Göran Wahnström

MOLECULAR DYNAMICS
Lecture notes

Göteborg, 27 October 2021

Molecular dynamics simulation

Molecular dynamics (MD) is a computer simulation technique where the time evolution of a set of interacting particles is followed by integrating their equation of motion. The technique has been applied to systems of several hundreds to millions of particles and has given much insight into the behaviour of interacting classical many-particle systems. The physical movements of atoms and molecules are investigated by numerically solving Newton's equation of motion using a description for the inter-atomic interaction. Equilibrium and transport properties are obtained using the theoretical approach developed within statistical mechanics.

The method was originally conceived within theoretical physics in the late 1950s and is today applied extensively in various scientific disciplines as materials science and nanoscience, and for biomolecular systems. For a general background to the technique we refer to the excellent books by Frenkel and Smit [1] and Allen and Tildesley [2]. A more elementary introduction can be found in the book by Haile [3].

Contents

1 Classical mechanics

1.1 Newton's equation of motion

In molecular-dynamics (MD) simulations Newton's equation of motion is solved for a set of interacting particles

$$m_i \, \ddot{\boldsymbol{r}}_i(t) = \boldsymbol{F}_i \; ; \quad i = 1, \ldots, N \tag{1}$$

where m_i is the mass of particle i, $\boldsymbol{r}_i(t)$ its position at time t, $\boldsymbol{F}_i$ the force acting on particle i, and N the number of particles. We will mainly consider the case when the force is conservative. It can then be expressed in terms of the gradient of a potential $V_{\rm pot}$,

$$\boldsymbol{F}_i(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = -\nabla_i V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \tag{2}$$

where $\nabla_i = \partial / \partial \boldsymbol{r}_i$. The total energy for the system

$$E = \sum_{i=1}^{N} \frac{m_i \boldsymbol{v}_i^2}{2} + V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \tag{3}$$

will then be conserved. The first term is the kinetic energy, where $\boldsymbol{v}_i$ is the velocity for particle i, and the second term is the potential energy. Example of a non-conservative system is if e.g. friction is introduced through a velocity dependent force as in Brownian dynamics.

Eq. (1), together with the force in Eq. (2), corresponds to a set of $3N$ coupled second order ordinary differential equations. Given a set of initial conditions for the positions $(\boldsymbol{r}_1(0), \ldots, \boldsymbol{r}_N(0))$ and velocities $(\boldsymbol{v}_1(0), \ldots, \boldsymbol{v}_N(0))$ a unique solution can formally be obtained.

1.2 Hamilton's formulation of classical mechanics

It is often more convenient to use the Hamilton's formulation of classical mechanics [4]. The connection to statistical mechanics becomes more direct and transparent as well as the transition to quantum mechanics. The system is then defined in terms of a set of F generalized coordinates q_α and generalized momenta p_α, where F is the number of degrees of freedom. For a system with N particles in three dimensions $F = 3N$. The *Hamilton's equations of motion* is

$$\dot{q}_\alpha = \frac{\partial \mathcal{H}}{\partial p_\alpha} \; , \qquad \dot{p}_\alpha = -\frac{\partial \mathcal{H}}{\partial q_\alpha} \; ; \qquad \alpha = 1, \ldots, F \tag{4}$$

where $\mathcal{H}$ is the Hamiltonian for the system and is given by the total energy

$$\mathcal{H}(q_\alpha, p_\alpha) = \mathcal{E}_{\rm kin} + \mathcal{E}_{\rm pot} \; , \tag{5}$$

the sum of the kinetic $\mathcal{E}_{\rm kin}$ and potential $\mathcal{E}_{\rm pot}$ energies. The total energy for the system is a conserved quantity. This can shown be taking the time-derivative of the Hamiltonian

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(q_\alpha, p_\alpha) &= \sum_{\alpha=1}^{F}\left[\frac{\partial\mathcal{H}}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial\mathcal{H}}{\partial p_\alpha}\dot{p}_\alpha\right] \\ &= \sum_{\alpha=1}^{F}\left[\frac{\partial\mathcal{H}}{\partial q_\alpha}\frac{\partial\mathcal{H}}{\partial p_\alpha} - \frac{\partial\mathcal{H}}{\partial p_\alpha}\frac{\partial\mathcal{H}}{\partial q_\alpha}\right] = 0 \end{aligned} \tag{6}$$

where in the second line the Hamilton's equation of motion has been used.

For a system of N particles in three dimensions using Cartesian coordinates we have

$$\mathcal{H} = \sum_{i=1}^{N}\frac{m_i\boldsymbol{v}_i^2}{2} + V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \tag{7}$$

where, in this case, $\boldsymbol{r}_i$ is the generalized coordinates and $m_i\boldsymbol{v}_i$ the generalized momenta. Using Eq. (4) we obtain

$$\dot{\boldsymbol{r}}_i = \boldsymbol{v}_i \ , \qquad m_i\dot{\boldsymbol{v}}_i = -\frac{\partial}{\partial\boldsymbol{r}_i}V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \ ; \qquad i = 1, \ldots, N \tag{8}$$

or

$$m_i\ \ddot{\boldsymbol{r}}_i(t) = -\nabla_i V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \ ; \qquad i = 1, \ldots, N \tag{9}$$

which is equal to Newton's equation of motion (1) with the force given by Eq. (2)

Hamilton's equation of motion describes the unique evolution of the coordinates and momenta subject to a set of initial conditions. More precisely, Eq. (4) specifies a trajectory

$$\boldsymbol{x}(t) \equiv (q_1(t), \ldots, q_F(t), p_1(t), \ldots, p_F(t)) \tag{10}$$

in phase-space, starting from an initial point $\boldsymbol{x}(0)$. For a system of N particles in three dimensions the phase-space is $6N$-dimensional. The energy conservation condition restricts the motion on a ($6N$-1)-dimensional surface in phase-space, known as the constant-energy hypersurface or simply the *constant-energy surface*, see Fig. 1. We also notice that the Hamiltonian in Eq. (7) is invariant when replacing $\boldsymbol{v}_i$ with $-\boldsymbol{v}_i$ which implies that the time evolution is reversible in time.

Hamilton's equation of motion can also be expressed using a matrix or *symplectic* notation [4]. The time derivative of the phase space vector $\boldsymbol{x}(t)$ can be written as

$$\dot{\boldsymbol{x}}(t) = \left(\frac{\partial\mathcal{H}}{\partial p_1}, \ldots, \frac{\partial\mathcal{H}}{\partial p_F}, -\frac{\partial\mathcal{H}}{\partial q_1}, \ldots, -\frac{\partial\mathcal{H}}{\partial q_F}\right) \tag{11}$$

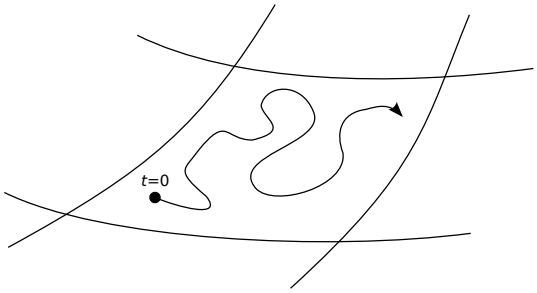

Figure 1: A trajectory $\boldsymbol{x}(t)$ in $6N$-dimensional phase-space that describes the time-evolution of the system. The trajectory is restricted to a ($6N$-1)-dimensional constant-energy surface.

and hence the Hamiltion's equation of motion can be recast as

$$\dot{\boldsymbol{x}} = \mathbf{M}\,\frac{\partial \mathcal{H}}{\partial \boldsymbol{x}} \tag{12}$$

where $\mathbf{M}$ is a matrix expressible in block form as

$$\mathbf{M} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \tag{13}$$

where $\mathbf{0}$ and $\mathbf{I}$ are the $F \times F$ zero and identity matrices, respectively. Dynamical systems expressible in this form are said to possess a symplectic structure.

The time-dependence of a property $\mathcal{A}(t)$, that is a function of phase-space $\mathcal{A}(t) = \mathcal{A}(\boldsymbol{x}(t))$, is formally given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}(\boldsymbol{x}(t)) &= \sum_{\alpha=1}^{F}\left[\frac{\partial \mathcal{A}}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial \mathcal{A}}{\partial p_\alpha}\dot{p}_\alpha\right] \\ &= \sum_{\alpha=1}^{F}\left[\frac{\partial \mathcal{A}}{\partial q_\alpha}\frac{\partial \mathcal{H}}{\partial p_\alpha} - \frac{\partial \mathcal{A}}{\partial p_\alpha}\frac{\partial \mathcal{H}}{\partial q_\alpha}\right] \\ &= \{\mathcal{A}, \mathcal{H}\} \end{aligned} \tag{14}$$

where we have introduced the *Poisson bracket* notation

$$\{\mathcal{A}, \mathcal{B}\} \equiv \sum_{\alpha=1}^{F}\left[\frac{\partial \mathcal{A}}{\partial q_\alpha}\frac{\partial \mathcal{B}}{\partial p_\alpha} - \frac{\partial \mathcal{B}}{\partial q_\alpha}\frac{\partial \mathcal{A}}{\partial p_\alpha}\right] \tag{15}$$

The Poisson bracket can be used to define the *Liouville operator* $i\mathcal{L}$ according to

$$i\mathcal{L}\ldots \equiv \{\ldots, \mathcal{H}\} \tag{16}$$

where $i = \sqrt{-1}$. More explicitly, the Liouville operator can written as the differential operator

$$i\mathcal{L} = \sum_{\alpha=1}^{F} \left[\dot{q}_\alpha \frac{\partial}{\partial q_\alpha} + \dot{p}_\alpha \frac{\partial}{\partial p_\alpha} \right] \tag{17}$$

where Hamilton's equation of motion has been used once again. The time-dependence of an arbitrary phase-space function $\mathcal{A}(\boldsymbol{x}(t))$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A} = i\mathcal{L}\mathcal{A} \tag{18}$$

with the formal solution

$$\mathcal{A}(t) = e^{i\mathcal{L}t}\mathcal{A}(0) \tag{19}$$

The operator $\exp(i\mathcal{L}t)$ is known as the classical propagator. By introducing the imaginary unit i into the definition of the Liouville operator the classical propagator $\exp(i\mathcal{L}t)$ resembles the quantum propagator $\exp(-i\hat{H}t/\hbar)$. The time evolution of the phase-space vector $\boldsymbol{x}(t)$ can be written as

$$\boldsymbol{x}(t) = e^{i\mathcal{L}t}\boldsymbol{x}(0) \tag{20}$$

This equation describes the central numerical problem in molecular dynamics simulation, to obtain the time-dependent trajectory $\boldsymbol{x}(t)$ in phase space, given an initial condition $\boldsymbol{x}(0)$.

2 Statistical averaging

Molecular-dynamics (MD) simulations generate a very detailed information of the system at the microscopic level, positions and momenta for all particles as function of time. To obtain useful information on the macroscopic level one has to make some averaging. This is the province of statistical mechanics [5, 6]. In molecular dynamics one evaluates the average quantities by performing time averaging, an average along the generated trajectory in phase-space.

Consider some macroscopic equilibrium property A. It could be, for instance, the temperature or pressure of a system. Suppose that a microscopic, instantaneous value $\mathcal{A}$ can be defined, which is a function of phase-space $\mathcal{A}(\mathbf{x})$ and when averaging the macroscopic observable quantity A is obtained. We define the *time-average* according to

$$A = \langle \mathcal{A} \rangle_{\text{time}} = \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \int_0^{\mathsf{T}} \mathcal{A}(\mathbf{x}(t)) \mathrm{d}t \tag{21}$$

where we let the observation time T goes to infinity.

Conventional statistical mechanics is not based on the time averaging approach used in MD simulations. To facilitate a more analytical approach Gibbs introduced the concept of an *ensemble* (see appendix A). An ensemble can be viewed as a collection of systems described by the same microscopic interactions and sharing a common set of macroscopic properties. The most fundamental ensemble is the microcanonical ensemble. It represents an isolated N-particle system with constant energy E and constant volume V. The corresponding probability distribution function is proportional to

$$\delta(\mathcal{H}(\mathbf{x}) - E)$$

where the delta function is supposed to select out all those states in phase space of an N-particle system in a volume V that have the desired energy E. The ensemble average of a quantity $\mathcal{A}$ is then given by

$$\langle \mathcal{A} \rangle_{NVE} = \frac{\int_V \mathrm{d}\mathbf{x}\ \mathcal{A}(\mathbf{x}) \delta(\mathcal{H}(\mathbf{x}) - E)}{\int_V \mathrm{d}\mathbf{x}\ \delta(\mathcal{H}(\mathbf{x}) - E)} \tag{22}$$

where we have indicated by the subscript NVE that it is an ensemble average with fixed macroscopic parameters NVE, the microcanonical ensemble average.

The time average in Eq. (21) is also taken under the condition of constant energy, particle number and volume. If the system, given an infinite amount of time, is able to visit all states on the constant-energy surface it is said to be *ergodic*. The time and ensemble average are then equal

$$\langle \mathcal{A} \rangle_{\text{time}} = \langle \mathcal{A} \rangle_{NVE} \tag{23}$$

This is called the *ergodic hypothesis*. It is generally believed that most systems are ergodic and the ergodic hypothesis can be used for systems studied using molecular dynamics. However, one should be aware of that some systems are not ergodic *in practise*, such as glasses and metastable phases, or even *in principle* such as nearly harmonic solids. To formally prove ergodicity has turned out to be a very difficult task.

In statistical mechanics several different ensembles are introduced, appropriate at different conditions. The most important is the *canonical ensemble* which describes a system at constant particle number N, volume V and temperature T. The corresponding probability distribution is proportionell to

$$\exp\left[-\beta\mathcal{H}(\boldsymbol{x})\right]$$

and the ensemble average is given by

$$\langle\mathcal{A}\rangle_{NVT} = \frac{\int_V \mathrm{d}\mathbf{x}\ \mathcal{A}(\mathbf{x})\exp\left[-\beta\mathcal{H}(\boldsymbol{x})\right]}{\int_V \mathrm{d}\mathbf{x}\ \exp\left[-\beta\mathcal{H}(\boldsymbol{x})\right]} \tag{24}$$

where the subscript NVT indicate a canonical ensemble average.

Analytical calculations are often more easy to perform using the canonical ensemble compared with the microcanonical ensemble. The precise nature of the ensemble is often not so important. For large systems the average values from the microcanonical and the canonical ensembles will be the same

$$\langle\mathcal{A}\rangle_{NVE} = \langle\mathcal{A}\rangle_{NVT} + \mathcal{O}\left(\frac{1}{N}\right) \tag{25}$$

but the fluctuations from the average values are in general different. In appendix D we derive microscopic expressions for temperature and pressure for a system using the canonical ensemble.

3 Physical models of the system

To perform an MD simulation we need a description of the forces on the particles. That is usually obtained from the gradient of a potential

$$V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$$

known as the *potential energy surface*. We will consider systems where the particles are atoms and, hence, $(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ are atomic coordinates. The size of the system also has to be specified together with suitable boundary conditions.

3.1 The potential energy surface

The origin of the potential energy surface $V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ is quantum mechanical, but it is a highly nontrivial problem to determine $V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ quantum mechanically. Density functional theory, quantum Monte Carlo and quantum chemistry techniques are being used but all these methods are computationally very demanding.

Considerable effort has therefore been directed to develop simplified descriptions of the potential energy surface, *e.g.* models for the inter-atomic interactions. They often differ in terms of the type of bonding they are supposed to mimic: covalent, metallic, ionic or van der Waals type. We will here consider a few different types of models which all can be classified as *classical potentials*, while they do not include the electronic degrees of freedom explicitly.

A natural way to represent $V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ is to expand it in terms of one-body, two-body, three-body and higher order terms,

$$\begin{aligned} V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = &\sum_i v_1(\boldsymbol{r}_i) + \sum_i \sum_{j>i} v_2(\boldsymbol{r}_i, \boldsymbol{r}_j) \\ &+ \sum_i \sum_{j>i} \sum_{k>j>i} v_3(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) + \ldots \end{aligned} \tag{26}$$

where the first term represents the effect of an external field and the remaining terms inter-atomic interactions. Often we only keep the two-body term and set v_1 equal to zero. This approximation is used in describing systems with predominantly ionic or van der Waals bonding. For covalent bonding systems it is necessary to include the directional bonding through higher order terms in the expansion and in metals the conduction electrons introduce a many-body term into the description of the inter-atomic interaction.

3.2 Noble gas systems

In liquids and solids composed of noble gas atoms the electronic distribution is only slightly perturbed from the stable atomic closed-shell configuration.

The interaction can be quite well represented by a sum of pair-wise interactions

$$V_{\rm pot}(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N) = \sum_{i=1}^{N}\sum_{j>i}^{N} v_2(r_{ij}) \tag{27}$$

with $r_{ij} =\mid \boldsymbol{r}_j - \boldsymbol{r}_i \mid$. The attractive part of the pair interaction $v(r)$ is caused by mutual polarization of each atom and can be modelled using interacting fluctuating dipoles. It results in an attraction which varies as $1/r^6$ at large inter-nuclear separations. For small separations the potential is strongly repulsive due to overlap between the electron clouds and the Pauli exclusion principle. The most commonly used form for $v_2(r)$ is the Lennard-Jones potential

$$v_{\rm LJ}(r) = 4\epsilon \left[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6} \right] \tag{28}$$

The parameters ϵ and σ measure the strength of the interaction and the radius of the repulsive core, respectively. There is no particular physical reason for choosing the exponent in the repulsive term to be 12, other than the resulting simplicity. Based on physical reasons we may argue in favour of a more steeply rising repulsive potential.

The Lennard-Jones potential has been extensively used in computer simulations and may be viewed as the "hydrogen atom" for computer simulators. It was used by Aneesur Rahman in the first molecular dynamics simulation study using continuous potentials [7]. Often one has concentrated on some more generic features in many-body systems and the particular form of the model potential is in that case not crucial.

	ϵ/k_B (K)	σ (nm)
He	10.2	0.228
Ne	47.0	0.272
Ar	119.8	0.341
Kr	164.0	0.383

Table 1: Suitable Lennard-Jones pair potential parameters for simulation studies [2].

The model describes quite well the interaction in liquids and solids composed of rare gas atoms. The pair potential should be viewed as an effective pair potential which, to some extent, includes contributions from higher order terms. Typical numbers for ϵ and σ used when comparing results from simulations with real systems are given in Tab. 1 [2]. In Fig. 2 the Lennard-Jones potential used for liquid argon in simulations is compared with the "true" pair-potential for two isolated argon atoms with nuclear distance r [2]. The latter has been determined using both theoretical calculations and experimental data and represents the "true" pair-potential for argon. The Lennard-Jones potential used in simulations has been fitted to reproduce

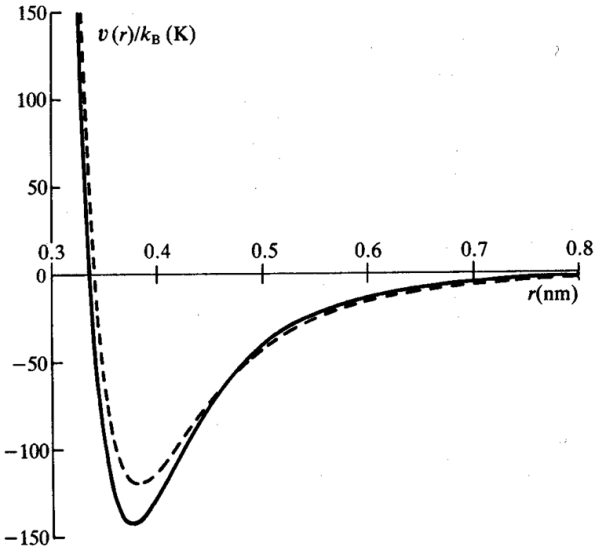

Figure 2: Argon pair potentials [2]. The Lennard-Jones pair potential used in simulation studies (dashed line) compared with the "true" pair potential for two isolated argon atoms (solid line).

liquid data and therefore it incorporates in an approximate way three-body and higher order terms. The difference between the two potentials reflects the magnitude of the non-additive terms in the true many-body potential.

