

Sorbonne Université

2021-2022

## Stage de Master 2

Noé Bassel

(Non)-equilibrium molecular dynamics  
and a Norton method for the estimation  
of transport coefficients

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

July 27, 2022

This document is the report for an internship which took place from February 1st to July 31st 2022, at the CERMICS laboratory in the École des Ponts. This is a research unit in applied mathematics, comprising of teams working on problems arising from probability theory and optimization, as well as, more directly relevant to us, material science and molecular simulations. It was conducted under the advisement of Gabriel Stoltz, and had two explicitly stated aims.

One was to engage with Julia and its molecular simulation ecosystem by implementing various methods from equilibrium and non-equilibrium molecular dynamics inside the Molly package. Eventually, this approach resulted in a one-week stay with Molly's main author, to integrate some of the implementations developed during the course of the internship inside Molly.

A second aim, more prospective from the scientific point of view, was to advance the understanding of the Norton method, which is a novel method for the computation of transport coefficients from molecular simulations, based on a dual approach from the standard non-equilibrium method, whereby the response is fixed and the average forcing needed to induce it is measured, instead of the usual reverse situation. As of the end of this internship, we have proposed a numerical integration strategy for a class of Norton dynamics, and applied it to the cases of mobility and shear viscosity computations. Our numerical results suggest several avenues for future theoretical work, which will be continued in the PhD work of Shiva Darshan, starting from the fall of 2022.

The report is divided into five chapters and one appendix. The first chapter is dedicated to a basic introduction to some concepts in statistical mechanics which are relevant to molecular simulation. The second and third chapter are dedicated to equilibrium averages, with the second's focus on a presentation of the different numerical methods involved, and the third's on answering a question that arose when examining's Molly native Langevin integrator, which coincidentally was also under investigation at the same time by a team of theoretical chemists, Bettina Keller and Stefanie Kieninger. It is essentially compiled from the written communication we sent to them, presenting our understanding of the BAOA scheme. The fourth and fifth chapters are dedicated to the non-equilibrium setting, with the fourth centered on a discussion of the standard methods (Thévenin and Green-Kubo), and the fifth on the presentation of the Norton method. All chapters are supplemented with numerical examples, which are destined to illustrate some theoretical properties, or the viability of a given numerical method. We conclude the report by a short appendix which highlight some of the thinking that went into the choice of Molly as a molecular simulation package, as well as pointing the reader to relevant source code.

Every example on a realistic system was implemented within Molly, and thus we wish to thank Joe Greener for creating this very flexible and pleasant to work with package, as well as for inviting us to stay for a fascinating week in Cambridge. We also take advantage of this short introduction to thank Gabriel Stoltz for trusting us with this subject, for his precious advice, and more generally for introducing us to the fun and mathematically rich subject that is molecular simulation.

### Notational conventions

We convene that the gradient of a differentiable 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

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

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

$$(1) \quad \nabla f = \begin{pmatrix} \nabla^\top f_1 \\ \vdots \\ \nabla^\top f_n \end{pmatrix} = (\nabla \otimes f)^\top, \quad \operatorname{div} f = \partial_{x_1} f_1 + \cdots + \partial_{x_n} f_n = \nabla^\top f$$

are respectively the Jacobian matrix and divergence of  $f$ .

# Contents

<b>1</b>	<b>Introduction to molecular dynamics</b>	<b>6</b>
1.1	The microscopic description of atomic systems . . . . .	6
1.1.1	Classical phase space . . . . .	6
1.1.2	Dynamical description . . . . .	7
1.1.3	Reduced units . . . . .	8
1.2	Statistical ensembles . . . . .	9
1.2.1	Microcanonical ensemble . . . . .	10
1.2.2	Canonical ensemble . . . . .	11
1.2.3	From microscopic dynamics to macroscopic observables . . . . .	12
1.2.4	Examples of instantaneous observables . . . . .	12
1.2.5	Ergodic averages . . . . .	13
<b>2</b>	<b>Sampling equilibrium properties</b>	<b>14</b>
2.1	Microcanonical averages . . . . .	14
2.1.1	Elementary properties of Hamiltonian dynamics . . . . .	14
2.1.2	Numerical schemes for Hamiltonian dynamics . . . . .	16
2.2	Canonical averages . . . . .	22
2.2.1	Langevin dynamics . . . . .	22
2.2.2	Invariance of the canonical measure . . . . .	23
2.2.3	Overdamped limit of Langevin dynamics . . . . .	25
2.2.4	Convergence to equilibrium . . . . .	26
2.2.5	Asymptotic variance for ergodic averages . . . . .	27
2.2.6	Splitting schemes for the Langevin dynamics . . . . .	28
2.2.7	Error analysis for splitting schemes . . . . .	31
2.2.8	Unbiased sampling . . . . .	32
2.2.9	Numerical illustrations . . . . .	34
<b>3</b>	<b>Study of the BAOA scheme</b>	<b>38</b>
3.0.1	Definitions and notations . . . . .	38
3.0.2	Relating invariant measures of discretization schemes . . . . .	39
3.0.3	Error estimate on the phase space measure . . . . .	40
3.0.4	Error estimates on the kinetic marginal distributions . . . . .	41
3.0.5	Analysis of the second order error term for kinetic averages under $\mu_{\Delta t, P}$ . . . . .	42
3.0.6	Analysis of the discrepancy between the dominant error terms on the kinetic marginals. . . . .	44
3.1	Numerical results . . . . .	46
3.1.1	Models . . . . .	46
3.1.2	Equality of marginal configurational distributions . . . . .	47
3.1.3	Comparison of marginal kinetic distributions . . . . .	47
3.1.4	Verification of the first-order expansion . . . . .	49
3.1.5	Example of first-order bias in a BAOA average . . . . .	49
3.1.6	Effect of the friction parameter . . . . .	49

<b>4 Non-equilibrium Molecular Dynamics</b>	<b>53</b>
4.1 Non-equilibrium molecular dynamics . . . . .	53
4.1.1 General framework . . . . .	53
4.1.2 Numerical implementation . . . . .	55
4.1.3 Mobility . . . . .	56
4.1.4 Shear viscosity . . . . .	58
4.2 The Green-Kubo method . . . . .	59
4.2.1 Numerical implementation . . . . .	61
4.2.2 Mobility . . . . .	62
4.2.3 Shear viscosity . . . . .	67
<b>5 Norton dynamics</b>	<b>69</b>
5.1 Introduction . . . . .	69
5.2 General framework . . . . .	69
5.3 Norton dynamics for transport coefficients . . . . .	71
5.3.1 Splitting schemes for numerical integration . . . . .	72
5.3.2 Physical interpretation . . . . .	76
5.4 Numerical results . . . . .	78
5.4.1 Mobility . . . . .	78
5.4.2 Shear viscosity . . . . .	79
<b>A Implementation details</b>	<b>84</b>

# 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.

#### 1.1.1 Classical phase space

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. It is often the case in computer simulations that we consider positions which are restricted to a bounded domain by the use of periodic boundary conditions. To that effect, let

$$\mathcal{D} = (LT)^{dN} \text{ or } \mathbb{R}^{dN},$$

where  $\mathbb{T}$  is the one-dimensional torus. We call  $\mathcal{D}$  the configuration space.

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

$$q = (q_{1,1}, \dots, q_{1,d}, \dots, q_{N,1}, \dots, q_{N,d})^\top \in \mathcal{D},$$

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

where  $q_i := (q_{i,1}, \dots, q_{i,d})^\top$  is the position vector of the  $i$ -th particle, and similarly for  $p$ . Let

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

Central objects in our study will be trajectories

$$(q_t, p_t)_{t \geq 0} \subset \mathcal{E},$$

which describe the time evolution of a physical system.

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.

### 1.1.2 Dynamical description

In order to describe the evolution of the system's state, one must specify a dynamical law. This is done by giving a function  $V : \mathcal{E} \rightarrow \mathbb{R}$  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 our case we will always take the potential to be independent of the momentum, so that we can think of  $V$  as having domain  $\mathcal{D}$ . 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. The function  $V$  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 pair-interaction potential

$$(1.3) \quad V_{\text{LJ}}(q) = \sum_{1 \leq i < j \leq N} v(|q_i - q_j|),$$

where  $v$  is a radial function

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

The reference energy  $\varepsilon$  and length  $\sigma$  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 interparticular 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  $\sigma$  and  $\varepsilon$  must be chosen empirically to describe the behavior of a particular atomic species. For Argon, values of reference are:  $\sigma = , \varepsilon =$ .

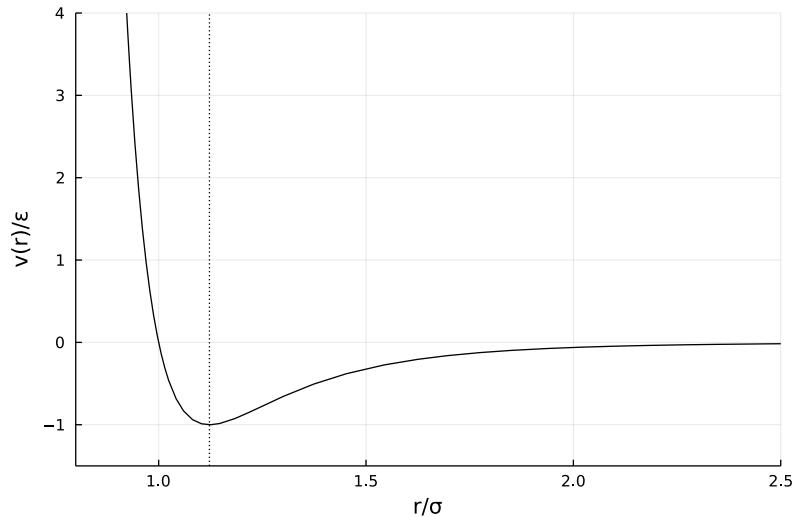


Figure 1.1: The pair potential  $v$ , with lengths and energy given in reduced units. The equilibrium interparticular distance is indicated by the vertical dotted line.

### 1.1.3 Reduced units

It is convenient, given an atomic system, to describe quantities therein within a system of units in which they are of order one. This has several advantages. Firstly, like any reasonable system of units, they make quantities easier and more intuitive to reason about. Secondly, from the computational point of view, numerical artifacts due to loss of precision at very large or very small scales and overflow errors can be avoided more often. Thirdly, they may help transfer knowledge about one system to another. For instance, in the Lennard-Jones system, expressing a

quantity in reduced units, and knowing the dependency of these units on the parameters  $(\varepsilon, \sigma, M)$  of the system, one can infer properties about a system with parameters  $(\tilde{\varepsilon}, \tilde{\sigma}, \tilde{M})$  by applying the inverse transformation, thus effectively yielding equivalence of different systems under different conditions, and sparing the cost of running redundant simulations.

We describe a general procedure to construct a set of reduced units. Our choice, although not necessarily unique, is quite natural, especially when dealing with the Lennard-Jones system. Let us fix a reference mass  $m_*$ , a reference energy  $\varepsilon_*$  and a reference length  $\sigma_*$ . Then various reference quantities can be derived by natural conversions and dimensional analysis:

$$\begin{aligned} (\text{time}) \quad t_* &= \sqrt{\sigma_*^2 m_* \varepsilon_*^{-1}}, \\ (\text{temperature}) \quad T_* &= \varepsilon_* k_B^{-1}, \\ (\text{velocity}) \quad v_* &= \sigma_* t_*^{-1}, \\ (\text{volume}) \quad V_* &= \sigma_*^3, \\ (\text{area}) \quad A_* &= \sigma_*^2, \\ (\text{density}) \quad \rho_* &= V_*^{-1}, \\ (\text{force}) \quad F_* &= m_* \sigma_* t_*^{-2}, \\ (\text{pressure}) \quad P_* &= F_* A_*^{-1}, \end{aligned}$$

and one can of course go on. The point is that if  $X$  is a quantity, one can obtain its reduced value  $X_{\text{red}}$  by dividing  $X$  by the reference quantity  $X_*$  which is dimensionally compatible with  $X$ , and which can be derived as above. For instance, given a pressure  $P$ , we obtain

$$P_{\text{red}} = P \sigma_*^3 \varepsilon_*^{-1}.$$

Note this is an adimensional quantity. Note, also that, due to the definition of the reference temperature  $T_*$  relative to the reference energy  $\varepsilon_*$ , energy are now commensurate to temperatures. The Boltzmann factor, which is the conversion factor between units of temperature and units of energy simplifies when expressed in reduced units, a fact we can capture with the maxim  $k_B* = 1$ . This simplifies many formulae.

For a monoatomic Lennard-Jones system, a natural choice for  $m_*$  is the atomic mass of the considered species. We also take  $\varepsilon_* = \varepsilon$  and  $\sigma_* = \sigma$ , (although the equilibrium length  $\sigma = 2^{1/6}\sigma$  is another possible choice). In the case of Argon, we use the following values:

$$m_* = 6.634 \times 10^{-26} \text{ kg}, \quad \sigma_* = 3.405 \times 10^{-10} \text{ m}, \quad \varepsilon_* = 1.66 \times 10^{-21} \text{ J}.$$

Unless explicitly specified, all numerical results will be in this system of reduced units.

## 1.2 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 ( $6.02 \times 10^{23}$ ). 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, loosing 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  $\rho$  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 the following two examples.

### 1.2.1 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  $H$ :

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

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  $A$ . Alternatively,  $\mu_{mc,E}$  is uniquely defined by the relation

$$\int_{\mathcal{E}} g(H(q, p)) f(q, p) dq dp = \int_{\mathbb{R}} g(E) \int_{\mathcal{E}} f(q, p) \mu_{mc,E}(dq, dp),$$

for all test functions  $g : \mathbb{R} \rightarrow \mathbb{R}$  and  $f : \mathcal{E} \rightarrow \mathbb{R}$ . It is possible, using the coarea formula, to derive an explicit formula for  $\mu_{mc,E}$ , namely,

$$(1.5) \quad \mu_{mc,E}(dq, dp) = \frac{\sigma_E(dq, dp)}{|\nabla H(q, p)|},$$

where  $\sigma_E$  is the surface measure induced by the Lebesgue measure on the constant energy manifold

$$(1.6) \quad S(E) = \{(q, p) \in \mathcal{E} \mid H(q, p) = E\}.$$

### 1.2.2 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 the canonical measure as the maximizer:

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

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

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

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

$$(1.7) \quad \mu(q, p) := \frac{1}{Z} e^{-\beta H(q, p)}.$$

Furthermore, one can show that  $\mu$  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  $\beta$  is a tuning parameter related to the value of  $\bar{E}$ . The physical interpretation of  $\beta$  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 NV $\bar{E}$ )

One could of course go further and remark 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  $\mu$ V $\bar{T}$  ensemble. This, and many other constructions are possible, but we will restrict our attention to the NVE and NVT cases.

**Remark 1** (Independence of canonical momenta and configurations). *We can make an observation on  $\mu$  using the fact that the Hamiltonian 1.1 is separable: it is the sum of a kinetic term involving  $p$  and a configurational term involving  $q$ . Thus we can write*

$$e^{-\beta H(q, p)} = e^{-\frac{\beta}{2} p^\top M^{-1} p} e^{-\beta V(q)},$$

which implies that the canonical measure  $\mu$  is of tensor form:

$$(1.8) \quad \mu = \kappa \otimes \nu,$$

where  $\kappa$  is a probability measure on  $\mathbb{R}^{dN}$  has a density proportional to  $e^{-\frac{\beta}{2} p^\top M^{-1} p}$  and  $\nu$  has a density proportional to  $e^{-\beta V(q)}$  on  $\mathcal{D}$ . Recognizing a multivariate Gaussian density, we can further write, abusing the notations  $\kappa$  and  $\nu$  to denote both the laws and their densities,

$$(1.9) \quad \kappa(p) = \det\left(\frac{M}{2\pi\beta}\right)^{\frac{1}{2}} e^{-\frac{\beta}{2} p^\top M^{-1} p}$$

$$(1.10) \quad \nu(p) = \frac{1}{Z_\nu} e^{-\beta V(q)},$$

with

$$Z_\nu = Z \det \left( \frac{M}{2\pi\beta} \right)^{\frac{1}{2}}.$$

This implies in particular that the marginal distribution in  $p$  of the canonical measure can be sampled easily, using standard algorithms for generating i.i.d. Gaussian variables, such as the Box-Muller method.

The difficult part is sampling  $\nu$ . Why we should even care to do so is the object of the following paragraph.

### 1.2.3 From microscopic dynamics to macroscopic observables

The main interest in describing the macroscopic state of a system as a probability distribution over its microscopic configurations is that one can then express constant macroscopic quantities as averages of fluctuating microscopic observables with respect to the ensemble measure. In our context, an observable is simply a function defined over phase-space. Given a macroscopic system defined by an equilibrium ensemble  $\mu$  and an observable  $\varphi$ , we are interested in computing its average value over the ensemble,

$$(1.11) \quad \mathbb{E}_\mu[\varphi] := \int_{\mathcal{E}} \varphi(q, p) \mu(dq, dp).$$

Two problems arise. One is, for a given macroscopic quantity of interest, how to correctly define a microscopic observable  $\varphi$  which underlies the macroscopic behavior. The second, once we have properly defined  $\varphi$ , is how to compute averages (??).

### 1.2.4 Examples of instantaneous observables

Let us give some examples of observables which will be of interest to us. Obvious examples include the kinetic energy,

$$E_\kappa(q, p) = \frac{1}{2} p^\top M^{-1} p,$$

the potential energy

$$E_p(q, p) = V(q),$$

and the Hamiltonian  $H$ . We will generally be considering real-valued observables, although vector-valued observables such as the velocity  $v = M^{-1}p$  may be of interest. We will give two additional examples. The kinetic temperature is defined by the following expression:

$$T_\kappa(q, p) = \frac{2}{k_B d N} E_\kappa(q, p).$$

It is, up to a conversion factor of Boltzmann's constant, twice the kinetic energy per degree of freedom. In the canonical ensemble, it is easily shown that  $\mathbb{E}_\mu[T_\kappa] = T$ , justifying the terminology. (reference also equipartition theorem)

We will also be considering the instantaneous pressure, which is given, for a pair interaction potential of the form (1.3) and a periodic domain  $\mathcal{D}$

$$P(q, p) = \frac{1}{|\mathcal{D}|} \left( N k_B T_\kappa - \frac{1}{d} \sum_{i=1}^N q_i^\top \nabla_{q_i} V(q) \right).$$

Neglecting the right-hand side gives the famous ideal gas law,  $PV = Nk_B T$ , which is a good approximation at low densities. The right hand side, otherwise known as the virial, appears to be problematic since the  $q_i$  are not periodic functions of  $q$ , which would suggest  $P$  is not a well-defined observable on  $\mathcal{D}$ . However, using the particular form of the potential, which is a function of the displacements  $|q_i - q_j|$ , and symmetry arising from Newton's third law, we can arrive at the following expression:

$$(1.12) \quad P(q, p) = \frac{1}{d|\mathcal{D}|} \left( \sum_{i=1}^N \frac{|p_i|}{m_i} - \sum_{i \neq j} |q_i - q_j| v'(|q_i - q_j|) \right),$$

which is indeed a periodic function of  $q$ . In the NVT ensemble, the average of the kinetic part of the pressure is a known parameter, so we may replace it by its value, considering the observable:

$$\frac{N}{\beta|\mathcal{D}|} - \frac{1}{d|\mathcal{D}|} \sum_{i=1}^N q_i^\top \nabla_{q_i} V(q).$$

### 1.2.5 Ergodic averages

This second problem is one of sampling a probability measure in many dimensions, which is a difficult problem in general. Our broad strategy, which will remain the same in the NVE and NVT case, is the following. We define a process  $(q_t, p_t)_{t \geq 0}$  on  $\mathcal{E}$ , either deterministic or stochastic, which is invariant for the target measure  $\mu$ , in the following sense:

$$(1.13) \quad \forall t \geq 0, \varphi \in B^\infty(\mathcal{E}), \int_{\mathcal{E}} \mathbb{E}^{(q,p)} [\varphi(q_t, p_t)] \mu(dq, dp) = \int_{\mathcal{E}} \varphi(q, p) \mu(dq, dp),$$

where the superscript in the expectation denotes that the process has value  $(q, p)$  at  $t = 0$ , and the expectation is over values of  $(q_t, p_t)$ . In other words, this is a process which, given that its initial condition is distributed according to  $\mu$ , remains distributed according to  $\mu$  at any later time. It is then natural to consider average values of  $\varphi$  over the trajectories:

$$(1.14) \quad \frac{1}{T} \int_0^T \varphi(q_t, p_t) dt,$$

which we may hope will converge to the target value. The convergence of ergodic averages to the ensemble average 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. If the underlying dynamic is stochastic, then the variance of the random variables (1.14) becomes an issue, which one must keep in check by ensuring that time averages are taken over long enough trajectories. Furthermore, since the true invariant dynamics is in general in continuous time, one must devise discrete in time approximations to the true trajectories. 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, and even in the absence of theoretical guarantees.

# Chapter 2

## Sampling equilibrium properties

### 2.1 Microcanonical averages

Our aim in this section will be to describe methods to sample microcanonical averages. We first describe a few qualitative properties of Hamiltonian dynamics. These properties will serve as litmus tests to determine viable candidate numerical schemes, which we will require to preserve, either exactly or asymptotically, these qualitative properties.

#### 2.1.1 Elementary 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

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

Applying 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 3** (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*

$$(2.2) \quad \mathcal{L}_{\text{ham}} \varphi = (\nabla_p H \cdot \nabla_q - \nabla_q H \cdot \nabla_p) \varphi = (J \nabla H)^\top \nabla \varphi$$

We can split the generator as the sum of two elementary operators,

$$\mathcal{L}_{\text{ham}} = A + B,$$

with

$$(2.3) \quad A = (M^{-1}p) \cdot \nabla_q \quad B = -\nabla V(q) \cdot \nabla_p.$$

The generator allows us to quantify the rate of change of an observable  $\varphi$  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  $\Phi$  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 unique solution to (1.2) with initial conditions  $(q_0, p_0)$ , then we have formally:

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

In the following result, we collect certain qualitative properties of Hamiltonian dynamics.

**Proposition 1** (Properties of Hamiltonian dynamics). *Assume that the Hamiltonian  $H$  is  $C^2$  on  $\mathcal{E}$  and that the flow  $\Phi_t$  is globally defined for  $t \in \mathbb{R}$ . Then the following properties hold.*

i) *Group structure:*

$$\forall t, s \in \mathbb{R}, \Phi_t \circ \Phi_s = \Phi_{t+s}, \Phi_0 = \text{Id}.$$

ii) *Energy preservation:*

$$\frac{dH(q_t, p_t)}{dt} = 0.$$

iii) *Conservation of the Lebesgue measure:*

$$\forall D \subseteq \mathcal{E}, \forall t \geq 0, |\Phi_t(D)| = |D|.$$

iv) *Symplecticity:*

$$\forall t \geq 0, \nabla \Phi_t^\top J \nabla \Phi_t = J.$$

v) *Time reversibility:*

$$\Phi_t \circ \mathcal{R} \circ \Phi_t = \mathcal{R}.$$

The notation  $\nabla$  corresponds to (1), and the map  $\mathcal{R}$  is the momentum-reversing involution

$$\mathcal{R}(q, p) = (q, -p).$$

*Hints of proof.* i) This property expresses the fact that the Hamiltonian evolution is autonomous, and follows from uniqueness in the Cauchy-Lipschitz theorem. This allows one to formally interpret the flow as a group action of  $\mathbb{R}$  on  $\mathcal{E}$ .

ii) The energy conservation property simply follows from applying  $\mathcal{L}_{\text{ham}}$  to  $H$ .

iii) This property, known as Liouville's theorem, holds generally for any divergent-free flow. Its proof is based on a time differentiation of the determinant  $\det(\nabla \Phi_t)$ , and observing that the Hamiltonian vector field is divergence free:

$$\text{div}(J \nabla H) = \text{div}_q(\nabla_p H) - \text{div}_p(\nabla_q H) = 0.$$

- iv) The property is trivially verified at time  $t = 0$ . A straightforward calculation shows that the time derivative of the symplecticity condition for  $\Phi_t$  is 0, which proves the claim. Let us remark that this property also implies property iii), since it shows  $\det(\nabla \Phi_t)^2 \det(J) = \det(J) \implies |\det(\nabla \Phi_t)| = 1$ .
- v) This again follows by uniqueness of trajectories, by observing that, fixing an initial condition  $(q_0, p_0)$ , time differentiating the trajectory  $(q_{-t}, -p_{-t}) = \mathcal{R} \circ \Phi_{-t}(q_0, p_0)$  shows it is Hamiltonian. Thus it must coincide with  $\Phi_t(q_0, -p_0) = \Phi_t \circ \mathcal{R}(q_0, p_0)$ , or as an equality of mappings,  $\mathcal{R} \circ \Phi_{-t} = \Phi_t \circ \mathcal{R}$ . Precomposing by  $\Phi_t$  on each side yields the result using property i).

□

**Remark 2.** Properties i) to iv) above are valid for any dynamics of the form (2.1), thus it is possible to consider dynamics with more general Hamiltonians, which still obey them, disregarding issues of well-posedness. Property v) requires the additional condition that the Hamiltonian be separable into a kinetic and potential part and that the kinetic part be an even function of  $p$ .

Property *ii*) in the proposition above asserts that Hamiltonian trajectories remain on the constant energy manifold  $S(H(q_0, p_0))$  defined in (1.6). Since this is the support of the microcanonical measure  $\mu_{mc,E}$ , it is natural to ask whether the Hamiltonian dynamics can be used to sample the microcanonical measure, by means of ergodic averages. A minimum requirement for this to hold is that the measure is left invariant under Hamiltonian evolution. Indeed, this is the case. We consider a Hamiltonian trajectory  $(q_t, p_t)$  such that  $H(q_0, p_0) = E$ . For  $g$  and  $f$  test functions,

$$\begin{aligned} \int_{\mathbb{R}} g(E) \int_{\mathcal{E}} f(q_t, p_t) d\mu_{mc,E}(dq, dp) dE &= \int_{\mathcal{E}} g(H(q, p)) f \circ \Phi_t(q, p) dq dp \\ &= \int_{\mathcal{E}} g(H \circ \Phi_{-t}(\tilde{q}, \tilde{p})) f(\tilde{q}, \tilde{p}) d\tilde{q} d\tilde{p} \\ &= \int_{\mathcal{E}} g(H(\tilde{q}, \tilde{p})) f(\tilde{q}, \tilde{p}) d\tilde{q} d\tilde{p} \\ &= \int_{\mathbb{R}} g(E) \int_{\mathcal{E}} f(\tilde{q}, \tilde{p}) d\mu_{mc,E}(d\tilde{q}, d\tilde{p}) dE. \end{aligned}$$

the absence of a Jacobian determinant term in the change of variables from the second to the third line follows from property *iii*), while the passage from the second to the third line is justified by the energy conservation property. This shows that the microcanonical measure is invariant under the Hamiltonian dynamics, but it is easy to construct examples where ergodic averages fail to converge to the correct value, as the following simple example shows.

**Example 2** (A non-ergodic system). *We consider the following one-dimensional system with  $\mathcal{E} = \mathbb{T} \times \mathbb{R}$ , and the potential given by*

$$V(q) = \cos\left(4\pi(q - \frac{1}{2})\right).$$

*The constant-energy manifold  $S(0)$  consists of two disjoint compact connected components (see Figure 2.1). The observable*

$$\varphi(q, p) = \mathbb{1}_{q > \frac{1}{2}} - \mathbb{1}_{q < \frac{1}{2}}$$

*is constant on each component of  $S(0)$ , and has zero average with respect to  $\mu_{mc,0}$  by symmetry, but ergodic averages do not converge, since*

$$\frac{1}{T} \int_0^T \varphi(q_t, p_t) dt = \frac{1}{T} \int_0^T \varphi(q_0, p_0) dt = \varphi(q_0, p_0) \in \{\pm 1\}$$

The issue of ergodicity is in general very hard to tackle for realistic systems, and in practice is overlooked. Besides, it may happen that for observables respectful of symmetries in the constant energy manifold converge to the correct microcanonical average even in the absence of ergodicity. This is, for example, the case for the kinetic energy observable in our Example 2.

### 2.1.2 Numerical schemes for Hamiltonian dynamics

It is not impossible, except for a restricted class of systems, which do not typically arise in practice, to analytically integrate Hamilton's equation (1.2). For this reason, one must revert to numerical schemes, which can and will interpret as approximations of the flow map over one timestep. More precisely, for a fixed timestep  $\Delta t$ , if we possess an approximation of the flow

$$\tilde{\Phi}_{\Delta t} \approx \Phi_{\Delta t},$$

we will deduce discrete approximations of the evolution by iteration:

$$(2.4) \quad (q^n, p^n) := \tilde{\Phi}_{\Delta t}^n(q_0, p_0) \approx (q_{n\Delta t}, p_{n\Delta t}),$$

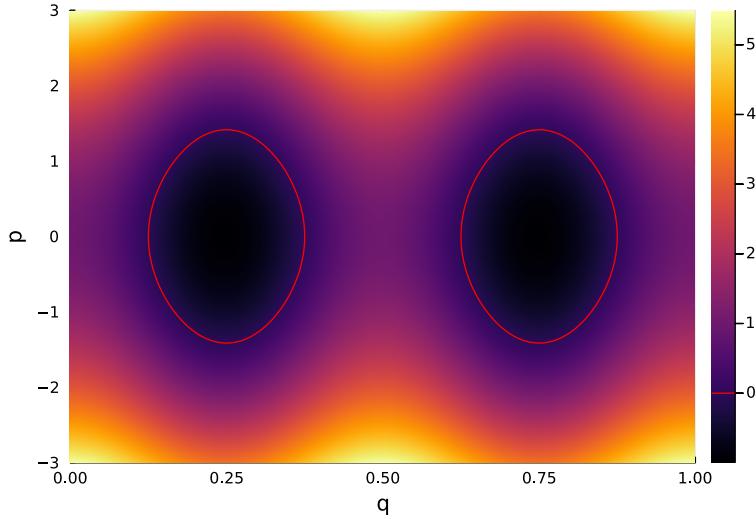


Figure 2.1: The Hamiltonian landscape for the potential  $V$  of Example 2 , with  $S(0)$  plotted in red.

which can then be used as sample points for the computation of empirical averages, discrete counterparts to the ergodic averages (1.14),

$$(2.5) \quad \frac{1}{n} \sum_{k=0}^n \varphi(q^n, p^n).$$

In most common applications involving ODEs, the aim is to approximate the exact solution as precisely as possible over a given domain. In the case of sampling trajectory averages in 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 the evolution's sensitivity to initial conditions will cause small initial errors to rapidly blow-up. Furthermore, one does not really *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 constant energy manifold associated with a given initial condition. It can be shown through eigenanalysis that 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 splitting schemes, based on operator splitting approximations of the Hamiltonian evolution operator over one timestep, preserve crucial qualitative properties of the Hamiltonian evolution.

An important observation is that if one considers each part of (2.3) as a generator in its own right, the corresponding dynamics is analytically solvable as ballistic trajectory.

**Remark 3.** Consider the two dynamics defined by

$$(2.6) \quad \begin{cases} dq_t^A = M^{-1} p_t^A dt, & dp_t^A = 0, \\ dq_t^B = 0, & dp_t^B = -\nabla V(q_t^B) dt. \end{cases}$$

These are easily solved, namely

$$(2.7) \quad \begin{cases} (q_t^A, p_t^A) = (q_0^A + tp_0^A, p_0^A), \\ (q_t^B, p_t^B) = (q_0^B, p_0^B - tV(q_0^B)). \end{cases}$$

Moreover, these evolutions are of Hamiltonian form, with Hamiltonians corresponding respectively to the kinetic part and the configurational part only, and have corresponding generators  $A$  and  $B$ . We denote by  $(\Phi_t^A)_{t \in \mathbb{R}}$  and  $(\Phi_t^B)_{t \in \mathbb{R}}$  their respective flow maps.

This observation suggests the following general recipes to construct a class of numerical schemes for the Hamiltonian dynamics, named splitting schemes. We consider approximations of the form

$$(2.8) \quad \Phi_{\Delta t} \approx \Phi_{\Delta t_k}^{G_k} \circ \cdots \circ \Phi_{\Delta t_1}^{G_1},$$

where  $G_i \in \{A, B\}$  for all  $i$  and  $\sum_{G_i=A} \Delta t_i = \sum_{G_i=B} \Delta t_i = 1$ . We will be considering three schemes, the simplest of which are the symplectic Euler schemes.

The symplectic Euler schemes are defined by the following update equations.

$$(2.9) \quad \begin{cases} p^{n+1} = p^n - \nabla V(q^n) \Delta t \\ q^{n+1} = q^n + M^{-1} p^{n+1} \Delta t \end{cases}$$

$$(2.10) \quad \begin{cases} q^{n+1} = q^n + M^{-1} p^n \Delta t \\ p^{n+1} = p^n - \nabla V(q^{n+1}) \Delta t \end{cases}$$

These correspond respectively to the splitting

$$\Phi_{\Delta t}^A \circ \Phi_{\Delta t}^B := \Phi_{\Delta t}^{BA}$$

and

$$\Phi_{\Delta t}^B \circ \Phi_{\Delta t}^A := \Phi_{\Delta t}^{AB}.$$

The velocity Verlet scheme is based on the symmetric splitting

$$\Phi_{\Delta t/2}^B \circ \Phi_{\Delta t}^A \circ \Phi_{\Delta t/2}^B := \Phi_{\Delta t}^{BAB}.$$

Its update equation is given by

$$(2.11) \quad \begin{cases} p^{n+\frac{1}{2}} = p^n - \frac{\Delta t}{2} \nabla V(q^n) \\ q^{n+1} = q^n + \Delta t M^{-1} p^{n+\frac{1}{2}} \\ p^{n+1} = p^{n+\frac{1}{2}} - \frac{\Delta t}{2} \nabla V(q^{n+1}). \end{cases}$$

We have advertised above that these numerical schemes preserve some qualitative properties of the Hamiltonian dynamics. We now turn to making this statement precise. Let us fix an evolution operator  $\tilde{\Phi}_{\Delta t}$  corresponding to a splitting of the form (2.8) with a timestep  $\Delta t > 0$ . We refer to these properties using the same indexing as in Proposition 1, but where  $\Phi_t$  is replaced by  $\tilde{\Phi}_{\Delta t}$  regardless of  $t$ . We further say a splitting is *symmetric* if the corresponding order of operators  $A$  and  $B$  is a palindrome. The velocity Verlet splitting is symmetric, while the symplectic Euler splittings are not.

