

## A guide to statistical physics issues in molecular simulations

J.-L. Barrat<sup>1</sup> and R. Vuilleumier<sup>2</sup>

<sup>1</sup>Laboratoire de Physique de la Matière Condensée et Nanostructures, UMR 5586, CNRS et Université Lyon 1, France

<sup>2</sup>École normale supérieure, Département de chimie, UMR CNRS-ENS-UPMC, Paris 06, 8640 PASTEUR, 24 rue Lhomond, 75231 Paris, France

**Abstract.** The aim of this short chapter is to remind the reader of some of the statistical physics ideas that underlie the use of simulation methods, starting from classical Monte Carlo (MC) and Molecular Dynamics (MD) and extrapolating to more advanced simulation methods that permit an exploration of more sophisticated issues, such as long time scales or spectroscopic response. The objective is not an exhaustive description of the relevant issues, but rather to provide a guide to the relevant literature – which has now to a large extent appeared in the form of textbooks.

### 1. INTRODUCTION

The use of computers to explore the transport or thermodynamic properties of systems of interacting particles is almost as old as computers themselves, with the pioneering work of Fermi, Pasta and Ulam [1] attempting to understand the role of anharmonicity in the thermal conductivity of harmonic chains. Interestingly this specific example, which revealed extremely interesting mathematical properties of anharmonic chains at low energy, could serve as a counter example to the properties that are implicitly expected from any one using a simulation program: instead of sampling the phase space of the system in the typical way that will be described below, the system exhibited “recurrences” and an absence of energy equipartition between the modes. Fortunately, such cases are infrequent, but they should remind the simulation practitioner that the “miracle” he observes in his daily use of simulation are associated with the validity of a number of statistical mechanical assumptions and properties of the corresponding systems. By “miracle”, we mean here the fact that properties relevant to actual laboratory experiments can be described accurately using a ridiculously small number of particles, simulated over times that are far from being macroscopic. While we tend now to use “big” (at least a few thousands of atoms) systems even in routine applications, it should be remembered that some of the most spectacular successes of simulation were obtained using systems containing less than a few hundred atoms, and “super” computers that could not compete with a modern laptop. The understanding of the statistical physics and algorithms behind such successes is important, if only to understand the intrinsic limitations and to make the casual user aware of possible dangers of a blind use of simulation packages.

Before we start our discussion, we would like to emphasize that the material presented here is extensively covered in a number of textbooks, among which we particularly recommend those by Allen and Tildesley [2, 3], Frenkel and Smit [4], and Krauth [5].

## 2. STATISTICAL ENSEMBLES AND AVERAGES

### 2.1 Gibbs and Boltzmann averages

Thermodynamics [6] describes a system through a very limited number of macroscopic variables, which are sufficient to specify the equilibrium properties of a system of interacting particles. For a fluid, these variables are for example the number of particles, the volume and the energy. For a given choice of variables, the equilibrium state corresponds to the extremum with respect to all internal variables of an appropriate thermodynamic potential, the entropy in this case. At constant volume and temperature, the appropriate potential is the free energy. Statistical physics allows one to justify the thermodynamics approach by using a microscopic description of the system in a classical “phase space”, and assumptions concerning the statistical weight of various configurations in phase space. In statistical mechanics introductory lectures, it is often stated that any attempt to compute the trajectory of the system in phase space is both unfeasible and useless, and that only the statistical distribution is relevant. This in a sense reflects the huge success of the statistical physics developed by Gibbs, with the notion of “ensemble” of systems characterized by a statistical weight in phase space, appropriate for the thermodynamic constraints under consideration. The statistical weight is denoted by  $w(\Gamma)$ , where  $\Gamma = (q_i, p_i)$  denotes a set of positions ( $q$ ) and momenta ( $p$ ) in phase space. The average value of any observable, i.e. of any function  $A(\Gamma)$  defined on phase space, is then given by an integral of the form

$$\langle A \rangle = \int d\Gamma w(\Gamma) A(\Gamma) \quad (2.1)$$

At equilibrium, the statistical weight depends on phase space through an energy function  $H(\Gamma)$ , which is the Hamiltonian governing the time evolution of the trajectories, i.e.  $w(\Gamma) = f(H(\Gamma))$ . One goal of equilibrium statistical mechanics is to develop appropriate calculation schemes to estimate such highly multidimensional integrals. Numerical simulation provides such a scheme, which can proceed in two different ways: either an attempt to sample as directly as possible the statistical weight  $w(\Gamma)$ , which is the idea behind Monte Carlo methods. Or use the “useless” method mentioned above, i.e. calculate a trajectory of the system using the microscopic equation of motions, with the hope that the phase space will be sampled according to one of Gibbs distributions. This is the basis of the molecular dynamics method, which typically samples the “microcanonical”, constant energy, statistical ensemble – with variants that allow one to sample constant temperature or constant pressure ensembles. Averages in this case will be computed in a completely different manner, namely

$$\langle A \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt A(\Gamma(t)) \quad (2.2)$$

That the two kind of averages coincide can be considered as a “practical proof” of the ergodic hypothesis, and it is always surprising that this equivalence actually works for very short simulation times – typically much shorter than the duration of any scattering experiment, and much, much shorter than any conceivable “ergodic time” that would allow for Poincaré recurrences.

### 2.2 Ensemble equivalence

For reasons that will become clear below, molecular dynamics typically samples a constant energy ensemble, while Monte-Carlo is most appropriate for sampling of constant temperature, Boltzmann-Gibbs distributions. It is therefore important to estimate how the results may depend on the specific method that is used to estimate the observable, or in other words their ensemble dependence. This topic is a classical issue in equilibrium statistical physics, and was discussed with specific emphasis on computer simulations in a celebrated paper by Lebowitz [7]. Let us start with the formal computation of the free energy in the canonical ensemble,  $F(N, V, T)$ , from the knowledge of the entropy  $S(N, V, E)$  in the microcanonical ensemble. Denoting by  $\Omega(E) = \exp(S(E)/k_B)$  the density of states (i.e. the volume

of phase space available per unit of energy), one has

$$\exp(-\beta F) = Z = \int dE \Omega(E) \exp(-\beta E) = \int dE \exp(S(E)/k_B - \beta E) \quad (2.3)$$

where  $\beta = 1/(k_B T)$  is the inverse temperature. If  $\bar{E}$  is the energy that makes the integrand maximum it is easily shown, using a saddle point approximation valid for large systems, that

$$\beta F(\beta) = -S(\bar{E})/k_B + \beta \bar{E} - \ln \sqrt{2\pi k_B T c} \quad (2.4)$$

where  $c$  is the heat capacity defined by  $\frac{\partial^2 S}{\partial E^2} = \frac{1}{c}$ .

