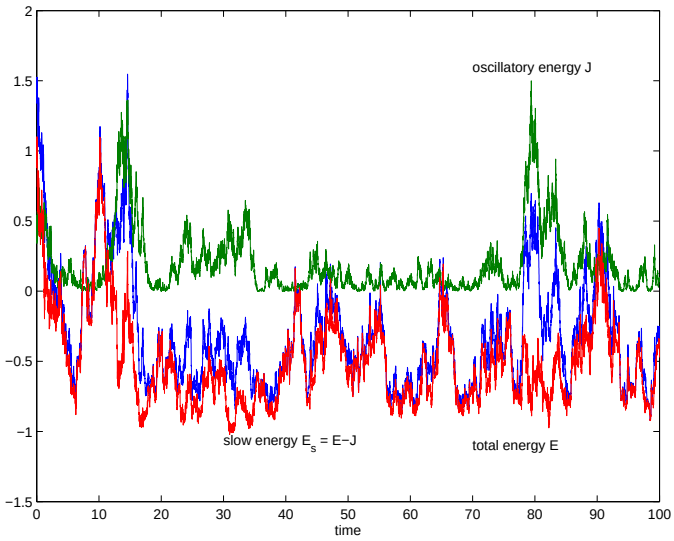

FIG. 3: Total energy $E(t)$, fast oscillatory energy $J(t)$, and slow rotational energy $E_s(t) = E(t) - J(t)$ as a function of time for the Langevin model of the stiff pendulum system.

Fig. x**4** (should go into Section §3.4)

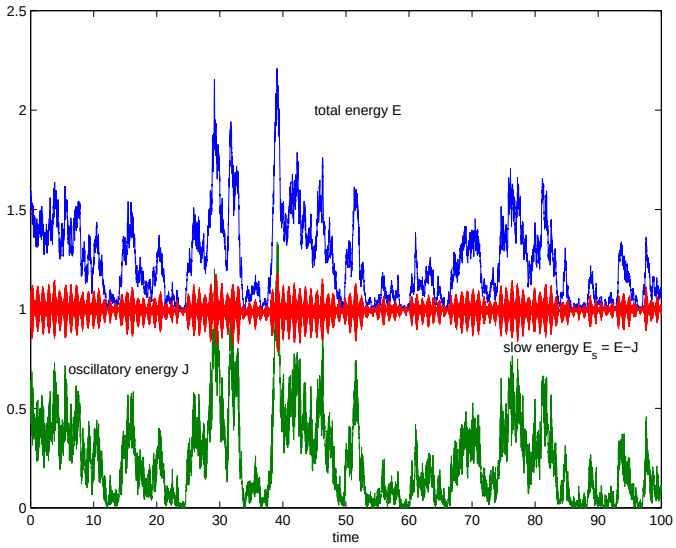

FIG. 4: Total energy $E(t)$, fast oscillatory energy $J(t)$, and slow rotational energy $E_s(t) = E(t) - J(t)$ as a function of time for the DPD stiff spring pendulum model. Note the approximate conservation of the slow (rotational) energy E_s.

Fig. x**5** (should go into Section §3.4)

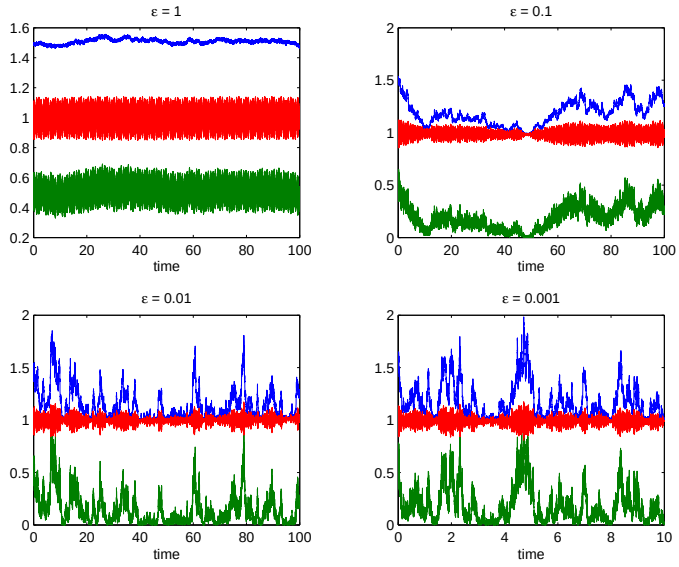

FIG. 5: Total energy $E(t)$, fast oscillatory energy $J(t)$, and slow rotational energy $E_s(t) = E(t) - J(t)$ as a function of time for several different decay rates $\alpha = 1/\varepsilon$ of the auto-covariance matrix in the extended DPD model.