

Sorbonne Université

2021-2022

Stage de Master 2

Noé Blassel

Titre

Projet réalisé en collaboration avec le CERMICS
Ecole Nationale des Ponts et Chaussées
6 et 8 avenue Blaise Pascal
Cité Descartes - Champs sur Marne
77455 Marne la Vallée Cedex 2

Tuteur : Gabriel Stoltz

March 23, 2022

Contents

1	Introduction to molecular dynamics	3
1.1	The microscopic description of atomic systems	3
1.2	Reduced units	6
1.3	Statistical ensembles	6
1.4	From microscopic dynamics to macroscopic observables	7
2	Sampling equilibrium ensembles	9
2.1	Microcanonical averages	9
2.1.1	Properties of Hamiltonian dynamics	9
2.1.2	Numerical schemes for Hamiltonian dynamics	10
2.1.3	Energy conservation properties	11
2.1.4	Examples of instantaneous observables	11
2.1.5	Shortcomings of the Hamiltonian approach	11
2.2	Canonical averages	11
2.2.1	Properties of the Langevin dynamics	12
2.2.2	Numerical schemes for the Langevin dynamics	13
2.2.3	Choosing a scheme	14
2.2.4	Unbiased sampling	16
2.2.5	Illustration: the equation of state of Argon	16
3	Transport coefficients	17
4	A Norton method	18
A	Performance analysis	19
B	Conventions	20

Chapter 1

Introduction to molecular dynamics

1.1 The microscopic description of atomic systems

Molecular dynamics, and computational statistical physics at large, aim at simulating on the computer the behavior of physical systems. The hope is that one can infer quantities and properties of real-life interest from observing the results of numerical simulations, which may be relevant to understand the material properties of many-particle systems, or the nature of interactions in complex systems such as those found in biology. Computational simulations can thus act as surrogate experiments in cases where experimental setups are hard to achieve, or measurements are impossible. They can also be seen as surrogate tests of theoretical models, as they allow to test the validity of a mathematical description by comparing numerical predictions to experimental data. Molecular dynamics, in particular, is concerned with simulating atomic systems, most often (and as we shall systematically do) using a classical description.

We consider a system of N particles evolving in d -dimensional space. The classical description contends that the *state* of a system is the datum of the positions and momenta of every particle in the system. We can interpret this as the statement that, given full knowledge of the positions and momenta at some initial time, and of the forces at play, one can deduce exactly the positions and momenta at any future time.

Definition 1 (Phase space). *We describe the positions and momenta of the atoms as vectors*

$$q = (q_{1,1}, \dots, q_{1,d}, \dots, q_{N,1}, \dots, q_{N,d})^\top \in \mathbb{R}^{dN},$$

$$p = (p_{1,1}, \dots, p_{1,d}, \dots, p_{N,1}, \dots, p_{N,d})^\top \in \mathbb{R}^{dN},$$

where $q_i := (q_{i,1}, \dots, q_{i,d})^\top$ is the position vector of the i -th particle, and similarly for p . In practice, it is often the case that each q_i is restricted to some d -dimensional manifold \mathcal{D} , called the configuration space. For our purposes, we will always take $\mathcal{D} = \mathbb{R}^d$ or $\mathcal{D} = (L\mathbb{T})^d$, where $L > 0$ is some size parameter. Phase space, then, is the set of possible microscopic states of the system, that is, the set

$$\mathcal{E} = \mathcal{D}^N \times \mathbb{R}^{dN}$$

Trajectories through phase space, that is functions

$$\begin{cases} \mathbb{R}_+ \mapsto \mathcal{E} \\ t \mapsto (p_t, q_t) \end{cases},$$

can be seen as describing time evolutions of the system, objects which will be central to our study.

It is not clear *a priori* why we should choose momenta to describe the kinetic quality of the system, rather than velocities. However it is of no importance since we can change from one description to the other via the relation

$$v = M^{-1}p,$$

where $M \in \mathbb{R}^{dN \times dN}$ is a diagonal matrix recording the masses of each particle (d times per particle), and v is the velocity vector.

In order to describe the evolution of the system's state, one must specify a dynamical law. This is done by giving a function

$$\begin{cases} \mathcal{D} \mapsto \mathbb{R} \\ q \mapsto V(q) \end{cases},$$

V whose gradient in the i -th particle's coordinates

$$\nabla_{q_i} V := (\partial_{q_{i,d}}, \dots, \partial_{q_{i,d}})^\top$$

gives minus the force vector acting on the i -th particle. In the case where $\mathcal{D} = (L\mathbb{T})^{dN}$, it will be convenient to think of V as a function from \mathbb{R}^{dN} to \mathbb{R} which is C^1 and L -periodic in each direction.

V is called the potential, and, as it encodes the dynamics of the system, it is of paramount importance. The time evolution of the system, then, is described by Newton's second law:

$$\frac{dp}{dt} = -\nabla V(q)$$

It will be convenient for our analysis to use of reformulation of Newton's equations, based on the Hamiltonian of a system.