3.2.1 Ionic systems

Ionic systems can also be represented quite well with a sum of pair-potentials. The particles are electrically charged ions and the interaction is dominated by the interionic Coulomb interaction, which varies as the inverse first power of the interionic distance. Special techniques have to be used to sum this long-range interaction in a computer simulation [1]. A common form is [2]

$$v_2(r_{ij}) = \frac{z_i z_j}{r_{ij}} + A_{ij} \exp\left(-B_{ij} r_{ij}\right) - \frac{C_{ij}}{r_{ij}^6} - \frac{D_{ij}}{r_{ij}^8} \tag{29}$$

where z_i is the ionic charge of species i. The first term is the Coulomb interaction, the second repulsive term prevents ions to overlap and the remaining terms represents the weaker dispersion interactions. Appropriate numbers for the potential parameters A_{ij}, B_{ij}, C_{ij} and D_{ij} can be found in the literature.

3.2.2 Metallic systems

Many-body aspects of the inter-atomic interaction is apparent in metals. Assuming a two-body potential the vacancy formation energy is by necessity equal to the sublimation energy, while it is known that in metals the vacancy formation energy is about one third of the sublimation energy. The Cauchy relation for two of the elastic constants $c_{12} = c_{44}$ is satisfied to a good approximation in van der Waals solids and often in ionic crystals but never in metals. A description in terms of two-body potentials always leads to the relation $c_{12} = c_{44}$ for cubic structures. The many-body effects are also clearly present in many properties related to the surfaces of metals.

A surprisingly simple modification of the two-body description seems to capture many of the essential many-body effects in metals. The inter-atomic interaction is written on the form

$$V_{\mathrm{pot}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \sum_{i=1}^{N} g_i(\rho_i) + \sum_{i=1}^{N} \sum_{j>i}^{N} v_2(r_{ij}) \tag{30}$$

where $v_2(r)$ is a usual two-body potential and

$$\rho_i = \sum_{j \neq i}^{N} h_i(r_{ij}) \tag{31}$$

a pair function describing the local environment of atom i in terms of contributions from its neighbours. This model has been called pair functional model [8] or glue model [9]. It can be justified both as an approximation to either the density functional or tight binding theory. In the former case the physical interpretation of the function g_i is related to the local electron density, while in the latter case it is related to the electron bandwidth. Various names are associated with the model as effective medium theory [10], embedded atom method [11], and Finnis-Sinclair model [12].

The pair functional models generalize the two-body description and introduce a truly many-body part in the expression for the inter-atomic interaction. These models have roughly the same computational speed as simple two-body potentials and provide a much improved description of a broad range of inhomogeneous environments. The pair functional models have had a significant impact on the simulation of metals, but they do have limitations. They are most accurate for metals with completely empty or filled d bands, but are less reliable for transition metals, near the center of the transition series. The partial filling of the d bands implies an angular character to the interactions which has turned out to be difficult to incorporate into a simplified description.

3.2.3 Covalent systems

The situation in covalent bonded systems is more complicated. The directional bonding has to be included through angular terms but how that should be done in practise is intricate. Silicon is an important materials where the covalent bonding is predominant. An early attempt to model the condensed phases of silicon was put forward by Stillinger and Weber [13]. Silicon atoms form bonds with the four nearest atoms with a local tetrahedron structure. Stillinger-Weber included an explicit term that forced the angle between nearby triples to stay close to 109 degrees. The potential was written as a sum of a two-body and a three-body term and it was adjusted to the lattice constant, the cohesive energy, the melting point and the liquid structure.

Many other potentials have been suggested for silicon, but the general wisdom is that it is difficult to construct classical potentials which are totally transferable. With transferability is meant that the model should be applicable in various atomic surroundings, in particular in situations which has not been used in the fitting procedure. By explicitly including the electronic degrees of freedom the accuracy of the description of the potential energy surface can be enhanced to the price of more extensive computations.

3.3 The system size

MD simulations are usually performed on systems containing a few thousand particles and sometimes up to a few millions are used. These number of particles are appropriate for clusters of atoms but not for bulk systems. A non-negligible number of atoms will be on the surface of the system and those atoms will have a very different surrounding compared with the bulk atoms. For a simple cubic system about half of the particles are at the surface with $N = 10^3$. For $N = 10^6$ it has decreased to about 6%, still a non-negligible number of particles.

In order to simulate bulk systems it is important to choose boundary conditions that mimics the presence of an infinite bulk surrounding. This is achieved by employing *periodic boundary conditions*. The simulation cell is replicated throughout space to form an infinite lattice (see Fig. 3). The most commonly used simulation box is the cubic box. The volume $V = L^3$ of the box together with the number of particles N defines the number density $n = N/V$. The number density in the central box, and hence in the entire system, is conserved. However, it is not necessary to store the coordinates of all the images, an infinite number, just the particles in the central box have to be stored. When a particle leaves the central box by crossing a "boundary", attention can be switched to the image just entering or one can follow the motion of the particle leaving the central box. A given particle will interact with all other particles in the infinite periodic system,

i.e. all other particles in the same periodic cell as well as all particles in all other cells, including its own periodic images. It is important realize that the boundary of the periodic box itself has no special significance. The origin of the periodic lattice of the primitive cells may be chosen anywhere, and this choice will not affect any property of the model system under study. In contrast, what *is* fixed is the shape of the periodic cell and its orientation.

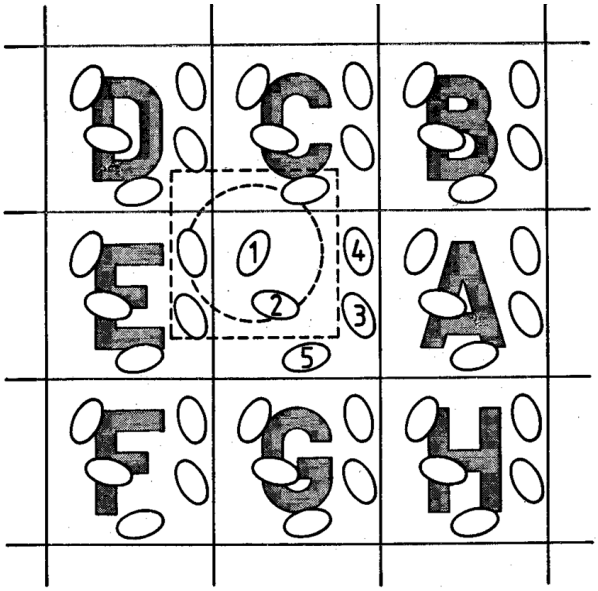

Figure 3: Illustration of the minimum image convention [2]. The central box contains five atoms. The 'box' constructed with atom 1 at its centre, the dashed square, also contain five atoms. It includes all nearest periodic images to particle 1. The dashed circle represents a potential cutoff.

In practise, we are often dealing with relatively short-range pair interactions, as for instance the Lennard-Jones potential. In that case it is permissible to truncate the inter-atomic interaction at a finite radial cut-off distance r_c and a particles will now only interact with a finite number of surrounding particles. In particular, if $r_c < L/2$ it will only interact with at most the nearest periodic image of a surrounding particle. This is called the *minimum image convention*. In Fig. 3 the dashed square encloses all nearest periodic images to the particle located in the middle of the dashed square. The dashed circle represents a potential cutoff. The minimum image convention can be coded in an efficient way. Denote the position of particle i and j with $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$. Before the pair-interaction between i and j is calculated the distance is calculated according to

$$\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$$
$$\boldsymbol{r}_{ij} = \boldsymbol{r}_{ij} - L * [\boldsymbol{r}_{ij}/L]$$

where $[X]$ denotes the nearest integer to X. This code will give the minimum image distance, no matter how many "box lengths" apart the original particles may be. The code works both if the periodic boundary condition has been implemented for the particle positions or if the motion of a particle leaving the central box is explicitly followed.

Systems with long range interactions are computationally more demanding to simulate. This is the case for instance in ionic systems with Coulomb interactions (*cf.* Eq. (29)). Various techniques have been developed as Ewald summation, fast multipole methods, and particle-mesh-based techniques [1].

The simulated system, an infinite periodic system, is not identical to a macroscopic system which it is supposed to mimic. The differences will depend on both the range of the inter-atomic interaction and the phenomena under investigation. In most cases the use of periodic boundary conditions proves to be a surprisingly effective method to simulate homogeneous bulk systems. However, it inhibits the occurrence of long-wavelength fluctuations, which are crucial in the vicinity of a second order phase transition. The same limitation applies to the simulation of long-wavelength phonons in solids. The boundary condition also enforces a certain symmetry for the system. That is devastating if one would like to study symmetry changing phase transitions in solids. For that case methods has been developed that allow the simulation box to dynamically change shape as well as size [14].

4 The time integration algorithm

The core of an MD program is the numerical solution of Newton's equation of motion. We would like to determine a trajectory in $6N$-dimensional phase-space numerically.

4.1 The Verlet algorithm

In 1967 Loup Verlet introduced a central difference based algorithm into molecular simulations [15]. In many cases, this simple algorithm has turned out to be the best to use in MD simulations and it is extensively used.

We can derive the algorithm by making a Taylor expansion of the position coordinate both forward and backward in time,

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 + \ldots \quad (32)$$

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 + \ldots \quad (33)$$

where $\boldsymbol{v}_i = \dot{\boldsymbol{r}}_i$, $\boldsymbol{a}_i = \ddot{\boldsymbol{r}}_i = \boldsymbol{F}_i/m_i$ and Δt the time step in the numerical scheme. By adding these two equations we obtain

$$\mathbf{r}(t+\Delta t) + \mathbf{r}(t-\Delta t) = 2\mathbf{r}(t) + \mathbf{a}(t)\Delta t^2 + \mathcal{O}(\Delta t^4)$$

or

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \mathbf{a}(t)\Delta t^2 + \mathcal{O}(\Delta t^4) \quad (34)$$

This equation is known as the *Verlet algorithm*. The algorithm is properly centered, $\boldsymbol{r}_i(t+\Delta t)$ and $\boldsymbol{r}_i(t-\Delta t)$ play symmetrical roles, making it time-reversible and it shows excellent energy-conserving properties over long times. It is assumed that the forces only depend on the position coordinates. The velocities do not enter explicitly in the algorithm, however, they are needed for estimating the kinetic energy and hence the temperature. They can be obtained by subtracting Eq. (32) with Eq. (33),

$$\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t) = 2\mathbf{v}(t)\Delta t + \mathcal{O}(\Delta t^3)$$

or

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^2) \quad (35)$$

A drawback with the original Verlet algorithm is that it is not "self-starting". Usually the initial conditions are given at the same time, $\boldsymbol{r}_i(0)$ and $\boldsymbol{v}_i(0)$, but for the Verlet algorithm we need the positions at the present $\boldsymbol{r}_i(0)$ and previous $\boldsymbol{r}_i(-\Delta t)$ time steps. There is also some concerns [16] that roundoff errors may arise when implementing Eq. (34).

The original Verlet algorithm, Eqs (34) and (35), does not handle the velocities in a fully satisfactory manner. A Verlet-equivalent algorithm that

store positions, velocities and accelerations at the same time and which also minimize round-off errors was introduced in Ref. [17]. Consider Eq. (33) at the next time step

$$\mathbf{r}(t) = \mathbf{r}(t + \Delta t) - \mathbf{v}(t + \Delta t)\Delta t + \frac{1}{2}\mathbf{a}(t + \Delta t)\Delta t^2 + \ldots$$

By adding these with Eq. (32) and solving for $\boldsymbol{v}_i(t + \Delta t)$ yields

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{1}{2}\left[\mathbf{a}(t) + \mathbf{a}(t + \Delta t)\right]\Delta t \tag{36}$$

Eqs (32) and (36) are referred to as the *velocity Verlet algorithm*. It is equivalent to the original Verlet algorithm and produces the same trajectories. It involves two stages with a force (acceleration) evaluation in between. A time step

$$\mathbf{r}(t), \mathbf{v}(t), \mathbf{a}(t) \rightarrow \mathbf{r}(t + \Delta t), \mathbf{v}(t + \Delta t), \mathbf{a}(t + \Delta t)$$

can be written in the following way:

$$\begin{aligned}
&\mathbf{v}(t + \Delta t/2) = \mathbf{v}(t) + \frac{1}{2}\mathbf{a}(t)\Delta t \\
&\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t + \Delta t/2)\Delta t \\
&\texttt{calculate new accelerations/forces} \\
&\mathbf{v}(t + \Delta t) = \mathbf{v}(t + \Delta t/2) + \frac{1}{2}\mathbf{a}(t + \Delta t)\Delta t
\end{aligned}$$

The algorithm only requires storage of positions, velocities and accelerations at one time point and it can be coded as

$$\begin{aligned}
&v = v + 0.5\ a\ dt \\
&r = r + v\ dt \\
&a = \texttt{accel}(r) \\
&v = v + 0.5\ a\ dt
\end{aligned}$$

where `accel` is a subroutine that returns the accelerations given the positions. The numerical stability, convenience and simplicity makes the velocity Verlet algorithm attractive for MD simulation studies and is highly recommended.

4.2 Accurate predictions and the Lyaponov instability

To obtain accurate predictions using the MD technique is obvious that we need a good numerical algorithm to solve the equation of motion. However, it is not obvious what is meant by a good algorithm.

An intriguing fact is that the systems that we study by MD simulations in general show chaotic behaviour. The trajectory in phase-space depends

sensitively on the initial conditions. This means that two trajectories that initially are very close will diverge exponentially from each other as time progresses. In the same way, any small perturbation, even the tiny error associated with finite precision arithmetic, will tend to cause a computer-generated trajectory to diverge from the true trajectory with which it is initially coincident. This is shown in Fig. 4 where the time dependence of the difference $\sum_{i=1}^{N} \mid \mathbf{r}_i(t) - \mathbf{r}'_i(t) \mid^2$ between a reference trajectory $\mathbf{r}_i(t)$ and a perturbed trajectory $\mathbf{r}'_i(t)$ is shown. The perturbed trajectory differs from the reference trajectory in that one of the particles has been displaced $10^{-10}\sigma$ in one direction. As can been seen in Fig. 4 the measure of the distance increases expontentially with time. Presumably, both the reference trajectory and the perturbed trajectory are diverging from the true solution of Newton's equation of motion. We can not predict the true trajectory over long times in an MD simulation.

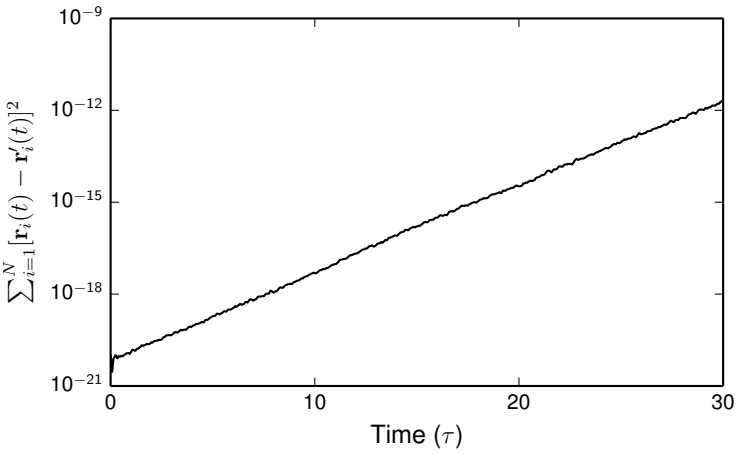

Figure 4: Illustration of the Lyaponov instability in a simulation of a Lennard-Jones system. The figure shows the time dependence of the distance between two trajectories that were initially very close to each other (see text). Note, that within this relatively short time, the two trajectories become essentially uncorrelated.

This seems to be devastating, but it is not. It is important to realize that what we need is to obtain a trajectory that gives accurate predictions of average quantities, such as defined in Eq. (21), and that it provides a time-dependence accurate over times-scales comparable with the correlation times of interest, so that we may correctly calculate correlation functions of interest, such as defined in Eq. (69). Experience has shown that it is possible to obtain such trajectories numerically. MD simulations can be used to obtain accurate results for average properties. If the MD simulation is repeated the same, but not identical, results are obtained.

So what is important? To obtain accurate results from an MD simulation we have to generate a trajectory that conserves the energy very well. That is of primary importance in an MD simulation. The generated trajectory has to stay on the appropriate constant-energy surface in phase-space. Of particular importance is that the algorithm should show absence of long-term energy drift.

The true Hamiltonian dynamics is time reversible. If we reverse the momenta of all particles at a given instant, the system would retrace its trajectory in phase space. Time reversible algorithms seem to show good energy conservation over long time periods and we would like the numerical algorithm to be time reversible. The true Hamiltonian dynamics also leaves the magnitude of any volume element in phase-space unchanged. This volume or area preserving property leads to the Liouville's theorem (112), derived in appendix A. A non-area-conserving algorithm will map a constant-energy-surface area on a different, probably larger area, and that is not consistent with energy conservation. Hence it is likely that reversible, area-preserving algorithms will show little long-term energy drift.

This has lead to a more formal development of efficient algorithms to be used in MD simulations. Based on the geometric structure of Hamilton's equation so called symplectic integrators have been developed. In appendix B the formal development of symplectic algorithms is presented based on the classical time evolution operator approach. A symplectic algorithm is also area preserving. We show that the velocity Verlet algorithm is a time-reversible area-preserving algorithm (see appendix B).

4.3 The time step

The size of the time step has to be determined. The optimum value is a compromise. It should be large to sample as much of phase space as possible, but it has to be small enough to generate trajectories that produce accurate predictions.

One may expect that sophisticated higher-order algorithms would allow the use of large time steps while keeping the error in energy small. However, general higher-order algorithms tend to have very good energy conservation for short times, but the energy drift over long times can be substantial. In contrast, the Verlet algorithm has moderate short time energy conservation but very little energy drift over long times, even with a quite large time step. This seems to be a general feature for time-reversible area-preserving algorithms. Absence of long-term energy drift is important.

Furthermore, the most time-consuming part in the time integration algorithm is the evaluation of the forces. Higher-order algorithms implies the evaluation of forces and higher order derivatives, which often makes those algorithms less efficient. Inter-atomic interactions are also quite steep for short distances. A high-order algorithm with a large time step may generate

configurations with short and therefore unfavourable inter-atomic distances. The conclusion is that the low order time-reversible area-preserving velocity Verlet algorithm is usually the best to use in MD simulations.

A typical size of the time step for an atomic system is a few femtoseconds. With heavier atoms a larger time step can be used. With a tentative value for Δt test runs can be done to study the effect of changing the size of the time step. Test of energy conservation is then vital.

