

# Introduction to Molecular Dynamics Simulation

Michael P. Allen

published in

*Computational Soft Matter: From Synthetic Polymers to Proteins,  
Lecture Notes,  
Norbert Attig, Kurt Binder, Helmut Grubmüller, Kurt Kremer (Eds.),  
John von Neumann Institute for Computing, Jülich,  
NIC Series, Vol. 23, ISBN 3-00-012641-4, pp. 1-28, 2004.*

© 2004 by John von Neumann Institute for Computing

Permission to make digital or hard copies of portions of this work for personal or classroom use is granted provided that the copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise requires prior specific permission by the publisher mentioned above.

<http://www.fz-juelich.de/nic-series/volume23>



# Introduction to Molecular Dynamics Simulation

Michael P. Allen

Centre for Scientific Computing and Department of Physics,  
University of Warwick, Coventry CV4 7AL, United Kingdom  
*E-mail: m.p.allen@warwick.ac.uk*

In this chapter a summary is given of the key ingredients necessary to carry out a molecular dynamics simulation, with particular emphasis on macromolecular systems. We discuss the form of the intermolecular potential for molecules composed of atoms, and of non-spherical sub-units, giving examples of how to compute the forces and torques. We also describe some of the MD algorithms in current use. Finally, we briefly refer to the factors that influence the size of systems, and length of runs, that are needed to calculate statistical properties.

## 1 The Aims of Molecular Dynamics

We carry out computer simulations in the hope of understanding the properties of assemblies of molecules in terms of their structure and the microscopic interactions between them. This serves as a complement to conventional experiments, enabling us to learn something new, something that cannot be found out in other ways. The two main families of simulation technique are molecular dynamics (MD) and Monte Carlo (MC); additionally, there is a whole range of hybrid techniques which combine features from both. In this lecture we shall concentrate on MD. The obvious advantage of MD over MC is that it gives a route to dynamical properties of the system: transport coefficients, time-dependent responses to perturbations, rheological properties and spectra.

Computer simulations act as a bridge (see Fig. 1) between microscopic length and time scales and the macroscopic world of the laboratory: we provide a guess at the interactions between molecules, and obtain ‘exact’ predictions of bulk properties. The predictions are ‘exact’ in the sense that they can be made as accurate as we like, subject to the limitations imposed by our computer budget. At the same time, the hidden detail behind bulk measurements can be revealed. An example is the link between the diffusion coefficient and velocity autocorrelation function (the former easy to measure experimentally, the latter much harder). Simulations act as a bridge in another sense: between theory and experiment. We may test a theory by conducting a simulation using the same model. We may test the model by comparing with experimental results. We may also carry out simulations on the computer that are difficult or impossible in the laboratory (for example, working at extremes of temperature or pressure).

Ultimately we may want to make direct comparisons with experimental measurements made on specific materials, in which case a good model of molecular interactions is essential. The aim of so-called *ab initio* molecular dynamics is to reduce the amount of fitting and guesswork in this process to a minimum. On the other hand, we may be interested in phenomena of a rather generic nature, or we may simply want to discriminate between good and bad theories. When it comes to aims of this kind, it is not necessary to have a perfectly realistic molecular model; one that contains the essential physics may be quite suitable.

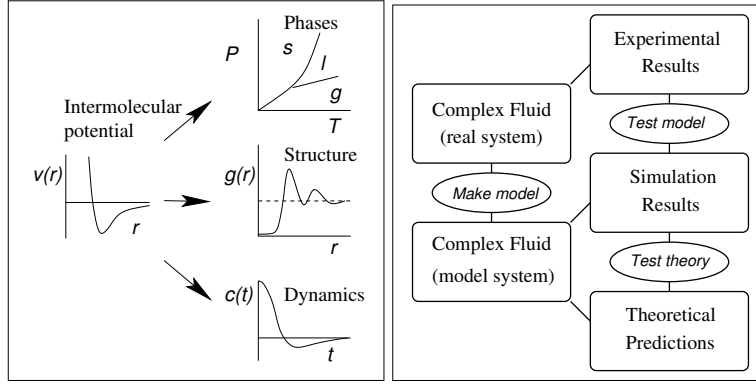


Figure 1. Simulations as a bridge between (a) microscopic and macroscopic; (b) theory and experiment.

## 2 Molecular Interactions

Molecular dynamics simulation consists of the numerical, step-by-step, solution of the classical equations of motion, which for a simple atomic system may be written

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i \quad \mathbf{f}_i = -\frac{\partial}{\partial \mathbf{r}_i} \mathcal{U} \quad (1)$$

For this purpose we need to be able to calculate the forces  $\mathbf{f}_i$  acting on the atoms, and these are usually derived from a potential energy  $\mathcal{U}(\mathbf{r}^N)$ , where  $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  represents the complete set of  $3N$  atomic coordinates. In this section we focus on this function  $\mathcal{U}(\mathbf{r}^N)$ , restricting ourselves to an atomic description for simplicity. (In simulating soft condensed matter systems, we sometimes wish to consider non-spherical rigid units which have rotational degrees of freedom: rotational equations of motion and interaction potentials will be considered in section 5).

### 2.1 Non-bonded Interactions

The part of the potential energy  $\mathcal{U}_{\text{non-bonded}}$  representing non-bonded interactions between atoms is traditionally split into 1-body, 2-body, 3-body ... terms:

$$\mathcal{U}_{\text{non-bonded}}(\mathbf{r}^N) = \sum_i u(\mathbf{r}_i) + \sum_i \sum_{j>i} v(\mathbf{r}_i, \mathbf{r}_j) + \dots \quad (2)$$

The  $u(\mathbf{r})$  term represents an externally applied potential field or the effects of the container walls; it is usually dropped for fully periodic simulations of bulk systems. Also, it is usual to concentrate on the pair potential  $v(\mathbf{r}_i, \mathbf{r}_j) = v(r_{ij})$  and neglect three-body (and higher order) interactions. There is an extensive literature on the way these potentials are determined experimentally, or modelled theoretically<sup>1-4</sup>.

In some simulations of complex fluids, it is sufficient to use the simplest models that faithfully represent the essential physics. In this chapter we shall concentrate on continuous, differentiable pair-potentials (although discontinuous potentials such as hard spheres

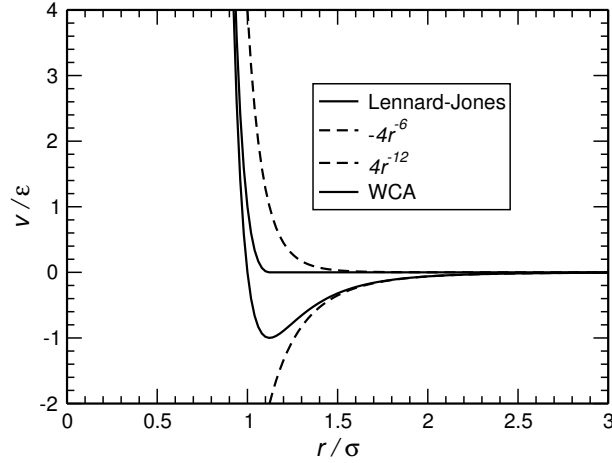


Figure 2. Lennard-Jones pair potential showing the  $r^{-12}$  and  $r^{-6}$  contributions. Also shown is the WCA shifted repulsive part of the potential.

and spheroids have also played a role, see e.g. 5). The Lennard-Jones potential is the most commonly used form:

$$v^{\text{LJ}}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right]. \quad (3)$$

with two parameters:  $\sigma$ , the diameter, and  $\epsilon$ , the well depth. This potential was used, for instance, in the earliest studies of the properties of liquid argon<sup>6,7</sup> and is illustrated in Fig. 2. For applications in which attractive interactions are of less concern than the excluded volume effects which dictate molecular packing, the potential may be truncated at the position of its minimum, and shifted upwards to give what is usually termed the WCA model<sup>8</sup>. If electrostatic charges are present, we add the appropriate Coulomb potentials

$$v^{\text{Coulomb}}(r) = \frac{Q_1 Q_2}{4\pi\epsilon_0 r}, \quad (4)$$

where  $Q_1, Q_2$  are the charges and  $\epsilon_0$  is the permittivity of free space. The correct handling of long-range forces in a simulation is an essential aspect of polyelectrolyte simulations, which will be the subject of the later chapter of Holm<sup>9</sup>.

## 2.2 Bonding Potentials

For molecular systems, we simply build the molecules out of site-site potentials of the form of Eq. (3) or similar. Typically, a single-molecule quantum-chemical calculation may be used to estimate the electron density throughout the molecule, which may then be modelled by a distribution of partial charges via Eq. (4), or more accurately by a distribution of electrostatic multipoles<sup>4,10</sup>. For molecules we must also consider the intramolecular bonding interactions. The simplest molecular model will include terms of the following kind:

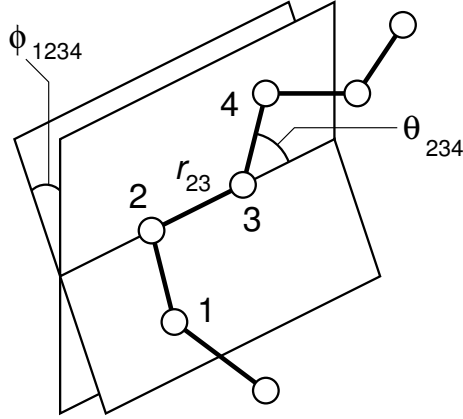


Figure 3. Geometry of a simple chain molecule, illustrating the definition of interatomic distance  $r_{23}$ , bend angle  $\theta_{234}$ , and torsion angle  $\phi_{1234}$ .

$$\mathcal{U}_{\text{intramolecular}} = \frac{1}{2} \sum_{\text{bonds}} k_{ij}^r (r_{ij} - r_{\text{eq}})^2 \quad (5a)$$

$$+ \frac{1}{2} \sum_{\text{bend angles}} k_{ijk}^\theta (\theta_{ijk} - \theta_{\text{eq}})^2 \quad (5b)$$

$$+ \frac{1}{2} \sum_{\text{torsion angles}} \sum_m k_{ijkl}^{\phi, m} (1 + \cos(m\phi_{ijkl} - \gamma_m)) \quad (5c)$$

The geometry is illustrated in Fig. 3. The “bonds” will typically involve the separation  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  between adjacent pairs of atoms in a molecular framework, and we assume in Eq. (5a) a harmonic form with specified equilibrium separation, although this is not the only possibility. The “bend angles”  $\theta_{ijk}$  are between successive bond vectors such as  $\mathbf{r}_i - \mathbf{r}_j$  and  $\mathbf{r}_j - \mathbf{r}_k$ , and therefore involve three atom coordinates:

$$\cos \theta_{ijk} = \hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{r}}_{jk} = (\mathbf{r}_{ij} \cdot \mathbf{r}_{ij})^{-1/2} (\mathbf{r}_{jk} \cdot \mathbf{r}_{jk})^{-1/2} (\mathbf{r}_{ij} \cdot \mathbf{r}_{jk})$$

where  $\hat{\mathbf{r}} = \mathbf{r}/r$ . Usually this bending term is taken to be quadratic in the angular displacement from the equilibrium value, as in Eq. (5b), although periodic functions are also used. The “torsion angles”  $\phi_{ijkl}$  are defined in terms of three connected bonds, hence four atomic coordinates:

$$\cos \phi_{ijkl} = -\hat{\mathbf{n}}_{ijk} \cdot \hat{\mathbf{n}}_{jkl}, \quad \text{where} \quad \mathbf{n}_{ijk} = \mathbf{r}_{ij} \times \mathbf{r}_{jk}, \quad \mathbf{n}_{jkl} = \mathbf{r}_{jk} \times \mathbf{r}_{kl},$$

and  $\hat{\mathbf{n}} = \mathbf{n}/n$ , the unit normal to the plane defined by each pair of bonds. Usually the torsional potential involves an expansion in periodic functions of order  $m = 1, 2, \dots$ , Eq. (5c).