Definition 2 (Hamiltonian). *The Hamiltonian of a classical system is its total energy, which is the sum of a kinetic energy term depending only on the momenta and a potential energy term depending only on the positions.*

$$(1.1) \quad H(q, p) = \frac{1}{2} p^\top M^{-1} p + V(q)$$

Using the Hamiltonian, we can rewrite the classical equations of motion as

$$(1.2) \quad \begin{cases} dq_t = M^{-1} p_t dt = \nabla_p H(q_t, p_t) dt \\ dp_t = -\nabla V(q_t) dt = -\nabla_q H(q_t, p_t) dt \end{cases},$$

The potential is the most important part of the microscopic description, and accordingly, the main problem in establishing a physical model of this kind is to determine potential functions which adequately capture the dynamic behavior of a given system. The choice of a classical description automatically implies a degree of approximation, since behavior arising from the laws of quantum mechanics, which may be relevant at a microscopic level, are described by Newton's law. Furthermore, if the aim is to simulate such systems numerically, computational constraints imply that some compromise has to be reached between theoretical accuracy and computational cost. If, for small systems, it may be possible to simulate all atomic interactions, for larger or more complex systems, it is often to use potential functions which are both cheap from a computational point of view and empirically shown to be accurate enough for the purpose of a simulation.

Our main numerical example will be the system given by the following potential, which is of this empirical form, and which is often used to describe the microscopic behavior of chemically inert fluids, such as Argon.

Example 1 (The Lennard-Jones fluid). We fix $L > 0$, $d = 3$, and N the number of particles. The Lennard-Jones fluid is the classical system given by the potential

$$(1.3) \quad V_{\text{LJ}}(q) = \sum_{i=1}^N \sum_{j < i} 4\varepsilon \left(\left(\frac{|q_i - q_j|}{\sigma} \right)^{-12} - \left(\frac{|q_i - q_j|}{\sigma} \right)^{-6} \right).$$

Note that V_{LJ} is given by a sum over pairs of particles,

$$V_{\text{LJ}}(q) = \sum_{1 \leq i < j \leq N} v(|q_i - q_j|),$$

where v is a radial function

$$v(r) = 4\varepsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right).$$

ε , an energy, and σ , a length, are shape parameters which respectively control the depth of the potential well of v and the equilibrium distance $2^{1/6}\sigma$. As seen on Figure 1.1, the potential combines two effects. At small interparticle distances, the dominant term is in r^{-12} , which translates into a strongly repulsive force between close pairs of particles, and makes individual particles essentially impenetrable. At long range, the dominant term is in $-r^6$, which translates into a weakly attractive force between distant particles. Contrary to the repulsive term, which is empirical, this scaling has a theoretical origin in the Van der Waals forces. From a computational standpoint, the fact that v is an even function of r allows one to compute the normalized force while sparing the expense of computing a square root, while the identity $r^{12} = (r^6)^2$ allows further economy. The shape parameters σ and ε must be chosen empirically to describe the behavior of a particular atomic species. For Argon, values of reference are: $\sigma = 3.4$, $\varepsilon = 1.65 \times 10^{-21}$ J.

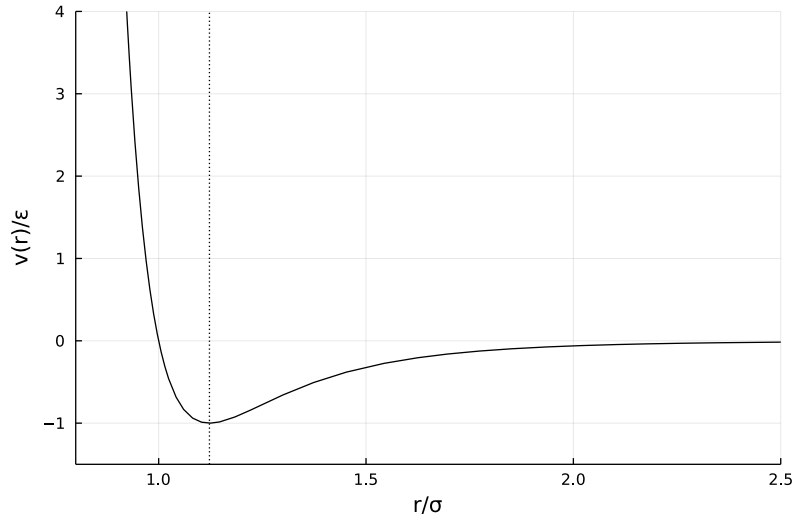


Figure 1.1: The pair potential v , with lengths and energy given in reduced units. The equilibrium interparticle distance is indicated by the vertical dotted line.

1.2 Reduced units

1.3 Statistical ensembles

The microscopic description is interesting from a theoretical standpoint, but it fails to be relevant when attempting to describe the behavior of atomic systems with a macroscopic number of particles, of the order of Avogadro's number (6.02×10^{23}). Besides the technical impossibility of measuring to a high accuracy the configuration of such systems, and that of recording the information required to track it (coincidentally, the total amount of digitally stored information on Earth is estimated to be 10^{23} bytes as of 2022), it is also the case that knowledge of a system at this level of detail is unnecessary to describe the quantities which are relevant to our macroscopic experience. In the instance of a gas at thermal equilibrium, examples of relevant quantities are total energy, pressure, temperature, density, which, while of course resulting from the internal state of the system, are independent of the minutiae of individual atomic motions: loosely speaking, one may describe the macroscopic state of a system by only a handful of macroscopic variables, losing track of the myriad of microscopic degrees of freedom. An important point is that for a given macroscopic state, there are many microscopic configurations which are compatible with our observations. This motivates defining the macroscopic state of a system as a probability distribution over phase space, which we may interpret as assigning to each microscopic configuration a likelihood that this configuration underlies the macroscopic state.