In more complex systems the forces may generate motion with widely different time scales. It could be a molecular system with fast motion associated with the intra-molecular forces and slower inter-molecular motion. Multiple time-step techniques have been developed for those situations. The power of the Liouville operator approach (see appendix B) was demonstrated by Tuckerman *et al.* [18] in their derivation of a multiple time-steps method.

5 Average properties

We now turn to the problem of determining average properties. In MD simulations we perform an average over a finite time interval T_{prod}. This time interval has to be long enough so that a sufficient region of phase-space is explored by the system to yield a satisfactory time average according to Eq. (21). It is also essential to let the system "equilibrate" in time before the actual averaging is performed. At the end of the equilibration period T_{eq} the memory of the initial configuration should have been lost and a typical equilibrium configurations should have been reached. The MD average can therefore be written as

$$A \approx \langle \mathcal{A} \rangle_{\text{MD}} = \frac{1}{\mathsf{T}_{\text{prod}}} \int_{\mathsf{T}_{\text{eq}}}^{\mathsf{T}_{\text{eq}}+\mathsf{T}_{\text{prod}}} \mathcal{A}(t) dt \tag{37}$$

The actual averaging is done as a summation. The integration procedure creates a sequence of values $\{\mathcal{A}_i\}$ with timespacing Δt and

$$\langle \mathcal{A} \rangle_{\text{MD}} = \frac{1}{M} \sum_{i=1}^{M} \mathcal{A}_i \tag{38}$$

where $\mathcal{A}_i = \mathcal{A}(\mathsf{T}_{\text{eq}} + i\Delta t)$ and $M = \mathsf{T}_{\text{prod}}/\Delta t$.

The result $\langle \mathcal{A} \rangle_{\text{MD}}$ from the MD study is subject to systematic and statistical errors. Sources of systematic errors could be system size-dependence or poor equilibration. These errors should be reduced as much as possible. It is also essential to obtain and estimate of the statistical significance of the results. The MD results is obtained over a finite time period and that causes statistical imprecision in the obtained average value. By assuming Gaussian statistics the error can be estimated from the variance of the mean value. If we assume that each data point $\mathcal{A}_i$ is independent, the variance of the mean is simply given by $\sigma^2(\langle \mathcal{A} \rangle_{\text{MD}}) = \sigma^2(\mathcal{A})/M$. However, the data points are not independent, they are highly correlated, and the number of uncorrelated data points

$$M_{\text{eff}} = M/s \tag{39}$$

has to be estimated, where s is called the *statistical inefficiency*, the number of time steps for which the correlations effectively persist. The variance of the mean value can then be written as

$$\sigma^2(\langle \mathcal{A} \rangle_{\text{MD}}) = \frac{1}{M/s} \sigma^2(\mathcal{A}) \tag{40}$$

with

$$\sigma^2(\mathcal{A}) = \left\langle \delta\mathcal{A}^2 \right\rangle_{\text{MD}} = \frac{1}{M} \sum_{i=1}^{M} (\mathcal{A}_i - \langle \mathcal{A} \rangle_{\text{MD}})^2 = \left\langle \mathcal{A}^2 \right\rangle_{\text{MD}} - \langle \mathcal{A} \rangle_{\text{MD}}^2 \tag{41}$$

Two different methods to estimate the statistical inefficiency s are to determine the correlation function or to perform block averaging. These two techniques are presented in appendix C. We can now write our final result as

$$A = \langle \mathcal{A} \rangle_{\mathrm{MD}} \pm \frac{1}{\sqrt{M/s}} \sigma(\mathcal{A}) \tag{42}$$

Here, the true value will be within the given error bars with 68% probability. If we instead use two standard deviations the probability increases to 95%.

5.1 Kinetic, potential and total energies

The instantaneous kinetic and potential energies are given by

$$\mathcal{E}_{\mathrm{kin}}(t) = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2(t)}{2m_i} \tag{43}$$

and

$$\mathcal{E}_{\mathrm{pot}}(t) = V_{\mathrm{pot}}(\boldsymbol{r}_1(t), \ldots, \boldsymbol{r}_N(t)) \tag{44}$$

Hence

$$E_{\mathrm{kin}} = \langle \mathcal{E}_{\mathrm{kin}}(t) \rangle_{\mathrm{time}} = \left\langle \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2(t)}{2m_i} \right\rangle_{\mathrm{time}} \tag{45}$$

and

$$E_{\mathrm{pot}} = \langle \mathcal{E}_{\mathrm{pot}}(t) \rangle_{\mathrm{time}} = \langle V_{\mathrm{pot}}(\boldsymbol{r}_1(t), \ldots, \boldsymbol{r}_N(t)) \rangle_{\mathrm{time}} \tag{46}$$

The total energy

$$E = E_{\mathrm{kin}} + E_{\mathrm{pot}} \tag{47}$$

is a conserved quantity.

5.2 Temperature

In appendix D microscopic expressions for the temperature and pressure were derived within the canonical ensemble. The instantaneous expression for the temperature

$$\mathcal{T}(t) = \frac{2}{3Nk} \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2(t)}{2m_i}$$

is used to determine the temperature for the system

$$kT = \frac{2}{3N} \left\langle \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2(t)}{2m_i} \right\rangle_{\mathrm{time}} = \frac{2}{3N} \langle \mathcal{E}_{\mathrm{kin}} \rangle_{\mathrm{time}} \tag{48}$$

All $3N$ momentum coordinates are usually not independent in a MD simulation. The total momentum is conserved and those three degrees of freedom

should be subtracted. Only the internal momenta contribute to the temperature. The definition of temperature then becomes

$$kT = \frac{2}{3N-3} \langle \mathcal{E}_{\text{kin}} \rangle_{\text{time}} \tag{49}$$

Using the microcanonical ensemble one can show [19] that

$$kT = \frac{2}{3N-5} \left\langle \frac{1}{\mathcal{E}_{\text{kin}}} \right\rangle_{\text{time}}^{-1} \tag{50}$$

The difference between the definitions in Eqs (48), (49), and (50) is of the order $1/N$ and for large systems they will give essentially the same result. For convenience, we will stick to the definition in Eq. (48).

5.3 Pressure

The microscopic expression for the pressure, derived in appendix D, is

$$\mathcal{P}(t) = \frac{1}{3V} \sum_{i=1}^{N} \left[\frac{\boldsymbol{p}_i^2(t)}{m} + \boldsymbol{r}_i(t) \cdot \boldsymbol{F}_i(t) \right]$$

and

$$PV = NkT + W$$

where

$$W = \langle \mathcal{W}(t) \rangle_{\text{time}} = \left\langle \frac{1}{3} \sum_{i=1}^{N} \boldsymbol{r}_i(t) \cdot \boldsymbol{F}_i(t) \right\rangle_{\text{time}}$$

is the virial function.

For pair-wise interactions it is convenient to rewrite the virial $\mathcal{W}$. The force on particle i can be written as a sum of contributions from the surrounding particles

$$\boldsymbol{F}_i = \sum_{j \neq i} \boldsymbol{f}_{ij}$$

where $\boldsymbol{f}_{ij}$ is the force from particle j acting on particle i. We then have

$$\sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i = \sum_i \sum_{j \neq i} \boldsymbol{r}_i \cdot \boldsymbol{f}_{ij} = \frac{1}{2} \sum_i \sum_{j \neq i} [\boldsymbol{r}_i \cdot \boldsymbol{f}_{ij} + \boldsymbol{r}_j \cdot \boldsymbol{f}_{ji}]$$

where the second equality follows because the indices i and j are equivalent. Newton's third law can then be used and

$$\sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i = \frac{1}{2} \sum_i \sum_{j \neq i} \boldsymbol{r}_{ij} \cdot \boldsymbol{f}_{ji} = \sum_i \sum_{j > i} \boldsymbol{r}_{ij} \cdot \boldsymbol{f}_{ji}$$

where $\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$. The force can be expressed in terms of the pair-potential $v(r)$

$$\boldsymbol{f}_{ji} = -\nabla_{\boldsymbol{r}_{ij}} v(\boldsymbol{r}_{ij})$$

This implies that the virial can be expressed as

$$\mathcal{W} = \frac{1}{3} \sum_i \sum_{j>i} \boldsymbol{r}_{ij} \cdot \boldsymbol{f}_{ji} = -\frac{1}{3} \sum_i \sum_{j>i} w(r_{ij}) \tag{51}$$

where the pair virial function $w(r)$ is

$$w(r) = r \frac{dv(r)}{dr} \tag{52}$$

6 A program

The most important input in a MD simulation is a description of the inter-particle interaction. The Lennard-Jones model has been used extensively for atomic systems. The total potential energy is represented as a sum of pair-wise interactions, the Lennard-Jones potential

$$v_{\mathrm{LJ}}(r) = 4\epsilon \left[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6} \right] \tag{53}$$

It is common to use ϵ and σ as units for energy and length, respectively. By further using the mass m as unit for mass the time unit becomes $\tau = \sqrt{m\sigma^2/\epsilon}$. In applications, it is customary to introduce a cutoff radius r_c and disregard the interactions between particles separated by more than r_c. A simple truncation of the potential creates problem with test of energy conservation. Whenever a particle "crosses" the cutoff distance, the potential energy makes a jump and spoil energy conservation. It is therefore common to use a truncated and shifted potential

$$v(r) = \begin{cases} v_{\mathrm{LJ}}(r) - v_{\mathrm{LJ}}(r_c) & \text{if } r < r_c \\ 0 & \text{if } r > r_c \end{cases} \tag{54}$$

The energy then becomes continuous, however not the forces.

6.1 Initialization

In the initialization part the number of particles, the size of the simulation box and the initial positions and velocities for all particles are selected. For a crystalline solid the atoms are usually placed according to the lattice structure. For a liquid it is often most convenient to place the atoms on a lattice to circumvent overlap between atoms. On can then either introduce small deviations from the regular lattice positions and/or give the particles some random initial velocities. This will fix the total energy for the system. It is also convenient to enforce that the total linear momentum is zero in the initial configuration.

6.2 Equilibration

The initialization procedure fixes the external parameters for the system; the total energy, the volume, the particle number and the total linear momentum, which define the thermodynamic state for the system. However, the initial configuration will not be in a typical equilibrium configuration and one has to allow for a certain number of time-steps in the integration procedure before the actual simulation can be started. This corresponds to equilibrating the system.

The equilibration time can be a substantial part of the total simulation time. Some system will never equilibrate and find its true equilibrium state

on the computer, as glassy states and various metastable phases. To determine the time for equilibration is not easy. One can investigate various thermodynamic quantities and look for approach to constant, time-independent values. In Fig. 5 this is illustrated for the potential and kinetic energies. Notice that the total energy is conserved.

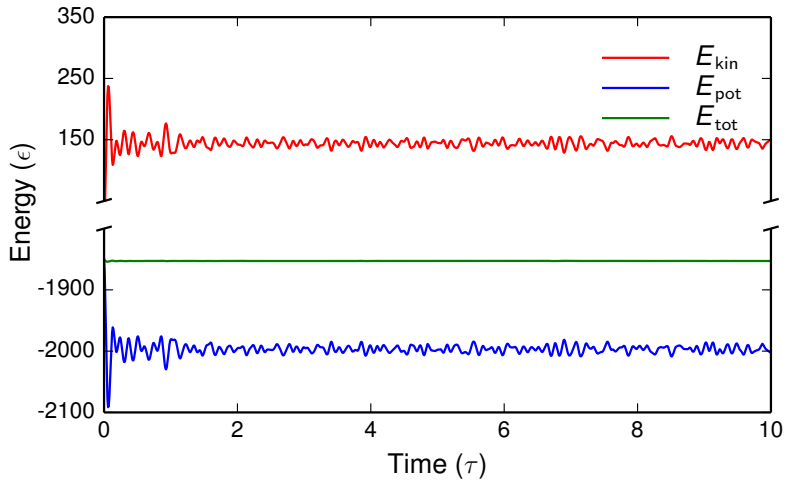

Figure 5: The initial time-evolution for the potential, kinetic and total energies. The data are taken from a MD simulation of a Lennard-Jones system.

Often one would like to reach another thermodynamic state, not the one specified in the initial configuration. One can then scale the velocities and/or the box-size to change the temperature and pressure for the system. Techniques to perform scaling is discussed in App. E. In Fig. 6 we show the result using the scaling suggested in App. E.

6.3 Production

Next follows the actual simulation where the output data are produced. The equation of motion is solved for a large number of time-steps. The most time-consuming part is the force evaluation, not the actual stepping in the integration procedure. It is recommended to store the configurations from the production run. That amounts to a huge set of data. Some properties are often calculated during the actual production simulation but often one would like to reanalyse the data and determine properties which were not considered during the production run. Too much direct evaluations of various properties during the simulation may also make the simulation program unnecessary complicated. It is also recommended to store sufficient information at the end of the simulation which could be used to restart the simulation.

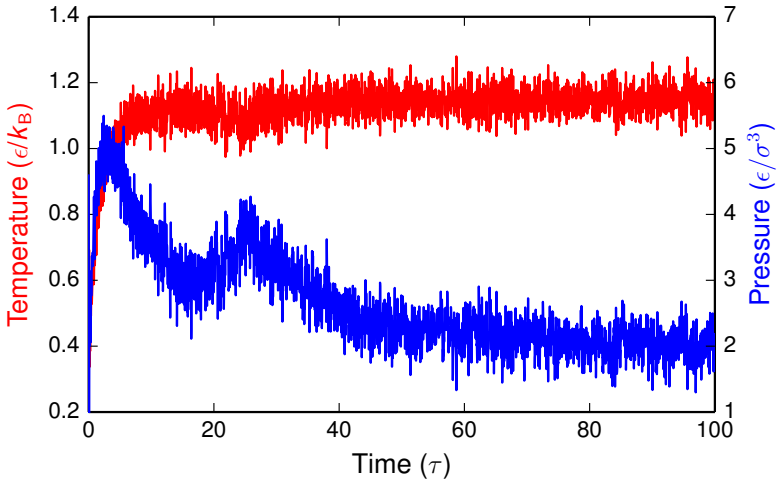

Figure 6: The time-evolution of the instantaneous temperature and pressure where these are scaled to the temperature 1.15 ϵ/k_B and pressure 2.0 ϵ/σ^3. The data are taken from a MD simulation of a Lennard-Jones system.

6.4 Analysis

In the last step one analyses the output data. The quantities that have not been determined during the actual simulation can be evaluated provided the configurations have been stored. Various thermodynamic quantities, structural properties, time-dependent correlation functions and transport coefficients can all be determined from a simulation of an equilibrium system. An important aspect is evaluation of proper error bars. The results may be subjected to both systematic and statistical errors. Systematic errors should be eliminated where possible and statistical errors have to be estimated. How the statistical error can be estimated is discussed in App. C.

7 Static properties

We now turn to the problem of analysing the results from the MD simulation, to obtain properties from the generated phase-space trajectory. We can divide this into static and dynamic properties. Static properties will depend on one time-point while dynamic properties depend on correlations between two different time-points. The latter will be treated in the next section.

Static properties can be divided into two different categories, sometimes called *mechanic* and *entropic* quantities. Mechanic properties can be directly expressed as an average over some phase-space quantity $\mathcal{A}(\mathbf{x})$, as in Eq. (21), while entropic quantities are more difficult to evaluate.

The static structure of matter is measured in diffraction experiments. This can be related to the pair distribution function which describes the spatial organization of particles about a central particle. This function is important in the theory for dense fluids, disordered materials and it can be used to identify the lattice structure of a crystalline solid.

7.1 Mechanic properties

7.1.1 Simple average properties

We have already introduced several mechanic average properties in Sec. 5, as for instance the potential energy

$$E_{\text{pot}} = \langle V_{\text{pot}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \rangle_{\text{time}}$$

and the virial

$$W = -\frac{1}{3} \left\langle \sum_{i=1}^{N} \boldsymbol{r}_i \cdot \nabla_i V_{\text{pot}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \right\rangle_{\text{time}}$$

which is related to the force, the first order derivative of the potential. We can also introduce an effective force constant k_{eff}, given by the second order derivative of the potential

$$k_{\text{eff}} = \left\langle \frac{1}{N} \sum_{i=1}^{N} \nabla_i^2 V_{\text{pot}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \right\rangle_{\text{time}} \tag{55}$$

Properties related to the momenta have also been introduced, as the kinetic energy

$$E_{\text{kin}} = \left\langle \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2(t)}{2m} \right\rangle_{\text{time}}$$

and the temperature

$$T = \frac{2}{3Nk_{\text{B}}} \left\langle \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2(t)}{2m} \right\rangle_{\text{time}}$$

7.1.2 Fluctuations

One can also study fluctuations from the average quantities. These are related to thermodynamic response functions as heat capacities, compressibilities and thermal expansion. In deriving these relations using statistical mechanics one has to be careful with the type of ensemble that is used. For average quantities, *cf.* Eq. (38), the result is independent on the ensemble used for large systems ($N \to \infty$). However, the fluctuations, *cf.* Eq. (41), will depend on the ensemble used, even for large systems. For instance, using the canonical ensemble where N, V and T are kept constant, the heat capacity at constant volume is given by the fluctuations in the total energy $\mathcal{E} \equiv \mathcal{E}_{\text{kin}} + \mathcal{E}_{\text{pot}}$ according to

$$C_V = \frac{1}{k_{\text{B}} T^2} \left\langle (\delta \mathcal{E})^2 \right\rangle_{NVT} \tag{56}$$

This can not be true in the micro-canonical ensemble where the total energy is kept fix by definition. In the micro-canonical ensemble one can show that

$$\left\langle (\delta \mathcal{E}_{\text{kin}})^2 \right\rangle_{NVE} = \left\langle (\delta \mathcal{E}_{\text{pot}})^2 \right\rangle_{NVE} = \frac{3Nk_{\text{B}}^2 T^2}{2} \left(1 - \frac{3Nk_{\text{B}}}{2C_V} \right)$$

This can be used to determine the heat capacity in a MD simulation (at constant N, V and E) from the fluctuations in the potential energy

$$C_V = \frac{3Nk_{\text{B}}}{2} \left[1 - \frac{2}{3Nk_{\text{B}}^2 T^2} \left\langle (\delta \mathcal{E}_{\text{pot}})^2 \right\rangle_{NVE} \right]^{-1} \tag{57}$$

or from the fluctuations in the kinetic energy

$$C_V = \frac{3Nk_{\text{B}}}{2} \left[1 - \frac{2}{3Nk_{\text{B}}^2 T^2} \left\langle (\delta \mathcal{E}_{\text{kin}})^2 \right\rangle_{NVE} \right]^{-1} \tag{58}$$

7.2 Entropic properties

Entropic properties as entropy free energies and chemical potentials are more difficult to evaluate. They can not be directly expressed as a time average over some phase-space quantity. Instead they are related to the phase-space volume. Special techniques have been developed and we refer to Refs. [1] and [2].