A simulation package force-field will specify the precise form of Eq. (5), and the various strength parameters  $k$  and other constants therein. Actually, Eq. (5) is a considerable oversimplification. Molecular mechanics force-fields, aimed at accurately predicting

structures and properties, will include many cross-terms (e.g. stretch-bend): MM3<sup>11–13</sup> and MM4<sup>14–16</sup> are examples. Quantum mechanical calculations may give a guide to the “best” molecular force-field; also comparison of simulation results with thermophysical properties and vibration frequencies is invaluable in force-field development and refinement. A separate family of force fields, such as AMBER<sup>17,18</sup>, CHARMM<sup>19</sup> and OPLS<sup>20</sup> are geared more to larger molecules (proteins, polymers) in condensed phases; their functional form is simpler, closer to that of Eq. (5), and their parameters are typically determined by quantum chemical calculations combined with thermophysical and phase coexistence data. This field is too broad to be reviewed here; several molecular modelling texts<sup>21–23</sup> (albeit targeted at biological applications) should be consulted by the interested reader. The modelling

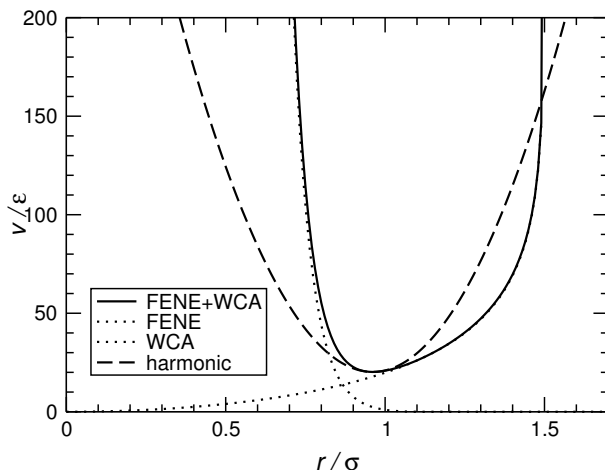


Figure 4. The FENE+WCA potential, with its separate FENE (attractive) and WCA (repulsive) components, between bonded atoms in a coarse-grained polymer chain. Also shown is the equivalent harmonic potential. Unlike the harmonic spring, the FENE potential cannot be extended beyond a specified limit, here  $R_0 = 1.5\sigma$ . For more details see Ref. 24.

of long chain molecules will be of particular interest to us, especially as an illustration of the scope for progressively simplifying and “coarse-graining” the potential model. Various explicit-atom potentials have been devised for the  $n$ -alkanes<sup>25</sup>. More approximate potentials have also been constructed<sup>26–28</sup> in which the  $\text{CH}_2$  and  $\text{CH}_3$  units are represented by single “united atoms”. These potentials are typically less accurate and less transferable than the explicit-atom potentials, but significantly less expensive; comparisons have been made between the two approaches<sup>29</sup>. For more complicated molecules this approach may need to be modified. In the liquid crystal field, for instance, a compromise has been suggested<sup>30</sup>: use the united-atom approach for hydrocarbon chains, but model phenyl ring hydrogens explicitly.

In polymer simulations, there is frequently a need to economize further and coarse-grain the interactions more dramatically: significant progress has been made in recent years in approaching this problem systematically<sup>31,32</sup>. Finally, the most fundamental properties, such as the entanglement length in a polymer melt<sup>33</sup>, may be investigated using a

simple chain of pseudo-atoms or beads (modelled using the WCA potential of Fig. 2, and each representing several monomers), joined by an attractive finitely-extensible non-linear elastic (FENE) potential<sup>24</sup> which is illustrated in Fig. 4.

$$v^{\text{FENE}}(r) = \begin{cases} -\frac{1}{2}kR_0^2 \ln(1 - (r/R_0)^2) & r < R_0 \\ \infty & r \geq R_0 \end{cases} \quad (6)$$

The key feature of this potential is that it cannot be extended beyond  $r = R_0$ , ensuring (for suitable choices of the parameters  $k$  and  $R_0$ ) that polymer chains cannot move through one another.

### 2.3 Force Calculation

Having specified the potential energy function  $\mathcal{U}(\mathbf{r}^N)$ , the next step is to calculate the atomic forces

$$\mathbf{f}_i = -\frac{\partial}{\partial \mathbf{r}_i} \mathcal{U}(\mathbf{r}^N)$$

For site-site potentials this is a simple exercise. For the intramolecular part of the potential, it is a little more involved, but still a relatively straightforward application of the chain rule. Examples of how to do it are given in appendix C of Ref. 34. As a simple illustration, consider one of the bending potential terms for the polymer of Fig. 3, supposing that it may be written

$$v = -k \cos \theta_{234} = -k(\mathbf{r}_{23} \cdot \mathbf{r}_{23})^{-1/2}(\mathbf{r}_{34} \cdot \mathbf{r}_{34})^{-1/2}(\mathbf{r}_{23} \cdot \mathbf{r}_{34})$$

This will contribute to the forces on all three atoms. To calculate these, we need:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}_2}(\mathbf{r}_{23} \cdot \mathbf{r}_{34}) &= \mathbf{r}_{34} & \frac{\partial}{\partial \mathbf{r}_3}(\mathbf{r}_{23} \cdot \mathbf{r}_{34}) &= \mathbf{r}_{23} - \mathbf{r}_{34} & \frac{\partial}{\partial \mathbf{r}_4}(\mathbf{r}_{23} \cdot \mathbf{r}_{34}) &= -\mathbf{r}_{23} \\ \frac{\partial}{\partial \mathbf{r}_2}(\mathbf{r}_{23} \cdot \mathbf{r}_{23}) &= 2\mathbf{r}_{23} & \frac{\partial}{\partial \mathbf{r}_3}(\mathbf{r}_{23} \cdot \mathbf{r}_{23}) &= -2\mathbf{r}_{23} & \frac{\partial}{\partial \mathbf{r}_4}(\mathbf{r}_{23} \cdot \mathbf{r}_{23}) &= 0 \\ \frac{\partial}{\partial \mathbf{r}_2}(\mathbf{r}_{34} \cdot \mathbf{r}_{34}) &= 0 & \frac{\partial}{\partial \mathbf{r}_3}(\mathbf{r}_{34} \cdot \mathbf{r}_{34}) &= 2\mathbf{r}_{34} & \frac{\partial}{\partial \mathbf{r}_4}(\mathbf{r}_{34} \cdot \mathbf{r}_{34}) &= -2\mathbf{r}_{34} \end{aligned}$$

and hence

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}_2} \cos \theta_{234} &= r_{23}^{-1} r_{34}^{-1} \left( \mathbf{r}_{34} - \frac{\mathbf{r}_{23} \cdot \mathbf{r}_{34}}{r_{23}^2} \mathbf{r}_{23} \right) \\ \frac{\partial}{\partial \mathbf{r}_3} \cos \theta_{234} &= r_{23}^{-1} r_{34}^{-1} \left( \frac{\mathbf{r}_{23} \cdot \mathbf{r}_{34}}{r_{23}^2} \mathbf{r}_{23} - \frac{\mathbf{r}_{23} \cdot \mathbf{r}_{34}}{r_{34}^2} \mathbf{r}_{34} + \mathbf{r}_{23} - \mathbf{r}_{34} \right) \\ \frac{\partial}{\partial \mathbf{r}_4} \cos \theta_{234} &= r_{23}^{-1} r_{34}^{-1} \left( \frac{\mathbf{r}_{23} \cdot \mathbf{r}_{34}}{r_{34}^2} \mathbf{r}_{34} - \mathbf{r}_{23} \right) \end{aligned}$$

A similar approach applied to the torsional potential gives the forces on all four involved atoms.



### 3 The MD Algorithm

Solving Newton’s equations of motion does not immediately suggest activity at the cutting edge of research. The molecular dynamics algorithm in most common use today may even have been known to Newton<sup>35</sup>. Nonetheless, the last decade has seen a rapid development in our understanding of numerical algorithms; a forthcoming review<sup>36</sup> and a book<sup>37</sup> summarize the present state of the field.

Continuing to discuss, for simplicity, a system composed of atoms with coordinates  $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  and potential energy  $\mathcal{U}(\mathbf{r}^N)$ , we introduce the atomic momenta  $\mathbf{p}^N = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$ , in terms of which the kinetic energy may be written  $\mathcal{K}(\mathbf{p}^N) = \sum_{i=1}^N |\mathbf{p}_i|^2 / 2m_i$ . Then the energy, or hamiltonian, may be written as a sum of kinetic and potential terms  $\mathcal{H} = \mathcal{K} + \mathcal{U}$ . Write the classical equations of motion as

$$\dot{\mathbf{r}}_i = \mathbf{p}_i / m_i \quad \text{and} \quad \dot{\mathbf{p}}_i = \mathbf{f}_i \quad (7)$$

This is a system of coupled ordinary differential equations. Many methods exist to perform step-by-step numerical integration of them. Characteristics of these equations are: (a) they are ‘stiff’, i.e. there may be short and long timescales, and the algorithm must cope with both; (b) calculating the forces is expensive, typically involving a sum over pairs of atoms, and should be performed as infrequently as possible. Also we must bear in mind that the advancement of the coordinates fulfils two functions: (i) accurate calculation of dynamical properties, especially over times as long as typical correlation times  $\tau_a$  of properties  $a$  of interest (we shall define this later); (ii) accurately staying on the constant-energy hypersurface, for much longer times  $\tau_{\text{run}} \gg \tau_a$ , in order to sample the correct ensemble.

To ensure rapid sampling of phase space, we wish to make the timestep as large as possible consistent with these requirements. For these reasons, simulation algorithms have tended to be of *low order* (i.e. they do not involve storing high derivatives of positions, velocities etc.): this allows the time step to be increased as much as possible without jeopardizing energy conservation. It is unrealistic to expect the numerical method to accurately follow the true trajectory for very long times  $\tau_{\text{run}}$ . The ‘ergodic’ and ‘mixing’ properties of classical trajectories, i.e. the fact that nearby trajectories diverge from each other exponentially quickly, make this impossible to achieve.

All these observations tend to favour the Verlet algorithm in one form or another, and we look closely at this in the following section. For historical reasons only, we mention the more general class of predictor-corrector methods which have been optimized for classical mechanical equations<sup>38,39</sup>; further details are available elsewhere<sup>34,40,41</sup>.