Averages of various observables at a fixed temperature  $T$  can be obtained in a similar manner from the observable calculated at the energy  $\bar{E}$  defined above, as:

$$\langle A \rangle_T = \langle A \rangle_E + \frac{1}{2} \frac{\partial^2 A}{\partial E^2} \langle \delta E^2 \rangle \quad (2.5)$$

Hence the typical difference that can be expected between observables measured in molecular dynamics and Monte Carlo simulations is of the order of  $1/N$ , where  $N$  is the number of particles in the system. The situation is completely different for fluctuations in the value of the observables, as is quite obvious from the consideration of energy fluctuations which are absent in a molecular dynamics simulation, and proportional to the specific heat in a Monte-Carlo calculation. The appropriate formula in this case is

$$\langle \delta A^2 \rangle_E = \langle \delta A^2 \rangle_T - \frac{k_B T^2}{Nc} \left( \frac{\partial A}{\partial T} \right)^2 \quad (2.6)$$

which shows that the difference in the variance of the observable is typically of the same order as the variance itself. Note that such formulae can be directly exploited to obtain thermodynamic coefficients in various ensembles, an interesting example being that of the heat capacity per particle in the microcanonical ensemble, that can be related to the fluctuations in the kinetic energy  $K$  in a constant energy molecular dynamics simulations using [7] :

$$\langle \delta K^2 \rangle = \frac{3N(k_B T)^2}{2} \left( 1 - \frac{3k_B}{2c} \right). \quad (2.7)$$

### 2.3 Correlation times and statistical accuracy

In the above we mentioned that very short simulation times could allow one to determine rather accurately the values of thermodynamic observables, and thus ensure a reasonable sampling of phase space. The minimal requirement for an efficient sampling is, of course, that the system explores a sufficiently large number  $\mathcal{N}$  of *independent* configurations, as the statistical error will decrease as  $1/\sqrt{\mathcal{N}}$ . The independence of configurations, with respect to the calculation of a specific observable  $A$ , can be assessed by estimating the correlation time of this observable

$$\tau_A = \int_0^\infty dt \frac{\langle \delta A(t) \delta A(0) \rangle}{\langle \delta A(0) \delta A(0) \rangle} \quad (2.8)$$

with  $\delta A = A - \langle A \rangle$ . An order of magnitude of  $\tau_A$  may be obtained by computing the time correlation function along the trajectory, which may be a molecular dynamics or Monte-Carlo trajectory (in the latter case  $\tau_A$  is interpreted as a number of cycles rather than as a physical time). The trajectory should always be sufficiently long to cover a large number of  $\tau_A$ , so that the calculation of the correlation function is possible – even if not very accurate – and can be used to get at least an estimate of the order of magnitude of the correlation time. To this end, it is important to realize that a time dependent correlation function should always be estimated using a “sliding average” along the trajectory: all pairs of states separated by a time lag  $t$  are used to obtain the correlation function at time  $t$ .

Based on general statistical grounds one expects that the number of independent samples for an observable  $A$  is of the order of  $\mathcal{N} = t/\tau_A$  for a run of total duration  $t$ , so that if one defines the quantity

$$A_t = \frac{1}{t} \int_0^t ds A(s) \quad (2.9)$$

the mean squared error on  $A_t$  is of the order of  $\langle \delta A^2 \rangle^{1/2} \times \sqrt{\tau_A/t}$ . The argument can be formalized by writing explicitly the expression for the variance of  $A_t$ :

$$\langle \delta A_t^2 \rangle = \left\langle \frac{1}{t^2} \int_0^t ds \int_0^t ds' \delta A(s) \delta A(s') \right\rangle = \frac{1}{t} \int_0^t ds \left(1 - \frac{s}{t}\right) \langle \delta A(s) \delta A(0) \rangle \quad (2.10)$$

In the limit of large  $t$  and for a fast enough (typically exponential) decay of the correlation function this behaves as  $2\tau_A/t$ , as expected.

From a practical point of view an alternative approach is to subdivide the total run into  $k$  subruns of duration  $\Delta = t/k$ , and to compute the variance of  $A_\Delta$  on the corresponding statistical sample with  $k$  realizations. This variance is expected to scale linearly with  $k$ , with a prefactor proportional to  $\tau_A$ .

## 2.4 Finite size

Simulations necessarily deal with small numbers of particles; even a large scale simulation involving one million particles corresponds to an extremely small ‘‘sample’’. This intrinsic limitation is to some extent bypassed by using periodic boundary conditions, which convert a system with a small degrees of freedom into an infinitely extended one, albeit with a periodic density. The main advantage of periodic boundary conditions (PBC), which are systematically used in studies of bulk systems, is that the system is free of any interface (remember that the ‘‘boxes’’ that limit each periodic image do not have any physical meaning, all systems of space spanning boxes with the same periodicity are strictly equivalent). Hence the size effects that are a topic of current interest in nanosciences are essentially avoided, with corrections to the thermodynamic limit that are typically of order  $1/N$  rather than  $1/N^{1/3}$ .

Still, it should be remembered that the finite size and the use of PBC implies some important limitations: first, the system is intrinsically periodic with a Bravais lattice associated with the periodicity of the boundary conditions. Therefore, a quantity such as the particle density has, strictly speaking, non-zero Fourier components  $\rho_{\mathbf{k}}$  only for those wavevectors that belong to the associated reciprocal lattice, typically  $\mathbf{k} = \mathbf{G} = \frac{2\pi}{L}(\mathbf{n}_x, \mathbf{n}_y, \mathbf{n}_z)$  for a cubic box of side  $L$ . Hence the structure factor  $S(\mathbf{k})$  can be computed only for the corresponding values of  $\mathbf{k}$ , and in particular any extrapolation towards small wavevector (for example by Fourier transforming the pair correlation function) should be handled very carefully. Clearly this is a limitation for studying systems in which correlations become long ranged, in particular close to second order phase transitions. A related problem is the study of defects in solids, or solute molecules in liquids, which should systematically be interpreted as periodic arrays of impurities with a concentration that is often unphysically high. Whether or not this should be considered as a serious problem depends on the particular problem under consideration, and can be analysed by performing systematic studies using different box sizes.

In the field of phase transitions, the apparent drawback of small size has been turned into an advantage by using the ideas of finite size scaling, developed during the 1980’s and described extensively in the book by Privman [8]. Both for first order and second order transitions, the finite size induces a peculiar behaviour, which, when carefully analysed, gives access to relevant physical characteristics of the transition. In first order transition, small size gives rise to hysteretic behaviour, which vanishes in the thermodynamic limit. For example, in a fluid, a histogram of the volume at a given pressure, or of the energy at a given temperature, will give rise to a double peaked structure when the coexistence conditions are reached. An analysis of this double peak as a function of system size can be used to locate the coexistence conditions (see for example the chapter by D. Tildesley in [3]). In second order transition, one takes advantage of the fact that quantities that diverge in the limit of infinite system sizes,

such as the susceptibility  $\chi$  (or compressibility for a fluid) display a rounded peak in the vicinity of the critical temperature for finite systems. As the only relevant length is the correlation length  $\xi(T)$  of fluctuations, the simplest finite size scaling analysis consists in writing, on dimensional grounds

$$\frac{\chi(T, L)}{\chi(T, L = \infty)} = f\left(\frac{L}{\xi_\infty(T)}\right) \quad (2.11)$$