7.3 Static structure

7.3.1 Pair distribution function

Detailed information on the structure of an atomic system can be obtained from different static distribution functions. The most important is the pair

distribution function

$$g(\boldsymbol{r}', \boldsymbol{r}'') = \frac{1}{n^2} \left\langle \sum_{i=1}^{N} \sum_{j \neq i}^{N} \delta(\boldsymbol{r}' - \boldsymbol{r}_i(t)) \delta(\boldsymbol{r}'' - \boldsymbol{r}_j(t)) \right\rangle$$

which is equal to the probability of finding a particle at $\boldsymbol{r}''$ provided there is another particle at $\boldsymbol{r}'$, and relative to the probability expected for a completely random distribution at the same density. In the case of spatially homogeneous systems, only relative separations are meaningful, leading to sum over atom pairs

$$g(\boldsymbol{r}) = \frac{V}{N^2} \left\langle \sum_{i=1}^{N} \sum_{j \neq i}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_{ij}(t)) \right\rangle \tag{59}$$

where $\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$ and $\boldsymbol{r} = \boldsymbol{r}'' - \boldsymbol{r}'$. If the system is also isotropic the function can be averaged over angles without loss of information. The result is the radial distribution function $g(r)$. For large separations between $\boldsymbol{r}'$ and $\boldsymbol{r}''$ the probability to find a particle at $\boldsymbol{r}'$ and at $\boldsymbol{r}''$, respectively, becomes independent of each other and

$$\begin{aligned} g(r \to \infty) &= \frac{1}{n^2} \left\langle \sum_{i=1}^{N} \delta(\boldsymbol{r}' - \boldsymbol{r}_i(t)) \right\rangle \left\langle \sum_{j \neq i}^{N} \delta(\boldsymbol{r}'' - \boldsymbol{r}_j(t)) \right\rangle \\ &= \frac{N(N-1)}{n^2} \frac{1}{V^2} = 1 - \frac{1}{N} \simeq 1 \end{aligned}$$

The quantity $ng(r)$ is equal to the average density of particles at r, given that a particle is located at the origin. By integrating the radial distribution function one obtains the number of particles within a distance r_m from the particle at the origin

$$I(r_m) = n \int_0^{r_m} g(r) 4\pi r^2 dr \tag{60}$$

This value can be used to define the coordination number by choosing r_m equal to the first minimum of $g(r)$.

Calculation algorithm The radial distribution function $g(r)$ can be determined by evaluating the distance between all pairs of particles and sort these into a histogram. Assume that each bin has the size Δr,

$$r_k = (k - \frac{1}{2})\Delta r \quad ; \quad k = 1, 2, \ldots \tag{61}$$

and denote the average number of particles whose distance from a given particle lies in the interval $[(k-1)\Delta r, k\Delta r]$ with $\langle N(r_k) \rangle$. This average is done

over all identical particles in the system together with a time-average along the phase-space trajectory. The average number of particles in the spherical shell assuming a completely random distribution at the same density is

$$\begin{aligned} N^{ideal}(r_k) &= \frac{(N-1)}{V}\frac{4\pi}{3}[k^3-(k-1)^3]\Delta r^3 \\ &= \frac{(N-1)}{V}\frac{4\pi}{3}[3k^2-3k+1]\Delta r^3 \end{aligned} \tag{62}$$

and the radial distribution function can be evaluated as

$$g(r_k) = \frac{\langle N(r_k)\rangle}{N^{ideal}(r_k)} \tag{63}$$

By using the factor $(N-1)/V$ in the definition of $N^{ideal}(r_k)$ we ensure that $g(r_k)$ approaches exactly one when k becomes large.

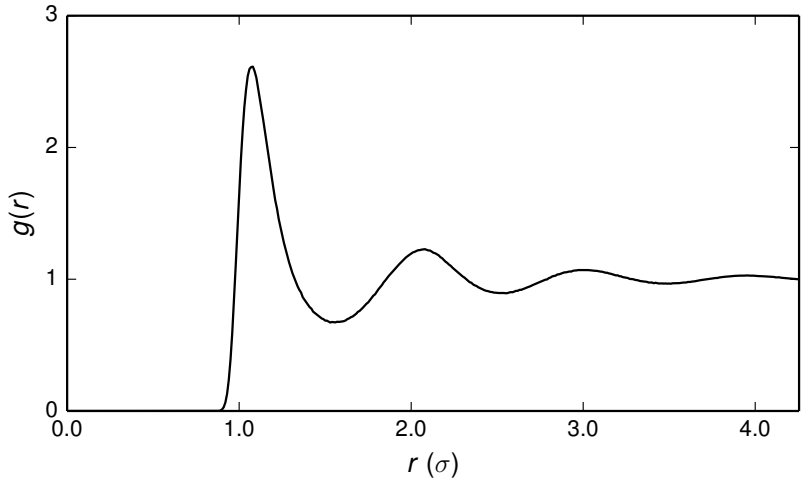

Figure 7: The pair distribution function for a Lennard-Jones system in the liquid phase. Data taken from a MD simulation study.

7.3.2 Static structure factor

The measured intensity in a diffraction experiment is proportional to the so called static structure factor

$$S(\boldsymbol{q}) = \frac{1}{N}\left\langle \mid \sum_{i=1}^{N} e^{i\boldsymbol{q}\cdot\boldsymbol{r}_i(t)} \mid^2 \right\rangle \tag{64}$$

In a crystalline solid it will contain delta-function contributions at the reciprocal lattice vectors. In a liquid it will only contain a delta-function

contribution for $\boldsymbol{q} = 0$, forward scattering. We will consider the liquid case. The static structure factor can be expressed as the Fourier transform of $g(\boldsymbol{r})$ according to

$$\begin{aligned} S(\boldsymbol{q}) &= \frac{1}{N}\left\langle \sum_i \sum_j e^{-i\boldsymbol{q}\cdot(\boldsymbol{r}_j(t)-\boldsymbol{r}_i(t))} \right\rangle \\ &= 1 + \frac{1}{N}\left\langle \sum_i \sum_{j\neq i} e^{-i\boldsymbol{q}\cdot(\boldsymbol{r}_j(t)-\boldsymbol{r}_i(t))} \right\rangle \\ &= 1 + n\int g(\boldsymbol{r}) e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} d\boldsymbol{r} \\ &= 1 + n\int [g(\boldsymbol{r}) - 1] e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} d\boldsymbol{r} + n(2\pi)^3\delta(\boldsymbol{q}) \end{aligned} \tag{65}$$

Conversely, $g(\boldsymbol{r})$ is given by the Fourier transform of $S(\boldsymbol{q})$

$$g(\boldsymbol{r}) = \frac{1}{n}\frac{1}{(2\pi)^3}\int [S(\boldsymbol{q}) - 1] e^{i\boldsymbol{q}\cdot\boldsymbol{r}} d\boldsymbol{q} \tag{66}$$

For an isotropic system, neglecting the delta-function contribution, Eq. (65) reduces to

$$S(q) = 1 + 4\pi n\int r^2 [g(r) - 1]\frac{\sin qr}{qr} dr \tag{67}$$

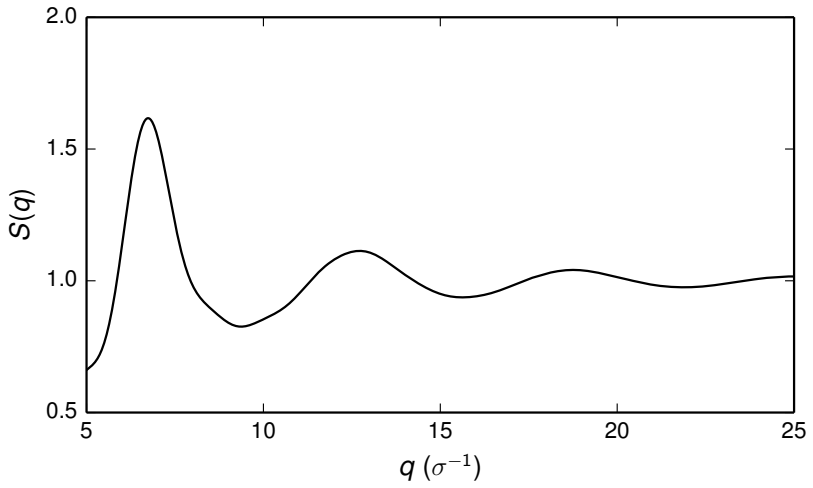

Figure 8: The static structure factor for a Lennard-Jones system in the liquid phase. Data taken from a MD simulation study.

Calculation algorithm The static structure factor for an isotropic system can be calculated from the radial distribution function $g(r)$ according to Eq. (67). However, the upper limit for the integration is limited to half the length of the periodic simulation cell. This often prevents an accurate Fourier transform to be performed. One can try to extend $g(r)$ to larger distances, but that involves some further modeling [2]. Alternatively, one can stick to the definition in Eq. (64). This can be rewritten as

$$\begin{aligned} S(\boldsymbol{q}) &= \frac{1}{N}\left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} e^{-i\boldsymbol{q}\cdot(\boldsymbol{r}_i(t)-\boldsymbol{r}_j(t))} \right\rangle \\ &= \frac{1}{N}\left\langle \left(\sum_{i=1}^{N}\cos\left[\boldsymbol{q}\cdot\boldsymbol{r}_i(t)\right]\right)^2 + \left(\sum_{i=1}^{N}\sin\left[\boldsymbol{q}\cdot\boldsymbol{r}_i(t)\right]\right)^2 \right\rangle \end{aligned} \quad (68)$$

$S(\mathbf{q})$ is then evaluated on a 3-dimensional $\boldsymbol{q}$-grid. The grid has to be consistent with the periodic boundary conditions, *i.e.* $\boldsymbol{q} = (2\pi/L)(n_x, n_y, n_z)$ where n_x, n_y and n_z are integers. To obtain S(q) in 1-dimensional q-space a spherical average has to be done.

8 Dynamic properties

We now turn to dynamic properties, time correlation functions and transport coefficients.

8.1 Time-correlation function

A time-correlation function measures how the value of some dynamic quantity $\mathcal{A}$ at time t' depends on some other (or the same) dynamic quantity $\mathcal{B}$ at time t''. For a system at equilibrium the correlation function will only depend on the time difference $t = t'' - t'$, the tim lag, not on the separate times t' and t''. We define the time-correlation function as an ordinary time-average according to

$$C_{\mathcal{AB}}(t) = \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \int_0^{\mathsf{T}} \mathcal{B}(t+t')\mathcal{A}(t')dt' = \left\langle \mathcal{B}(t+t')\mathcal{A}(t') \right\rangle_{\text{time}} \tag{69}$$

If $\mathcal{A}$=$\mathcal{B}$ it is called an auto-correlation function and[1]

$$C(t) = \left\langle \mathcal{A}(t+t')\mathcal{A}(t') \right\rangle \tag{70}$$

For large time differences the quantities usually become uncorrelated and $C(t \to \infty) = \langle \mathcal{A} \rangle \langle \mathcal{A} \rangle$. If $\langle \mathcal{A} \rangle \neq 0$ it is convenient to define the correlation function in terms of the fluctuations

$$\delta\mathcal{A}(t) = \mathcal{A}(t) - \langle \mathcal{A} \rangle$$

instead, *i.e.*

$$C(t) = \left\langle \delta\mathcal{A}(t+t')\, \delta\mathcal{A}(t') \right\rangle \tag{71}$$

The correlation function then has the initial value

$$C(0) = \left\langle \delta\mathcal{A}^2 \right\rangle$$

and it decays to zero for large times

$$C(t \to \infty) = 0$$

The time integral defines a correlation or relaxation time τ_{rel} according to

$$\tau_{\text{rel}} = \frac{1}{C(0)} \int_0^\infty C(t)dt \tag{72}$$

provided the correlation function decays sufficiently rapidly. For a system in equilibrium the correlation function is stationary, *i.e.* it is invariant under a translation of the time origin,

$$C(t) = \left\langle \mathcal{A}(t+t')\mathcal{A}(t') \right\rangle = \left\langle \mathcal{A}(t+t'+s)\mathcal{A}(t'+s) \right\rangle$$

[1] For convenience, we will drop the subscript "time" in most cases.

If we let $s = -t$ it is easy to show[2] that $C(t)$ is an even function in time

$$C(-t) = C(t) \tag{73}$$

Correlation functions can be divided into two classes: the one-particle functions, in which the dynamic quantity $\mathcal{A}(t)$ is a property of individual particles, and the collective functions, in which $\mathcal{A}(t)$ depends on the accumulated contributions from all particles in the system.

Calculational algorithm The direct approach to the calculation of time-correlation functions is based directly on the definition Eq. (70) (or Eq. (71)). Assume that $\mathcal{A}(t)$ is available at equal intervals of time

$$\mathcal{A}_m \equiv \mathcal{A}(m\Delta\tau) \quad ; \quad m = 0, 1, \ldots, (M-1)$$

of size $\Delta\tau$. Typically $\Delta\tau$ is a small multiple of the time step Δt used in the simulation. The correlation function can then be expressed as

$$C_l = \frac{1}{M-l} \sum_{m=0}^{M-l-1} \mathcal{A}_{m+l}\mathcal{A}_m \quad ; \quad l = 0, 1, \ldots, (M-1) \tag{74}$$

We average over M time origins, but the statistical accuracy becomes less for large time lags $t = l\ \Delta\tau$. Each successive data point is used as time origin. This is not necessary, and indeed may be inefficient, since successive origins will be highly correlated. Also the total simulation time is often considerable longer than the typical relaxation time for the correlation function. The index l in Eq. (74) can then be restricted to a number considerable less than M.

8.1.1 Velocity correlation function

A typical example of a one-particle correlation function is the velocity autocorrelation function

$$\Phi(t) = \big\langle \boldsymbol{v}_i(t+t') \cdot \boldsymbol{v}_i(t') \big\rangle$$

It describes how the motion of a tagged particle is correlated with itself after a time t. If particles are identical the statistical precision of a calculation can be improved by averaging over those particles

$$\Phi(t) = \frac{1}{N} \sum_{i=1}^{N} \big\langle \boldsymbol{v}_i(t+t') \cdot \boldsymbol{v}_i(t') \big\rangle \tag{75}$$

For a system in equilibrium the initial value is given by the equipartition theorem

$$\Phi(0) = \frac{3k_{\mathrm{B}}T}{m}$$

[2]We consider only classical dynamic variables $\mathcal{A}(t)$, so that they commute with themselves at different times.

and for long times the initial and final velocities are expected to be completely uncorrelated and

$$\Phi(t \to \infty) = 0$$

The short time behaviour can be obtained by making a Taylor expansion of $\boldsymbol{v}_i(t)$ around $t = 0$. For the velocity correlation function we get

$$\Phi(t) = \frac{3k_{\mathrm{B}}T}{m}[1 - \frac{(\omega_{\mathrm{E}}t)^2}{2} + \ldots] \tag{76}$$

with

$$\omega_{\mathrm{E}}^2 = \frac{1}{3mk_{\mathrm{B}}T} \left\langle |\, \boldsymbol{F}_i(t) \,|^2 \right\rangle \tag{77}$$

where $\boldsymbol{F}_i(t)$ is the force acting on the tagged particle. The frequency ω_E is an effective frequency known as the *Einstein frequency*. It can be rewritten as [2]

$$\omega_{\mathrm{E}}^2 = \frac{1}{3m} \left\langle \nabla^2 V_{\mathrm{pot}} \right\rangle \tag{78}$$

which shows that it represents the frequency at which the tagged particle would vibrate if it was undergoing small oscillations in the potential well produced by the surrounding particles when they are maintained at their mean equilibrium positions around the tagged particle. If the interaction potential V_{pot} is a sum of pairwise interactions with a pair potential $v(r)$

$$\omega_{\mathrm{E}}^2 = \frac{n}{3m} \int \nabla^2 v(r) g(r) d\boldsymbol{r} \tag{79}$$

where $n = N/V$ is the mean density.

According to the transport theory of Boltzmann and Enskog the velocity correlation function should decay exponentially at long times. It then came as a big surprise when Alder and Wainwright [20] found in a molecular dynamics simulation of a dense gas of hard spheres that the velocity correlation function did not decay exponentially but more as a power law at long times. Alder and Wainwright analysed their results and showed that the tagged particle created a vortex pattern in the velocity field of the nearby surrounding particles that made the velocity correlation function to develop a long-time tail. Further analysis using theory and simulations showed that [21].

$$\Phi(t \to \infty) \propto t^{-3/2} \tag{80}$$

in three dimensions. The magnitude of the long-time tails is very small and it is not easy to detect them in a molecular simulation study. The discovery of the long-time tails raised basic questions such as the assumption about molecular chaos in the derivation of Boltzmann's transport equation. The foundation of hydrodynamics, which is based on the existence of two time regimes, a short-time regime where molecular molecular relaxation takes and a long-time regime where only macroscopic relaxation is important, also had to be clarified.

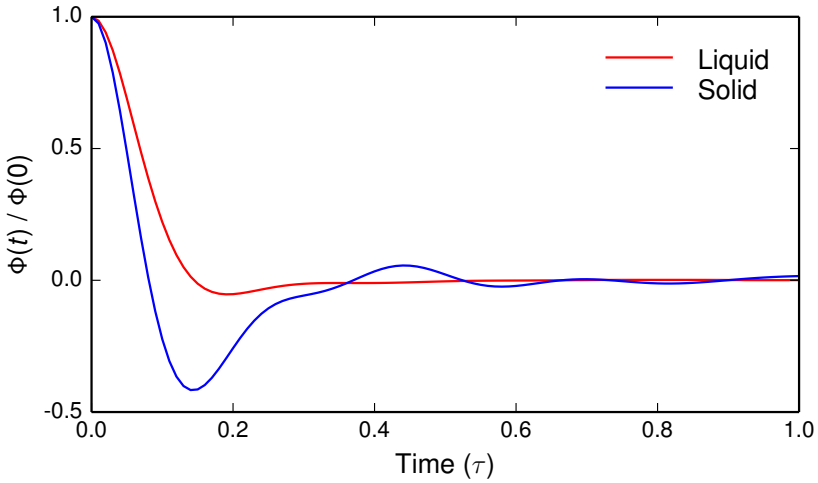

Figure 9: The velocity correlation function. Data taken from a MD simulation study of a liquid and a solid Lennard Jones system.

8.2 Spectral function and power spectrum

It is also useful to consider the Fourier transform of a correlation function, the *spectral function*, defined by

$$\hat{C}(\omega) = \int_{-\infty}^{\infty} C(t) e^{i\omega t} dt \tag{81}$$

where $C(t) = \langle \mathcal{A}(t+t')\mathcal{A}(t') \rangle$. The spectral function can be directly related to the power spectrum of the corresponding dynamic quantity $\mathcal{A}(t)$ through what is known as the Wiener-Khintchin's theorem (see appendix F). Let

$$\mathcal{A}_{\mathsf{T}}(t) = \begin{cases} \mathcal{A}(t) & -\mathsf{T}/2 < t < \mathsf{T}/2 \\ 0 & \text{otherwise} \end{cases}$$

Introduce the Fourier transform

$$\hat{\mathcal{A}}(\omega) = \int_{-\infty}^{\infty} \mathcal{A}_{\mathsf{T}}(t)\, e^{i\omega t}\, \mathrm{d}t$$

The power spectrum is then defined as

$$\mathcal{P}_{\mathcal{A}}(\omega) = \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \mid \hat{\mathcal{A}}(\omega) \mid^2 \tag{82}$$

The Wiener-Khintchin's theorem states that power spectrum is equal to the spectral function $\hat{C}(\omega)$, the Fourier transform of the corresponding correlation function $C(t)$,

$$\mathcal{P}_{\mathcal{A}}(\omega) = \hat{C}(\omega) \tag{83}$$

or equivalently, the correlation function $C(t)$ is equal to the inverse Fourier transform of the corresponding power spectrum

$$C(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{P}_{\mathcal{A}}(\omega) e^{-i\omega t} \, \mathrm{d}\omega \tag{84}$$

Calculational algorithm The Wiener-Khintchin's theorem can be used to calculate the time correlation function in an efficient way by making use of the fast Fourier Transform (FFT) algorithm. The method is often called the Fast Correlation Algorithm (see appendix G). In using the discrete version of the Wiener-Khintchin's theorem, the circular correlation theorem Eq. (170), some care has to be taken, the circular correlation theorem assumes periodic input data. Otherwise, the Fast Correlation Algorithm is a direct application of the relation between the correlation function and the corresponding power spectrum in Eq. (84).

