The error in the invariant measure of numerical discretization schemes for canonical sampling of molecular dynamics

Charles Matthews

Doctor of Philosophy
University of Edinburgh
2013

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Charles Matthews*)

Contents

Abstract **1**

1 Introduction **2**

2 Deterministic molecular dynamics **9**
2.1 Hamiltonian dynamics 9
2.1.1 General numerical methods for autonomous ODEs 12
2.1.2 Symplectic maps 17
2.1.3 Geometric integration 22
2.1.4 Higher order symplectic methods 32
2.1.5 Scale separation 36
2.2 Sampling the canonical ensemble 40
2.2.1 Statistical mechanics 40
2.2.2 The Boltzmann-Gibbs distribution 44
2.2.3 The Nosé-Hoover thermostat 50

3 Stochastic thermostats **54**
3.1 Stochastic dynamics 54
3.1.1 The Fokker-Planck equation 54
3.1.2 Hörmander's condition 58
3.2 Nosé-Hoover Langevin dynamics 60
3.3 Non-Newtonian dynamical sampling 63

4 Splitting methods for Langevin dynamics **68**
4.1 Foundations for stochastic splitting 68
4.1.1 Preliminaries 69
4.1.2 Decomposition into pieces A, B and C 71
4.1.3 Computing error estimates for general systems 82
4.2 Performance for harmonic systems 96
4.2.1 Quadratic potential energy functions 96
4.2.2 Perturbed harmonic oscillator 99
4.2.3 Stability for systems with multiple scales 101
4.3 The overdamped regime 105
4.3.1 Error estimates in the limit of infinite friction 105
4.3.2 Large and finite friction 113
4.3.3 Numerical experiment 122

5 Applications in molecular sampling **124**
5.1 Error estimation using linear response techniques 124
5.2 Modified Langevin dynamics equation 130
5.3 Application of Langevin dynamics schemes to configurational sampling . 132
5.3.1 A description of the alanine dipeptide molecule 136
5.3.2 Numerical results for vacuum simulation 138
5.3.3 Numerical results from simulations using solvent 141

Summary and conclusion **145**

Bibliography **147**

A Numerical Methods **157**

Abstract

Molecular dynamics (MD) computations aim to simulate materials at the atomic level by approximating molecular interactions classically, relying on the Born-Oppenheimer approximation and semi-empirical potential energy functions as an alternative to solving the difficult time-dependent Schrödinger equation. An approximate solution is obtained by discretization in time, with an appropriate algorithm used to advance the state of the system between successive timesteps. Modern MD simulations simulate complex systems with as many as a trillion individual atoms in three spatial dimensions.

Many applications use MD to compute ensemble averages of molecular systems at constant temperature. Langevin dynamics approximates the effects of weakly coupling an external energy reservoir to a system of interest, by adding the stochastic Ornstein-Uhlenbeck process to the system momenta, where the resulting trajectories are ergodic with respect to the canonical (Boltzmann-Gibbs) distribution. By solving the resulting stochastic differential equations (SDEs), we can compute trajectories that sample the accessible states of a system at a constant temperature by evolving the dynamics in time. The complexity of the classical potential energy function requires the use of efficient discretization schemes to evolve the dynamics.

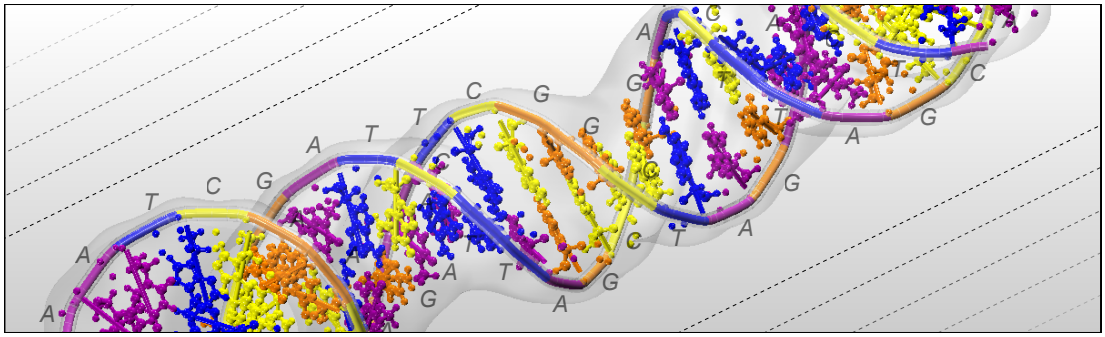
In this thesis we provide a systematic evaluation of splitting-based methods for the integration of Langevin dynamics. We focus on the weak properties of methods for configurational sampling in MD, given as the accuracy of averages computed via numerical discretization. Our emphasis is on the application of discretization algorithms to high performance computing (HPC) simulations of a wide variety of phenomena, where configurational sampling is the goal.

Our first contribution is to give a framework for the analysis of stochastic splitting methods in the spirit of backward error analysis, which provides, in certain cases, explicit formulae required to correct the errors in observed averages. A second contribution of this thesis is the investigation of the performance of schemes in the overdamped limit of Langevin dynamics (Brownian or Smoluchowski dynamics), showing the inconsistency of some numerical schemes in this limit. A new method is given that is second-order accurate (in law) but requires only one force evaluation per timestep.

Finally we compare the performance of our derived schemes against those in common use in MD codes, by comparing the observed errors introduced by each algorithm when sampling a solvated alanine dipeptide molecule, based on our implementation of the schemes in state-of-the-art molecular simulation software. One scheme is found to give exceptional results for the computed averages of functions purely of position.

Chapter 1

Introduction



Molecular dynamics (MD) is a technique for probing the behaviour of aggregations of matter using computer simulation. The form of the aggregation is dependent on the discipline under study, but in general we imagine a collection of individual atoms, represented by point masses, whose instantaneous positions and momenta are stored to high accuracy in a machine. This disregard for the uncertainty principle may be concerning, but one can rest assured that we will in fact disregard almost all quantum effects equally.

Our best understanding of how molecular systems evolve and interact at the atomic level is given by quantum mechanics and the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi = \mathcal{H}\Psi,$$

for Planck's constant $\hbar$, imaginary unit i and with the system's global wavefunction (or instantaneous state function) Ψ and Hamiltonian operator $\mathcal{H}$. The state Ψ is a function of the positions of all of the electrons and atomic nuclei in the system, as well as time. In practice, the analytic solution to this partial differential equation (PDE) is known only for a smattering of small, isolated cases. Numerical treatment of the problem is also overwhelmed by the high dimensionality of the space in which the equation is posed.

If we wish to predict or interpret the changes in a large molecular system over time, outside of special cases where significant reduction is possible, the Schrödinger equation

as it stands is simply too complex an object for us to tackle.

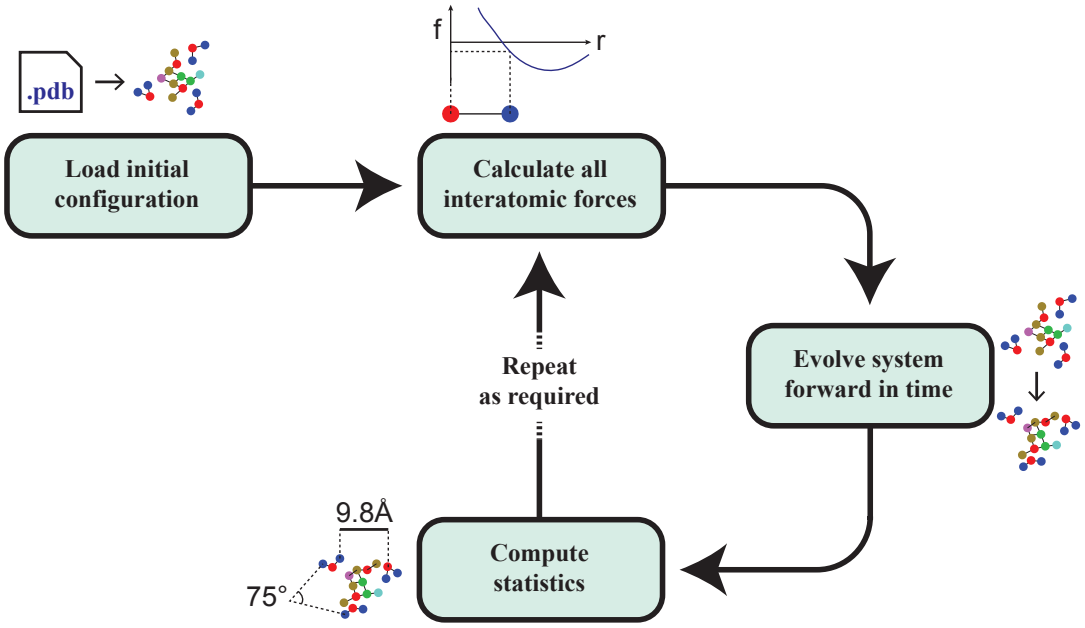
We instead gain traction in the problem by making the Born-Oppenheimer approximation. We first assume that we can treat atomic and electronic motions as independent, owing to the enormous difference in mass between a proton and electron. Second, we abandon any computations involving the individual electrons, in favour of approximating their influence with a potential energy function. The movement of atoms becomes classical–moving in space according to Newton's equations of motion, like billiard balls on a pool table. Quite unlike billiard balls however, forces felt are not due to elastic collisions between particles (the atoms have no physical radius), but instead come from traversing a potential energy landscape derived either from the electronic part of the Schrödinger equation, or fitted from experimental data (or even chosen for computational efficiency as in the case of the Lennard Jones potential). Internuclear interaction within the model is derived from the potential energy itself: states with a lower potential energy are favoured and higher potential energy regions (such as nuclei being in very close proximity) are shunned.

For large systems, the colossal number of intramolecular force terms gives an extremely complex potential energy landscape, with a very large number of metastable minima. Therefore it is no surprise that if we are interested in the system's evolution between two states (or regions in the phase space), we may have to wait for a long time to observe this change just by evolving the equations of motion.

As we have assumed the system's behaviour is classical, its behaviour in time is governed by Newton's second law, just as the celestial motion of planets and stars. The evolution is wholly described by a set of ordinary differential equations (ODEs), though in general we still cannot solve the equations analytically the underlying structure of the ODEs makes them considerably more amenable to numerical treatment than the original Schrödinger equation. The positions of all of the atoms q, along with their momenta p, is assumed to be known exactly and stored in a computer. Solution trajectories are approximated by computing the total forces on all atoms at any one instant, giving the atoms a kick in the direction of its force, and then allowing all the atoms to drift for a small step in time. Assuming the time steps are chosen small enough (typically of the order of a femtosecond: 10^{-15}s), the results of these MD simulations yield a good approximation to the dances of the individual atoms, and allow the simulator to elucidate the behaviour of the system under study.

The dominant computational cost in the simulation is in the force calculation (required at every step), although this is ironically the mathematician's biggest asset when studying molecular dynamics methods. It grants a license to devise all manner of algorithms for evolving the dynamics, as the only cost that matters in simulation is the evaluation of the instantaneous forces.

The dramatic increase in available scientific computer power after World War II made molecular dynamics simulations accessible and practical for many material scientists, with the first studies probing the behaviour of small clusters of particles (modelled



The standard workflow of molecular dynamics.

as hard spheres) using computers based at Los Alamos [2, 3, 84, 128]. Computer simulation of more realistic models (using more complicated potential energy functions) followed swiftly [46]. Whereas Rahman's original trailblazing experiment [96] used 864 atoms (with each timestep taking around 45 seconds), the scope of modern MD is considerably more vast, employing many billions [61] or even a trillion particles [45] in simulations. Molecular dynamics simulations are prevalent in biology, chemical physics, material science, engineering and a plethora of other scientific disciplines. It is frequently used as a tool to test new theories in place of physical experiments (describing the computer simulations as numerical experiments) which would otherwise be too costly or impermissible due to the constraints of realizing it in a lab (for example applications of extreme temperature or pressure). By contrast to physical experiments, the limits on numerical experiments are proportional to the computer hardware available.

Such limits restrict the amount of data we can reasonably store, and hence the size of systems (in terms of the dimensionality or number of atoms) we can tackle in a numerical experiment. The speed and architecture of the computer will decide how many time steps we can take, and hence dictate the time interval we can observe, as we cannot push the timestep beyond certain limits without risking instability in our propagation algorithm.

Molecular dynamics has surfed the wave of computer revolution for the past decades, enabling larger and more complex simulations to be run with upgraded hardware and little change in algorithms and approaches [24, 53]. The relation between processor speed and power consumption (with Dennard's scaling law) puts paid to this philosophy–current machine architectures have hit fundamental physical constraints,

meaning computers do not get faster, only more numerous [34].

This presents a problem for molecular dynamics. The most computationally expensive part of the simulation is (and has always been) the calculation of the forces between the atoms at any instant. While these forces are well suited to parallel computation, the simulation itself is sequential and temporally contiguous: all atoms must be moved together [23, 92, 95]. This results in repeated reduction operations, forcing data to be shunted between physical parts of the computer, which itself gives rise to latency issues slowing simulation. Parallelism can offer gains, but it is no magic bullet.

Breakthroughs in the field are unlikely to come from advancing technical specifications of hardware, but instead new algorithms (tuned to particular applications) are the driving force behind the progress in efficient numerical simulation. Currently, many of the algorithms in general use for generating system trajectories have changed very little since early numerical work. The Verlet algorithm, ubiquitous in modern MD codes, can be traced back to geometrical ideas in Newton's original *Principia Mathematica* (see the discussion in [48, Section I.2.1]).

While the use of MD has always been guided by the utility and practical nature of simulations, in the late 1960s many began to question the usefulness and futility of attempting to approximate atomic trajectories for realistic applications [40]. The chaotic nature of high dimensional systems means that trajectories that start close together diverge exponentially fast in time, so small errors in the trajectory (introduced from the very first step of an algorithm) will compound to render our simulation inaccurate very quickly. As a result, the conservation of the total system energy is destroyed. Wnergies reported in simulation exhibit oscillations or drift, which threatens the veracity of the simulation itself [33, 40, 127].

A revolution came in 1983 after pioneering work by Ruth [101] introduced the notion of geometric integration and symplecticness first present in work by de Vogelaere [26]. Subsequent work by Channell [21], Menyuk [83], Suris [113] and Feng [36, 37, 38], followed by work in physics and numerical analysis, laid the groundwork for algorithms that preserve a *nearby* energy function exactly, and hence maintain the desired qualitative (geometrical) qualities of the original system under study.

But are maintaining these qualitative properties the key? While beneficial, far more useful would be the ability to relate simulation data to quantitative results pertaining to the molecular models being studied. However, as we are evolving a chaotic system in time, we expect that each successive point of the approximated trajectory will be further away from the true trajectory than the previous point. This leads to a change of perspective about our MD simulations: rather than attempting to approximate the exact motion in time, we instead view our trajectory as a representative sample of the average behaviour of the system (or a system with energy function very close to the one we are interested in).