As the susceptibility remains finite for finite systems, and knowing that  $\chi(T, \infty)$  and  $\xi_\infty(T)$  diverge respectively as  $|T - T_c|^{-\gamma}$  and  $|T - T_c|^{-\nu}$  as the temperature approaches  $T_c$ , it follows immediately that

$$\chi(T, L) = L^{-\gamma/\nu} g(L|T - T_c|^\nu) \quad (2.12)$$

where  $g$  is some unspecified, regular function, that does not depend on  $L$ . For large, but finite systems, it is expected that  $\chi(T, L)$ , plotted as a function of temperature, has a maximum in the vicinity of  $T_c$ . This maximum occurs for a fixed value of the argument of  $g$ , so that the corresponding temperature changes with size as  $T_{max} = T_c + A/L^{1/\nu}$ . Similarly, the height of the maximum will change as a function of system size as  $L^{-\gamma/\nu}$ , and the width as  $L^{-1/\nu}$ . Hence playing with the system size, which can indeed be achieved rather easily in numerical simulations, becomes an elegant and powerful manner of identifying critical properties.

## 2.5 Linear response

Very often simulations aim at determining not only the average values of thermodynamic observables, but also functions that can be considered as response functions of the system to some external perturbation, which may be static (independent of time), such as the specific heat, the thermal expansion or the compressibility, or time dependent, as for example the frequency dependent dielectric response. Very generally, such response functions can, at the level of a linear response to a perturbation, be obtained in two different manners: one can either monitor the spontaneous fluctuations in the system at equilibrium, as illustrated above for the specific heat, or actually apply the external perturbation (which in the case of specific heat may just be a change in temperature) and measure the change in the relevant observables. These two approaches are formally related through the fluctuation dissipation theorem, which states that for a perturbation of the form  $-f(t)A(\Gamma)$ , the response function  $\chi_{AB}$  of the observable  $B$ , defined through

$$\langle \delta B(t) \rangle = \int_{-\infty}^t \chi_{AB}(t-s) f(s) ds \quad (2.13)$$

is related to the equilibrium time dependent correlation function  $\langle \delta A(t) \delta B(0) \rangle$

$$\chi_{AB}(t) = -\frac{1}{k_B T} \frac{d}{dt} \langle \delta A(t) \delta B(0) \rangle. \quad (2.14)$$

In most experiments, one or the other possibility is used to study the dynamics of the response. For example, neutron scattering measures a time dependent correlation function without perturbing the system, while a dielectric experiment will probe the response function directly. In simulations, a choice must be made between the application of an external perturbation, which has the advantage of selecting the relevant fluctuations (at the price of perturbing the trajectory), and following an unperturbed trajectory, during which all kinds of different correlation functions can be obtained simultaneously.

Simulations have the great flexibility of allowing one to impose external perturbations on the system that do not necessarily correspond to experimentally realisable external fields. This leads to a very powerful panoply of “synthetic” algorithms, that can be used to determine all kinds of correlations. Beautiful examples and a detailed discussion of algorithms used to determine, for example, viscosity (by imposing an external shear flow) or thermal conductivities are given in the book of Evans and Morriss [9]. An important point here is that one should never be afraid of being “unphysical” in applying

external perturbations to a system. In a sense, the possibility of submitting a system to “unphysical” perturbations, i.e. perturbations that could not be realized in the lab, is indeed a strength of simulation methods. As examples we may mention the possibility to “force” an external shear flow in the  $x$  direction (with a velocity gradient in the  $z$  direction) using a perturbation to the Hamiltonian of the form  $\sum_i p_{ix}q_{iz}$ . Another interesting possibility is to assign random “charges”  $\epsilon_i = \pm 1$  to individual particles that interact with an external potential  $V(q)$ . The resulting perturbation is of the form  $\sum_i \epsilon_i V(q_i)$ , and it can be easily shown that a statistical average over the random charge distribution allows one to recover the individual, “tagged particle” response, with a statistics that actually involves all particles in the system.

### 3. MOLECULAR DYNAMICS

Molecular dynamics is in some sense the most “brute force” of simulation methods. All the simulator has to do is to integrate numerically a (large) number of coupled equations of motion, a task that can be carried out using any elementary integration algorithm. Still, it turns out that not all algorithms are the same, and that the game can be made much more subtle with trajectories that sample a constant temperature statistical weight rather than constant energies. These aspects are discussed below. We do not discuss here the tricks that allow one to transform a system of coupled equations that should imply  $O(N^2)$  calculations (one for each pair interaction) into a system that only requires  $O(N)$  calculations per step, owing to the short range nature of interatomic potentials.

#### 3.1 Constant energy MD

One particularity of the second order differential equations that govern the time evolution of an ensemble of interacting particle is that the energy  $H$  is, rigorously, a constant of motion. Hence the numerical scheme used to integrate these equations should, in principle, conserve this essential property. One possible route here is to choose a high order, highly accurate algorithm that ensures energy conservation at short times to a high degree of accuracy. Most simulation studies however make a different choice, i.e. the choice of a robust algorithm first proposed by Verlet in this context. For a time step  $h$ , the positions  $q(t)$  are updated using the forces at time  $t$  according to

$$q(t+h) = 2q(t) - q(t-h) + F(t) \quad (3.1)$$

This algorithm is easily shown to be accurate only to order  $h^4$ , which is far from being exceptional. Still it has the very important property of being exactly time reversible, meaning that the operations that lead from  $q(t-h)$  to  $q(t+h)$  are exactly symmetric to those that would evolve the system from  $q(t+h)$  to  $q(t-h)$ . In other words, the trajectory of the system can be traced backward in time without error other than the rounding off errors. This property can be shown [4] to imply that the computed trajectory is *exactly* the trajectory that would be generated by an approximate Hamiltonian, that differs from the real one at the order of  $h^2$ . Hence the energy conservation for this pseudo-Hamiltonian is essentially preserved by the numerical scheme (apart from round off errors), which ensures excellent properties of energy conservation (absence of drift) for long times, even if the short time accuracy is not particularly good.

Whatever the algorithm, round off errors will always imply that the simulated trajectory is not an exact solution of the equations of motion, even for the approximate Hamiltonian. The essential point is therefore that the statistical properties of this trajectory are quite robust, and can be obtained from the simulation in a way that is largely independent of these numerical aspects, provided a “reasonable” algorithm is used. Fortunately, these statistical properties appear, for systems at thermal equilibrium, to be extremely robust, so that in general the reader should have not to worry too much about the mathematical details of the integration procedure; still, it is always recommended to check that your results do not depend on the time step used in the simulation!