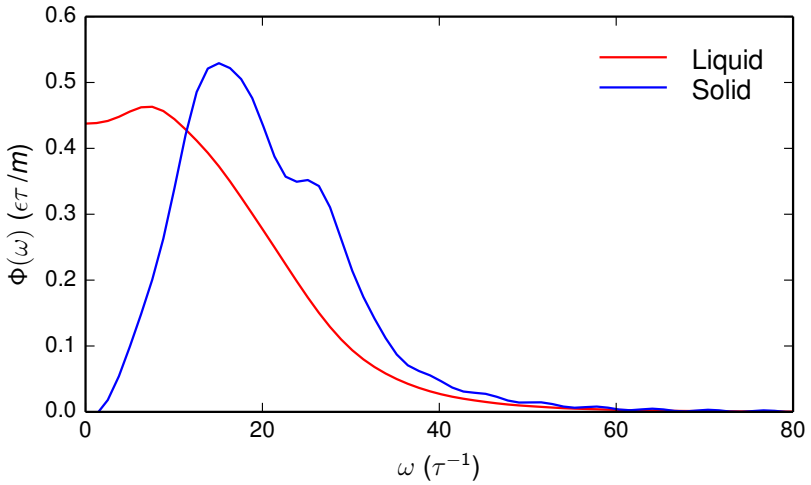

Figure 10: The spectral function, the Fourier transform of the velocity correlation function. Data taken from a MD simulation study of a liquid and a solid Lennard Jones system.

8.3 Space-time correlation functions

Time-correlation functions can be generalised to correlation functions in space and time. Consider the microscopic number density

$$n(\boldsymbol{r}, t) = \sum_{i=1}^{N} \delta[\boldsymbol{r} - \boldsymbol{r}_i(t)]$$

with the average defined by

$$\langle n(\boldsymbol{r},t)\rangle = \lim_{\mathsf{T}\to\infty}\frac{1}{\mathsf{T}}\int_0^{\mathsf{T}} dt\ \frac{1}{V}\int_V d\boldsymbol{r}\ n(\boldsymbol{r},t)$$

and given by

$$\langle n(\boldsymbol{r},t)\rangle = \frac{N}{V} = n$$

As a generalisation of an ordinary time-correlation function (cf. Eq. (71)), we construct the density-density correlation

$$G(\boldsymbol{r}'',\boldsymbol{r}';t'',t') = \frac{1}{n}\left\langle n(\boldsymbol{r}'',t'')n(\boldsymbol{r}',t')\right\rangle$$

For a homogeneous system at equilibrium it will depend only on the space and time differences, $\boldsymbol{r}\equiv\boldsymbol{r}''-\boldsymbol{r}'$ and $t\equiv t''-t'$, and

$$G(\boldsymbol{r},t)) = \frac{1}{N}\left\langle\sum_{i=1}^{N}\sum_{j=1}^{N}\delta[\boldsymbol{r}-(\boldsymbol{r}_i(t+t')-\boldsymbol{r}_j(t'))]\right\rangle$$

This function was introduced by van Hove 1954 [22] to characterise the dynamic structure measured in inelastic neutron scattering experiments. It is known as the *van Hove correlation function*. It naturally separates into two terms, usually called the "self" (s) and the "distinct" (d) part, according to

$$G(\boldsymbol{r},t)) = G_s(\boldsymbol{r},t)) + G_d(\boldsymbol{r},t))$$

where

$$G_s(\boldsymbol{r},t)) = \frac{1}{N}\left\langle\sum_{i=1}^{N}\delta[\boldsymbol{r}-(\boldsymbol{r}_i(t+t')-\boldsymbol{r}_i(t'))]\right\rangle$$

and

$$G_d(\boldsymbol{r},t)) = \frac{1}{N}\left\langle\sum_{i=1}^{N}\sum_{j\neq i}^{N}\delta[\boldsymbol{r}-(\boldsymbol{r}_i(t+t')-\boldsymbol{r}_j(t'))]\right\rangle$$

The self-part is an example of a single-particle correlation function and the distinct part a collective correlation function. The normalization is given by

$$\int_V d\boldsymbol{r}G_s(\boldsymbol{r},t) = 1$$

and

$$\int_V d\boldsymbol{r}G_d(\boldsymbol{r},t) = N-1$$

The physical interpretation of the van Hove function is that it is proportional to the probability to find a particle i at position $\boldsymbol{r}''$ at time t'', provided that a particle j was located at position $\boldsymbol{r}'$ at time t'. More precisely, it is related

to find a particle in a small region $d\boldsymbol{r}'$ around a point $\boldsymbol{r}'$, *etc.* The division into self and distinct parts then corresponds to the possibilities that i and j are the same particles or different ones, respectively. Furthermore, if the system is isotropic it will only depend on the scalar quantity r.

The distinct part can be viewed as a direct dynamic generalisation of the pair-distribution function $g(\boldsymbol{r})$. Initially $G_d(\boldsymbol{r}, t)$ is given by

$$G_d(\boldsymbol{r}, 0) = ng(\boldsymbol{r})$$

and for large times t it approaches a constant value, the mean density

$$\lim_{t\to\infty} G_d(\boldsymbol{r}, t) = (\frac{N}{V} - \frac{1}{V}) \simeq n$$

The self part follows the motion of a single particle. Initially, it is a delta function in space

$$G_s(\boldsymbol{r}, 0) = \delta(\boldsymbol{r})$$

and it will then broaden in space when time proceeds and finally it will reach the value

$$\lim_{t\to\infty} G_s(\boldsymbol{r}, t) = \frac{1}{V} \simeq 0$$

The broadening can be described by its second moment

$$\Delta_{MSD}(t) = \int_V r^2 G_s(\boldsymbol{r}, t) d\boldsymbol{r} = \frac{1}{N} \left\langle \sum_{i=1}^{N} [\boldsymbol{r}_i(t + t') - \boldsymbol{r}_i(t')]^2 \right\rangle \tag{85}$$

This function is called the *mean-squared displacement* and its time dependence gives information about the single particle dynamics. For short times it increases as

$$\Delta_{MSD}(t) = \frac{3k_B T}{m} t^2 \left[1 - \frac{(\omega_E t)^2}{12} + \ldots \right]$$

The first term describes the initial free particle motion while the second term describes the initial effect of the restoring force from the surrounding particles. At long times it will approach a constant value for a solid

$$\lim_{t\to\infty} \Delta_{MSD}(t) = 6\, d_{\text{th}}^2$$

where d_{th} is the mean thermal displacement of a particle from its lattice position. However, in a liquid or dense gas it will diffuse away from its initial position. Assume that the position of the particle becomes uncorrelated after a time-scale τ_c and divide the total tim lag t into M uncorrelated pieces

$$\Delta \boldsymbol{r}_i(j) = \boldsymbol{r}_i(t_j + \tau_c) - \boldsymbol{r}_i(t_j) \; ; \quad j = 1, \ldots, M \; , \quad t_1 = t'$$

with $t = M\tau_c$. The mean squared displacement can now be written as

$$\begin{aligned}\Delta_{MSD}(t) &= \frac{1}{N}\sum_{i=1}^{N}\left\langle\left[\sum_{j=1}^{M}\Delta\boldsymbol{r}_i(j)\right]^2\right\rangle = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{j'=1}^{M}\left\langle\Delta\boldsymbol{r}_i(j)\Delta\boldsymbol{r}_i(j')\right\rangle \\ &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}\left\langle[\Delta\boldsymbol{r}_i(j)]^2\right\rangle \equiv M(\Delta\boldsymbol{r})^2 = \frac{(\Delta\boldsymbol{r})^2}{\tau_c}\,t \propto t\end{aligned}$$

i.e. for a liquid or dense gas we have the limiting behaviour

$$\lim_{t\to\infty}\Delta_{MSD}(t) \propto t$$

8.4 Transport coefficients

The transport of mass, energy, momentum and charge in matter is described phenomenologically by various *transport coefficients*. They are introduced by linear relations of the form

$$\text{Flux} = -\ \text{coefficient} \times \text{gradient}$$

The flux measures the transfer per unit area and unit time, the gradient provides the driving force for the flux, and the coefficient characterizes the resistance to flow. Examples includes Newton's law of viscosity, Fick's law of diffusion, Fourier's law of heat conduction and Ohm's law of electrical conduction. We normally think of these laws applied to nonequilibrium situations were we apply a gradient which then results in a flux. An example could be that we apply a electric potential to a material. This gives rise to a current, which magnitude is determined by the resistivity of the material. However, in addition to nonequilibrium situations linear transport relations also apply to microscopic fluctuations that occur in a system at equilibrium [23]. Thus, transport coefficients, which are properties of matter, can be extracted from equilibrium molecular dynamics simulations.

8.4.1 Generalized Einstein relation

To illustrate a typical transport phenomenon, consider diffusion of a tagged particle, so called *self-diffusion*. It is described by Fick's law of diffusion. The microscopic density of the tagged particle with position $\boldsymbol{r}_i(t)$ is described by

$$n_s(\boldsymbol{r},t) = \delta(\boldsymbol{r} - \boldsymbol{r}_i(t))$$

with the corresponding current or flux,

$$\boldsymbol{j}_s(\boldsymbol{r},t) = \boldsymbol{v}_i(t)\ \delta(\boldsymbol{r} - \boldsymbol{r}_i(t))$$

The *self-diffusion coefficient* D_s is introduced through Fick's law of diffusion

$$\boldsymbol{j}_s(\boldsymbol{r},t) = -D_s \nabla n_s(\boldsymbol{r},t) \tag{86}$$

We expect this equation to be valid in the hydrodynamic regime, *i.e.* for distances $r \gg l$ and times $t \gg \tau$, where l is of the order an inter-particle distance and τ a typical collision time. If we combine Fick's law of diffusion with the continuity equation

$$\frac{\partial}{\partial t} n_s(\boldsymbol{r},t) + \nabla \cdot \boldsymbol{j}_s(\boldsymbol{r},t) = 0 \tag{87}$$

we obtain the diffusion equation

$$\frac{\partial}{\partial t} n_s(\boldsymbol{r},t) = D_s \nabla^2 n_s(\boldsymbol{r},t)$$

or

$$\frac{\partial}{\partial t} G_s(\boldsymbol{r},t) = D_s \nabla^2 G_s(\boldsymbol{r},t) \ , \quad \text{valid if} \quad r \gg l \quad \text{and} \quad t \gg \tau \tag{88}$$

We can solve this by making a Fourier transform in space

$$\begin{aligned} G_s(\boldsymbol{r},t) &= \int \frac{d\boldsymbol{q}}{(2\pi)^3} e^{-D_s q^2 t} e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \\ &= \frac{1}{\sqrt{4\pi D_s t}} \exp\left[-\frac{r^2}{4D_s t}\right] \end{aligned}$$

This equation is valid if $r \gg l$ and $t \gg \tau$ and hence

$$D_s = \lim_{t\to\infty} \frac{1}{6t} \Delta_{MSD}(t) \tag{89}$$

8.4.2 Green-Kubo relations

Consider the dynamic quantity $\mathcal{A}(t)$. At time t the displacement of $\mathcal{A}$ from its value at $t = 0$ is

$$[\mathcal{A}(t) - \mathcal{A}(0)] = \int_0^t dt' \dot{\mathcal{A}}(t')$$

Squaring both sides and averaging over time origins gives

$$\begin{aligned} \left\langle [\mathcal{A}(t) - \mathcal{A}(0)]^2 \right\rangle &= \int_0^t dt'' \int_0^t dt' \left\langle \dot{\mathcal{A}}(t'') \dot{\mathcal{A}}(t') \right\rangle \\ &= 2 \int_0^t dt'' \int_0^{t''} dt' \left\langle \dot{\mathcal{A}}(t'') \dot{\mathcal{A}}(t') \right\rangle \end{aligned}$$

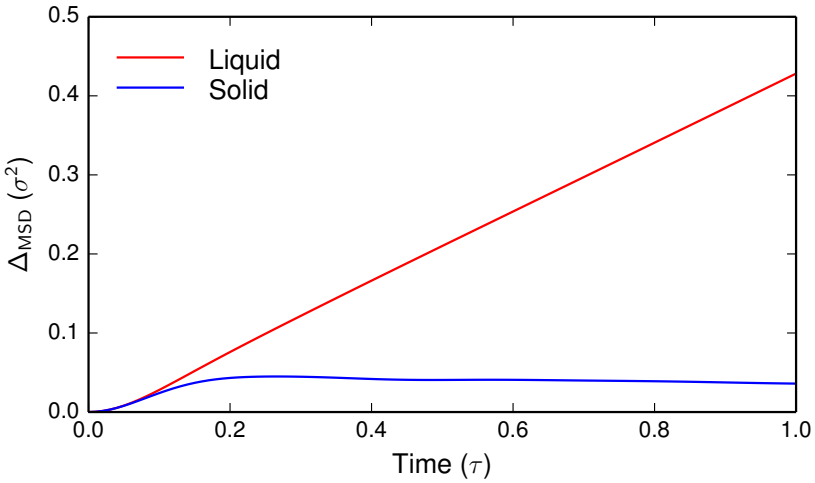

Figure 11: The mean squared displacment. Data taken from a MD simulation study of a liquid and a solid Lennard Jones system.

where we in the last line have used the fact that the integrand is symmetric in t' and t''. The correlation function $\left\langle \dot{\mathcal{A}}(t'')\dot{\mathcal{A}}(t')\right\rangle$ only depends on the time difference and we introduce $\tau \equiv t'' - t'$ and we can write the inegral as

$$\begin{aligned}\left\langle [\mathcal{A}(t) - \mathcal{A}(0)]^2 \right\rangle &= 2\int_0^t d\tau \int_0^{t-\tau} dt' \left\langle \dot{\mathcal{A}}(\tau)\dot{\mathcal{A}}(0)\right\rangle \\ &= 2\int_0^t d\tau \left\langle \dot{\mathcal{A}}(\tau)\dot{\mathcal{A}}(0)\right\rangle \int_0^{t-\tau} dt' \\ &= 2\int_0^t d\tau \left\langle \dot{\mathcal{A}}(\tau)\dot{\mathcal{A}}(0)\right\rangle (t-\tau)\end{aligned}$$

or

$$\frac{\left\langle [\mathcal{A}(t) - \mathcal{A}(0)]^2 \right\rangle}{2t} = \int_0^t d\tau \left\langle \dot{\mathcal{A}}(\tau)\dot{\mathcal{A}}(0)\right\rangle \left(1 - \frac{\tau}{t}\right) \tag{90}$$

Taking the long-time limit, we find

$$\lim_{t\to\infty} \frac{\left\langle [\mathcal{A}(t) - \mathcal{A}(0)]^2 \right\rangle}{2t} = \int_0^\infty d\tau \left\langle \dot{\mathcal{A}}(\tau)\dot{\mathcal{A}}(0)\right\rangle \tag{91}$$

8.4.3 The self-diffusion coefficient

We can apply this to the case $\mathcal{A}(t) = \boldsymbol{r}_i(t)$. This implies that $\dot{\mathcal{A}}(t) = \boldsymbol{v}_i(t)$ and hence

$$3D_s = \int_0^\infty d\tau \Phi(\tau)$$

We now have two different ways to determine the self-diffusion coefficient, either from the mean-squared displacement according to Eq. (89)

$$D_s = \lim_{t\to\infty} \frac{1}{6t} \Delta_{MSD}(t)$$

or from the time-integral of the velocity correlation function

$$D_s = \frac{1}{3} \int_0^\infty d\tau \Phi(\tau) = \frac{1}{6} \hat{\Phi}(\omega = 0) \tag{92}$$

where $\hat{\Phi}(\omega)$ is the spectral function, the Fourier transform of the correlation function.

A Elements of ensemble theory

In Eq. (21) we introduced a time average procedure to obtain macroscopic properties of a system based on the underlying microscopic equation of motion. This is what is used in the molecular dynamics (MD) simulation technique. However, it is not so commonly used in more conventional statistical mechanics due to the complexity of the time evolution of a system with many degrees of freedom. Gibbs suggested to replace the time average with an ensemble average, an average over a set of "mental copies" of the real system [5]. All copies are governed by the same set of microscopic interactions and they are all characterized by the same macrostate as the original system (e.g. the same total energy, volume and number of particles), and the time evolution of each member of the ensemble is governed by classical mechanics, the Hamilton's equation of motion.