#### 3.1 The Verlet Algorithm

There are various, essentially equivalent, versions of the Verlet algorithm, including the original method<sup>7,42</sup> and a ‘leapfrog’ form<sup>43</sup>. Here we concentrate on the ‘velocity Verlet’ algorithm<sup>44</sup>, which may be written

$$\mathbf{p}_i(t + \tfrac{1}{2}\delta t) = \mathbf{p}_i(t) + \tfrac{1}{2}\delta t \mathbf{f}_i(t) \quad (8a)$$

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \delta t \mathbf{p}_i(t + \tfrac{1}{2}\delta t) / m_i \quad (8b)$$

$$\mathbf{p}_i(t + \delta t) = \mathbf{p}_i(t + \tfrac{1}{2}\delta t) + \tfrac{1}{2}\delta t \mathbf{f}_i(t + \delta t) \quad (8c)$$

After step (8b), a force evaluation is carried out, to give  $\mathbf{f}_i(t + \delta t)$  for step (8c). This scheme advances the coordinates and momenta over a timestep  $\delta t$ . A piece of pseudo-code illustrates how this works:

```
do step = 1, nstep
  p = p + 0.5*dt*f
  r = r + dt*p/m
  f = force(r)
  p = p + 0.5*dt*f
enddo
```

As we shall see shortly there is an interesting theoretical derivation of this version of the algorithm. Important features of the Verlet algorithm are: (a) it is *exactly* time reversible; (b) it is *symplectic* (to be discussed shortly); (c) it is *low* order in time, hence permitting long timesteps; (d) it requires just one (expensive) force evaluation per step; (e) it is easy to program.

### 3.2 Constraints

It is quite common practice in classical computer simulations *not* to attempt to represent intramolecular bonds by terms in the potential energy function, because these bonds have very high vibration frequencies (and arguably should be treated in a quantum mechanical way rather than in the classical approximation). Instead, the bonds are treated as being constrained to have fixed length. In classical mechanics, constraints are introduced through the Lagrangian<sup>45</sup> or Hamiltonian<sup>46</sup> formalisms. Given an algebraic relation between two atomic coordinates, for example a fixed bond length  $b$  between atoms 1 and 2, one may write a constraint equation, plus an equation for the time derivative of the constraint

$$\chi(\mathbf{r}_1, \mathbf{r}_2) = (\mathbf{r}_1 - \mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) - b^2 = 0 \quad (9a)$$

$$\dot{\chi}(\mathbf{r}_1, \mathbf{r}_2) = 2(\mathbf{v}_1 - \mathbf{v}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) = 0 \quad (9b)$$

In the Lagrangian formulation, the constraint forces acting on the atoms will enter thus:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i + \Lambda \mathbf{g}_i$$

where  $\Lambda$  is the undetermined multiplier and

$$\mathbf{g}_1 = -\frac{\partial \chi}{\partial \mathbf{r}_1} = -2(\mathbf{r}_1 - \mathbf{r}_2) \quad \mathbf{g}_2 = -\frac{\partial \chi}{\partial \mathbf{r}_2} = 2(\mathbf{r}_1 - \mathbf{r}_2)$$

It is easy to derive an exact expression for the multiplier  $\Lambda$  from the above equations; if several constraints are imposed, a system of equations (one per constraint) is obtained. However, this exact solution is not what we want: in practice, since the equations of motion are only solved approximately, in discrete time steps, the constraints will be increasingly violated as the simulation proceeds. The breakthrough in this area came with the proposal to determine the constraint forces in such a way that the constraints are satisfied *exactly* at the end of each time step<sup>47-49</sup>. For the original Verlet algorithm, this scheme is called SHAKE. The appropriate version of this scheme for the velocity Verlet algorithm is called RATTLE<sup>50</sup>. Formally, we wish to solve the following scheme, in which we combine

$(\mathbf{r}_1, \mathbf{r}_2)$  into  $\mathbf{r}$ ,  $(\mathbf{p}_1, \mathbf{p}_2)$  into  $\mathbf{p}$ , etc. for simplicity:

$$\mathbf{p}(t + \frac{1}{2}\delta t) = \mathbf{p}(t) + \frac{1}{2}\delta t \mathbf{f}(t) + \lambda \mathbf{g}(t)$$

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{p}(t + \frac{1}{2}\delta t)/m$$

$$\text{choosing } \lambda \text{ such that: } 0 = \chi(\mathbf{r}(t + \delta t)) \quad (10a)$$

$$\mathbf{p}(t + \delta t) = \mathbf{p}(t + \frac{1}{2}\delta t) + \frac{1}{2}\delta t \mathbf{f}(t + \delta t) + \mu \mathbf{g}(t + \delta t)$$

$$\text{choosing } \mu \text{ such that: } 0 = \dot{\chi}(\mathbf{r}(t + \delta t), \mathbf{p}(t + \delta t)) \quad (10b)$$

Step (10a) may be implemented by defining *unconstrained* variables

$$\bar{\mathbf{p}}(t + \frac{1}{2}\delta t) = \mathbf{p}(t) + \frac{1}{2}\delta t \mathbf{f}(t), \quad \bar{\mathbf{r}}(t + \delta t) = \mathbf{r}(t) + \delta t \bar{\mathbf{p}}(t + \frac{1}{2}\delta t)/m$$

then solving the nonlinear equation for  $\lambda$

$$\chi(t + \delta t) = \chi(\bar{\mathbf{r}}(t + \delta t) + \lambda \delta t \mathbf{g}(t)/m) = 0$$

and substituting back

$$\mathbf{p}(t + \frac{1}{2}\delta t) = \bar{\mathbf{p}}(t + \frac{1}{2}\delta t) + \lambda \mathbf{g}(t), \quad \mathbf{r}(t + \delta t) = \bar{\mathbf{r}}(t + \delta t) + \delta t \lambda \mathbf{g}(t)/m$$

Step (10b) may be handled by defining

$$\bar{\mathbf{p}}(t + \delta t) = \mathbf{p}(t + \frac{1}{2}\delta t) + \frac{1}{2}\delta t \mathbf{f}(t + \delta t)$$

solving the equation for the second Lagrange multiplier  $\mu$

$$\dot{\chi}(t + \delta t) = \dot{\chi}(\mathbf{r}(t + \delta t), \bar{\mathbf{p}}(t + \delta t) + \mu \mathbf{g}(t + \delta t)) = 0$$

(which is actually linear, since  $\dot{\chi}(\mathbf{r}, \mathbf{p}) = -\mathbf{g}(\mathbf{r}) \cdot \mathbf{p}/m$ ) and substituting back

$$\mathbf{p}(t + \delta t) = \bar{\mathbf{p}}(t + \delta t) + \mu \mathbf{g}(t + \delta t)$$

In pseudo-code this scheme may be written

```
do step = 1, nstep
  p = p + (dt/2)*f
  r = r + dt*p/m
  lambda_g = shake(r)
  p = p + lambda_g
  r = r + dt*lambda_g/m
  f = force(r)
  p = p + (dt/2)*f
  mu_g = rattle(r,p)
  p = p + mu_g
enddo
```

The routine called `shake` here calculates the constraint forces  $\lambda \mathbf{g}_i$  necessary to ensure that the end-of-step positions  $\mathbf{r}_i$  satisfy Eq. (9a). For a system of many constraints, this calculation is usually performed in an iterative fashion, so as to satisfy each constraint in turn until convergence; the original SHAKE algorithm was framed in this way. These constraint forces are incorporated into both the end-of-step positions and the mid-step momenta. The routine called `rattle` calculates a new set of constraint forces  $\mu \mathbf{g}_i$  to ensure that the end-of-step momenta satisfy Eq. (9b). This also may be carried out iteratively.

It is important to realize that a simulation of a system with rigidly constrained bond lengths, is not equivalent to a simulation with, for example, harmonic springs representing the bonds, even in the limit of very strong springs. A subtle, but crucial, difference lies in the distribution function for the other coordinates. If we obtain the configurational distribution function by integrating over the momenta, the difference arises because in one case a set of momenta is set to zero, and not integrated, while in the other an integration is performed, which may lead to an extra term depending on particle coordinates. This is frequently called the ‘metric tensor problem’; it is explained in more detail in the references<sup>34,51</sup>, and there are well-established ways of determining when the difference is likely to be significant<sup>52</sup> and how to handle it, if necessary<sup>53</sup>. Constraints also find an application in the study of rare events<sup>51</sup>, or for convenience when it is desired to fix, for example, the director in a liquid crystal simulation<sup>54</sup>.

An alternative to constraints, is to retain the intramolecular bond potentials and use a multiple time step approach to handle the fast degrees of freedom. We discuss this shortly.

### 3.3 Periodic Boundary Conditions

Small sample size means that, unless surface effects are of particular interest, periodic boundary conditions need to be used. Consider 1000 atoms arranged in a  $10 \times 10 \times 10$  cube. Nearly half the atoms are on the outer faces, and these will have a large effect on the measured properties. Even for  $10^6 = 100^3$  atoms, the surface atoms amount to 6% of the total, which is still nontrivial. Surrounding the cube with replicas of itself takes care of this problem. Provided the potential range is not too long, we can adopt the *minimum image convention* that each atom interacts with the nearest atom or image in the periodic array. In the course of the simulation, if an atom leaves the basic simulation box, attention can be switched to the incoming image. This is shown in Figure 5. Of course, it is important to bear in mind the imposed artificial periodicity when considering properties which are influenced by long-range correlations. Special attention must be paid to the case where the potential range is not short: for example for charged and dipolar systems.

### 3.4 Neighbour Lists

Computing the non-bonded contribution to the interatomic forces in an MD simulation involves, in principle, a large number of pairwise calculations: we consider each atom  $i$  and loop over all other atoms  $j$  to calculate the minimum image separations  $r_{ij}$ . Let us assume that the interaction potentials are of short range,  $v(r_{ij}) = 0$  if  $r_{ij} > r_{\text{cut}}$ , the potential cutoff. In this case, the program skips the force calculation, avoiding expensive calculations, and considers the next candidate  $j$ . Nonetheless, the time to examine all pair separations is proportional to the number of distinct pairs,  $\frac{1}{2}N(N-1)$  in an  $N$ -atom system, and for every pair one must compute at least  $r_{ij}^2$ ; this still consumes a lot of time.