This does not tell one how to choose the distribution over microscopic states. However, it seems reasonable to assign positive probabilities to states compatible with the macroscopic constraints, and in such a way as to make the weakest possible assumptions on this microscopic state, or in other words contain the least amount of information about the system, given the macroscopic constraints. The mathematical translation of this idea is given by the principle of maximal entropy. Given a class of probability distributions compatible with the macroscopic constraints, define the macroscopic state as the one which maximizes the entropy, which is defined for a probability distribution ρ by

$$(1.4) \quad \mathfrak{S}(\rho) = - \int_{\mathcal{E}} \rho(x) \ln(\rho(x)) dx.$$

The specification of a probability distribution over states is called a thermodynamic ensemble. We will be considering two ensembles:

Example 2 (Microcanonical ensemble). *The microcanonical ensemble is the suitable model for an isolated system in thermodynamic equilibrium, evolving under Hamiltonian dynamics. The number of particles N , the volume $V = L^3$, and the energy E is fixed. We will alternatively refer to the microcanonical ensemble as the NVE ensemble. Because the constant energy condition constrains the compatible microstates to level sets of H , which in general will be negligible subsets of \mathcal{E} , some care must be taken in defining the microcanonical measure, since one cannot express the macroscopic constraints by a family of probability densities. However, under suitable assumptions on V , one can define the microcanonical measure as a weak limit of uniform distributions over level “shells” of H :*

$$\int_{\mathcal{E}} \varphi d\mu_{NVE} := \lim_{\varepsilon \rightarrow 0} \frac{1}{|S(E, \varepsilon)|} \int_{S(E, \varepsilon)} \varphi(q, p) dq dp,$$

where

$$S(E, \varepsilon) = \{(q, p) \in \mathcal{E} \mid H(q, p) \in [E - \varepsilon, E + \varepsilon]\}.$$

This is consistent with the fact that, for a set A with finite Lebesgue measure, the probability distribution on A which maximizes the entropy is the uniform distribution on A . It is possible, using the coarea formula, to derive a precise expression for this limit, which is, however, of limited practical interest.

Example 3 (Canonical ensemble). *Isolated systems in thermal equilibrium are not typically those that we encounter in experiments. Instead, it is more common to observe systems which are in thermal equilibrium with respect to their environment, an ambient heat bath at a fixed temperature. The total energy of such systems is not fixed: small fluctuations can occur as energy is exchanged back and forth between the heat bath and the system. However, the average energy \bar{E} is fixed. This is the macroscopic constraint that defines the canonical ensemble. For a fixed N, V, \bar{E} , define the density of the canonical measure as the maximizer:*

$$\operatorname{argmax}_{\rho \in \mathcal{A}} \mathfrak{S}(\rho)$$

where \mathcal{A} is the set of admissible densities

$$\mathcal{A} = \{\rho : \mathcal{E} \mapsto \mathbb{R}_+ \mid \int_{\mathcal{E}} \rho = 1, \int_{\mathcal{E}} H(q, p) \rho(q, p) dq dp = E\}.$$

Solving the Euler-Lagrange equation associated with this constrained optimization problem yields that the only admissible solution can be written under the form:

$$\rho^*(q, p) = \frac{1}{Z} e^{-\beta H(q, p)}.$$

Furthermore, one can show that ρ^* is indeed the unique maximizer. Here, $-\beta$ and $1 + \ln Z$ are the critical Lagrange multipliers associated respectively with the energy constraint and the normalization constraint. Thus

$$Z = \int_{\mathcal{E}} e^{-\beta H(q, p)} dq dp$$

is a normalization constant called the partition function, and β is a tuning parameter related to the value of \bar{E} . The physical interpretation of β is that of an inverse temperature,

$$\beta = \frac{1}{k_B T},$$

where $k_B = 1.38 \times 10^{-23} \text{ J} \cdot \text{K}^{-1}$ is Boltzmann's constant. For obvious reasons, we prefer to refer to the canonical ensemble as the NVT ensemble (rather than the NVE)

Remark 1. One can go further and observe that when observing a fixed volume of unconfined gas in thermal equilibrium, the total number of particles N is not fixed. Instead, this fluctuates as particles are constantly exchanged with an ambient particle reservoir. Instead, the average number of particles \bar{N} is fixed. The resulting ensemble is called the grand canonical or μVT ensemble. This, and many other constructions are possible, but we will restrict our attention to the NVE and NVT cases.

The main interest in obtaining such a description is that one can then express the macroscopic state of a system in terms of averages of microscopic observables with respect to the ensemble measure.

1.4 From microscopic dynamics to macroscopic observables

Definition 3 (Observables). *An observable is simply a measurable function*

$$\varphi : \mathcal{E} \mapsto \mathbb{R}$$

One can also consider vector-valued observables, such as the velocity. The potential energy V (viewed as a function on \mathcal{E}), the kinetic energy $(q, p) \mapsto p^\top M^{-1} p / 2$ and the Hamiltonian H are obvious examples of observables. Observables relate the microscopic state of the system to macroscopic quantities, through ensemble averages. Given a macroscopic system defined by an

equilibrium ensemble μ and an observable φ , we are interested in computing its average over the ensemble

$$(1.5) \quad \mathbb{E}_\mu[\varphi] := \int_{\mathcal{E}} \varphi(q, p) \mu(dq, dp).$$