The challenge is then to accurately recover these averages from computed trajectories, and correct any errors we have introduced through our imperfect evolution of the

system.

Our motivation in this thesis will be to develop new algorithms for computing canonical (constant temperature) averages in molecular dynamics. Given some system we are interested in, we imagine immersing it in a large heat bath at some fixed temperature. The heat bath interacts with the target system of interest, equilibrating it to a matching temperature. For many applications this is a more realistic setting for simulation than considering the system sealed off in a vacuum with constant energy.

Simulating the heat bath itself presents a wealth of problems. We can choose to simulate a large number of particles outwith our target system, and imbue an interaction potential between them to facilitate energy exchange. Of course, the computational cost of the simulation balloons as the large number of additional particles introduced causes the evaluation of the interatomic forces to become significantly more expensive. Ideally we would prefer to approximate the heat bath particle's interactions, as we are really only interested in the behaviour of the target system (which will be dimensionally much smaller). A common strategy is to omit the heat bath completely from the simulation, and instead approximate its interaction through the use of an auxiliary device requiring a significantly smaller dimensional footprint, alleviating the volume of computation required per timestep.

There are many forms of auxiliary device that can be shown to perturb the evolution of the atoms so that computing an average along the simulation trajectory (such as the average total potential energy) will give the same answer as if we sampled the system while it was in equilibrium with the heat bath. The methods we shall be considering for canonical (constant temperature) sampling will be dynamics-based, perturbing Newton's constant-energy equations in a manner consistent with the approximation of the heat bath interactions. For obvious reasons, the dynamical equations that give commensurate sampling to constant-temperature simulations are called *thermostats*. They can be implemented deterministically, but in general it is more reliable to introduce a stochastic perturbation, transforming the ODEs into stochastic differential equations (SDEs). The resulting SDEs will be just as insoluble as in the deterministic case,

We will be primarily motivated by applications in biomolecular modelling that involve the spectrum of complex nonlinear characteristics (such as long-ranged forces, steep potentials and stiff harmonic components). However, the molecular dynamics techniques we develop will be no less helpful for other examples. We shall think of our system of interest as being a collection of atoms modelling a protein, with a scientist seeking to calculate average values of some function of the position or momentum of the system (such function we will call observables, for example the total kinetic energy). Additionally the system often includes a soup of atoms surrounding the protein (typically water molecules), replicating the agitating effects of the blood or solvent.

This thesis offers a new strategy for the development and analysis of numerical methods for the stochastic differential equations that appear in molecular dynamics thermostats. The focus will be on observed long-time averages computed from trajec-

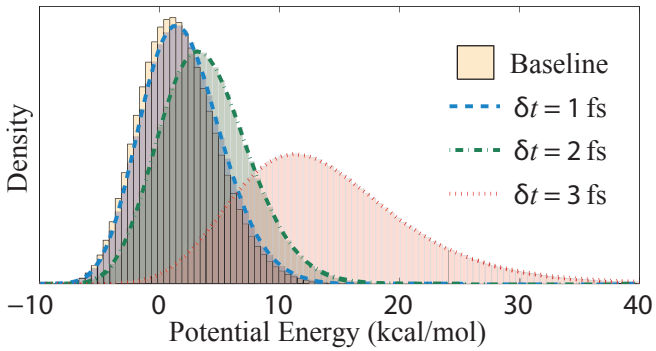
tory data obtained through the numerical schemes. Wherever possible we will supplement theory with numerical examples demonstrating the results for model problems.

Chapter 2 presents our notation and reviews necessary background for the analysis of deterministic numerical methods. The concepts of symplecticness and geometric integration will be introduced, before turning to statistical mechanics and ergodicity. We develop the ideas behind using trajectories to sample an ensemble corresponding to a particular distribution, leading to a change of perspective on the goals of the simulation. The use of thermostats to sample a system at constant temperature is discussed, while a lack of ergodicity is demonstrated (at reasonable parameter choices) for the popular deterministic Nosé-Hoover chains thermostat. These results provide motivation for stochastic dynamics as an alternative giving ergodic and efficient sampling.

The subsequent chapter develops stochastic (Itō) calculus, giving details on the convergence of distributions and the concept of a spectral gap. We introduce stochastic thermostats based on perturbations of Newtonian dynamics and prove ergodicity using a Hörmander condition. Both Langevin and Nosé-Hoover Langevin dynamics will be discussed, with examples of thermostats resulting from perturbing non-Newtonian dynamics.

The fourth chapter builds on the splitting approach for deterministic methods, introduced in Chapter 2. We detail a number of available splitting strategies for Langevin dynamics present in the literature, before extending the GLA splitting scheme of [12] to create a new class of method. One such new method is found to have a fortunate cancellation at leading order, making it considerably more efficient than existing schemes. We examine the specific case of harmonic potential energy functions, that is perticularly relevant for molecular dynamics applications. We also consider the case of overdamped Langevin (Brownian) dynamics, and discuss the behaviour of the resulting limiting schemes. The focus of this chapter is to rigorously ground our work originally published in [68].

Chapter 5 utilizes linear response techniques to numerically compute and correct the



The computed potential energy distribution for a 1ns simulation of alanine dipeptide in a vacuum at 300K. The timestep used to evolve the simulation has a dramatic effect on the resulting distribution (the baseline result comes from a simulation with extremely small timestep). The differences are the result of errors due to discretization.

error in observed averages, through the evaluation of integrated correlation functions (similar to Green-Kubo-type formulas). We additionally consider intentionally perturbing the equations of motion in an effort to provide higher–order sampling possibilities for a particular observable. We additionally summarize our work in [69], where we apply the newly developed methods to model problems and compare the performance of the algorithms with those already present in the literature, for a small solvated molecular problem.

We conclude this thesis by summarizing our results and presenting our perspective on their potential impact.

Chapter 2

Deterministic molecular dynamics

In this chapter we provide a review of standard mathematical tools for the analysis of dynamical systems and numerical schemes. In particular we introduce the concept of the numerical scheme preserving a perturbed quantity exactly, and illustrate the benefits afforded to such schemes in the deterministic setting. This is a central theme of the thesis that we revisit in Chapter 4 in the context of stochastic dynamics.

We explain the benefits of using splitting schemes in the context of symplecticness, before introducing our notation for splitting methods (which we shall rely heavily on in later chapters). We finish the chapter by changing our perspective on our numerical method: namely that we think of our trajectory as sampling a distribution, rather than necessarily giving us meaningful dynamical information.

The material covered in this section can be found in most textbooks on classical and statistical mechanics, in particular we present material from [4, 6, 42, 48, 72, 74, 104, 106].

2.1 Hamiltonian dynamics

The class of problems we shall consider in this thesis are canonical Hamiltonian systems in an even dimensional phase space. Consider n atoms existing in d-dimensional space, where for most applications $d = 3$, but in simple models we may choose to lower the dimensionality, confining the atoms to the plane or a line. The i^{th} particle's instantaneous position is given by the vector

$$q_{(i)} = \left[q_{(i,1)}, q_{(i,2)}, \ldots, q_{(i,d)}\right] \in \mathbb{R}^d,$$

though we will more often refer to the position data of all of the system's particles, given in a compact form as the vector $q = \left[q_{(1)}, \ldots, q_{(n)}\right]^T \in \mathbb{R}^N$ for total position dimension $N = dn$. The i^{th} particle's instantaneous momenta (its mass m_i multiplied

by its velocity $\dot{q}$) is similarly denoted $p_{(i)}$. Any particle's momentum is independent of its position, with the overall state of the entire system at any one instant denoted $z = [q, p]^T \in \mathbb{R}^{2N}$.

We call $\Omega = \mathbb{R}^{2N}$ the phase space of the system, representing all possible states of the system. What a molecular scientist is typically interested in is averages of functions of z with respect to the evolution of the system in phase space. What we will be concerned with in this thesis is the evolution of autonomous Hamiltonian systems, corresponding to the initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}t}z = J\nabla H(z), \quad z(0) = z_0, \tag{2.1}$$

where $H(z) : \Omega \to \mathbb{R}$ is at least twice continuously differentiable on some open subset. The skew-symmetric structure matrix $J \in \mathbb{R}^{2N\times 2N}$ is given as

$$J = \begin{bmatrix} 0 & I_N \\ -I_N & 0 \end{bmatrix}, \tag{2.2}$$

where I_k is the $k \times k$ identity matrix. Differentiating $H(z)$ with respect to time and using (2.1) we find

$$\frac{\mathrm{d}}{\mathrm{d}t}H(z) = (\nabla H(z))^T J \nabla H(z) = 0,$$

and so $H(z) = \text{const}$ is a conserved quantity of the system for all initial conditions, corresponding to the conservation of the total system energy.

In molecular dynamics, the Hamiltonian is typically separable into functions of position and momentum,

$$H(z) = T(p) + U(q), \tag{2.3}$$

where $U(q) : \mathbb{R}^N \to \mathbb{R}$ is the instantaneous potential energy of the system and $T(p)$ is the kinetic energy, given as

$$T(p) = \frac{p^T M^{-1} p}{2},$$

with $M = \mathrm{diag}(m_1 I_d, m_2 I_d, \ldots, m_n I_d)$ the diagonal mass matrix. The system (2.1) hence reduces to Newton's classical equations of motion

$$\begin{aligned} \frac{\mathrm{d}q}{\mathrm{d}t} &= M^{-1}p, \\ \frac{\mathrm{d}p}{\mathrm{d}t} &= -\nabla U(q), \end{aligned} \tag{2.4}$$

which can be interpreted as a rewriting of Newton's second law

$$\begin{array}{ccccc} \text{mass} & \times & \text{acceleration} & = & \text{force}, \\ M & \times & \dfrac{\mathrm{d}^2 q}{\mathrm{d}t^2} & = & -\nabla U(q). \end{array}$$

The ODEs (2.4) will describe the entirety of the evolution of our system, and it may be

surprising that all of the physics, chemistry and flavour we wish to simulate must be encapsulated solely within the potential energy function $U(q)$. It is therefore often a very complicated *semi-empirical* function of all N variables, defined from experimental data or from solutions to the Schrödinger equation. Typically in MD, we think of the force as a 'black box' that we have little control over. In simulation, a force field parameter file is used to parameterize the force terms appropriately for simulation.

As an example, consider a single particle in one dimension (so $n = d = 1$) with unit mass, evolving with respect to the Hamiltonian (2.3) with the potential energy function given as

$$U(q) = \frac{1}{2}\omega^2 q^2,$$

where $\omega > 0$ is a parameter and $q \in \mathbb{R}$ is a scalar. This potential corresponds to the one-dimensional harmonic oscillator, often used to approximate bond stretches in atomic systems, with governing ODEs

$$\begin{aligned} \frac{\mathrm{d}q}{\mathrm{d}t} &= p, \\ \frac{\mathrm{d}p}{\mathrm{d}t} &= -\omega^2 q, \end{aligned} \tag{2.5}$$

which are Hooke's familiar equations of motion for a particle connected to the origin by a linear spring, with spring constant ω^2. We can rewrite this as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix}\begin{bmatrix} q \\ p \end{bmatrix},$$

which, given an initial value $[q(0), p(0)]$, we can solve to find oscillatory solutions

$$\begin{bmatrix} q(t) \\ p(t) \end{bmatrix} = \begin{bmatrix} \cos(\omega t) & \omega^{-1}\sin(\omega t) \\ \omega^{-1}\sin(\omega t) & \cos(\omega t) \end{bmatrix}\begin{bmatrix} q(0) \\ p(0) \end{bmatrix},$$

with frequency ω. We know these solutions will lie along curves of constant energy

$$H(z(t)) = \frac{1}{2}p(t)^2 + \frac{1}{2}\omega^2 q(t)^2 = \frac{1}{2}p(0)^2 + \frac{1}{2}\omega^2 q(0)^2, \tag{2.6}$$

which give a family of concentric ellipses in the phase space $\Omega = \mathbb{R}^2$, parameterized by the initial condition and centred on the origin.

It is perhaps difficult to envisage such a small model with such a mathematically well-understood solution to have much significance or applicability to molecular modelling, though in practice under the assumptions of the classical nature of atomic motions such harmonic potential functions are ubiquitous. A bond stretch between (bonded) atoms i and j is assumed to have a linear restoring force, and hence due to the linearity of the ∇ operator in (2.4), this force can be added to a molecular model by adding the harmonic potential energy to the Hamiltonian (where it becomes part of the total potential energy).

The total potential energy of the system is chosen to be a superposition of functions that best fit some prescribed experimental data. Given the natural system one wishes to emulate, the laws of physics are essentially 'plug and play' in the sense that additional potential energy terms can be summed to tailor the equations of motion to the specific elements or environmental conditions involved. In the framework we have described, $U(q)$ is a blank canvas for whatever physical phenomena one can dream up.

If we wish to describe a physical system, the forces (and hence potential energies) involved in MD simulations will usually be some functions of the distances between atoms, although some static (non-moving) boundary terms can also be included. Most numerous are pair potentials, giving either the relative potential energy from a bonded pair (usually coming from the harmonic oscillator term) or the non-bonded intermolecular forces (such as the van der Waals or Coulombic charge terms).

Additionally, three or many-body potentials may be added to the system to simulate structural features in the system, such as angular predispositions.

2.1.1 General numerical methods for autonomous ODEs

The high dimensionality and nonlinearity of our evolution equations (2.1) make numerical treatment inevitable, as exactly-solvable cases such as (2.5) seldom occur. Consider the initial value problem,

$$\frac{\mathrm{d}}{\mathrm{d}t}z = f(z), \qquad z(0) = z_0 \in \mathbb{R}^{2N}, \tag{2.7}$$

with solution written as $z(t; z_0)$ at time t. We define the flow map, or solution operator, for this system to be the map $\Phi_t : \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ such that

$$\Phi_t(\eta) = z(t; \eta),$$

which takes a point η in the phase space and evolves it forward t units of time, with respect to the flow of (2.7). The flow map also has a semigroup property; evolving the flow successively by a time t then s is equivalent to evolution by time $t + s$, such that

$$\Phi_t \circ \Phi_s = \Phi_{t+s}.$$

The flow map can be extremely challenging to compute analytically even for simple, small dimensional examples. While being able to write Φ_t down would allow us to sample as many trajectories as we can find initial conditions, we are fortunate in that molecular dynamics applications do not require such fidelity. The numerous approximations and simplifications we make in formulating the classical case permits some margin of error in our computed solutions.