### 3.2 Constant temperature MD

In constant energy simulation, temperature is an outcome of the simulation, which may be obtained for example from the kinetic energy of the particles. It is therefore highly desirable to devise methods that allow one to impose a given temperature in Molecular Dynamics. Many methods have been devised to achieve this goal, the simplest one being “velocity rescaling” that rescales velocities regularly by a uniform factor in order to impose a value of the kinetic energy that corresponds to the desired temperature. The resulting phase space distribution, however, is an unusual “constant kinetic energy” ensemble. Other methods involve adding “Langevin thermostats” that effectively transform the molecular dynamics into Brownian dynamics. A very elegant method, proposed by Nosé, allows one to preserve the Hamiltonian nature of the equations of motion and to produce a phase space distribution which is exactly the canonical, Gibbs-Boltzmann distribution. The so-called Nosé thermostat introduces an extended phase space with a new variable  $s$  and an associated momentum  $p_s$ , with the idea that by integrating out the extra variables in a microcanonical (constant energy surface) distribution of the extended phase space, one recovers the Gibbs-Boltzmann distribution in the original phase space. Technically, the extended Hamiltonian is written (for a single particle  $(p, q)$ ) as

$$H^*(p, q, p_s, s) = A \ln s + \frac{p_s^2}{2Q} + \frac{p^2}{2sm} + V(q) \quad (3.2)$$

where  $A$  and  $Q$  are two parameters, yet unspecified. The idea is to ensure that

$$\int ds dp_s dp dq f(p/\sqrt{s}, q) \delta(H^*(p, q, p_s, s) - E_0) \sim \int dp dq f(p, q) \exp\left(-\beta\left(V(q) + \frac{p^2}{2m}\right)\right) \quad (3.3)$$

which can easily be shown to be the case when  $A = 3N + 1$  is the number of degrees of freedom in the (extended) system. The choice of  $Q$ , sometimes described as the “mass of the thermostat” is on the other hand arbitrary. While any choice of  $Q$  ensures that the statistical distribution will be given by the Gibbs-Boltzmann expression, the dynamical properties will in general depend on  $Q$ , which can be written in the form  $Q = (3N + 1)k_B T \tau^2$  where  $\tau$  is a typical relaxation time for temperature fluctuations. Therefore, for the calculation of dynamical properties in the presence of a thermostat, the influence of the thermostat mass should be checked carefully.

## 4. MONTE-CARLO

### 4.1 Monte-Carlo integration

Monte-Carlo integration is based on a very simple idea. In order to calculate an integral of the form  $\int d\Gamma f(\Gamma)w(\Gamma)$ , where  $f$  is a function and  $w$  a statistical weight, a convenient method is to compute the average value of  $f$  over a set of points sampled according to the distribution  $w$  (i.e. the local density of points should be proportional to  $w$ ). A textbook example illustrated in ref. [5]: consider the case where  $w$  is a uniform probability in a square of side 1, and  $f$  is a function which is 1 inside a circle of diameter 1 drawn within this square, 0 outside. To obtain the integral, which is simply equal to  $\pi$  in this specific case, one draws at random points in the square using a random number generator that generates the uniform distribution  $w$ . For a large number of points, the integral is obtained by taking the average value of  $f$  for the sample, i.e. the number of points that fall within the circle divided by the total number of trials.

This rather trivial example is, in a space with a small number of dimensions, a rather inefficient method for calculating integrals. In a space of high dimensions on the contrary, it turns out to be one of the less demanding from a computational standpoint, provided one knows how to produce a sample of points according to the prescribed weight  $w(\gamma)$ . This is, in general, a non trivial task, and the Metropolis

algorithm described below aims at achieving it for the phase space distributions encountered in systems at thermal equilibrium. The general idea is to generate a “Markov chain”, i.e. a statistical process that will produce a series of points with a probability distribution that converges towards  $w$ . In the preceding example, this can be done by generating a random walk inside the square: one starts with a random walker inside the square, which performs random steps of size  $\delta \ll 1$ . If one of those steps brings the random walker outside the square, the attempt is a rejected step and the walkers stays in the same place. An important feature, as will become evident in the discussion of the Metropolis algorithm, is that this point has to be counted in the averaging procedure.

## 4.2 Metropolis algorithm

A Markov chain in phase space is a stochastic process (the random walk described above being a simple example) that can be defined by a set of *transition probabilities*  $\pi(\Gamma \rightarrow \Gamma')$ . The evolution equation for the probability density of an ensemble of random walkers at step  $i$ ,  $f(\Gamma, i)$  is a so called master equation of the form

$$f(\Gamma, i + 1) - f(\Gamma, i) = -f(\Gamma, i) \int d\Gamma' \pi(\Gamma \rightarrow \Gamma') + \int d\Gamma' \pi(\Gamma' \rightarrow \Gamma) f(\Gamma', i) \quad (4.1)$$

The aim of the algorithm is to generate a probability density that, in the limit of a large number of steps  $i \rightarrow \infty$ , converges to the prescribed density  $w(\Gamma)$ . As the distribution  $w(\Gamma)$  is, in general, defined over all phase space, this implies that the transition probabilities must allow the random walkers to visit all phase space. In other words, one should avoid a Markov chains that does not allow every point in phase space to be connected to any other in a finite number of steps: the chain has to be *ergodic*. The second requirement is that the right hand side of equation (4.1) vanishes for  $f = w$ , a condition that is fulfilled whenever the (even stronger) property known as *detailed balance* is realized, i.e. for any  $\Gamma, \Gamma'$ :

$$w(\Gamma')\pi(\Gamma' \rightarrow \Gamma) = w(\Gamma)\pi(\Gamma \rightarrow \Gamma') \quad (4.2)$$

The Metropolis algorithm was proposed in 1953 [10], interestingly in the same Los Alamos context as the Molecular Dynamics method. In fact the name “Monte-Carlo method” had been introduced slightly earlier by Metropolis and Ulam [11], in describing a stochastic method for simulating trajectories of neutrons in fissile or absorbing materials. The algorithm is usually described as consisting of two steps:

- Starting from any point  $\Gamma$ , choose a point  $\Gamma'$  at random in a small neighbourhood around  $\Gamma$ . This can be achieved by modifying by a small amount some of the atomic coordinates, and is described as a “trial move”. This trial move must be such that the probability of reaching  $\Gamma'$  from  $\Gamma$ , let us say  $\pi_0(\Gamma \rightarrow \Gamma')$ , is identical to the probability  $\pi_0(\Gamma' \rightarrow \Gamma)$  of reaching  $\Gamma$  from  $\Gamma'$ .
- Compute the energy  $H(\Gamma')$  after the trial move, and the energy difference  $\Delta H = H(\Gamma') - H(\Gamma)$ . Accept the move and go to  $\Gamma'$  with a probability

$$\pi_a(\Gamma \rightarrow \Gamma') = \min(1, \exp(-\beta\Delta H)) \quad (4.3)$$

i.e. the trial move is always accepted if the energy decreases, and accepted with a finite probability  $\exp(-\beta\Delta H)$  if it increases the energy.

The corresponding rates can easily be shown to verify the ergodicity and detailed balance conditions for the Gibbs distribution  $w(\Gamma) = \exp(-\beta H(\Gamma))$ .