Two problems arise. One is, for a given macroscopic quantity of interest, how to define the microscopic observable φ which corresponds to this quantity. The second, once we have properly defined φ , is how to compute averages (1.5). This second problem is one of statistically sampling a high-dimensional measure, which is a difficult problem in general. Our broad strategy, which will remain the same in the NVE and NVT case, is to define dynamical processes $(X_t)_{t \geq 0}$ on \mathcal{E} which are invariant for the target measure μ , and then consider estimates of $\mathbb{E}_\mu[\varphi]$ through the computation of ergodic averages:

$$(1.6) \quad \frac{1}{T} \int_0^T \varphi(X_t) dt,$$

which we may hope will converge to the target value. The convergence of ergodic averages can be shown not to hold in generality, and is something which must be proven on a case by case basis, though general criteria can be derived. Furthermore, if the underlying dynamic is stochastic, then the variance of the random variables (1.6) becomes an issue, which one must control. However, empirical practice shows that ergodic averages obtained from computer simulations, even for a modest number of atoms, agrees very well with experimental data for certain types of systems, even in the absence of theoretical guarantees. We will demonstrate an example in the next chapter.

Chapter 2

Sampling equilibrium ensembles

2.1 Microcanonical averages

2.1.1 Properties of Hamiltonian dynamics

The Hamiltonian dynamics (1.2) rewrites in matrix form, writing $X_t = (q_t, p_t)$:

$$(2.1) \quad dX_t = J \nabla H(X_t) dt,$$

where J is the symplectic matrix

$$J = \begin{pmatrix} 0_{dN} & \text{Id}_{dN} \\ -\text{Id}_{dN} & 0_{dN} \end{pmatrix}$$

This will be useful to investigate properties of the Hamiltonian dynamics. First, using the chain rule, we obtain the following property.

Proposition 1 (Energy conservation).

$$(2.2) \quad dH(X_t) = dX_t^\top \nabla H(X_t) = (J \nabla H(X_t))^\top \nabla H(X_t) dt = 0$$

This relation expresses the fact that the Hamiltonian is invariant under the flow of (1.2). This, in turn, is the mathematical translation of the physical principle of conservation of energy.

More generally, we may apply the chain rule to any smooth function $\varphi : \mathcal{S} \mapsto \mathbb{R}$. We obtain

$$d\varphi(X_t) = dX_t^\top \nabla \varphi(X_t) = (J \nabla H(X_t))^\top \nabla \varphi(X_t) dt = (\nabla_p H \cdot \nabla_q - \nabla_q H \cdot \nabla_p) \varphi(X_t) dt$$

This motivates the following.

Definition 4 (Generator of the Hamiltonian dynamics). *We define the generator associated with the Hamiltonian dynamics to be the operator \mathcal{L}_H defined on smooth functions by*

$$(2.3) \quad \mathcal{L}_H \varphi = (\nabla_p H \cdot \nabla_q - \nabla_q H \cdot \nabla_p) \varphi$$

The generator quantifies the rate of change of a property φ under the evolution of the system. If we define, for $t \geq 0$, the evolution operators

$$P_t \varphi(q_0, p_0) = \varphi(\Phi_t(q_0, p_0)),$$

where Φ is the flow associated with the Hamiltonian dynamics, that is the collection of maps $(\Phi_t)_{t \geq 0}$, defined by $\Phi_t(q_0, p_0) = (q_t, p_t)$, the solution to (1.2) with initial conditions (q_0, p_0) , then we have formally:

$$\frac{\partial}{\partial t} P_t \varphi(q, p) = \partial_t \varphi(q_t, p_t) = \mathcal{L}_{\text{ham}} \varphi(q_t, p_t) = \mathcal{L}_{\text{ham}} P_t \varphi(q, p) = P_t \mathcal{L}_{\text{ham}} \varphi(q, p).$$

This motivates the alternative definition for the semi-group of evolutions operators:

$$(2.4) \quad P_t = e^{t\mathcal{L}_{\text{ham}}}$$

Remark 2. The property (2.2) is only due to the form of J , and not to the specific expression for H . Thus any H , we may consider any dynamics of the form (2.1), to devise a dynamical system whose orbits are restricted to the level set $H^{-1}\{H(q_0, p_0)\}$.

Conversely, given a differential dynamical system, if through a change of coordinates one is able to write the system under this form, one has found a conservation law.

One key property of the classical Hamiltonian is that it is *separable*: it splits into a kinetic part involving only the momentum variable and a potential part involving only the coordinate variable, a property which is especially useful for constructing explicit numerical schemes for the Hamiltonian dynamics.

2.1.2 Numerical schemes for Hamiltonian dynamics

It is impossible, except for a very restricted class of systems, which do not occur in practical settings anyhow, to analytically integrate Hamilton's equation (1.2). For this reason, one must revert to numerical schemes, which we may interpret as discrete approximations of the Hamiltonian flow. In most common applications, the aim is to approximate the exact solution of an evolution equation as precisely as possible over a given time domain. In the case of molecular dynamics, however, the time domain is usually very large, because simulating long trajectories is a requirement to ensure that a representative portion of phase space is explored. As a consequence, it is in practice impossible to obtain precise solutions over a long time, because of the evolution's sensitivity to the initial conditions. Furthermore, one does not even care about the exact evolution, since the dynamics are merely used as a sampling device. Instead, one key requirement is that the dynamics stay on or close to the initial level set of the Hamiltonian. It can be shown through eigenanalysis that even for simple linear systems, this requirement is not satisfied by standard ODE numerical methods such as the explicit and implicit Euler schemes, or the RK4 method, for which the energy may explode or implode geometrically. This has the practical effect that for reasonably sized atomic systems, numerical instabilities render the simulations nonsensical after only a few time steps, a far cry from what is needed to obtain good estimates. One must then devise dedicated numerical methods, guided by the aim to preserve qualitative properties of the Hamiltonian evolution.