**Proposition 2** (Properties of splitting schemes for Hamiltonian dynamics). *We go through the properties in Proposition 1, listing those which apply, and those which have to be modified.*

- i) *Group structure: the analogous statement is a group action of  $\mathbb{Z}$  on  $\mathcal{E}$ . This holds if the splitting is symmetric: then,  $\tilde{\Phi}_{\Delta t}^{-1} = \tilde{\Phi}_{-\Delta t}$ .*

- ii) *Energy preservation:* this does not hold as is, but does in a weakened sense that we discuss below.
- iii) *Conservation of the Lebesgue measure:* this holds for any splitting.
- iv) *Symplecticity:* similarly, this holds for all splittings.
- v) *Time-reversibility:* this holds for any symmetric splitting.

*Hints of proofs.* For every one of these properties, the strategy is the same. From Proposition 1, each one of them holds for  $\Phi_{\Delta t}^A$  and for  $\Phi_{\Delta t}^B$ . The aim is then to show that these properties are stable, at best under composition, and at worst under symmetric composition.

- i) This simply follows from writing

$$(g_1 \circ g_2 \circ \cdots \circ g_2 \circ g_1)^{-1} = g_1^{-1} \circ g_2^{-1} \circ \cdots \circ g_2^{-1} \circ g_1^{-1},$$

and using property i) applied to  $\Phi_{\Delta t}^A$  and  $\Phi_{\Delta t}^B$ .

- ii) This is a crucial and slightly subtle point, which we return to in more detail below.
- iii) This follows trivially by composition (or alternatively by symplecticity). For any measure preserving measurable maps  $f$  and  $g$ ,

$$|f \circ g(D)| = |g(D)| = |D|.$$

- iv) This follows from the fact that any composition of symplectic maps is symplectic. This, in turn, follows from the multivariate chain rule below, and applying the symplectic property twice:

$$\nabla(g \circ f) = ((\nabla g) \circ f) \nabla f \implies [((\nabla g) \circ f) \nabla f]^\top J [((\nabla g) \circ f) \nabla f] = \nabla f^\top J \nabla f = J.$$

- v) Fixing  $f \circ \mathcal{R} \circ f = g \circ \mathcal{R} \circ g = \mathcal{R}$ , we write

$$f \circ g \circ \mathcal{R} \circ g \circ f = f \circ \mathcal{R} \circ f = \mathcal{R},$$

and conclude by induction that the property holds for any symmetric splitting.

□

We have shown that splitting schemes inherit some nice geometrical properties from the underlying Hamiltonian flow, and all the more for symmetric splittings. However our final aim is to sample from the microcanonical measure, hence we should aim to sample points which remain close to the constant energy manifold  $S(E)$ . Ideally, we would want to guarantee that the Hamiltonian is perfectly preserved under the discrete evolution induced by these schemes. This, it turns out, is too high a hope. It happens, however, that, for each of these schemes, a *perturbed* Hamiltonian is (almost) exactly conserved, which we can interpret to mean that the discrete dynamics (2.4) are regularly and (almost) exactly sampled points from a *perturbed* dynamics corresponding to the new Hamiltonian. The order of this perturbation in the timestep  $\Delta t$  then allows one to quantify over finite time intervals the drift. The statements and proofs of this kind of results fall under the broad umbrella of backward numerical analysis. For a limpid and more detailed introduction, one can consult section 4 of [10]. The general idea of backward numerical analysis, given an evolution equation and an associated numerical method:

$$\dot{y} = F_0(y), \quad y^{n+1} = \tilde{\Phi}_{\Delta t}(y),$$

is to reinterpret the numerical solution  $y^1$  not as an approximation of the exact solution  $y_{\Delta t}$ , but as the exact solution  $\tilde{y}_{\Delta t}$  of an approximate evolution  $\tilde{y}$ , given by the ODE

$$\dot{\tilde{y}} = \tilde{F}(y),$$

where  $\tilde{F}$  is given by a perturbative expansion

$$\tilde{F} = F_0 + \Delta t F_1 + \Delta t^2 F_2 + \dots$$

To avoid any convergence issue related to infinite expansions, precise statements usually truncate the expansion at some finite order  $\alpha > 0$ , and require the exactitude of the numerical trajectory up to order  $\alpha + 1$ , say

$$\tilde{F} = F_0 + \Delta t F_1 + \dots + \Delta t^\alpha F_\alpha, \quad |\tilde{\Phi}_{\Delta t}(y_0) - \tilde{y}_{\Delta t}| = O(\Delta t^{\alpha+2}).$$

Comparing Taylor expansions in powers  $\Delta t$  of  $\tilde{\Phi}_{\Delta t}$  and  $\tilde{y}_{\Delta t}$  then allows us to explicitly compute the terms in the expansion of  $\tilde{F}$ . As an example we compute the first correction term for the symplectic Euler scheme  $\Phi_{\Delta t}^{AB}$ .

**Example 3** (Leading-order correction for  $\Phi_{\Delta t}^{AB}$ ). *Following our strategy, we Taylor-expand our scheme as*

$$\Phi_{\Delta t}^{AB}(q, p) = \begin{pmatrix} q + \Delta t M^{-1} p \\ p - \Delta t \nabla V(q) - \Delta t^2 \nabla^2 V(q) M^{-1} p \end{pmatrix} + O(\Delta t^3).$$

Similarly, expanding the exact solution with initial condition  $(q, p) = y_0$  to the second order yields

$$\begin{aligned} \Phi_{\Delta t}(q, p) &= y_0 + \Delta t J \nabla H(y_0) + \frac{\Delta t^2}{2} \nabla [J \nabla H(y_0)] J \nabla H(y_0) + O(\Delta t^3) \\ &= y_0 + \Delta t J \nabla H(y_0) + \frac{\Delta t^2}{2} J \nabla^2 H(y_0) J \nabla H(y_0) + O(\Delta t^3) \\ &= \begin{pmatrix} q + \Delta t M^{-1} p - \frac{\Delta t^2}{2} M^{-1} \nabla V(q) \\ p - \Delta t \nabla V(q) - \frac{\Delta t^2}{2} \nabla^2 V(q) M^{-1} p \end{pmatrix} + O(\Delta t^3). \end{aligned}$$

This shows the scheme is of order one, hence

$$|\Phi_{\Delta t}(q, p) - \Phi_{\Delta t}^{AB}(q, p)| = O(\Delta t^2).$$

Comparing the two expansions allows us to recover the discrepancy term, showing that the solution of the modified equation at time  $\Delta t$

$$\begin{cases} \frac{d}{dt} q = M^{-1} p + \frac{\Delta t}{2} M^{-1} \nabla V(q) \\ \frac{d}{dt} p = -\nabla V(q) + \frac{\Delta t}{2} \nabla^2 V(q) M^{-1} p \end{cases}$$

agrees with  $\Phi_{\Delta t}^{AB}(q, p)$  up to order 2 in  $\Delta t$ . Crucially, this equation is still of Hamiltonian form, for the modified Hamiltonian

$$(2.12) \quad \tilde{H}(q, p) = H(q, p) - \frac{\Delta t}{2} V(q)^\top M^{-1} p.$$

In fact, it is a general fact that all truncations of the modified dynamics for a symplectic method is of Hamiltonian form. For a precise statement and proof, we refer to the book by Hairer, [demander a regis pour la ref], but for now we simply make note of the fact that given a numerical method, we can construct a modified Hamiltonian equation for which this numerical order is of arbitrarily high order of local consistency. This is important, because, as the following result shows, the order of the numerical method is directly linked to the long-time energy conservation properties along numerical trajectories.

**Theorem 1.** *Let  $H$  be an analytic Hamiltonian, and  $\tilde{\Phi}_{\Delta t}$  be a symplectic numerical method of order  $\alpha$ . If there exists  $K \subset \mathcal{E}$  a compact set such that  $(q^n, p^n) \in K \forall n$ , there exists  $\tau > 0$  such that for  $\Delta t$  small enough,*

$$(2.13) \quad H(q^n, p^n) = H(q^0, p^0) + O(\Delta t^\alpha)$$

for times  $n \Delta t \leq e^{\tau/\Delta t}$ .

For a proof, see Theorem IX.8.1 in [meme ref]. The fact that the order of the scheme is linked to the local conservation in time of the Hamiltonian is unsurprising. The main content of the above result is that this conservation is valid over very long times, provided the numerical trajectory does not explode and that the timestep is chosen to be small enough.

We have already seen that the symplectic Euler methods are of order one, and it can straightforwardly be shown that the Verlet scheme is of order two by Taylor expansion. Hence we expect the fluctuation of the Hamiltonian to be of order  $\Delta t$  for the symplectic Euler methods and of order  $\Delta t^2$  for the Verlet scheme. Moreover, by construction, the symplectic Euler methods are of order two for the first-order modified Hamiltonian dynamics, so we expect the fluctuation of the first-order modified Hamiltonian computed in Example 3 to be of order two for the symplectic Euler methods. The leading-order correction term for the other Euler symplectic scheme can be computed using the same method, and is given by the opposite of (2.12).

In Figure 2.2, we verify this result numerically. A Lennard-Jones system of 1000 particles was simulated with a temperature  $T = 1.5$ , for a time  $\tau = 1.0$ , and at density  $\rho = 0.7$ . The Lennard-Jones potential was cutoff using a cubic spline interpolation between  $r_a = 2.0$  and  $r_c = 2.5$ . We plot the maximum fluctuation of the Hamiltonian

$$\Delta H := \max_{0 \leq n \leq \lceil \tau / \Delta t \rceil} H(q^n, p^n) - \min_{0 \leq n \leq \lceil \tau / \Delta t \rceil} H(q^n, p^n)$$

as a function of the timestep  $\Delta t$  for each of the symplectic splitting schemes. We also plot the maximum fluctuation of the modified Hamiltonian (2.12) for the symplectic Euler schemes, which we denote in the legend by a lowercase  $m$ . The legend gives the order of the operators in the splitting, along with the slope of the least squares regression line in log-log space, which we superimpose in dotted line. In order for the trajectories to be comparable, every simulation was started from the same precomputed equilibrium starting configuration. The results concur with theoretical prediction.

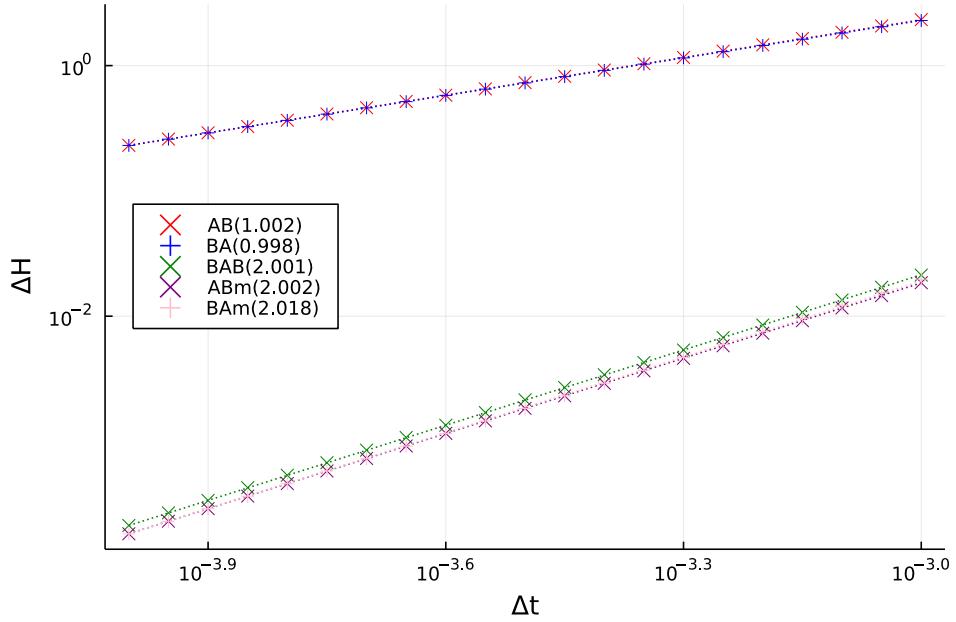


Figure 2.2: Maximum fluctuation of the Hamiltonians and leading-order modified Hamiltonians for the Symplectic Euler and Verlet schemes.

From a practical point of view, we will always use the Verlet scheme, since it offers the better energy conservation property without any computational overhead compared to the symplectic

Euler schemes. Indeed, the force calculation in the last step of (2.11) can be repurposed for the first step of the next iteration, so that there is only one force calculation per iteration, as in the symplectic Euler method. Finally let us mention that the splitting based on the ordering *ABA* of the elementary generators yields another symplectic method called the *position* Verlet scheme, which enjoys the same conservation properties as standard (velocity) Verlet. It is, however, rarely used in practice.

## 2.2 Canonical averages

We now turn our attention to methods to compute canonical averages, which are expectations of observables with respect to the canonical measure (1.7).

### 2.2.1 Langevin dynamics

We first consider the inertial Langevin dynamics, defined by the following stochastic differential equation (SDE), where  $\gamma, \beta$  are fixed positive constants.

$$(2.14) \quad \begin{cases} dq_t = M^{-1} p_t dt, \\ dp_t = -\nabla V(q_t) dt - \gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t, \end{cases}$$

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 4** (Generalized Langevin dynamics). *There are several ways to generalize this process: one is to consider a generic, possibly non-separable, Hamiltonians, as in Remark 2, rather than the classical Hamiltonian used above. The other is to allow the fluctuation-dissipation term to be parametrized by coefficients  $\gamma$  and  $\sigma$  depending on the state variable, and which obey a relation ensuring the invariance of  $\mu$ . Hence in full generality, we could consider the following system of SDEs:*

$$(2.15) \quad \begin{cases} dq_t = \nabla_p H(q_t, p_t) dt, \\ dp_t = -\nabla_q H(q_t, p_t) dt - \gamma(q_t, p_t) \nabla_p H(q_t, p_t) dt + \sigma(q_t, p_t) dW_t, \end{cases}$$

where  $\gamma$  and  $\sigma$  are  $dN \times dN$  matrix-valued functions. In fact one can also consider the case where  $W_t$  is a  $r$ -dimensional Brownian motion and  $\sigma$  is a  $dN \times r$  matrix-valued function. The Dissipative Particle Dynamics (DPD) is a generalized Langevin equation of this form, where  $\gamma$  and  $\sigma$  are position-dependent and the Brownian motion is  $dN(N-1)/2$ -dimensional, which corresponds to the number of pairs of non-orthogonal momentum degrees of freedom.

The generator of the Langevin dynamics is the operator

$$(2.16) \quad \mathcal{L}_\gamma = M^{-1} p \cdot \nabla_q - \nabla V(q) \cdot \nabla_p - \gamma M^{-1} p \cdot \nabla_p + \frac{\gamma}{\beta} \Delta_p,$$

and we denote the evolution operator using exponential notation:

$$(2.17) \quad e^{t\mathcal{L}_\gamma} \varphi(q, p) := \mathbb{E}_{(q,p)} [\varphi(q_t, p_t)],$$

where the expectation is over all trajectories of the dynamics (2.14), conditioned on  $(q_0, p_0) = (q, p)$ .

### 2.2.2 Invariance of the canonical measure

Using the generator, one can easily express the evolution of a probability distribution under the Langevin dynamics. We assume that the solution  $(q_t, p_t)_{t \geq 0}$  to (2.14) has a distribution with a smooth density  $\rho_0$  over  $\mathcal{E}$  at time  $t = 0$ , and denote  $\rho_t$  the probability density of  $(q_t, p_t)$ . For any test observable  $\varphi$ , we have

$$\int_{\mathcal{E}} \varphi(q, p) \rho_t(q, p) dq dp = \int_{\mathcal{E}} \mathbb{E}^{(q,p)} [\varphi(q_t, p_t)] \rho_0(q, p) dq dp = \int_{\mathcal{E}} e^{t\mathcal{L}_\gamma} \varphi(q, p) \rho_0(q, p) dq dp,$$

where the superscript is as in (1.13). Thus,

$$\frac{\partial}{\partial t} \int_{\mathcal{E}} \varphi(q, p) \rho_t(q, p) dq dp = \int_{\mathcal{E}} e^{t\mathcal{L}_\gamma} \mathcal{L}_\gamma \varphi(q, p) \rho_0(q, p) dq dp = \int_{\mathcal{E}} \mathcal{L}_\gamma \varphi(q, p) \rho_t(q, p) dq dp$$

If we define  $\mathcal{L}_\gamma^\dagger$  as the adjoint of  $\mathcal{L}_\gamma$  on the flat space  $L^2(\mathcal{E})$ , that is,

$$(2.18) \quad \int_{\mathcal{E}} \mathcal{L}_\gamma \varphi \psi = \int_{\mathcal{E}} \varphi \mathcal{L}_\gamma^\dagger \psi \quad \text{for all test functions } \phi, \psi,$$

we have the Fokker-Planck equation,

$$(2.19) \quad \frac{\partial}{\partial t} \int_{\mathcal{E}} \varphi(q, p) \rho_t(q, p) dq dp = \int_{\mathcal{E}} \varphi(q, p) \mathcal{L}_\gamma^\dagger \rho_t(q, p) dq dp,$$

which rewrites formally as

$$(2.20) \quad \frac{\partial}{\partial t} \rho_t = \mathcal{L}_\gamma^\dagger \rho_t.$$

Using this equation, we can easily show that the canonical distribution is invariant under this dynamics, which is equivalent to the condition

$$\mathcal{L}_\gamma^\dagger \mu = 0.$$

In fact, it is useful to reformulate this condition in the weighted space  $L^2(\mu)$ . Indeed, the stationary Fokker-Planck equation rewrites

$$\int_{\mathcal{E}} \mathcal{L}_\gamma \varphi d\mu = 0 \quad \forall \varphi,$$

or equivalently,

$$\mathcal{L}_\gamma^* \mathbb{1}_{\mathcal{E}} = 0,$$

where  $\mathcal{L}_\gamma^*$  is the adjoint of  $\mathcal{L}_\gamma$  in  $L^2(\mu)$  under the scalar product

$$\langle \varphi, \psi \rangle_\mu := \int_{\mathcal{E}} \varphi \psi d\mu.$$

This, in turn, follows easily from the following lemma.

**Lemma 1.** *The  $L^2(\mu)$  adjoints of the elementary differential operators are given by the formulae*

$$(2.21) \quad \begin{cases} \partial_{q_i}^* = -\partial_{q_i} + \beta \partial_{q_i} V, \\ \partial_{p_i}^* = -\partial_{p_i} + \beta (M^{-1}p)_i. \end{cases}$$

These are easily found by integration by parts. In particular, we find that

$$\partial_{q_i} \partial_{p_i}^* - \partial_{p_i} \partial_{q_i}^* = \beta ((M^{-1}p)_i \partial_{q_i} - \partial_{q_i} V \partial_{p_i}),$$

whence, by summing over  $i$ , we get

$$(2.22) \quad \mathcal{L}_{\text{ham}} = \frac{1}{\beta} (\nabla_q \cdot \nabla_p^* - \nabla_p \cdot \nabla_q^*),$$

which is an antisymmetric operator. Similarly,

$$\partial_{p_i} \partial_{p_i}^* = \beta (M^{-1}p)_i \partial_{p_i} - \partial_{p_i}^2,$$

hence

$$(2.23) \quad C = -\frac{1}{\beta} \nabla_p \cdot \nabla_p^*,$$

which is a symmetric operator. In summary, we have that

$$(2.24) \quad \mathcal{L}_\gamma^* = -\mathcal{L}_{\text{ham}} + \gamma C = -(A + B) + \gamma C.$$

It follows immediately that  $\mathcal{L}_\gamma^* \mathbb{1}_{\mathcal{E}} = 0$ . Notice that since  $\mathcal{L}_{\text{ham}}^* \mathbb{1}_{\mathcal{E}} = 0$ , the canonical measure is also invariant under the Hamiltonian dynamics. However, because the latter is restricted to a manifold with zero measure with respect to  $\mu$ , Hamiltonian ergodic averages cannot in general converge to the correct value.

**Remark 5** (Fluctuation-dissipation relation for generalized Langevin dynamics). *We come back to the general Langevin dynamics (2.15). In this case the generator is given by*

$$\mathcal{L}_{\gamma, \sigma} = \mathcal{L}_{\text{ham}} - (\gamma \nabla_p H) \cdot \nabla_p + \frac{1}{2} (\sigma \sigma^\top) : \nabla_p^2.$$

*The canonical measure is invariant under the action of the Hamiltonian part, so having*

$$\tilde{\mathcal{L}}_{\text{FD}} := -(\gamma \nabla_p H) \cdot \nabla_p + \frac{1}{2} (\sigma \sigma^\top) : \nabla_p^2$$

*such that  $\tilde{\mathcal{L}}_{\text{FD}}^* \mathbb{1}_{\mathcal{E}} = 0$  is enough to guarantee the invariance of the canonical measure. If  $\gamma$  and  $\sigma$  are momentum-dependent, then this condition is a complicated differential-in- $p$  relation between the coefficients of  $\gamma$  and  $\sigma$ , which can be explicitly computed by integration by parts in the expression*

$$\int_{\mathcal{E}} (\tilde{\mathcal{L}}_{\text{FD}} \varphi) \psi e^{-\beta H},$$

*with  $\varphi, \psi \in C_c^\infty(\mathcal{E})$  test functions. In the case where  $\gamma$  and  $\sigma$  are only position-dependent however, as is often the case in practice, then the expression simplifies greatly, and becomes simply an algebraic relationship between  $\gamma$  and  $\sigma$ ,*

$$(2.25) \quad \sigma \sigma^\top = \frac{2\gamma}{\beta}.$$

### 2.2.3 Overdamped limit of Langevin dynamics

As already pointed out, the fact that the kinetic marginal of  $\mu$  is a Gaussian distribution makes sampling canonical momenta trivial. Instead, the main problem is sampling from  $\nu$ . It follows directly from the invariance of  $\mu$  under trajectories of the Langevin dynamics that  $\nu$  is invariant under the configurational trajectories of the Langevin dynamics. It would be convenient, however, to have at our disposal a dynamics on  $\mathcal{D}$  which has  $\nu$  as an invariant measure. It turns out this is possible, by observing that the invariance of  $\mu$  is independent of the parameter  $\gamma$ , and taking the limit  $\gamma \rightarrow \infty$ . This requires a bit of care. Notice the SDE on the momenta in (2.14) rewrites

$$dp_t = -\nabla V(q_t)dt - \gamma dq_t + \sqrt{\frac{2\gamma}{\beta}}dW_t,$$

thus integrating gives

$$q_t - q_0 = \frac{p_0 - p_t}{\gamma} - \frac{1}{\gamma} \int_0^t \nabla V(q_s)ds + \sqrt{\frac{2}{\gamma\beta}}W_t.$$

The scaling invariance of the Brownian motion  $(\sqrt{\alpha}W_{t/\alpha^2})_{t \geq 0} \sim (W_t)_{t \geq 0}$  suggests considering the timescale  $\gamma\beta t$ , thus

$$q_{\gamma\beta t} - q_0 = \frac{p_0 - p_{\gamma\beta t}}{\gamma} - \frac{1}{\gamma} \int_0^{\gamma\beta t} \nabla V(q_s)ds + \sqrt{2}\widetilde{W}_t,$$

where  $\widetilde{W}$  is again a Brownian motion. Using the change of variables  $s = \gamma\beta u$  in the integral term yields

$$(2.26) \quad q_{\gamma\beta t} - q_0 = \frac{p_0 - p_{\gamma\beta t}}{\gamma} - \beta \int_0^t \nabla V(q_{\gamma\beta u})du + \sqrt{2}\widetilde{W}_t,$$

At this point, we formally take  $\gamma \rightarrow \infty$ , which suggests the following SDE for the rescaled in time process,

$$(2.27) \quad dq_t = -\beta \nabla V(q_t)dt + \sqrt{2}dW_t.$$

This equation defines the overdamped Langevin, or Brownian, dynamics. To justify the limit in a rigorous manner, one would hope to show that the rescaled process (2.26) converges in law to a weak solution of the SDE (2.27), in some functional space. However, this is technical overkill, since, we only need to consider dynamics as sampling devices. In fact, the physical interpretation of this equation is not entirely clear in terms of the dimensions of the quantities involved. We can just as well take equation (2.27) as given, and be satisfied by the following fact.

**Proposition 3.** *The configurational Gibbs measure  $\nu$  is invariant under the dynamics (2.27).*

This follows along the same lines as for the Langevin dynamics. The generator (now acting on observables defined on  $\mathcal{D}$ ) is the operator

$$(2.28) \quad \mathcal{L}\varphi = -\beta \nabla V \cdot \nabla \varphi + \Delta \varphi.$$

Again, we consider the weighted space  $L^2(\nu)$ . Adjoints of elementary differential operators are still given by the first line of (2.21), and it is then easily seen that

$$(2.29) \quad \mathcal{L} = -\nabla^* \cdot \nabla$$

is a symmetric operator. Again we have  $\mathcal{L}^* \mathbb{1}_{\mathcal{D}} = 0$ , so  $\nu$  satisfies the stationary Fokker-Planck equation under this dynamics.

**Remark 6.** Instead of rescaling time by  $\beta\gamma$ , we could have rescaled by  $\gamma$ , which would yield the dynamics

$$(2.30) \quad dq_t = -\nabla V(q_t)dt + \sqrt{\frac{2}{\beta}}dW_t.$$

Which formulation to choose is a matter of preference, since both yield a dynamics invariant under  $\nu$ , as seen from the identity (where we still write  $\mathcal{L}$  for the generator)

$$\mathcal{L} = -\frac{1}{\beta}\nabla^* \cdot \nabla.$$

We end this chapter by stating some key properties of the continuous dynamics, before moving to the description of splitting schemes.

## 2.2.4 Convergence to equilibrium

Once we have shown that the Langevin dynamics and its overdamped limit are reasonable candidates to sample from the canonical measure and its configurational marginal, we should ensure that the target measure is reached, and not simply invariant for the dynamics, as the cautionary example of the Hamiltonian dynamics shows. This is the question of ergodicity, and results of this nature chiefly come in two flavors.

- (i) Probabilistic ergodic theorems that ask about the convergence of random variables like (1.14), that give analogs of the Law of Large Numbers for the correlated process  $\varphi(q_t, p_t)$ .
- (ii) More analytic results which express the convergence of the law of the process at time  $t$  towards a stationary solution to the Fokker-Planck equation. These can usually be expressed as decay estimates on the evolution Semigroup (2.17), in a judiciously chosen functional setting.

We will not get into details as these questions can quickly get rather technical, but rather give vague ideas of the setting, and point the reader to more thorough sources. Results such as (i) typically leverage the strong Law of Large Numbers by dissecting the trajectory into a discrete number of (loosely) *i.i.d.* excursions through phase space, guaranteeing the almost sure convergence of trajectory averages. As such, proving the positive recurrence of the dynamics is crucial to showing that these excursions are well-behaved, while also implying that the invariant measure is unique. One idea, exploited by Klieman in [14], is to recast the Langevin dynamics as a control equation (where  $W_t$  acts as the control). He is thus able to leverage criteria on  $\mathcal{L}_\gamma$  from geometric control theory to ensure the almost sure convergence of trajectory averages. Results such as (ii) depend on the functional setting. Let us simply state the following result, based on hypocoercive estimates. For an introduction to these ideas, and references, we point to Section 2.1.1 in [16].

**Proposition 4** (Exponential decay rate on the Semigroup). *Let  $H^1(\mu)$  be the weighted Sobolev space*

$$(2.31) \quad H^1(\mu) = \{f \in L^2(\mu) \mid \nabla f \in L^2(\mu)^{2dN}\},$$

with the norm

$$\|f\|_{H^1(\mu)}^2 = \|f\|_{L^2(\mu)}^2 + \|\nabla_p f\|_{L^2(\mu)}^2 + \|\nabla_q f\|_{L^2(\mu)}^2.$$

We define  $\mathcal{H}^1(\mu)$  as the subspace

$$\mathcal{H}^1(\mu) = \left\{ f \in H^1(\mu) \mid \int_{\mathcal{E}} f d\mu = 0 \right\}$$

of centered observables, endowed with the same norm. Then there exist constants  $C_\gamma, \lambda_\gamma > 0$  such that

$$(2.32) \quad \|e^{t\mathcal{L}_\gamma}\|_{\mathcal{B}(\mathcal{H}(\mu))} \leq C_\gamma e^{-t\lambda_\gamma}.$$

For the overdamped case, exponential decay rates can be obtained more directly. In both cases, conditions have to be met by  $V$ , such as a Poincaré inequality for  $\nu$ , which are however trivially verified in the case where  $\mathcal{D}$  is compact.

### 2.2.5 Asymptotic variance for ergodic averages

Since ergodic averages are computed over a finite time-interval, the corresponding random variable will have some variance. Let us show how to relate this variance to the dynamics, provided a Central Limit Theorem holds. For simplicity, we assume that  $(q_0, p_0) \sim \mu$ , and let  $\varphi \in H^1(\mu)$  be an observable of interest. We also denote by

$$(2.33) \quad \Pi\varphi = \varphi - \int_{\mathcal{E}} \varphi d\mu$$

the centering projector associated with  $\mu$ . The Central Limit Theorem asserts that the following convergence in law holds,

$$(2.34) \quad \frac{1}{\sqrt{T}} \int_0^T \Pi\varphi(q_t, p_t) dt \xrightarrow{\text{law}} \mathcal{N}(0, \sigma_\varphi^2),$$

where  $\sigma_\varphi^2$  is the asymptotic variance associated to  $\varphi$  under the dynamics, which is thus given by the limit of the variance on the left-hand side. Let us compute:

$$\sigma_T := \mathbb{E}_\mu \left[ \left( \frac{1}{\sqrt{T}} \int_0^T \Pi\varphi(q_t, p_t) dt \right)^2 \right] = \frac{1}{T} \int_0^T \int_0^T \mathbb{E}_\mu [\Pi\varphi(q_t, p_t) \Pi\varphi(q_s, p_s)] ds dt.$$

By stationarity, for  $t > s$ ,  $\Pi\varphi(q_t, p_t) \Pi\varphi(q_s, p_s) \sim \Pi\varphi(q_{t-s}, p_{t-s}) \Pi\varphi(q_0, p_0)$ , hence we may write, by Fubini,

$$\begin{aligned} \sigma_T &= \frac{2}{T} \int_0^T \int_0^t \mathbb{E}_\mu [\Pi\varphi(q_{t-s}, p_{t-s}) \Pi\varphi(q_0, p_0)] ds dt \\ &= \frac{2}{T} \int_0^T \int_s^T \mathbb{E}_\mu [\Pi\varphi(q_s, p_s) \Pi\varphi(q_0, p_0)] dt ds \\ &= 2 \int_0^T \mathbb{E}_\mu [\Pi\varphi(q_s, p_s) \Pi\varphi(q_0, p_0)] \left(1 - \frac{s}{T}\right) ds \\ &= 2 \int_0^T \mathbb{E}_\mu [(e^{s\mathcal{L}_\gamma} \Pi\varphi)(\Pi\varphi)] \left(1 - \frac{s}{T}\right) ds \end{aligned}$$

Using Cauchy-Schwartz in  $L^2(\mu)$  and an exponential decay estimate on the evolution semigroup like (2.32), we obtain a bound of the form

$$\int_0^T \left| \mathbb{E}_\mu [(e^{s\mathcal{L}_\gamma} \Pi\varphi)(\Pi\varphi)] \frac{s}{T} \right| ds \leq \int_0^\infty C_\gamma e^{-\lambda_\gamma s} \|\Pi\varphi\|_{L^2(\mu)}^2 \frac{s}{T} ds$$

for some positive constants  $C$  and  $\alpha$ , which converges uniformly to 0 as  $T \rightarrow \infty$ . It follows that

$$(2.35) \quad \sigma_\varphi^2 = \int_0^\infty \mathbb{E}_\mu [(e^{s\mathcal{L}_\gamma} \Pi\varphi)(\Pi\varphi)] ds.$$

We use the following equality of bounded operators on  $H^1(\mu)$ , which is the continuous analog of a Neumann series:

$$(2.36) \quad (-\mathcal{L}_\gamma)^{-1} = \int_0^\infty e^{s\mathcal{L}_\gamma} ds,$$

which again is justified by the exponential decay rate of evolution semigroup, and where the integral on the right is in the Bochner sense. Using this identity, we can write the asymptotic variance more concisely,

$$(2.37) \quad \sigma_\varphi^2 = \mathbb{E}_\mu [(\Pi\varphi)(-\mathcal{L}_\gamma)^{-1}(\Pi\varphi)].$$

Note that the exact same computations can be performed for the overdamped Langevin dynamics. It follows from [2], (Theorem 2.1) that a sufficient condition for a CLT to hold is that  $\Pi\varphi \in \text{Ran } \mathcal{L}_\gamma$ , which is the case provided we can express the inverse of  $\mathcal{L}_\gamma$  using (2.36).