The simple Metropolis idea is still the basis of most elementary MC simulations; it can however be refined considerably, in particular by enlarging the notion of “trial moves” to moves that are totally “unphysical” (i.e. do not correspond to a motion that could be expected to take place in an actual system simulated using molecular dynamics). Examples of such unphysical moves are for example the reptation move in polymers (in which a polymer at the end of a chain is “cut” and “pasted” at the other end) or the double bridging move [12], in which two chains are cut and recombined to form two new chains.

The only essential constraint in inventing moves is that one should always check for detailed balance in the transition probabilities from one state to another.

Ideally, a Monte-Carlo simulation that would sample directly the distribution  $w(\Gamma)$  would be 100% efficient, i.e. all calculations would contribute to the desired averages. On the contrary, the Metropolis algorithm inevitably introduces rejections, i.e. calculations of the energy at new state points that are eventually not part of the Markov chains, and can therefore be considered as useless. The rule of thumb in this field is that a rejection rate of about 50% is somehow “efficient”: smaller rejection rates simply indicate that the moves that are attempted are not very efficient and produce configurations that are statistically correlated, while high rejection rates indicates on the contrary moves that are too bold and result in the system staying in the same configuration. A way to partially avoid rejections is to generalize the Metropolis algorithm by realizing that the algorithm allows a complete freedom as to the choice of trial moves. Instead of choosing  $\pi_0(\Gamma' \rightarrow \Gamma)$ , the probability distribution for trial moves, as being a probability of uniformly moving atoms within a small volume around the starting point, it can indeed be any distribution that can be exactly sampled, and therefore incorporate at least some part of the energy. The formula for the transition probability is then generalized in the form:

$$\pi(\Gamma \rightarrow \Gamma') = \pi_0(\Gamma \rightarrow \Gamma') \min \left( 1, \exp(-\beta\Delta H) \times \frac{\pi_0(\Gamma' \rightarrow \Gamma)}{\pi_0(\Gamma \rightarrow \Gamma')} \right) \quad (4.4)$$

If, for example,  $H = H_0 + H_1$  where  $H_0$  is simple enough that the distribution  $\exp(-\beta H_0)$  can be sampled exactly, then clearly the acceptance of the moves will be made much more efficient. As a simple example, one may consider moves of a polymer chain with a simple quadratic energy associated with the angular degrees of freedom. Clearly if the position of a new monomer is selected by a priori taking into account this part of the energy (which is easy as the corresponding distribution is Gaussian), many useless rejections will be avoided.

Many other possibilities of cleverly extending the original Metropolis Monte-Carlo method are described in [4, 5]. The crucial point to keep in mind before trying a “clever” Monte-Carlo is however, detailed balance, which should be respected in all cases.

## 5. MODELLING THE PHYSICAL SYSTEM

What statistical mechanics then tells us is that the macroscopic properties of a system at equilibrium depend only on the data of the system Hamiltonian, defined on the system phase space. Performing a computer simulation thus involves the choice of the integration method (Molecular Dynamics or Monte-Carlo, constant energy or constant temperature etc.), and the definition of the particles in the system and the force-field acting on them (the phase-space and the Hamiltonian). In the reminder of this introduction, we will discuss different current approaches to the modelling of force-fields.

It should be noted however that a third ingredient is also necessary to integrate the evolution equation of Molecular Dynamics or the Markov chain of the Monte-Carlo method: an initial condition for the system. The choice of this initial configuration is not harmless. For the “miracle” discussed above to occur, it is indeed important to start from a configuration representative of the equilibrium ensemble. This may be easily achieved for a low viscosity fluid for example. It is enough in that case to start from a crystal structure and run a high enough temperature simulation: the system will quickly melt and rapidly reach a representative liquid state. After throwing away this initial equilibration period, one can rather safely compute averages over the rest of the simulation. If the system however contains slow dynamics, this transient equilibration time may extend to prohibitive times and the choice of the initial configuration becomes crucial. When simulating a protein for example, it would be in the large majority of cases unfeasible to start from a denaturated configuration and wait for the protein to fold spontaneously. The construction of the initial configuration is however very case specific and cannot be treated in this short introduction to computational methods; the interested reader should consult the literature pertinent to his problem.

### 5.1 The Born-Oppenheimer approximation

Atomic and molecular systems are made of positive nuclei (ions) and electrons, interacting through the coulombic interaction. Molecular simulations thus aim at describing this composite system. In most applications however, this problem is greatly simplified using the Born-Oppenheimer approximation. It reduces the description of the system to nuclei interacting through a complex potential energy, a mechanistic framework that we are very much used to.

To see how this happens, let's start from the quantum description of the full system, nuclei and electrons. The instantaneous quantum state of such a system is given by the wavefunction  $\Psi_S(\mathbf{R}_I, \mathbf{r}_1, \dots, \mathbf{r}_N)$ , which is an integrable function of the coordinates of the nuclei and the electrons,  $\mathbf{R}_I$  and  $\mathbf{r}_1, \dots, \mathbf{r}_N$  respectively. The dynamics of the wavefunction  $\Psi_S$  is then governed by the Hamiltonian of the total system:

$$\hat{H}_T = \hat{T}_N + \hat{V}_{NN} + \hat{H}_e, \quad (5.1)$$

where  $\hat{T}_N$  is the kinetic energy operator for the nuclei,  $\hat{V}_{NN}$  is the potential energy operator acting on nuclear degrees of freedom only and arising from the nuclei-nuclei coulombic interaction and  $\hat{H}_e$ ,

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{Ne}, \quad (5.2)$$

is the Hamiltonian for the electrons, including the interaction between the electrons and the nuclei.

Considering that the nuclear mass is much larger than the electron mass, we can in many cases approximate well the total system wavefunction by a wavefunction of the form [13]:

$$\Psi_S(\mathbf{R}_I, \mathbf{r}_1, \dots, \mathbf{r}_N) = \chi(\mathbf{R}_I) \Psi_0(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{R}_I), \quad (5.3)$$

where  $\Psi_0(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{R}_I)$  is the ground state electronic wavefunction of the electronic Hamiltonian  $\hat{H}$  at fixed nuclear (ionic) configuration  $\{\mathbf{R}_I\}$  and  $\chi(\mathbf{R}_I)$  constitutes the nuclear wavefunction part of the system wavefunction. We will note  $E_0(\mathbf{R}_I)$

$$E_0(\mathbf{R}_I) = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \quad (5.4)$$

the (ground-state) energy of the electrons at this ionic configuration  $\{\mathbf{R}_I\}$ . This factorisation of the total system wavefunction assumes that the electrons, which move and relax much faster than the nuclei due to the large mass ratio, are always found in their ground state and follow adiabatically the nuclear motion [13]. This constitutes the Born-Oppenheimer approximation which greatly simplifies the problem of finding the dynamics of the ions+electrons system. Cases where this approximation fails are thus cases where the electron motion is not fast enough, due to the presence of a low lying excited state, and cannot follow adiabatically the nuclear motion and thus give rise to spontaneous population of the low lying electronic excited states.