It turns out that the symplecticity of the Hamiltonian flow is the key property one should aim to preserve. Many symplectic schemes may then be constructed using splitting approximations of the evolution operator over one time step associated with the Hamiltonian flow. This is motivated by the following straightforward lemma:

Lemma 1 (Composition of symplectic maps). *Let $U \subset \mathbb{R}^d$ be an open set, $g : U \mapsto \mathbb{R}^d$ and $f : g(U) \mapsto \mathbb{R}^d$ be C^1 symplectic mappings. Then the composition $f \circ g$ is symplectic.*

Let us fix a timestep $\Delta t > 0$. The evolution operator over one timestep associated with the Hamiltonian flow is the operator defined by

$$e^{\Delta t \mathcal{L}_{\text{ham}}} \varphi = \varphi \circ \Phi_{\Delta t}$$

Symplectic Euler schemes

$$(2.5) \quad \begin{cases} p^{n+1} = p^n - \nabla V(q^n) \Delta t \\ q^{n+1} = q^n + M^{-1} p^{n+1} \Delta t \end{cases}$$

$$(2.6) \quad \begin{cases} q^{n+1} = q^n + M^{-1}p^n \Delta t \\ p^{n+1} = p^n - \nabla V(q^{n+1})\Delta t \end{cases}$$

Verlet scheme

2.1.3 Energy conservation properties

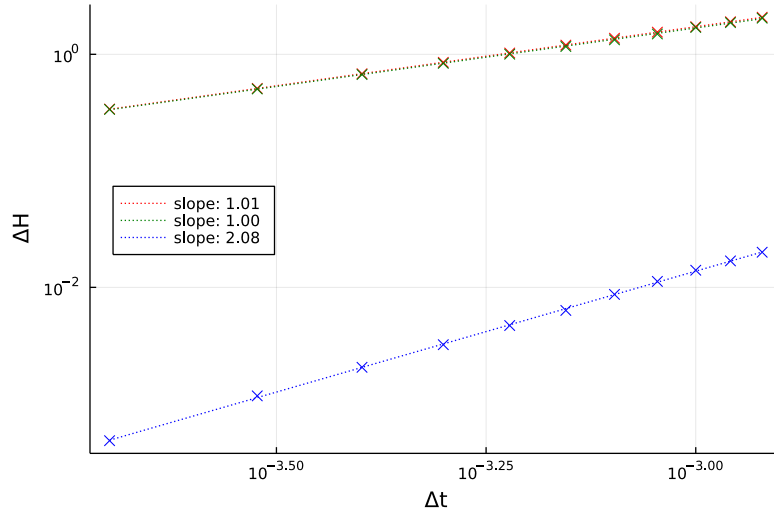


Figure 2.1: Effect of the time step on absolute variation of the Hamiltonian for the symplectic Euler (red and green) and the Verlet (blue) schemes. As expected, total variation scales as Δt for symplectic Euler, and as Δt^2 for Verlet.

2.1.4 Examples of instantaneous observables

2.1.5 Shortcomings of the Hamiltonian approach

2.2 Canonical averages

Langevin dynamics We consider a special case of the inertial Langevin dynamics, defined by the following stochastic differential equation (SDE), where γ, β are set real constants.

$$(2.7) \quad \begin{cases} dq_t = M^{-1}p_t dt, \\ dp_t = -\nabla V(q_t)dt - \gamma M^{-1}p_t dt + \sqrt{\frac{2\gamma}{\beta}}dW_t, \end{cases}$$

where $(W_t)_{t \geq 0}$ is a standard dN -dimensional Brownian motion.

This process is a combination of a Hamiltonian evolution with an additional action on the momenta which, if isolated, defines a dN -dimensional Ornstein-Uhlenbeck process. This additional term be interpreted physically as the combination of two effects: a dissipation term

$$-\gamma M^{-1}p_t dt,$$

which can be understood as the effect of a viscous friction force on the particles, and a fluctuation term,

$$\sqrt{\frac{2\gamma}{\beta}}dW_t,$$

which corresponds to the input of kinetic energy into the system as thermal agitation induced by a surrounding heat bath at temperature $1/(k_B\beta)$.

However, the physical meaning can be forgotten thanks to the fact that, *in fine*, we only require that the canonical measure be invariant under this dynamic: as we shall shortly see, this is indeed the case.

Remark 3. *There are several ways to generalize this process: one is to consider more general, possibly non-separable, Hamiltonians, as in (2), rather than the classical Hamiltonian used above. The other is to allow the fluctuation-dissipation term to be parametrized by coefficients γ and σ depending on the state variable, and which obey a relation ensuring the invariance of μ . Hence in full generality, we could consider the following Langevin dynamic:*

$$(2.8) \quad \begin{cases} dq_t = \nabla_p H(q_t, p_t) dt, \\ dp_t = -\nabla_q H(q_t, p_t) dt - \gamma(q_t, p_t) \nabla_p H(q_t, p_t) dt + \sigma(q_t, p_t) dW_t. \end{cases}$$

2.2.1 Properties of the Langevin dynamics

To investigate some of the properties of the dynamics, it is useful to introduce the notion of a generator for a process defined by a possibly inhomogeneous SDE.