A.1 The Liouville equation

Consider an ensemble of systems. Each system in the ensemble is defined by its position $\boldsymbol{x}$ in phase-space

$$\boldsymbol{x} = (q_1, \ldots, q_F, p_1, \ldots, p_F) \tag{93}$$

and each system is moving in phase-space according to Hamilton's equation of motion

$$\dot{q}_\alpha = \frac{\partial \mathcal{H}}{\partial p_\alpha} \ , \qquad \dot{p}_\alpha = -\frac{\partial \mathcal{H}}{\partial q_\alpha} \ ; \qquad \alpha = 1, \ldots, F \tag{94}$$

We denote the number of systems in a small volume $d\boldsymbol{x}$ at a point $\boldsymbol{x}$ in phase-space and at time t as $\rho'(\boldsymbol{x}, t)d\boldsymbol{x}$. We can then introduce the probability to find a system in a small volume $d\boldsymbol{x}$ at a point $\boldsymbol{x}$ and at time t according to

$$\rho(\boldsymbol{x}, t)d\boldsymbol{x} = \frac{1}{N_{ens}}\rho'(\boldsymbol{x}, t)d\boldsymbol{x} \tag{95}$$

where N_{ens} is the total number of systems in the ensemble. The *density function* $\rho(\boldsymbol{x}, t)$ is a probability distribution function, properly normalized over the entire phase-space volume

$$\int \rho(\boldsymbol{x}, t) \, \mathrm{d}\boldsymbol{x} = 1 \qquad \rho(\boldsymbol{x}, t) \geq 0 \tag{96}$$

To derive the Liouville equation we consider a small arbitrary volume ω in phase-space and we denote the surface enclosing this volume by σ. The change

$$\frac{\partial}{\partial t} \int_\omega \rho(\boldsymbol{x}, t)d\boldsymbol{x}$$

is given by the difference of the flux of systems into and out from the volume ω. There are no sources or sinks of systems in phase space. The flux is given by the vector

$$\boldsymbol{v} = \dot{\boldsymbol{x}} = (\dot{q}_1, \ldots, \dot{q}_F, \dot{p}_1, \ldots, \dot{p}_F) \tag{97}$$

hence,

$$\frac{\partial}{\partial t}\int_\omega \rho(\boldsymbol{x}, t) d\boldsymbol{x} = -\int_\sigma \rho(\boldsymbol{x}, t)(\boldsymbol{v}\cdot\hat{n}) d\boldsymbol{s} \tag{98}$$

where $\hat{n}$ is the outward unit vector normal to the surface element $d\boldsymbol{s}$. Using the divergence theorem the surface integral can be converted to a volume integral

$$\int_\sigma \rho(\boldsymbol{v}\cdot\hat{n}) d\boldsymbol{s} = \int_\omega \mathrm{div}(\rho\boldsymbol{v}) d\boldsymbol{x} \tag{99}$$

The volume ω is arbitrary and hence we obtain the *equation of continuity in phase-space*

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho\boldsymbol{v}) = 0 \tag{100}$$

More explicitly $\mathrm{div}(\rho\boldsymbol{v})$ can be written as

$$\begin{aligned}
\mathrm{div}(\rho\boldsymbol{v}) &= \sum_{\alpha=1}^{F}\left[\frac{\partial}{\partial q_\alpha}(\rho\dot{q}_\alpha) + \frac{\partial}{\partial p_\alpha}(\rho\dot{p}_\alpha)\right] \\
&= \sum_{\alpha=1}^{F}\left[\dot{q}_\alpha\frac{\partial\rho}{\partial q_\alpha} + \dot{p}_\alpha\frac{\partial\rho}{\partial p_\alpha}\right] + \rho\sum_{\alpha=1}^{F}\left[\frac{\partial}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial}{\partial p_\alpha}\dot{p}_\alpha\right]
\end{aligned}$$

We now use the fact that each system in the ensemble is moving according to Hamilton's equation of motion. The first term in the expression for the divergence can then be written as

$$\sum_{\alpha=1}^{F}\left[\dot{q}_\alpha\frac{\partial\rho}{\partial q_\alpha} + \dot{p}_\alpha\frac{\partial\rho}{\partial p_\alpha}\right] = \sum_{\alpha=1}^{F}\left[\frac{\partial\mathcal{H}}{\partial p_\alpha}\frac{\partial\rho}{\partial q_\alpha} - \frac{\partial\mathcal{H}}{\partial q_\alpha}\frac{\partial\rho}{\partial p_\alpha}\right] = \{\rho, \mathcal{H}\}$$

where we have used the *Poisson bracket* notation

$$\{\mathcal{A}, \mathcal{B}\} \equiv \sum_{\alpha=1}^{F}\left[\frac{\partial\mathcal{A}}{\partial q_\alpha}\frac{\partial\mathcal{B}}{\partial p_\alpha} - \frac{\partial\mathcal{B}}{\partial q_\alpha}\frac{\partial\mathcal{A}}{\partial p_\alpha}\right] \tag{101}$$

The second term is equal to zero, according to

$$\rho\sum_{\alpha=1}^{F}\left[\frac{\partial}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial}{\partial p_\alpha}\dot{p}_\alpha\right] = \rho\sum_{\alpha=1}^{F}\left[\frac{\partial}{\partial q_\alpha}\frac{\partial\mathcal{H}}{\partial p_\alpha} - \frac{\partial}{\partial p_\alpha}\frac{\partial\mathcal{H}}{\partial q_\alpha}\right] = 0$$

To summarize, we have derived the *Liouville equation*

$$\frac{\partial\rho}{\partial t} + \{\rho, \mathcal{H}\} = 0 \tag{102}$$

the most fundamental equation of statistical mechanics. The Liouville equation is often written as

$$\frac{\partial \rho}{\partial t} + i\mathcal{L}\rho = 0 \tag{103}$$

where

$$i\mathcal{L}\ldots \equiv \{\ldots, \mathcal{H}\} \tag{104}$$

is the *Liouville operator*.

For a system with N particles and in Cartesian coordinates the Liouville equation takes the form

$$\left[\frac{\partial}{\partial t} + \sum_{i=1}^{N} \boldsymbol{v}_i \cdot \frac{\partial}{\partial \boldsymbol{r}_i} + \sum_{i=1}^{N} \boldsymbol{a}_i \cdot \frac{\partial}{\partial \boldsymbol{v}_i}\right] \rho(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_N, t) = 0 \tag{105}$$

where

$$\boldsymbol{a}_i = \frac{\boldsymbol{F}_i}{m_i} = -\frac{1}{m_i}\nabla_i V_{\text{pot}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \tag{106}$$

is the acceleration for particle number i and the Liouville operator is given by

$$i\mathcal{L} = \sum_{i=1}^{N} \left[\boldsymbol{v}_i \cdot \frac{\partial}{\partial \boldsymbol{r}_i} + \boldsymbol{a}_i \cdot \frac{\partial}{\partial \boldsymbol{v}_i}\right] \tag{107}$$

In the quantum case the probability distribution is replaced by the density matrix $\hat{\rho}$ and the Liouville equation by the equation

$$\frac{\partial}{\partial t}\hat{\rho} + \frac{1}{i\hbar}\left[\hat{\rho}, \hat{H}\right] = 0 \tag{108}$$

where

$$\left[\hat{\mathcal{A}}, \hat{\mathcal{B}}\right] \equiv \hat{\mathcal{A}}\hat{\mathcal{B}} - \hat{\mathcal{B}}\hat{\mathcal{A}} \tag{109}$$

is the commutator.

A.2 Liouville's theorem

Since $\rho = \rho(\boldsymbol{x}, t)$ its total time derivative can be written as

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\partial \rho}{\partial t} + + \sum_{\alpha=1}^{F}\left[\frac{\partial \rho}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial \rho}{\partial p_\alpha}\dot{p}_\alpha\right] \tag{110}$$

$$= \frac{\partial \rho}{\partial t} + \{\rho, \mathcal{H}\} \tag{111}$$

By using the Liouville equation (102) we obtain

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = 0 \tag{112}$$

which is known as the *Liouville's theorem* and is of fundamental importance for statistical mechanics. It states that the change in density as viewed by an observer moving along with the trajectory is zero. The density function along a flow line stays constant in time. Thus, the swarm of the systems in an ensemble moves in phase-space in essentially the same manner as an incompressible fluid moves in physical space. The occupied volume in phase-space does not change in time.

On the other hand, $\partial\rho/\partial t$ is the change of the density function at a fixed point in phase-space and that may change in time. The Liouville's theorem follows from classical mechanics, the Hamilton's equation of motion. It is the volume in phase-space that is constant in time. This shows the importance to use the phase-space in describing the microscopic state of a system. Phase-space plays a privileged role in statistical mechanics.

A.3 Equilibrium distribution functions

The Liouville equation (102) is quite generally true. If we are interested in a system in equilibrium the density function has to be stationary in time

$$\frac{\partial \rho}{\partial t} = 0$$

i.e. it can not have an explicit dependence on time, $\rho = \rho(\boldsymbol{x})$. The Liouville equation (102) is reduced to

$$\{\rho, \mathcal{H}\} = 0 \tag{113}$$

The ensemble average of some dynamical quantity $\mathcal{A}(\boldsymbol{x})$ is then given by

$$\langle \mathcal{A} \rangle_{\text{ens}} = \int \mathcal{A}(\boldsymbol{x}) \rho(\boldsymbol{x}) d\boldsymbol{x} \tag{114}$$

where the integration is over all phase space.

The condition in Eq. (113) can be satisfied in many different ways. If we assume that the dependence of ρ on $\boldsymbol{x}$ comes only through an explicit dependence on the Hamiltonian $\mathcal{H}(\boldsymbol{x})$ i.e.

$$\rho(\boldsymbol{x}) = f(\mathcal{H}(\boldsymbol{x})) \tag{115}$$

where $f(\ldots)$ is some arbitrary function, the condition in Eq. (113) will be automatically satisfied.

A.4 Microcanonical ensemble

Consider a system with N particles in a volume V. We assume that the system is isolated from the surrounding and hence, its energy E is conserved. If there are no further restrictions on the system its macroscopic state is

therefore characterized by N, V and E. We can then construct an ensemble of systems that all have the particle number N, volume V and energy E. This ensemble is called the *microcanonical ensemble*. We then make the fundamental assumption that each accessible microscopic state in phase space has the same probability

$$\rho(\boldsymbol{x}) \propto \delta(\mathcal{H}(\boldsymbol{x}) - E) \tag{116}$$

The ensemble average is then given by

$$\langle \mathcal{A} \rangle_{NVE} = \frac{\int_V \mathcal{A}(\boldsymbol{x}) \delta(\mathcal{H}(\boldsymbol{x}) - E) d\boldsymbol{x}}{\int_V \delta(\mathcal{H}(\boldsymbol{x}) - E) d\boldsymbol{x}} \tag{117}$$

where the integration is restricted to the volume V and to N particles. The expression for $\rho(\boldsymbol{x})$ in Eq. (116) depends only explicitly on $\mathcal{H}(\boldsymbol{x})$ and will therefore correspond to an equilibrium ensemble according to the Liouville equation (113). For mathematical reasons one sometimes consider a system not with a definite energy E but with an energy in a thin shell ΔE around E and one then let ΔE goes to zero.

The fundamental assumption that each accessible microscopic state has the same probability is known as the *assumption of equal a priori probabilities*. This assumption is very reasonable. We know from the Liouville's theorem that there is no tendency for phase points to crowd into one region in phase space rather than another, and hence, it would be arbitrary to make any assumption other than that of equal *a priori* probabilities for different regions of equal extent in the phase space. The assumption of equal *a priori* probabilities is the only hypothesis that can reasonably be chosen [24].

If Hamilton's equation of motion is being solved in time that would generate microscopic configurations with constant energy E. Suppose that given an infinite amount of time the system is able to visit all configurations on the constant energy surface. A system with this property is said to be *ergodic*. The time and ensemble average are then equal

$$\langle \mathcal{A} \rangle_{\text{time}} = \langle \mathcal{A} \rangle_{NVE} \tag{118}$$

This is called the *ergodic hypothesis*. It is generally believed that most systems are ergodic but it has turned out to be very difficult to formally prove ergodicity.

A.5 Canonical ensemble

Many more ensembles can be introduced that can be used to study systems in equilibrium. The condition (115) provides a recipe to obtain stationary distributions.

The microcanonical ensemble is of fundamental importance. However, for applications other ensembles are more important. To keep a system

at constant energy is not easy. It is more natural to consider a system at constant temperature T by keeping it in contact with an appropriate heat reservoir. The corresponding ensemble is the *canonical ensemble*. In this case the macroscopic state of the system is defined by the parameters N, V and T. The energy of the system can vary be exchanging energy with the surrrounding, the heat reservoir. The precise nature of the reservoir is not important, it only has to be much larger than the system

Starting with the microcanonical description one can then derive that for this system the density function is given by

$$\rho(\boldsymbol{x}) \propto \exp\left[-\beta \mathcal{H}(\boldsymbol{x})\right] \tag{119}$$

where $\beta = 1/kT$ and k is Boltzmann's constant. Also in this case the expression for $\rho(\boldsymbol{x})$ is on the form in (115) and it will describe an equilibrium distribution. The ensemble average is given by

$$\langle \mathcal{A} \rangle_{NVT} = \frac{\int_V \mathcal{A}(\boldsymbol{x}) \exp\left[-\beta \mathcal{H}(\boldsymbol{x})\right] d\boldsymbol{x}}{\int_V \exp\left[-\beta \mathcal{H}(\boldsymbol{x})\right] d\boldsymbol{x}} \tag{120}$$

Analytical calculations are often more easy to perform using the canonical ensemble compared with the microcanonical ensemble. The precise nature of the ensemble is often not so important. For large systems the average values from the microcanonical and the canonical ensembles will be the same

$$\langle \mathcal{A} \rangle_{NVE} = \langle \mathcal{A} \rangle_{NVT} + \mathcal{O}\left(\frac{1}{N}\right) \tag{121}$$

but the fluctuations from the average values can be different.

B Symplectic integrators

In Sec. 4 numerical integration methods where derived based on Taylor expansions. A more formal and more powerful technique has been developed, based on the classical time evolution operator approach.

The time evolution of the phase-space point $\boldsymbol{x}(t)$ is formally given by

$$\boldsymbol{x}(t) = e^{i\mathcal{L}t}\boldsymbol{x}(0)$$

where $i\mathcal{L}$ is the classical Liouville operator. For a system of N point particles using Cartesian coordinates we have

$$i\mathcal{L} = \sum_{i=1}^{N}\left[\boldsymbol{v}_i \cdot \frac{\partial}{\partial \boldsymbol{r}_i} + \boldsymbol{a}_i \cdot \frac{\partial}{\partial \boldsymbol{v}_i}\right] \tag{122}$$

We can divide the operator into two parts

$$i\mathcal{L} = i\mathcal{L}_r + i\mathcal{L}_v \tag{123}$$

where

$$i\mathcal{L}_r = \sum_{i=1}^{N} \boldsymbol{v}_i \cdot \frac{\partial}{\partial \boldsymbol{r}_i} \tag{124}$$

$$i\mathcal{L}_v = \sum_{i=1}^{N} \boldsymbol{a}_i \cdot \frac{\partial}{\partial \boldsymbol{v}_i} \tag{125}$$

The two corresponding propagators, $e^{i\mathcal{L}_r t}$ and $e^{i\mathcal{L}_v t}$, can be evaluated analytically. Consider the operator $\exp\left(c\partial/\partial z\right)$ acting on an arbitrary function $g(z)$, where c is independent on z. The action of the operator can be worked out by expanding the exponential in a Taylor series

$$\begin{aligned}\exp\left(c\frac{\partial}{\partial z}\right)g(z) &= \sum_{k=0}^{\infty}\frac{1}{k!}\left(c\frac{\partial}{\partial z}\right)^k g(z) \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\, c^k\, \frac{d^k}{dz^k}\, g(z)\end{aligned}$$

The second line is just the Taylor expansion of $g(z+c)$ about $c=0$. Thus, we have the general result

$$\exp\left(c\frac{\partial}{\partial z}\right)g(z) = g(z+c)$$

which shows that the operator gives rise to a pure translation. Therefore, the propagator $\exp\left(i\mathcal{L}_r t\right)$ only translates all position coordinates

$$\boldsymbol{r}_i \to \boldsymbol{r}_i + \boldsymbol{v}_i t \quad \forall i$$

and the propagator $\exp(i\mathcal{L}_v t)$ only all velocity coordinates

$$\boldsymbol{v}_i \to \boldsymbol{v}_i + \boldsymbol{a}_i t \quad \forall i$$

We can now write that

$$\boldsymbol{x}(t) = e^{i(\mathcal{L}_r + \mathcal{L}_v)t}\boldsymbol{x}(0) \tag{126}$$

However, the two operators $i\mathcal{L}_r$ and $i\mathcal{L}_v$ do not commute

$$i\mathcal{L}_r i\mathcal{L}_v \neq i\mathcal{L}_v i\mathcal{L}_r$$

and hence

$$e^{i(\mathcal{L}_r+\mathcal{L}_v)t} \neq e^{i\mathcal{L}_r t}e^{i\mathcal{L}_v t}$$

and Eq. (126) is not easily solved.

To show that they do not commute consider a single particle moving in one dimension. Its phase-space is described by two coordinates (x, v) and

$$i\mathcal{L}_x = v\frac{\partial}{\partial x} \tag{127}$$

$$i\mathcal{L}_v = a(x)\frac{\partial}{\partial v} \tag{128}$$

The action of $i\mathcal{L}_x i\mathcal{L}_v$ on some arbitrary function $g(x, v)$ is

$$\left[v\frac{\partial}{\partial x}a(x)\frac{\partial}{\partial v}\right]g(x,v) = \left[va'(x)\frac{\partial}{\partial v} + va(x)\frac{\partial^2}{\partial x \partial v}\right]g(x,v)$$

and the action of $i\mathcal{L}_v i\mathcal{L}_x$ is

$$\left[a(x)\frac{\partial}{\partial v}v\frac{\partial}{\partial x}\right]g(x,v) = \left[a(x)\frac{\partial}{\partial x} + a(x)v\frac{\partial^2}{\partial x \partial v}\right]g(x,v)$$

The function $g(x, v)$ is arbitrary, hence

$$[i\mathcal{L}_x, i\mathcal{L}_v] = va'(x)\frac{\partial}{\partial v} - a(x)\frac{\partial}{\partial x}$$

and they do not in general commute.

For noncommuting operators A and B,

$$e^{(A+B)} \neq e^A e^B$$

we can make use of the Trotter identity

$$e^{(A+B)} = \lim_{P\to\infty}\left[e^{A/P}e^{B/P}\right]^P \tag{129}$$

where P is an integer. It is somewhat better to use the symmetric version

$$e^{(A+B)} = \lim_{P\to\infty}\left[e^{B/2P}e^{A/P}e^{B/2P}\right]^P \tag{130}$$

We can apply this to the classical propagator $\exp(i\mathcal{L}t)$. If we define the time step $\Delta t = t/P$, we can write

$$e^{i\mathcal{L}t} = \lim_{\Delta t \to 0 (P \to \infty)} \left[e^{i\mathcal{L}_v \Delta t/2} e^{i\mathcal{L}_r \Delta t} e^{i\mathcal{L}_v \Delta t/2} \right]^P \tag{131}$$

For large but finite P we then get an approximation to $\exp(i\mathcal{L}t)$,

$$e^{i\mathcal{L}t} = \left[e^{i\mathcal{L}_v \Delta t/2} e^{i\mathcal{L}_r \Delta t} e^{i\mathcal{L}_v \Delta t/2} \right]^P + \mathcal{O}(P\Delta t^3) \tag{132}$$

which can be used in a numerical propagation scheme. To see how this works in practise, consider again a single particle moving in one dimension. One time step is the given by

$$\begin{aligned}
\begin{bmatrix} x(t+\Delta t) \\ v(t+\Delta t) \end{bmatrix} &= e^{i\mathcal{L}_v \Delta t/2} e^{i\mathcal{L}_r \Delta t} e^{i\mathcal{L}_v \Delta t/2} \begin{bmatrix} x(t) \\ v(t) \end{bmatrix} \\
&= e^{i\mathcal{L}_v \Delta t/2} e^{i\mathcal{L}_r \Delta t} \begin{bmatrix} x(t) \\ v(t) + \frac{\Delta t}{2} a[x(t)] \end{bmatrix} \\
&= e^{i\mathcal{L}_v \Delta t/2} \begin{bmatrix} x(t) + \Delta t v(t) \\ v(t) + \frac{\Delta t}{2} a[x(t) + \Delta t v(t)] \end{bmatrix} \\
&= \begin{bmatrix} x(t) + \Delta t\{v(t) + \frac{\Delta t}{2} a[x(t)]\} \\ v(t) + \frac{\Delta t}{2} a[x(t)] + \frac{\Delta t}{2} a\left[x(t) + \Delta t\left\{v(t) + \frac{\Delta t}{2} a[x(t)]\right\}\right] \end{bmatrix} \\
&= \begin{bmatrix} x(t) + \Delta t v(t) + \frac{\Delta t^2}{2} a[x(t)] \\ v(t) + \frac{\Delta t}{2}\{a[x(t)] + a[x(t+\Delta t)]\} \end{bmatrix}
\end{aligned}$$

The resulting algorithm is identical to the velocity Verlet algorithm, presented in Eq. (133)

$$\begin{aligned}
v &= v + 0.5\ a\ dt \\
r &= r + v\ dt \\
a &= \texttt{accel}(r) \\
v &= v + 0.5\ a\ dt
\end{aligned}$$

The above analysis demonstrates how the velocity Verlet algorithm can be obtained via the powerful Trotter factorization scheme. The first step is a velocity translation (a half time-step). It is sometimes called a "kick", since it impulsively changes the velocity without altering the positions. The second step is a position translation (a full time-step). This is often denoted a "drift" step, because it advances the positions without changing the velocities. The accelerations are then updated and, finally, the velocities are translated (a half time-step) with the new accelerations. These are just the steps required by the above operator factorization scheme. The fact that the instructions in the computer code can be written directly from the operator

factorization scheme, by-passing the lengthy algebra needed to derive explicit finite-difference equations, has created the powerful direct translation method [25].