Instead, let us consider a multitude of discrete points along a single trajectory, equally spaced in time, like the slits around a zoetrope or a collection of frames in a strip of film. We fix the time between snapshots and define this as our timestep $\delta t > 0$, with a

discretized trajectory written as the ordered set $\{z(k\,\delta t; z_0)\}_{k=0}^{\infty} = \{\Phi_{\delta t}^k(z_0)\}_{k=0}^{\infty}$, where the notation Φ^k represents repeated composition of the flow map k times. Our goal is to approximate the points in the discretized trajectory, writing successive approximations as $z_k \approx z(k\,\delta t; z_0)$.

Time	0	δt	$2\,\delta t$	$\cdots$	$k\,\delta t$	$\cdots$
Solution	$z(0; z_0) = z_0$	$z(\delta t; z_0)$	$z(2\,\delta t; z_0)$	$\cdots$	$z(k\,\delta t; z_0)$	$\cdots$
Flow	$\Phi_{\delta t}^0(z_0) = z_0$	$\Phi_{\delta t}(z_0)$	$\Phi_{\delta t}^2(z_0)$	$\cdots$	$\Phi_{\delta t}^k(z_0)$	$\cdots$
Approximation	z_0	z_1	z_2	$\cdots$	z_k	$\cdots$

We think of our successive approximations as being generated through the iteration of a map $\widehat{\Phi}_{\delta t} : \Omega \to \Omega$, such that

$$z_k = \widehat{\Phi}_{\delta t}(z_{k-1}) = \widehat{\Phi}_{\delta t}^k(z_0).$$

The mapping $\widehat{\Phi}$ will necessarily depend on the system dynamics (2.7) and hence will be dependent on the (typically nonlinear) vector field f. Integrating (2.7) between times t and $t + \delta t$, we find

$$z(t + \delta t; z_0) = z(t; z_0) + \int_t^{t+\delta t} f(z(\tau; z_0))\,\mathrm{d}\tau, \tag{2.8}$$

where evaluation of the integral $\int_t^{t+\delta t} f(z(\tau; z_0))\,\mathrm{d}\tau$ by quadrature or interpolation gives an update scheme. For instance, the left-endpoint approximation to the integral

$$\int_t^{t+\delta t} f(z(\tau; z_0))\,\mathrm{d}\tau \approx \delta t\, f(z(t; z_0)),$$

yields the explicit Euler (eE) scheme, given as

$$z_{n+1} = z_n + \delta t f(z_n). \tag{eE}$$

Whereas using a right-endpoint approximation instead gives yields the implicit Euler scheme (iE)

$$z_{n+1} = z_n + \delta t f(z_{n+1}). \tag{iE}$$

We can see that the implicit Euler scheme makes it slightly more challenging to compute the update for nonlinear $f(z)$, as we must use additional computation to solve (iE) implicitly. Newton or fixed point iteration are practical options to find a solution, but this extra computational work makes implicit schemes often less efficient than explicit schemes.

Endpoint approximations provide a suitable starting point, but a wealth of numerical quadrature schemes are available to estimate the integral in (2.8), so it makes sense at this point to discuss the quantification of error in our simulation. The simplest way to grade the quality of our approximation is to consider the difference between the

solution and approximation at an arbitrary point $\eta \in \Omega$, written as

$$\left|\Phi_{\delta t}(\eta) - \widehat{\Phi}_{\delta t}(\eta)\right|,$$

which we call the local error of the scheme. For the explicit Euler scheme (eE), the local error is

$$\left|\Phi_{\delta t}(\eta) - \widehat{\Phi}_{\delta t}(\eta)\right| = \left|z(\delta t; \eta) - (\eta + \delta t f(\eta))\right|,$$

whence utilizing a Maclaurin series expansion of the form

$$z(\delta t) = z(0) + \delta t \dot{z}(0) + \frac{1}{2}\delta t^2 \ddot{z}(0) + \dots$$

with the equation $\dot{z}(t) = f(z(t))$, the local error at η becomes

$$\delta t^2 \left|\frac{1}{2} f'(\eta) f(\eta) + \mathcal{O}(\delta t)\right|.$$

If the local error term is of order δt^s, then we say that the scheme is of local order s. Hence explicit Euler is a scheme of local order 2.

As our goal is to compute a long trajectory, the local error is perhaps not readily useful to us. What matters more will be how these errors compound over successive iterations, affecting the long-term behaviour of z_k as k gets large. The global error of a scheme after k steps, denoted e_k, is defined as the error in the scheme at the time $k\delta t$, written as

$$\mathrm{e}_k := \left|\Phi^k_{\delta t}(\eta) - \widehat{\Phi}^k_{\delta t}(\eta)\right| = \left|z(k\delta t; z_0) - z_k\right|.$$

This can be difficult to quantify in generality. For Euler's method, we have

$$\begin{aligned}
\mathrm{e}_{k+1} &= \left|z((k+1)\delta t; z_0) - z_{k+1}\right|, \\
&= \left|z(k; z_0) + \delta t \dot{z}(k\delta t; z_0) - (z_k + \delta t f(z_k)) + \frac{1}{2}\delta t^2 \ddot{z}(\tau_k; z_0)\right|, \\
&= \left|z(k; z_0) - z_k + \delta t \left[f(z_k; z_0)) - f(z_k)\right] + \frac{1}{2}\delta t^2 \ddot{z}(\tau_k; z_0)\right|, \qquad (2.9)
\end{aligned}$$

for some $\tau_k \in [k\delta t, (k+1)\delta t]$. If we assume that $f(z)$ obeys a Lipschitz condition

$$|f(\mu) - f(\eta)| \leq L|\mu - \eta|,$$

for some constant $L > 0$, then applying the triangle inequality to (2.9) gives

$$\mathrm{e}_{k+1} \leq (1 + \delta t L)\mathrm{e}_k + \frac{1}{2}\delta t^2 |\ddot{z}(\tau; z_0))|.$$

Using the relation $(1 + \delta t L)^n \leq e^{n\delta t L}$, we can solve the recurrence to get a global error bound of the form

$$\mathrm{e}_{k+1} \leq K e^{TL} \delta t,$$

where $T = k\delta t$ is the total time of integration, and K is a constant that depends on the vector field f and the initial condition z_0. In practice K may be large, but this bound does give us a quantitative strategy to reduce error; given a fixed time interval, we can choose our step size sufficiently small such that the global error in our solution is below any desired threshold. However, we will expect exponential divergence away from our exact solution: the error grows exponentially when increasing the time interval T. A scheme is of order s if it has global error proportional to δt^s, thus the explicit Euler scheme is first-order (as T is fixed). It can be shown that if a scheme has local error $\mathcal{O}(\delta t^{p+1})$, then it has global order p [72, 104].

Given the global error in trajectories, it is natural to ask what happens to observable functions (such as the energy) under the discretized dynamics. One important example is the behaviour of first integrals, which are non-constant functions whose derivative vanishes on solutions to the ODEs. If we define $I : \mathbb{R}^{2N} \to \mathbb{R}$ to be a first integral of the system (2.7), then by the chain rule

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} I(z) = \nabla I(z) \cdot \dot{z} = \nabla I(z) \cdot f(z), \tag{2.10}$$

for all $z \in \Omega$. Using an order p method and propagating forward in time, we can apply the mean value theorem to show that

$$I(z_k) = I(z_0) + \nabla I(\widetilde{z}) \cdot |z_k - z_0|,$$

for some intermediary point $\widetilde{z}$ on the line connecting z_k and z_0. If we assume that $\nabla I(z)$ is bounded in some open domain, such that

$$|\nabla I(\widetilde{z})| < \widetilde{K},$$

for some constant $\widetilde{K} > 0$, then we have

$$|I(z_k) - I(z_0)| \leq \widetilde{K} \mathrm{e}_k,$$

and hence

$$|I(z_k) - I(z_0)| \propto e^{TL} \delta t^p.$$

Therefore for a given a fixed time interval, reducing the value of δt will allow the value of $I(z)$ (which is constant for the exact dynamics) to remain in a desired tolerance corridor. This is useful for Hamiltonian systems where the total energy should be preserved under the dynamics. An alternative way to proceed is to increase p, the order of the method. The two Euler methods (iE) and (eE) encountered so far are both of order $p = 1$. We can seek higher order estimations of (2.8), rather than the endpoint approximations used in the Euler schemes. One such technique is the midpoint rule, employing an implicit Euler estimate for the centrepoint of the time interval $z_{n+1/2}$,

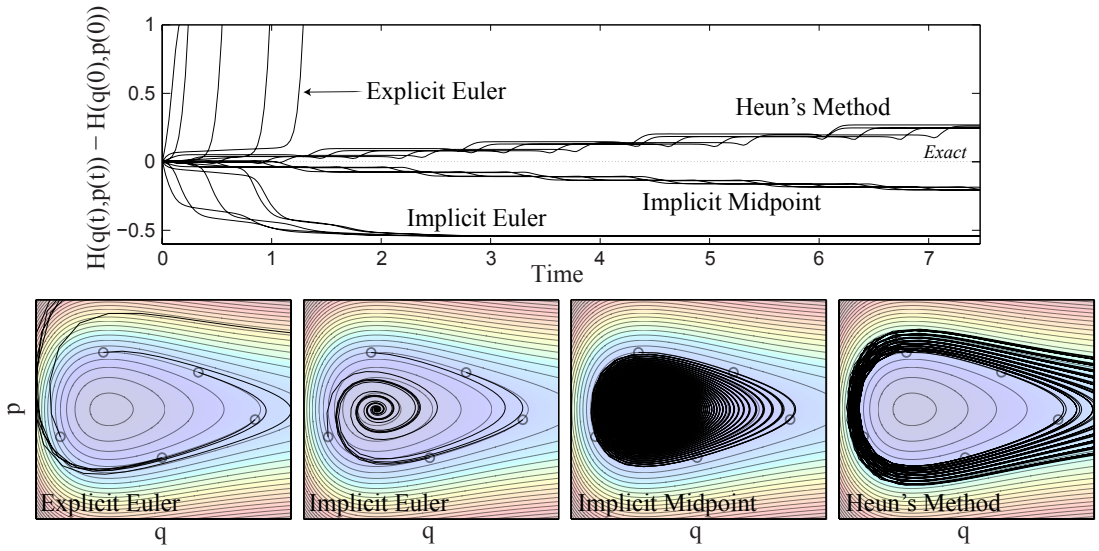


Figure 2.1: Energy drift is demonstrated in numerically computed trajectories for the Lennard-Jones oscillator. Five initial conditions (marked with circles) belonging to the levelset $H(q,p) = 0.45$ are propagated forward in time with timestep $\delta t = 0.03$. Trajectories are plotted in phase space $(q,p) \in \mathbb{R}^2$ for $t \in [0, 300]$.

where $z_{n+1/2} \approx z_n + \delta t f(z_{n+1})/2$. This gives the implicit midpoint rule

$$z_{n+1} = z_n + \delta t\, f(z_n + \delta t f(z_n)/2). \tag{imr}$$

This is a second-order implicit method. Similarly we can use a different second-order approximation of the integral, such as the trapezoidal rule. To estimate the endpoint, we can this time use an explicit Euler approximation for $f(z_{n+1})$, giving Heun's method (Hm) as

$$z_{n+1} = z_n + \delta t \left[f(z_n) + f(z_n + \delta t f(z_n))\right]/2. \tag{Hm}$$

This is now an explicit second-order method. We show the qualitative difference in first and second order methods in Figure 2.1, for the Hamiltonian system $f(z) = J\nabla H(z)$, with Hamiltonian as in (2.3) for $N = 1$. We model the Lennard Jones (6-12) oscillator with total potential energy function given as

$$U(q) = \phi_{\text{lj}}(q) = q^{-12} - 2q^{-6}, \tag{2.11}$$

with scalar $q > 0$. The dynamics has a single equilibrium point at $(q^*, p^*) = (1, 0)$, which solution trajectories will orbit. The results of the experiment show a rapid decay in the energy for the first-order methods as time increases. The second-order methods do much better, with a much slower drift in the energy.

Higher order general methods

We briefly mention that one can consider methods for a higher order approximation of the solutions $z(t; z_0)$ than those given so far, as these will help minimize the global error drift of the numerical solution. The most common class of high order methods is the s–stage Runge-Kutta method,

$$z_{n+1} = z_n + \delta t \sum_{i=1}^{s} b_i f(Z^{[i]}), \quad Z^{[i]} = z_n + \delta t \sum_{j=1}^{s} a_{ij} f(Z^{[j]}),$$

where the constants b_i and a_{ij} define the method. Runge-Kutta methods can be either implicit or explicit depending on the choice of the constants, giving a more general framework for the interpolation of the integral in (2.8). Methods can be constructed to arbitrary order, providing minimal energy drift in a giving time interval. Obviously the total number of stages s used can balloon with the order, so it may be that the required numerical fidelity comes at the cost of a slower running algorithm, due to the high computational cost (equivalently large number of evaluations of $f(z)$).

We can use these general methods to solve Newton's equations (2.4), written in vector form

$$\frac{\mathrm{d}}{\mathrm{d}t} z = \frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} M^{-1} p \\ -\nabla U(q) \end{bmatrix} = J \nabla H(q,p) = f(z),$$

where $H(q,p)$ is the total energy, given in (2.3). The update scheme for the explicit Euler method (eE) becomes

$$\begin{aligned} q_{n+1} &= q_n + \delta t M^{-1} p_n, \\ p_{n+1} &= p_n - \delta t \nabla U(q_n), \end{aligned}$$

giving us an error of order δt in the long-time dynamics (global error) of the system. This is a good starting point, but perhaps given our focus is on Hamiltonian systems, we can cook up a better method by emulating the properties of the exact flow map.