The evolution of the system is thus the evolution of the pure nuclear wavefunction, which is governed by the Hamiltonian [13]:

$$\hat{H}_N = \hat{T}_N + \hat{V}_{NN} + E_0(\hat{\mathbf{R}}_I). \quad (5.5)$$

The electrons enter  $\hat{H}_N$  only through a potential energy term,  $E_0(\hat{\mathbf{R}}_I)$ , added to the bare nuclei-nuclei interaction  $\hat{V}_{NN}$ . This potential energy term due to the electrons is the ground state energy of the electronic system at a fixed nuclear (ionic) configuration.

Taking now the classical limit for the nuclei, the potential energy surface for the nuclear classical motion is thus  $U_0(\{\mathbf{R}_I\}) = V_{NN}(\{\mathbf{R}_I\}) + E_0(\{\mathbf{R}_I\})$ . The forces acting on the nuclei are extracted from the gradient of this electronic energy with respect to the nuclear positions,  $-\frac{\partial U_0}{\partial \mathbf{R}_I}$ .

The description of this system of classical nuclei can now be handled using the statistical mechanics framework described above. The phase-space is that of the positions  $\mathbf{R}_I \equiv q$  and momenta  $\mathbf{P}_I = M_I \mathbf{V}_I \equiv p$  of the nuclei with mass  $M_I$ . The interaction between those is the potential energy  $U_0(\{\mathbf{R}_I\})$ , which involves the electronic ground-state energy at this nuclear configuration.

## 5.2 First-Principle Molecular Dynamics

Two main routes are possible at this point: to evaluate for each nuclear configuration the electronic ground state energy (and eventually the resulting forces on the nuclei) either from an explicit electronic state calculation or from an approximate analytical model of it. The number of electronic state calculations required in the first case is enormous as it is equal to the number of atomic configurations sampled in the Molecular Dynamics or Monte-Carlo simulation. Furthermore, statistical mechanics is mainly interested in condensed phase systems and the number of atoms in a simulation cannot be very small in order to represent such state well; a single electronic state calculation for one configuration of the simulated system is thus already quite computationally expensive. This route then seems out of reach. Nevertheless, in 1985, Car and Parrinello showed that this is quite feasible, employing Density Functional Theory as electronic state method and a clever trick to obtain the electronic ground-state energies and forces along a Molecular Dynamics trajectory [14]. They have thus opened the way for First-Principle Molecular Dynamics (FPMD) simulations.

First-principle Molecular Dynamics allows for important rearrangements of the electronic structure in the system, like covalent bond breaking and forming. It is then well suited for the simulation of systems undergoing chemical reactions or at extreme conditions of pressure and temperature [15, 16]. Among the many applications of FPMD, one can quote for example the description of the structure and transport of an excess proton in liquid water [17]. Very frequent proton transfers occur in this system with the excess proton being essentially shared between two water molecules; the change of identity of this so-called “special bond” is then responsible for the fast proton transport in liquid water [17].

One should however be aware of the defect of the quantum chemistry method employed, namely DFT. Although based on exact principles, its implementation for the calculation of ground state energies of atomic configurations is subject to approximations. One common defect of the most popular functionals is the lack of long range dispersions, decreasing as the 6<sup>th</sup> power of the inverse of the interatomic distance. This has notorious effects on the description of liquid water by FPMD. The main microscopic aspects are very well reproduced by FPMD, as well as solvation properties [15], but the melting temperature is much too high [18] and the density at 1 bar much too low (about 0.8 g.cm<sup>-3</sup>!) [19]. At 300 K and 1 g.cm<sup>-3</sup>, FPMD then leads to an overstructured liquid with too low a diffusion constant. Progress has been achieved to compensate for this error of the common implementations of DFT but this is still a very active field of research [20–22].

The main limitation of FPMD is however the small sample sizes and timescales accessible. With today’s computers one can simulate a few hundreds of atoms for a few hundreds of picoseconds. Although this will surely evolve in the near future with the increasing computer power available and with the advent of fast algorithms (see for example the presentation of Thierry Deutch in this school), classical force-fields allow, in comparison to this, simulations at much larger scales.

## 5.3 Molecular mechanics – classical force-fields

Classical<sup>1</sup> force-fields are an analytic model of the Born-Oppenheimer potential energy surface. Instead of calculating the electronic energy for each atomic configuration, we use a much simpler analytic model to approximate it. It is historically the first approach to the modelling of system Hamiltonians in statistical mechanics, even before simulation appeared: this is for example the Lennard-Jones potential for the modelling of rare gases [23], or even the harmonic force-field to describe the role of phonons in the heat capacity of crystals by Einstein or Debye [6]. Classical force-fields are well adapted to situations where chemical bonding is maintained (no chemical reactions) and gentle thermodynamic conditions: whenever the electronic structure is not subject to important modifications.

---

<sup>1</sup> “Classical” in contrast to the “quantum” character of the FPMD that involve a quantum chemistry calculation at each time-step.

Let's take as an example, the popular SPC-E force-field for water [24]. The force-field first specifies the geometry of the water molecule: an HOH angle of  $109.47^\circ$  and an OH distance of  $1.0 \text{ \AA}$ . The atoms of the water molecules are assigned partial charges and interact through the Coulomb potential. These partial charges are not integer charges and they aim at describing in a simplified manner the charge distribution (nuclei+electrons) in the molecule. These charges are  $q_O = -0.8476 e$  and  $q_H = +0.4238 e$  for the SPC-E model. There is a further interaction between the oxygen atoms having the form of a Lennard-Jones potential:

$$U_{LJ}(r) = 4\epsilon \left( \frac{\sigma^{12}}{r^{12}} - \frac{\sigma^6}{r^6} \right), \quad (5.6)$$

with  $\epsilon = 0.65 \text{ kJ.mol}^{-1}$  and  $\sigma = 3.166 \text{ \AA}$ . The first term describes the short-range repulsion ( $1/r^{12}$ ) between water molecules and the long range dispersion attraction is proportional to  $1/r^6$ . This apparently very simplified description of the water molecules as a soft sphere with three partial charges turns out to be very powerful and able to describe most of the very complex behaviour of liquid water: tetrahedral structure (not forced a priori by the force-field), high dielectric constant etc. All popular force-fields for liquid water, with some refinements, share more or less the same simplified description of the interactions.

In this molecular mechanics pictures, the interactions between particles are often classified as bonded and non-bonded:

$$U_0(\{\mathbf{R}_I\}) = U_{bonded}(\{\mathbf{R}_I\}) + U_{non-bonded}(\{\mathbf{R}_I\}). \quad (5.7)$$

The bonded interactions hold together the atoms in the molecules. A bond between two atoms can be represented by a harmonic potential:

$$U_r(r) = \frac{1}{2} k_r (r - r_0)^2, \quad (5.8)$$

where  $r_0$  is the equilibrium bond length and  $k_r$  is the force constant related to the frequency of the bond vibrations, holding together the two atoms. Further contributions from the bends (angle between two bonds with one common atom) and torsions (of three consecutive bonds) allow to describe the atom and the bond arrangement within the molecules. As in the SPC-E case, it may be favourable to constrain the molecular geometry (constraining for example the bond length or bending angles) rather than allowing bond vibrations. As vibrations are very fast, longer time-step can then be used in such MD simulations. The SHAKE algorithm is very efficient in integrating the equation of motions with holonomic constraints [25].