Moreover, it is now clear that the velocity Verlet algorithm constitutes a symplectic integrator, that preserves the important symmetries of classical mechanics. It is area preserving [1] and time reversible. Each individual part of the product implied by Eq. (131), is symplectic and hence the overall time evolution is symplectic.

Finally, let us try to understand the absence of long-term energy drift in the Verlet algorithm. When we use the Verlet algorithm the true Liouville operator $e^{(i\mathcal{L}t)}$ is replaced by the factorization in Eq. (132) In doing so we make an approximation. We can write

$$e^{i\mathcal{L}_v\Delta t/2}e^{i\mathcal{L}_r\Delta t}e^{i\mathcal{L}_v\Delta t/2} = e^{(i\mathcal{L}\Delta t+\epsilon)}$$

where ϵ is an operator that can be expressed in terms of the commutators of $\mathcal{L}_r$ and $\mathcal{L}_v$. We can then define a pseudo-Liouville operator

$$i\mathcal{L}_{pseudo}\Delta t \equiv i\mathcal{L}\Delta t + \epsilon$$

that corresponds to a pseudo Hamiltonian. The energy for this pseudo Hamiltonian is rigorously conserved by the Verlet style algorithm and the difference between the pseudo Hamiltonian and the true Hamiltonian can be made as small as we like by choosing Δt sufficiently small. As the true Hamiltonian is forced to remain close to a conserved quantity, we can now understand why there is no long-term drift in the energy with Verlet type of algorithms.

C Error estimate

It is important to find error bounds associated with evaluated quantities in a simulation. Consider a variable f. Assume that M measurements have been made $\{f_i\}$ and denote the average as

$$I = \frac{1}{M}\sum_{i=1}^{M} f_i \tag{133}$$

We would like to determine the error bounds for I, its variance. If the values $\{f_i\}$ are independent on each others, *i.e.* uncorrelated data, the variance for I is given by $\mathrm{Var}[I] = \frac{1}{M}\mathrm{Var}[f]$ where

$$\mathrm{Var}[f] = \sigma^2(f) = \left\langle (f - \langle f\rangle)^2 \right\rangle = \left\langle f^2 \right\rangle - \langle f\rangle^2 \tag{134}$$

However, in a simulation subsequent data are often highly correlated. The variance of I will depend on the number of independent samples M_{eff} generated by the simulation. We can introduce the *statistical inefficiency* s according to

$$M_{\mathrm{eff}} = M/s$$

The variance of I can then be written as

$$\mathrm{Var}[I] = \frac{1}{M_{\mathrm{eff}}}\mathrm{Var}[f] = \frac{s}{M}\mathrm{Var}[f] \tag{135}$$

and the problem is reduced to determine s. We will consider two methods, one based on a direct evaluation of the corresponding correlation function and one based on data blocking.

The variance of I can be written as

$$\begin{aligned}
\mathrm{Var}[I] &= \left\langle \left(\frac{1}{M}\sum_{i=1}^{M} f_i - \langle f\rangle \right)^2 \right\rangle \\
&= \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} \left[\langle f_i f_j\rangle - \langle f\rangle^2 \right]
\end{aligned}$$

If the data are uncorrelated we have that $\langle f_i f_j\rangle - \langle f\rangle^2 = [\langle f^2\rangle - \langle f\rangle^2]\delta_{ij}$ and $\mathrm{Var}[I] = \frac{1}{M}\mathrm{Var}[f]$. If the data are correlated we introduce the correlation function

$$\Phi_k = \frac{\langle f_i f_{i+k}\rangle - \langle f\rangle^2}{\langle f^2\rangle - \langle f\rangle^2} \tag{136}$$

This is normalized such that

$$\Phi_{k=0} = 1$$

We also assume that we study a stationary system and hence $\Phi_k = \Phi_{-k}$. For large k, $k > M_c$, Φ_k will decay,

$$\Phi_{k>M_c} \to 0$$

We assume that the total length of the simulation is considerably longer, $M > M_c$. By introducing $k = i - j$ we can now write the variance as

$$\begin{aligned}
\mathrm{Var}[I] &= \frac{1}{M^2} \sum_{i=1}^{M} \sum_{k=-(M-1)}^{M-1} \left(1 - \frac{\mid k \mid}{M}\right) \left[\langle f_i f_{i+k} \rangle - \langle f \rangle^2\right] \\
&= \mathrm{Var}[f] \, \frac{1}{M^2} \sum_{i=1}^{M} \sum_{k=-(M-1)}^{M-1} \left(1 - \frac{\mid k \mid}{M}\right) \Phi_k \\
&= \mathrm{Var}[f] \, \frac{1}{M^2} \sum_{i=1}^{M} \sum_{k=-M_c}^{M_c} \Phi_k \\
&= \mathrm{Var}[f] \, \frac{1}{M} \sum_{k=-M_c}^{M_c} \Phi_k
\end{aligned}$$

and hence

$$s = \sum_{k=-M_c}^{M_c} \Phi_k \tag{137}$$

By comparing with the definition of the relaxation time τ_{rel} in Eq. (72) we find that the statistical inefficiency s is equal to 2 times the relaxation time

$$s = 2\tau_{rel} \tag{138}$$

If we assume that the correlation function decays exponentially, $\Phi_k = \exp\left(-k/\tau_{rel}\right)$, we notice that

$$\Phi_{k=s} = e^{-2} = 0.135 \sim 0.1$$

The statistical inefficiency can then be determined as the "time" when the corresponding correlation function has decayed to about 10% of its initial value.

Another way to determine the statistical inefficiency s is to use so called *block averaging*. Divide the total length M of the simulation into M_B blocks of size B,

$$M = BM_B$$

Determine the average in each block

$$F_j = \frac{1}{B} \sum_{i=1}^{B} f_{i+(j-1)B} \quad \text{for } j = 1, \ldots, M_B \tag{139}$$

and the corresponding variance Var[F]. If the block size B is larger than s $\{F_j\}$ will be uncorrelated and hence

$$\mathrm{Var}[I] = \frac{1}{M_B}\mathrm{Var}[F] \quad \text{if } B > s$$

However, if the block size is smaller than s we have that

$$\mathrm{Var}[I] > \frac{1}{M_B}\mathrm{Var}[F] \quad \text{if } B < s$$

and we obtain the following relation for s

$$\begin{aligned} \mathrm{Var}[I] &\geq \frac{1}{M_B}\mathrm{Var}[F] \\ \frac{s}{M}\mathrm{Var}[f] &\geq \frac{B}{M}\mathrm{Var}[F] \\ s &\geq \frac{B\mathrm{Var}[F]}{\mathrm{Var}[f]} \end{aligned}$$

We can then obtain the statistical inefficiency

$$s = \lim_{B\ large} \frac{B\mathrm{Var}[F]}{\mathrm{Var}[f]} \tag{140}$$

by plotting $B\mathrm{Var}[F]/\mathrm{Var}[f]$ as function of the block size B. In Fig. 12 we show a typical result from a simulation.

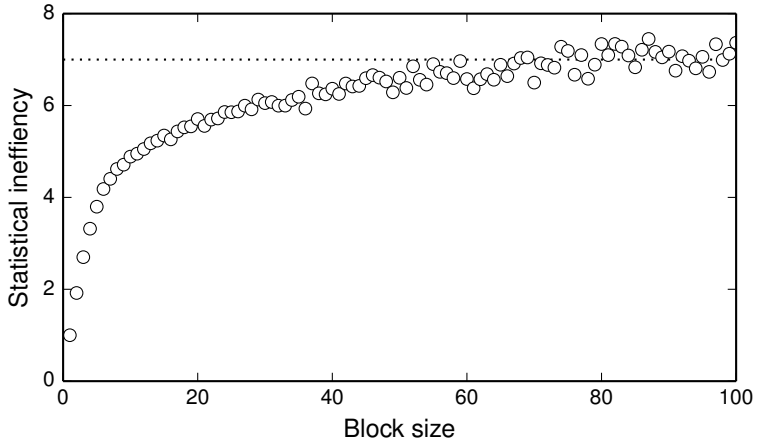

Figure 12: Illustration of the calculation of s using block averaging. The figure shows the approach to the plateau value $s = 7$ when the block size B is increased.

D Temperature and pressure

Expressions for temperature and pressure can be derived using the canonical ensemble. Consider a system of N identical particles with mass m moving in a volume V at temperature T and assume that the interaction is given by the potential $V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$. In the classical limit the partition function is then given by

$$Q(N,V,T) = \frac{1}{N!}\int \frac{d\boldsymbol{x}}{h^{3N}} \exp\left[-\beta \mathcal{H}(\boldsymbol{x})\right]$$

where $d\boldsymbol{x} = d\boldsymbol{r}_1, \ldots, d\boldsymbol{r}_N, d\boldsymbol{p}_1, \ldots, d\boldsymbol{p}_N$ and

$$\mathcal{H}(\boldsymbol{x}) = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} + V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$$

The momentum part of the partition function can be done analytically

$$\begin{aligned} Q(N,V,T) &= \frac{1}{N!}\frac{1}{\lambda^{3N}} \int d\boldsymbol{r}_1 \ldots d\boldsymbol{r}_N \exp\left[-\beta V_{\rm pot}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)\right] \quad (141) \\ &= \frac{1}{N!}\frac{1}{\lambda^{3N}} Z(N,V,T) \quad (142) \end{aligned}$$

where then $Z(N,V,T)$ is the configurational part of the partition function and

$$\lambda = \sqrt{h^2/2\pi m k T}$$

is the thermal deBroglie wavelength.

D.1 The temperature

The temperature is related to the random motion of the particles, contained in the kinetic part of the total energy. Consider the average of the kinetic energy

$$\begin{aligned} \left\langle \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} \right\rangle_{NVT} &= \frac{1}{Q}\frac{1}{N!}\int \frac{d\boldsymbol{x}}{h^{3N}} \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} \exp\left[-\beta \mathcal{H}(\boldsymbol{x})\right] \\ &= \left(\frac{\lambda}{h}\right)^{3N} \int d\boldsymbol{p}_1 \ldots d\boldsymbol{p}_N \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} \exp\left[-\beta \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m}\right] \\ &= \frac{3NkT}{2} \quad (143) \end{aligned}$$

We can then define the instantaneous temperature as

$$\mathcal{T} = \frac{1}{3Nk} \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{m}$$

and

$$T = \langle \mathcal{T} \rangle_{NVT}$$

D.2 The pressure

The pressure P is given by the expression

$$P = -\left(\frac{\partial F}{\partial V}\right)_{N,T}$$

where the Helmholtz's energy is given by

$$F(N,V,T) = -k_{\mathrm{B}}T \ln Q(N,V,T)$$

This implies that

$$P = k_{\mathrm{B}}T\frac{1}{Q}\left(\frac{\partial Q}{\partial V}\right)_{N,T} = k_{\mathrm{B}}T\frac{1}{Z}\left(\frac{\partial Z}{\partial V}\right)_{N,T}$$

To evaluate the volume derivate we introduce scaled coordinates $\boldsymbol{s}_i$ according to

$$\boldsymbol{r}_i = L\boldsymbol{s}_i \qquad i = 1,\ldots,N$$

where L is the length of the cubic box, $V = L^3$. The scaled coordinates are not influenced by a change of the volume V, only the real coordinates $\boldsymbol{r}_i$ change when the volume changes. The configurational partition function can be writtes as

$$Z(N,V,T) = V^N \int d\boldsymbol{s}_1 \ldots d\boldsymbol{s}_N \exp\left[-\beta\ V_{\mathrm{pot}}(V^{1/3}\boldsymbol{s}_1,\ldots,V^{1/3}\boldsymbol{s}_N)\right]$$

which implies that

$$\begin{aligned}
P &= \frac{k_{\mathrm{B}}T}{Z}NV^{N-1}\int d\boldsymbol{s}_1 \ldots d\boldsymbol{s}_N \exp\left[-\beta\ V_{\mathrm{pot}}\right] \\
&+ k_{\mathrm{B}}T\frac{V^N}{Z}\int d\boldsymbol{s}_1 \ldots d\boldsymbol{s}_N \left(-\beta\frac{dV_{\mathrm{pot}}}{dV}\right)\exp\left[-\beta\ V_{\mathrm{pot}}\right] \\
&= \frac{Nk_{\mathrm{B}}T}{V} - \frac{V^N}{Z}\int d\boldsymbol{s}_1 \ldots d\boldsymbol{s}_N \frac{dV_{\mathrm{pot}}}{dV}\exp\left[-\beta\ V_{\mathrm{pot}}\right] \\
&= \frac{Nk_{\mathrm{B}}T}{V} - \frac{V^N}{Z}\int d\boldsymbol{s}_1 \ldots d\boldsymbol{s}_N \frac{V^{-2/3}}{3}\sum_{i=1}^{N}(\nabla_i V_{\mathrm{pot}}\cdot\boldsymbol{s}_i)\exp\left[-\beta\ V_{\mathrm{pot}}\right] \\
&= \frac{Nk_{\mathrm{B}}T}{V} + \frac{1}{3V}\frac{V^N}{Z}\int d\boldsymbol{s}_1 \ldots d\boldsymbol{s}_N \sum_{i=1}^{N}(\boldsymbol{r}_i\cdot\boldsymbol{F}_i)\exp\left[-\beta\ V_{\mathrm{pot}}\right] \\
&= \frac{Nk_{\mathrm{B}}T}{V} + \frac{1}{3V}\frac{1}{Z}\int d\boldsymbol{r}_1 \ldots d\boldsymbol{r}_N \sum_{i=1}^{N}(\boldsymbol{r}_i\cdot\boldsymbol{F}_i)\exp\left[-\beta\ V_{\mathrm{pot}}\right] \\
&= \frac{Nk_{\mathrm{B}}T}{V} + \frac{1}{3V}\left\langle\sum_{i=1}^{N}\boldsymbol{r}_i\cdot\boldsymbol{F}_i\right\rangle_{NVT} \qquad (144)
\end{aligned}$$

We can then define the instanteneous pressure as

$$\mathcal{P} = \frac{1}{3V} \sum_{i=1}^{N} \left[\frac{\boldsymbol{p}_i^2}{m} + \boldsymbol{r}_i \cdot \boldsymbol{F}_i \right]$$

and

$$PV = N k_{\mathrm{B}} T + W$$

where

$$W = \langle \mathcal{W} \rangle_{NVT} = \left\langle \frac{1}{3} \sum_{i=1}^{N} \boldsymbol{r}_i \cdot \boldsymbol{F}_i \right\rangle_{NVT}$$

is called the virial function.

E Equilibration

To start a simulation we need to give the initial positions and velocities for all particles. This will give rise to a certain temperature and pressure after some initial equilibration time T_{eq}. However, there is no explicit expression for the obtained temperature and pressure expressed in terms of the initial positions and velocities. To obtain a certain temperature and pressure one can scale the position coordinates and velocities during the equilibration run and then turn off the scaling, during the production run when average quantities are being computed.

We consider here one type of scaling technique to obtain a certain temperature T_{eq} and pressure P_{eq}. Consider first the temperature. Initially, it is equal to $\mathcal{T}(0)$, where $\mathcal{T}(t)$ is the instantaneous temperature. We would like it to decay exponentially to T_{eq},

$$\mathcal{T}(t) = T_{eq} + (\mathcal{T}(0) - T_{eq})e^{-t/\tau_T}$$

with some decay time constant τ_T. We can change the instantaneous temperature by scaling the velocities at each time-step according to

$$\boldsymbol{v}_i^{new} = \alpha_T^{1/2} \boldsymbol{v}_i^{old}$$

By choosing

$$\alpha_T(t) = 1 + \frac{2\Delta t}{\tau_T} \frac{T_{eq} - \mathcal{T}(t)}{\mathcal{T}(t)} \tag{145}$$

the instantaneous temperature will approximately decay exponentially with time constant τ_T to the desired temperature T_{eq}.

In the same way we can modify the pressure. In this case we have to scale the positions (and the box size). We again would like to obtain an exponential decay

$$\mathcal{P}(t) = P_{eq} + (\mathcal{P}(0) - P_{eq})e^{-t/\tau_P}$$

to the desired pressure P_{eq}. The positions are scaled according to

$$\boldsymbol{r}_i^{new} = \alpha_P^{1/3} \boldsymbol{r}_i^{old}$$

and by choosing

$$\alpha_P(t) = 1 - \kappa_T \frac{\Delta t}{\tau_P} \left[P_{eq} - \mathcal{P}(t)\right] \tag{146}$$

the desired exponential decay is obtained. Here

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \tag{147}$$

is the isothermal compressibility.

F The Fourier Transform

In this appendix the Fourier Transform is presented, which is used in many different areas in science and engineering. The Discrete Fourier Transform (DFT) is introduced, which is instrumental in computational approaches, as well as the Fast Fourier Transform (FFT) technique, that evaluates the DFT in a very efficient way and has therefore revolutionized computational approaches in several areas of science and engineering.

F.1 The continuous Fourier Transform (FT)

Consider a physical process in time domain described by the values of some quantity

$$h(t) \quad \text{with} \quad -\infty < t < \infty$$

The same process can be described in frequency space. It is then specified by an amplitude $H(\omega)$, generally a complex number, as function of angular frequency ω. The change between these two representations are made using the *Fourier transform* equations

$$H(\omega) = \int_{-\infty}^{\infty} h(t) e^{i\omega t}\, \mathrm{d}t \tag{148}$$

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) e^{-i\omega t}\, \mathrm{d}\omega \tag{149}$$

where the orthogonality relation takes the form

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega\, e^{-i\omega t}\, e^{i\omega t'} = \delta(t - t') \tag{150}$$

The Fourier transform is also used for other quantities. For instance, if h is a function of position $\boldsymbol{r}$, the inverse will be a function of the wavevector $\boldsymbol{k}$. In Eq. (148) the Fourier transform is defined using the angular frequency ω. This is quite common among physicists and mathematicians. However, one can also use the ordinary frequency

$$f = \omega/2\pi \tag{151}$$

In that case the Fourier transform equations take the more symmetrical form

$$H(f) = \int_{-\infty}^{\infty} h(t) e^{2\pi i f t}\, \mathrm{d}t \tag{152}$$

$$h(t) = \int_{-\infty}^{\infty} H(f) e^{-2\pi i f t}\, \mathrm{d}f \tag{153}$$

where $H(f) \equiv [H(\omega)]_{\omega=2\pi f}$. The factor $1/2\pi$ disappears and the orthogonality relation takes the form

$$\int_{-\infty}^{\infty} \mathrm{d}f\, e^{-2\pi i f t}\, e^{2\pi i f t'} = \delta(t - t') \tag{154}$$

Furthermore, the transition to discretely sampled data and the discrete Fourier transform (DFT) becomes more straightforward. We will therefore here stick to this definition.