### 2.2.6 Splitting schemes for the Langevin dynamics

Similar to the Hamiltonian case, the generator (2.16) splits into three elementary generators, namely

$$\mathcal{L}_\gamma = A + B + \gamma C = \mathcal{L}_{\text{ham}} + \gamma C,$$

with

$$(2.38) \quad C = -M^{-1}p \cdot \nabla_p + \frac{1}{\beta}\Delta_p.$$

These generators individually give rise to dynamics which we can express explicitly, defined by the following evolution operators:

$$(2.39) \quad \begin{cases} e^{tA}\varphi(q, p) = \varphi(q + tM^{-1}p, p), \\ e^{tB}\varphi(q, p) = \varphi(q, p - t\nabla V(q)), \\ e^{t\gamma C}\varphi(q, p) = \mathbb{E} \left[ \varphi \left( q, e^{-\gamma M^{-1}t}p + \sqrt{\frac{M}{\beta}(1 - e^{-2\gamma M^{-1}t})}G \right) \right], \end{cases}$$

where  $G$  is a standard  $dN$ -dimensional Gaussian. The third equality translates an equality in law between an Itô integral and a Gaussian random variable, and follows by applying Itô's formula to the rescaled process

$$e^{\gamma M^{-1}t}p_t,$$

where  $p_t$  is the Ornstein-Uhlenbeck process:

$$(2.40) \quad dp_t = -\gamma M^{-1}p_t dt + \sqrt{\frac{2\gamma}{\beta}}dW_t,$$

then applying the following matrix form of Itô's isometric property

$$(2.41) \quad \int_0^t A_s dW_s \stackrel{\text{law}}{=} \left( \int_0^t A_s A_s^\top ds \right)^{\frac{1}{2}} \mathcal{G}$$

to obtain the equality in law. The dynamics associated with the  $A$  and  $B$  part are deterministic Hamiltonian dynamics already identified in (2.7). Just as in the Hamiltonian case, we can define schemes for the Langevin dynamics based on approximating the evolution operator (2.17) over one timestep by splitting the generator  $\mathcal{L}_\gamma$ , and combining the corresponding evolution operators (2.39) in a sequence. We refer to such a splitting approximation by the sequence in which the individual propagators are composed. It is useful at this point to introduce the stochastic flow map associated with the Ornstein-Uhlenbeck dynamics.

$$(2.42) \quad \Phi_t^C(q, p, \xi) = \left( q, e^{-\gamma M^{-1}t}p + \sqrt{\frac{M}{\beta}(1 - e^{-2\gamma M^{-1}t})}\xi \right),$$

where  $\xi \in \mathbb{R}^{dN}$ , the point being  $\mathbb{E}[\varphi(\Phi_t^C(q, p, G))] = e^{t\gamma C}\varphi(q, p)$  when  $G$  is standard Gaussian. Given an ordering of operators,

$$(2.43) \quad (R_1, \dots, R_k) \in \{A, B, \gamma C\}^k,$$

we can consider the mapping obtained by composing the elementary flows, noted as

$$(2.44) \quad \Phi^{R_1, \dots, R_k} := \Phi_{\Delta t/n_{R_k}}^{R_k} \circ \dots \circ \Phi_{\Delta t/n_{R_1}}^{R_1},$$

where

$$n_R := \# \{1 \leq j \leq n | R_j = R\}$$

for  $R \in \{A, B, \gamma C\}$ , and which we may consider by a slight abuse to be the mapping which takes a point in phase space and  $n_{\gamma C}$  vectors in  $\mathbb{R}^{dN}$  ( $\xi_1, \dots, \xi_{n_{\gamma C}}$ ), yielding a point in phase space by successively applying the flow, and, if need be, stochastic flow, mappings corresponding to the reverse ordering of (2.43). Applying this mapping with a vector of independent standard Gaussians yields a stochastic mapping, which defines the update rule for the splitting scheme associated with the ordering (2.43). An important property which follows from writing the update rule using the mapping (2.44) is that numerical trajectories formed by iterating this update rule with independent vectors of standard Gaussians form a Markov chain. The hope is that the invariant measure corresponding to this Markov chain (provided it is unique) is a close approximation to the canonical measure, as well as being ergodic. It is less cumbersome to write such schemes at the level of the evolution operator associated with the Markov chain,

$$(2.45) \quad P_{\Delta t} \varphi(q, p) = \mathbb{E} [\varphi(q^1, p^1) | q^0 = q, p^0 = p].$$

The evolution operator associated with the splitting (2.43) is

$$(2.46) \quad P_{\Delta t} = e^{\frac{\Delta t}{n_R} R_1} \cdots e^{\frac{\Delta t}{n_R} R_k}.$$

Note that the order is reversed compared to the stochastic flow formulation. We will refer to such schemes by the name obtained by concatenating the names of each operator appearing in the ordering, using  $O$  instead of  $\gamma C$ . For instance, the BAO scheme is given by the following update rule, which corresponds to applying the symplectic Euler scheme over one timestep, followed by one timestep of the Ornstein-Uhlenbeck stochastic flow (2.42).

**Example 4** (BAO scheme). *The update rule is given by the following equations, where we introduce the intermediate momentum variable  $p^{n+\frac{1}{2}}$ :*

$$(2.47) \quad \begin{cases} p^{n+\frac{1}{2}} = p^n - \Delta t \nabla V(q^n) \\ q^{n+1} = q^n + \Delta t M^{-1} p^{n+\frac{1}{2}} \\ p^{n+1} = \alpha_{\Delta t} p^{n+\frac{1}{2}} + \sigma_{\Delta t} G^n, \end{cases}$$

where  $G^n$  is a standard  $dN$ -dimensional Gaussian, and  $\alpha_{\Delta t}$ ,  $\sigma_{\Delta t}$  are given by

$$(2.48) \quad \alpha_{\Delta t} = e^{-\gamma M^{-1} \Delta t}, \quad \sigma_{\Delta t} = \sqrt{\frac{M}{\beta} (1 - \alpha_{\Delta t}^2)}.$$

Similarly, we define the BAOAB scheme.

**Example 5** (BAOAB scheme). *The update rule is given by the following equations, with additional intermediate coordinate and momentum variables:*

$$(2.49) \quad \begin{cases} p^{n+\frac{1}{3}} = p^n - \frac{\Delta t}{2} \nabla V(q^n) \\ q^{n+\frac{1}{2}} = q^n + \frac{\Delta t}{2} M^{-1} p^{n+\frac{1}{3}} \\ p^{n+\frac{2}{3}} = \alpha_{\Delta t} p^{n+\frac{1}{3}} + \sigma_{\Delta t} G^n \\ q^{n+1} = q^{n+\frac{1}{2}} + \frac{\Delta t}{2} M^{-1} p^{n+\frac{2}{3}} \\ p^{n+1} = p^{n+\frac{2}{3}} - \frac{\Delta t}{2} \nabla V(q^{n+1}), \end{cases}$$

where, again,  $G^n$  is a standard  $dN$ -dimensional Gaussian.

Now we have a recipe to make an infinite number of numerical schemes, which can easily be implemented in a computer. We could even go further and consider methods with an uneven distribution for the secondary timesteps, the introduction of negative secondary timesteps for the  $A$  and  $B$  steps, and so on. This room for creativity highlights the need for criteria to assess the quality of such schemes. Several considerations have to be weighed.

- (i) Our aim is to compute long trajectories, which are needed to ensure that phase space is properly explored, as well as to obtain better statistical properties for averages (2.5). Thus, for a fixed computational budget, we desire a scheme which allows us to take as large a timestep  $\Delta t$  as possible. This is the issue of numerical stability.
- (ii) The use of a positive timestep  $\Delta t$  implies in general that the invariant measure for the Markov chain corresponding to a given scheme is not the canonical measure. This issue is called systematic error, or bias, and one would desire a scheme which minimizes this bias.
- (iii) The main computational cost in computing iterates of these numerical schemes is the evaluation of the gradient of the potential used for the  $B$  steps. As such, it is desirable to have a scheme which requires as few evaluations of this gradient per iteration. Some care must be taken when implementing these, to ensure that already computed gradients are not re-computed: for instance, the gradient in the last step of the BOAB scheme, is equal to the one in the first step of the next iteration.
- (iv) Notice that the parameter  $\gamma$  is free for the practitioner to choose. A natural question is to determine the properties of the marginal dynamics in  $q$  in the limit  $\gamma \rightarrow +\infty$ , and in particular if we obtain a consistent discretization of the overdamped Langevin dynamics. Conversely, one could ask about properties of the dynamics as we take the Hamiltonian limit  $\gamma \rightarrow 0$ .

**Remark 7** (Overdamped and Hamiltonian limits in splitting schemes). *Concern (iv) is simple to address. Since*

$$\lim_{\gamma \rightarrow 0} \alpha_{\Delta t} = 1 \quad \lim_{\gamma \rightarrow +\infty} \alpha_{\Delta t} = 0,$$

*every O-step can simply be dropped as  $\gamma \rightarrow 0$  from the sequence of operators defining the splitting. This yields a symplectic scheme for the Hamiltonian parts of the dynamics, whose properties can be analyzed as before. For this reason, the ordering of the Hamiltonian parts of the splitting of most commonly used schemes correspond to that of a velocity Verlet method. As  $\gamma \rightarrow \infty$ , every O-step reduces to resampling the momentum according to the Maxwell-Boltzmann distribution. The properties of the resulting scheme depend on the particular ordering of the splitting at hand. For instance, if every A step is preceded by an O step, then the potential is entirely ignored by the evolution, and the trajectories form an isotropic random walk. This is for instance the case of the BOA scheme. On the other hand, it may happen that one obtains a discretization of the overdamped Langevin dynamics. For example, the update equation for the overdamped limit of the BAOA scheme in the case  $M = \text{Id}$  rewrites*

$$q^{n+1} = q^n - \frac{\Delta t^2}{2} \nabla V(q^n) + \frac{1}{2} \sqrt{\frac{\Delta t^2}{\beta}} (G^n + G^{n+1}),$$

*with  $(G^n)$  an i.i.d. standard Gaussian sequence. Because of the two-step correlations in the Gaussian increments, the numerical trajectories are not Markovian,. Nevertheless the scheme is reminiscent of a discretization of the overdamped equation (2.30), but with an effective timestep  $\frac{\Delta t^2}{2}$ . This quadratic rescaling of the timestep is common, and has to be related to the fact that the overdamped equation is defined through a diffusive rescaling. This discretization also corresponds to the overdamped limit of the BAOAB scheme, which is unsurprising in view of the results of the next chapter.*

### 2.2.7 Error analysis for splitting schemes

We turn our attention to analyzing the error arising from the estimation of canonical expectations by discrete ergodic averages. We consider estimations of the form

$$\widehat{\varphi}_{N_{\text{iter}}} := \frac{1}{N_{\text{iter}}} \sum_{i=0}^{N_{\text{iter}}-1} \varphi(q^i, p^i),$$

where  $(q^n, p^n)_{n \geq 0}$  is a numerical trajectory of a Markov chain which is ergodic with respect to a unique stationary distribution  $\pi_{\Delta t}$  on  $\mathcal{E}$ , which approximates  $\mu$ . For simplicity, we will assume that  $(q^0, p^0)$  is distributed according to  $\pi_{\Delta t}$ , so that the whole numerical trajectory is stationary. This requires in practice that the system be equilibrated before starting to sample the observables of interest. We decompose the error as follows:

$$\widehat{\varphi}_{N_{\text{iter}}} - \int_{\mathcal{E}} \varphi \, d\mu = \widehat{\varphi}_{N_{\text{iter}}} - \int_{\mathcal{E}} \varphi \, d\pi_{\Delta t} + \int_{\mathcal{E}} \varphi \, d\pi_{\Delta t} - \int_{\mathcal{E}} \varphi \, d\mu.$$

Two terms contribute to the error.

#### Statistical Error

The first term

$$\widehat{\varphi}_{N_{\text{iter}}} - \int_{\mathcal{E}} \varphi \, d\pi_{\Delta t}$$

is the statistical error due to the truncation of computed trajectories to a finite time  $T_{\text{sim}} := N_{\text{iter}} \Delta t$ . By the Central Limit Theorem for Markov Chains, asymptotically, the statistical error is of order

$$\frac{\sigma_{\varphi, \Delta t}}{\sqrt{N_{\text{iter}}}},$$

where  $\sigma_{\varphi, \Delta t}$  is the asymptotic variance of the Markov chain, which is given by the following expression, under suitable assumptions, by

$$\sigma_{\varphi, \Delta t}^2 = \text{Var}_{\pi_{\Delta t}}(\varphi) + 2 \sum_{k=1}^{\infty} \text{Cov}_{\pi_{\Delta t}}(\varphi(q^0, p^0) \varphi(q^k, p^k)).$$

The proof of the above expression is in spirit the same as the one in the continuous case, exploiting the stationarity in law of the trajectory. Using the evolution and projection operators associated with the Markov chain, which are respectively defined by

$$(2.50) \quad P_{\Delta t} \varphi(q, p) = \mathbb{E}[\varphi(q^1, p^1) | (q^0, p^0) = (q, p)], \quad \Pi_{\Delta t} \varphi = \varphi - \int_{\mathcal{E}} \varphi \, d\pi_{\Delta t},$$

we can rewrite the asymptotic variance in a more analytic form, namely

$$(2.51) \quad \sigma_{\varphi, \Delta t}^2 = 2 \int_{\mathcal{E}} (\Pi_{\Delta t} \varphi)(\text{Id} - P_{\Delta t})^{-1} (\Pi_{\Delta t} \varphi) \, d\pi_{\Delta t} - \int_{\mathcal{E}} (\Pi_{\Delta t} \varphi)^2 \, d\pi_{\Delta t},$$

where we used formally the Neumann series

$$\sum_{k=0}^{\infty} P_{\Delta t}^k = (\text{Id} - P_{\Delta t})^{-1},$$

which has to be justified rigorously, usually using a geometric decay estimate for  $P_{\Delta t}^n$ , seen as an operator on  $L_0^2(\pi_{\Delta t}) = \Pi_{\Delta t} L^2(\pi_{\Delta t})$ . This implies that

$$\Delta t \sigma_{\varphi, \Delta t}^2 = 2 \int_{\mathcal{E}} (\Pi_{\Delta t} \varphi) \left( \frac{\text{Id} - P_{\Delta t}}{\Delta t} \right)^{-1} (\Pi_{\Delta t} \varphi) \, d\pi_{\Delta t} + O(\Delta t).$$

For reasonable discretizations of the Langevin dynamics, we expect that

$$\frac{\text{Id} - P_{\Delta t}}{\Delta t} = -\mathcal{L}_\gamma + O(\Delta t),$$

which suggests that at dominant order in  $\Delta t$  approaching 0,  $\Delta t \sigma_{\varphi, \Delta t}^2$  behaves like

$$2 \int_{\mathcal{E}} (\Pi \varphi) (-\mathcal{L}_\gamma)^{-1} (\Pi \varphi) d\mu = \sigma_\varphi^2,$$

where we recall the asymptotic variance for the continuous dynamics (2.37). Hence the statistical error can be estimated, for  $\Delta t$  small enough, by

$$\frac{\sigma_{\varphi, \Delta t}}{\sqrt{N_{\text{iter}}} } = \frac{\sigma_{\varphi, \Delta t} \sqrt{\Delta t}}{\sqrt{T_{\text{sim}}} } \approx \frac{\sigma_\varphi}{\sqrt{T_{\text{sim}}}}.$$

Loosely speaking, the statistical error for the discrete ergodic estimator is governed at dominant order by the corresponding asymptotic variance for the underlying continuous dynamics, as well as the *physical* time of the simulation. This confirms that to minimize statistical error, and given a fixed budget of simulation steps, one should maximize  $T_{\text{sim}}$  and thus  $\Delta t$ , as discussed in consideration (i). However, increasing the timestep comes at the following cost.

### Systematic Error

The second term,

$$\int_{\mathcal{E}} \varphi d\pi_{\Delta t} - \int_{\mathcal{E}} \varphi d\mu,$$

is independent of the simulation time, and expresses the fact, highlighted in consideration (ii), that the invariant measure  $\pi_{\Delta t}$  for the discrete evolution will in general be offset from  $\mu$ , in the sense that the average of an observable  $\varphi$  under  $\pi_{\Delta t}$  can be expressed by an expansion of the form

$$(2.52) \quad \int_{\mathcal{E}} \varphi(q, p) \pi_{\Delta t}(dq, dp) = \int_{\mathcal{E}} \varphi(q, p) \mu(dq, dp) + \Delta t^\alpha \int_{\mathcal{E}} \varphi(q, p) f_\alpha(q, p) \mu(dq, dp) + O(\Delta t^{\alpha+1}),$$

where  $\alpha > 0$  and  $f_\alpha$  is the dominant correction term, which can be explicitly computed for splitting schemes. We postpone further discussion of these types of weak error estimates to the next chapter, in which we analyze the systematic error in the BAOA scheme.

### 2.2.8 Unbiased sampling

It turns out that one can devise schemes which have no systematic error: the Markov chain generating the trajectories has invariant measure exactly  $\mu$ . These methods are based on the Metropolis-Hastings algorithm, which gives a general method to sample a given target distribution.

#### The Metropolis-Hastings algorithm

We aim to sample from a given target measure on  $\mathbb{R}^d$ . We suppose we have at our disposal a way to generate proposal points from a given point  $\in \mathbb{R}^d$ . This amounts to defining a transition kernel, the *proposal*, which we may take to be a map

$$T : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}_+,$$

such that for any  $x \in \mathbb{R}^d$ ,  $T(x, \cdot)$  is a probability density on  $\mathbb{R}^d$ , and which is cheap to sample from (very often these are taken to be some form of Gaussian distribution). We also assume that we always have  $T(x, y) > 0$ . (This is always the case if the kernel is Gaussian). We also fix a function  $r : \mathbb{R}_+ \rightarrow (0, 1]$ , the *rule*, which satisfies the property

$$(2.53) \quad x \cdot r \left( \frac{1}{x} \right) = r(x)$$

We then define a Markov chain by iterating the following algorithm, starting from an arbitrary point  $q^0 \in \mathbb{R}^d$ .

**Algorithm 1** (Metropolis-Hastings). *From a given point  $q^n$*

(1) *Sample a proposal  $\tilde{q}^{n+1}$  according to the probability law  $T(q^n, \cdot)$ .*

(2) *Compute*

$$R(\tilde{q}^{n+1}, q^n) = r \left( \frac{\pi(\tilde{q}^{n+1})T(\tilde{q}^{n+1}, q^n)}{\pi(q^n)T(q^n, \tilde{q}^{n+1})} \right).$$

(3) *With probability  $R(\tilde{q}^{n+1}, q^n)$ , set  $q^{n+1} = \tilde{q}^{n+1}$ , otherwise, set  $q^{n+1} = q^n$ .*

(4) *Go back to step (1) with  $q^n \leftarrow q^{n+1}$ .*

Since  $T$  defines a Markov chain, we may always write  $\tilde{q}^{n+1} = \Phi(q^n, \xi^n)$  for some family of *i.i.d.* variables  $(\xi^n)_{n \geq 0}$ . We can then write  $q^{n+1}$  in a concise form:

$$(2.54) \quad q^{n+1} = \Phi(q^n, \xi^n) + \mathbb{1}_{U^n > R(\Phi(q^n, \xi^n), q^n)} (q^n - \Phi(q^n, \xi^n)) = \Psi(q^n, \xi^n, U^n),$$

where the  $U^n$  are *i.i.d.* uniform on  $[0, 1]$ , such that the  $(\xi^n, U^n)$  are an independent family. This shows that the algorithm defines a Markov chain. Furthermore, we may compute

$$\begin{aligned} \pi(x)\mathbb{P}(q^1 = y | q^0 = x) &= \pi(x)T(x, y)R(x, y) \\ &= \pi(x)T(x, y)r \left( \frac{\pi(y)T(y, x)}{\pi(x)T(x, y)} \right) \\ (2.55) \quad &= \pi(y)T(y, x)\frac{\pi(x)T(x, y)}{\pi(y)T(y, x)}r \left( \frac{\pi(y)T(y, x)}{\pi(x)T(x, y)} \right) \\ &= \pi(y)T(y, x)r \left( \frac{\pi(x)T(x, y)}{\pi(y)T(y, x)} \right) \text{ (Using (2.53))} \\ &= \pi(y)\mathbb{P}(q^1 = x | q^0 = y). \end{aligned}$$

Thus, the chain is reversible with respect to  $\pi$  which is then an invariant measure. Note that the algorithm is applicable even when we do not know how to evaluate  $\pi$ , but only the ratios  $\pi(x)/\pi(y)$ , which is in particular the case for Gibbs measures.

**Remark 8** (Rules for Metropolis-Hastings). *Possible choices for  $r$  are:*

1. *The Metropolis rule,*

$$r(x) = \min \{1, x\},$$

2. *The Barker rule,*

$$r(x) = \frac{x}{1+x},$$

3. *Any combination of these of the form, for  $\gamma > 0$ ,*

$$r(x) = \frac{x}{1+x} \left( 1 + 2 \left( \frac{1}{2} \min \left( r, \frac{1}{r} \right) \right)^\gamma \right).$$

**Remark 9** (Issues with Metropolized-schemes). *The Metropolis-Hastings algorithm provides a general recipe to define an unbiased Markov chain for the overdamped Langevin dynamics: one only needs to specify a proposition kernel, an acceptance rule, and a way to compute the ratio of the corresponding transition probabilities. For example, the so-called MALA scheme is very common, which combines a proposal function based on the Euler-Maruyama discretization of the dynamics (2.27), and a Metropolis acceptance rule. The use of the Metropolis-Hastings algorithm is not free. Because of the rejection rate, the correlations between samples decays at a slower rate,*

*resulting in a higher asymptotic variance for estimators based on ergodic averages. This implies that the use of a Metropolized scheme is only beneficial if the simulation time regime in which the systematic error overcomes the statistical error is computationally attainable. For most systems of interest, this is not the case. Furthermore, the possibility of rejection effectively slows down the dynamics, which may degrade the quality of estimates for dynamical quantities. The effect of the Metropolization procedure with the MALA scheme on the estimation of self-diffusion properties has been analyzed in [6], and improved rules and proposals for the computation of transport properties are discussed in detail in [7].*

### 2.2.9 Numerical illustrations

We now turn to a few numerical illustrations of the discussions above. In Figure 2.3, we compute numerically compute the pressure as a function of the density for the Lennard-Jones fluid, using the parameters for Argon, at two different temperatures. The resulting profile gives a numerical estimation of the equation of state of Argon, and comparison with experimental data shows that the agreement is very good at low densities and room temperature. The size of the systems was  $N = 2744$  atoms, using a sharp cutoff at  $r_c = 1.364$  nm, and a BOAB splitting scheme with  $\Delta t = 0.01$  ps. A reduced value of  $\gamma = 1.0$  was used for the friction parameter. We verified by using a block averaging procedure for the trajectories at both ends of the density range that the standard error bars are negligible.

In Figures 2.4 and 2.5, we highlight the systematic error in several observables. Simulations were run for systems of  $N = 27$  Lennard-Jones particles, at a temperature  $T = 1.25$  and at a reduced density  $\rho = 0.25$ . A cutoff of the potential at a distance  $r_c = 2.0$  was imposed, with a linear correction term ensuring that  $V$  be  $C^1$ . Additionally, regression lines were added by extrapolating the behavior at small  $\Delta t$  based on the theoretical expansions (2.52), and constraining the regression lines to converge in the limit  $\Delta t \rightarrow 0$ . This was achieved by a linear least squares regression procedure. The result also illustrate that a possibility to reduce the systematic error is to compute a desired average for several timesteps, and computing an extrapolated value at  $\Delta t = 0$  based on the theoretical knowledge of the behavior of the systematic error as  $\Delta t \rightarrow 0$ . This is known as Romberg extrapolation (ref?).

For the two configurational quantities we investigated, the virial and potential energy, we observe an overlap in the bias between the BAOAB and BAOA schemes. It so happens that this overlap can be simply explained by a result relating the invariant measures of certain pairs of numerical schemes (Lemma 2). In fact, a preprint [13] was recently posted, showing that a certain widely used scheme in the molecular dynamics community was equivalent to the BAOA scheme, and that the latter sampled the same configurational marginal as the BAOAB scheme. This prompted the need for a more thorough investigation of the properties of the BAOA scheme, which we discuss in the next chapter.

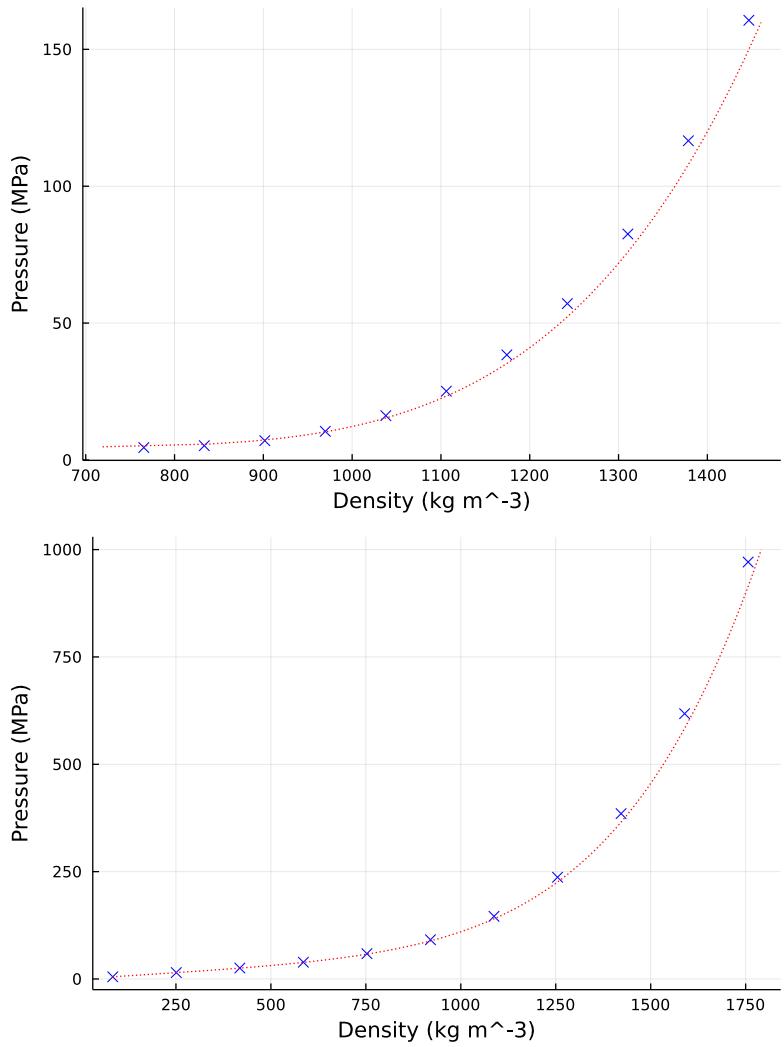


Figure 2.3: Simulated equations of state of Argon at 150 K (liquid phase, left) and 300 K (supercritical phase, right). Experimental reference curves are plotted in red, simulated data points are scattered in blue.

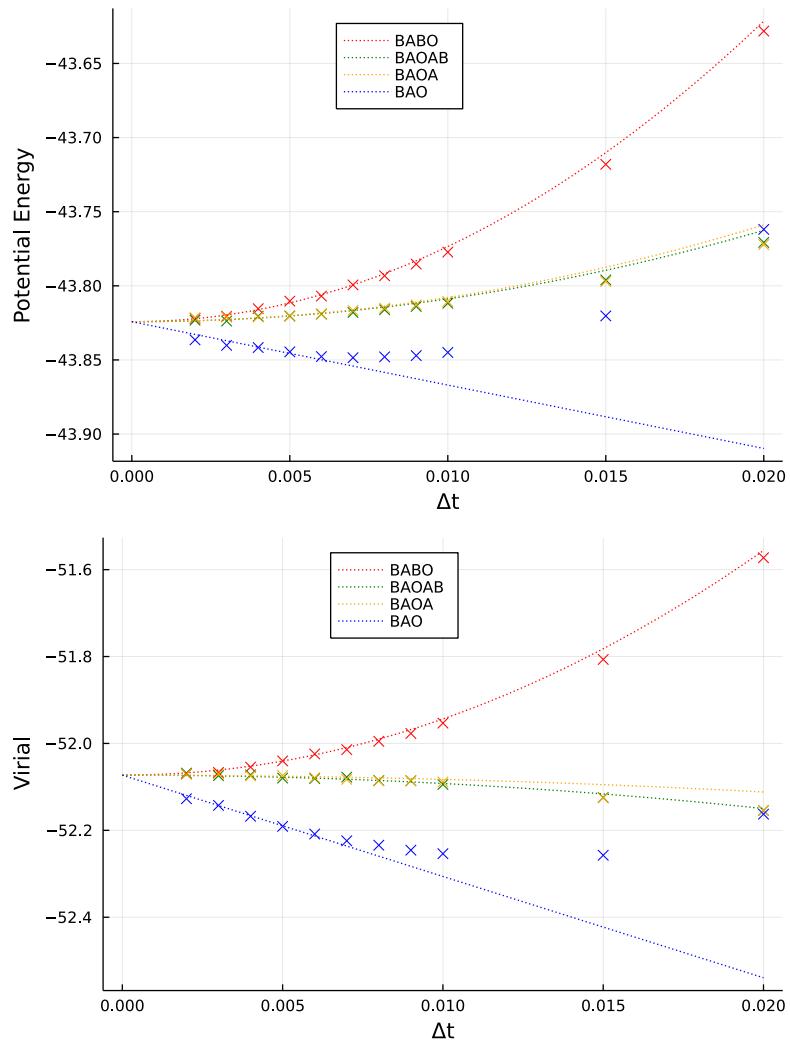


Figure 2.4: Systematic error in configurational quantities for a Lennard-Jones fluid.

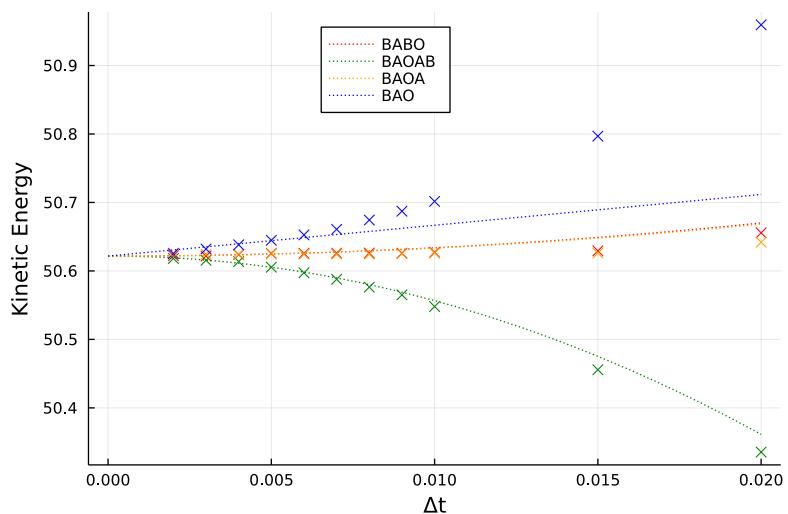


Figure 2.5: Systematic error in kinetic energy for a Lennard-Jones fluid.

## Chapter 3

# Study of the BAOA scheme

We consider time discretization schemes of underdamped Langevin dynamics known as the BAOA and BAOAB schemes, in order to compare the sampling bias induced by the timestep  $\Delta t$  for these two methods. Analysis of the timestep bias for the BAOAB scheme is given in [16], however it does not appear that the BAOA bias is so well understood. It has been observed numerically in [13] (Section III.B) that the bias on the kinetic marginal distribution is much lower using the BAOA method. We attempt to explain this from a theoretical point of view, before illustrating our results in numerical examples. Building on known results for the BAOAB scheme, we show the following results.

- (i) In section 3.0.2, we express the invariant measure of the BAOA scheme in terms of the invariant measure of the BAOAB scheme (Proposition 5), and using this expression, we show, as in [13] (Section II.C), the equality between their respective configurational marginal distributions (Corollary 1).
- (ii) In section 3.0.3, we show that the dominant error term for BAOA averages is only of order one in  $\Delta t$ , confirming that the BAOAB method is in general of higher order (Corollary 2).
- (iii) We show in section 3.0.4 that for kinetic observables, however, the error is of second order in  $\Delta t$ , so that both marginal distributions are second-order accurate (Corollary 3).
- (iv) In section 3.0.5, we give an expression for the dominant error term in the kinetic marginal distribution of the BAOA scheme (Proposition 6). In fact, we conjecture that, at least in dimension one, this term cancels, leading to an order of at least  $\Delta t^3$  (Conjecture 1).
- (v) Lastly, in section 3.0.6, we analyze the difference between the kinetic marginal distribution under the BAOA and BAOAB scheme (Proposition 7), and explain why this difference leads to a systematic underestimation of the kinetic variance in BAOAB trajectories (Remark 11).

### 3.0.1 Definitions and notations

We consider two splitting schemes for (2.14), as defined by the following evolution operators:

$$(3.1) \quad \begin{cases} P_{\Delta t} = e^{\Delta t B} e^{\frac{\Delta t}{2} A} e^{\Delta t \gamma C} e^{\frac{\Delta t}{2} A}, \\ Q_{\Delta t} = e^{\frac{\Delta t}{2} B} e^{\frac{\Delta t}{2} A} e^{\Delta t \gamma C} e^{\frac{\Delta t}{2} A} e^{\frac{\Delta t}{2} B}. \end{cases}$$

These correspond respectively to the BAOA and the BAOAB scheme. We also denote  $\mu_{\Delta t, P}, \mu_{\Delta t, Q}$  the invariant measures for the Markov chains associated with (3.1). We assume these have smooth densities which we also denote  $\mu_{\Delta t, P}, \mu_{\Delta t, Q}$ , and that a certain ergodicity condition holds (see Lemma 2). Additionally we denote by  $\nu_{\Delta t, P}, \nu_{\Delta t, Q}, \kappa_{\Delta t, P}, \kappa_{\Delta t, Q}$  the associated marginals and densities with obvious notation inspired by (1.8).

### 3.0.2 Relating invariant measures of discretization schemes

In this paragraph, we provide a formula for  $\mu_{\Delta t, P}$  in terms of  $\mu_{\Delta t, Q}$ . This result allows one to very simply show the equality in the configurational marginals between these two measures, as noted in [13]. The main tool is the following result, which is a reformulation of the TU lemma (Lemma 9 from [16]).

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

$$R_{\Delta t} P_{\Delta t}^n = Q_{\Delta t}^n S_{\Delta t},$$

where  $R_{\Delta t}$  and  $S_{\Delta t}$  are bounded operators on  $B^\infty(\mathcal{E})$ , such that  $R_{\Delta t} \mathbb{1} = \mathbb{1}$ , 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 have the relation  $\mu_{\Delta t, P}$  and  $\mu_{\Delta t, Q}$  via the following relation:

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

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

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

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

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

Since  $R P_{\Delta t}^n = Q_{\Delta t}^n S$ , identifying the two limits yields (2) □

Applying Lemma 2 to (3.1) yields the following result.