The non-bonded interactions, that act between atoms of different molecules (intermolecular potential) or between atoms far apart in a large molecule, involve different contributions [26]:

$$U_{non-bonded}(\{\mathbf{R}_I\}) = U_{rep}(\{\mathbf{R}_I\}) + U_{disp}(\{\mathbf{R}_I\}) + U_{coul}(\{\mathbf{R}_I\}) + U_{ind}(\{\mathbf{R}_I\}), \quad (5.9)$$

where  $U_{rep}(\{\mathbf{R}_I\})$  is the short range repulsion,  $U_{disp}(\{\mathbf{R}_I\})$  the long range dispersion attraction,  $U_{coul}(\{\mathbf{R}_I\})$  the Coulombic interaction between charge distributions, and finally  $U_{ind}(\{\mathbf{R}_I\})$  the induction or polarisation term describing the contribution from the relaxation of the electronic structure. This last term is often omitted and we then speak of non-polarisable force-fields. As for the SPC-E model, the charge distribution is often represented by partial charges on atoms and it then amounts to

$$U_{coul}(\{\mathbf{R}_I\}) = \frac{1}{2} \sum_{I \neq J} \frac{q_I q_J}{4\pi\epsilon_0 |\mathbf{R}_I - \mathbf{R}_J|}, \quad (5.10)$$

special care has to be taken however in case of the periodic boundary conditions often employed in simulations [4]. The two terms in the Lennard-Jones potential, eq. (5.6), correspond respectively to the repulsion and long-range dispersion contribution to  $U_{non-bonded}$ .

Apart from the general consideration above, there is no universal way of building a classical force-field. The analytical form and the parameters may be determined, by trial and error, so as to best reproduce available experimental data (microscopic or macroscopic); these are sometimes called empirical force-fields. On the contrary, one may wish to determine classical force-fields by directly fitting the forces or energy to explicit electronic state calculations of the Born-Oppenheimer energy surface (force-matching technique) [27]. A third route consists of evaluating the quantities appearing in the force-field description (partial charges, long-range dispersion coefficient etc.) from the properties of the electronic ground-state [28].

#### 5.4 Coarse-grained and continuous models

Classical force-fields allow the simulation of thousands to a few millions of atoms. In situations like suspensions of colloids or polymers, protein association, membranes etc. one may wish to reach even larger scales, while the description of all atoms in the system may not be necessary. To this aim, coarse-graining amounts to representing a group of atoms, from a few atoms (a chemical group) to the whole object (colloid, polymer or fatty acid) by one single interaction site [29, 30]. The solvent itself may be completely discarded if we are only interested in the behaviour of the solutes. At an intermediate level of description, the solvent can be coarse-grained or described as a continuous medium in which the system, colloids, polymers etc., is embedded. The simplest of such continuous descriptions is that of a dielectric medium with the dielectric constant of the solvent.

The same type of contributions as in molecular mechanics are involved in the description of the interactions between the coarse-grained particles[30]: short-range repulsion, intermediate and long-range attraction<sup>2</sup>, coulombic interactions, bonded terms etc. These interactions are effective; in the canonical ensemble they are associated with the free-energy as a function of the coarse-grained particle configuration. Similar to the determination of classical force-fields, the determination of a model effective interaction in coarse-grained model can then either rely on experimental data or be obtained from the description of the system at a lower scale. In the latter case, this involves the explicit calculation of free-energies at fixed geometries of the collective coarse-grained variables.

The difference between FPMD and classical force-fields is merely the approach to the evaluation of the Born-Oppenheimer potential energy surface; coarse-graining however involves discarding some degrees of freedom. This has the further consequence of modifying the effective dynamics of the remaining degrees of freedom of interest, introducing dissipative terms. For example, the solvent gives rise to friction forces that need to be introduced in the dynamics of embedded objects like colloids. An example of such dissipative dynamics is, taking the high friction limit, Brownian dynamics that involves dissipation and random noise [31]:

$$\Delta \mathbf{r} = \left( \frac{1}{k_B T} \mathbf{D} \mathbf{F} + \frac{\partial}{\partial \mathbf{r}} \mathbf{D} \right) \Delta t + \mathbf{R}, \quad (5.11)$$

where  $\Delta \mathbf{r}$  is the particle displacement during an interval  $\Delta t$ ,  $\mathbf{F}$  is the force acting on the particles at positions  $\mathbf{r}$ ,  $\mathbf{D}$  is a diffusion tensor and  $\mathbf{R}$  a random displacement, chosen from a Gaussian distribution with zero mean,  $\langle \mathbf{R} \rangle = 0$ , and a variance satisfying the fluctuation-dissipation relation  $\langle \mathbf{R} \mathbf{R}^T \rangle = 2 \mathbf{D} \Delta t$ . The friction force on one coarse-grained particle depends not only on its motion but also on that of the other coarse-grained particles through hydrodynamics interactions: the diffusion tensor in Brownian dynamics in principle involves coupling between the motion of all particles.

---

<sup>2</sup> An example of intermediate range attraction in this context is hydrogen bonding.

At even larger scale, we find continuous description of matter involving density and velocity fields, concentration fields etc. We leave here the field of molecular simulations.

### 5.5 Bridging the different scales

FPMD, classical force-fields, coarse-grained and Brownian dynamics approaches cover different scales of the representation of matter, both in length and time. As briefly discussed above, it is possible to bridge these different descriptions: force-matching to design force-fields from FPMD, free-energy calculation to design coarse-grained models... One can then simulate sequentially a system at different scales, starting from the smallest to the largest. This is particularly of interest for homogeneous, or not too heterogeneous systems. If, on the contrary, a small region of interest (e.g. the reactive center of an enzyme, a crack in a solid...) is to be treated at the electronic level, while the surroundings may be treated at a molecular-mechanics level, one wishes to perform a simulation involving simultaneously the two levels of description: this is the purpose of mixed Quantum-Mechanics/Molecular-Mechanics (QM/MM) methods [32]. Other methods have also been recently designed to provide an interface between molecular-mechanics and coarse-grained descriptions [33], and even continuous hydrodynamics [34]. These mixed techniques usually solve the problem of length scale but the largest time-scale achievable is that of the lower level of description employed (a few 10's of ps in QM/MM for example). We would like to mention here that acceleration techniques have been and are still being developed to solve this second problem [35]. These aim at more efficiently sampling rare but interesting regions of phase-space.