2.1.2 Symplectic maps

Let us consider the evolution of two solutions that start in close proximity. We define the flow map of the Hamiltonian system (2.1) to be $\Phi_{H,t}(z)$, and define

$$\delta z_0 := \widetilde{z_0} - z_0$$

to be a small deviation in an initial condition z_0. The evolution of this difference is given by simply comparing solutions at time t,

$$\delta z(t) = z(t; \widetilde{z_0}) - z(t; z_0) = \Phi_{H,t}(\widetilde{z_0}) - \Phi_{H,t}(z_0).$$

As we have a smooth Hamiltonian, the flow map is also smooth (see [72]), and we can then use a Taylor series expansion to obtain

$$\begin{aligned}\delta z(t) &= \Phi_{H,t}(z_0 + \delta z_0) - \Phi_{H,t}(z_0),\\ &\approx \delta z_0 \frac{\partial}{\partial z}\Phi_{H,t}(z_0),\end{aligned}$$

and hence the quantity

$$\xi(t) := \frac{\partial}{\partial z}\Phi_{H,t}(z(0))\delta z(0)$$

gives the approximation $\delta z(t) \approx \xi(t)$. As we have derived this from a Taylor series, this approximation is only valid for small initial deviations δz_0 and for short time intervals. To investigate the growth of this quantity, we can substitute $z(t) = \Phi_{H,t}(z(0))$ into (2.1), to find

$$\frac{\partial}{\partial t}\Phi_{H,t}(z(0)) = J\nabla H(\Phi_{H,t}(z(0))).$$

This gives

$$\begin{aligned}\frac{\partial}{\partial t}\xi(t) &= \frac{\partial}{\partial t}\left[\frac{\partial}{\partial z}\Phi_{H,t}(z(0))\,\delta z(0)\right],\\ &= \frac{\partial}{\partial z}\left[\frac{\partial}{\partial t}\Phi_{H,t}(z(0))\,\delta z(0)\right],\\ &= \frac{\partial}{\partial z}\left[J\nabla H(\Phi_{H,t}(z(0)))\,\delta z(0)\right],\\ &= JH''(\Phi_{H,t}(z(0)))\left[\frac{\partial}{\partial z}\Phi_{H,t}(z(0))\,\delta z(0)\right],\\ &= JH''(z(t))\,\xi(t),\end{aligned}$$

where $H''(z) := \nabla\nabla^T H(z)$ is the Hessian matrix of mixed partial second derivatives. This is known as the variational equation, written as

$$\frac{\partial}{\partial t}\xi(t) = JH''(z(t))\,\xi(t). \tag{2.12}$$

Other quantities can also satisfy a variational equation. Consider the Jacobian of the flow map $\Phi_{H,t}(z(0))$, which we abbreviate to f_t, i.e. let

$$f_t = \frac{\partial}{\partial z}\Phi_{H,t}(z(0)), \quad \text{where } {}_t : \mathbb{R}^{2N} \to \mathbb{R}^{2N}.$$

We can write $\xi(t) = f_t\delta z_0$, where $f_0 = I_{2N}$ from the definition of the flow map. It can be easily verified that the matrix f_t also satisfies a variational equation

$$\frac{\mathrm{d}}{\mathrm{d}t}f_t = JH''(z(t))\,f_t. \tag{2.13}$$

Variational equations are an important class of problems, and we will make use of these results to aid our understanding of the geometrical properties of the flow map.

Differentiating $H(z)$ with respect to time, recalling the formula for a first integral

(2.10), it is easy to see that the Hamiltonian is invariant under the flow $\Phi_{H,t}(z_0)$, i.e. that it is itself a first integral of the dynamics (2.1). We will now study an important geometrical property of Hamiltonian flow maps, relating to the preservation of area.

Consider two vectors $\mu, \eta \in \Omega$ in phase space, where

$$\mu = \begin{bmatrix} \mu_q \\ \mu_p \end{bmatrix}, \quad \eta = \begin{bmatrix} \eta_q \\ \eta_p \end{bmatrix},$$

with $(\mu_{q,(1)}, \ldots, \mu_{q,(N)})^T =: \mu_q \in \mathbb{R}^N$ and similarly defined $\mu_p, \eta_q, \eta_p \in \mathbb{R}^N$. If we suppose $N = 1$, then μ and η reduce to two planar vectors, and we define their oriented area $\bar{\omega} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ as

$$\bar{\omega}(\mu, \eta) = \mu_p \eta_q - \mu_q \eta_p,$$

whose magnitude is equal to the area of the parallelogram determined by these vectors. The 'oriented' part of $\bar{\omega}$ refers to the fact that the function is skew, in the sense that

$$\bar{\omega}(\mu, \eta) = -\bar{\omega}(\eta, \mu).$$

It might be expected that for $N > 1$ we will move to more abstract volumes of parallelepipeds, however our immediate purpose is served by considering instead a sum of all the oriented areas arising from projecting the coordinates of μ and η on to the (q, p) coordinate planes. We define this as the *symplectic two-form* $\omega : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$, given as

$$\omega(\mu, \eta) := \sum_{i=1}^{N} \bar{\omega}(\mu_{[i]}, \eta_{[i]}), \tag{2.14}$$

where $\mu_{[i]} = (\mu_{q,(i)}, \mu_{p,(i)})^T \in \mathbb{R}^2$ and similarly denoted $\eta_{[i]} = (\eta_{q,(i)}, \eta_{p,(i)})^T \in \mathbb{R}^2$. The description of this function as symplectic comes from the Greek adjective meaning a braiding together, first used mathematically by Weyl [126] to describe the symplectic area as a bundle (or complex) of lines. We can write (2.14) compactly in matrix form as

$$\omega(\mu, \eta) = \mu^T J^{-1} \eta, \tag{2.15}$$

where J is the structure matrix given in (2.2). The linear mapping $A : \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ is a *symplectic* map if the symplectic two-form is preserved under application, i.e. if

$$\omega(A\mu, A\eta) = \omega(\mu, \eta), \qquad \forall \mu, \eta \in \mathbb{R}^{2N}.$$

From (2.15) we can equivalently write this condition as

$$A^T J^{-1} A = J^{-1}, \tag{2.16}$$

which implies $|\det(A)| = 1$.

In the case of nonlinear differentiable mappings $\psi : U \to \mathbb{R}^{2N}$, where $U \subset \mathbb{R}^{2N}$ is

open, we will approximate ψ through linearization. Defining its Jacobian $\psi'(q,p) \in \mathbb{R}^{2N\times 2N}$, we call the map symplectic if

$$\psi'(q,p)^T J^{-1} \psi'(q,p) = J^{-1}, \qquad \forall q,p \in U. \tag{2.17}$$

Notably, the composition of two symplectic maps ψ and ϕ is itself a symplectic map, as

$$\begin{aligned}(\psi\circ\phi)'^T J^{-1}(\psi\circ\phi)' &= \psi'^T\phi'^T J^{-1}\phi'\psi',\\ &= \psi'^T J^{-1}\psi',\\ &= J^{-1}.\end{aligned} \tag{2.18}$$

This yields a natural geometric interpretation of symplecticness, namely that symplectic maps conserve area and orientation. To demonstrate this, consider a bounded, two-dimensional surface $S \subset \mathbb{R}^{2N}$, with some continuously differentiable reparameterization $\theta : K \to S$ where $K \subset \mathbb{R}^2$, such that $S = \theta(K)$. Consider a point $P = (u,v) \in K$ and construct a rectangle $R \subset K$ with a corner set at P and with side lengths Δu and Δv. The image of R under the reparameterization θ will be a two-dimensional subregion of S, potentially now with warped or curved boundaries. The co-ordinates of the corners of the image $\theta(R)$ can be estimated through linearization, where for example

$$\theta(u+\Delta u, v) \approx \theta(u,v) + \frac{\partial\theta}{\partial u}(u,v)\Delta u.$$

Shrinking Δu and Δv to infinitesimal size, we can express the entirety of the planar region K as exactly the union (or tessellation) of all such rectangles R. The image of each rectangle corresponds to a region on S, hence we can find the total oriented area of S, defined as $\alpha(S)$, by summing the oriented areas of each rectangle's image. In this infinitesimal limit the sum becomes an integral, and hence

$$\alpha(S) = \int_K \bar{\omega}\left(\frac{\partial\theta}{\partial u}(u,v)\mathrm{d}u, \frac{\partial\theta}{\partial v}(u,v)\mathrm{d}v\right) = \int_K \bar{\omega}\left(\frac{\partial\theta}{\partial u}(u,v), \frac{\partial\theta}{\partial v}(u,v)\right)\mathrm{d}u\,\mathrm{d}v. \tag{2.19}$$

Given a symplectic map $\psi : U \to \mathbb{R}^{2N}$, we know by definition that the symplectic two-form will be left invariant. In light of (2.19), we can calculate the oriented area of $\psi(S)$, as

$$\begin{aligned}\alpha(\psi(S)) = \alpha(\psi(\theta(K))) &= \int\int_K \bar{\omega}\left(\frac{\partial(\psi\circ\theta)}{\partial u}(u,v), \frac{\partial(\psi\circ\theta)}{\partial v}(u,v)\right)\mathrm{d}u\mathrm{d}v,\\ &= \int\int_K \bar{\omega}\left(\psi'(\theta(u,v))\frac{\partial\theta}{\partial u}(u,v), \psi'(\theta(u,v))\frac{\partial\theta}{\partial v}(u,v)\right)\mathrm{d}u\mathrm{d}v,\\ &= \int\int_K \bar{\omega}\left(\frac{\partial\theta}{\partial u}(u,v), \frac{\partial\theta}{\partial v}(u,v)\right)\mathrm{d}u\mathrm{d}v,\\ &= \alpha(S).\end{aligned}$$

Symplectic maps therefore preserve the symplectic two-form ω of a region of phase

space under transformation. This is an important result in dynamical systems, as the flow of Hamiltonian systems is symplectic.

Lemma 2.20 (Poincaré Theorem)**.** *The flow map $\Phi_{H,t}$ of a twice continuously differentiable Hamiltonian system* (2.3)*, is symplectic for all $t \in \mathbb{R}$ and all differentiable Hamiltonians $H(q,p)$.*

Proof. From the definition of symplecticness (2.17), we need to show that

$$f_t^T J^{-1} f_t = J^{-1}, \qquad \text{where } f_t = \frac{\partial}{\partial z}\Phi_{H,t}(z).$$

From the definition of the flow map we have $\Phi_{H,0}(z) = z$, and hence $f_0 = I_{2N}$ satisfies the condition. It is sufficient therefore to show

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t) = 0, \qquad \text{where} \quad g(t) := f_t^T J^{-1} f_t.$$

By the chain rule and using the variational equation (2.13), we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} g(t) &= f_t^T J^{-1} \dot{f}_t + \dot{f}(t)^T J^{-1} f_t, \\
&= f_t^T J^{-1} \left[J H''(z(t))\, f_t \right] + \left[f_t^T H''(z(t))^T J^T \right] J^{-1} f_t, \\
&= f_t^T H''(z(t))\, f_t - f_t^T H''(z(t))^T f_t = 0,
\end{aligned}$$

as required. □

The effect of Lemma 2.20 is readily apparent if we take $N = 1$, and consider repeatedly propagating a small region S of initial conditions in the phase space $\mathbb{R}^2$, with respect to some Hamiltonian $H(q,p)$. The symplecticness of the flow map implies the images $\Phi_{H,t}^k(S)$ (where $k \in \mathbb{N}$) will all have equal area. As an example, consider the evolution of the system with respect to the Hamiltonian $H(q,p) = p^2/2 + \phi_{\mathrm{dw}}(q)$, where $(q,p) \in \mathbb{R}^2$ and we define the double well potential as

$$U(q) = \phi_{\mathrm{dw}}(q) := (q^2 - 1)^2.$$

This potential is characterized by disjoint solution curves orbiting equilibria $(q_\pm^*, p^*) = (\pm 1, 0)$, unless initial conditions are sufficiently energetic. In Figure 2.2, we plot the propagation of three disconnected regions forward in time, with each region coloured by their energy level. The regions have an arrow-shaped hole in the centre that deforms according to the level set $H(q_0,p_0) = H(q,p)$. In spite of deformations, the three regions have commensurate areas under the flow for all time.

Though we have defined symplecticness as a conservation of the image's symplectic two-form, for $N > 1$ it can also be shown that volumes of the phase space are also conserved as a result of symplecticness (and hence under Hamiltonian flow maps). An alternate way to demonstrate this property follows from Liouville's theorem, which

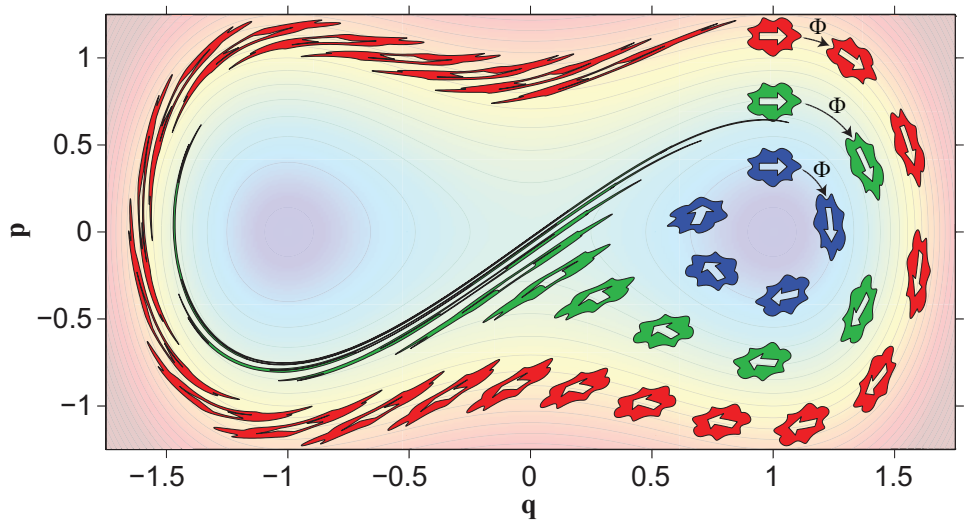


Figure 2.2: Area preservation under Hamiltonian flow maps due to symplecticness.

implies conservation of phase space volumes if the vector field in (2.7) in divergence-free. It is easily verified that $\nabla \cdot f(z) = 0$ for Hamiltonian systems (2.1).

Symplecticness implies the existence of other integral invariants in systems for $N > 1$, and we therefore consider it a stronger property that simply being volume conserving.

2.1.3 Geometric integration

In an effort to emulate the geometric properties of the Hamiltonian flow map we shall endeavour to find a numerical approximation that has an associated symplectic flow map, in order to obtain solution trajectories that are geometrically similar to the exact result. The resulting schemes will not necessarily preserve invariants (such as the Hamiltonian) to higher order than other schemes considered, but we will show that the methods preserve invariants 'close' to those in the original system. This is of central importance in this thesis, as when we move to schemes for stochastic dynamics in Chapter 4 we will use the same machinery to find 'close' invariant properties of the discretization as a perturbation in powers of the step size.

Additionally, we introduce the strategy we shall employ to create our integrators in the remaining chapters: splitting methods. The advantages of such methods for deterministic systems will be presented in this section.