**Proposition 5.** *The following relation between the densities  $\mu_{\Delta t, P}$  and  $\mu_{\Delta t, Q}$  holds.*

$$(3.2) \quad \mu_{\Delta t, P}(q, p) = \mu_{\Delta t, Q} \left( q, p - \frac{\Delta t}{2} V(q) \right).$$

*Proof.* From the expressions (3.1), we immediately get:

$$(3.3) \quad P_{\Delta t}^n e^{\frac{\Delta t}{2}B} = e^{\frac{\Delta t}{2}B} Q_{\Delta t}^n,$$

whereby applying Lemma 2, we get for any test function  $\varphi$ ,

$$(3.4) \quad \int_{\mathcal{E}} e^{\frac{\Delta t}{2}B} \varphi d\mu_{P,\Delta t} = \int_{\mathcal{E}} \varphi d\mu_{Q,\Delta t}$$

Using equation (3.4) with  $\psi = e^{-\frac{\Delta t}{2}B} \varphi$  yields an exact expression for  $\mu_{\Delta t,P}$  in terms of  $\mu_{\Delta t,Q}$ :

$$(3.5) \quad \int_{\mathcal{E}} \varphi d\mu_{\Delta t,P} = \int_{\mathcal{E}} e^{-\frac{\Delta t}{2}B} \varphi d\mu_{\Delta t,Q}.$$

Since  $\varphi$  is arbitrary, we infer that at the level of densities,

$$(3.6) \quad \mu_{\Delta t,P}(q,p) = \left( e^{-\frac{\Delta t}{2}B} \right)^{\dagger} \mu_{\Delta t,Q}(q,p),$$

where  $\dagger$  denotes the adjoint on the flat space  $L^2(\mathcal{E})$ . A simple computation shows that

$$e^{-\frac{\Delta t}{2}B^{\dagger}} = e^{\frac{\Delta t}{2}B},$$

since  $B^{\dagger} = -B$ . Hence,

$$(3.7) \quad \mu_{\Delta t,P}(q,p) = e^{\frac{\Delta t B}{2}} \mu_{\Delta t,Q}(q,p) = \mu_{\Delta t,Q} \left( q, p - \frac{\Delta t}{2} \nabla V(q) \right),$$

which is the desired conclusion.  $\square$

Relation (3.7) is enough to show an equality between the configurational marginal distributions  $\nu_{\Delta t,P}$  and  $\nu_{\Delta t,Q}$ , as noted in [13].

**Corollary 1.** *The marginal distributions in the  $q$  variable of  $\mu_{\Delta t,P}$  and  $\mu_{\Delta t,Q}$  coincide:*

$$(3.8) \quad \nu_{\Delta t,Q}(q) = \nu_{\Delta t,P}(q).$$

*Proof.* Write, for any  $q \in \mathcal{D}$ ,

$$\begin{aligned} \nu_{\Delta t,Q}(q) &= \int_{\mathbb{R}^{dN}} \mu_{\Delta t,Q}(q,p) dp \\ &= \int_{\mathbb{R}^{dN}} \mu_{\Delta t,Q} \left( q, p - \frac{\Delta t}{2} \nabla V(q) \right) dp \\ &= \int_{\mathbb{R}^{dN}} \mu_{\Delta t,P}(q,p) dp \\ &= \nu_{\Delta t,P}(q), \end{aligned}$$

which proves the claim.  $\square$

### 3.0.3 Error estimate on the phase space measure

We now turn to obtaining the dominant order in the sampling bias of  $\mu_{\Delta t,P}$ , building on previously known results for  $\mu_{\Delta t,Q}$ , and the relation (3.2). Error estimates on  $\mu_{\Delta t,Q}$  have been investigated in [16] (Section 1.4). In particular, the following expansion of  $\mu_{\Delta t,Q}$  is derived, which will be central in our analysis.

**Theorem 2** (Theorem 13 in [16]). *There exists a smooth function  $f_2$  such that for any smooth  $\psi$ ,*

$$(3.9) \quad \int_{\mathcal{E}} \psi(q, p) \mu_{\Delta t, Q}(q, p) dq dp = \int_{\mathcal{E}} \psi(q, p) \mu(q, p) dq dp + \Delta t^2 \int_{\mathcal{E}} \varphi(q, p) f_2(q, p) \mu(q, p) dq dp + \Delta t^4 r_{\psi, \gamma, \Delta t},$$

where the remainder  $r_{\psi, \gamma, \Delta t}$  is uniformly bounded for  $\Delta t$  sufficiently small. Moreover, an expression for the dominant error term is obtained,

$$(3.10) \quad \begin{cases} f_2 = \tilde{f}_2 - \frac{1}{8}(A + B)g \\ \mathcal{L}_\gamma^* \tilde{f}_2 = \frac{1}{12}(A + B) \left[ \left( A + \frac{B}{2} \right) g \right], \\ g := \beta(M^{-1}p) \cdot \nabla V(q) \end{cases}$$

where  $\mathcal{L}_\gamma^*$  is the adjoint of  $\mathcal{L}_\gamma$  on the weighted space  $L^2(\mu)$ .

Smoothness here is meant in a technical sense (see Definition 8 in [16]), which we refrain from detailing here. Using this expansion, one can derive the dominant order error for the BAOA scheme.

**Corollary 2.** *For any smooth observable  $\varphi$ ,*

$$(3.11) \quad \int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, P}(q, p) dq dp = \int_{\mathcal{E}} \varphi(q, p) \left( 1 + \frac{\Delta t}{2} g(q, p) \right) \mu(q, p) dq dp + O(\Delta t^2),$$

where  $g$  is given by (3.10).

*Proof.* Combining (3.9) with (3.2), we get the following estimation for averages with respect to  $\mu_{\Delta t, P}$ :

$$(3.12) \quad \int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, P}(q, p) dq dp = \int_{\mathcal{E}} \varphi(q, p) \mu \left( q, p - \frac{\Delta t}{2} \nabla V(q) \right) dq dp + O(\Delta t^2).$$

Taylor expanding  $\mu$  gives

$$\mu \left( q, p - \frac{\Delta t}{2} \nabla V(q) \right) = \mu(q, p) \left( 1 + \frac{\Delta t}{2} \beta(M^{-1}p) \cdot \nabla V(q) + O(\Delta t^2) \right) = \mu(q, p) \left( 1 + \frac{\Delta t}{2} g + O(\Delta t^2) \right),$$

hence we get

$$(3.13) \quad \int_{\mathcal{E}} \varphi(q, p) \mu_{\Delta t, P}(q, p) dq dp = \int_{\mathcal{E}} \varphi(q, p) \mu(q, p) \left( 1 + \frac{\Delta t}{2} g(q, p) + O(\Delta t^2) \right) dq dp,$$

which proves the claim.  $\square$

### 3.0.4 Error estimates on the kinetic marginal distributions

Equation (3.11) expresses the fact that the invariant measure  $\mu_{\Delta t, P}$  is only exact at first order in  $\Delta t$ , which is one less than  $\mu_{\Delta t, Q}$ . So in full generality, one can expect an error of order  $\Delta t$  on averages obtained from BAOA trajectories, versus  $\Delta t^2$  for averages computed from BAOAB trajectories. However, if we restrict ourselves to marginal observables, that is observables which only depend on the configurational coordinate or the kinetic coordinate, the first order error term vanishes. Indeed, we have the following.

**Corollary 3.** *Let  $\varphi(q, p) = \varphi(q)$  or  $\varphi(q, p) = \varphi(p)$  be a marginal observable. Then*

$$\int_{\mathcal{E}} \varphi dq d\mu_{\Delta t, P} = \int_{\mathcal{E}} \varphi dq d\mu + O(\Delta t^2).$$

*Proof.* By (3.11), it is sufficient to show

$$(3.14) \quad \int_{\mathcal{E}} \varphi g \, d\mu = 0.$$

This follows from the following cancellations.

$$(3.15) \quad \int_{\mathbb{R}^{dN}} g(q, p) \mu(q, p) \, dp = \int_{\mathcal{D}} g(q, p) \mu(q, p) \, dq = 0.$$

Indeed,

$$\int_{\mathbb{R}^{dN}} \beta(M^{-1}p) \cdot \nabla V(q) \mu(q, p) \, dp = 0,$$

since the integrand is an odd function of  $p$ , and the marginal of  $\mu$  in  $p$  is a centered Gaussian density. Also,

$$\int_{\mathcal{D}} \beta(M^{-1}p) \cdot \nabla V(q) \mu(q, p) \, dq = - \int_{\mathcal{D}} (M^{-1}p) \cdot \nabla_q \mu(q, p) \, dq = 0,$$

by an integration by parts. By first integrating (3.14) over the coordinate independent of  $\varphi$ , one of the cancellations (3.15) yields the result.  $\square$

Corollary 2 gives no new information concerning configurational observables, since we already know by Corollary 1 that these have the same averages under  $\mu_{\Delta t, P}$  and  $\mu_{\Delta t, Q}$ , and that by Theorem 2, these have error of order  $\Delta t^2$ . However, kinetic observables may yield different averages.

### 3.0.5 Analysis of the second order error term for kinetic averages under $\mu_{\Delta t, P}$

It was observed numerically in [13] that the averages of the kinetic and configurational temperatures computed with a BAOA scheme have a bias of order greater than  $\Delta t$ , as expected from the argument above. In fact, for the kinetic temperature, the order appears to be greater than  $\Delta t^2$ , in contrast to averages computed with the BAOAB method. Understanding this behavior theoretically requires comparing second order error terms.

We show the following result, which identifies the second-order error term for kinetic observables.

**Proposition 6.** *Let  $\psi(q, p) = \psi(p)$  be a smooth kinetic observable. Then,*

$$(3.16) \quad \int_{\mathcal{E}} \psi(p) \mu_{\Delta t, P}(q, p) \, dq \, dp = \int_{\mathcal{E}} \psi(p) \mu(q, p) \, dp \, dq + \Delta t^2 \int_{\mathcal{E}} \psi(p) \tilde{f}_2(q, p) \mu(q, p) \, dq \, dp + O(\Delta t^3),$$

where  $\tilde{f}_2$  is given by (3.10).

From Theorem 13 of [16], this error term is identical to the dominant error term for OBABO averages, and minus the dominant error term for OABAO averages.

*Proof.* By writing

$$\int_{\mathcal{E}} \psi(q, p) \mu_{\Delta t, P}(q, p) \, dq \, dp - \int_{\mathcal{E}} \psi(q, p) \mu(q, p) \, dp \, dq = \int_{\mathcal{E}} \left( \psi(q, p) - \int_{\mathcal{E}} \psi \, d\mu \right) \mu_{\Delta t, P}(q, p) \, dq \, dp,$$

we may assume without loss of generality that  $\psi$  has average 0 with respect to  $\mu$ .

Using (3.9), we get

$$\int_{\mathcal{E}} \psi(p) \mu_{\Delta t, Q}(q, p) \, dq \, dp = \Delta t^2 \int_{\mathcal{E}} \psi(p) f_2(q, p) \mu(q, p) \, dq \, dp + O(\Delta t^3),$$

so that using our relation (3.2), we get

$$\int_{\mathcal{E}} \psi(p) \mu_{\Delta t, P}(q, p) dq dp = \int_{\mathcal{E}} \psi(p) e^{\frac{\Delta t}{2} B} \mu(q, p) dq dp + \Delta t^2 \int_{\mathcal{E}} \psi(p) e^{\frac{\Delta t}{2} B} [f_2(q, p) \mu(q, p)] dq dp + O(\Delta t^3).$$

This rewrites, at dominant order,

$$\int_{\mathcal{E}} \psi(p) \mu \left( q, p - \frac{\Delta t}{2} \nabla V(q) \right) dq dp + \Delta t^2 \int_{\mathcal{E}} \psi(p) f_2(q, p) \mu(q, p) dq dp + O(\Delta t^3).$$

Expanding  $\mu$  to the second order yields

$$\begin{aligned} & \mu \left( q, p - \frac{\Delta t}{2} \nabla V(q) \right) + O(\Delta t^3) \\ &= \mu(q, p) \left[ 1 + \beta \frac{\Delta t}{2} (M^{-1}p) \cdot \nabla V(q) + \frac{\Delta t^2}{8} [(\beta M^{-1}p) \otimes (\beta M^{-1}p) \nabla V(q)] \cdot \nabla V(q) - \beta \frac{\Delta t^2}{8} (M^{-1} \nabla V(q)) \cdot \nabla V(q) \right] \\ &= \mu(q, p) \left[ 1 + \beta \frac{\Delta t}{2} (M^{-1}p) \cdot \nabla V(q) + \frac{\Delta t^2}{8} (g^2(q, p) - \beta (M^{-1} \nabla V(q)) \cdot \nabla V(q)) \right]. \end{aligned}$$

Using  $\int \psi d\mu = 0$  and the cancellation (3.15) on  $q$  to remove the first order term, we obtain:

$$(3.17) \quad \int_{\mathcal{E}} \psi(p) \mu_{\Delta t, P}(q, p) dq dp = \Delta t^2 \int_{\mathcal{E}} \psi(p) \left( \frac{1}{8} (g^2(q, p) - \beta (M^{-1} \nabla V(q)) \cdot \nabla V(q)) + f_2(q, p) \right) \mu(q, p) dq dp + O(\Delta t^3).$$

Simplifications are possible. First, observe that

$$-\beta (M^{-1} \nabla V(q)) \cdot \nabla V(q) = B g(q, p),$$

so that, using the expression for  $f_2$  given in (3.10), we get

$$(3.18) \quad \int_{\mathcal{E}} \psi(p) \mu_{\Delta t, P}(q, p) dq dp = \Delta t^2 \int_{\mathcal{E}} \psi(p) \left( \frac{1}{8} (g^2(q, p) - Ag(q, p)) + \tilde{f}_2(q, p) \right) \mu(q, p) dq dp + O(\Delta t^3).$$

Next, we examine the term

$$(g^2(q, p) - Ag(q, p)) \mu(q, p) = [\beta^2 ((M^{-1}p) \cdot \nabla V(q))^2 - \beta (M^{-1}p) \cdot (\nabla^2 V(q) M^{-1}p)] \mu(q, p),$$

by a straightforward calculation, where  $\nabla^2$  denotes the Hessian matrix. This expression is a finite sum of diagonal terms coming from both terms inside the brackets, and off-diagonal terms coming only from the rightmost term inside the bracket. Importantly, these all vanish when integrated against the configurational marginal of  $\mu$ . To make this precise, we index  $p$  and  $q$  as

$$p = (p_i)_{1 \leq i \leq dN}, \quad q = (q_i)_{1 \leq i \leq dN}.$$

Fixing indices  $i \neq j$ , the diagonal term corresponding to  $i$  is

$$(3.19) \quad \left[ \beta^2 (M^{-1}p)_i^2 \left( \frac{\partial}{\partial q_i} V(q) \right)^2 - \beta (M^{-1}p)_i^2 \frac{\partial^2}{\partial q_i^2} V(q) \right] \mu(q, p) = (M^{-1}p)_i^2 \frac{\partial^2}{\partial q_i^2} \mu(q, p),$$

and the off-diagonal term corresponding to  $(i, j)$  is

$$(3.20) \quad -\beta (M^{-1}p)_i (M^{-1}p)_j \frac{\partial}{\partial q_i} V(q) \frac{\partial}{\partial q_j} V(q) \mu(q, p) = -\frac{1}{\beta} (M^{-1}p)_i (M^{-1}p)_j \frac{\partial^2}{\partial q_i \partial q_j} \mu(q, p).$$

Factoring out the  $q$ -independent terms, and using the cancellations

$$(3.21) \quad \int_{\mathcal{D}} \frac{\partial^2}{\partial q_i^2} \mu(q, p) dq = \int_{\mathcal{D}} \frac{\partial^2}{\partial q_i \partial q_j} \mu(q, p) dq = 0,$$

which follow by integration by parts, we infer

$$(3.22) \quad \int_{\mathcal{E}} \psi(p) \mu_{\Delta t, P}(q, p) dq dp = \Delta t^2 \int_{\mathcal{E}} \psi(p) \tilde{f}_2(q, p) \mu(q, p) dq dp + O(\Delta t^3),$$

which concludes the proof.  $\square$

**Remark 10.** Using exponential decay estimates on the evolution semigroup  $(e^{t\mathcal{L}_\gamma})_{t \geq 0}$  like (2.32), one can show that the inverse operator  $\mathcal{L}_\gamma^{-1}$  is well-defined for smooth centered observables. Thus  $\mathcal{L}_\gamma^{-1}\psi$  is well defined, say  $\mathcal{L}_\gamma\Psi(q, p) = \psi(p)$ . Hence, (3.16) rewrites

$$\begin{aligned} \int_{\mathcal{E}} \psi(p) \mu_{\Delta t, P}(q, p) dq dp &= \Delta t^2 \int_{\mathcal{E}} \mathcal{L}_\gamma\Psi(q, p) \tilde{f}_2(q, p) \mu(q, p) dq dp + O(\Delta t^3) \\ &= \Delta t^2 \int_{\mathcal{E}} \Psi(q, p) \mathcal{L}_\gamma^* \tilde{f}_2(q, p) \mu(q, p) dq dp + O(\Delta t^3) \\ &= \frac{\Delta t^2}{12} \int_{\mathcal{E}} \Psi(q, p) \left[ (A + B) \left( A + \frac{B}{2} \right) g \right] (q, p) \mu(q, p) dq dp + O(\Delta t^3), \end{aligned}$$

using (3.10), which provides an alternative expression for the dominant error term. Numerical evidence (see Figure 3.11) suggests that the error on BAOA and BAOAB averages is at dominant order independent of  $\gamma$ . Since the error term on BAOA given in (3.16) depends on  $\gamma$ , this suggests that this term is zero, motivating the following conjecture.

**Conjecture 1.** For any smooth centered kinetic observable  $\psi(p)$ , we have

$$(3.23) \quad \int_{\mathcal{E}} (\mathcal{L}_\gamma^{-1}\psi)(q, p) \left[ (A + B) \left( A + \frac{B}{2} \right) g \right] (q, p) \mu(q, p) dq dp = 0.$$

This would in particular imply that the kinetic marginal  $\kappa_{\Delta t, P}$  is correct at order at least three in  $\Delta t$ , and is the subject of further investigation.

### 3.0.6 Analysis of the discrepancy between the dominant error terms on the kinetic marginals.

Numerical evidence presented in [13] shows a significant discrepancy between  $\kappa_{\Delta t, P}$  and  $\kappa_{\Delta t, Q}$ . Specifically,  $\kappa_{\Delta t, Q}$  in the case  $d = N = 1$  tends to present a sharper peak than  $\kappa_{\Delta t, P}$ , thus underestimating the variance in the kinetic marginal. We show this in this paragraph this behavior is generic, in the sense that it does not, up to a shape parameter, depend on  $V$ . The arguments above show that

$$(3.24) \quad \begin{aligned} \int_{\mathcal{E}} \psi(p) \mu_{\Delta t, P}(q, p) dq dp &= \int_{\mathbb{R}^{dN}} \psi(p) \kappa_{\Delta t, P}(p) dp \\ &= \int_{\mathbb{R}^{dN}} \psi(p) \kappa(p) dp + \Delta t^2 \int_{\mathbb{R}^{dN}} \psi(p) \left( \int_{\mathcal{D}} \tilde{f}_2(q, p) \nu(q) dq \right) \kappa(p) dp + O(\Delta t^3), \end{aligned}$$

where we used the product form (1.8) for  $\mu$ . Similarly,

$$(3.25) \quad \int_{\mathcal{E}} \psi(p) \kappa_{\Delta t, Q}(p) dp = \int_{\mathbb{R}^{dN}} \psi(p) \kappa(p) dp + \Delta t^2 \int_{\mathbb{R}^{dN}} \psi(p) \left( \int_{\mathcal{D}} f_2(q, p) \nu(q) dq \right) \kappa(p) dp + O(\Delta t^3),$$

so that

$$\begin{aligned} \int_{\mathcal{E}} \psi(p) (\kappa_{\Delta t, P}(p) - \kappa_{\Delta t, Q}(p)) dp &= \Delta t^2 \int_{\mathbb{R}^{dN}} \psi(p) \left( \int_{\mathcal{D}} (\tilde{f}_2(q, p) - f_2(q, p)) \nu(q) dq \right) \kappa(p) dp + O(\Delta t^3) \\ &= \frac{\Delta t^2}{8} \int_{\mathbb{R}^{dN}} \psi(p) \left( \int_{\mathcal{D}} (A + B) g(q, p) \nu(q) dq \right) \kappa(p) dp + O(\Delta t^3). \end{aligned}$$

Hence at the level of densities, we have at dominant order,

$$\kappa_{\Delta t, P}(p) - \kappa_{\Delta t, Q}(p) = \frac{\Delta t^2 \kappa(p)}{8} \int_{\mathcal{D}} (A + B) g(q, p) \nu(q) dq + O(\Delta t^3),$$

using the expressions for  $f_2$  and  $\tilde{f}_2$  given in (3.10). The following proposition gives an alternative expression for this discrepancy term.

**Proposition 7.** *We have the following expression for the discrepancy term.*

$$(3.26) \quad \kappa_{\Delta t, P}(p) - \kappa_{\Delta t, Q}(p) = \frac{\Delta t^2}{8} \text{Tr} \left( \left( (\beta M^{-1} p)^{\otimes 2} - \beta M^{-1} \right)^T \text{Cov}_{\nu}(\nabla V) \right) \kappa(p) + O(\Delta t^3).$$

*Proof.* For simplicity we assume  $Z_{\nu} = \frac{\Delta t^2}{8} = \kappa(p) = 1$ . This has no incidence on our computations. We write:

$$(3.27) \quad (A + B)g(q, p)\nu(q) = \beta \left[ (M^{-1}p) \cdot (\nabla^2 V(q)M^{-1}p) - (M^{-1}\nabla V(q)) \cdot \nabla V(q) \right] e^{-\beta V(q)}.$$

Setting  $\tilde{p} = M^{-1}p$ , we get

$$(3.28) \quad (A + B)g(q, M\tilde{p})\nu(q) = \beta \left[ \tilde{p} \cdot (\nabla^2 V(q)\tilde{p}) - (M^{-1}\nabla V(q)) \cdot \nabla V(q) \right] e^{-\beta V(q)}.$$

This is a sum of terms of the form

$$T_{ij}(q, p) = \left[ \beta \tilde{p}_i \tilde{p}_j \frac{\partial^2}{\partial q_i \partial q_j} V(q) - \beta M_{i,j}^{-1} \frac{\partial}{\partial q_i} V(q) \frac{\partial}{\partial q_j} V(q) \right] e^{-\beta V(q)}.$$

Upon integrating this term over  $\mathcal{D}$ , we can integrate the left-most term by parts (boundary terms cancel out by periodicity or by growth conditions on  $V$ ), to obtain

$$\int_{\mathcal{D}} T_{ij}(q, p) dq = \int_{\mathcal{D}} \left[ \beta^2 \tilde{p}_i \tilde{p}_j \frac{\partial}{\partial q_i} V(q) \frac{\partial}{\partial q_j} V(q) - \beta M_{i,j}^{-1} \frac{\partial}{\partial q_i} V(q) \frac{\partial}{\partial q_j} V(q) \right] e^{-\beta V(q)} dq.$$

Hence,

$$\int_{\mathcal{D}} T_{ij}(q, p) dq = (\beta^2 \tilde{p}_i \tilde{p}_j - \beta M_{i,j}^{-1}) \int_{\mathcal{D}} \frac{\partial}{\partial q_i} V(q) \frac{\partial}{\partial q_j} V(q) e^{-\beta V(q)} dq,$$

so that

$$\int_{\mathcal{D}} (A + B)g(q, p)\nu(q) dq = \sum_{i,j} (\beta^2 \tilde{p}_i \tilde{p}_j - \beta M_{i,j}^{-1}) \int_{\mathcal{D}} \frac{\partial}{\partial q_i} V(q) \frac{\partial}{\partial q_j} V(q) e^{-\beta V(q)} dq,$$

which we rewrite

$$(3.29) \quad \left( (\beta M^{-1} p)^{\otimes 2} - \beta M^{-1} \right) : \int_{\mathcal{D}} (\nabla V \otimes \nabla V)(q) \nu(q) dq = \text{Tr} \left( \left( (\beta M^{-1} p)^{\otimes 2} - \beta M^{-1} \right)^T \text{Cov}_{\nu}(\nabla V) \right),$$

using the fact that  $\nabla V$  is a centered observable with respect to  $\nu$ , and concluding the proof.  $\square$

**Remark 11.** This expression for the discrepancy term is not particularly wieldy, however it does explain the behavior observed in [13]. In the case  $d = N = \beta = M = 1$ , it becomes,

$$\kappa_{\Delta t, P}(p) - \kappa_{\Delta t, Q}(p) = \frac{\Delta t^2}{8}(p^2 - 1)\text{Var}_\nu(V')\kappa(p) + O(\Delta t^3),$$

which is, up to a constant, the same correction term for any potential  $V$ . We plot this correction profile in figure 3.1. The shape of this profile explains the higher peak observed in  $\kappa_{\Delta t, Q}$ .

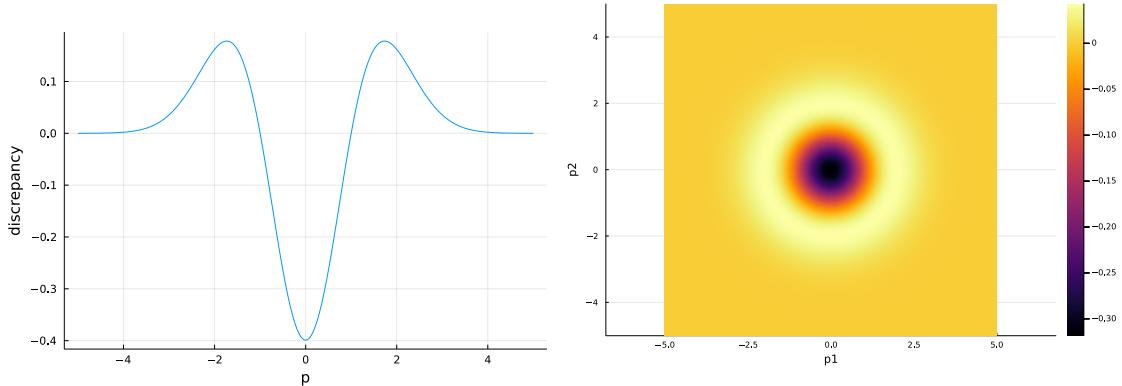


Figure 3.1: Profile of the discrepancy term in one and two dimension, in the case of identity covariances for  $\nabla V$ .

### 3.1 Numerical results

We propose illustrating our computations with numerical examples, on toy one dimensional systems.

- (i) In section 3.1.1, we define the potentials used for all the following experiments, and describe the sampling method used.
- (ii) In section 3.1.2, we verify numerically the relation (3.8).
- (iii) In section 3.1.3, we show that there is a significant discrepancy between the two kinetic marginal distributions. We also pinpoint the main, and possibly only source of this error, namely the  $\gamma$ -independent term (3.26).
- (iv) In section 3.1.4, we numerically verify that the first order behavior (3.13) is correct.
- (v) In section 3.1.5, we give an explicit example of an observable for which the BAOA scheme yields a bias of order  $\Delta t$ .
- (vi) Finally, in section 3.1.6, we show that the effect of the parameter  $\gamma$  is undetectable at the level of the kinetic marginals, motivating Conjecture 1.

#### 3.1.1 Models

We take  $\beta = 1$ ,  $M = \text{Id}$ , and consider four potentials:

- Periodic potential

$$\mathcal{D} = L(\mathbb{R}/\mathbb{Z}), \quad L = 1, \quad V(q) = \sin(2\pi q/L),$$

- Quadratic potential

$$\mathcal{D} = \mathbb{R}, V(q) = \alpha \frac{q^2}{2}, \alpha = 1,$$

- Double well potential

$$\mathcal{D} = \mathbb{R}, V(q) = \alpha \frac{q^2}{2} + \beta e^{-\frac{q^2}{2\sigma^2}}, \alpha = 1, \beta = 4, \sigma = 0.5,$$

- Tilted double well potential

$$\mathcal{D} = \mathbb{R}, V(q) = \alpha \frac{q^2}{2} + \gamma q + \beta e^{-\frac{q^2}{2\sigma^2}}, \alpha = 1, \beta = 4, \gamma = 1, \sigma = 0.5.$$

Analytically unknown normalizing constants and reference quantities were obtained through numerical integration of  $\mu$ , using trapezoid rules with a mesh size of  $10^{-6}$ . For unbounded coordinates, we truncated the domain to the interval  $[-5, 5]$ . Approximations of  $\mu_{\Delta t, P}, \mu_{\Delta t, Q}$  were computed by recording the states of 10,000 independently evolving trajectories over  $2 \times 10^6$  timesteps in a  $1000 \times 1000$  two-dimensional histogram on the truncated domain. The rare sample points outside of the truncated domain were discarded.

### 3.1.2 Equality of marginal configurational distributions

On Figures 3.2 and 3.3, we verify numerically the equality (3.8) between the configurational marginal distributions  $\nu_{\Delta t, Q}$  and  $\nu_{\Delta t, P}$ , which holds for any  $\Delta t$ . This point was demonstrated in [13].

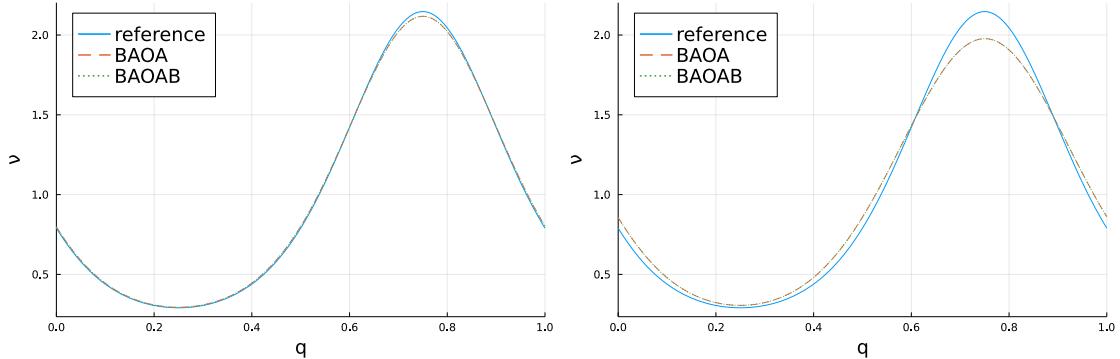


Figure 3.2: Marginal configurational distributions for the periodic potential. Left:  $\Delta t = 0.1$ . Right:  $\Delta t = 0.2$ . Even for large timesteps, the distributions coincide perfectly.

### 3.1.3 Comparison of marginal kinetic distributions

We observe, as in [13], that the kinetic marginal distribution  $\kappa_{\Delta t, Q}$  departs from the reference at a faster rate than  $\kappa_{\Delta t, P}$ , and more precisely appears to underestimate the variance, leading to a sharper distribution. Additionally we observe that removing the part of the bias on BAOAB due to the discrepancy term (3.26) leads to a significant improvement. These corrected marginals are plotted under the label "correction". See Figures 3.4 and 3.5.

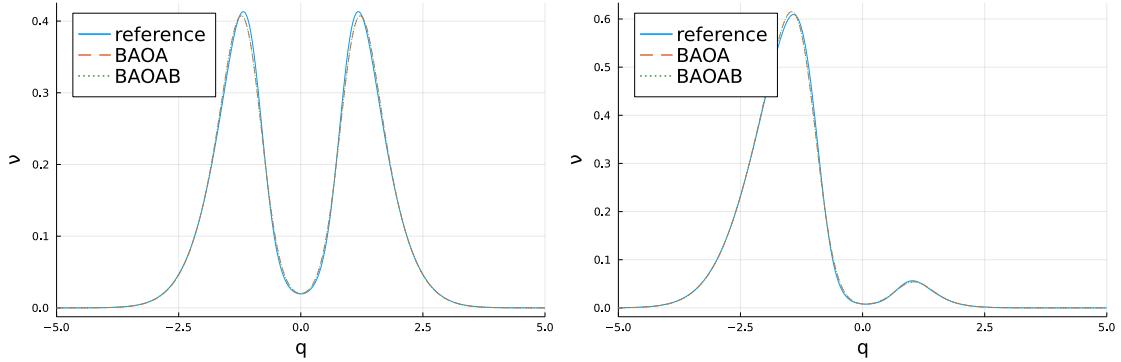


Figure 3.3: Marginal configurational distributions for  $\Delta t = 0.4$ . Left: double well potential. Right: tilted double well potential.

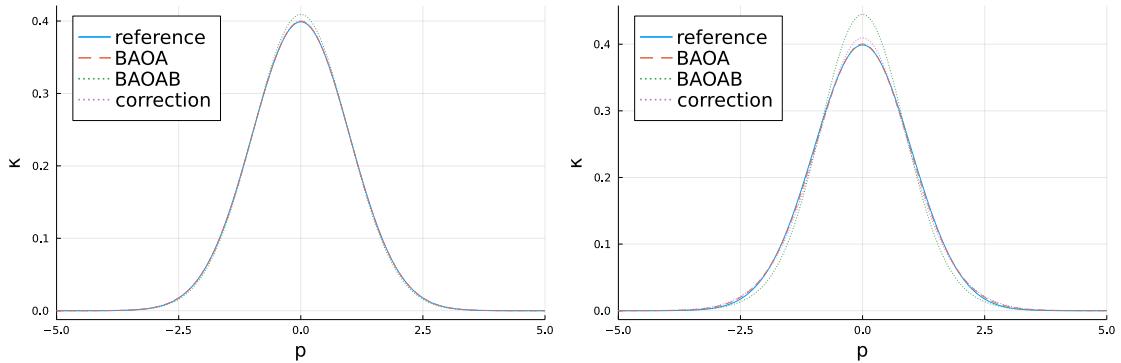


Figure 3.4: Marginal kinetic distributions for the periodic potential. Left:  $\Delta t = 0.1$ . Right:  $\Delta t = 0.2$ .

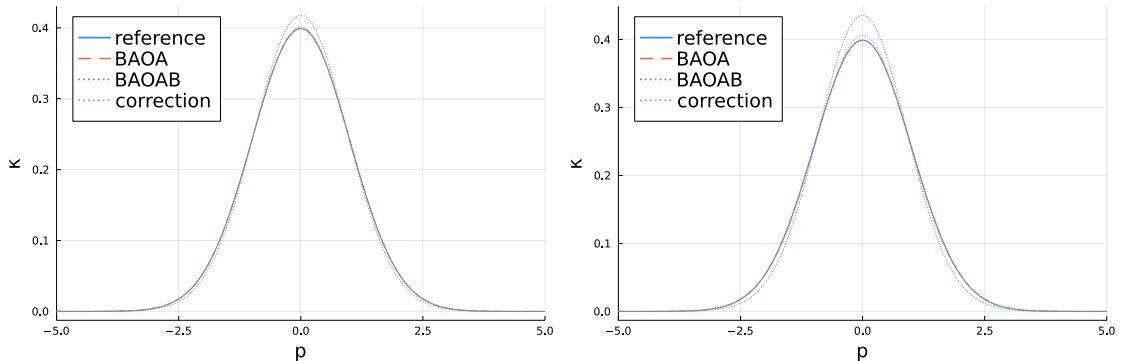


Figure 3.5: Marginal kinetic distributions for the double well potential. Left:  $\Delta t = 0.3$ . Right:  $\Delta t = 0.4$ .