Some economies result from the use of lists of nearby pairs of atoms. Verlet<sup>7</sup> suggested such a technique for improving the speed of a program. The potential cutoff sphere, of radius  $r_{\text{cut}}$ , around a particular atom is surrounded by a ‘skin’, to give a larger sphere of radius  $r_{\text{list}}$  as shown in Figure 6. At the first step in a simulation, a list is constructed of all the neighbours of each atom, for which the pair separation is within  $r_{\text{list}}$ . Over the next few MD time steps, only pairs appearing in the list are checked in the force routine. From

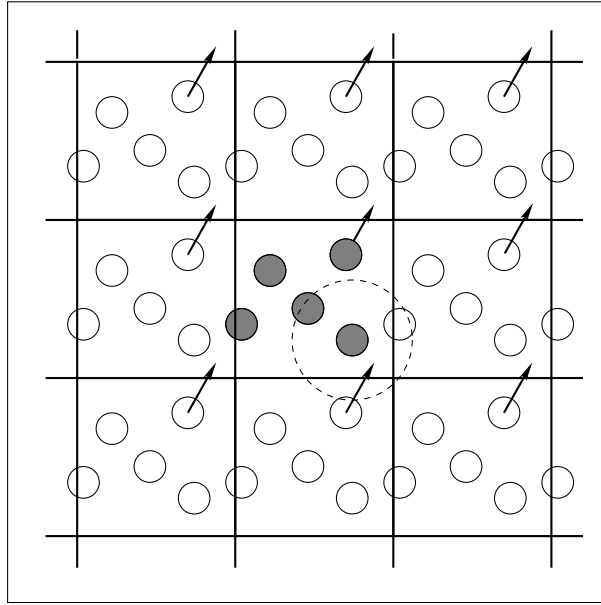


Figure 5. Periodic boundary conditions. As a particle moves out of the simulation box, an image particle moves in to replace it. In calculating particle interactions within the cutoff range, both real and image neighbours are included.

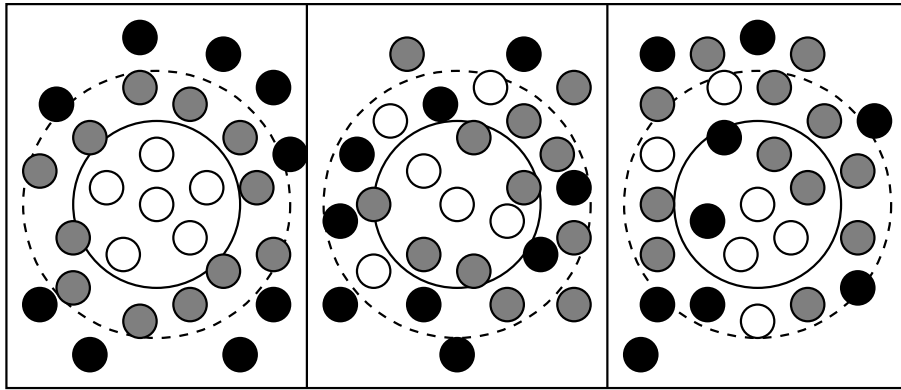


Figure 6. The Verlet list on its construction, later, and too late. The potential cutoff range (solid circle), and the list range (dashed circle), are indicated. The list must be reconstructed before particles originally outside the list range (black) have penetrated the potential cutoff sphere.

time to time the list is reconstructed: it is important to do this before any unlisted pairs have crossed the safety zone and come within interaction range. It is possible to trigger the list reconstruction automatically, if a record is kept of the distance travelled by each atom since the last update. The choice of list cutoff distance  $r_{\text{list}}$  is a compromise: larger lists

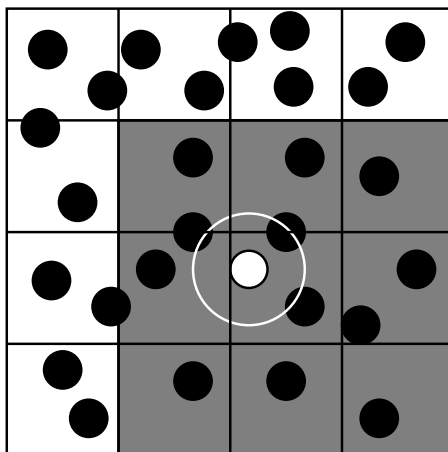


Figure 7. The cell structure. The potential cutoff range is indicated. In searching for neighbours of an atom, it is only necessary to examine the atom's own cell, and its nearest-neighbour cells (shaded).

will need to be reconstructed less frequently, but will not give as much of a saving on cpu time as smaller lists. This choice can easily be made by experimentation.

For larger systems ( $N \geq 1000$  or so, depending on the potential range) another technique becomes preferable. The cubic simulation box (extension to non-cubic cases is possible) is divided into a regular lattice of  $n_{\text{cell}} \times n_{\text{cell}} \times n_{\text{cell}}$  cells; see Figure 7. These cells are chosen so that the side of the cell  $\ell_{\text{cell}} = L/n_{\text{cell}}$  is greater than the potential cutoff distance  $r_{\text{cut}}$ . If there is a separate list of atoms in each of those cells, then searching through the neighbours is a rapid process: it is only necessary to look at atoms in the same cell as the atom of interest, and in nearest neighbour cells. The cell structure may be set up and used by the method of linked lists<sup>55,43</sup>. The first part of the method involves sorting all the atoms into their appropriate cells. This sorting is rapid, and may be performed every step. Then, within the force routine, pointers are used to scan through the contents of cells, and calculate pair forces. This approach is very efficient for large systems with short-range forces. A certain amount of unnecessary work is done because the search region is cubic, not (as for the Verlet list) spherical.

## 4 Time Dependence

A knowledge of time-dependent statistical mechanics is important in three general areas of simulation. Firstly, in recent years there have been significant advances in the understanding of molecular dynamics algorithms, which have arisen out of an appreciation of the formal operator approach to classical mechanics. Second, an understanding of equilibrium time correlation functions, their link with dynamical properties, and especially their connection with transport coefficients, is essential in making contact with experiment. Third, the last decade has seen a rapid development of the use of nonequilibrium molecular dynamics, with a better understanding of the formal aspects, particularly the link between the dynamical algorithm, dissipation, chaos, and fractal geometry. Space does not permit a full

description of all these topics here: the interested reader should consult Refs 56,57,51 and references therein.

The Liouville equation dictates how the classical statistical mechanical distribution function  $\varrho(\mathbf{r}^N, \mathbf{p}^N, t)$  evolves in time. From considerations of standard, Hamiltonian, mechanics<sup>45</sup> and the flow of representative systems in an ensemble through a particular region of phase space, it is easy to derive the Liouville equation

$$\frac{\partial \varrho}{\partial t} = - \left\{ \sum_i \dot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} + \dot{\mathbf{p}}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} \right\} \varrho \equiv -i\mathcal{L}\varrho, \quad (11)$$

defining the Liouville operator  $i\mathcal{L}$  as the quantity in braces. Contrast this equation for  $\varrho$  with the time evolution equation for a dynamical variable  $A(\mathbf{r}^N, \mathbf{p}^N)$ , which comes directly from the chain rule applied to Hamilton's equations

$$\dot{A} = \sum_i \dot{\mathbf{r}}_i \cdot \frac{\partial A}{\partial \mathbf{r}_i} + \dot{\mathbf{p}}_i \cdot \frac{\partial A}{\partial \mathbf{p}_i} \equiv i\mathcal{L}A. \quad (12)$$

The formal solutions of the time evolution equations are

$$\varrho(t) = e^{-i\mathcal{L}t} \varrho(0) \quad \text{and} \quad A(t) = e^{i\mathcal{L}t} A(0) \quad (13)$$

where, in either case, the exponential operator is called the *propagator*. A number of manipulations are possible, once this formalism has been established. There are useful analogies both with the Eulerian and Lagrangian pictures of incompressible fluid flow, and with the Heisenberg and Schrödinger pictures of quantum mechanics (see e.g. Ref. 58, Chap. 7, and Ref. 59, Chap. 11). These analogies are particularly useful in formulating the equations of classical response theory<sup>60</sup>, linking transport coefficients with both equilibrium and nonequilibrium simulations<sup>56</sup>.

The Liouville equation applies to any ensemble, equilibrium or not. Equilibrium means that  $\varrho$  should be *stationary*, i.e. that  $\partial \varrho / \partial t = 0$ . In other words, if we look at any phase-space volume element, the rate of incoming state points should equal the rate of outflow. This requires that  $\varrho$  be a function of the constants of the motion, and especially  $\varrho = \varrho(\mathcal{H})$ . Equilibrium also implies  $d\langle A \rangle / dt = 0$  for any  $A$ . The extension of the above equations to nonequilibrium ensembles requires a consideration of entropy production, the method of controlling energy dissipation (thermostatting) and the consequent non-Liouville nature of the time evolution<sup>56</sup>.

#### 4.1 Propagators and the Verlet Algorithm

The velocity Verlet algorithm may be derived by considering a standard approximate decomposition of the Liouville operator which preserves reversibility and is *symplectic* (which implies that volume in phase space is conserved). This approach<sup>61</sup> has had several beneficial consequences.

The Liouville operator of equation (13) may be written<sup>62</sup>

$$e^{i\mathcal{L}t} = \left( e^{i\mathcal{L}\delta t} \right)_{\text{approx}}^{n_{\text{step}}} + \mathcal{O}(n_{\text{step}}\delta t^3)$$

where  $\delta t = t/n_{\text{step}}$  and an approximate propagator, correct at short timesteps  $\delta t \rightarrow 0$ , appears in the parentheses. This is a formal way of stating what we do in molecular dynamics, when we split a long time period  $t$  into a large number  $n_{\text{step}}$  of small timesteps  $\delta t$ ,

using an *approximation* to the true equations of motion over each timestep. It turns out that useful approximations arise from splitting  $iL$  into two parts<sup>61</sup>

$$iL = iL_p + iL_r . \quad (14)$$

where

$$iL_p = \sum_i \dot{\mathbf{p}}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} = \sum_i \mathbf{f}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} \quad (15a)$$

$$iL_r = \sum_i \dot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} = \sum_i m_i^{-1} \mathbf{p}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} . \quad (15b)$$

The following approximation

$$e^{iL\delta t} = e^{(iL_p + iL_r)\delta t} \approx e^{iL_p\delta t/2} e^{iL_r\delta t} e^{iL_p\delta t/2} \quad (16)$$

is asymptotically exact in the limit  $\delta t \rightarrow 0$ . For nonzero  $\delta t$  this is an approximation to  $e^{iL\delta t}$  because in general  $iL_p$  and  $iL_r$  do not commute, but it is still exactly time reversible and symplectic.

Effectively, we are propagating the equation of motion in steps which ignore, in turn, the kinetic part and the potential part of the hamiltonian. A straightforward derivation<sup>51</sup> shows that the effect of each operator on a dynamical variable  $A(\mathbf{r}^N, \mathbf{p}^N)$  is to advance, respectively, the coordinates and the momenta separately:

$$e^{iL_r\delta t} A(\mathbf{r}, \mathbf{p}) = A(\mathbf{r} + m^{-1}\mathbf{p}\delta t, \mathbf{p}) \quad (17a)$$

$$e^{iL_p\delta t} A(\mathbf{r}, \mathbf{p}) = A(\mathbf{r}, \mathbf{p} + \mathbf{f}\delta t) \quad (17b)$$

where, to avoid clutter, we have written  $\mathbf{r}, \mathbf{p}$  for  $\mathbf{r}^N, \mathbf{p}^N$ . It is then easy to see that the three successive steps embodied in equation (16), with the above choice of operators, generate the velocity Verlet algorithm.