Consider a Hamiltonian $H(q, p)$, and suppose we can additively decompose it into two pieces, such that

$$H(q, p) = H_1(q, p) + H_2(q, p).$$

We can think of each piece as a Hamiltonian in its own right, with dynamics such as

$$\frac{\mathrm{d}}{\mathrm{d}t} z = J \nabla H_1(z), \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t} z = J \nabla H_2(z).$$

If we are able to solve each independent dynamics exactly, then we can compute the flow map for each corresponding piece of the Hamiltonian. These flow maps will themselves be symplectic, by Lemma 2.20. As we have shown in (2.18) that the composition of symplectic maps is itself a symplectic map, composing the flow maps for each piece will give us a symplectic map. Note that these flow maps do not commute in general, so

$$\Phi_{H_1,t} \circ \Phi_{H_2,t} \neq \Phi_{H_2,t} \circ \Phi_{H_1,t} \neq \Phi_{H_1+H_2,t}.$$

Of course, it is not at all obvious that the composition will be consistent; how do we know that the composition of the flow maps approximates the full dynamics? Taking the Taylor series of the composition at an arbitrary point,

$$\begin{aligned}
\Phi_{H_1,\delta t} \circ \Phi_{H_2,\delta t}(z_0) &= \Phi_{H_2,\delta t}(z_0) + \delta t J \nabla H_1(\Phi_{H_2,\delta t}(z_0)) + \mathcal{O}(\delta t^2), \\
&= z_0 + \delta t J \nabla H_2(z_0) + \delta t J \nabla H_1(z_0 + \delta t J \nabla H_2(z_0)) + \mathcal{O}(\delta t^2), \\
&= z_0 + \delta t J \nabla H(z_0) + \mathcal{O}(\delta t^2), \\
&= \Phi_{H,\delta t}(z_0) + \mathcal{O}(\delta t^2),
\end{aligned}$$

and hence we see it is consistent, with local error of at least order 2. Given previous results, we would expect that this method gives us a scheme with global order 1, but now crucially we have a scheme that is symplectic.

Though it seems like we have doubled our workload in computing two flow maps rather than one, they key point here is that we are free to choose our splitting of the Hamiltonian. If we can express the Hamiltonian as a sum of pieces, each one corresponding to an integrable set of ODEs, then we can decompose the problem into a multitude of simpler subproblems.

In molecular dynamics, the Hamiltonians we shall consider are of the form

$$H(q,p) = p^T M^{-1} p + U(q).$$

This gives a natural splitting strategy, defining the pieces as

$$H_A(p) := p^T M^{-1} p, \qquad H_B(q) := U(q), \qquad H(q,p) = H_A(p) + H_B(q). \tag{2.21}$$

The flow maps for each piece are then simply linear drift terms

$$\Phi_{H_A,t}\left(\begin{bmatrix} q \\ p \end{bmatrix}\right) = \begin{bmatrix} q + tM^{-1}p \\ p \end{bmatrix}, \qquad \Phi_{H_B,t}\left(\begin{bmatrix} q \\ p \end{bmatrix}\right) = \begin{bmatrix} q \\ p - t\nabla U(q) \end{bmatrix}.$$

Of course other splittings can be considered, though these are perhaps the simplest building blocks for which to construct a splitting method. The methods we will consider will result from compositions of these basic flow maps. We will introduce a notation for writing these compositions, based on the notation used in [117].

Notation

We will use a shorthand to define symplectic methods, coding them by the order in which we perform our symplectic updates. A symplectic numerical method using the splitting (2.21) is coded by a string of alternating letters A and B, for example $[\![\mathrm{ABABA}]\!]$. Such methods corresponds to a composition of flow maps, for $a_i, b_i \in \mathbb{R}$ the string

$$[\![\mathrm{A}^{\mathrm{a}_1}\mathrm{B}^{\mathrm{b}_1}\mathrm{A}^{\mathrm{a}_2}\mathrm{B}^{\mathrm{b}_2}\cdots\mathrm{A}^{\mathrm{a}_\mathrm{m}}\mathrm{B}^{\mathrm{b}_\mathrm{m}}]\!]$$

corresponds to the method with flow map

$$\Phi_{H_B,\widehat{b}_m\delta t}\circ\Phi_{H_A,\widehat{a}_m\delta t}\circ\cdots\circ\Phi_{H_B,\widehat{b}_2\delta t}\circ\Phi_{H_A,\widehat{a}_2\delta t}\circ\Phi_{H_B,\widehat{b}_1\delta t}\circ\Phi_{H_A,\widehat{a}_1\delta t}.$$

Note that the step sizes used for each step are normalized to ensure the consistency of the overall method, with

$$\widehat{a}_i := \frac{a_i}{a_1+a_2+\ldots+a_m} \quad \text{and} \quad \widehat{b}_i := \frac{b_i}{b_1+b_2+\ldots+b_m}.$$

If a particular a_i or b_i are 0, then we omit the letter from the string, while if a_i or b_i are 1 then we omit the superscript (but keep the letter).

As a simple example, let us consider the method coded as $[\![\mathrm{AB}]\!]$. Composing the flow maps, we write the overall flow map of the numerical method as

$$\widehat{\Phi}_{[\![\mathrm{AB}]\!],\delta t} := \Phi_{H_B,\delta t}\circ\Phi_{H_A,\delta t},$$

where the subscript $[\![\mathrm{AB}]\!]$ now denotes the sequential composition of the splitting pieces. The update scheme reads as

$$\begin{aligned} q_{n+1} &= q_n + \delta t M^{-1} p_n, \\ p_{n+1} &= p_n - \delta t \nabla U(q_{n+1}), \end{aligned} \tag{sE-AB}$$

which is the symplectic Euler method. We will refer to this method as symplectic Euler-AB (or sE-AB) as there is the similar method $[\![\mathrm{BA}]\!]$ resulting from the reverse ordering of the pieces,

$$\widehat{\Phi}_{[\![\mathrm{BA}]\!],\delta t} := \Phi_{A,\delta t}\circ\Phi_{B,\delta t},$$

where the update scheme instead reads

$$\begin{aligned} p_{n+1} &= p_n - \delta t \nabla U(q_n), \\ q_{n+1} &= q_n + \delta t M^{-1} p_{n+1}. \end{aligned} \tag{sE-BA}$$

These are both explicit first-order symplectic methods, but aside from the geometric properties (i.e. oriented area preservation), it is not obvious why symplectic methods should be superior for numerical integration. Indeed the differences between the Euler methods (eE) and (sE-AB) seem primarily cosmetic at a first glance. We can motivate the use of symplectic methods by returning to the harmonic oscillator, with Hamiltonian

$$H(q,p) = p^2/2 + q^2/2, \tag{2.22}$$

where $q, p \in \mathbb{R}$. We have already seen that for a function $I(z)$ to be a first-integral of the dynamics (2.1) we require

$$\nabla I(z) \cdot J \nabla H(z) = 0,$$

for all z in our space. It is clear that the function

$$\widetilde{H}_{\delta t}(q,p) = \frac{1}{2}p^2 + \frac{1}{2}q^2 + \frac{\delta t}{2}qp \tag{2.23}$$

is not a first integral of our system defined by (2.22) for $\delta t > 0$, however examining the behaviour of this function when propagating using the symplectic scheme (sE-AB), we see

$$\begin{aligned}
\widetilde{H}_{\delta t}(q_{n+1}, p_{n+1}) &= \frac{1}{2}p_{n+1}^2 + \frac{1}{2}q_{n+1}^2 + \frac{\delta t}{2}p_{n+1}q_{n+1}, \\
&= \frac{1}{2}(p_n - \delta t q_{n+1})^2 + \frac{1}{2}q_{n+1}^2 + \frac{\delta t}{2}(p_n - \delta t q_{n+1})q_{n+1}, \\
&= \frac{1}{2}p_n^2 + \frac{1}{2}q_{n+1}^2 - \frac{\delta t}{2}p_n q_{n+1}, \\
&= \frac{1}{2}p_n^2 + \frac{1}{2}(q_n + \delta t p_n)^2 - \frac{\delta t}{2}p_n(q_n + \delta t p_n), \\
&= \frac{1}{2}p_n^2 + \frac{1}{2}q_n^2 + \frac{\delta t}{2}p_n q_n = \widetilde{H}_{\delta t}(q_n, p_n),
\end{aligned}$$

and hence the function is preserved under the method. This suggests that the discretized solutions given by the scheme $[\![\mathrm{AB}]\!]$ lie on level sets of $\widetilde{H}_{\delta t}(q,p)$, which correspond to conic sections whose type depends on the step size parameter δt. For $0 < \delta t < 2$, the discretization gives points on skewed concentric ellipses centred at the origin, whereas for $\delta t > 2$ we have a multitude of hyperbolae. This suggests that choosing a step size $\delta t < 2$ our discretized approximations z_n will lie on closed curves, and hence our computed solution will be drift-free with respect to the surface $\widetilde{H}_{\delta t}(q_{n+1}, p_{n+1}) = \text{const}$, giving a bounded energy error. An example is shown in Figure 2.3 for $\delta t = 0.1$.

One can think of the symplectic Euler methods as providing an exact flow map for a nearby Hamiltonian to our target system (where 'nearby' assumes of course that the step size δt is sufficiently small). The method solves exactly the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}z(t) = J \nabla \widetilde{H}_{\delta t}(z),$$

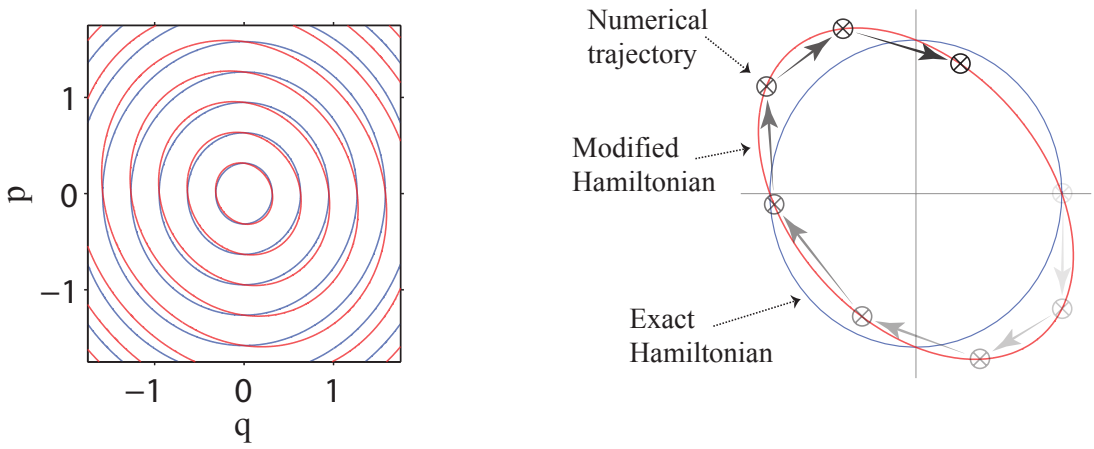


Figure 2.3: Numerical and exact solutions for the harmonic oscillator (2.22) are compared for the symplectic Euler method (sE-AB). The left image plots in phase space the level sets of $\widetilde{H}_{\delta t}$ (red) and H (blue) for $\delta t = 0.1$. The right plot shows how a numerical trajectory is confined to the perturbed (modified) Hamiltonian.

for some fixed step size δt. What is perhaps surprising, is that this is true for all symplectic discretization methods and for Hamiltonians with far more general potential energy functions $U(q)$. For an order s symplectic method, we can think of a perturbed function $\widetilde{H}_{\delta t}(q,p) = H(q,p) + \mathcal{O}(\delta t^s)$ preserved after iteration of the numerical method, or equivalently under applying the method's symplectic map ($\widetilde{H}_{\delta t}$ is sometimes called the scheme's *shadow Hamiltonian*).

To find this shadow Hamiltonian, we could consider perturbing the Hamiltonian H by powers of δt, and solving successively to find functions preserved by the method. A more sophisticated technique is known as backward error analysis. Consider splitting the general vector field in (2.7) into two pieces,

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t) = f_1(z) + f_2(z), \tag{2.24}$$

and define the Lie derivative as the differential operator

$$\mathcal{L}_f := f \cdot \nabla. \tag{2.25}$$

For some differentiable function $\theta(z)$, by taking its Lie derivative we obtain a new function

$$\left(\mathcal{L}_f \theta\right)(z) = f(z) \cdot \nabla\theta(z).$$

Note the linearity of the Lie derivative, such that

$$\mathcal{L}_{f_i + f_j}\theta = (f_i + f_j)\cdot\nabla\theta = f_i\cdot\nabla\theta + f_j\cdot\nabla\theta = \mathcal{L}_{f_i}\theta + \mathcal{L}_{f_j}\theta. \tag{2.26}$$

The Lie derivative is directly related to the evolution of functions along solutions to

the corresponding ODEs. Consider some function $\theta(z)$ of solutions along one piece of the vector field

$$\frac{\mathrm{d}}{\mathrm{d}t}z(t) = f_i(z),$$

then taking the derivative of $\theta(z)$ with respect to time, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta(z(t)) = \nabla\theta(z(t)) \cdot \dot{z}(t) = \left(\mathcal{L}_{f_i}\theta\right)(z),$$

by chain rule. Consequently we can use Taylor's identity and the formula for the exponential to obtain

$$\theta(z(t)) = \sum_{j=0}^{\infty} \frac{t^j}{j!}\frac{\mathrm{d}^j}{\mathrm{d}t^j}\theta(z(t))\bigg|_{z(t)=z_0} = \sum_{j=0}^{\infty} \frac{t^j}{j!}(\mathcal{L}_{f_i}^j\theta)(z_0) = \exp\left(t\mathcal{L}_{f_i}\right)\theta(z_0).$$

Note in particular that for the identity map $\theta(z) = \mathrm{Id}(z) = z$, we find that the exponential of the Lie derivative is related to the flow map itself, namely

$$\Phi_{f_i,t}(z_0) = z(t) = \exp\left(t\mathcal{L}_{f_i}\right)\mathrm{Id}(z_0). \tag{2.27}$$

Returning to our original question, what can we say about the composition of such methods? Writing the flow map for the scheme denoted $[\![\mathrm{f_i f_j}]\!]$ as $\widehat{\Phi}_{[\![\mathrm{f_i f_j}]\!],t}$, we find

$$\begin{aligned}\widehat{\Phi}_{[\![\mathrm{f_i f_j}]\!],t}(z_0) &= \left(\Phi_{f_j,t} \circ \Phi_{f_i,t}\right)(z_0),\\ &= \exp\left(t\mathcal{L}_{f_i}\right)\Phi_{f_j,t}(z_0),\\ &= \exp\left(t\mathcal{L}_{f_i}\right)\exp\left(t\mathcal{L}_{f_j}\right)\mathrm{Id}(z_0).\end{aligned} \tag{2.28}$$

More generally, splitting the ODE into m pieces such that

$$\frac{\mathrm{d}}{\mathrm{d}t}z(t) = \sum_{i=1}^{m} f_i(z) = f(z),$$

we can apply this reasoning inductively to find that, for the method coded $[\![\mathrm{f_1 f_2 \ldots f_m}]\!]$, we have

$$\begin{aligned}\widehat{\Phi}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],t}(z_0) &= \left(\Phi_{f_m,t} \circ \ldots \circ \Phi_{f_2,t} \circ \Phi_{f_1,t}\right)(z_0),\\ &= \exp\left(t\mathcal{L}_{f_1}\right)\exp\left(t\mathcal{L}_{f_2}\right)\ldots\exp\left(t\mathcal{L}_{f_m}\right)\mathrm{Id}(z_0).\end{aligned}$$

The apparent reversal of the order of the semigroups, compared to the order of the flow maps, is known as the *Vertauschungssatz* (see the discussion in [48, Section III.5.1]). This is the reason behind our notation for a method reading left to right, as our interest will primarily be in the sequence of the exponential terms, rather than the flow maps.