The generator We consider a general process defined by a SDE of the form:

$$(2.9) \quad dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$

Where b is a \mathbb{R}^n -valued function, W is a standard d -dimensional Brownian motion and σ is a $n \times d$ matrix-valued function.

For φ a smooth bounded function, Itô's lemma allows us to compute:

$$\begin{aligned} d\varphi(t, X_t) &= \frac{\partial \varphi}{\partial t}(t, X_t)dt + \nabla^\top \varphi(t, X_t)dX_t + \frac{1}{2} \text{Tr}(\nabla^2 \varphi(t, X_t)d\langle X, X \rangle_t) \\ &= \left(\frac{\partial \varphi}{\partial t} + \nabla^\top \varphi b + \frac{1}{2} \text{Tr}(\nabla^2 \varphi \sigma \sigma^\top) \right)(t, X_t)dt + (\nabla^\top \varphi \sigma)(t, X_t)dW_t, \end{aligned}$$

where ∇, ∇^2 are with respect to the spatial coordinates. In other words,

$$(2.10) \quad \varphi(t, X_t) = \varphi(0, X_0) + \int_0^t \left(\frac{\partial \varphi}{\partial t} + \nabla^\top \varphi b + \frac{1}{2} \text{Tr}(\nabla^2 \varphi \sigma \sigma^\top) \right)(s, X_s)ds + \int_0^t \nabla^\top \varphi \sigma(s, X_s)dW_s.$$

Definition 5 (Generator of an Itô process). *Let X_t be a \mathbb{R}^n -valued process defined by 2.9. We define its generator at time t as the operator defined by*

$$(2.11) \quad \mathcal{L}_t \varphi(x) = \left(\frac{\partial \varphi}{\partial t} + \nabla^\top \varphi b + \frac{1}{2} \text{Tr}(\nabla^2 \varphi \sigma \sigma^\top) \right)(t, x)$$

In view of (2.10), we have, provided regularity conditions on σ and φ ,

$$\mathbb{E}[\varphi(t, X_t) | X_s = x] = x + \int_s^t \mathbb{E}[\mathcal{L}_u \varphi(X_u)]du$$

so that, at least formally,

$$(2.12) \quad \frac{\partial}{\partial t} \mathbb{E}[\varphi(t, X_t) | X_s = x] = \mathbb{E}[\mathcal{L}_t \varphi(X_t) | X_s = x] = \mathcal{L}_t \mathbb{E}[\varphi(t, X_t) | X_s = x]$$

If we define a family of evolution operators $(P_{s,t})_{s \leq t}$ by the formula

$$P_{s,t} \varphi(x) = \mathbb{E}[\varphi(t, X_t) | X_s = x]$$

(2.12) rewrites

$$\frac{\partial}{\partial t} P_{s,t} \varphi(x) = P_{s,t} \mathcal{L}_t \varphi(x) = \mathcal{L}_t P_{s,t} \varphi(x)$$

An important special case occurs when b , σ and φ do not depend on time. In this case the generator is a single operator \mathcal{L} , defined by

$$\mathcal{L}\varphi = \nabla^\top \varphi b + \frac{1}{2} \text{Tr}(\nabla^2 \varphi \sigma \sigma^\top)$$

The evolution operators $P_t := P_{0,t}$ ($= P_{s,s+t} \forall s$ by stationarity) form a semi-group, and act on the space of smooth functions as

$$P_t \varphi(x) = \mathbb{E}[\varphi(X_t) | X_0 = x]$$

The formal derivative is given by

$$\frac{\partial}{\partial t} P_t = P_t \mathcal{L} = \mathcal{L} P_t$$

As in (2.1.1), we may write, by analogy with the finite-dimensional setting, $e^{t\mathcal{L}} := P_t$

Invariance of the canonical measure The Langevin dynamics (2.7), when written under the form (2.9), corresponds to the case

$$b(q, p) = \begin{pmatrix} M^{-1}p \\ -\nabla V(q) - \gamma M^{-1}p \end{pmatrix}, \quad \sigma(q, p) = \sqrt{\frac{2\gamma}{\beta}} \begin{pmatrix} 0_{dN} & 0_{dN} \\ 0_{dN} & \text{I}_{dN} \end{pmatrix}$$

Hence, applying 2.11, we obtain the generator for the dynamics

$$\mathcal{L} = \mathcal{L}_q + \mathcal{L}_p$$

where

$$\begin{aligned} \mathcal{L}_q \varphi &= \nabla_q^\top \varphi M^{-1}p \\ \mathcal{L}_p \varphi &= -\nabla_p^\top \varphi (\nabla V(q) + \gamma M^{-1}p) + \frac{\gamma}{\beta} \Delta_p \varphi \end{aligned}$$

which we may rewrite, recognizing the generator for the Hamiltonian dynamics 2.3

$$\mathcal{L} = \mathcal{L}_{\text{ham}} + \gamma \mathcal{L}_O$$

where $\mathcal{L}_O \varphi = -M^{-1} \nabla_p^\top \varphi p + \frac{1}{\beta} \Delta_p \varphi$

Convergence to equilibrium

2.2.2 Numerical schemes for the Langevin dynamics

Just as in the Hamiltonian case, we can define schemes for the Langevin dynamics based on approximating the evolution operator over one timestep by splitting the full dynamics into elementary integrable dynamics, and combining their corresponding evolution operators in a sequence.