Such a formal approach may seem somewhat abstract, but has been invaluable in understanding the excellent performance of Verlet-like algorithms in molecular dynamics, and in extending the range of algorithms available to us. It may be shown that, although the trajectories generated by the above scheme are approximate, and will not conserve the true energy  $\mathcal{H}$ , nonetheless, they do exactly conserve a “pseudo-hamiltonian” or “shadow hamiltonian”  $\bar{\mathcal{H}}$  which differs from the true one by a small amount (vanishing as  $\delta t \rightarrow 0$ ). This means that no drift in the energy will occur: the system will remain on a hypersurface in phase space which is “close” (in the above sense) to the true constant-energy hypersurface. Such a stability property is extremely useful in molecular dynamics, since we wish to sample constant-energy states.

An example may make things clearer. Consider a simple harmonic oscillator<sup>63</sup>, of natural frequency  $\omega$ , representing perhaps an interatomic bond in a diatomic molecule. The equations of motion are

$$\dot{x} = p/m \quad \dot{p} = -m\omega^2 x$$

For these equations, a few lines of algebra shows that the following shadow hamiltonian

$$\bar{\mathcal{H}}(x, p) = \frac{p^2/2m}{1 - (\omega\delta t/2)^2} + \frac{1}{2}m\omega^2 x^2$$



is *exactly* conserved by the velocity Verlet algorithm. In a phase portrait, the simulated system remains on a constant- $\bar{\mathcal{H}}$  ellipse which differs only slightly (for small  $\omega\delta t$ ) from the true constant- $\mathcal{H}$  ellipse. This is illustrated in Fig. 8, where we deliberately choose a high step size  $\delta t = 0.7/\omega$  to accentuate the differences. Note that, even for this value of  $\delta t$ , the energy conservation is very good (the deviation is  $\mathcal{O}(\delta t^2)$ ). At the same time, the frequency of oscillation is shifted from the true value, so the trajectories steadily diverge from each other.

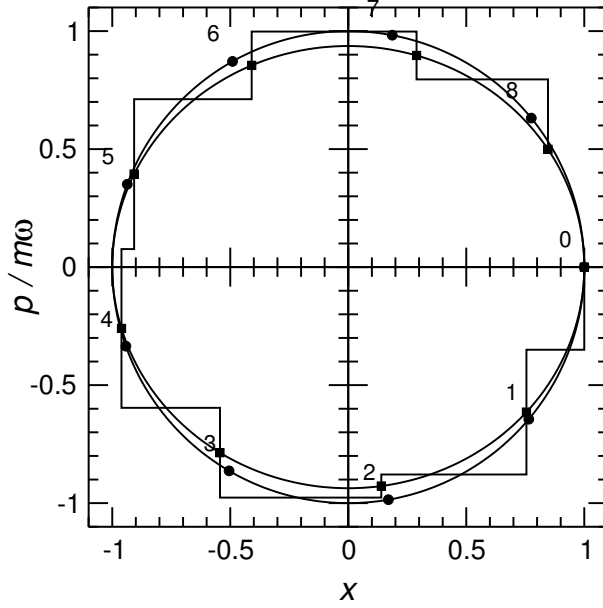


Figure 8. Velocity Verlet algorithm for simple harmonic oscillator with initial conditions  $x(0) = 1, p(0) = 0$ . The outer circle shows the exact trajectory, conserving the true hamiltonian  $\mathcal{H}$ ; the inner ellipse is a contour of constant shadow hamiltonian  $\bar{\mathcal{H}}$  for a (relatively large) timestep  $\delta t = 0.7/\omega$ . The circles show the exact solutions at regular intervals  $\delta t$ ; the squares show the corresponding velocity Verlet phase points, connected by straight sections representing the intermediate steps in the algorithm.

## 4.2 Multiple Timesteps

An important extension of the MD method allows it to tackle systems with multiple time scales: for example, molecules which have very strong internal springs representing the bonds, while interacting externally through softer potentials; molecules having strongly-varying short-range interactions but more smoothly-varying long-range interactions; or perhaps molecules consisting of both heavy and light atoms. A simple MD algorithm will have to adopt a timestep short enough to handle the fastest variables.

An attractive improvement is to handle the fast motions with a shorter timestep<sup>64,61,65</sup>. A time-reversible Verlet-like multiple-timestep algorithm may be generated using the Liouville operator formalism described above. Here we suppose that there are two types of

force in the system: “slow”  $F_i$ , and “fast”  $f_i$ . The momentum satisfies  $\dot{p}_i = f_i + F_i$ . Then we break up the Liouville operator  $iL = iL_p + iL_r$ :

$$iL_p = \sum_i F_i \cdot \frac{\partial}{\partial p_i} \quad (18a)$$

$$iL_r = \sum_i f_i \cdot \frac{\partial}{\partial p_i} \quad (18b)$$

$$iL_r = \sum_i m^{-1} p_i \cdot \frac{\partial}{\partial r_i} \quad (18c)$$

The propagator approximately factorizes

$$e^{iL\Delta t} \approx e^{iL_p\Delta t/2} e^{i(L_p+L_r)\Delta t} e^{iL_p\Delta t/2}$$

where  $\Delta t$  represents a long time step. The middle part is then split again, using the conventional separation, and iterating over small time steps  $\delta t = \Delta t/n_{\text{step}}$ :

$$e^{i(L_p+L_r)\Delta t} \approx \left[ e^{iL_p\delta t/2} e^{iL_r\delta t} e^{iL_p\delta t/2} \right]^{n_{\text{step}}}.$$

So the fast-varying forces must be computed many times at short intervals; the slow-varying forces are used just before and just after this stage, and they only need be calculated once per long timestep.

This translates into a fairly simple algorithm, based closely on the standard velocity Verlet method. Written in a Fortran-like pseudo-code, it is as follows. At the start of the run we calculate both rapidly-varying ( $f$ ) and slowly-varying ( $F$ ) forces, then, in the main loop:

```
do STEP = 1, NSTEP
  p = p + (DT/2)*F
  do step = 1, nstep
    p = p + (dt/2)*f
    r = r + dt*p/m
    f = force(r)
    p = p + (dt/2)*f
  enddo
  F = FORCE(r)
  p = p + (DT/2)*F
enddo
```

The entire simulation run consists of  $NSTEP$  long steps; each step consists of  $nstep$  shorter sub-steps.  $DT$  and  $dt$  are the corresponding timesteps,  $DT = nstep*dt$ .

A particularly fruitful application, which has been incorporated into computer codes such as ORAC<sup>66</sup>, is to split the interatomic force law into a succession of components covering different ranges: the short-range forces change rapidly with time and require a short time step, but advantage can be taken of the much slower time variation of the long-range forces, by using a longer time step and less frequent evaluation for these. Having said this, multiple-time-step algorithms are still under active study<sup>67</sup>, and there is some concern that resonances may occur between the natural frequencies of the system and the

various timesteps used in schemes of this kind<sup>68,69</sup>. Significant efforts have been made in recent years to overcome these problems and achieve significant increases in step size by alternative methods<sup>70–72</sup>. The area remains one of active research<sup>73,36</sup>.

## 5 Rigid Molecule Rotation

In certain applications, particularly in the simulation of liquid crystals, colloidal systems, and polymers, it is advantageous to include non-spherical rigid bodies in the molecular model. This means that we must calculate intermolecular torques as well as forces, and implement the classical dynamical equations for rotational motion.

If the intermolecular forces are expressed as sums of site-site (or atom-atom) terms, the conversion of these into centre-of-mass forces, and torques about the centre of mass, is easily performed. Consider two molecules A and B, centre-of-mass position vectors  $\mathbf{R}_A$ ,  $\mathbf{R}_B$ . Define the intermolecular vector  $\mathbf{R}_{AB} = \mathbf{R}_A - \mathbf{R}_B$ , and suppose that the interaction potential may be expressed

$$v_{AB} = \sum_{i \in A} \sum_{j \in B} v(r_{ij})$$

where  $i$  and  $j$  are atomic sites in the respective molecules. Then we may compute

$$\text{force on } A \text{ due to } B: \quad \mathbf{F}_{AB} = \sum_{i \in A} \sum_{j \in B} \mathbf{f}_{ij} = -\mathbf{F}_{BA}$$

$$\text{torque on } A \text{ due to } B: \quad \mathbf{N}_{AB} = \sum_{i \in A} \sum_{j \in B} \mathbf{r}_{iA} \times \mathbf{f}_{ij}$$

$$\text{torque on } B \text{ due to } A: \quad \mathbf{N}_{BA} = \sum_{i \in A} \sum_{j \in B} \mathbf{r}_{jB} \times \mathbf{f}_{ji}$$

where  $\mathbf{r}_{iA} = \mathbf{r}_i - \mathbf{R}_A$  is the position of site  $i$  relative to the centre of molecule A (and likewise for  $\mathbf{r}_{jB}$ ). Note that  $\mathbf{N}_{AB} \neq -\mathbf{N}_{BA}$  (a common misconception), but that the above equations give directly

$$\mathbf{N}_{AB} + \mathbf{N}_{BA} + \mathbf{R}_{AB} \times \mathbf{F}_{AB} = 0 \quad (19)$$

provided the forces satisfy  $\mathbf{f}_{ij} = -\mathbf{f}_{ji}$  and act along the site-site vectors  $\mathbf{r}_{ij}$ . This is the expression of local angular momentum conservation, which follows directly from the invariance of the potential energy  $v_{AB}$  to a rotation of the coordinate system. (Note, in passing, that, in periodic boundaries, angular momentum is not globally conserved).

There is also a trend to use rigid-body potentials which are defined explicitly in terms of centre-of-mass positions and molecular orientations. An example is the Gay-Berne potential<sup>74</sup>

$$v_{AB}^{\text{GB}}(\mathbf{R}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) = 4\epsilon(\hat{\mathbf{R}}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) [\varrho^{-12} - \varrho^{-6}] \quad (20a)$$

$$\text{with} \quad \varrho = \frac{R - \sigma(\hat{\mathbf{R}}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) + \sigma_0}{\sigma_0} \quad (20b)$$

which depends upon the molecular axis vectors  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{b}}$ , and on the direction  $\hat{\mathbf{R}}$  and magnitude  $R$  of the centre-centre vector  $\mathbf{R}_{AB}$ , which we write  $\mathbf{R}$  here and henceforth. The

parameter  $\sigma_0$  determines the smallest molecular diameter and there are two orientation-dependent quantities in the above shifted-Lennard-Jones form: a diameter  $\sigma(\hat{\mathbf{R}}, \hat{\mathbf{a}}, \hat{\mathbf{b}})$  and an energy  $\varepsilon(\hat{\mathbf{R}}, \hat{\mathbf{a}}, \hat{\mathbf{b}})$ . Each quantity depends in a complicated way (not given here) on parameters characterizing molecular shape and structure. This potential has been extensively used in the study of molecular liquids and liquid crystals<sup>75,76,54,77-79</sup> and will be discussed further in a later chapter<sup>80</sup>. Generalizations to non-uniaxial rigid bodies<sup>81-83</sup> have also been studied: here, the diameter and energy parameters depend on the full orthogonal orientation matrices  $\hat{\mathbf{a}}, \hat{\mathbf{b}}$ , which convert from space-fixed ( $xyz$ ) to molecule-fixed (123) coordinates, and whose rows are the three molecule-fixed orthonormal principal axis vectors  $\hat{\mathbf{a}}_\alpha, \hat{\mathbf{b}}_\beta$ ,  $\alpha, \beta = 1, 2, 3$ .