Given a composition method written $[\![\mathrm{f_1 f_2 \ldots f_m}]\!]$, fixing the time interval to be the step size $\mathrm{d}t$, we define a differential operator $\widehat{\mathcal{L}}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],\mathrm{d}t}$ as

$$\exp\left(\delta t\,\widehat{\mathcal{L}}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],\mathrm{d}t}\right) := \exp\left(\delta t\,\mathcal{L}_{f_1}\right)\exp\left(\delta t\,\mathcal{L}_{f_2}\right)\ldots\exp\left(\delta t\,\mathcal{L}_{f_m}\right), \tag{2.29}$$

and hence for any smooth differentiable function $\theta(z)$, its evolution under the method is given by

$$\theta(\hat{z}(\delta t)) = \exp\left(\delta t\, \widehat{\mathcal{L}}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],\delta t}\right) \theta(z(0)).$$

From the Taylor series of $\exp\left(\delta t\, \widehat{\mathcal{L}}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],\delta t}\right)$, we can see that for a function $\theta(z)$ to be invariant under iteration (for any fixed δt), we require $\widehat{\mathcal{L}}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],\delta t}\theta(z) \equiv 0$. In order to compute the operator, we will make use of the Baker-Campbell-Hausdorff (BCH) formula to compute the product of exponentials in (2.29) [48], as for general linear operators X and Y, we have

$$\exp(\delta t\, X)\exp(\delta t\, Y) \neq \exp(\delta t\, (X+Y)),$$

unless X and Y commute. The BCH formula gives that

$$\exp(\delta t\, X)\exp(\delta t\, Y) = \exp(\delta t\, Z_{\delta t}) = \exp(\delta t\, Z_{[1]} + \delta t^2 Z_{[2]} + \delta t^3 Z_{[3]} + \mathcal{O}(\delta t^4)), \quad (2.30)$$

where $Z_{[i]}$ are linear operators given by combinations of X and Y;

$$\begin{aligned} Z_{[1]} &= X + Y, \\ Z_{[2]} &= \frac{1}{2}[X, Y], \\ Z_{[3]} &= \frac{1}{12}\left([X,[X,Y]] + [Y,[Y,X]]\right), \end{aligned}$$

with the commutator of X and Y defined as $[X,Y] := XY - YX$. By repeatedly applying the BCH formula we can compute $\widehat{\mathcal{L}}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],\delta t}$ as required, however it immediately becomes apparent that we can expect the terms at higher orders of δt to become extremely complicated linear combinations of $\mathcal{L}_{f_i}$, particularly as a method becomes more complex (i.e. as m increases). An additional worry comes from the formal nature of this approach, given the unbounded nature of each Lie derivative we cannot expect the BCH series to converge in the general case.

Rather than thinking of the characteristic operator $\widehat{\mathcal{L}}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],\delta t}$ as a perturbation of the exact operator $\mathcal{L}_{J\nabla H}$, we can instead consider it as the exact operator of a modified Hamiltonian $\widetilde{H}_{\delta t}(q,p)$, such that

$$\exp\left(\delta t\, \widehat{\mathcal{L}}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!],\delta t}\right) = \exp\left(\delta t\, \mathcal{L}_{J\nabla H} + \delta t^2 \widehat{\mathcal{L}}^{[1]}_{[\![\mathrm{f_1 f_2 \ldots f_m}]\!]} + \mathcal{O}(\delta t^3)\right) = \exp\left(\delta t\, \mathcal{L}_{\widetilde{H}_{\delta t}}\right).$$

The function $\widetilde{H}_{\delta t}(z)$ will be invariant under the numerical method, as we think of the numerical method solving exactly the Hamiltonian system

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t) = J\nabla \widetilde{H}_{\delta t}(z).$$

Defining the Poisson bracket of two functions $F(z)$ and $G(z)$ as

$$\{F(z), G(z)\} := \nabla F(z) \cdot J\nabla G(z),$$

the Lie derivative can be reformulated elegantly as

$$\mathcal{L}_{J\nabla H}\phi = \nabla\phi \cdot J\nabla H = \{\phi, H\},$$

with any first-integral $I(z)$ of the Hamiltonian system (2.1) satisfying $\{I, H\} = 0$.

Lemma 2.31. *Consider two smooth Hamiltonians $G_1(z)$ and $G_2(z)$, with corresponding vector fields*

$$g_1(z) = J\nabla G_1(z), \quad g_2(z) = J\nabla G_2(z).$$

Given their associated Lie derivatives $\mathcal{L}_{g_1}$ and $\mathcal{L}_{g_2}$, we can write their commutator as

$$[\mathcal{L}_{g_1}, \mathcal{L}_{g_2}] = \mathcal{L}_{J\nabla\{G_2,G_1\}}.$$

Proof. We make use of the Jacobi identity

$$\{\{A, B\}, C\} + \{\{B, C\}, A\} + \{\{C, A\}, B\} = 0,$$

where here A, B and C are arbitrary functions. Replacing these with G_1, G_2 and ϕ, we can take advantage of the skew-symmetry of the Poisson bracket to write

$$\{\{\phi, G_2\}, G_1\} - \{\phi, \{G_2, G_1\}\} - \{\{\phi, G_1\}, G_2\} = 0.$$

Using the definition of the Lie derivative, this becomes

$$\{\mathcal{L}_{g_2}\phi, G_1\} - \mathcal{L}_{J\nabla\{G_2,G_1\}}\phi - \{\mathcal{L}_{g_1}\phi, G_2\} = 0,$$

and hence

$$\mathcal{L}_{g_1}\mathcal{L}_{g_2}\phi - \mathcal{L}_{g_2}\mathcal{L}_{g_1}\phi = \mathcal{L}_{J\nabla\{G_2,G_1\}}\phi,$$

as required. □

Using the correspondence between Poisson brackets and Lie derivatives, we can write down the shadow Hamiltonian $\widetilde{H}_{\delta t}$ for a method. As an example, we return to the symplectic Euler method $[\![\mathrm{AB}]\!]$. Using the BCH formula we can compute

$$\widehat{\mathcal{L}}_{[\![\mathrm{AB}]\!],\delta t} = \mathcal{L}_A + \mathcal{L}_B + \frac{\delta t}{2}[\mathcal{L}_A, \mathcal{L}_B] + \frac{\delta t^2}{12}([\mathcal{L}_A, [\mathcal{L}_A, \mathcal{L}_B]] + [\mathcal{L}_B, [\mathcal{L}_B, \mathcal{L}_A]]) + \mathcal{O}(\delta t^3),$$

where we have simplified our notation from $\mathcal{L}_{J\nabla H_A}$ to $\mathcal{L}_A$. Using the skew properties of the Poisson bracket, as well as Lemma 2.31 and (2.26), we can write this as $\mathcal{L}_{J\widetilde{H}_{\delta t}}$

where

$$\widetilde{H}_{\delta t} = H - \frac{\delta t}{2}\{H_A, H_B\} + \frac{\delta t^2}{12}\left(\{H_A, \{H_A, H_B\}\} + \{H_B, \{H_B, H_A\}\}\right) + \mathcal{O}(\delta t^3).$$

For the Hamiltonian splitting $H(q,p) = H_A(p) + H_B(q)$, where

$$H_A(q,p) = \frac{1}{2}p^T M^{-1} p, \qquad H_B(q,p) = U(q),$$

we have

$$\begin{aligned}\{H_A, H_B\} &= \nabla H_A \cdot J \nabla H_B, \\ &= \begin{bmatrix} 0 \\ M^{-1}p \end{bmatrix} \cdot \begin{bmatrix} 0 & I_N \\ -I_N & 0 \end{bmatrix} \begin{bmatrix} \nabla U(q) \\ 0 \end{bmatrix}, \\ &= -p^T M^{-1} \nabla U(q),\end{aligned}$$

as the leading term in the expansion for $[\![\mathrm{AB}]\!]$. Similarly we can calculate

$$\{H_A, \{H_A, H_B\}\} = p^T M^{-1} \nabla^2 U(q) M^{-1} p, \quad \{H_B, \{H_B, H_A\}\} = \nabla U(q)^T M^{-1} \nabla U(q), \tag{2.32}$$

where we denote the Hessian matrix $\nabla^2 U(q) := \nabla \nabla^T U(q)$.

Returning to our motivating example, we can see that in a planar phase space for $U(q) = q^2$ and $M = 1$, the perturbation given in (2.23) reduces to exactly $-\{H_A, H_B\}$. It is not at all obvious where the higher order δt terms in the expansion have disappeared to, but it can be shown that the expansion converges (owing to the simplicity of H_A and H_B) to

$$\widetilde{H}_{\delta t} = \lambda_{\delta t}\left(H + \frac{\delta t}{2} p^T M^{-1} q\right),$$

for some finite nonzero value $\lambda_{\delta t}$ that depends only on δt. Fixing the step size, we can simply divide through by $\lambda_{\delta t}$ to recover (2.23) as a first integral of the perturbed dynamics (as all functions of first-integrals are themselves first integrals). This is a special case however, for simple (in this case, quadratic) Hamiltonians, many of the Poisson brackets are zero due to high derivatives, whereas for more general systems we cannot expect the formal series for $\widetilde{H}_{\delta t}$ to converge.

For a symplectic order p numerical method, its flow map $\widehat{\Phi}_{\delta t}$ is equivalent to the exact flow map for a perturbed Hamiltonian $\widetilde{H}_{\delta t}$, such that

$$\widehat{\Phi}_{\delta t} = \Phi_{J\nabla \widetilde{H}_{\delta t}, \delta t}.$$

For general systems, we can find the perturbed Hamiltonian using the BCH formula, though it is purely formal approach, making any information we gain suspicious at best or misleading at worst. We can imagine truncating to order δt^i (for $i > s$) writing

$$\widetilde{H}_{\delta t,[i]} = H + \delta t^s \widetilde{H}^{[s]} + \delta t^{s+1} \widetilde{H}^{[s+1]} + \delta t^{s+2} \widetilde{H}^{[s+2]} + \ldots \delta t^i \widetilde{H}^{[i]}.$$

We then hope to find an estimate for the difference between the numerical method and the exact flow of this truncated system. Estimates can be given [48, 72] such that

$$\left|\Phi_{J\nabla\widetilde{H}_{\delta t,[i]},\delta t} - \widehat{\Phi}_{\delta t}\right| \leq c_1\delta t\left(c_2\delta t\left(i+1\right)\right)^{i+1},$$

for positive constants c_1 and c_2 independent of the step size δt and the truncation index i. The behaviour of the right hand side suggests that, for fixed small δt, if we take more terms in the perturbation by increasing i then we will initially converge towards the flow map of the numerical method. However, we find a strong divergence effect by increasing i too far, with this divergence tipping point occurring for larger i as $\delta t \to 0$. If we choose the truncation index $i = i^*$, where

$$i^* = \left\lfloor \frac{e^{-1}}{c_2\delta t} - 1 \right\rfloor \leq \frac{e^{-1}}{c_2\delta t} - 1,$$

then this gives

$$\begin{aligned}\left|\Phi_{J\nabla\widetilde{H}_{\delta t,[i^*]},\delta t} - \widehat{\Phi}_{\delta t}\right| &\leq c_1\delta t\left(c_2\delta t\left(i^*+1\right)\right)^{i^*+1},\\ &\leq c_1\delta t e^{-(i^*+1)},\\ &\leq \kappa_1\delta t e^{-\kappa_2/\delta t},\end{aligned}$$

for positive constants $\kappa_1 = ec_1$ and $\kappa_2 = e^{-1}/c_2$. The drift in the truncated Hamiltonian over a discretized trajectory of n steps is then simply

$$\begin{aligned}\left|\widetilde{H}_{\delta t,[i^*]}(z_n) - \widetilde{H}_{\delta t,[i^*]}(z_0)\right| &= \left|\sum_{i=1}^{n}\widetilde{H}_{\delta t,[i^*]}(z_i) - \widetilde{H}_{\delta t,[i^*]}(z_{i-1})\right|,\\ &\leq \sum_{i=1}^{n}\left|\widetilde{H}_{\delta t,[i^*]}(z_i) - \widetilde{H}_{\delta t,[i^*]}(z_{i-1})\right|,\\ &\leq \sum_{i=1}^{n}\left|\widetilde{H}_{\delta t,[i^*]}\left(\widehat{\Phi}_{\delta t}(z_{i-1})\right) - \widetilde{H}_{\delta t,[i^*]}\left(\Phi_{J\nabla\widetilde{H}_{\delta t,[i^*]},\delta t}(z_{i-1})\right)\right|,\\ &\leq \sum_{i=1}^{n}K\left|\widehat{\Phi}_{\delta t}(z_{i-1}) - \Phi_{J\nabla\widetilde{H}_{\delta t,[i^*]},\delta t}(z_{i-1})\right|,\\ &\leq n\delta t K\kappa_1 e^{-\kappa_2/\delta t},\end{aligned}$$

for $K > 0$ the Lipschitz constant of $\widetilde{H}_{\delta t,[i^*]}$, and making the use of the face that $\widetilde{H}_{\delta t,[i^*]}$ is preserved under the flow of the exact flow map $\Phi_{J\nabla\widetilde{H}_{\delta t,[i^*]},\delta t}$. Thus the numerically computed trajectory will remain exponentially close to the perturbed dynamics given by the truncated shadow Hamiltonian, over an exponentially long time interval

$$T < T^* = n^*\delta t \leq e^{\kappa_2/\delta t}.$$

The expected drift in the Hamiltonian over this timescale is then

$$|H(z_n) - H(z_0)| = \left|\widetilde{H}_{\delta t,[i^*]}(z_n) - \widetilde{H}_{\delta t,[i^*]}(z_0)\right| + \mathcal{O}(\delta t^p) = \mathcal{O}(\delta t^p), \tag{2.33}$$

where here $n < n^*$.