Lie-Trotter splittings

Example 4 (BAO scheme).

Strang splittings

Example 5 (BAOAB scheme).

2.2.3 Choosing a scheme

Asymptotic variance

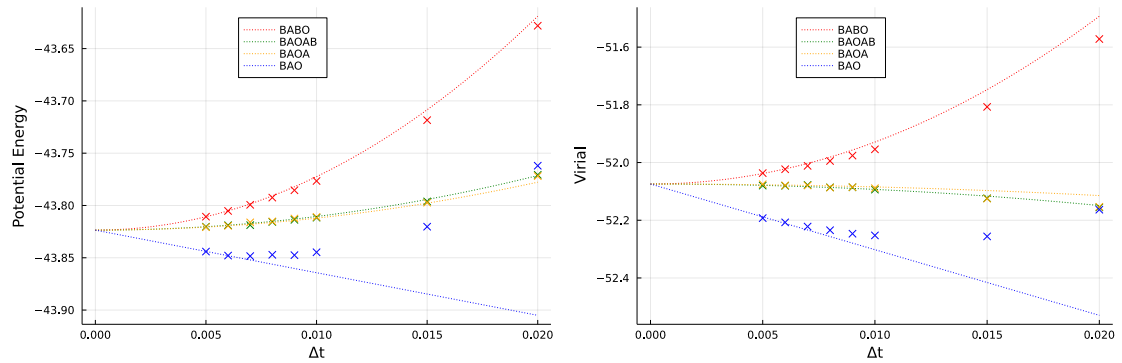


Figure 2.2: Effect of the time step on average potential energy and virial for a Lennard-Jones system of 27 particles.

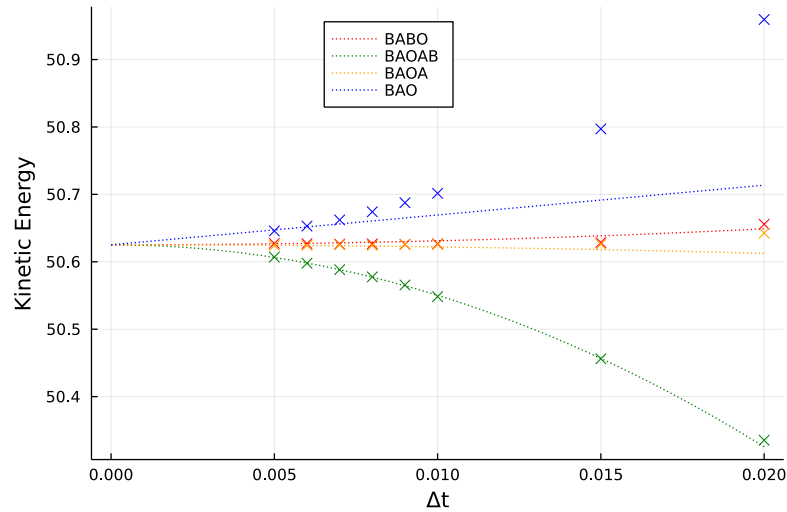


Figure 2.3: Effect of the time step on average kinetic energy for a Lennard-Jones system of 27 particles.

Bias To explain the overlap of bias between the BAOAB and the BAOA schemes observed for the potential energy and virial on Figure ??, we use the following result, which is a variation on the TU lemma.

Lemma 2. Let $P_{\Delta t}, Q_{\Delta t}$ be bounded operators on $B^\infty(\mathcal{E})$. Assume that, for any $n \geq 1$,

$$AP_{\Delta t}^n = Q_{\Delta t}^n B,$$

where A and B are bounded operators on $B^\infty(\mathcal{E})$, such that $A\mathbb{1} = \mathbb{1}$, and that the following ergodic condition holds: for any $\varphi \in B^\infty(\mathcal{E})$, and almost all $(q, p) \in \mathcal{E}$,

$$\begin{aligned} \lim_{n \rightarrow \infty} P_{\Delta t}^n \varphi(q, p) &= \int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, P}(dq, dp) \\ \lim_{n \rightarrow \infty} Q_{\Delta t}^n \varphi(q, p) &= \int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, Q}(dq, dp). \end{aligned}$$

Then we can relate $\mu_{\Delta t, P}$ and $\mu_{\Delta t, Q}$ via the following relation:

$$(2.13) \quad \int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, P}(dq, dp) = \int_{\mathcal{E}} B\varphi(q, p) \mu_{\Delta t, Q}(dq, dp)$$

Proof. Fix an initial probability measure ρ on \mathcal{E} , absolutely continuous with respect to the Lebesgue measure. Then we may write, using dominated convergence to pass to the limit:

$$\begin{aligned} & \int_{\mathcal{E}} AP_{\Delta t}^n \varphi(q, p) \rho(dq, dp) \\ &= \int_{\mathcal{E}} P_{\Delta t}^n \varphi(q, p) A^\dagger \rho(dq, dp) \\ &\xrightarrow{n \rightarrow \infty} \int_{\mathcal{E}} \left(\int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, P}(dq, dp) \right) A^\dagger \rho(d\tilde{q}, d\tilde{p}) \\ &= \int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, P}(dq, dp) \int_{\mathcal{E}} A\mathbb{1} d\rho \\ &= \int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, P}(dq, dp) \end{aligned}$$