## 6. SUMMARY AND CONCLUSION

Thanks to the ever increasing availability of efficient numerical resources, molecular simulations are being used more and more on a routine basis, often in the form of "simulation packages", to interpret and complement data obtained from various scattering experiments – in particular neutrons. The link between molecular simulation and neutron scattering, which typically access information on similar space and time scales, has indeed always been strong, as illustrated e.g. by the pioneering works of A. Rahman at Argonne. With this very short guide, we have attempted to remind the reader of two important features of molecular simulations. First, the success of molecular simulation is based on an understanding of statistical physics, of numerical algorithms and of the electronic structure of matter. A basic understanding of these aspects is required to properly understand the nature of the models that are implemented, and the manner in which the data is obtained from these models. While all these aspects could easily serve as pedagogical illustrations of general concepts in physics, chemistry or material science, it is regrettable that in many cases the description of computer simulations is essentially absent from the traditional curriculum. Hence computer simulations are often seen by non-practitioners as a rather obscure and esoteric field, while they are in fact relatively easy to understand and to implement, even for non-specialists.

A second point, which is not always realized, is the fact that the field is in constant evolution and development, with a continued effort to develop more accurate models and more efficient methods to explore the behavior of more complex materials. This effort involves development in quantum physics and chemistry (with e.g. the treatment of excited states, or of the interaction of radiation with matter) as well as in statistical physics, with the development of coarse graining methods that allow one to explore larger length and time scales, or of "multiscale" methods in which several levels of description are used simultaneously. These developments often imply sophisticated theoretical tools, and make the development of computational methods a challenging field from the standpoint of theoretical physics and chemistry.

## References

- [1] E Fermi, J Pasta, and S Ulam. Studies of nonlinear problems. *Los Alamos report LA-1940 (1955) published later in Collected Papers of Enrico Fermi, E. Segré (Ed.), University of Chicago Press (1965)*, 1965.
- [2] M.P. Allen and D.J. Tildesley. *Computer simulation of liquids*. Oxford University Press, Oxford, 1987.
- [3] M.P. Allen (Ed.). *Computer simulation in chemical physics*. Kluwer, Dordrecht, 1993.
- [4] D. Frenkel and B. Smit. *Understanding molecular simulation*. Academic Press, New York, 2002.
- [5] W Krauth. *Statistical Mechanics, Algorithms and computations*. Oxford University Press, Oxford, 2006.
- [6] H.B. Callen. *Thermodynamics and an Introduction to Thermostatistics*. Wiley, New-York, 1985.
- [7] JL Lebowitz, JK Percus, and L Verlet. Ensemble dependence of fluctuations with applications to machine computations. *Physical Review*, 153:250–257, 1967.
- [8] V. Privman (Ed.). *Finite Size Scaling and Numerical Simulation of Statistical Systems*. World Scientific, Singapore, 1990.
- [9] D.J. Evans and G.P. Morriss. *Statistical Mechanics of Nonequilibrium Liquids, 2nd edition*. Cambridge University Press, 2008.
- [10] N Metropolis, AW Rosenbluth, MN Rosenbluth, AH Teller, and E Teller. Equation of state calculation by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.
- [11] N Metropolis and S Ulam. The Monte-Carlo method. *Journal of the American Statistical Association*, 44:335–341, 1949.
- [12] PVK Pant and DN Theodorou. Variable connectivity method for the atomistic Monte-Carlo simulation of polydisperse polymer melts. *Macromolecules*, 28:7224–7234, 1995.
- [13] Albert Messiah. *Quantum Mechanics, 2 volumes Bound as One*, chapter XVIII. Dover, New York, 2000.
- [14] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55:2471, 1985.
- [15] D. Marx and J. Hutter. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge University Press, 2009.
- [16] J. Kohanoff. *Electronic Structure Calculations For Solids And Molecules: Theory And Computational Methods*. Cambridge University Press, 2006.
- [17] D. Marx, M.E. Tuckerman, J. Hutter, and M. Parrinello. The nature of the hydrated excess proton in water. *Nature*, 397:601–604, 1999.
- [18] Soohaeng Yoo, Xiao Cheng Zeng, and Sotiris S. Xantheas. On the phase diagram of water with density functional theory potentials: The melting temperature of ice  $i_h$  with the Perdew–Burke–Ernzerhof and Becke–Lee–Yang–Parr functionals. *The Journal of Chemical Physics*, 130:221102, 2009.
- [19] Jochen Schmidt, Joost VandeVondele, I.-F. William Kuo, Daniel Sebastiani, J. Ilja Siepmann, Juerg Hutter, and Christopher J. Mundy. Isobaric-isothermal molecular dynamics simulations utilizing density functional theory: An assessment of the structure and density of water at near-ambient conditions. *The Journal of Physical Chemistry B*, 113:11959–11964, 2009.
- [20] O. Anatole von Lilienfeld, Ivano Tavernelli, Ursula Rothlisberger, and Daniel Sebastiani. Optimization of effective atom centered potentials for london dispersion forces in density functional theory. *Phys. Rev. Lett.*, 93:153004, 2004.
- [21] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of Chemical Physics*, 132:154104, 2010.
- [22] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. Van der Waals density functional for general geometries. *Phys. Rev. Lett.*, 92:246401, 2004.

- [23] J. E. Lennard-Jones. On the determination of molecular fields. *Proc. R. Soc. Lond. A*, 106:463, 1924.
- [24] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91:6269–6271, 1987.
- [25] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. Numerical integration of the cartesian equation of motion of a system with constraints: molecular dynamics of n-alkanes. *J. of Computational Physics*, 23:327, 1977.
- [26] A.J. Stone. *Theory of intermolecular forces*. Oxford University Press, Oxford, 1996.
- [27] Sergei Izvekov, Michele Parrinello, Christian J. Burnham, and Gregory A. Voth. Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: A new method for force-matching. *The Journal of Chemical Physics*, 120:10896–10913, 2004.
- [28] Benjamin Rotenberg, Mathieu Salanne, Christian Simon, and Rodolphe Vuilleumier. From localized orbitals to material properties: Building classical force fields for nonmetallic condensed matter systems. *Phys. Rev. Lett.*, 104:138301, 2010.
- [29] F Muller-Plathe. Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *CHEMPHYSCHEM*, 3:754–769, 2002.
- [30] Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries. The martini force field: a coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry B*, 111:7812–7824, 2007.
- [31] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland, Amsterdam, 1987.
- [32] U. Roethlisberger, M. Sprik, and J. Hutter. Time and length scales in ab initio molecular dynamics. In P. Niebala, M. Maraschal, and G. Ciccotti, editors, *Bridging time scales: Molecular simulations for the next decade*, page 413. Springer Verlag, 2002.
- [33] Matej Praprotnik, Silvina Matysiak, Luigi Delle Site, Kurt Kremer, and Cecilia Clementi. Adaptive resolution simulation of liquid water. *Journal of Physics: Condensed Matter*, 19:292201, 2007.
- [34] Rafael Delgado-Buscalioni, Kurt Kremer, and Matej Praprotnik. Coupling atomistic and continuum hydrodynamics through a mesoscopic model: Application to liquid water. *The Journal of Chemical Physics*, 131:244107, 2009.
- [35] M. Ferrario, G. Ciccotti, and K. Binder, editors. *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, volume 1-2 of *Lecture Notes in Physics*. Springer, Berlin, 2006.