We may hope that we can construct a symplectic method such that the Hamiltonian itself is preserved under iteration, so

$$\{\widetilde{H}_{\delta t}, H\} = 0.$$

If we assume that the only first integrals of the dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t) = J\nabla H(z),$$

are of the form $F(H(q,p))$, for arbitrary F, then this implies (fixing step size) that the shadow Hamiltonian itself must be a function of the Hamiltonian,

$$\widetilde{H}_{\delta t}(q,p) = F_{\delta t}(H(q,p)),$$

for some function $F_{\delta t}$. Hence on level sets $H(q,p) = E \neq 0$, we have

$$\widetilde{H}_{\delta t}(q,p) = F_{\delta t}(E) = \mu H(q,p),$$

for some fixed value μ, depending on the energy level E and step size δt. Hence the numerically computed solutions $\hat{z}(t)$ are such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{z}(t) = J\nabla\widetilde{H}_{\delta t}(z) = \mu J\nabla H(z) = \frac{\mathrm{d}}{\mathrm{d}t} z(\mu\, t),$$

suggesting that our perturbed dynamics solves the problem exactly after applying a time-transformation. Finding a symplectic integrator with this property for a general Hamiltonian will not be possible, and hence conservation of $H(q,p)$ and symplecticness are mutually exclusive properties for numerical methods. We elucidate some strategies for more accurate approximation of $H(q,p)$ using symplectic methods.

2.1.4 Higher order symplectic methods

We can exploit certain features of the BCH series to obtain higher order symplectic methods. For linear operators X and Y, consider symmetrizing a composition of exponentials, such that

$$\exp\left(\frac{t}{2}X\right)\exp\left(tY\right)\exp\left(\frac{t}{2}X\right) = \exp(t\widetilde{Z}_t) = \exp(t\widetilde{Z}_{[1]} + t^2\widetilde{Z}_{[2]} + t^3\widetilde{Z}_{[3]} + t^4\widetilde{Z}_{[4]} + \ldots).$$

Multiplying by the same expansion using a different time step s, we have

$$\exp\left(\frac{s}{2}X\right)\exp\left(sY\right)\exp\left(\frac{s}{2}X\right)\exp\left(\frac{t}{2}X\right)\exp\left(tY\right)\exp\left(\frac{t}{2}X\right) = \exp(s\widetilde{Z}_s)\exp(t\widetilde{Z}_t), \tag{2.34}$$

where $\widetilde{Z}_t$ commutes with $\widetilde{Z}_s$, giving

$$\begin{aligned}\exp\left(s\widetilde{Z}_s\right)&\exp\left(t\widetilde{Z}_t\right)\\ &= \exp\left(s\widetilde{Z}_s + t\widetilde{Z}_t\right),\\ &= \exp\left((s+t)\widetilde{Z}_{[1]} + (s^2+t^2)\widetilde{Z}_{[2]} + (s^3+t^3)\widetilde{Z}_{[3]} + (s^4+t^4)\widetilde{Z}_{[4]} + \ldots\right).\end{aligned}$$

Setting $s = -t$, the left hand side of (2.34) collapses to the identity. Hence the surviving even-powered terms in the expansion on the right hand side must be zero. Evidently, composing linear operators symmetrically leave only the odd order terms in the exponent, giving the symmetric BCH formula

$$\exp\left(\frac{t}{2}X\right)\exp\left(tY\right)\exp\left(\frac{t}{2}X\right) = \exp(t\left(X+Y\right) + t^3\widetilde{Z}_{[3]} + \ldots), \tag{2.35}$$

where

$$\widetilde{Z}_{[3]} = \frac{1}{12}[Y,[Y,X]] - \frac{1}{24}[X,[X,Y]]. \tag{2.36}$$

Hence we would expect that we can easily find a second-order method using a symmetric (or a Strang) splitting scheme. Returning to Hamiltonian dynamics, recall splitting the vector field into pieces defined in (2.21), denoted A and B. In light of (2.35), the simplest splitting schemes giving the desired symmetric structure are denoted simply $[\![\text{ABA}]\!]$ and $[\![\text{BAB}]\!]$, which correspond to the position and velocity Verlet methods respectively:

Position Verlet: $[\![\text{ABA}]\!]$	Velocity Verlet: $[\![\text{BAB}]\!]$	
$q_{n+1/2} = q_n + (\delta t/2)\ M^{-1}p_n,$	$p_{n+1/2} = p_n - (\delta t/2)\ \nabla U(q_n),$	(Verlet)
$p_{n+1} = p_n - \delta t\,\nabla U(q_{n+1/2}),$	$q_{n+1} = q_n + \delta t\,M^{-1}p_{n+1/2},$	
$q_{n+1} = q_{n+1/2} + (\delta t/2)\ M^{-1}p_{n+1}.$	$p_{n+1} = p_{n+1/2} - (\delta t/2)\ \nabla U(q_{n+1}).$	

More complicated symmetric second-order schemes can be devised by proceeding almost arbitrarily and maintaining a symmetric composition, but it is not immediately clear how these would improve on the two Verlet schemes. The Verlet methods are seen as the gold-standard for molecular dynamics computations - both require only one evaluation of $\nabla U(q)$ per iteration (where the velocity Verlet scheme can reuse $\nabla U(q_{n+1})$ for the next iteration), and offer a second-order symplectic evolution. We can use the Poisson brackets computed in (2.32) to find the perturbation to the Hamiltonian by substituting them into (2.36).

For example, for the velocity Verlet scheme $[\![\mathrm{BAB}]\!]$, defining its shadow Hamiltonian as $\widetilde{H}_{\delta t} = H + \delta t^2 \widetilde{H}^{[2]} + \mathcal{O}(\delta t^4)$, we can use the BCH formula to find the leading perturbation as

$$\begin{aligned}\widetilde{H}^{[2]} &= \frac{1}{12}\left\{H_A, \{H_A, H_B\}\right\} - \frac{1}{24}\left\{H_B, \{H_B, H_A\}\right\}, \\ &= \frac{1}{12} p^T M^{-1} \nabla^2 U(q) M^{-1} p - \frac{1}{24} \nabla U(q)^T M^{-1} \nabla U(q).\end{aligned}$$

Yoshida [129] gives an elegant method for creating a scheme of arbitrarily high even order. If we have some scheme with order $2s$ and perturbation terms of even order (such as arises from a symmetric composition), then the evolution of the system under this method is given by $\exp(\delta t\, \widehat{\mathcal{L}}_{\delta t})$, where

$$\widehat{\mathcal{L}}_{\delta t} = \mathcal{L}_{J\nabla H} + \delta t^{2s} \widehat{\mathcal{L}}^{[2s]} + \delta t^{2s+2} \widehat{\mathcal{L}}^{[2s+2]} + \delta t^{2s+4} \widehat{\mathcal{L}}^{[2s+4]} + \mathcal{O}(\delta t^{2s+6})$$

is the method's characteristic operator. The linear operators $\widehat{\mathcal{L}}^{[2i]}$ are compositions of Lie derivatives, and can be computed using the symmetric BCH formula.

We now consider the product

$$\exp\left(\tau_0 \delta t\, \widehat{\mathcal{L}}_{\tau_0 \delta t}\right) \exp\left(\tau_1 \delta t\, \widehat{\mathcal{L}}_{\tau_1 \delta t}\right) \exp\left(\tau_0 \delta t\, \widehat{\mathcal{L}}_{\tau_0 \delta t}\right) = \exp\left(\delta t\, \bar{Z}_{\delta t}\right),$$

where, using the fact that $\widehat{\mathcal{L}}_{\delta t}$ commutes with itself, we find

$$\begin{aligned}\bar{Z}_{\delta t} &= 2\tau_0\, \widehat{\mathcal{L}}_{\tau_0 \delta t} + \tau_1\, \widehat{\mathcal{L}}_{\tau_1 \delta t}, \\ &= (2\tau_0 + \tau_1)\, \mathcal{L}_{J\nabla H} + (2\tau_0^{2s+1} + \tau_1^{2s+1}) \delta t^{2s} \widehat{\mathcal{L}}^{[2s]} + \mathcal{O}(\delta t^{2s+2}).\end{aligned}$$

We have free reign over constants τ_0 and τ_1, though for the method to be consistent we require $2\tau_0 + \tau_1 = 1$. Hence we have an opportunity to annihilate the perturbation operator at order δt^{2s} by choosing $2\tau_0^{2s+1} + \tau_1^{2s+1} = 0$ as well. Solving simultaneously, there exists a unique real solution

$$\tau_0 = \frac{1}{2 - \kappa}, \qquad \tau_1 = -\frac{\kappa}{2 - \kappa}, \qquad \kappa^{2s+1} = 2,$$

giving us a scheme of order $2s + 2$. We can then proceed recursively, as this new order $2s + 2$ scheme can be composed similarly to wipe out successive higher order terms.

Though the fourth-order version of the scheme was given first by Forest and Ruth [39] (and discovered independently by Yoshida [129] and Candy and Rozmus [19]), we shall simply refer to these higher-order schemes as Yoshida methods, owing to the elegant derivation of schemes of arbitrary order.

Using $[\![\mathrm{ABA}]\!]$ as a base second-order method ($2s = 2$), a Yoshida fourth-order method is written in our notation as $[\![\mathrm{ABAA}^{-\sqrt[3]{2}}\mathrm{B}^{-\sqrt[3]{2}}\mathrm{A}^{-\sqrt[3]{2}}\mathrm{ABA}]\!]$. This scheme requires three evaluations of the force $\nabla U(q)$ per iteration, making it significantly more expensive than the vanilla second-order Verlet method. Note as well that because

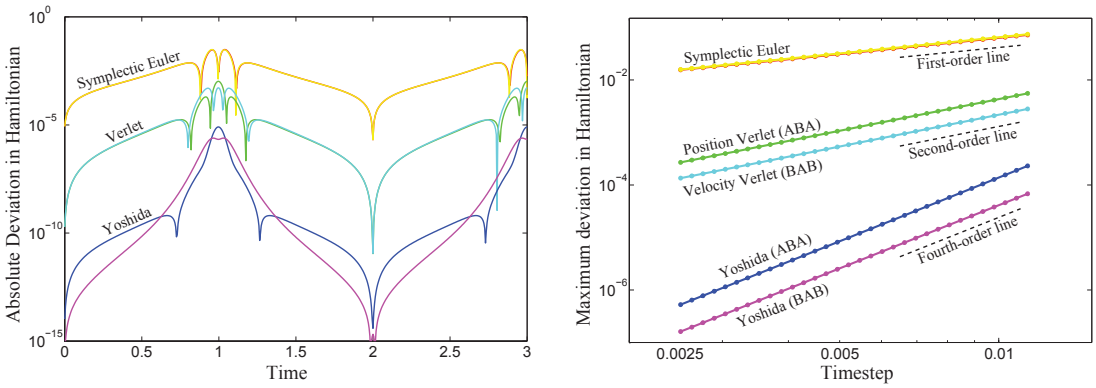


Figure 2.4: We repeat the experiment Figure 2.1 with the symplectic Euler, Verlet and Yoshida schemes. We plot the absolute deviation in the computed Hamiltonian (left) as a function of time, for each scheme at a fixed timestep $\delta t = 0.005$. Additionally we vary the timestep and compute the maximum deviation in the Hamiltonian (right), comparing the results with guide lines of powers of the step size.

$|\tau_i| > 1$, we expect a shrink in the stability region for this method relative to the second-order method it is built upon, as we perform iterations of our 'base' method at timesteps $\tau_i \delta t$.

We compare the symplectic methods developed so far in Figure 2.4, for $t \in [0, 100]$ using planar phase space and the Lennard Jones oscillator (2.11). As expected, there is no visible drift in the Hamiltonian as the simulation time increases, compared to the same problem studied in Figure 2.1. Instead, we see a periodic oscillation in the fluctuation of the total energy. Additionally we compute trajectories propagated at different step sizes, and plot the maximum deviation

$$\max |H(q(t), p(t)) - H(q(0), p(0))|$$

over the entire trajectory. We can see that for an order s method, the maximum deviance in the Hamiltonian is of order s, as is expected from the analysis in (2.33). Yoshida fourth-order schemes are tested using both the position and velocity Verlet schemes as their base second-order methods (the base method is denoted in parenthesis in Figure 2.4), giving similar results.

Of course we need not stop at a fourth-order scheme: we can use the constructed method to target any higher order terms we wish by composing the new scheme in a symmetric form and proceeding identically to the above. For the purposes of molecular sampling however, the computational price-tag of higher order methods outweighs potential numerical gains. After all, the potential energy functions used in the ODEs are not perfect reflections of the laws of physics, and past a certain point using a high order method merely finds a better solution to an approximate model.

For more general applications, such methods are extremely useful. For our purposes

though, far more important than high quantitative fidelity is the preservation of the qualitative nature of solutions to Hamiltonian systems. In that respect, second-order methods provide a 'good enough' option for numerical integration, with symplectic methods the de facto choice for molecular modelling.

2.1.5 Scale separation

Recall that for the harmonic oscillator (2.22), the symplectic Euler method $[\![\mathrm{AB}]\!]$ preserves the quantity

$$\widetilde{H}_{\delta t}(q,p) = \frac{1}{2}p^2 + \frac{1}{2}q^2 + \frac{\delta t}{2}qp, \tag{2.37}$$

for all $\delta t > 0$. As discussed previously, level sets of this function in the plane are conic sections parameterized by δt, whose type changes from a family of ellipses to hyperbolae when $\delta t > 2$. Therefore the qualitative merit of numerical solutions is only valid for this problem in the region $\delta t \in (0, 2)$, introducing the notion of the step size threshold - a maximum value the step size can take before computed solutions unbounded, or the numerical method itself becomes unstable.

The presence of such a step size limit is unsurprising; we cannot expect to be able to push our discretization as far as we would like. The discretized points (z_k) are only an approximation to the true trajectory, and it is perhaps asking too much that the predictions made by the numerical method are valid for all timesteps. It is common for systems to exhibit a separation of time scales, where we have two (or more) coupled process in the overall system that have natural frequencies orders of magnitude apart. Such systems are ubiquitous in nature, and typically the slowest acting system (with the longest relaxation timescale) is of primary interest. However, in simulation the step size threshold is set by the fastest frequencies of the system [4], meaning that one must take a very large number of small timesteps in order to resolve the behaviour of interest. Where the separation of timescales is very large, such simulation is infeasible.