F.1.1 Power spectrum

The power spectral density

$$P(f) = |H(f)|^2$$

of a time series $h(t)$ describes the distribution of power into frequency components composing that signal. If the signal $h(t)$ goes endlessly from $-\infty$ to ∞ then its power spectral density will, in general, be infinite and one then consider the power spectral density per unit time. Introduce

$$h_{\mathsf{T}}(t) = \begin{cases} h(t) & -\mathsf{T}/2 < t < \mathsf{T}/2 \\ 0 & \text{otherwise} \end{cases}$$

with the Fourier transform

$$\begin{aligned} H_{\mathsf{T}}(f) &= \int_{-\infty}^{\infty} h_{\mathsf{T}}(t) e^{2\pi i f t}\, \mathrm{d}t \\ &= \int_{-T/2}^{T/2} h(t) e^{2\pi i f t}\, \mathrm{d}t \end{aligned}$$

The *power spectral density per unit time* is then defined as

$$\mathcal{P}(f) = \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \mid H_{\mathsf{T}}(f) \mid^2 \tag{155}$$

F.1.2 Time-correlation function

Another important quantity is the *time-correlation function.* The autocorrelation of $h(t)$ at different times is defined as

$$C(t) = \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \int_0^{\mathsf{T}} h(t+t') h^*(t')\, \mathrm{d}t' \tag{156}$$

For a system in equilibrium the correlation function is stationary, *i.e.* it is invariant under a translation of the time origin.

F.1.3 Wiener-Kintchin's theorem

There is an important relation between the power spectrum and the time-correlation function, known as *the Wiener-Khinchin's theorem* or *the correlation theorem.* It states that the power spectral density is equal to the Fourier transform of the corresponding time-correlation function,

$$\mathcal{P}(f) = \int_{-\infty}^{\infty} C(t) e^{2\pi i f t}\, \mathrm{d}t \tag{157}$$

To prove the Wiener-Khinchin's theorem consider the Fourier transform of the power spectral density

$$\begin{aligned}
&\int_{-\infty}^{\infty} \mathcal{P}(f) e^{-2\pi i f t}\, \mathrm{d}f \\
&= \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \int_{-\infty}^{\infty} \mathrm{d}f\, e^{-2\pi i f t} H_{\mathsf{T}}(f) H_{\mathsf{T}}^*(f) \\
&= \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \int_{-\infty}^{\infty} \mathrm{d}f\, e^{-2\pi i f t} H_{\mathsf{T}}(f) \int_{-\infty}^{\infty} \mathrm{d}t'\, e^{-2\pi i f t'} h_{\mathsf{T}}^*(t') \\
&= \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \int_{-\infty}^{\infty} \mathrm{d}t'\, h_{\mathsf{T}}^*(t') \int_{-\infty}^{\infty} \mathrm{d}f\, e^{-2\pi i f (t+t')} H_{\mathsf{T}}(f) \\
&= \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \int_{-\infty}^{\infty} \mathrm{d}t'\, h_{\mathsf{T}}(t+t') h_{\mathsf{T}}^*(t')
\end{aligned}$$

Notice that

$$\begin{aligned}
&\int_{-\infty}^{\infty} \mathrm{d}t'\, h_{\mathsf{T}}(t+t') h_{\mathsf{T}}^*(t') \\
&= \begin{cases} \int_{-T/2}^{T/2} h(t+t') h^*(t')\, \mathrm{d}t' - \int_{-T/2}^{-T/2+t} h(t+t') h^*(t')\, \mathrm{d}t' & \text{if } 0 < t < T \\ \int_{-T/2}^{T/2} h(t+t') h^*(t')\, \mathrm{d}t' - \int_{T/2+t}^{T/2} h(t+t') h^*(t')\, \mathrm{d}t' & \text{if } -T < t < 0 \\ 0 & \text{otherwise} \end{cases} \\
&= \int_{-T/2}^{T/2} h(t+t') h^*(t')\, \mathrm{d}t' - |t|\mathcal{O}(1) \quad \text{if} |t| < T
\end{aligned}$$

For a fix t with $|t| < T$ we get

$$\begin{aligned}
&\int_{-\infty}^{\infty} \mathcal{P}(f) e^{-2\pi i f t}\, \mathrm{d}f \\
&= \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \left[\int_{-T/2}^{T/2} h(t+t') h^*(t')\, \mathrm{d}t' - |t|\mathcal{O}(1) \right] \\
&= C(t)
\end{aligned}$$

where we have used that the correlation function is stationary in time, *i.e.* it only depends on the time lag t. This proves the Wiener-Khintchin's theorem; the inverse Fourier transform of the power spectral density is equal to the corresponding correlation function.

By integrating the Wiener-Khinchin's theorem in Eq. (157) over all frequencies

$$\int_{-\infty}^{\infty} \mathcal{P}(f)\, \mathrm{d}f = \int_{-\infty}^{\infty} C(t)\, \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}f e^{2\pi i f t} = C(0)$$

we obtain the *Parseval's theorem*

$$\int_{-\infty}^{\infty} \mathcal{P}(f)\, \mathrm{d}f = \lim_{\mathsf{T}\to\infty} \frac{1}{\mathsf{T}} \int_{-T/2}^{T/2} |h(t')|^2\, \mathrm{d}t' \tag{158}$$

for the total power. This last identity makes it clear that, given any two frequencies f_1 and f_2, the quantity

$$\int_{f_1}^{f_2} \mathcal{P}(f)\,\mathrm{d}f$$

represents the portion of the total signal power contained in the signal between the frequencies f_1 and f_2 and hence, $\mathcal{P}(f)$ is indeed a "spectral density".

F.2 The Discrete Fourier Transform (DFT)

The most common situation is that one has access to data for $h(t)$ sampled evenly spaced in time. We denote the sampling interval in time with Δt and write

$$h_k = h(k\Delta t) \quad , \quad k = \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

The reciprocal of the time interval Δt is called the sampling rate.

F.2.1 Sampling theorem and aliasing

For each sampling interval Δt there is a special frequency f_c, the *Nyquist critical frequency*, given by

$$f_c \equiv \frac{1}{2\Delta t} \tag{159}$$

Consider a signal with frequency f_c, $h(t) \propto \sin(2\pi f_c t)$, which is sampled with the sampling interval $\Delta t = 1/2f_c$. If a sample point is located at the positive peak value of this signal then the next sample point will be at its negative trough value, and the next at the positive peak again, and so on. This is the critical sampling interval for a component with frequency f_c.

If a continuous signal $h(t)$ is bandwidth limited to frequencies less than f_c, *i.e.* $H(f) = 0\ \forall |f| > f_c$, then the function $h(t)$ is completely determined if it is sampled discretely with the sampling interval $\Delta t = 1/2f_c$. The continuous function $h(t)$ can be exactly recovered using the so called *sampling theorem*

$$h(t) = \Delta t \sum_{k=-\infty}^{\infty} h_k \frac{\sin\left[2\pi f_c(t - k\Delta t)\right]}{\pi(t - k\Delta t)} \tag{160}$$

The question is what happens if the signal is not bandwidth limited to less than the Nyquist critical frequency. Unfortunately, all power spectral density that lies outside the frequency range $|f| \leq f_c$ will then be spuriously moved into that range. This phenomenon is called *aliasing*. Any frequency component outside the range $|f| \leq f_c$ is aliased (falsely translated) into that frequency range by the very act of discrete sampling. There is little one can do to remove aliased power once the signal has been discretely sampled. To overcome aliasing one can either redo the sampling with a sufficiently small

sampling rate if the bandwidth of the signal is known or otherwise one can enforce a known limit by analog filtering of the continuous signal $h(t)$ and then resample the signal with the appropriate time interval Δt.

F.2.2 The transform

We now want to estimate the Fourier transform of a function $h(t)$ from a finite value a sampled points. The number of sampled points is N and they are sampled equidistant with the sampling interval Δt, *i.e.*

$$h_k = h(t_k), \quad \text{with } t_k \equiv k\Delta t \ \text{ and } k = 0, 1, 2, \ldots, N-1$$

This corresponds to the maximum time

$$t_{\max} = N\Delta t \tag{161}$$

and the Nyquist critical frequency $f_c = 1/(2\Delta t)$. With N input numbers in time we can generate at maximum N output numbers in frequency. These are chosen equidistant between $-f_c$ and f_c, *i.e.*

$$f_n = n\,\Delta f; \quad n = -\frac{N}{2}, \ldots, \frac{N}{2} - 1$$

where

$$\Delta f = \frac{2f_c}{N} = \frac{1}{N\Delta t} = \frac{1}{t_{\max}} \tag{162}$$

and where we have assumed N to be even. The Fourier transform in Eq. (152) can now be approximated by

$$H(f_n) = \int_{-\infty}^{\infty} h(t) e^{2\pi i f_n t}\,\mathrm{d}t \approx \Delta t \sum_{k=0}^{N-1} h_k\, e^{2\pi i f_n t_k} = \Delta t \sum_{k=0}^{N-1} h_k\, e^{2\pi i k n/N}$$

where the last sum is called the *discrete Fourier transform* of the sequence h_k. We denote this as

$$H_n = \sum_{k=0}^{N-1} h_k\, e^{2\pi i k n/N} \tag{163}$$

We notice that H_n is periodic in N, $H_{n+N} = H_n$. Above we assumed that n varied from $-N/2$ to $N/2-1$ but we can also, based on the periodicity of H_n, let n varies from 0 to $N-1$. The formula for the discrete *inverse* Fourier transform then takes the from

$$h_k = \frac{1}{N} \sum_{n=0}^{N-1} H_n\, e^{-2\pi i k n/N} \tag{164}$$

Essentially the same computer routine can then be used to evaluate discrete Fourier transform and its inverse. However, one then has to remember that zero frequency $f = 0$ corresponds to $n = 0$, positive frequencies $0 < f < f_c$ to $1 \le n \le N/2-1$ and negative frequencies $-f_c \le f < 0$ to $N/2 \le f \le N-1$.

F.2.3 Summary - DFT

Time domain

Sampling rate (time spacing): Δt

$$h_k = h(t_k); \quad t_k = k\Delta t, \quad k = 0, \ldots, N-1$$

Nyqvist frequency: $f_c = 1/(2\Delta t)$

Frequency domain

Frequency spacing: $\Delta f = 1/(N\Delta t)$

$$H_n = \frac{1}{\Delta t} H(f_n); \quad f_n = \begin{cases} n\Delta f & n = 0, \ldots, N/2-1 \\ (n-N)\Delta f & n = N/2, \ldots, N-1 \end{cases}$$

Nyqvist frequency: $f_c = (N/2)\Delta f = 1/(2\Delta t)\ , \quad -f_c < f < f_c$

Discrete Fourier transform

$$H_n = \sum_{k=0}^{N-1} h_k \, e^{2\pi i k n/N} \tag{165}$$

Inverse Discrete Fourier transform

$$h_k = \frac{1}{N} \sum_{n=0}^{N-1} H_n \, e^{-2\pi i k n/N} \tag{166}$$

Orthogonality relation

$$\frac{1}{N} \sum_{n=0}^{N-1} e^{-2\pi i k n/N} \; e^{2\pi i k' n/N} = \delta_{k,k'} \tag{167}$$

Power spectrum

$$P_n = \frac{1}{N} |H_n|^2 \tag{168}$$

Parseval's theorem

$$\sum_{k=0}^{N-1} |h_k|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |H_n|^2 \tag{169}$$

Periodic extension

The function h_k and its inverse H_n are both periodic with period N

$$h_k = h_{k+N}$$
$$H_n = H_{n+N}$$

The circular correlation theorem

$$\sum_{k'=0}^{N-1} g_{k+k'}\, h^*_{k'} = \frac{1}{N} \sum_{n=0}^{N-1} G_n\, H^*_n\, e^{-2\pi i k n/N} \tag{170}$$

The circular convolution theorem

$$\sum_{k'=0}^{N-1} g_{k'}\, h_{k-k'} = \frac{1}{N} \sum_{n=0}^{N-1} G_n\, H_n\, e^{-2\pi i k n/N} \tag{171}$$

F.3 The Fast Fourier Transform (FFT)

A main reason for the wide applicability of Fourier methods in science and engineering is the existence of the *Fast Fourier Transform* (FFT) algorithm. FFT simply computes the Discrete Fourier Transform (DFT), but in a very efficient way. The FFT algorithm was chosen as one of the 10 most important algorithms of the 20th century [26].

The Discrete Fourier Transform (DFT) is given by the expression

$$H_n = \sum_{k=0}^{N-1} W_N^{nk} h_k$$

where we have introduced the complex number

$$W_N \equiv e^{2\pi i/N}$$

W_N^{nk} is a complex matrix with $N \times N$ elements. To obtain H_n for $n = 0, 1, \ldots, N-1$ we therefore essentially need to do N^2 complex multiplications. So the DFT appears to be an $\mathcal{O}(N^2)$ process.

The different Fast Fourier Transform (FFT) algorithms make use of the fact that W_N^{nk} is periodic in N, *i.e.* $W_N^{n'k'} = W_N^{nk}$ if $n'k' = nk+N$. W_N^{nk} does not contain N^2 different elements, only N, and the FFT avoids calculating a number that already has been calculated. Using clever reorganisation of the different terms the FFT then reduces the computational complexity to $\mathcal{O}(N \log_2 N)$ operations. One can show that an algorithm with less operations than the FFT can not be constructed in evaluating the DFT [27]. In that respect the FFT is most efficient algorithm to determine the DFT.

The difference between $\mathcal{O}(N^2)$ and $\mathcal{O}(N \log_2 N)$ matters. Consider a sequence with $N = 10^6$ and a Mflops computer. In that case

$$\begin{aligned} &\text{DFT} : N^2 && \Rightarrow 10^{12} \text{ operations} \sim 10^6\,\text{s} \sim 278\,\text{h} \sim 12\,\text{days} \\ &\text{FFT} : N \log_2 N && \Rightarrow 10^6 \log_2(10^6) \text{ operations} \sim 20\,\text{s} \end{aligned}$$

a substantial payoff in efficiency.

F.3.1 The Fastest Fourier Transform in the West (FFTW)

The FFT is the most efficient algorithm to determine the DFT. Still there is room for improvements. The speed also depends on how efficiently the code make use of the hardware. Matteo Frio and Steven G. Johnson at MIT developed the "Fastest Fourier Transform in the West" (FFTW)and won the prestigious J. H. Wilkinson Numerical Software Prize 1999.

The FFTW project is an attempt to implement a self-optimising lirary for computing FFTs. It is freely available at http://www.fftw.org. The FFTW package tries all factorisations of N and selects the best for the particular hardware. It makes efficient use of th ememory (cache sizes, access patterns, and data locality). FFTW often outperforms the manufactures own highly tuned codes.

The FFTW package was implemented in MATLAB version 6. Test calculations using $N = 10^6$ and a 266 MHz Pentium laptop gave the result:

MATLAB version 5.3 :	6s
MATLAB version 6 :	1.2s

Not too bad.

G The Fast Correlation Algorithm

It is possible to improve on the speed of calculating the time correlation function in Eq. (74) by making use of the fast Fourier Transform (FFT) technique. The discrete version of the Wiener-Khintchin's theorem, the circular correlation theorem Eq. (170), is then being used. The technique is often called the *Fast Correlation Algorithm* and it reduces the number of operations greatly.

It is important to realize that the circular correlation theorem presumes that the input trajectory is periodic in time. That is not the case for the real data. When applying the circular correlation theorem it will falsely correlate input data from the beginning of the data set with some wrapped-around data from the end of the data set. To circumvent this one has to add a buffer zone with zero-padded values at the end of the input data set.

Consider a dynamical quantity $\mathcal{A}(t)$ that is sampled equidistant in time with sampling distance $\Delta\tau$,

$$\mathcal{A}_k = \mathcal{A}(k\Delta\tau) \quad ; \quad k = 0, 1, \ldots, (M-1)$$

The corresponding auto-correlation function is then given by

$$C_l = \begin{cases} \frac{1}{M-l} \sum_{m=0}^{M-l-1} \mathcal{A}_{m+l}\mathcal{A}_m \; ; & l = 0, 1, \ldots, M-1 \\ \frac{1}{M-|l|} \sum_{m=|l|}^{M-1} \mathcal{A}_{m-|l|}\mathcal{A}_m \; ; & l = -(M-1), \ldots, -1 \end{cases}$$

where $t = l\Delta\tau$ is the time lag and we have used the fact the correlation function is even in time. To avoid spurious correlation we add M zeros at the end of the input data set and define

$$h_k = \begin{cases} \mathcal{A}_k \; ; & k = 0, 1, \ldots, M-1 \\ 0 \; ; & k = M, \ldots, N-1 \end{cases}$$

where $N = 2M$. We assume that h_k is periodic in k, *i.e*

$$h_{k+N} = h_k$$

and introduce the correlation function

$$S_l = \sum_{k=0}^{N-1} h_{l+k} h_k^* = (M - |l|)\; C_l$$

From the circular correlation theorem we then have

$$S_l = \frac{1}{N} \sum_{n=0}^{N-1} |H_n|^2 \, e^{2\pi i l n/N}$$

where H_n is the inverse discrete Fourier transform of h_k.

References

[1] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, 2nd ed., Academic Press, 2002.

[2] M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids*, Clarendon Press, 1989.

[3] J. M. Haile, *Molecular Dynamics Simulation*, John Wiley & Sons, 1992.

[4] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed., Addison Wesley, 2002.

[5] R. K. Pathria and P. D. Beale, *Statistical Mechanics*, Academic Press, 2011.

[6] S.-K. Ma, *Statistical Mechanics*, World Scientific, 1985.

[7] A. Rahman, Phys. Rev. **136A**, 405 (1964).

[8] A. E. Carlsson, Solid State Phys. **43**, 1 (1990).

[9] F. Ercolessi, E. Tosatti, and M. Parrinello, Phys. Rev. Lett. **57**, 719 (1986).

[10] K. W. Jacobsen, Comments Cond. Mat. Phys. **14**, 129 (1988).

[11] M. S. Daw, S. M. Folies, and M. I. Baskes, Materials Sci. Rep. **9**, 251 (1993).

[12] M. Finnis and J. Sinclair, Philos. Mag. A **50**, 45 (1984).

[13] F. H. Stillinger and T. A. Weber, Phys. Rev. B **31**, 5262 (1985).

[14] M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).

[15] L. Verlet, Phys. Rev. **159**, 98 (1967).

[16] G. Dahlquist and A. Björk, *Numerical methods* (Prentice Hall, 1974).

[17] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).

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

[19] J. M. Thijssen, *Computational Physics*, 2nd ed., Cambridge, 2007.

[20] B. J. Alder and T. E. Wainwright, Phys. Rev. Lett. **18**, 988 (1967).

[21] J. P. Boon and S. Yip, *Molecular Hydrodynamics*, Dover 1992.

[22] L. van Hove, Phys. Rev. **95**, 249 (1954).

[23] R. Zwanzig, Ann. Rev. Phys. Chem. **16**, 67 (1965).

[24] R. C. Tolman, *The principles of statistical mechanics*, Clarendon Press, 1938.

[25] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, Mol. Phys. **87**, 1117 (1996).

[26] J. Dongarra and F. Sullivan, (January 2000). "Guest Editors Introduction to the top 10 algorithms". Computing in Science Engineering **2** 22 (2000).

[27] S. Winograd, CBMSNSF Reeional Conf. Series in Applied Mathematics, Vol. 33, SIAM, Philadelphia, 1980.