### 3.1.4 Verification of the first-order expansion

We verify the correctness first-order expansion of  $\mu_{\Delta,P}$  obtained in (3.13), by comparing the joint distributions obtained from Monte-Carlo simulations with a reference calculation of the first-order expansion for  $\mu_{\Delta t,P}$ . Additionally, we plot the empirical estimate of  $\mu_{\Delta t,Q}$  and  $\mu$ . The plots show joint likelihoods as a function of the state, using a color mapping. Empirical joint distributions for BAOA and BAOAB trajectories are plotted on the top row of each figure. On the bottom row, a reference computation of  $\mu$  is plotted on the right, as well as a reference computation of

$$\left(1 + \frac{\Delta t}{2} g\right) \mu$$

on the left. The results visually confirm our result, while suggesting that, as a whole,  $\mu_{\Delta t,Q}$  is the superior approximation of  $\mu$ . See Figures 3.6, 3.7 and 3.8.

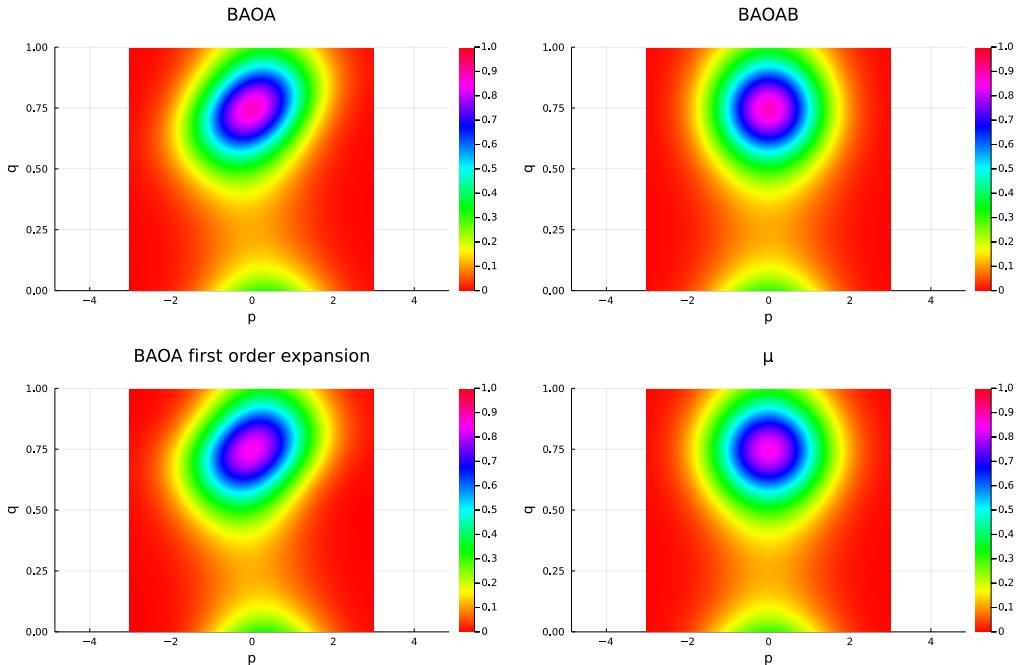


Figure 3.6: Joint distributions for the periodic potential,  $\Delta t = 0.1$ .

### 3.1.5 Example of first-order bias in a BAOA average

We demonstrate that for certain observables, BAOA is drastically outperformed by BAOAB, by calculating the average of  $g$  for increasing timesteps. Note by (3.15), the true average is 0. Figures 3.9 and 3.10 show the estimated averages as a function of the timestep on the left, and the same data on a log-log plot on the right. The order of the error on the BAOAB averages suggest that the second order error term

$$\int_{\mathcal{E}} g f_2 \, d\mu$$

given in (3.9) cancels out, yielding a fourth-order bias in  $\Delta t$  for BAOAB averages of  $g$ .

### 3.1.6 Effect of the friction parameter

All experiments shown above used a value of  $\gamma = 1$  for the friction parameter. In this final experiment, we examine the effect of changing  $\gamma$ . We show the marginal kinetic distributions for

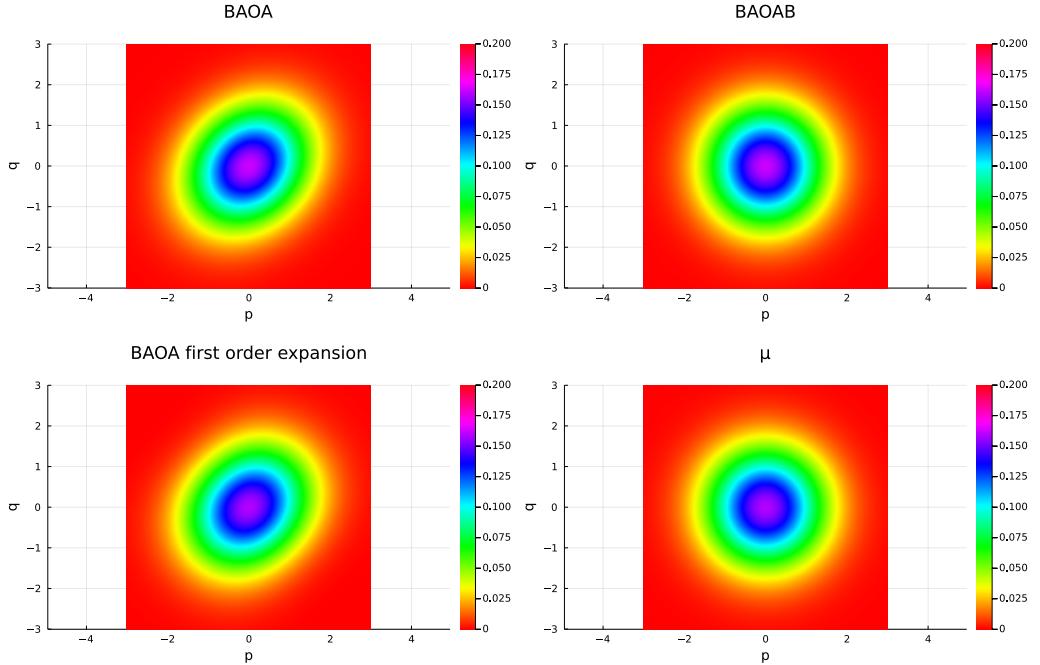


Figure 3.7: Joint distributions for the quadratic potential,  $\Delta t = 0.4$ .

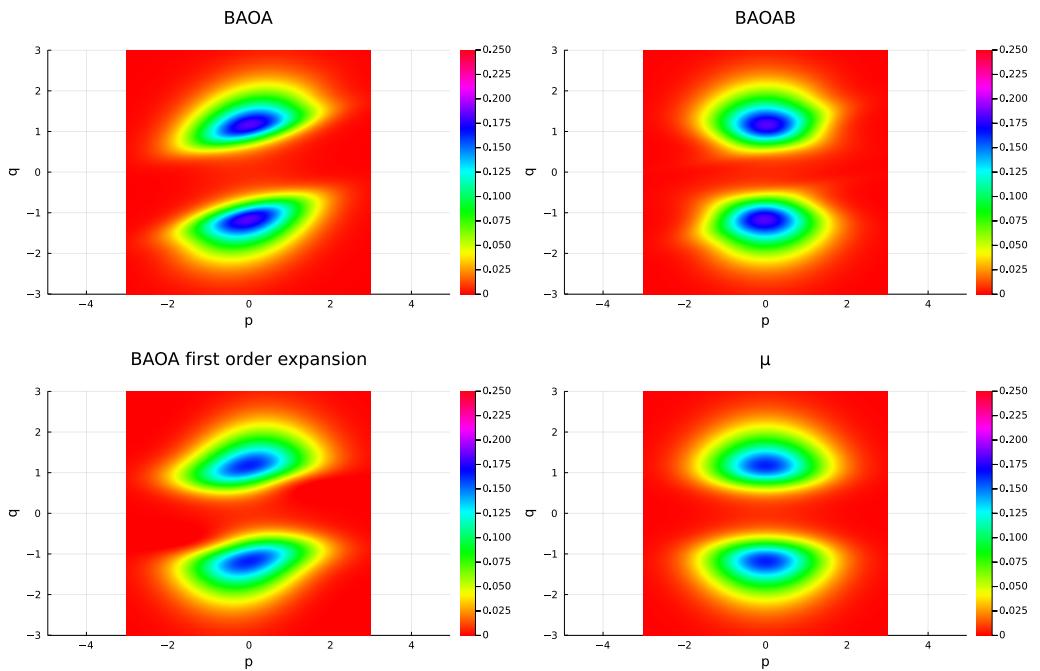


Figure 3.8: Joint distributions for the double well potential,  $\Delta t = 0.4$ .

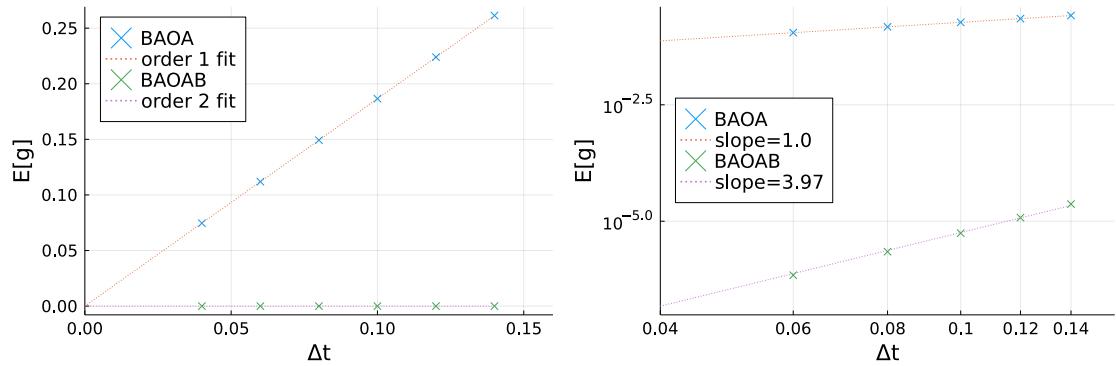


Figure 3.9: Averages of  $g$  for the double well potential.

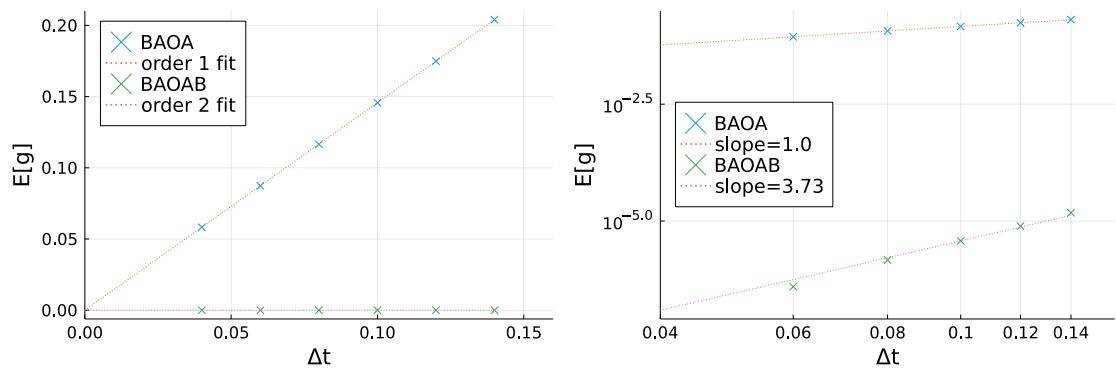


Figure 3.10: Averages of  $g$  for the tilted double well potential.

three values of  $\gamma \in \{0.1, 1, 10\}$ . The results show that there is no visually discernable effect of the parameter  $\gamma$ : all  $\kappa_{\Delta t, P}$ s are superposed close to the reference curve, and all  $\kappa_{\Delta t, Q}$ s are superposed above. This suggest that most of the error on  $\kappa_{\Delta t, Q}$  arises from the additional term

$$-\frac{\Delta t^2}{8} \int_{\mathcal{E}} \varphi(A + B) g d\mu,$$

which is the dominant discrepancy term in (3.26), and which is independent of  $\gamma$ . This is the fact we observed numerically on figures 3.5 and 3.4. See figure 3.11.

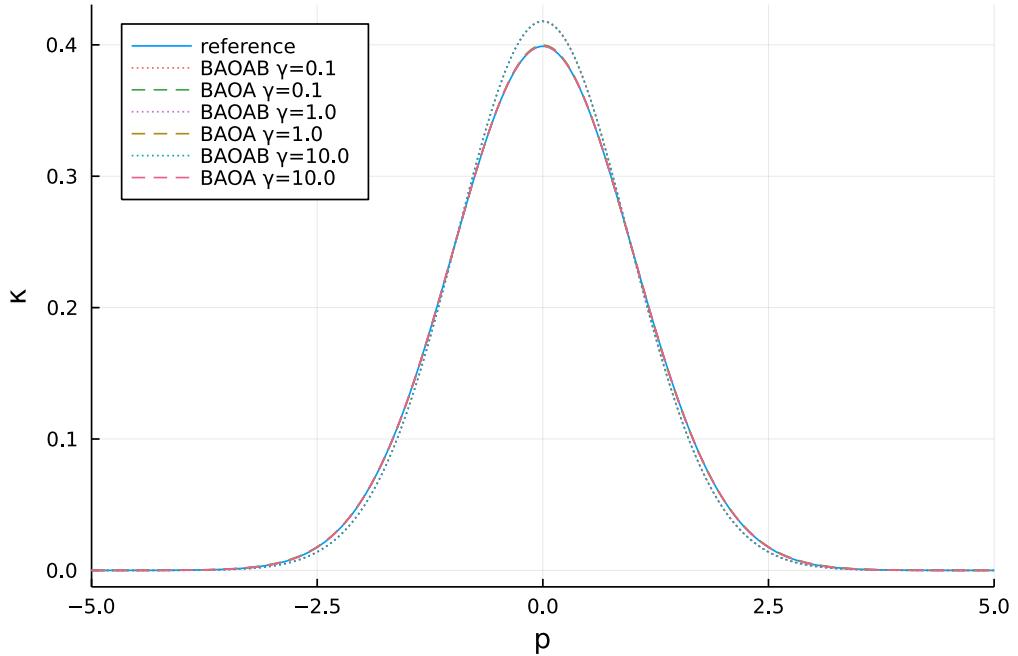


Figure 3.11: Kinetic marginal distributions for  $\Delta t = 0.3$  on the double well potential.

# Chapter 4

## Non-equilibrium Molecular Dynamics

So far, we have only considered methods to sample static averages, which concern quantities at thermodynamic equilibrium. Such techniques yield information about the bulk macroscopic properties of the system, for which there is no discernable macroscopic evolution. We now turn to the next natural question, which is to consider systems in which there is such an evolution, which typically arises from a perturbation of the equilibrium dynamics, either by the introduction of a non-gradient forcing term, or a modification of the fluctuation-dissipation part for which the fluctuation-dissipation relation (2.25) is not verified. We will not consider the latter case, which is relevant for instance to the modelling of heat transport within an atomic system, to concentrate on the first case. In general, systems undergoing such a perturbation of the dynamics will reach a new steady state in which there is a net flux in some observable. Mathematically, this translates into the existence of a response observable  $R$  which has zero average with respect to the canonical measure  $\mu$ , but which has a positive average with respect to the perturbation steady-state. A natural question is that of the sensitivity of the system to the perturbation: one way to quantify this is to modulate the strength of the perturbation by a positive real parameter  $\eta > 0$ , and, assuming that the average response is asymptotically linear as  $\eta \rightarrow 0$ , to compute the linear coefficient linking  $\eta$  and the average response. This quantity is called a *transport coefficient*, and we will dedicate the next two chapters to different methods of computing them using molecular simulation.

### 4.1 Non-equilibrium molecular dynamics

#### 4.1.1 General framework

We first consider the most natural method, which is to directly apply a forcing term which is not the gradient of a periodic function. The general framework is that of a classical Langevin dynamics

$$(4.1) \quad \begin{cases} dq_t = M^{-1} p_t dt, \\ dp_t = -\nabla V(q_t) dt - \gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t + \eta F(q_t) dt, \end{cases}$$

perturbed by the configuration-dependent forcing term  $F$ , and where the strength of the perturbation is modulated by the parameter  $\eta > 0$ . Its generator is given by the operator

$$(4.2) \quad \mathcal{L}_{\gamma,F} = \mathcal{L}_\gamma + \eta F \cdot \nabla_p = \mathcal{L}_\gamma + \eta \tilde{\mathcal{L}}.$$

It is also possible to consider non-equilibrium generalized Langevin equations, which are analogous perturbations of (2.15). Given a response of interest  $R$ , the transport coefficient is given by the following definition:

$$(4.3) \quad \rho_{R,F} = \lim_{\eta \rightarrow 0} \frac{\mathbb{E}[R]}{\eta},$$

where  $\mathbb{E}_\eta$  denotes the expectation with respect to the steady-state probability distribution. If the context makes  $R$  clear from the knowledge of  $F$ , we will drop it from the notation and simply write  $\rho_F$  for the transport coefficient. For this definition to make sense, one has to show that the steady-state with respect to which we take the expectation is well-defined. Since the steady-state is defined by the fact that it is invariant with respect to the dynamics (4.1), this translates into the fact that it is the solution to a stationary Fokker-Planck equation, in the sense of distributions:

$$(4.4) \quad (\mathcal{L}_\gamma + \eta \tilde{\mathcal{L}})^\dagger \psi_\eta = 0,$$

where  $\psi_\eta$  is the steady state measure. Using analytic properties of  $\mathcal{L}_{\gamma,F}$  (such as hypoellipticity), one can hope to infer regularity results of its solutions, such as the existence of a smooth density. Existence and unicity can in principle be inferred following Klieman's method [14]. The steady-state measure  $\mathbb{E}_\eta$  is a high-dimensional measure on phase space for which a closed form is generally unavailable. Thus, as usual, one has to resort to ergodic averages under the dynamics (4.1) to compute ensemble averages. This poses another theoretical difficulty, that of showing that the steady-state measure is ergodic.

In fact, for our purposes,  $F$  will always be  $L\mathbb{T}^{dN}$ -periodic, and so we can invoke a special case of Proposition 1 in [11] with time-constant forcings, provided  $V$  and  $F$  are smooth.

**Theorem 3** (Existence of a unique ergodic measure with smooth density). *Let  $\eta_* > 0$ . For all  $\eta \in [-\eta_*, \eta_*]$ , the Fokker-Planck equation (4.4) admits a unique solution with a  $C^\infty$  density  $\psi_\eta$ . Additionally, the evolution semigroup decays exponentially in a Lyapunov sense: for all  $n \geq 1$ , there exist  $C_n, \lambda_n > 0$  such that*

$$(4.5) \quad \|e^{t\mathcal{L}_{\gamma,F}} \varphi - \mathbb{E}_\eta[\varphi]\|_{L^\infty(\mathcal{E})} \leq C_n e^{-\lambda_n t} \|\varphi\|_{L^\infty_{\mathcal{K}_n}},$$

where we define the Lyapunov weight functions

$$\mathcal{K}_n(q, p) = 1 + |p|^{2n},$$

and the corresponding weighted  $L^\infty$  spaces by the norm

$$\|f\|_{L^\infty_{\mathcal{K}_n}} = \left\| \frac{f}{\mathcal{K}_n} \right\|_{L^\infty(\mathcal{E})}.$$

Because of the continuous injections  $L^\infty \subset L^\infty_{\mathcal{K}_n}$ , this result implies in particular that the operator  $(-\mathcal{L}_{\gamma,F})^{-1}$  is well-defined on

$$L^\infty_{\mathcal{K}_n,0} = \{f \in L^\infty_{\mathcal{K}_n} \mid \mathbb{E}_\eta[f] = 0\},$$

with an inverse given by the formula (2.36). Proposition 2 from the same papers also gives the almost-sure convergence of ergodic averages.

**Remark 12** (Size of the linear regime). *As  $\eta \rightarrow 0$ , it is reasonable to expect that the statistical error in ergodic averages for  $\mathbb{E}_\eta[R]$  arise at dominant order in  $\eta$  from the asymptotic variance at equilibrium  $\sigma_R^2$  (2.37). In particular the finite difference estimator for  $\rho_{R,F}$*

$$(4.6) \quad \frac{\mathbb{E}_\eta[R]}{\eta}$$

has, for a fixed simulation time, a variance which scales like  $\frac{1}{\eta^2}$  as  $\eta \rightarrow 0$ . To achieve an acceptable statistical error at minimal cost, one should thus aim to take  $\eta$  as large as possible. On the other hand, if the perturbation is too large, then non-linear effects on  $R$  will be observed and (4.6) will give a poor estimation of the transport coefficient. Thus, if one manages to devise another forcing  $F$  which does not change the steady-state, but which extends the range of linearity of the response, it may be very beneficial to do so from a computational standpoint. These kinds of methods, sometimes called synthetic forcings, are an area of ongoing research. (ref...)

#### 4.1.2 Numerical implementation

To integrate the perturbed dynamics, we again rely on a splitting strategy. We split the dynamics into elementary evolutions, whose respective evolution semigroups are given analogously to (2.39). The only difference is in the additional  $\eta F$  term, which we incorporate into the  $B$  step, yielding a  $B_{\eta F}$  step, which corresponds to the evolution semigroup

$$(4.7) \quad e^{tB_{\eta F}} \varphi(q, p) = \varphi(q, p - t(\nabla V(q) - \eta F(q))).$$

Note it would also have been possible to incorporate the forcing term into the Ornstein-Uhlenbeck part of the dynamics, which then corresponds to an Ornstein-Uhlenbeck process with constant drift. However, we will always choose the method described above. We will use the same terminology for these methods, referring to the scheme with evolution operator

$$(4.8) \quad e^{\frac{\Delta t}{2}B_{\eta F}} e^{\frac{\Delta t}{A}} e^{\Delta t \gamma C} e^{\frac{\Delta t}{2}A} e^{\frac{\Delta t}{2}B_{\eta F}}$$

as the BAOAB scheme, for example. Numerical estimates for  $\mathbb{E}_\eta[R]$  are then obtained through

$$(4.9) \quad \widehat{R}_{\eta, N_{\text{iter}}} = \frac{1}{N_{\text{iter}}} \sum_{n=0}^{N_{\text{iter}}-1} R(q^n, p^n),$$

where  $(q^n, p^n)_{n \geq 0}$  denote the numerical trajectory under the Markov chain obtained for a chosen splitting of the non-equilibrium dynamics (4.1). To estimate the transport coefficient, we fix different forcing intensities

$$0 < \eta_1 < \dots < \eta_k, \quad \eta := (\eta_i)_{1 \leq i \leq k}$$

all in the linear response regime, and given corresponding estimators

$$\widehat{R} := (\widehat{R}_i)_{1 \leq i \leq k}$$

of the form (4.9), we estimate  $\rho_F$  by a least squares linear fit

$$(4.10) \quad \widehat{\rho}_F = |\eta|^{-2} \eta \cdot \widehat{R} = \underset{\rho \in \mathbb{R}}{\operatorname{argmin}} \left| \rho \eta - \widehat{R} \right|^2.$$

Note the case  $k = 1$  recuperates the finite difference estimator (4.6). Note that the variance of this estimator is likely to be much smaller than for the latter estimator. In fact, provided  $|\eta|^2 \geq 1$ , the pitfall of the NEMD technique mentioned in Remark 12 is completely avoided. Indeed, assume that each  $\widehat{R}_i$  is an ergodic estimator simulated with  $N_{i,\text{iter}}$  timesteps, say

$$\widehat{R}_i = \frac{S_i}{\eta_i N_{i,\text{iter}}}.$$

We assume that the asymptotic regime is reached, thus we may assume that  $\widehat{R}_i$  is Gaussian with variance

$$\frac{\sigma_i^2}{\eta_i^2 N_{i,\text{iter}}},$$

where  $\sigma_i^2$  is the asymptotic variance associated with the ergodic mean  $S_i/N_{i,\text{iter}}$ . Thus,

$$(4.11) \quad \text{Var}(\hat{\rho}_F) = |\eta|^{-4} \sum_{i=1}^k \frac{\sigma_i^2}{\eta_i^2 N_{i,\text{iter}}}.$$

If we make the approximation  $\sigma_i^2 \approx \sigma_R^2$ , which is good at first order in  $\eta$ , and assume  $N_{i,\text{iter}} \geq N_{\text{iter}}$ , we can estimate an upper bound on the variance,

$$\text{Var}(\hat{\rho}_F) \lesssim \frac{\sigma_R^2}{|\eta|^2 N_{\text{iter}}}.$$

### 4.1.3 Mobility

As a first and simplest example, we introduce a method for computing the mobility. In this case the forcing is simply a constant vector, and the response is velocity in the direction  $F$ , which we may think of as the particle flux through  $F$ 's orthogonal hyperplane. The transport coefficient corresponding to the forcing

$$(4.12) \quad F \in \mathbb{R}^{dN} \quad R(q, p) = F \cdot M^{-1}p$$

is called the mobility. Let us assume once and for all that  $|F|^2 = 1$ . For practical computations, we will be considering two cases:

- (i) **Single drift:** this corresponds to a perturbation where the force acts on a single component of the momentum, which we can assume by indistinguishability of the particles to be the x component of the first particle,

$$F_S = (1, 0, \dots)^T \in \mathbb{R}^{dN}.$$

- (ii) **Color drift:** this corresponds to a perturbation in which we the force acts on half of the particles in one direction, and on half of the particles in the opposite direction. By isotropy, we may assume this direction is the x direction,

$$F_C = \frac{1}{\sqrt{N}} \underbrace{(1, 0, \dots)}_{d \text{ components}}, -1, 0, \dots, 1, 0 \dots)^T \in \mathbb{R}^{dN}.$$

The color drift method derive its name from the fact that it divides the particles into two categories: the particles with a positive color charge (corresponding to those whose x-component of momentum see a positive force contribution from  $F$ ), and those with a negative color charge. The equations of motion are analogous to a system of charged particles under a constant electric field, but where the charges do not interact between themselves. The idea is taken from [4], Chapter 6. Intuitively, we expect the color drift method to be more economical than the single drift method, since the latter should give essentially equilibrium dynamics for the majority of particles outside of the first particle's sphere of influence. On the other hand, the normalization condition implies that the forcing is more dilute in the color drift case, which may make the response more difficult to measure. In fact, numerical evidence presented below shows that the two effects roughly compensate one another.

In order for the two methods to be useful, we need to be able to relate them to a shared dynamical property of the considered system. Doing is more conveniently done upon reformulating the linear response in terms of integrated autocorrelation functions, a point we postpone to the next section. For now, let us show a few numerical results. In order to compare different methods, we fix a thermodynamical condition, which we give in reduced units below.

$$(4.13) \quad T = 1.25, \quad \rho = 0.6, \quad \gamma = 1.0, \quad N = 1000.$$

Let us also mention here that all simulations for mobility were performed using a BAOAB splitting, a timestep  $\Delta t = 10^{-3}$  and a linearly corrected cutoff at distance  $r_c = 2.5$ . In Figure 4.1, we plot the response profile as a function of the forcing intensity. It appears that the linear response regime is longer for the color drift method. Estimates of  $\rho_F$  seem to roughly agree, although it is dubious that the small discrepancy observed is due to a systematic effect rather than statistical fluctuations persisting after very long integration times (all simulations for  $\eta \leq 1$  ran for reduced physical times of over  $2.5 \times 10^5$ ).

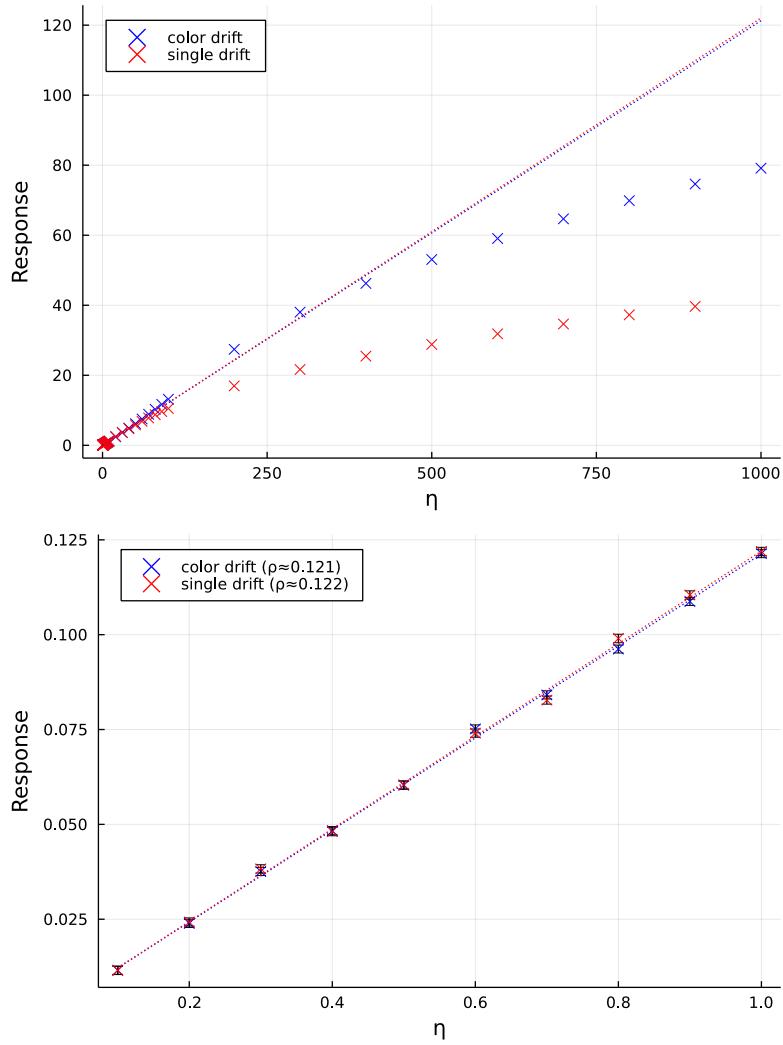


Figure 4.1: Average mobility response against forcing intensity for the Lennard-Jones system (4.13). Extrapolations of the linear response are plotted in dotted lines.

Estimating the transport coefficients using (4.9) and the variance through (4.11), we obtain

$$(4.14) \quad \rho_{F_S} = 0.12204 \pm 0.0009, \quad \rho_{F_C} = 0.12128 \pm 0.0009,$$

where  $\pm$  denotes one standard deviation.

#### 4.1.4 Shear viscosity

As a second example, we discuss the methods of [12] for computing shear-viscosity. We refer to this paper for thorough statements and proofs. Let us assume for notational simplicity that  $M = m\text{Id}$ . We consider a case of the dynamics (4.1) with a forcing which acts on the longitudinal (x) momenta, and is dependent on the transverse (y) positions. More precisely, we fix a reference periodic function  $F_y : L\mathbb{T} \rightarrow \mathbb{R}$ , and index the state as

$$(q_{i\alpha})_{\substack{1 \leq i \leq N \\ 1 \leq \alpha \leq d}}, \quad (p_{i\alpha})_{\substack{1 \leq i \leq N \\ 1 \leq \alpha \leq d}}.$$

The non-equilibrium dynamics is defined by the following expression for  $F$ :

$$(4.15) \quad \forall 1 \leq i \leq N, \forall 2 \leq \alpha \leq d, \quad F(q)_{i1} = f_y(q_{i2}), \quad F(q)_{i\alpha} = 0.$$

The idea is that by imposing a forcing *profile* in the x-direction, we observe a velocity profile in response, which depends on the *y*-coordinate.

**Remark 13** (Anisotropic friction). *In fact the precise model considered in [12] also imposes a separate friction coefficient  $\gamma_x$  for the longitudinal fluctuation-dissipation part of the dynamics. This can be interpreted as a simple case of a generalized Langevin dynamics where the friction coefficient  $\gamma$  is an anisotropic diagonal matrix, which does not pose any difficulties from the theoretical point of view. We therefore restrict our attention to the case where  $\gamma$  is scalar.*

More precisely, we define

$$(4.16) \quad u_x(Y) := \lim_{\epsilon \rightarrow 0} \lim_{\eta \rightarrow 0} \frac{L_y \mathbb{E}_\eta \left[ \sum_{i=1}^N p_{i1} \chi_\epsilon(q_{i2} - Y) \right]}{\eta m N}$$

to be the linear response profile in the longitudinal velocity, where  $(\chi_\epsilon)_{\epsilon > 0}$  is an approximation to the identity (that is a sequence of smooth compactly supported test functions which converge to the Dirac delta function in the sense of distributions). In practice,  $u_x$  can be estimated from numerical trajectories by decomposing the domain  $\mathcal{D}$  in a finite number of transverse slices, and measuring the average longitudinal velocity in each of these slices. The shear viscosity  $\sigma$  is then defined by the differential equation

$$(4.17) \quad \sigma u''_x(Y) + \gamma \rho u_x(Y) = \rho f_y(Y),$$

where  $\rho = N/L^3$  is the particle density. In fact, the solutions to (4.17) are periodic, thus the magnitude of the linear response can be estimated through Fourier-like coefficients. This has the advantage of giving an estimation of the linear response in the general framework defined above, and avoiding the discretization error arising from the finite number of transverse slices. To this effect, we define the response observables as empirical Fourier coefficients:

$$(4.18) \quad R_k(q, p) = \frac{1}{N} \sum_{i=1}^N \frac{p_{i1}}{m} \exp \left( \frac{2ik\pi q_{i2}}{L} \right),$$

and the corresponding transport coefficients

$$(4.19) \quad \rho_{F,k} = \lim_{\eta \rightarrow 0} \frac{\mathbb{E}_\eta[R_k(q, p)]}{\eta}.$$

It is easy to show that for the sinusoidal forcing, the response velocity profile, solution to (4.17), is still sinusoidal, so that the Fourier response is a meaningful measure of the magnitude of the response in the velocity profile. In practice, it is sufficient to consider  $k = 1$ .