We go through the following derivation in some detail, as it is seldom presented. A very common case is when the pair potential may be written in the form<sup>84</sup>

$$v_{AB} = v_{AB}(R, \{\hat{\mathbf{a}}_\alpha \cdot \hat{\mathbf{R}}\}, \{\hat{\mathbf{b}}_\beta \cdot \hat{\mathbf{R}}\}, \{\hat{\mathbf{a}}_\alpha \cdot \hat{\mathbf{b}}_\beta\}) \quad (21)$$

i.e. a function of the centre-centre separation  $R$ , and all possible scalar products of the unit vectors  $\hat{\mathbf{R}}, \hat{\mathbf{a}}_\alpha$  and  $\hat{\mathbf{b}}_\beta$ . Using the chain rule, we may write the force on A:

$$\begin{aligned} \mathbf{F}_{AB} &= -\frac{\partial v_{AB}}{\partial \mathbf{R}} = -\frac{\partial v_{AB}}{\partial R} \frac{\partial R}{\partial \mathbf{R}} - \sum_{\hat{\mathbf{e}}=\hat{\mathbf{a}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{R}})} \frac{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{R}})}{\partial \mathbf{R}} \\ &= -\frac{\partial v_{AB}}{\partial R} \hat{\mathbf{R}} - \sum_{\hat{\mathbf{e}}=\hat{\mathbf{a}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{R}})} \frac{\hat{\mathbf{e}} - (\hat{\mathbf{e}} \cdot \hat{\mathbf{R}}) \hat{\mathbf{R}}}{R}. \end{aligned} \quad (22)$$

The sum ranges over all the orientation vectors on both molecules,  $\hat{\mathbf{e}} = \{\hat{\mathbf{a}}_\alpha, \hat{\mathbf{b}}_\beta\}$ . The derivatives of the potential are easily evaluated, assuming that it has the general form of Eq. (21). To calculate the torques, we follow the general approach of Ref. 84. Consider the derivative of  $v_{AB}$  with respect to rotation of molecule A through an angle  $\psi$  about any axis  $\hat{\mathbf{n}}$ . By definition, this gives:

$$\hat{\mathbf{n}} \cdot \mathbf{N}_{AB} = -\frac{\partial v_{AB}}{\partial \psi} = -\sum_{\alpha} \sum_{\hat{\mathbf{e}}=\hat{\mathbf{R}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)} \frac{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)}{\partial \psi}. \quad (23)$$

The sum is over all combinations of unit vectors  $(\hat{\mathbf{a}}_\alpha, \hat{\mathbf{e}})$  for which one,  $\hat{\mathbf{a}}_\alpha$ , rotates with the molecule while the other,  $\hat{\mathbf{e}} = \hat{\mathbf{R}}$  or  $\hat{\mathbf{b}}_\beta$ , remains stationary. This has the effect<sup>45</sup>

$$\frac{\partial \hat{\mathbf{a}}_\alpha}{\partial \psi} = \hat{\mathbf{n}} \times \hat{\mathbf{a}}_\alpha \quad \Rightarrow \quad \frac{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)}{\partial \psi} = \hat{\mathbf{e}} \cdot \hat{\mathbf{n}} \times \hat{\mathbf{a}}_\alpha = -\hat{\mathbf{n}} \cdot \hat{\mathbf{e}} \times \hat{\mathbf{a}}_\alpha.$$

Then Eq. (23) gives

$$\hat{\mathbf{n}} \cdot \mathbf{N}_{AB} = \hat{\mathbf{n}} \cdot \sum_{\alpha} \sum_{\hat{\mathbf{e}}=\hat{\mathbf{R}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)} \hat{\mathbf{e}} \times \hat{\mathbf{a}}_\alpha. \quad (24)$$

Choosing  $\hat{\mathbf{n}}$  to be each of the coordinate directions in turn allows us to identify every component of the torque:

$$\mathbf{N}_{AB} = \sum_{\alpha} \sum_{\hat{\mathbf{e}}=\hat{\mathbf{R}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)} \hat{\mathbf{e}} \times \hat{\mathbf{a}}_\alpha. \quad (25)$$

Writing this out explicitly for molecules A and B:

$$\mathbf{N}_{AB} = \sum_{\alpha} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{a}}_{\alpha} \cdot \hat{\mathbf{R}})} \hat{\mathbf{R}} \times \hat{\mathbf{a}}_{\alpha} - \sum_{\alpha\beta} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{a}}_{\alpha} \cdot \hat{\mathbf{b}}_{\beta})} \hat{\mathbf{a}}_{\alpha} \times \hat{\mathbf{b}}_{\beta} \quad (26a)$$

$$\mathbf{N}_{BA} = \sum_{\beta} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{b}}_{\beta} \cdot \hat{\mathbf{R}})} \hat{\mathbf{R}} \times \hat{\mathbf{b}}_{\beta} + \sum_{\alpha\beta} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{a}}_{\alpha} \cdot \hat{\mathbf{b}}_{\beta})} \hat{\mathbf{a}}_{\alpha} \times \hat{\mathbf{b}}_{\beta}. \quad (26b)$$

Note that eqns (22), (26a) and (26b) give  $\mathbf{N}_{AB} + \mathbf{N}_{BA} + \mathbf{R} \times \mathbf{f}_{AB} = 0$  as before.

If the potential is not (easily) expressible in terms of scalar products, a similar derivation gives the expressions

$$\mathbf{f}_{AB} = -\frac{\partial v_{AB}}{\partial \mathbf{R}} = -\frac{\partial v_{AB}}{\partial R} \hat{\mathbf{R}} - \frac{\partial v_{AB}}{\partial \hat{\mathbf{R}}} \left[ \frac{\mathbf{1} - \hat{\mathbf{R}}\hat{\mathbf{R}}}{R} \right] \quad (27a)$$

$$\mathbf{N}_{AB} = -\sum_{\alpha} \hat{\mathbf{a}}_{\alpha} \times \frac{\partial v_{AB}}{\partial \hat{\mathbf{a}}_{\alpha}} \quad (27b)$$

$$\mathbf{N}_{BA} = -\sum_{\beta} \hat{\mathbf{b}}_{\beta} \times \frac{\partial v_{AB}}{\partial \hat{\mathbf{b}}_{\beta}} \quad (27c)$$

which may be more convenient. Specific examples of the use of these equations in the context of liquid crystal simulations are given elsewhere<sup>85</sup>.

Methods for integrating the rotational equations of motion, using a symplectic splitting method akin to that of section 4.1 have been described elsewhere. These tend to fall into two categories: those based on the rotation matrix<sup>86-88</sup>, and those based on quaternion parameters<sup>89</sup>. Here, we present briefly the former approach. Consider molecule *A*, and drop its identifying suffix. Assuming that the inertia tensor  $\mathbf{I}$  is diagonal in the frame defined by the  $\hat{\mathbf{a}}_{\alpha}$  vectors, the body-fixed angular momentum vector is defined by  $\boldsymbol{\pi} = \mathbf{I} \cdot \boldsymbol{\omega}$ , and the rotation matrix satisfies the equation

$$\frac{d\hat{\mathbf{a}}}{dt} = \hat{\mathbf{a}} \cdot \boldsymbol{\Omega} \quad \text{where} \quad \boldsymbol{\Omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}$$

The angular momentum satisfies

$$\frac{d\boldsymbol{\pi}}{dt} = \boldsymbol{\pi} \times \mathbf{I}^{-1} \boldsymbol{\pi} + \mathbf{N}$$

where the torque  $\mathbf{N}$  is expressed in body-fixed coordinates. Then, the essential idea is to split the rotational propagator

$$e^{i\mathbf{L}_{\pi}\delta t/2} e^{i\mathbf{L}_{\text{free}}\delta t} e^{i\mathbf{L}_{\pi}\delta t/2}$$

where the first and last components advance the angular momentum, using the computed torque, and the central part corresponds to free rotation governed by the kinetic energy part of the hamiltonian. In some special cases, the free rotation problem can be solved exactly, but in the general case it is split further into three separate rotations, each corresponding to a single element of the diagonal inertia tensor.

## 6 Molecular Dynamics in Different Ensembles

In this section we briefly discuss molecular dynamics methods in the constant- $NVT$  ensemble; the reader should be aware that analogous approaches exist for other ensembles, particularly to simulate at constant pressure or stress.

There are three general approaches to conducting molecular dynamics at constant temperature rather than constant energy. One method, simple to implement and reliable, is to periodically reselect atomic velocities at random from the Maxwell-Boltzmann distribution<sup>90</sup>. This is rather like an occasional random coupling with a thermal bath. The resampling may be done to individual atoms, or to the entire system; some guidance on the reselection frequency may be found in Ref. 90.

A second approach<sup>91,92</sup>, is to introduce an extra ‘thermal reservoir’ variable into the dynamical equations:

$$\dot{\mathbf{r}}_i = \mathbf{p}_i/m \quad (28a)$$

$$\dot{\mathbf{p}}_i = \mathbf{f}_i - \zeta \mathbf{p}_i \quad (28b)$$

$$\dot{\zeta} = \frac{\sum_{i\alpha} p_{i\alpha}^2/m - gk_B T}{Q} \equiv \nu_T^2 \left[ \frac{\sum_{i\alpha} p_{i\alpha}^2/m}{gk_B T} - 1 \right] = \nu_T^2 \left[ \frac{T}{T} - 1 \right]. \quad (28c)$$

Here  $\zeta$  is a friction coefficient which is allowed to vary in time;  $Q$  is a thermal inertia parameter, which may be replaced by  $\nu_T$ , a relaxation rate for thermal fluctuations;  $g \approx 3N$  is the number of degrees of freedom.  $T$  stands for the instantaneous ‘mechanical’ temperature. It may be shown that the distribution function for the ensemble is proportional to  $\exp\{-\beta\mathcal{W}\}$  where  $\mathcal{W} = \mathcal{H} + \frac{1}{2}3Nk_B T \zeta^2 / \nu_T^2$ . These equations lead to the following time variation of the system energy  $\mathcal{H} = \sum_{i\alpha} p_{i\alpha}^2/2m + \mathcal{U}$ , and for the variable  $\mathcal{W}$ :

$$\begin{aligned} \dot{\mathcal{H}} &= \sum_{i\alpha} p_{i\alpha} \dot{p}_{i\alpha}/m - \sum_{i\alpha} f_{i\alpha} \dot{r}_{i\alpha} = -\zeta \sum_{i\alpha} p_{i\alpha}^2/m \\ \dot{\mathcal{W}} &= -3Nk_B T \dot{\zeta}. \end{aligned}$$

If  $T > T$ , i.e. the system is too hot, then the ‘friction coefficient’  $\zeta$  will tend to increase; when it is positive the system will begin to cool down. If the system is too cold, the reverse happens, and the friction coefficient may become negative, tending to heat the system up again. In some circumstances, this approach generates non-ergodic behaviour, but this may be ameliorated by the use of chains of thermostat variables<sup>93</sup>. Ref. 94 gives an example of the use of this scheme in a biomolecular simulation.