Furthermore, applying the ergodic condition to the bounded function $B\varphi$ gives

$$\int_{\mathcal{E}} Q_{\Delta t}^n (B\varphi)(q, p) \rho(dq, dp) dq dp \xrightarrow{n \rightarrow \infty} \int_{\mathcal{E}} \left(\int_{\mathcal{E}} B\varphi(q, p) \mu_{\Delta t, Q}(dq, dp) \right) \rho(d\tilde{q}, d\tilde{p}) = \int_{\mathcal{E}} B\varphi(q, p) \mu_{\Delta t, Q}(dq, dp).$$

Since $AP_{\Delta t}^n = Q_{\Delta t}^n B$, identifying the two limits yields (2.13) \square

Corollary 1. Let $\pi_{\Delta t}$ and $\pi'_{\Delta t}$ be the invariant measures for the Markov transition operators defined respectively by the BAOA and BAOAB schemes for a fixed timestep Δt . Then the corresponding marginal distributions on \mathcal{D} are equal.

Proof. We denote by $P_{\Delta t}$ the transition operator for the BAOA scheme, and similarly $Q_{\Delta t}$ for the BAOAB scheme. It is straightforward to check that

$$e^{\frac{\Delta t}{2} B} P_{\Delta t}^n = Q_{\Delta t}^n e^{\frac{\Delta t}{2} B}.$$

Assuming the ergodic condition of Lemma 2 (TODO check doebelin condition), we can apply the result to get, for any bounded measurable observable φ ,

$$\int_{\mathcal{E}} \varphi(q, p) \pi_{\Delta t}(dq, dp) = \int_{\mathcal{E}} e^{\frac{\Delta t}{2} B} \varphi(q, p) \pi'_{\Delta t}(dq, dp)$$

Now if $\varphi(q, p) = \varphi(q, 0) := \varphi(q)$ for all p , then $e^{\frac{\Delta t}{2} B} \varphi = \varphi$, which yields the desired conclusion:

$$\forall \varphi \in B^\infty(\mathcal{D}), \quad \int_{\mathcal{E}} \varphi(q) \pi_{\Delta t}(dq, dp) = \int_{\mathcal{E}} \varphi(q) \pi'_{\Delta t}(dq, dp).$$

\square

The corollary also gives an relation between the marginal distributions in momentum space, which explains the difference displayed on (Figure 2.3) between those two schemes.

2.2.4 Unbiased sampling

The Metropolis-Hastings algorithm

2.2.5 Illustration: the equation of state of Argon

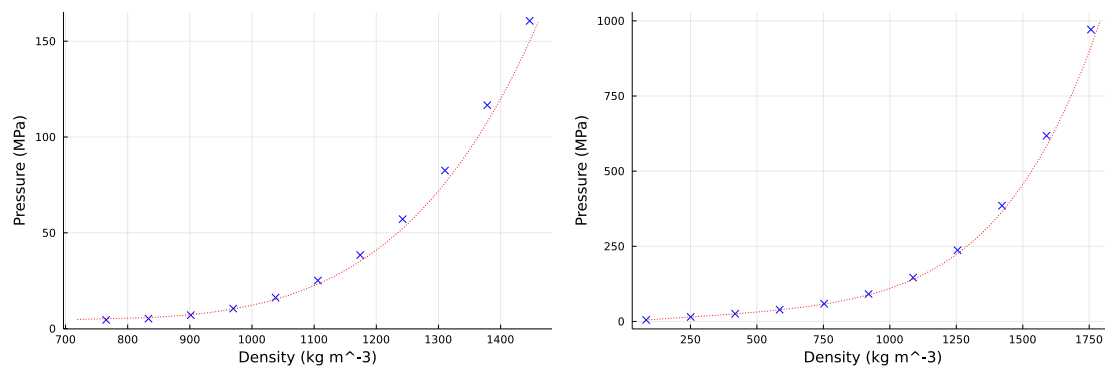


Figure 2.4: Simulated equations of state of Argon at 150 K (liquid phase, left) and 300 K (supercritical phase, right). Experimental reference curves are plotted in red, simulated data points are scattered in blue.

Chapter 3

Transport coefficients

Chapter 4

A Norton method

Appendix A

Performance analysis

Appendix B

Conventions

Notational conventions

We convene that the gradient of a function $\varphi : \mathbb{R}^n \mapsto \mathbb{R}$ is a column vector-valued function

$$\nabla\varphi : \mathbb{R}^n \mapsto \mathbb{R}^n := \mathbb{R}^{n \times 1}$$

Notationally,

$$\nabla = \begin{pmatrix} \partial x_1 \\ \vdots \\ \partial x_n \end{pmatrix}$$

So that the Hessian operator writes

$$\nabla^2 := \nabla \nabla^\top = \begin{pmatrix} \partial x_1 \partial x_1 & \cdots & \partial x_1 \partial x_n \\ \vdots & \ddots & \vdots \\ \partial x_n \partial x_1 & \cdots & \partial x_n \partial x_n \end{pmatrix}$$

And for $f = (f_1, \dots, f_n)^\top : \mathbb{R}^n \mapsto \mathbb{R}^n$

$$Jf = \begin{pmatrix} \nabla^\top f_1 \\ \vdots \\ \nabla^\top f_n \end{pmatrix} = (\nabla f^\top)^\top \qquad \operatorname{div} f = \partial_{x_1} f_1 + \cdots + \partial_{x_n} f_n = \nabla^\top f$$

are respectively the Jacobian and divergence of f .