**Remark 14.** *To obtain better statistics, one may for the purpose of numerical simulations want to consider a periodic domain whose unit cell is still a cuboid, but with a side length in the longitudinal direction that is longer than in the transverse direction. In this case we simply replace  $L$  by  $L_y$  in the response observables, where  $L_y$  is the length of the unit cell in the transverse direction.*

Forcing	Sinusoidal	Piecewise constant	Piecewise linear
$F_1$	$i/2$	$2i/\pi$	$-4/\pi^2$

Figure 4.2: First Fourier coefficients for transverse forcing profiles.

The shear viscosity can then be computed from equation (4.17), through

$$(4.20) \quad \sigma = \rho \left( \frac{F_k}{\rho_{F,k}} - \gamma \right) \left( \frac{L}{2k\pi} \right)^2,$$

by taking Fourier coefficients in equation (4.17), where  $F_k$  is the counterpart Fourier coefficient for the forcing,

$$F_k = \frac{1}{L} \int_0^L f_y(s) \exp \left( \frac{2ik\pi s}{L} \right) ds,$$

and again only considering  $k = 1$  in practice. Numerically,  $\rho_{F,k}$  can be approximated by an estimator of the form (4.10). In our numerical experiments, we consider the same three forcing profiles as in [12], namely:

(i) A sinusoidal forcing profile

$$f_y(u) = \sin \left( \frac{2\pi u}{L} \right),$$

(ii) a piecewise constant forcing profile

$$f_y(u) = \mathbb{1}_{u < \frac{L}{2}} - \mathbb{1}_{u \geq \frac{L}{2}},$$

(iii) and a piecewise linear forcing profile

$$f_y(u) = \frac{4}{L} \left( u - \frac{L}{4} \right) \mathbb{1}_{u < \frac{L}{2}} - \frac{4}{L} \left( \frac{3L}{4} - u \right) \mathbb{1}_{u \geq \frac{L}{2}}.$$

These forcings have the advantage that the Fourier coefficients can be analytically computed. We record in Figure 4.2 the value of the first Fourier coefficient  $F_1$  for each of these forcings. As in the case of mobility, we fix a reference thermodynamic condition,

$$(4.21) \quad T = 0.8, \quad \rho = 0.7, \quad \gamma = 1.0, \quad N = 1000.$$

## 4.2 The Green-Kubo method

An alternative route to the perturbation method described above leverages a famous expression for the transport coefficient in terms of an integrated correlation function, or, in less prosaic language, in terms of the fluctuations at equilibrium of the response observable. This is the Green-Kubo method, which we describe in this section. We consider the invariant measure for the non-equilibrium dynamics (4.1). By Theorem 3, there exists a unique invariant measure with density  $\psi_\eta$ . For notational consistency, let us write  $\psi_0$  for the density of the equilibrium measure  $\mu$ . Then the following result holds

**Theorem 4** (Series expansion for the non-equilibrium steady state). *There exists  $r > 0$  such that for all  $0 < \eta < r$ ,*

$$(4.22) \quad \frac{\psi_\eta}{\psi_0} = \left( 1 + \eta (\tilde{\mathcal{L}} \Pi \mathcal{L}_\gamma^{-1} \Pi)^* \right)^{-1} \mathbb{1}_{\mathcal{E}} = \left( 1 + \sum_{k=1}^{\infty} (-\eta)^k \left[ (\tilde{\mathcal{L}} \Pi \mathcal{L}_\gamma^{-1} \Pi)^* \right]^k \right) \mathbb{1}_{\mathcal{E}},$$

where  $\tilde{\mathcal{L}}$  is defined by (4.2),  $\Pi$  is the equilibrium centering projector defined in (2.33), and the adjoint is taken in  $L^2(\mu)$ .

*Sketch of proof.* The second equality identifies the Neumann series on the right as the resolvent of  $-(\tilde{\mathcal{L}}\Pi\mathcal{L}_\gamma^{-1}\Pi)^*$ , provided  $\eta$  is taken small enough. In fact, from the spectral theory of bounded operators,  $r$  can be determined as the spectral radius of  $(\tilde{\mathcal{L}}\Pi\mathcal{L}_\gamma^{-1}\Pi)^*$  in the space of bounded operators  $\mathcal{B}(L_0^2(\mu))$ . The core of the argument lies in making the ansatz

$$(4.23) \quad \psi_\eta = \psi_0(1 + \eta f_1 + \eta^2 f_2 + \dots).$$

The stationary Fokker-Planck equation (4.4) writes

$$(\mathcal{L}_\gamma + \eta \tilde{\mathcal{L}})^\dagger \psi_0(1 + \eta f_1 + \eta^2 f_2 + \dots) = (\mathcal{L}_\gamma + \eta \tilde{\mathcal{L}})^*(1 + \eta f_1 + \eta^2 f_2 + \dots) = 0.$$

By formally separating terms by degree in  $\eta$ , we obtain

$$\begin{aligned} \mathcal{L}_\gamma^* \mathbb{1}_{\mathcal{E}} &= 0, \\ \tilde{\mathcal{L}}^* \mathbb{1}_{\mathcal{E}} + \mathcal{L}_\gamma^* f_1 &= 0, \\ \tilde{\mathcal{L}}^* f_1 + \mathcal{L}_\gamma^* f_2 &= 0, \end{aligned}$$

and so on. Note that the first equality is the equilibrium Fokker-Planck equation. Thus, by induction, again formally, we obtain.

$$\begin{aligned} f_1 &= (-\mathcal{L}_\gamma^*)^{-1} \tilde{\mathcal{L}}^* \mathbb{1}_{\mathcal{E}}, \\ f_2 &= (-\mathcal{L}_\gamma^*)^{-1} \tilde{\mathcal{L}}^* f_1, \\ &\dots \\ f_n &= \left[ (-\mathcal{L}_\gamma^*)^{-1} \tilde{\mathcal{L}}^* \right]^n \mathbb{1}_{\mathcal{E}}, \end{aligned}$$

whence the formal proof follows, by observing that we can write

$$(-\mathcal{L}_\gamma^*)^{-1} \tilde{\mathcal{L}}^* = -(\tilde{\mathcal{L}}\Pi\mathcal{L}_\gamma^{-1}\Pi)^*.$$

For a rigorous proof, several points should be made precise:

- (i) The convergence of the series (4.22) for sufficiently small  $\eta$ , which can be obtained by showing the boundedness of the operator

$$(\tilde{\mathcal{L}}\Pi\mathcal{L}_\gamma^{-1}\Pi)^*$$

on  $L_0(\mu)$ .

- (ii) The fact that

$$\psi_0 \left( 1 + \eta (\tilde{\mathcal{L}}\Pi\mathcal{L}_\gamma^{-1}\Pi)^* \right)^{-1} \mathbb{1}_{\mathcal{E}}$$

is indeed a solution to the stationary Fokker-Planck equation.

- (iii) The fact that it is indeed a probability density.

One can then conclude by unicity of the steady-state measure.  $\square$

Strikingly, we can read the linear response directly from the first term of the series expansion.

**Corollary 4** (Green-Kubo formula). *Let  $R$  be any response observable such that  $\mathbb{E}_\mu[R] = 0$ , and  $R \in L_{\mathcal{K}_n}^\infty$  for some  $n$ . Then we have the following formula for the linear response*

$$(4.24) \quad \lim_{\eta \rightarrow 0} \frac{\mathbb{E}_\eta[R]}{\eta} = \int_0^\infty \mathbb{E}_\mu[R(q_t, p_t) S(q_0, p_0)] dt,$$

where the expectation on the right hand side is with respect to all equilibrium dynamics trajectories with canonical initial distribution, and  $S$  is defined by

$$(4.25) \quad S = \tilde{\mathcal{L}}^* \mathbb{1}_{\mathcal{E}} = \beta F(q) \cdot M^{-1} p,$$

and is called the conjugate response function. The latter expression follows from a simple integration by parts.

*Proof.* By Theorem 4, we can write

$$\begin{aligned} \lim_{\eta \rightarrow 0} \frac{\mathbb{E}_\eta[R]}{\eta} &= \int_{\mathcal{E}} R(q, p) (-\mathcal{L}_\gamma^*)^{-1} \tilde{\mathcal{L}}^* \mathbb{1}_{\mathcal{E}} \psi_0(q, p) dq dp \\ &= \int_{\mathcal{E}} (-\mathcal{L}_\gamma)^{-1} R(q, p) S(q, p) \psi_0(q, p) dq dp \\ &= \int_{\mathcal{E}} \int_0^\infty \mathbb{E}^{(q,p)}[R(q_t, p_t) S(q_0, p_0)] \psi_0(q, p) dt dq dp \\ &= \int_0^\infty \mathbb{E}_\mu[R(q_t, p_t) S(q_0, p_0)] dt, \end{aligned}$$

where we rely on an expression like (2.36) for  $(-\mathcal{L}_\gamma)^{-1}$ .  $\square$

The Green-Kubo formula has a great advantage, in that it allows us to estimate the transport coefficients for as many different perturbations and response observables as we want **from a single equilibrium trajectory**. Indeed, one only needs to compute the corresponding integrated correlation functions (4.24).

#### 4.2.1 Numerical implementation

We now describe a method to compute correlation functions necessary to the Green-Kubo method from a single long numerical trajectory. This method is described by Tuckerman in [18], section 13.4.2 for Hamiltonian trajectories. In fact we consider a slight extension in which we let  $R$  and  $S$  to be vector-valued observables. For notational simplicity, we assume that  $R$  and  $S$  are component-wise centered. The quantities we want to estimate are

$$(4.26) \quad C(t) = \mathbb{E}_\mu[R(q_t, p_t)^\top S(q_0, p_0)]$$

we let  $(q^n, p^n)_{n \geq 0}$  be a numerical trajectory, which we see as the random iterates of the Markov chain associated with a numerical scheme, with a regular timestep  $\Delta t > 0$ . We further assume that the trajectory is stationary for this Markov chain, which is a realistic assumption if we take care of equilibrating the system using our numerical scheme before recording states. By stationarity, the equality in law

$$(q^n, p^n, q^0, p^0) \xrightarrow{\text{law}} (q^{n+k}, p^{n+k}, q^k, p^k)$$

for all  $k$ . Hence we define the following estimator for  $C(n\Delta t)$ , for  $0 \leq n \leq N_{\text{iter}}$ :

$$(4.27) \quad \widehat{C}_{N_{\text{iter}}}(n\Delta t) = \frac{1}{N_{\text{iter}} - n + 1} \sum_{k=0}^{N_{\text{iter}}-n} R(q^{n+k}, p^{n+k})^\top S(q^k, p^k).$$

Note the quality of these estimators degrades with  $n$ . Thus in practice, we fix  $N_{\text{corr}} \ll N_{\text{iter}}$ , and compute these estimators for  $0 \leq n \leq N_{\text{corr}}$ . For implementation details, we refer to the Appendix and links therein to relevant Julia code.

Using these estimators, we can estimate the transport coefficient through the Green-Kubo formula (4.24). The simplest way is to use a naive Riemann sum, or rectangle rule:

$$\rho_F \approx \Delta t \sum_{k=0}^{N_{\text{corr}}} \widehat{C}_{N_{\text{iter}}}(k\Delta t).$$

In fact, analysis shows that using a trapezoidal rule reduces the error. However, these procedures introduce a truncation in time of the integral in (4.24). Another approach consists of extrapolating the behavior of  $\widehat{C}_{N_{\text{iter}}}$  by fitting a parametric model

$$(4.28) \quad C_\theta(t) = \left( \sum_{k=1}^m a_k e^{-\lambda_k t} \cos(f_k t + \omega_k) \right),$$

where

$$\theta = (\lambda_k, a_k, f_k, \omega_k) \in (\mathbb{R}_{>0} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R})^m$$

is the parameter. The form of this model is justified empirically, although a formal argument based on a diagonalisation of the evolution semigroup can be made, and using the conjugate symmetry property  $\mathcal{L}_\gamma \bar{\psi} = \overline{\mathcal{L}_\gamma \psi}$  for  $\psi$  a complex-valued observable. Rigorously justifying and quantifying the accuracy of this model requires fine knowledge of the spectrum of  $\mathcal{L}_\gamma$ . At any rate, we can then fit the model in a least-squares sense,

$$(4.29) \quad \theta^* = \operatorname{argmin}_{\theta \in (\mathbb{R}_{>0} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R})^m} \sum_{n=0}^{N_{\text{corr}}} \left| C_\theta(n\Delta t) - \hat{C}_{N_{\text{iter}}}(n\Delta t) \right|^2,$$

using gradient descent, or the Gauss-Newton method, and deduce an estimator for the transport coefficient,

$$(4.30) \quad \hat{\rho}_{F, N_{\text{iter}}}^{\text{GK}} = \int_0^\infty C_{\theta^*}(t) dt,$$

leveraging the simple identity

$$(4.31) \quad a \int_0^\infty e^{-\lambda t} \cos(ft + \omega) dt = a \frac{\lambda \cos \omega - f \sin \omega}{\lambda^2 + f^2},$$

we get a closed form for the estimator,

$$\hat{\rho}_{F, N_{\text{iter}}}^{\text{GK}} = \sum_{k=1}^m \frac{a_k \lambda_k}{\lambda_k^2 + f_k^2}.$$

#### 4.2.2 Mobility

The Green-Kubo formula asserts that

$$(4.32) \quad \lim_{\eta \rightarrow 0} \frac{\mathbb{E}_\eta[F \cdot M^{-1} p]}{\eta} = \beta \int_0^\infty \mathbb{E}_\mu[(F \cdot M^{-1} p_t) (F \cdot M^{-1} p_0)] dt.$$

Using this expression, in the case of a mass-homogeneous system where the potential  $V$  is of pair interaction form (1.3), we can relate the transport coefficients for different forcings. As a useful example, we compute an equation relating the transport coefficients for the single drift and color drift forcings. The argument is taken from unpublished notes by Julien Roussel.

**Example 6** (Relating linear responses). *We assume  $M = m\text{Id}$ . Let us define, for  $1 \leq i, j \leq N$ ,*

$$c_{ij} = \frac{\beta}{m^2} \int_0^\infty \mathbb{E}_\mu[p_{i1,t} p_{j1,0}] dt.$$

*By the form of the potential (Newton's third law), for all  $q$ ,*

$$\sum_{i=1}^N \frac{\partial}{\partial q_{i1}} V(q) = 0.$$

*This implies upon summing over  $i$  the longitudinal  $p$ -components of the SDE (2.14) and integrating*

$$\sum_{i=1}^N p_{i1,t} = \sum_{i=1}^N \left[ p_{i1,0} + \int_0^t \left( -\frac{\partial}{\partial q_{i1}} V(q_s) - \frac{\gamma}{m} p_{i1,s} ds + \sqrt{\frac{2\gamma}{\beta}} dW_{i1,s} \right) \right] = \sum_{i=1}^N \left[ p_{i1,0} - \frac{\gamma}{m} \int_0^t p_{i1,s} ds + \sqrt{\frac{2\gamma}{\beta}} W_{i1,t} \right].$$

*Multiply by  $p_{11,0}$  and take the expectation with respect with the canonical initial distribution, the Brownian terms vanish, and we get*

$$(4.33) \quad \mathbb{E}_\mu \left[ \left( \sum_{i=1}^N p_{i1,t} \right) p_{11,0} \right] = \sum_{i=1}^N \left[ \mathbb{E}_\mu [p_{i1,0} p_{11,0}] - \frac{\gamma}{m} \int_0^t \mathbb{E}_\mu [p_{i1,s} p_{11,0}] ds \right].$$

By the decay properties of the evolution semigroup, the left hand side converges to 0 as  $t \rightarrow \infty$ , while the integral is well-defined. Since  $p_0$  has diagonal covariance with respect to  $\mu$ , we get

$$(4.34) \quad \sum_{i=1}^N \frac{\gamma}{m} \int_0^\infty \mathbb{E}_\mu [p_{i1,s} p_{11,0}] ds = \mathbb{E}_\mu [p_{11,0}^2] = \frac{m}{\beta}.$$

Equivalently,

$$\sum_{i=1}^N c_{i1} = \frac{1}{\gamma}.$$

Using the indistinguishability property

$$c_{ii} = c_{11}, \quad c_{ij} = c_{12}$$

for all  $i \neq j$ , we can rewrite this identity as

$$c_{12} = \frac{1}{N-1} \left( \frac{1}{\gamma} - c_{11} \right).$$

Using this computation, we can relate the linear responses of the single drift and the color drift. By the Green-Kubo formula, the transport coefficient for the single drift is given by  $c_{11}$ ,

$$\rho_{F_S} = c_{11}.$$

For the color drift, we expand, by the Green-Kubo formula,

$$\rho_{F_C} = \frac{\beta}{m^2} \int_0^\infty \mathbb{E}_\mu [(F_C \cdot p_t) (F_C \cdot p_0)] dt = \frac{1}{N} \left( \sum_{i=1}^n c_{ii} + \sum_{i \neq j} (-1)^{i+j} c_{ij} \right)$$

By indistinguishability, and using

$$\sum_{1 \leq i \neq j \leq N} (-1)^{i+j} = -2 \left\lfloor \frac{N}{2} \right\rfloor,$$

we get

$$(4.35) \quad \rho_{F_C} = c_{11} - \frac{2 \lfloor N/2 \rfloor}{N(N-1)} \left( \frac{1}{\gamma} - c_{11} \right).$$

Note that in our numerical example, we expect  $\rho_{F_C} \approx \rho_{F_S} - \frac{0.88}{N}$ , which is consistent with what we observe numerically, although because of persisting statistical uncertainty, much longer trajectories (or smaller systems) have to be considered to confirm this relation numerically. Let us also, the discrepancy vanishes in the thermodynamic limit, at rate  $O(N^{-1})$ . Since the system we consider is isotropic, we also have

$$c_{11} = \frac{\beta}{m^2} \int_0^\infty \mathbb{E}_\mu [p_{i\alpha,s} p_{i\alpha,0}] ds$$

for any  $1 \leq i \leq N$  and  $1 \leq \alpha \leq d$ . Thus, the mobility can be computed through an expression of the form (4.26):

$$(4.36) \quad \rho_{F_S} = c_{11} = \frac{\beta}{m^2 d N} \int_0^\infty \mathbb{E}_\mu [p_s^\top p_0] ds,$$

which is the Green-Kubo relation we use in our numerical experiments. In Figure 4.3, we plot the velocity autocorrelation function used in the estimator (4.36), and its integral. We used the same thermodynamics conditions as for the NEMD method (4.13). In Figure 4.4, we show the fitted

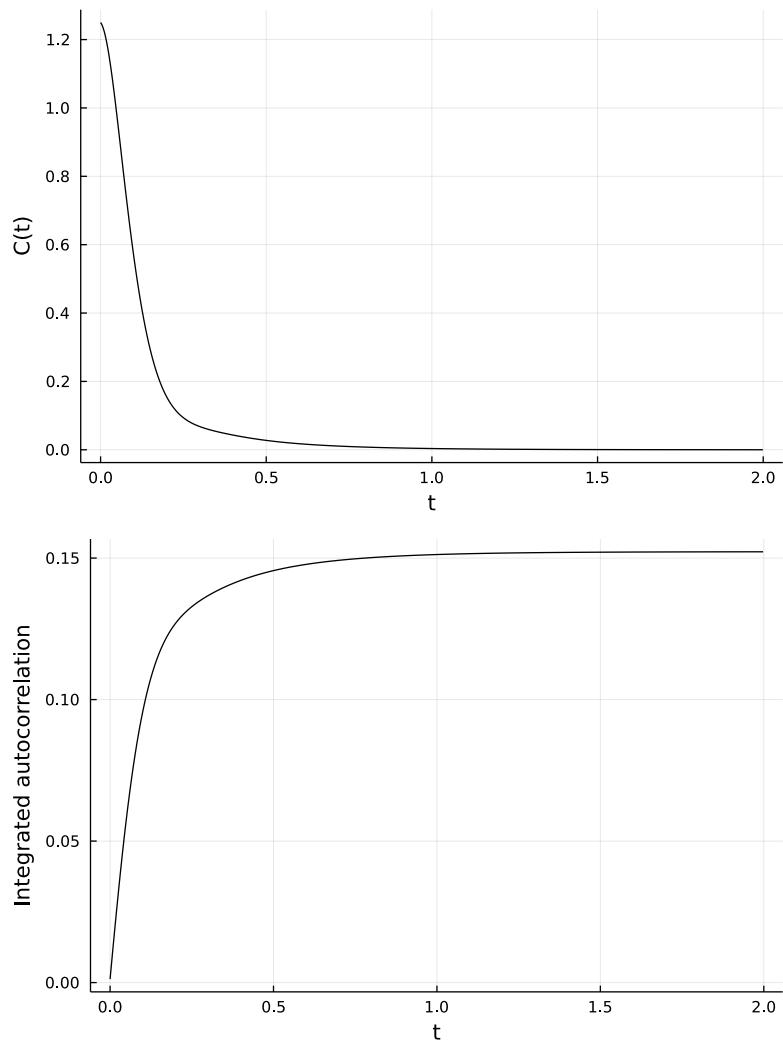


Figure 4.3: Velocity autocorrelation function (top) and its integral (bottom) for the Lennard-Jones system (4.13).

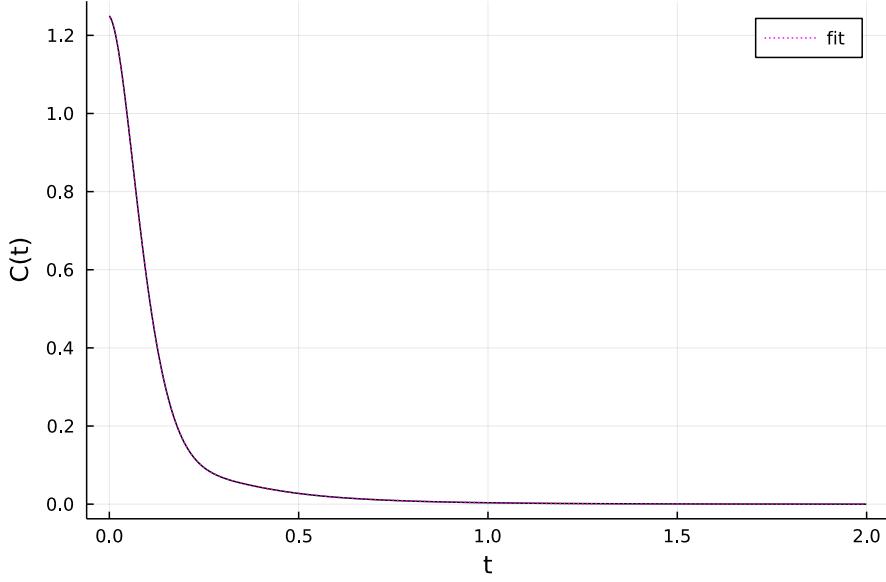


Figure 4.4: Fitted parametric model (4.28) (in dotted line) superimposed on autocorrelation function from Figure 4.3 (in black) for the Lennard-Jones system (4.13).

non-linear least squares model (4.28) for the autocorrelation function, using  $m = 4$  modes, trained using the Gauss-Newton algorithm. The estimated correlation function was obtained from a single numerical trajectory with a physical reduced time of  $t_{\text{fin}} = 1.628 \times 10^5$ . The corresponding estimated values for the mobility are 0.1218 based on a trapezoid quadrature of the truncated autocorrelation function, and 0.1217 for the analytic integral of the parametric model.

We conclude this section on mobility computations by the discussion of a final method, which yields a nice physical interpretation for the mobility.

**Remark 15** (Einstein relation for mobility). *Recalling the computations in Section 2.2.5, we can write, for  $R$  a centered response observable,*

$$(4.37) \quad \mathbb{E}_\mu \left[ \left( \frac{1}{\sqrt{T}} \int_0^T R(q_t, p_t) dt \right)^2 \right] = 2 \int_0^T \mathbb{E}_\mu [R(q_t, p_t) R(q_0, p_0)] \left( 1 - \frac{t}{T} \right) dt.$$

*Assuming an exponential decay of the evolution semigroup, we get*

$$(4.38) \quad \int_0^\infty \mathbb{E}_\mu [R(q_t, p_t) R(q_0, p_0)] dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \mathbb{E}_\mu \left[ \left( \int_0^T R(q_t, p_t) dt \right)^2 \right].$$

*This is the Einstein formula. In the case of mobility  $R(q, p) = F \cdot M^{-1}p$ , the right-hand side rewrites*

$$(4.39) \quad \lim_{T \rightarrow \infty} \frac{1}{2T} \mathbb{E}_\mu \left[ (F \cdot (Q_T - Q_0))^2 \right],$$

*where*

$$(4.40) \quad Q_t = Q_0 + \int_0^t M^{-1} p_s ds = \int_0^t dq_s$$

is the so called **self-diffusion** process, which formally satisfies the same SDE as  $q$ , but in the unperiodic domain  $\mathbb{R}^{dN}$ . Its trajectories corresponds to the unwrapped or unperiodicized trajectories of the coordinates, which makes them particularly easy to keep compute throughout a numerical simulation. This terminology is justified by a result (see Theorem 1 in [17]) asserting that the diffusively rescaled self-diffusion process

$$\sqrt{\varepsilon}(Q_{\varepsilon^{-1}t} - Q_0)$$

converges in law to a Brownian motion with diffusion matrix  $\mathfrak{D}$  characterized by

$$(4.41) \quad F \cdot \mathfrak{D} F = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_\mu \left[ (F \cdot (Q_T - Q_0))^2 \right].$$

In the case of an isotropic system, the diffusion matrix is characterized by a single number, the diffusion coefficient, which is given by the half normalized trace of the diffusion matrix.

$$(4.42) \quad D = \frac{1}{2dN} \text{Tr}(\mathfrak{D}) = \lim_{T \rightarrow \infty} \frac{1}{2dNT} \mathbb{E}_\mu \left[ |Q_T - Q_0|^2 \right]$$

In view of the Einstein relation (4.39) and of the Green-Kubo relation (4.32), the mobility is linked to the diffusion coefficient through

$$(4.43) \quad \rho_{F_S} = \beta D.$$

This yields a new family of estimators for the mobility, based on a computation of the self-diffusion coordinates expression (4.42). In Figure 4.5 we illustrate the strategy we used to compute the mobility using the Einstein formula, which we proceed to explain. We define the mean-squared deviation as the process

$$(4.44) \quad M_t = \frac{|Q_t - Q_0|^2}{dN}.$$

Fixing a number of independent realizations  $N_{\text{run}}$  and a number of simulation iterations  $N_{\text{iter}}$ , we consider, for each  $i = 1, \dots, N_{\text{run}}$ , a numerical trajectory of regularly sampled points from the mean-squared deviation,  $M^i = (M_i^k)_{0 \leq k \leq N_{\text{iter}}}$ . These correspond to points sampled at regular time intervals  $T_{\text{samp}} = N_{\text{samp}} \Delta t$ , where  $\Delta t$  is the simulation timestep, and computed from the numerical trajectory of the self-diffusion coordinates. We then obtain  $N_{\text{run}}$  estimators for  $D$  by a linear regression on the mean-squared deviation trajectories,

$$(4.45) \quad \hat{D}_{N_{\text{iter}}}^i = \frac{1}{2|\tau|^2} \tau \cdot M^i,$$

where

$$\tau = (kT_{\text{samp}})_{0 \leq k \leq N_{\text{iter}}}$$

is the vector of sampling times. An estimator for  $D$ , and thus for  $\rho_{F_S}$ , is obtained by averaging over the number of independent runs:

$$(4.46) \quad \hat{D}_{N_{\text{iter}}} = \frac{1}{N_{\text{run}}} \sum_{i=1}^{N_{\text{run}}} \hat{D}_{N_{\text{iter}}}^i.$$

Independent runs can be computed over a single trajectory, using the Markov property,

$$(Q_T - Q_0)_{T \geq 0} \sim (Q_{T+t} - Q_t)_{T \geq 0} \perp\!\!\!\perp (Q_s)_{0 \leq s \leq t} \forall T, t \geq 0.$$

This can be very simply exploited in a simulation, by resetting the self-diffusion coordinates to zero following each sample run. In practice, one has to take care to take  $N_{\text{samp}}$  large enough so that the mean-squared deviation behaves approximately linearly over the sample time range  $[0, T_{\text{samp}}]$ . In Figure 4.5, we illustrate the principle underlying this method. For our final computation, we used  $N_{\text{run}} = 545$ ,  $N_{\text{iter}} = 10^6$ ,  $\Delta t = 10^{-3}$ ,  $N_{\text{samp}} = 1$ . The estimated mobility is

$$(4.47) \quad \rho_{F_S} = 0.12157 \pm 0.00011.$$

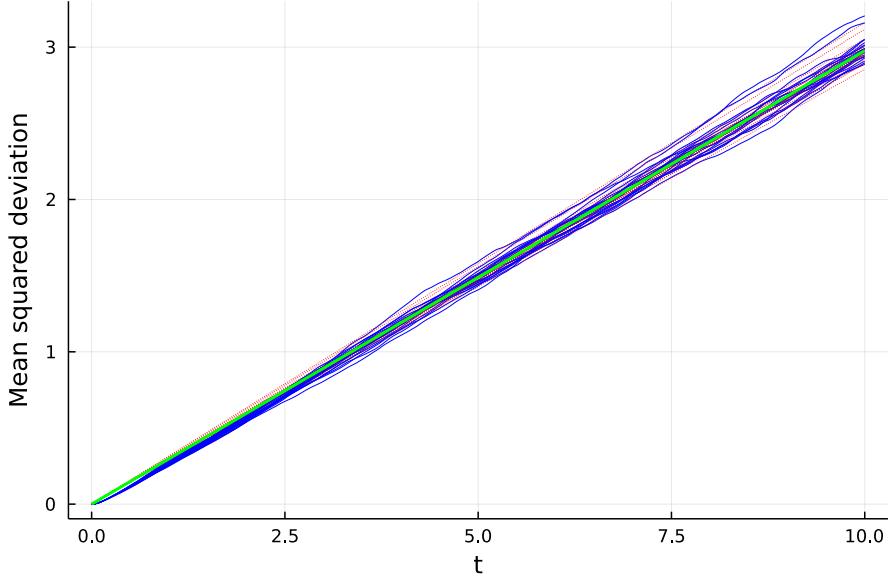


Figure 4.5: Twenty independent realizations of the mean squared-deviation (in blue) superimposed with their corresponding linear regression lines (in red dotted lines). The regression line corresponding to the corresponding estimator (4.46) is plotted in bright green.

### 4.2.3 Shear viscosity

We conclude by applying the Green-Kubo formula to shear viscosity computations. It asserts, with  $\rho_{F,k}$  given by (4.19),

$$(4.48) \quad \rho_{F,k} = \beta \int_0^\infty \mathbb{E}_\mu [R_k(q_t, p_t) (F(q_0) \cdot M^{-1} p_0)],$$

where  $R_k$  is given by (4.18). At this point, let us remark that we can, similar to the mobility case in (4.36), exploit the isotropy of the equilibrium measure to our statistical advantage. To do this let us define versions of the transverse forcing method for any pair of orthogonal canonical directions in  $\mathbb{R}^d$ . More precisely, we define, for any  $1 \leq i \neq j \leq d$ ,

$$R_{k,\alpha\beta}(q, p) = \frac{1}{N} \sum_{i=1}^N \frac{p_{i\alpha}}{m} \exp\left(\frac{2ik\pi q_{i\beta}}{L}\right),$$

$$F_{\alpha\beta}(q)_{i\alpha} = f_y(q_{i\beta}), \quad F_{\alpha\beta}(q)_{i\delta} = 0, \quad \forall 1 \leq i \leq N, \forall \delta \neq \beta.$$

Then, by isotropy, we have

$$(4.49) \quad \rho_{F,k} = \frac{\beta}{d(d-1)} \sum_{1 \leq \alpha \neq \beta \leq d} \int_0^\infty \mathbb{E}_\mu [R_{k,\alpha\beta}(q_t, p_t) (F_{\alpha\beta}(q_0) \cdot M^{-1} p_0)].$$

This can be rewritten as an expression of the form (4.26), with  $d(d-1)$ -dimensional observables. This is slightly theoretical since in practice  $d$  is always 2 or 3. When  $d = 3$ , the speedup in convergence is appreciable. We also take advantage of this opportunity to showcase the power of the Green-Kubo method by computing transport coefficients for all three forcing profiles using a single numerical trajectory. In Figure (4.6), we plot the correlation functions and corresponding integral for the three different forcing profiles. We note that the results are roughly consistent with the analytic computations of Figure 4.2, in the sense that the normalized transport coefficient

$$\frac{\rho_{F,k}}{F_k}$$

should be real and positive. The persistance of an imaginary component in the estimation of this normalized transport coefficient can be attributed to statistical error, and is useful as a direct empirical (albeit unrigorous) monitor for the convergence of the Green-Kubo estimator. Computations were run under the reference thermodynamic condition (4.21), for a physical simulation time of  $t_{\text{fin}} = 2.46 \times 10^4$ , with a timestep  $\Delta t = 10^{-3}$  and a number of decorrelation steps  $N_{\text{corr}} = 10^4$ . The corresponding estimates for the normalized transport coefficients are respectively 0.644, 0.645 and 0.655 for the sinusoidal, piecewise constant and piecewise linear forcing profiles.