It is also possible to extend the Liouville operator-splitting approach to generate algorithms for molecular dynamics in these ensembles<sup>65</sup>. Some care needs to be taken, because eqns (28) are not hamiltonian, but it turns out to be possible to correct this using a suitable Poincaré transformation<sup>95</sup> and to implement the resulting symplectic method in an elegant fashion<sup>96,36</sup>.

## 7 How Long? How Large?

Molecular dynamics evolves a finite-sized molecular configuration forward in time, in a step-by-step fashion. There are limits on the typical time scales and length scales that can

be investigated and the consequences must be considered in analyzing the results. Simulation runs are typically short: typically  $t \sim 10^3\text{--}10^6$  MD steps, corresponding to perhaps a few nanoseconds of real time, and in special cases extending to the microsecond regime<sup>97</sup>. This means that we need to test whether or not a simulation has reached equilibrium before we can trust the averages calculated in it. Moreover, there is a clear need to subject the simulation averages to a statistical analysis, to make a realistic estimate of the errors. How long should we run? This depends on the system and the physical properties of interest.

Suppose one is interested in a variable  $a$ , defined such that  $\langle a \rangle = 0$ . The time correlation function  $\langle a(t_0)a(t_0 + t) \rangle$  relates values calculated at times  $t$  apart; assuming that the system is in equilibrium, this function is independent of the choice of time origin and may be written  $\langle a(0)a(t) \rangle$ . It will decay from an initial value  $\langle a(0)a(0) \rangle \equiv \langle a^2 \rangle$  to a long-time limiting value

$$\lim_{t \rightarrow \infty} \langle a(0)a(t) \rangle = \lim_{t \rightarrow \infty} \langle a(0) \rangle \langle a(t) \rangle = 0$$

as the variables  $a(0)$  and  $a(t)$  become uncorrelated; this decay occurs over a characteristic time  $\tau_a$ . Formally we may define a correlation time

$$\tau_a = \int_0^\infty dt \langle a(0)a(t) \rangle / \langle a^2 \rangle.$$

Alternatively, if time correlations decay exponentially at long time,  $\tau_a$  may be identified approximately from the limiting form

$$\langle a(0)a(t) \rangle \propto \exp\{-t/\tau_a\}.$$

Similarly, define a spatial correlation function  $\langle a(0)a(r) \rangle$  relating values computed at different points  $r$  apart. Spatial isotropy allows us to write this as a function of the distance between the points,  $r$ , rather than the vector  $\mathbf{r}$ : note that this symmetry is broken in a liquid crystal. Spatial homogeneity, which applies to simple liquids (but not to solids or smectic liquid crystals) allows us to omit any reference to an absolute origin of coordinates. This function decays from a short-range nonzero value to zero over a characteristic distance  $\xi_a$ , the correlation length.

It is almost essential for simulation box sizes  $L$  to be large compared with  $\xi_a$ , and for simulation run lengths  $\tau$  to be large compared with  $\tau_a$ , for all properties of interest  $a$ . Only then can we guarantee that reliably-sampled statistical properties are obtained. Roughly speaking, the statistical error in a property calculated as an average over a simulation run of length  $\tau$  is proportional to  $\sqrt{\tau_a/\tau}$ : the time average is essentially a sum of  $\sim \tau/\tau_a$  independent quantities, each an average over time  $\tau_a$ . Within the time periods  $\tau_a$ , values of  $a$  are highly correlated. A similar statement can be made about properties which are effectively spatial averages over the simulation box volume  $L^3$ : root-mean-square variations of such averages are proportional to  $\sqrt{(\xi_a/L)^3}$ . This means that collective, system-wide properties deviate by only a relatively small amount from their thermodynamic, large-system, limiting values; the deviation becomes smaller as the averaging volume increases, and is also determined by the correlation length.

Near critical points, special care must be taken, in that these inequalities will almost certainly not be satisfied, and indeed one may see the onset of non-exponential decay of the correlation functions. In these circumstances a quantitative investigation of finite size effects and correlation times, with some consideration of the appropriate scaling laws, must

be undertaken. Phase diagrams of soft-matter systems often include continuous phase transitions, or weakly first-order transitions exhibiting significant pre-transitional fluctuations. One of the most encouraging developments of recent years has been the establishment of reliable and systematic methods of studying critical phenomena by simulation, although typically the Monte Carlo method is more useful for this type of study<sup>98–100, 34, 101, 51</sup>.

## 8 Conclusions

In this introduction, I have tried to focus on points that seem to me both topical and essential to the mainstream of complex fluid dynamics. Others might have chosen a different perspective, or focused on different aspects. Exciting areas which I have had to omit include the use of molecular dynamics to study rare events, the development of mesoscale modelling techniques such as dissipative particle dynamics, the incorporation of electronic degrees of freedom through ab initio molecular dynamics, and the efficient implementation of simulation algorithms on parallel computers. One could argue that these are advanced topics, but they are progressively entering the mainstream, and ready-written packages are steadily removing the need to explain some of the lower-level technical issues which I have included here. Nonetheless, I am a firm believer in understanding what is happening “under the hood”, even if one does not intend to become a “mechanic”, and hopefully the foregoing material will help towards that end.

## Acknowledgements

I have learnt much about the development of intermolecular potentials from Mark Wilson and members of his group. Guido Germano helped me understand several aspects of molecular motion and program design. Conversations with Sebastian Reich, Ben Leimkuhler, and others on the subject of algorithms have convinced me that this is still very much a live area. Finally, I wish to acknowledge the ongoing work of the CCP5 community in the UK.

## References

1. G. C. Maitland, M. Rigby, E. B. Smith, and W. A. Wakeham. *Intermolecular forces: their origin and determination*. Clarendon Press, Oxford, 1981.
2. C. G. Gray and K. E. Gubbins. *Theory of molecular fluids. 1. Fundamentals*. Clarendon Press, Oxford, 1984.
3. M. Sprik. Effective pair potentials and beyond. In Michael P. Allen and Dominic J. Tildesley, editors, *Computer simulation in chemical physics*, volume 397 of *NATO ASI Series C*, pages 211–259, Dordrecht, 1993. Kluwer Academic Publishers. Proceedings of the NATO Advanced Study Institute on ‘New Perspectives on Computer Simulation in Chemical Physics’, Alghero, Sardinia, Italy, September 14–24, 1992.
4. A. J. Stone. *The Theory of Intermolecular Forces*. Clarendon Press, Oxford, 1996.
5. Michael P. Allen, Glenn T. Evans, Daan Frenkel, and Bela Mulder. Hard convex body fluids. *Adv. Chem. Phys.*, 86:1–166, 1993.



6. A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev. A*, 136:405–411, 1964.
7. L. Verlet. Computer experiments on classical fluids. i. thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159:98–103, 1967.
8. J. Weeks, D. Chandler, and H. C. Andersen. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.*, 54:5237–5247, 1971.
9. C. Holm. Dealing with long range interactions: Polyelectrolytes. In *Computational Soft Matter: From Synthetic Polymers to Proteins*, 2004. this volume.
10. S. L. Price. Toward more accurate model intermolecular potentials for organic molecules. *Rev. Comput. Chem.*, 14:225–289, 2000.
11. N. L. Allinger, Y. H. Yuh, and J.-H. Lii. Molecular mechanics - the MM3 force-field for hydrocarbons. 1. *J. Am. Chem. Soc.*, 111:8551–8566, 1989.
12. J.-H. Lii and N. L. Allinger. Molecular mechanics - the MM3 force-field for hydrocarbons. 2. Vibrational frequencies and thermodynamics. *J. Am. Chem. Soc.*, 111:8566–8575, 1989.
13. J.-H. Lii and N. L. Allinger. Molecular mechanics - the MM3 force-field for hydrocarbons. 3. the Van der Waals potentials and crystal data for aliphatic and aromatic hydrocarbons. *J. Am. Chem. Soc.*, 111:8576–8582, 1989.
14. N. L. Allinger, K. S. Chen, and J.-H. Lii. An improved force field (MM4) for saturated hydrocarbons. *J. Comput. Chem.*, 17:642–668, 1996.
15. N. Nevins, K. S. Chen, and N. L. Allinger. Molecular mechanics (MM4) calculations on alkenes. *J. Comput. Chem.*, 17:669–694, 1996.
16. N. Nevins, J.-H. Lii, and N. L. Allinger. Molecular mechanics (MM4) calculations on conjugated hydrocarbons. *J. Comput. Chem.*, 17:695–729, 1996.
17. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. A new force-field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, 106:765–784, 1984.
18. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.
19. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. CHARMM - A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 4:187–217, 1983.
20. W. L. Jorgensen, D. S. Maxwell, and J. TiradoRives. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118:11225–11236, 1996.
21. M. J. Field. *A Practical Introduction to the Simulation of Molecular Systems*. Cambridge University Press, Cambridge, 1999.
22. A. R. Leach. *Molecular Modelling: Principles and Applications*. Prentice Hall, 2nd edition, 2001.
23. H.-D. Holtje, G. Folkers, W. Sippl, and D. Rognan. *Molecular Modeling: Basic Principles and Applications*. Wiley-VCH, 2003.
24. K. Kremer and G. S. Grest. Dynamics of entangled linear polymer melts - a molecular dynamics simulation. *J. Chem. Phys.*, 92:5057–5086, 1990.
25. B. Chen, M. G. Martin, and J. I. Siepmann. Thermodynamic properties of the

- williams, opls-aa, and mmff94 all-atom force fields for normal alkanes. *J. Phys. Chem. B*, 102:2578–2586, 1998.
26. S. K. Nath, F. A. Escobedo, and J. J. de Pablo. On the simulation of vapor-liquid equilibria for alkanes. *J. Chem. Phys.*, 108:9905–9911, 1998.
  27. M. G. Martin and J. I. Siepmann. Transferable potentials for phase equilibria. 1. united-atom description of n-alkanes. *J. Phys. Chem. B*, 102:2569–2577, 1998.
  28. J. R. Errington and A. Z. Panagiotopoulos. A new intermolecular potential model for the n-alkane homologous series. *J. Phys. Chem. B*, 103:6314–6322, 1999.
  29. M. Tsige, J. G. Curro, G. S. Grest, and J. D. McCoy. Molecular dynamics simulations and integral equation theory of alkane chains: comparison of explicit and united atom models. *Macromolecules*, 36:2158–2164, 2003.
  30. E. Garcia, M. A. Glaser, N. A. Clark, and D. M. Walba. HFF: a force field for liquid crystal molecules. *J. Molec. Struct. - THEOCHEM*, 464:39–48, 1999.
  31. D. Reith, H. Meyer, and F. Müller-Plathe. CG-OPT: A software package for automatic force field design. *Comput. Phys. Commun.*, 148:299–313, 2002.
  32. D. Reith, M. Putz, and F. Müller-Plathe. Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.*, 24:1624–1636, 2003.
  33. M. Pütz, K. Kremer, and G. S. Grest. What is the entanglement length in a polymer melt? *Europhys. Lett.*, 49:735–741, 2000.
  34. Michael P. Allen and Dominic J. Tildesley. *Computer simulation of liquids*. Clarendon Press, Oxford, hardback, 385pp edition, 1987.
  35. E. Hairer, C. Lubich, and Gerhard Wanner. Geometric numerical integration illustrated by the Störmer-Verlet method. *Acta Numerica.*, 12:399–450, 2003.
  36. C. J. Cotter and S. Reich. Time stepping algorithms for classical molecular dynamics. In M. Rieth and W. Schommers, editors, *Computational Nanotechnology*. American Scientific Publishers. to appear.
  37. B. Leimkuhler and S. Reich. *Geometric Numerical methods for Hamiltonian Mechanics*. Cambridge University Press, Cambridge, 2004. in press.
  38. C. W. Gear. The numerical integration of ordinary differential equations of various orders. ANL 7126, Argonne National Laboratory, 1966.
  39. C. W. Gear. *Numerical initial value problems in ordinary differential equations*. Prentice-Hall, Englewood Cliffs, NJ, 1971.
  40. J. M. Haile. *Molecular dynamics simulation: elementary methods*. Wiley, New York, 1992.
  41. D. C. Rapaport. *The art of molecular dynamics simulation*. Cambridge University Press, 1995.
  42. L. Verlet. Computer experiments on classical fluids. ii. equilibrium correlation functions. *Phys. Rev.*, 165:201–214, 1968.
  43. R. W. Hockney and J. W. Eastwood. *Computer simulations using particles*. Adam Hilger, Bristol, 1988.
  44. W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. *J. Chem. Phys.*, 76:637–649, 1982.
  45. H. Goldstein. *Classical Mechanics*. Addison Wesley, Reading, Massachusetts, second edition, 1980.

46. S. W. Deleeuw, J. W. Perram, and H. G. Petersen. Hamilton equations for constrained dynamic systems. *J. Stat. Phys.*, 61:1203–1222, 1990.
47. J-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, 23:327–341, 1977.
48. G. Ciccotti, M. Ferrario, and J.-P. Ryckaert. Molecular dynamics of rigid systems in cartesian coordinates. a general formulation. *Molec. Phys.*, 47:1253–1264, 1982.
49. G. Ciccotti and J. P. Ryckaert. Molecular dynamics simulation of rigid molecules. *Comput. Phys. Rep.*, 4:345–392, 1986.
50. H. C. Andersen. Rattle: a “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.*, 52:24–34, 1983.
51. D. Frenkel and B. Smit. *Understanding molecular simulation : from algorithms to applications*. Academic Press, San Diego, 2nd edition, 2002.
52. W. F. van Gunsteren. Constrained dynamics of flexible molecules. *Molec. Phys.*, 40:1015–1019, 1980.
53. M. Fixman. Classical statistical mechanics of constraints: a theorem and application to polymers. *Proc. Nat. Acad. Sci.*, 71:3050–3053, 1974.
54. Michael P. Allen, Mark A. Warren, Mark. R. Wilson, Alain Sauron, and William Smith. Molecular dynamics calculation of elastic constants in Gay-Berne nematic liquid crystals. *J. Chem. Phys.*, 105:2850–2858, 1996.
55. D. Knuth. *The art of computer programming*. Addison-Wesley, Reading MA, 2nd edition, 1973.
56. D. J. Evans and G. P. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. Academic Press, London, 1990.
57. B. L. Holian. The character of the nonequilibrium steady state: beautiful formalism meets ugly reality. In K. Binder and G. Ciccotti, editors, *Monte Carlo and molecular dynamics of condensed matter systems*, volume 49, pages 791–822. Italian Physical Society, Bologna, 1996. Proceedings of the Euroconference on ‘Monte Carlo and molecular dynamics of condensed matter systems’, Como, Italy, July 3–28, 1995.
58. L. E. Reichl. *A modern course in Statistical Physics*. University of Texas Press, Austin, 1980.
59. H. L. Friedman. *A course in statistical mechanics*. Prentice-Hall, Englewood Cliffs, NJ., 1985.
60. B. L. Holian and D. J. Evans. Classical response theory in the Heisenberg picture. *J. Chem. Phys.*, 83:3560–3566, 1985.
61. M. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.*, 97:1990–2001, 1992.
62. H. F. Trotter. *Proc. Amer. Math. Soc.*, 10:545, 1959.
63. S. Toxvaerd. Hamiltonians for discrete dynamics. *Phys. Rev. E*, 50:2271–2274, 1994.
64. H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten. Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Molec. Simul.*, 6:121, 1991.
65. G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein. Explicit reversible integrators for extended systems dynamics. *Molec. Phys.*, 87:1117–1157, 1996.
66. P. Procacci, T. A. Darden, E. Paci, and M. Marchi. ORAC: A molecular dynamics program to simulate complex molecular systems with realistic electrostatic interac-

- tions. *J. Comput. Chem.*, 18:1848–1862, 1997.
67. P. Deuffhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel, editors. *Computational Molecular Dynamics: Challenges, Methods, Ideas*, volume 4 of *Lect. Notes Comp. Sci. Eng.* Springer, Berlin, 1998.
  68. T. Schlick, M. Mandziuk, R. D. Skeel, and K. Srinivas. Nonlinear resonance artifacts in molecular dynamics simulations. *J. Comput. Phys.*, 140:1–29, 1998.
  69. Q. Ma, J. A. Izaguirre, and R. D. Skeel. Verlet-I/r-RESPA/Impulse is limited by nonlinear instabilities. *S.I.A.M. J. Sci. Comput.*, 24:1951–1973, 2003.
  70. B. Garcia-Archilla, J. M. Sanz-Serna, and R. D. Skeel. Long-time-step methods for oscillatory differential equations. *S.I.A.M. J. Sci. Comput.*, 20:930–963, 1998.
  71. J. A. Izaguirre, S. Reich, and R. D. Skeel. Longer time steps for molecular dynamics. *J. Chem. Phys.*, 110:9853–9864, 1999.
  72. J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, and R. D. Skeel. Langevin stabilization of molecular dynamics. *J. Chem. Phys.*, 114:2090–2098, 2001.
  73. T. Schlick, R. D. Skeel, A. T. Brunger, L. V. Kale, J. A. Board, J. Hermans, and K. Schulten. Algorithmic challenges in computational molecular biophysics. *J. Comput. Phys.*, 151:9–48, 1999.
  74. J. G. Gay and B. J. Berne. Modification of the overlap potential to mimic a linear site-site potential. *J. Chem. Phys.*, 74:3316–3319, 1981.
  75. E. de Miguel, L. F. Rull, M. K. Chalam, and K. E. Gubbins. Liquid crystal phase diagram of the Gay-Berne fluid. *Molec. Phys.*, 74:405–424, 1991.
  76. R. Berardi, A. P. J. Emerson, and C. Zannoni. Monte Carlo investigations of a Gay-Berne liquid crystal. *J. Chem. Soc. Faraday Trans.*, 89:4069–4078, 1993.
  77. M. A. Bates and G. R. Luckhurst. Computer simulation studies of anisotropic systems. 26. Monte Carlo investigations of a Gay-Berne discotic at constant pressure. *J. Chem. Phys.*, 104:6696–6709, 1996.
  78. M. R. Wilson. Molecular dynamics simulations of flexible liquid crystal molecules using a Gay-Berne / Lennard-Jones model. *J. Chem. Phys.*, 107:8654–8663, 1997.
  79. J. L. Billeter and R. A. Pelcovits. Simulations of liquid crystals. *Comput. Phys.*, 12:440–448, 1998.
  80. M. P. Allen. Liquid crystal systems. In *Computational Soft Matter: From Synthetic Polymers to Proteins*, 2004. this volume.
  81. R. Berardi, C. Fava, and C. Zannoni. A generalized Gay-Berne intermolecular potential for biaxial particles. *Chem. Phys. Lett.*, 236:462–468, 1995.
  82. D. J. Cleaver, C. M. Care, M. P. Allen, and M. P. Neal. Extension and generalization of the Gay-Berne potential. *Phys. Rev. E*, 54:559–567, 1996.
  83. R. Berardi, C. Fava, and C. Zannoni. A Gay-Berne potential for dissimilar biaxial particles. *Chem. Phys. Lett.*, 297:8–14, 1998.
  84. S. L. Price, A. J. Stone, and M. Alderton. Explicit formulae for the electrostatic energy, forces and torques between a pair of molecules of arbitrary symmetry. *Molec. Phys.*, 52:987–1001, 1984.
  85. M. P. Allen and G. Germano. Rigid body potentials, forces and torques. Technical report, University of Warwick, 2002.  
<http://eprints.csc.warwick.ac.uk/archive/00000064/>.
  86. A. Dullweber, B. Leimkuhler, and R. McLachlan. Symplectic splitting methods for rigid body molecular dynamics. *J. Chem. Phys.*, 107:5840–5851, 1997.

87. A. Kol, B. B. Laird, and B. J. Leimkuhler. A symplectic method for rigid-body molecular simulation. *J. Chem. Phys.*, 107:2580–2588, 1997.
88. T. Holder, B. Leimkuhler, and S. Reich. Explicit variable step-size and time-reversible integration. *Appl. Numer. Math.*, 39:367–377, 2001.
89. T. F. Miller, M. Eleftheriou, P. Pattnaik, A. Ndirango, D. Newns, and G. J. Martyna. Symplectic quaternion scheme for biophysical molecular dynamics. *J. Chem. Phys.*, 116:8649–8659, 2002.
90. H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72:2384–2393, 1980.
91. S. Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molec. Phys.*, 52:255–268, 1984.
92. W. G. Hoover. Canonical dynamics - equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, 1985.
93. G. J. Martyna, M. L. Klein, and M. Tuckerman. Nose-Hoover chains: the canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97:2635–2643, 1992.
94. D. J. Tobias, G. J. Martyna, and M. L. Klein. Molecular dynamics simulations of a protein in the canonical ensemble. *J. Phys. Chem.*, 97:12959–12966, 1993.
95. S. D. Bond, B. J. Leimkuhler, and B. B. Laird. The nose-poincare method for constant temperature molecular dynamics. *J. Comput. Phys.*, 151:114–134, 1999.
96. S. Nose. An improved symplectic integrator for nose-poincare thermostat. *J. Phys. Soc. Japan*, 70:75–77, 2001.
97. Y. Duan and P. A. Kollman. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, 282:740–744, 1998.
98. H. Müller-Krumbhaar and K. Binder. Dynamic properties of the Monte-Carlo method in statistical mechanics. *J. Stat. Phys.*, 8:1–24, 1973.
99. A. M. Ferrenberg, D. P. Landau, and K. Binder. Statistical and systematic errors in Monte-Carlo sampling. *J. Stat. Phys.*, 63:867–882, 1991.
100. K. Binder, editor. *The Monte Carlo Method in Condensed Matter Physics*, volume 71 of *Topics Appl. Phys.* Springer, Berlin, 2nd edition, 1995.
101. D. P. Landau and K. Binder. *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, Cambridge, 2000.