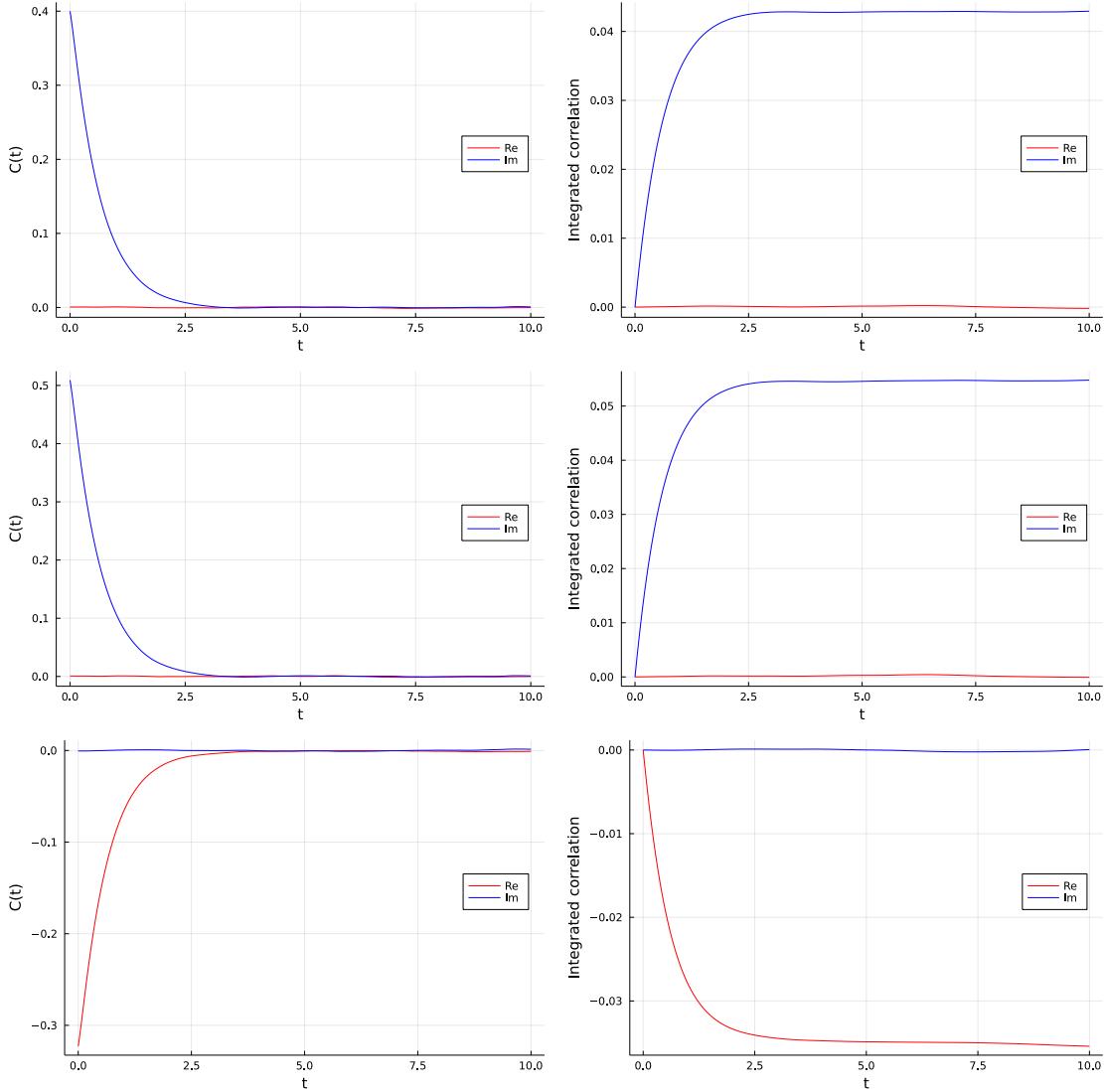


Figure 4.6: Correlation functions (left) and corresponding integrated correlation functions (right) for the shear viscosity dynamics. Top row: sinusoidal forcing, middle row: piecewise constant forcing, bottom row: piecewise linear forcing. Real parts are plotted in red and imaginary parts are plotted in blue. Integrals were obtained with a trapezoid rule.

# Chapter 5

## Norton dynamics

### 5.1 Introduction

We have considered so far two general methods for the computation of transport coefficients. The Green-Kubo method, which relies on analysis of autocorrelations in the fluctuation of zero-average equilibrium quantities, and the NEMD method, which relies in measuring the ratio in the average of a given equilibrium-centered response observable under a driven steady-state and the magnitude of the driving force. In the case of mobility, we apply a small constant force, and measure the resulting particle flux in the direction of the perturbation. We can thus think of the mobility  $\rho_F$  as measuring how *responsive* the flux is to the forcing on the system. A natural question to ask is whether it is possible to measure the dual quantity, that is, how *resistive* the system is to a given flux. One possible strategy to answer this question, in loose terms, would be to *constrain* the response to be constant, and measure the average magnitude of the forcing needed to maintain it. In the limit of a small response, the linear dependency between these quantities can be hoped to provide an equivalent and reciprocal measure of the transport coefficient. By analogy with the Thevenin and Norton circuit theorems, we will from now on refer to the standard, constant-forcing method as the Thevenin method, and the dual, constant-response method as the Norton method. We will again be using the mobility and the shear viscosity as our examples, so as to leverage our previous calculations as ways to validate our method.

### 5.2 General framework

We first express the method in full generality, before specializing to the non-equilibrium molecular dynamics context.

$$(5.1) \quad \begin{cases} dX_t = b(X_t)dt + \sigma(X_t)dW_t + d\Lambda_t F(X_t) \\ R(X_t) = r, \end{cases}$$

where  $b, \sigma$  and  $F$  are respectively  $R^D$ ,  $\mathbb{R}^{D \times E}$  and  $\mathbb{R}^D$ -valued functions, and  $W$  is a  $E$ -dimensional Brownian motion. We interpret  $F$  as a perturbation direction for an *equilibrium SDE*, and  $\Lambda$  is the perturbation intensity, which is determined in order to maintain the response observable  $R$  constant equal to  $r > 0$  along the dynamic's trajectories. For notational simplicity, let us introduce the following definitions:

$$(5.2) \quad b_t = b(X_t), \quad \sigma_t = \sigma(X_t), \quad F_t = F(X_t), \quad \nabla R_t = \nabla R(X_t), \quad \nabla^2 R_t = \nabla^2 R(X_t).$$

For  $A, B \in \mathbb{R}^D$ , we denote by  $P_{A,B}$  the projector onto  $A$  orthogonally to  $B$

$$(5.3) \quad P_{A,B} = \frac{A \otimes B}{A \cdot B}.$$

Finally, for any projector  $P$ , we denote by  $\bar{P}$  the complementary projector

$$(5.4) \quad \bar{P} = \text{Id} - P,$$

We can now state the following result, which follows by a simple application of Itô calculus, and gives an expression for the dynamics (5.1) without reference to  $\Lambda$ .

**Proposition 8.** *Provided it is well-posed, the solution of the following SDE solves (5.1)*

$$(5.5) \quad dX_t = \bar{P}_{F_t, \nabla R_t} [b_t dt + \sigma_t dW_t] - \frac{1}{2} \frac{\nabla^2 R_t : (\bar{P}_{F_t, \nabla R_t} \sigma_t \sigma_t^\top \bar{P}_{\nabla R_t, F_t})}{F_t \cdot \nabla R_t} F_t dt,$$

with  $\Lambda_t$  taken to be an Itô process defined by

$$(5.6) \quad d\Lambda_t = -\frac{\nabla R_t \cdot b_t}{F_t \cdot \nabla R_t} dt - \frac{\nabla R_t \cdot \sigma_t dW_t}{F_t \cdot \nabla R_t} - \frac{1}{2} \frac{\nabla^2 R_t : (\bar{P}_{F_t, \nabla R_t} \sigma_t \sigma_t^\top \bar{P}_{\nabla R_t, F_t})}{F_t \cdot \nabla R_t}.$$

*Proof.* The proof follows by analysis-synthesis. We first assume that  $\Lambda$  admits a decomposition as an Itô process,

$$(5.7) \quad d\Lambda_t = \lambda_t dt + d\tilde{\Lambda}_t,$$

where  $\lambda$  is a bounded variation process and  $\tilde{\Lambda}$  is a martingale, and that  $R$  sufficiently regular (say  $C^2$  and bounded). We can apply Ito's formula to the constant response condition, to get

$$(5.8) \quad \nabla R_t \cdot dX_t + \frac{1}{2} \nabla^2 R_t : \langle \sigma_t dW_t + d\tilde{\Lambda}_t F_t \rangle dt = 0,$$

where the brackets denote the quadratic covariation of an Itô process. By identifying the martingale and bounded variation parts of the above process with 0, we deduce an SDE for  $\Lambda_t$ ,

$$(5.9) \quad \begin{cases} \nabla R_t \cdot (\sigma_t dW_t + F_t d\tilde{\Lambda}_t) = 0 \\ \nabla R_t \cdot (b_t + \lambda_t F_t) + \frac{1}{2} \nabla^2 R_t : \langle \sigma_t dW_t + d\tilde{\Lambda}_t F_t \rangle = 0. \end{cases}$$

From the first equation, we deduce the expression of the martingale part of the forcing,

$$(5.10) \quad d\tilde{\Lambda}_t = -\frac{\nabla R_t \cdot \sigma_t dW_t}{\nabla R_t \cdot F_t},$$

and as a consequence, we may compute the quadratic covariation bracket in (5.8),

$$(5.11) \quad \langle \sigma_t dW_t + d\tilde{\Lambda}_t F_t \rangle = \left\langle \left( \text{Id} - \frac{F_t \otimes \nabla R_t}{F_t \cdot \nabla R_t} \right) \sigma_t dW_t \right\rangle = \bar{P}_{F_t, \nabla R_t} \sigma_t \sigma_t^\top \bar{P}_{F_t, \nabla R_t} dt,$$

where we use  $P_{A,B}^\top = P_{B,A}$ . Inserting equation (5.11) into (5.9) yields the full SDE for the forcing  $\Lambda_t$  (5.6) and the result follows by inserting the latter in (5.1).  $\square$

The Norton dynamics (??) should be thought of as the constant-response counterpart to a corresponding Thevenin dynamics

$$(5.12) \quad \begin{cases} dX_t = b(X_t)dt + \sigma(X_t)dW_t + \eta F(X_t), \\ \eta > 0, \end{cases}$$

which is perturbed in the same direction, but with a constant forcing intensity  $\eta$ . Both these dynamics have a corresponding equilibrium dynamics, respectively for  $r = 0$  and  $\eta = 0$ . As stated above, we propose the Norton dynamics as a possible and alternative means of computing the linear response to a small perturbation of the equilibrium dynamics. We finish this general introduction by evoking a few theoretical questions related to these dynamics.

- (i) Existence of a unique steady-state for the dynamics (5.1) and (5.12). As mentioned in the previous chapter, some results already exist in the Thevenin case.
- (ii) As before, proving the convergence of trajectory averages.
- (iii) With respect to the steady states, whose corresponding expectations we denote by  $\mathbb{E}_\eta$  in the Thevenin case, and  $\mathbb{E}_r$  in the Norton case, equality of the transport coefficients

$$\lim_{\eta \rightarrow 0} \frac{\mathbb{E}_\eta[R]}{\eta}$$

and

$$\lim_{r \rightarrow 0} \frac{r}{\mathbb{E}_r[\lambda]}.$$

This is the question of equivalence of the linear responses.

- (iv) Another natural question is that of the equivalence of the Norton and Thevenin equilibrium ensembles. This question asks about the existence of an asymptotic regime, or thermodynamic limit as  $D \rightarrow \infty$ , for which averages under the two equilibrium steady-states converge to a common value, for a sufficiently rich class of observables.
- (v) On a related note, investigating the equivalence of the non-linear responses, that is the existence of a well-defined  $r(\eta)$  such that the graphs  $(\eta, \mathbb{E}_\eta[R])$  and  $(\mathbb{E}_{r(\eta)}[\lambda], r(\eta))$  either agree or converge to a common value in some asymptotic regime when  $D \rightarrow \infty$ .
- (vi) Even more ambitiously, full equivalence of the non-equilibrium ensembles.
- (vii) Finally, we mention the question of finding a relation between the Norton equilibrium fluctuations of  $\lambda$  and the linear response, in an analogy with the Green-Kubo formula.

Let us mention that the question of the equivalence of non-equilibrium ensembles has already been investigated by Evans in [E93], although the setting is purely deterministic, and the proof is formal. The purpose of this chapter is to make somewhat rigorous a setting in which these questions are susceptible to find an answer, albeit a probably a difficult one to obtain. Furthermore, the numerical results which we present in the remainder of this report show that these questions do not have obviously negative answers, and thus should be worthy of investigation.

### 5.3 Norton dynamics for transport coefficients

We now turn to specialising the setting to the Norton counterpart of the dynamics (4.1). It reads

$$(5.13) \quad \begin{cases} dq_t = m^{-1} p_t dt, \\ dp_t = -\nabla V(q_t) dt - \gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t + d\Lambda_t F(q_t), \\ R(q_t, p_t) = r, \end{cases}$$

where  $\Lambda_t$  is the magnitude of the perturbation, which is defined by the constraint  $R(q_t, p_t) = r > 0$ . We may assume  $\gamma$  is a positive semi-definite diagonal matrix, as in the Thevenin case.  $R$  is the response observable, which we take of the form

$$R(q, p) = G(q) \cdot p,$$

with  $G$  a smooth vector field. Note in particular that  $\nabla_p R(q, p) = G(q)$ , and  $\nabla_p^2 R(q, p) = 0$ . Finally note that the forcing acts solely on the momenta. A SDE for  $\Lambda_t$  can be obtained by applying Proposition 8. It immediately implies

$$(5.14) \quad \begin{cases} dq_t = M^{-1} p_t dt \\ dp_t = \bar{P}_{F_t, G_t} \left[ -\nabla V(q_t) - dt \gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t \right] - \frac{\nabla_q R_t \cdot M^{-1} p_t}{F_t \cdot G_t} F_t dt, \end{cases}$$

with an expression for the forcing term

$$(5.15) \quad \begin{cases} d\Lambda_t = \lambda_t dt + d\tilde{\Lambda}_t, \\ \lambda_t = \frac{G_t \cdot [\nabla V(q_t) + \gamma M^{-1} p_t] - \nabla_q R_t \cdot M^{-1} p_t}{F_t \cdot G_t}, \\ d\Lambda_t = -\frac{G_t \cdot \sqrt{\frac{2\gamma}{\beta}} dW_t}{F_t \cdot G_t}. \end{cases}$$

We use notations identical to (5.3), (5.4) and (5.2) in the expressions above. The bounded and martingale parts of the forcing term can be identified in (5.15), and are given respectively by  $\lambda_t$  and  $\tilde{\Lambda}_t$ . We record the infinitesimal generator of the dynamics (5.14), which is given by

$$(5.16) \quad \mathcal{L}_\gamma^{\text{Norton}} \varphi = A\varphi - [\bar{P}_{F,G} (\nabla V + \gamma M^{-1} p)] \cdot \nabla_p \varphi - \frac{\nabla_q R \cdot M^{-1} p}{F \cdot G} F \cdot \nabla_p \varphi + \frac{1}{\beta} \nabla_p^2 \varphi : [\bar{P}_{F,G} \gamma \bar{P}_{G,F}].$$

Note  $\bar{P}_{F,G}$  is a function of the coordinate variable in the equation above.

**Remark 16.** Note that the  $G(q_t) \cdot F(q_t)$  term in the denominator may pose a question of well-posedness of the dynamics. Let us always suppose in our computations that  $G(q_t) \cdot F(q_t) > 0$ , but this is by no means automatic. Indeed, thinking of the extreme case when  $F$  and  $G$  are orthogonal everywhere and  $V = 0$  highlights the fact that this is an issue of controllability: in that case, by isotropy, the component of the momentum in the direction  $G$  will diffuse according to an Ornstein-Uhlenbeck process independent from any forcing applied in the direction  $F$ . In this case there is no way to control the response, and thus the dynamics is ill-defined.

Writing the Norton dynamics under the form (5.14) is instructive, since its structure appears clearly: its stochastic increments are comprised of the projected increments of an equilibrium Langevin dynamics, with an additional term involving  $\nabla_q R$ , which can be interpreted as a correction term enforcing the constant response constraint, in reaction to the configurational dynamics. This fact will become clearer in the following section, where we describe splitting schemes for the Norton dynamics.

### 5.3.1 Splitting schemes for numerical integration

As in every case we considered so far, we will again rely on the fact that the Norton dynamics (5.14) can be split into three simpler, structure-preserving dynamics, to construct a variety of numerical schemes.

**Definition 4** (Splitting of the Norton dynamics). Consider the following dynamics

$$(5.17) \quad \begin{cases} dq_t = M^{-1} p_t dt \\ dp_t = -\frac{\nabla_q R(q_t, p_t) \cdot M^{-1} p_t}{F(q_t) \cdot G(q_t)} F(q_t) dt \end{cases},$$

$$(5.18) \quad \begin{cases} dq_t = 0 \\ dp_t = -\bar{P}_{F_t, G_t} \nabla V(q_t) dt \end{cases},$$

and finally

$$(5.19) \quad \begin{cases} dq_t = 0 \\ dp_t = \bar{P}_{F_t, G_t} \left[ \gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t \right], \end{cases}$$

which additively combine to form the Norton dynamics (5.5). By analogy with the equilibrium setting, we will respectively refer to (5.17), (5.18) and (5.19) as the (Norton) A,B and O dynamics. At the level of the generator (5.16), this corresponds to a splitting into three operators,

$$\mathcal{L}_\gamma^{\text{Norton}} = \mathcal{L}_A^{\text{Norton}} + \mathcal{L}_B^{\text{Norton}} + \mathcal{L}_{\gamma,O}^{\text{Norton}},$$

with

$$(5.20) \quad \begin{cases} \mathcal{L}_A^{\text{Norton}}\varphi = A\varphi - \frac{\nabla_q R \cdot M^{-1} p}{F \cdot G} F \cdot \nabla_p \varphi, \\ \mathcal{L}_B^{\text{Norton}}\varphi = -(\bar{P}_{F,G} \nabla V) \cdot \nabla_p \varphi, \\ \mathcal{L}_{\gamma,O}^{\text{Norton}}\varphi = -(\bar{P}_{F,G} \gamma M^{-1} p) \cdot \nabla_p + \frac{1}{\beta} (\bar{P}_{F,G} \gamma \bar{P}_{G,F}) : \nabla_p^2 \varphi. \end{cases}$$

The claim that these dynamics are structure-preserving is summed up in the following straightforward result.

**Lemma 3.** *Let  $(q_t, p_t)$  refer to the solution of any of the Norton A,B or O dynamics. Then if  $R(q_0, p_0) = r$ ,  $R(q_t, p_t) = r$  for all  $t > 0$ .*

*Proof.* The result can be proven by three thoughtless applications of the chain rule or Ito's formula. A quicker and more instructive proof follows from observing that for any  $X, q, p \in \mathbb{R}^{dN}$ ,

$$\begin{aligned} G(q) \cdot \bar{P}_{F(q), G(q)} X &= \bar{P}_{F(q), G(q)}^\top G(q) \cdot X \\ &= \bar{P}_{G(q), F(q)} G(q) \cdot X \\ &= 0, \end{aligned}$$

since  $P_{G(q), F(q)}$  is a projector onto the one-dimensional subspace spanned by  $G(q)$  and  $\bar{P}_{G(q), F(q)}$  is its complement. Hence any dynamics of the form

$$dq_t = 0, \quad dp_t = \bar{P}_{G_t, F_t} [b(q_t, p_t) dt + \sigma(q_t, p_t) dW_t]$$

preserves  $R(q_t, p_t) = G(q_t) \cdot p_t$ , by applying the Itô formula, since  $\nabla_p R(q, p) = G(q)$  and  $\nabla_p^2 R = 0$ . This implies the result for the B and O dynamics. However, since the Norton dynamics also preserves  $R(q_t, p_t)$  by construction, the preservation property for the A dynamics follows immediately, because the contributions of the A, B and O dynamics to  $dR(q_t, p_t) = 0$  for the full dynamics combine additively, again by Itô's formula in the absence of a quadratic covariation term.  $\square$

As before, this splitting allows us to define a variety of numerical schemes for the Norton dynamics, following the same strategy as in Section 2.2.6. In order to do so, we must prescribe a way to integrate each of the A,B and O steps individually. This is what we now turn to. In the following, let us fix a timestep  $\Delta t > 0$ , and a response intensity  $r > 0$ . Similar to the equilibrium setting, the B step can be integrated analytically, since its trajectories are ballistic. Thus we have the following algorithm.

**Algorithm 2** (Norton B step).

$$(5.21) \quad \begin{cases} \tilde{p}^1 = p^0 - \nabla V(q^0) \\ p^1 = \tilde{p}^1 + \Delta t \lambda^1 F(q^0), \end{cases}$$

where  $\lambda^1$  is determined in order to conserve the response:

$$(5.22) \quad G(q^0) \cdot p^1 = r \iff \Delta t \lambda^1 = \frac{r - G(q^0) \cdot \tilde{p}^1}{F(q^0) \cdot G(q^0)}.$$

We next rely on the fact that the Norton O dynamics is in Ornstein–Uhlenbeck form to compute its exact solution over one timestep, using the following result.

**Proposition 9.** *We consider the following general form of the Norton Ornstein–Uhlenbeck process in  $\mathbb{R}^{dN}$ .*

$$(5.23) \quad dp_t = -P\gamma M^{-1} p_t dt + P\sqrt{\frac{2\gamma}{\beta}} dW_t,$$

where  $P^2 = P$  is a projector (not necessarily orthogonal),  $\gamma$  and  $M^{-1}$  are positive definite symmetric matrices, and  $W$  is a standard Brownian motion. We denote by  $\bar{P}$  the complementary projector  $\bar{P} = \text{Id} - P$ . Assume also that  $P$  and  $\gamma M^{-1}$  commute. Then, the analytical solution of (5.23) writes

$$(5.24) \quad p_t = \bar{P}p_0 + Pp_t^{\text{eq}},$$

where  $p_t^{\text{eq}}$  is the solution of the coupled equilibrium Ornstein–Uhlenbeck process,

$$(5.25) \quad \begin{cases} dp_t^{\text{eq}} = -\gamma M^{-1} p_t^{\text{eq}} dt + \sqrt{\frac{2\gamma}{\beta}} dW_t, \\ p_0^{\text{eq}} = p_0. \end{cases}$$

*Proof.* The proof follows the standard strategy of applying Itô’s formula to the process rescaled by  $e^{tP\gamma M^{-1}}$ . It yields

$$(5.26) \quad d(e^{tP\gamma M^{-1}} p_t) = e^{tP\gamma M^{-1}} \left( P\gamma M^{-1} p_t dt - P\gamma M^{-1} p_t dt + P\sqrt{\frac{2\gamma}{\beta}} dW_t \right),$$

whence, integrating in  $t$  and multiplying both sides by  $e^{-tP\gamma M^{-1}}$ ,

$$(5.27) \quad p_t = e^{-tP\gamma M^{-1}} p_0 + \int_0^t e^{-P\gamma M^{-1}(t-s)} P\sqrt{\frac{2\gamma}{\beta}} dW_s.$$

This yields an expression for the solution whatever the particular properties of  $\gamma$ ,  $M^{-1}$  and  $P$ . However, since  $P$  is a projector, we can make use of the following useful formula: if  $A$  and  $P$  are square matrices with  $AP = PA$ ,

$$(5.28) \quad e^{PA} = \sum_{k=0}^{\infty} \frac{(PA)^k}{k!} = \sum_{k=0}^{\infty} \frac{P^k A^k}{k!} = \text{Id} + P \sum_{k=1}^{\infty} \frac{A^k}{k!} = \text{Id} + P(e^A - \text{Id}) = \bar{P} + Pe^A.$$

The second equality follows from the fact that  $A$  and  $P$  commute, while the third follows from a repeated application of the projector identity  $P = P^2 = P^3 = \dots$ . This is sometimes referred to as Rodrigue’s formula (see for instance [15]chapter 8). Applying’s Rodrigue’s formula to our analytical form yields

$$(5.29) \quad p_t = \bar{P}p_0 + Pe^{-t\gamma M^{-1}} p_0 + \int_0^t (\bar{P} + Pe^{-\gamma M^{-1}(t-s)}) P\sqrt{\frac{2\gamma}{\beta}} dW_s.$$

expanding the product inside the integral, and using  $\bar{P}P = 0$ , we get

$$(5.30) \quad p_t = \bar{P}p_0 + Pe^{-t\gamma M^{-1}} p_0 + \int_0^t Pe^{-\gamma M^{-1}(t-s)} P\sqrt{\frac{2\gamma}{\beta}} dW_s.$$

By our commutativity assumption, we can factor out the  $P$ s from the integral sign and obtain (5.31)

$$p_t = \bar{P}p_0 + Pe^{-t\gamma M^{-1}} p_0 + P^2 \int_0^t e^{-\gamma M^{-1}(t-s)} \sqrt{\frac{2\gamma}{\beta}} dW_s = \bar{P}p_0 + P \left( e^{-t\gamma M^{-1}} p_0 + \int_0^t e^{-\gamma M^{-1}(t-s)} \sqrt{\frac{2\gamma}{\beta}} dW_s \right),$$

regrouping the terms in  $P$  and using again  $P^2 = P$ . The result follows by simply recognizing the parenthesized term as  $p_t^{\text{eq}}$ .  $\square$

**Remark 17.** The commutativity assumption may seem overly restrictive, but in fact it is enough in the cases we consider. For instance, in the case of shear viscosity computations with anisotropic friction,  $M$  is a scalar multiple of the identity and  $\gamma$  is diagonal and constant with respect to longitudinal coordinates, so that the commutativity condition is indeed verified.

As a consequence of this computation, we may define a numerical strategy to integrate the fluctuation-dissipation part of the Norton dynamics.

**Algorithm 3** (Norton O step).

$$(5.32) \quad \begin{cases} \tilde{p}^1 = \alpha_{\Delta t} p^0 + \sigma_{\Delta t} \mathcal{G}^1, \\ p^1 = \tilde{p}^1 + \lambda^1 F(q^0), \end{cases}$$

where  $\mathcal{G}^1$  is a standard  $dN$ -dimensional Gaussian,  $\alpha_{\Delta t}$ ,  $\sigma_{\Delta t}$  are given by (2.48), and  $\lambda^1$  is again determined by equation (5.22) in order to conserve the response.

Unfortunately, the A dynamics (5.17) cannot be solved analytically. However, its general form is that of an infinitesimal Hamiltonian increment on the position coordinate, with an additional correction term in the direction  $F$  on the momentum coordinate. Furthermore, we know the response is conserved by the A dynamics. This naturally suggests the following scheme.

**Algorithm 4.**

$$(5.33) \quad \begin{cases} q^1 = q^0 + \Delta t M^{-1} p^0, \\ p^1 = p^0 + \Delta t \lambda^1 F(q^1), \end{cases}$$

where  $\lambda^1$  is determined by

$$(5.34) \quad G(q^1) \cdot p^1 = r \iff \Delta t \lambda^1 = \frac{r - G(q^1) \cdot p^0}{F(q^1) \cdot G(q^1)}$$

to enforce the constant response constraint.

An advantage of schemes for the Norton dynamics formed by chaining and iterating steps of (5.33), (5.21) and (5.32) is that they immediately yield a discretization of the forcing intensity  $d\Lambda_t$ , and thus an estimation of the average  $\mathbb{E}_r[\lambda]$ , which is the quantity of prime interest to us. Indeed, observe from equation (5.15) that we can write

$$(5.35) \quad d\Lambda_t = \lambda_t^A dt + \lambda_t^B dt + d\Lambda_t^O,$$

with

$$(5.36) \quad \begin{cases} \lambda_t^A = -\frac{\nabla_q R(q_t, p_t) \cdot G(q_t)}{F(q_t) \cdot G(q_t)}, \\ \lambda_t^B = -\frac{\nabla V(q_t) \cdot G(q_t)}{F(q_t) \cdot G(q_t)}, \\ d\Lambda_t^O = -\frac{\left[ -\gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t \right] \cdot G(q_t)}{F(q_t) \cdot G(q_t)}. \end{cases}$$

These individual parts can be interpreted individually as the forcing intensities in the direction  $F(q_t)$  applied to each of the A,B and O dynamics, relative to the corresponding equilibrium dynamics. This yields a natural interpretation of the  $\lambda^n$  terms along numerical trajectories, as discretizations of (5.36), and thus one can use these ergodic averages for the average forcing. As a full example, we record the BABO-like scheme we used in our simulations, as well as the corresponding estimator of  $\mathbb{E}_r[\lambda]$ . This procedure can of course be generalized to other orderings of the operators in the splitting. The impact of such a choice on the discretization error of the mean force should be the subject of future investigation.

**Example 7** (Estimation of the mean force using a BABO scheme.). *The numerical scheme is implemented, for a fixed timestep  $\Delta t > 0$ , by iterating*

$$(5.37) \quad \left\{ \begin{array}{l} \tilde{p}^{n+\frac{1}{4}} = p^n - \frac{\Delta t}{2} \nabla V(q^n) \\ \frac{\Delta t}{2} \lambda^{n+\frac{1}{4}} = \frac{r - G(q^n) \cdot \tilde{p}^{n+\frac{1}{4}}}{F(q^n) \cdot G(q^n)} \\ p^{n+\frac{1}{4}} = \tilde{p}^{n+\frac{1}{4}} + \frac{\Delta t}{2} \lambda^{n+\frac{1}{4}} F(q^n) \\ q^{n+1} = q^n + \Delta t M^{-1} p^{n+\frac{1}{4}} \\ \Delta t \lambda^{n+\frac{1}{2}} = \frac{r - G(q^{n+1}) \cdot p^{n+\frac{1}{4}}}{F(q^{n+1}) \cdot G(q^{n+1})} \\ p^{n+\frac{1}{2}} = p^{n+\frac{1}{4}} + \Delta t \lambda^{n+\frac{1}{2}} F(q^{n+1}) \\ \tilde{p}^{n+\frac{3}{4}} = p^n - \frac{\Delta t}{2} \nabla V(q^{n+1}) \\ \frac{\Delta t}{2} \lambda^{n+\frac{3}{4}} = \frac{r - G(q^{n+1}) \cdot \tilde{p}^{n+\frac{3}{4}}}{F(q^{n+1}) \cdot G(q^{n+1})} \\ p^{n+\frac{3}{4}} = \tilde{p}^{n+\frac{3}{4}} + \frac{\Delta t}{2} \lambda^{n+\frac{3}{4}} F(q^{n+1}) \\ \tilde{p}^{n+1} = \alpha_{\Delta t} p^{n+\frac{3}{4}} + \sigma_{\Delta t} \mathcal{G}^{n+1} \\ \Delta t \lambda^{n+1} = \frac{r - G(q^{n+1}) \cdot \tilde{p}^{n+1}}{F(q^{n+1}) \cdot G(q^{n+1})} \\ p^{n+1} = \tilde{p}^{n+1} + \Delta t \lambda^{n+1} F(q^{n+1}). \end{array} \right.$$

If the scheme is iterated  $N_{\text{iter}}$  times, the mean force can be estimated by

$$(5.38) \quad \hat{\lambda}'_{N_{\text{iter}}} = \frac{1}{N_{\text{iter}}} \sum_{k=0}^{N_{\text{iter}}-1} \left( \frac{\lambda^{k+\frac{1}{4}} + \lambda^{k+\frac{3}{4}}}{2} + \lambda^{k+\frac{1}{2}} + \lambda^{k+1} \right).$$

Note that the  $\lambda^{k+1}$  also account for the part of the correction due to the Gaussian increment  $\mathcal{G}^{k+1}$ . In practice, it may be possible to replace  $\lambda^{k+1}$  with an estimation of the bounded variation part of  $d\Lambda^O$ ,

$$\lambda_t^O = \frac{\gamma M^{-1} p_t \cdot G(q_t)}{F(q_t) \cdot G(q_t)},$$

to obtain an estimator with a lower variance. For example, in our simulations, we took  $M = m\text{Id}$  and a constant  $\gamma$ , so that using the constant-response condition,

$$\lambda_t^O = \frac{\gamma m^{-1}}{F(q_t) \cdot G(q_t)}$$

, so that the mean force estimator we considered in practice was

$$(5.39) \quad \hat{\lambda}_{N_{\text{iter}}} = \frac{1}{N_{\text{iter}}} \sum_{k=0}^{N_{\text{iter}}-1} \left( \frac{\lambda^{k+\frac{1}{4}} + \lambda^{k+\frac{3}{4}}}{2} + \lambda^{k+\frac{1}{2}} + \frac{\gamma m^{-1}}{F(q^k) \cdot G(q^k)} \right).$$

### 5.3.2 Physical interpretation

We conclude this theoretical discussion by pointing out a property of the Norton dynamics (5.14) in the deterministic case  $\gamma = 0$ , and when the forcing direction  $F$  is proportional to the response direction  $G$ . Without loss of generality, we assume  $F = G$ , since changing the proportionality constant amounts to changing the scale of the response. We also assume that the system is

homogeneous, so that we may take  $M = \text{Id}$ . In this case the projectors  $P_{G_t, F_t}$  and the like all become orthogonal projectors onto the one-dimensional subspace spanned by  $F_t$ , or its orthogonal complement. We show that in this particular case there is a physical interpretation for the dynamics

**Proposition 10** (Gauss's principle of least constraint). *Consider the deterministic version of the Norton dynamics,*

$$(5.40) \quad \begin{cases} \dot{q} = p, \\ \dot{p} = -\bar{P}_{F(q)} \nabla V(q) - \frac{\nabla F(q)p \cdot p}{|F(q)|^2}, \end{cases}$$

where the constraint is given by

$$F(q_t) \cdot p_t = F(q_0) \cdot p_0 = r,$$

and where we write  $\bar{P}_F$  for  $\bar{P}_{F,F}$ . The rightmost term in the equation for  $\dot{p}$  comes from writing  $\nabla_q R(q, p) = \nabla G(q)p$ . Then Gauss's principle of least constraint is satisfied:

$$(5.41) \quad \begin{cases} \dot{q} = p, \\ \dot{p} = \underset{f \cdot F(q) + \nabla F(q)p \cdot p = 0}{\operatorname{argmin}} | -\nabla V(q) - f |^2. \end{cases}$$

In other words, the force applied to the system undergoing the Norton dynamics minimizes at each point in time the Euclidean distance to the force applied to the same system undergoing the Hamiltonian dynamics (1.2), subject to the constant response constraint.

*Proof.* Firstly, let us make clear that the constraint in the minimization problem actually expresses the constant response constraint. Since

$$r = F(q) \cdot p,$$

a differentiation in time yields

$$(5.42) \quad 0 = \nabla F(q)\dot{q} \cdot p + F(q) \cdot \dot{p} = \nabla F(q)p \cdot p + F(q) \cdot \dot{p}.$$

Since the force applied to the system is equal to  $\dot{p}$ , this effectively yields a constraint on the set of possible forces applied to the Norton system. This is analogous to the case of a Hamiltonian system subject to a set of holonomic constraints, which depend only on the position  $q$ . In this case, a so-called hidden or shadow constraint on the set of admissible forces can be obtained by twice-differentiating the constraint condition. For a pedagogical discussion of this case, we refer the reader to [15], section 4.3. In our non-holonomic case (the constraint also depends on  $p$ ), a single time-differentiation is enough. We can then compute the minimizer

$$(5.43) \quad \underset{f \cdot F(q) + \nabla F(q)p \cdot p = 0}{\operatorname{argmin}} | -\nabla V(q) - f |^2.$$

Since the constraint is affine in  $f$ , the unique solution, provided  $F(q) \neq 0$ , is given by an orthogonal projection of  $-\nabla V(q)$  onto the hyperplane (in the variable  $f$ )

$$f \cdot F(q) + \nabla F(q)p \cdot p = 0,$$

whose normal direction is given by  $F(q)$ . Hence the minimizer  $f^*$  can be written

$$(5.44) \quad f^* = -\nabla V(q) - \lambda F(q)$$

for some Lagrange multiplier  $\lambda$ , which is determined by

$$f^* \cdot F(q) + \nabla F(q)p \cdot p = 0.$$

Substituting our tentative expression for  $f^*$ , we get

$$(-\nabla V(q) - \lambda F(q)) \cdot F(q) + \nabla F(q)p \cdot p = 0.$$

This gives

$$(5.45) \quad \lambda = -\frac{\nabla V(q) \cdot F(q)}{|F(q)|^2} + \frac{\nabla F(q)p \cdot p}{|F(q)|^2},$$

and finally

$$f^* = -\bar{P}_{F(q)} \nabla V(q) - \frac{\nabla F(q)p \cdot p}{|F(q)|^2} F(q),$$

which concludes the proof.  $\square$

This result gives a nice physical interpretation of the deterministic Norton dynamics: it corresponds to a *physical* trajectory of a Hamiltonian system subject to the constant-response non-holonomic constraint. Indeed, Gauss's principle of least constraint [8] is equivalent to the Legendre-d'Alembert principle, which can be used to derive the equations of motion for constrained classical systems. The use of Gauss's principle of least constraint to enforce thermodynamic constraints for non-equilibrium molecular dynamics was already discussed in [5]. It is not yet clear to us how to extend this result to the case where  $F$  and  $G$  are non-collinear, and when the system is inhomogeneous, although it seems likely in this case Gauss's principle should have to be stated with respect to some modified scalar product. Another interesting direction would be to extend the principle to the properly Langevin case  $\gamma > 0$ . At any rate, we now turn to presenting our numerical results. As before, we will be using the two examples of mobility and shear viscosity computations.

## 5.4 Numerical results

### 5.4.1 Mobility

The Norton dynamics for mobility is recovered by setting  $F$  to be constant

$$F \in \{F_C, F_S\},$$

and taking

$$G = M^{-1}F$$

indeed, the response

$$R(q, p) = M^{-1}F \cdot p = F \cdot M^{-1}p,$$

since  $M$  is symmetric, which corresponds to the response observable for the mobility. Since  $F$  is constant in the  $q$  variable, the correction term in the A dynamics (5.17) is 0. Hence the A step (5.33) is analytically correct, with a vanishing Lagrange multiplier. We thus estimate the mean force through the estimator (5.39) with  $\lambda^{k+\frac{1}{2}} = 0$ . An estimator of

$$\lim_{r \rightarrow 0} \frac{\mathbb{E}_r[\lambda]}{r}$$

can then be obtained in the same way as (4.10), by

$$(5.46) \quad \hat{\rho}_F = |\hat{\lambda}|^{-2} \hat{\lambda} \cdot r = \underset{\rho \in \mathbb{R}}{\operatorname{argmin}} \left| \rho \hat{\lambda} - r \right|^2,$$

where  $\hat{\lambda} \in \mathbb{R}^k$  is a vector of  $k$  ergodic averages of the form (5.39), and  $r$  is a vector containing the corresponding response intensities. Assuming, for simplicity's sake, that all ergodic averages

are computed with the same number  $N_{\text{iter}}$  of time steps, and that  $\widehat{\lambda}$  has a (diagonal) matrix of asymptotic variances  $\Sigma_{\lambda}^2$ , applying the delta-method to (5.46) yields the following estimator for the asymptotic variance of  $\widehat{\rho}_F$

$$\sigma_{\rho_F}^2 = \nabla g(\widehat{\lambda})^\top \Sigma_{\lambda}^2 \nabla g(\widehat{\lambda}),$$

where

$$g(y) = |y|^{-2} y \cdot r,$$

which yields, after computation,

$$(5.47) \quad \sigma_{\rho_F}^2 = \sum_{i=1}^k \Sigma_{\lambda,ii}^2 \left( \frac{r_i |\widehat{\lambda}|^2 - 2\widehat{\lambda}_i (\widehat{\lambda} \cdot r)}{|\widehat{\lambda}|^4} \right)^2.$$

In the case  $k = 1$ , this becomes

$$(5.48) \quad \sigma_{\rho_F}^2 = \sigma_{\lambda}^2 r^2 \widehat{\lambda}^{-4}.$$

In Figure 5.1, we plot the response-forcing diagram in the linear regime. The estimated transport coefficient from the color drift is consistent with the results obtained from the Thevenin method. For the single drift, we observe a discrepancy in the linear response. In Figure 5.2, we plot the response-forcing curve in the non-linear regime, and superimpose the data obtained from the Thevenin method. For the color drift, the two computed equations of state coincide perfectly. For the single drift, we again observe a small discrepancy. This leads us to speculate that the equivalence of non-linear responses may only hold in the thermodynamic limit, and for forcings which act on the bulk of the system, contrarily to the single drift forcing. In Figure 5.3, we show the computed asymptotic variance for estimators of  $\rho_F$  based on the Thevenin methods and the Norton methods. Asymptotic variances for estimators of the response and mean force were computed using block averages, and the corresponding asymptotic variance for the Norton estimator of  $\rho_F$  was obtained using (5.48). It appears that the two methods are roughly equivalent in terms of asymptotic variance, and that the variance for the finite-difference estimator scales like the inverse square of the forcing intensity, consistent with the result expected in the Thevenin case, as noted in Remark 12.

The final estimates for the transport coefficients using the Norton method are

$$\rho_{FC} = 0.1211 \pm 0.0006, \quad \rho_{FS} = 0.1089 \pm 0.0005.$$

Estimates of the transport coefficients were obtained using (5.46), and statistical error was estimated using (5.47)

### 5.4.2 Shear viscosity

The NEMD response observable (4.18) for shear viscosity can be recovered in the form

$$R_k(q, p) = \tilde{G}_k(q) \cdot p$$

by setting

$$(5.49) \quad \forall 1 \leq i \leq N, \forall 2 \leq \alpha \leq d, \quad \tilde{G}_k(q)_{i1} = \frac{1}{m} \exp \left( \frac{2ik\pi q_{i2}}{L} \right), \quad \tilde{G}_k(q)_{i\alpha} = 0.$$

Thus we can apply the Norton method. In practice, to avoid dealing with a complex exponential, we rather define

$$(5.50) \quad \forall 1 \leq i \leq N, \forall 2 \leq \alpha \leq d, \quad G_k(q)_{i1} = \frac{1}{m} \sin \left( \frac{2k\pi q_{i2}}{L} \right), \quad G_k(q)_{i\alpha} = 0.$$

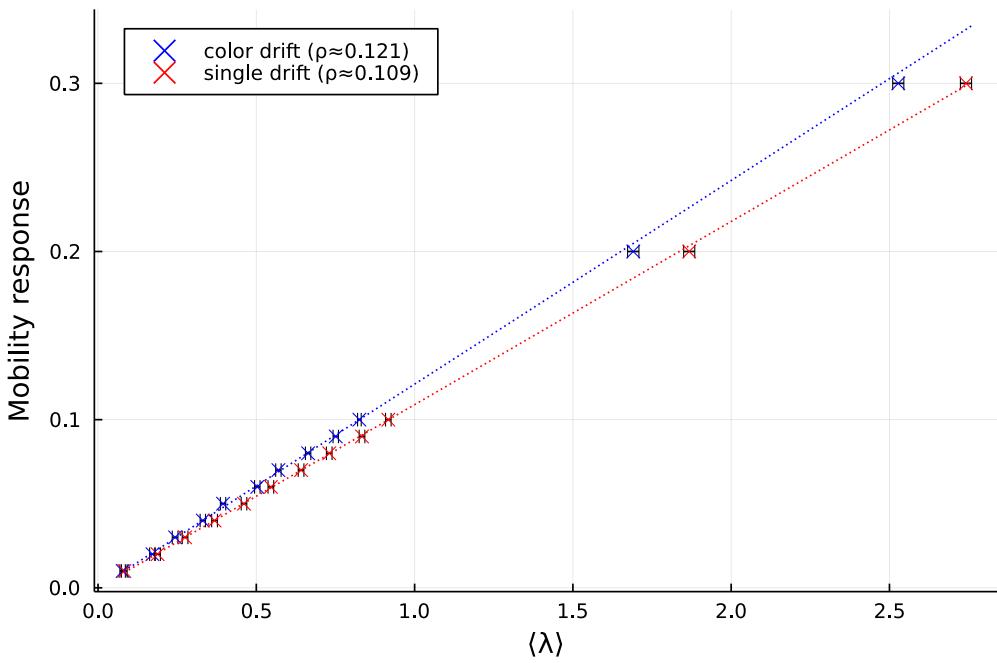


Figure 5.1: Mobility response versus average forcing intensity in the linear regime. Least squares linear regression lines are plotted in dotted line, and the estimated transport coefficients are indicated in the legend.

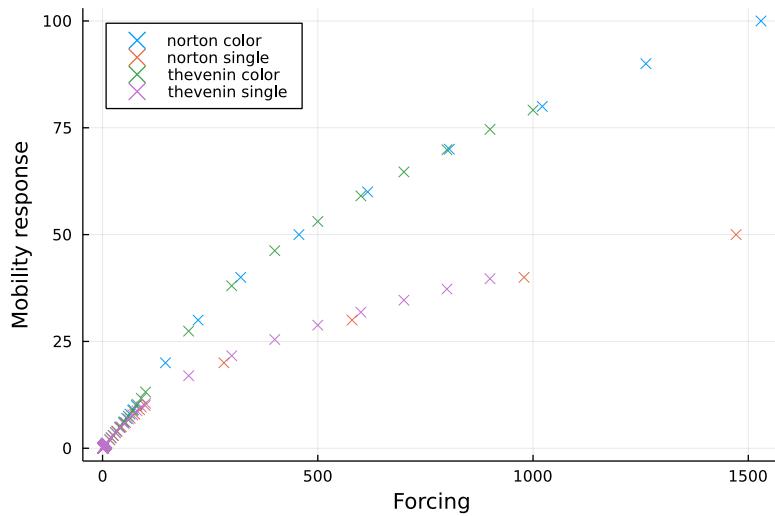


Figure 5.2: Comparison of the Thevenin and Norton mobility equations of state.

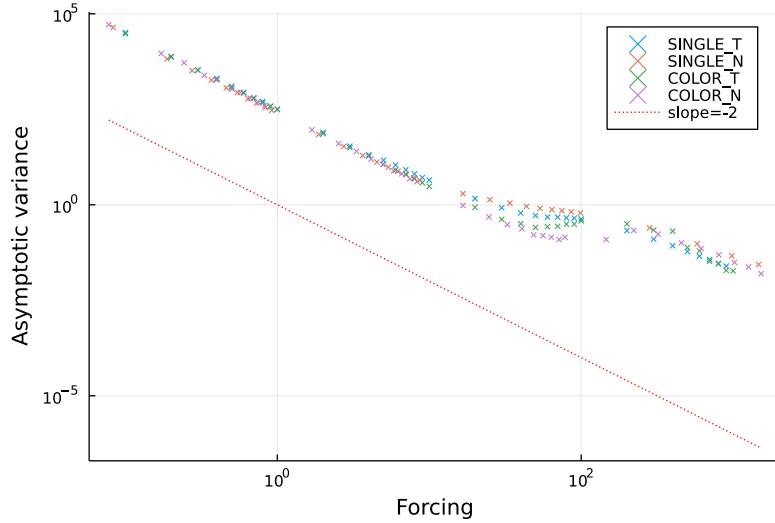


Figure 5.3: Plot of the asymptotic variance as a function of the forcing intensity for the different methods, on a log-log plot. Results corresponding to a Thevenin estimator are denoted with a T, and those corresponding to the Norton estimator are denoted with a N.

This can be achieved at no cost to generality by considering an appropriate translation of the forcing profile with respect to the results of Figure 4.2. Estimators for the transport coefficient (4.19) and estimates of the statistical error can be obtained identically to the mobility case. All simulations were run in the reference thermodynamics condition (4.21) for a minimum length of  $8 \times 10^7$  timesteps.

In Figure 5.4, we plot the normalized response

$$\frac{R_1}{F_1}$$

against the estimated average forcing  $\langle \lambda \rangle$  in the linear regime, for the different forcing profiles. The estimated transport coefficients are consistent with one another, and agree with those obtained in the Thevenin and Green–Kubo setting. In Figure 5.6, we compare the non-linear response profiles between the Thevenin and Norton methods, for each of the forcing profiles. In every case, the responses coincide perfectly throughout. In Figure 5.5 we compare asymptotic variances of the finite difference estimators for the normalized response using the Thevenin and Norton method. The asymptotic variance was estimated directly from the response time series in the Thevenin case, using a block averaging procedure, and for the Norton method, a block averaging procedure on the forcing time series was combined with a delta-method. We find that the Norton estimators consistently outperform their Thevenin counterparts, and that not all forcing profiles are equal in terms of statistical error.

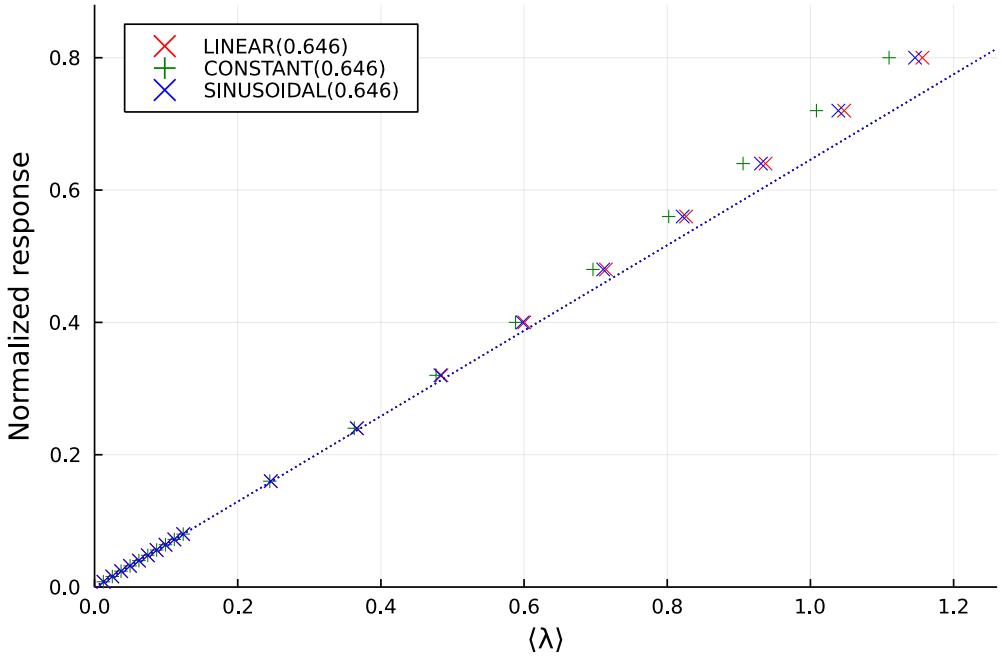


Figure 5.4: Normalized Fourier response versus forcing intensity for the three transverse forcing profiles. The size of the linear response regime is roughly the same for every type of forcing. Least squares linear regression lines on the 10 first values are plotted in dotted line, and estimated normalized transport coefficients are indicated in the legend.

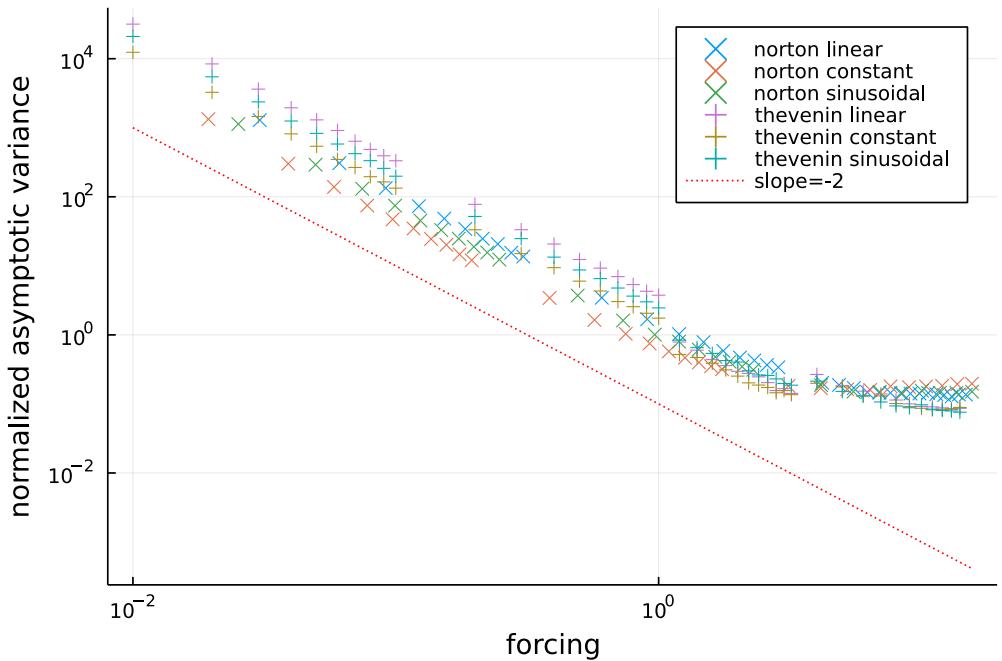


Figure 5.5: Scaling of the asymptotic variance for the finite difference estimator of the normalized transport coefficient for shear viscosity

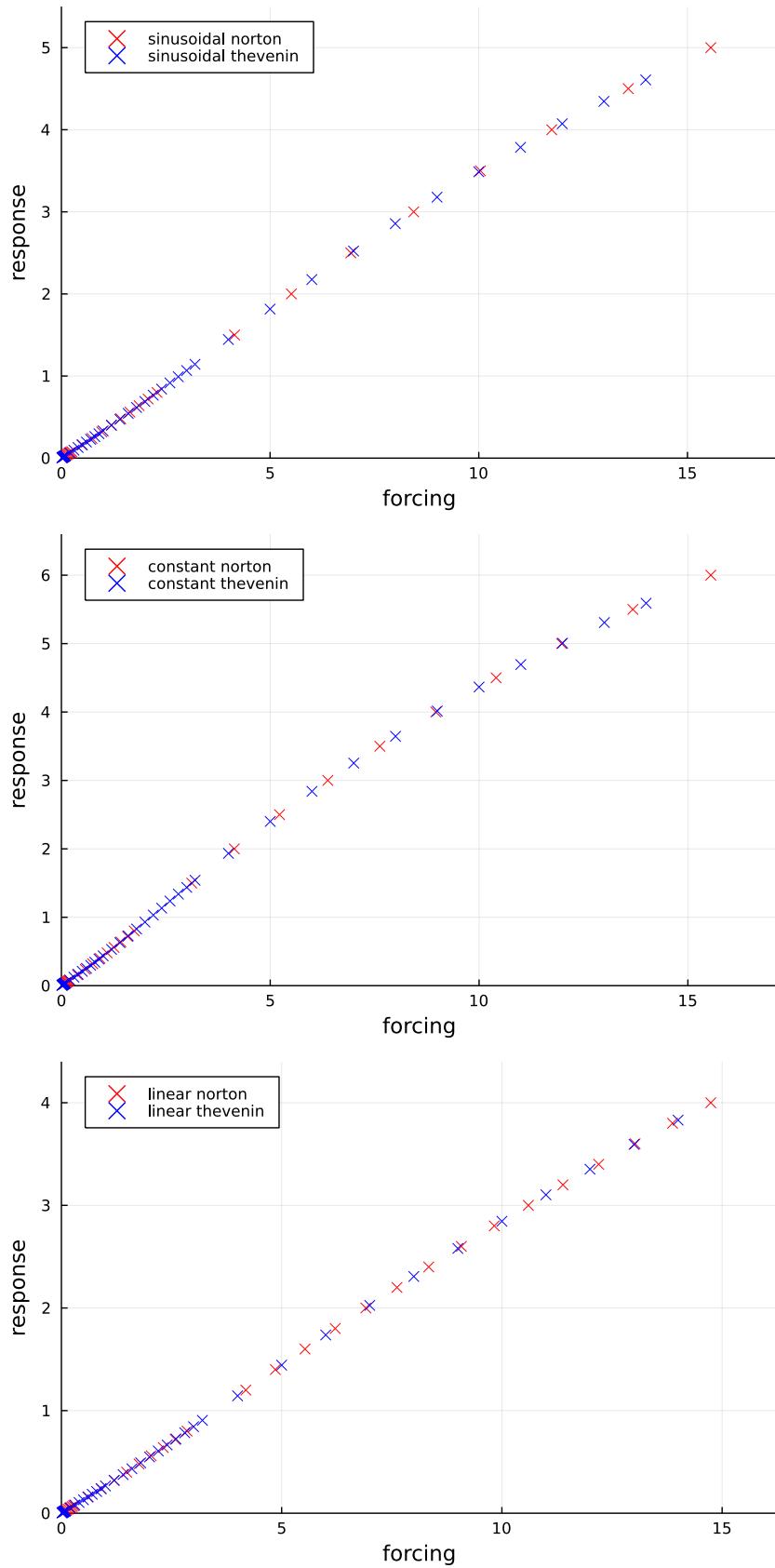


Figure 5.6: Norton and Thevenin equations of state for the shear-viscosity Fourier response, for each type of forcing profile.

## Appendix A

# Implementation details

We conclude this report by a short appendix dedicated to providing some details about the implementation of the methods we described in the previous chapters, both from a historical point of view, meaning that we will be documenting the process, as well as of implementing the methods we described, as well as filling in some methodological gaps. Throughout this document, the reader will notice that little is said about  $\nabla V$ , meaning the methods we described simply invoked  $\nabla V$  where it is needed, without getting into the specifics of *how* it is computed. For a pair interaction potential of the form (1.3), the naive method of summing over all pairs of particles incurs a cost of

$$\frac{N(N - 1)}{2} = O(N^2)$$

evaluations of the pair interaction function  $v$ . Another point we alluded to is that the statistical error on the computation of averages usually dissipates at a rate  $O(1/\sqrt{\Delta t N_{\text{iter}}})$ . Furthermore, for averages to be physically meaningful,  $N$  must be taken large enough for the thermodynamic limit to be considered reached. In practice, a quadratic scaling of the computational cost of computing  $\nabla V$  may render essentially impossible the computation of long numerical trajectories. Instead, one has to rely on approximate evaluations of  $\nabla V$ . For instance, in the case of a Lennard-Jones interaction, the pair interaction potential decays quickly, in  $O(r^{-6})$  as a function of the interparticular distance  $r$ . In this case, a reasonable approximation consists on imposing a fixed maximal interaction distance  $r_c$ , called the *cutoff* distance, and asserting that particles which are further apart than  $r_c$  do not interact. This amounts mathematically to replacing the pair interaction function  $v(r)$  by a cut-off version  $v(r)\mathbb{1}_{r < r_c}$ , effectively modifying the potential. There are various ways to alter this procedure in order to make the modified potential continuous or  $C^1$ , for instance using spline interpolation. Of course, this procedure alone does not solve anything, since one still needs to determine *which* pairs of particles are neighbors, meaning they are closer apart than  $r_c$ , which still amounts to  $O(N^2)$  computations of the (squared) Euclidean norm. To make this viable, one has to store a list of neighbor pairs in an efficient data structure which only needs to be updated once every few simulation iterations, since the set of neighboring pairs tend not to drastically change over short time intervals. In practice, and depending on the timestep, the meaning of the word *few* can sometimes be understood to be as large as 50. There are various strategies to implement such a data structure, such as Verlet lists or cell lists, which are described in detail in (citer Allen and Tildesley eg). At any rate, since the number of neighbors to a given particle should on average be close to  $4\pi\rho r_c^3/3 = O(1)$ , where  $\rho$  is the particle density, the cost of computing  $\nabla V$  becomes (ideally) **linear** in  $N$ , which makes long time simulations a practical possibility. We mention the issue of short-range interactions since all our molecular simulations used  $V$  of the Lennard-Jones form, but other methods exist in the case where  $V$  decays slowly like the Coulomb electrostatic potential, relying in this case on a fast decay property in Fourier space. Of course, many more tricks exists, which collectively form the whole craftsmanship of Molecular Dynamics (MD) simulations: finding efficient, approximate ways to compute  $V$  and  $\nabla V$ .

The process of implementing from scratch these methods is time consuming, and unnecessary

given the number of freely available MD packages. However, some of these packages are quite ancient and heavily optimized so that the process of implementing new methods comes with quite a heavy learning overhead, as well as being harder to debug. Instead, we chose to make use of a Julia package for MD simulation, thus taking this opportunity to learn the programming language Julia [1], which has been gaining a lot of traction in the scientific computing user base, in part because it offers the shared benefit of performance comparable to that of C or C++, as well as possessing a rich ecosystem of highly intercompatible packages, powered by a strong community and a design philosophy centered around generic, type-agnostic, interfaces.

The first step in the internship consisted in choosing a package between two contenders, NBodySimulator [19], and Molly [9]. NBodySimulator, in summary, acts as a wrapper around the native differential equations Julia framework, constructing a differential equation or stochastic differential equation from the specification of a molecular system. As such, it aims to compute *exact* solutions to the dynamics, or at least exact evaluations of  $V$ . On the other hand, Molly offers a number of neighbor-finding strategies, such as the cell list method, a parallelized naive iteration over all pairs of neighbors, as well as a tree-search based method. In Figure A.1, we show the results of a comparison of the scaling of the simulation time as a function of the system size, for NBodySimulator and Molly's various neighbor finding strategies. We find that, if for very small systems, NBodySimulator outperforms Molly, the advantage quickly disappears, for  $N \geq 750$  approximately, in the case of the cell list method. Interestingly, even the naive double loop for Molly overtakes NBodySimulator in performance. This is because by default, Molly parallelizes the force computation, while there is no parallelization in NBodySimulator.

On this basis, we chose to implement our methods in Molly. Along the way, some of our implementations were integrated in the Molly source, in large part during a one-week stay in the MRC Laboratory of Molecular Biology in Cambridge, where the main author of Molly's source code, Joe Greener, works. These integrations include:

- i) A cutoff strategy based on a cubic spline interpolation,
- ii) A general Langevin integrators, which works for all splitting orderings (2.43) and in the (non)-equilibrium setting,
- iii) A logger to efficiently estimate time correlations functions of the form (4.26),
- iv) A logger to track the self-diffusion coordinates (4.40) for Einstein mobility computations,
- v) A refactoring of Molly's general logging architecture, for more flexibility and extensibility,
- vi) The addition of a parameter for Boltzmann's constant, allowing for fully consistent sets of user-defined custom systems of units,
- vii) An online ergodic average and asymptotic variance estimator for a user-specified observable.

These were additionally documented, and non-regression tests were implemented for each of them. We refer the reader to the Molly documentation and source code for more details. Other implementations were too use-specific, not general or not mature enough to be integrated. These include

- i) A pressure observable based on (1.12),
- ii) The implementation of NEMD force fields for mobility and shear viscosity computations,
- iii) The implementation of an integrator for the Norton method,
- iv) Various logging and visualization utilities.

These implementations are available in the repository [3], although we apologize in advance to the interested reader for the poorly organized, largely uncommented mess that lies within. We expect, especially for the pressure observable, that a future integration in Molly will be possible, for

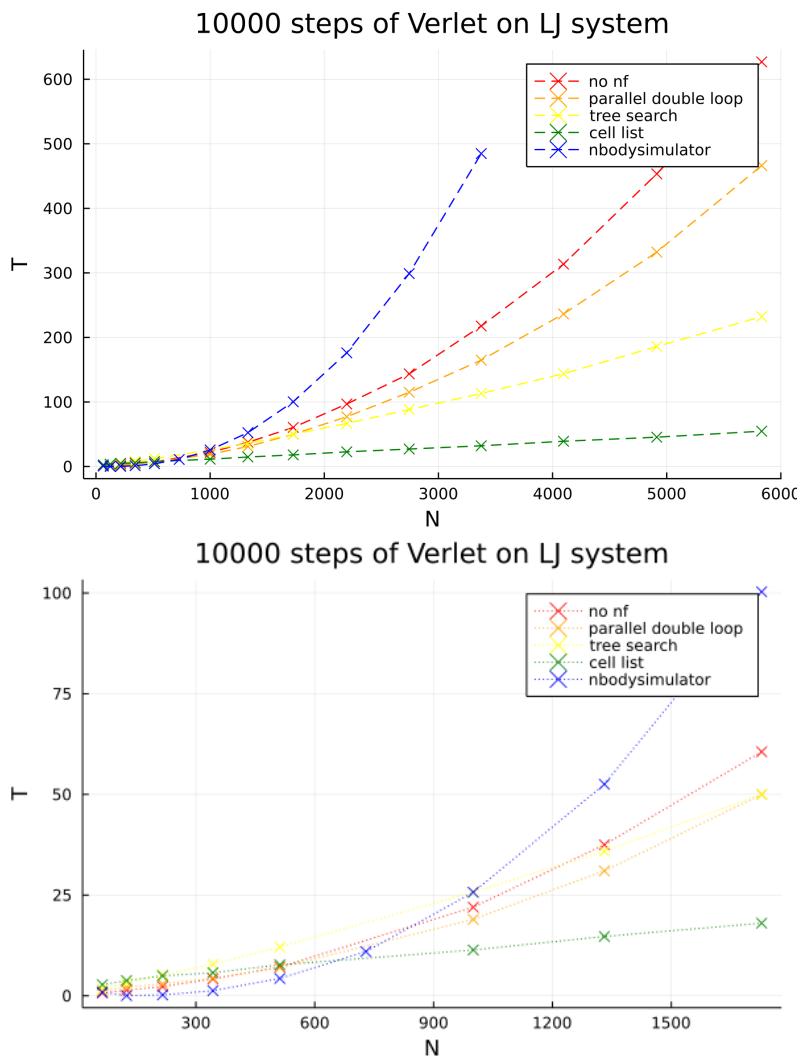


Figure A.1: Comparison of world clock time for the simulation of 10000 steps of a Lennard-Jones system. The "no nf" corresponds to the naive double loop. Top: full curve. Bottom: zoom on the small system regime.

instance once long-range force contributions are implemented, as well as a more flexible interface for force computations. This will be a necessary step in the implementation of barostats, which are important in the context of biomolecular simulation, a use case Molly is particularly attuned to. On the other hand, non-equilibrium methods are unlikely to have a place in a general-purpose MD package. Instead, we might consider publishing in the future a standalone package that extends Molly for transport coefficient computations, once our understanding of the Norton method has cohered.

# Bibliography

- [1] Jeff Bezanson et al. “Julia: A fresh approach to numerical computing”. In: *SIAM review* 59.1 (2017), pp. 65–98. URL: <https://doi.org/10.1137/141000671>.
- [2] R.N. Bhattacharya. “On the Functional Central Limit Theorem and the Law of the Iterated Logarithm for Markov Processes”. In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 60 (2 1982), pp. 185–201.
- [3] N. Bassel. *Internship repository*. 2022. URL: <https://github.com/noebassel/MolecularDynamicsJulia>.
- [4] D.J. Evans and G. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. Cambridge University Press, 2008.
- [5] D.J. Evans et al. “Nonequilibrium molecular dynamics via Gauss’s principle of least constraint”. In: *Physical Review A* 28 (2 1983), pp. 1016–1021.
- [6] M. Fathi, A.A. Homman, and G. Stoltz. “Error Analysis of the Transport Properties of Metropolized Schemes”. In: *ESAIM Proc.* 48 (2015), pp. 341–363.
- [7] M. Fathi and G. Stoltz. “Improving dynamical properties of metropolized discretizations of overdamped Langevin dynamics”. In: *Numerische Mathematik* 136 (2 2016), pp. 545–602.
- [8] Carl Friedrich Gauß. “Über ein neues allgemeines Grundgesetz der Mechanik.” In: (Jan. 1829). DOI: 10.1515/crll.1829.4.232. URL: <https://doi.org/10.1515/crll.1829.4.232>.
- [9] J. Greener. *Molly*. 2020. URL: <https://github.com/JuliaMolSim/Molly.jl>.
- [10] E. Hairer, C. Lubich, and G. Wanner. “Geometric numerical integration illustrated by the Störmer-Verlet method”. In: *Acta Numerica* 12 (2003), pp. 399–450.
- [11] R. Joubaud, G.A. Pavliotis, and G. Stoltz. “Langevin dynamics with Space-Time Periodic Nonequilibrium Forcing”. In: *Journal of Statistical Physics* 158 (2015), pp. 1–36.
- [12] R. Joubaud and G. Stoltz. “Nonequilibrium Shear Viscosity Computations With Langevin Dynamics”. In: *Multiscale Modelling & Simulation* 10 (1 2012), pp. 191–216. DOI: <https://doi.org/10.1137/110836237>.
- [13] B.G. Keller and S. Kieninger. “GROMACS Stochastic Dynamics and BAOAB are equivalent configurational sampling algorithms”. URL: <https://arxiv.org/abs/2204.02105v1>.
- [14] W. Kliemann. “Recurrence and Invariant Measures for Degenerate Diffusions”. In: *The Annals of Probability* 15 (2 1987), pp. 690–707.
- [15] B. Leimkuhler and C. Matthews. *Molecular Dynamics With Stochastic and Deterministic Methods*. Interdisciplinary Applied Mathematics. Springer, 2015.
- [16] B. Leimkuhler, C. Matthews, and G. Stoltz. “The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics”. In: *IMA Journal of Numerical Analysis* 36 (1 2016).
- [17] H. Rodenhausen. “Einstein’s relation between diffusion constant and mobility for a diffusion model”. In: *Journal of Statistical Physics* 55 (1989), pp. 1065–1088.
- [18] M. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford University Press, 2010.

- [19] M. Vaganov and C. Rackauckas. *NBodySimulator*. 2018. URL: <https://github.com/SciML/NBodySimulator.jl>.