We seek ways to increase the step size threshold in systems where there is a large timescale gap. Let us be given a Hamiltonian with $q, p \in \mathbb{R}^{2N}$, such that

$$H(q,p) = p^T M^{-1} p/2 + U_f(q) + U_s(q), \tag{2.38}$$

where we think of the total potential $U(q) = U_s(q) + U_f(q)$ being the sum of prescribed slow and fast potentials U_s and U_f respectively. We introduce the splitting

$$H_f(q,p) := p^T M^{-1} p/2 + U_f(q), \qquad H_s(q,p) := U_s(q), \tag{2.39}$$

and consider symplectic evolution schemes using these pieces. Just as with the Verlet schemes, we can use a symmetric composition to give a second-order method. One popular method is given by the r-RESPA (reversible-REference System Propagation

Algorithm) scheme [119, 120], written in our notation as ⟦sfs⟧, with dynamics

$$\begin{aligned} p_{n+1/3} &= p_n + (\delta t/2)\nabla U_S(q_n), \\ \left(q_{n+1}, p_{n+2/3}\right) &= \Phi_{J\nabla H_F,\delta t}(q_n, p_{n+1/3}), \\ p_{n+1} &= p_{n+2/3} + (\delta t/2)\nabla U_S(q_{n+1}), \end{aligned} \tag{r-RESPA}$$

where $\Phi_{J\nabla H_F,\delta t}$ is the flow map for the fast dynamics. We are forced to write the flow map into the algorithm as our choice of splitting for H_f in (2.39) means that we cannot solve the system exactly for general U_f. In practice, a high-order symplectic numerical method is often used, with a separate timestep $\delta t_f \ll \delta t$, ensuring that the dominant error in the scheme is a power of the slow (or exterior) timestep δt. This splitting technique is often known as multiple timestepping, due to the presence of these fast and slow timesteps. Additional mid-range timescales can be used to optimize efficiency [105], though we will focus on a binary splitting strategy for simplicity.

This may seem like simply dividing the workload for no real gain, however often the fast dynamics are harmonic potentials (where the flow map is known exactly), or a sum of local potentials only involving a few degrees of freedom. Either way, the computational cost of approximating the central 'fast' flow map is assumed to be far cheaper than the evaluation of the force $\nabla U_s(q)$.

We would expect that the r-RESPA scheme should benefit from an increased step size stability threshold, that should allow us to bridge disparities in timescale. However, in practice we can introduce artificial resonances in the system by choosing the timestep to be an integer multiple of the fast system's period [9, 72, 77].

For instance, consider a particle of unit mass in one-dimension, connected to the origin by two springs, with Hamiltonian as in (2.38) with fast and slow potentials

$$U_s(q) = q^2/2, \qquad U_f(q) = q^2/(2\epsilon^2),$$

where $0 < \epsilon \ll 1$ is a small constant. Using the symplectic Euler method ⟦AB⟧ on this problem with $U(q) = U_s(q) + U_f(q)$, the update scheme becomes

$$\begin{bmatrix} q_{n+1} \\ p_{n+1} \end{bmatrix} = \widehat{\Phi}_{\llbracket \mathrm{AB} \rrbracket,\delta t}\left(\begin{bmatrix} q_n \\ p_n \end{bmatrix}\right) = \Psi_{\llbracket \mathrm{AB} \rrbracket,\delta t}\begin{bmatrix} q_n \\ p_n \end{bmatrix},$$

where the matrix $\Psi_{\llbracket \mathrm{AB} \rrbracket,\delta t}$ is computed from the definition of the method as

$$\begin{aligned} \Psi_{\llbracket \mathrm{AB} \rrbracket,\delta t} &= \Psi_{B,\delta t}\,\Psi_{A,\delta t}, \\ &= \begin{bmatrix} 1 & 0 \\ -\delta t\left(1+\epsilon^{-2}\right) & 1 \end{bmatrix}\begin{bmatrix} 1 & \delta t \\ 0 & 1 \end{bmatrix}, \\ &= \begin{bmatrix} 1 & \delta t \\ -\delta t\left(1+\epsilon^{-2}\right) & 1-\delta t^2\left(1+\epsilon^{-2}\right) \end{bmatrix}. \end{aligned}$$

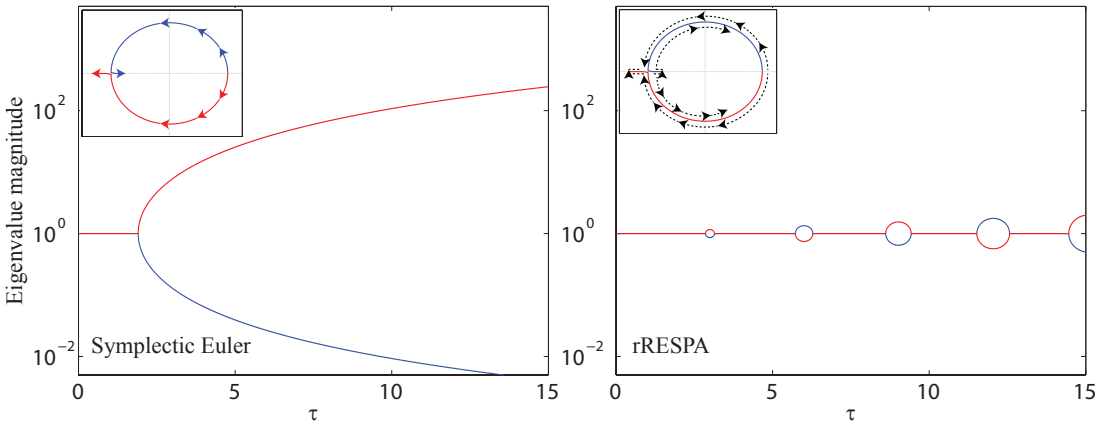


Figure 2.5: The magnitude of eigenvalues is plotted for the symplectic Euler (left) and rRESPA (right) methods, as a function of $\tau = \delta t/\epsilon$, where we fix $\epsilon = 1/3$. The inset for each figure plots the evolution of the eigenvalues in the complex plane, as we increase τ. In the symplectic Euler case we can see divergence along the real line, though the eigenvalues of the rRESPA scheme move in a more complicated way: the eigenvalues make a brief sortie into the real line before returning and continuing to orbit the origin around the unit circle.

The behaviour of the map under repeated iteration is given by the eigenvalues of the matrix. In order to ensure stability, we require its spectral radius (the supremum of the eigenvalues' absolute values) to be no greater than unity. As this is a symplectic map (and hence the determinant is one) mapping to $\mathbb{R}^2$, the eigenvalues are reciprocals of each other, and a necessary condition for the method's stability can be written as

$$\left|\mathrm{Tr}\left(\Psi_{[\![\mathrm{AB}]\!],\delta t}\right)\right| \leq 2,$$

and hence we require

$$\begin{aligned} \left(2 - \delta t^2\left(1+\epsilon^{-2}\right)\right)^2 &\leq 4, \\ \delta t^2\left(1+\epsilon^{-2}\right)\left(\delta t^2\left(1+\epsilon^{-2}\right) - 4\right) &\leq 0, \\ \delta t &\leq \frac{2}{\sqrt{(1+\epsilon^{-2})}} \leq 2\epsilon. \end{aligned} \tag{2.40}$$

The region of stable step size is therefore very small when using this general method, scaling with ϵ. We plot the eigenvalues in the inset in Figure 2.5, they lie on the unit circle until the diverge across the real line after half a revolution.

If we instead use the splitting in (2.39) and evolve the system with the r-RESPA scheme $[\![\mathrm{sfs}]\!]$, we can compute the propagation matrix similarly as

$$\begin{aligned} \Psi_{[\![\mathrm{sfs}]\!],\delta t} &= \Psi_{s,\delta t/2}\,\Psi_{f,\delta t}\,\Psi_{s,\delta t/2}, \\ &= \begin{bmatrix} 1 & 0 \\ -\delta t/2 & 1 \end{bmatrix} \begin{bmatrix} \cos(\delta t/\epsilon) & \epsilon\sin(\delta t/\epsilon) \\ -\sin(\delta t/\epsilon)/\epsilon & \cos(\delta t/\epsilon) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\delta t/2 & 1 \end{bmatrix}. \end{aligned}$$

The same condition must hold for stability as in the ⟦AB⟧ scheme, so

$$|2\cos(\delta t/\epsilon) - \delta t\epsilon \sin(\delta t/\epsilon)| \leq 2,$$

and rewriting $\tau = \delta t/\epsilon$, we have instability where

$$\left|2\cos(\tau) - \tau\epsilon^2 \sin(\tau)\right| > 2. \tag{2.41}$$

Using a Taylor expansion around $\tau = \pi$, we see

$$2\cos(\pi + \delta\tau) - \tau\epsilon^2 \sin(\pi + \delta\tau) = -2 + \tau\epsilon^2\delta\tau + \mathcal{O}(\delta\tau^2),$$

and hence we expect is instability to occur in a region where $\delta\tau < 0$, corresponding to an upper bound on the step size threshold for the r-RESPA scheme as

$$\delta t < \pi\epsilon.$$

As we are solving the 'fast' part exactly for all δt, it is potentially unintuitive where the instability could be coming from in the method. The problem is exactly that of a mechanical resonance effect: the fast part of the system has a very short oscillatory period of $2\pi\epsilon$, applying the 'slow' force in sync with this (i.e. setting τ to be just below an integer multiple of π) creates (numerical) resonance.

This is a disappointing result; despite our splitting strategy separating the two timescales of the system, the maximum stable timestep is still limited by the overall fastest frequency, scaling exactly as in the Euler method. If we wish to push our timestep far enough to bridge the timescale gap, then this method does not offer a clear line of attack. Given the oscillatory nature of (2.41), we may hope to choose a timestep that lies in a stable region sandwiched between instability (see Figure 2.5), though the timestep τ increases, the island stable regions shrink, making choosing them a challenge. Of course we only study here the linear case, in the more practical nonlinear case we expect qualitatively similar behaviour, but a thorough prediction of stability regions in the general case varies from difficult to impossible.

Resonance effects plague many different strategies for overcoming timescale separation [9]. There are several approaches for overcoming such effects [43], most popularly averaging or mollification of the dynamics is used to remove resonances [58, 71, 77]. Alternately, we can remove the fastest degrees of motion completely by locking their motion so that are fixed. This employs the use of a constraint algorithm to, for example, maintain fixed bond lengths or dihedral angles in a molecule. The use of such algorithms is well-studied [102, 121].

Often in biomolecular modelling, the highest-frequency motions occur in the solvent, ironically outwith the molecule of interest. We could simply delete the solvent molecules from the simulation, simulating the molecule as if in a vacuum. This is appealing, as

removing the solvent would significantly reduce the total number of degrees of freedom in the model (and hence the simulation cost, which scales with N). However, such a simulation is completely nonphysical, ignoring much of the thermodynamic behaviour that might be of interest: the solvent molecules constantly barrage and buffet the biomolecule, resulting in an exploration of its configuration space that is impossible to sample in a vacuum constant energy simulation.

The challenge is therefore to take our target Hamiltonian system (the archetype being a biomolecule), and embed a *pseudosolvent* into the simulated dynamics; we employ an artificial device that mimics the behaviour of the solvent on average. We will make precise how we do this, and to what end, in the next section.

2.2 Sampling the canonical ensemble

2.2.1 Statistical mechanics

We have so far considered evolving a Hamiltonian system along a level-set of constant energy, with an emphasis on requirements for effective computer simulation (noting concerns such as computational cost and a disparity of timescales). In order for our numerical simulations to be commensurate with experiments done in a laboratory, we would like to be able to compute averages of various quantities that might be observable to experimentalists. We cannot expect that the trajectories we can compute with our constant energy simulations will immediately match up with real measurements using a much larger sample in a Petri dish, taken over a long time.

While we can compute averages from constant energy simulation as well, advancing the system from an initial condition along its level-set gives a limited *microscopic* perspective. In order to establish more general information about the system under study, we can imagine taking a large collection of initial conditions and propagating them all forward in time. We can imagine this bundle of trajectories in the phase space as being a cloud of points, not necessarily equally distributed in space or energy. The coloured regions in Figure 2.2 are one example.

This distribution of points will evolve in time under the dynamics, just as a point in phase space does. In fact, we can think of one point in phase space as being a distribution characterized by a δ function. The discrete bundle of initial conditions will approximate a continuous density, allowing us to evolve the points forward in time in order to compute the density. In general, we write a density $\hat{\rho}$ of points at time t as $\hat{\rho}(z,t)$, where $(q,p) = z \in \Omega$.

In this section we shall adopt a change of perspective, and consider how a disctribution of points is evolved in time under the dynamics, rather than a single trajectory. Much of this material is standard, and can be found in many statistical mechanics textbooks [4, 63, 74].

It should be apparent that the evolving density as we have described it is directly related to a probability measure - each discrete trajectory packet can be thought of as

one sample of our distribution. In order to make it a probability measure we impose some normalizing conditions:

$$\hat{\rho}(z,t) \geq 0 \quad \forall z \in \Omega, t \in \mathbb{R}, \quad \text{and} \quad \int_\Omega \hat{\rho}(z,t)\mathrm{d}z < \infty \quad \forall t \in \mathbb{R}.$$

This allows us to normalize $\hat{\rho}(z,t)$, such that

$$\rho(z,t) = Z^{-1}\hat{\rho}(z,t), \quad Z := \int_\Omega \hat{\rho}(z,t)\mathrm{d}z, \qquad \int_\Omega \rho(z,t)\mathrm{d}z = 1.$$

We shall assume that $\rho \in L^2$, so that computed inner products at a time t of any smooth function $\phi \in L^2$ are bounded

$$\langle \phi(z), \rho(z,t) \rangle = \int_\Omega \phi(z)\, \rho(z,t)\mathrm{d}z < \infty.$$

The value of $\langle \phi(z), \rho(z,t) \rangle$ gives the average of some observable $\phi(z)$ at a time t, with respect to an evolving probability distribution $\rho(z,t)$.

For a system with governing ODE

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t) = f(z),$$

Liouville's equation for the evolution of the phase density is defined as

$$\frac{\partial}{\partial t}\rho(z,t) = \mathcal{L}_f^* \rho(z,t) = -\nabla \cdot (f(z)\rho(z,t)). \tag{2.42}$$

We can see that the macroscopic evolution (or simply the Liouville) operator $\mathcal{L}_f^*$ looks similar (both in notation and definition) to the microscopic evolution operator, the Lie derivative $\mathcal{L}_f$, given in (2.25). The superscript asterisk denotes that they are adjoints, in the L^2 sense. Indeed, if we have observables $\phi, \theta \in L^2$ then

$$\begin{aligned}
\langle \phi(z), \mathcal{L}_f^* \theta(z) \rangle &= \int_\Omega \phi(z) \left[\mathcal{L}_f^* \theta(z)\right] \mathrm{d}z, \\
&= -\sum_{i=1}^{2N} \int_\Omega \phi(z) \frac{\partial}{\partial z_{(i)}} \left[f_{(i)}(z)\theta(z)\right] \mathrm{d}z, \\
&= \sum_{i=1}^{2N} \int_\Omega \frac{\partial}{\partial z_{(i)}} \left[\phi(z)\right] f_{(i)}(z)\theta(z)\mathrm{d}z, \\
&= \int_\Omega \left[f \cdot \nabla \phi(z)\right] \theta(z)\mathrm{d}z, \\
&= \langle \mathcal{L}_f \phi(z), \theta(z) \rangle,
\end{aligned}$$

where subscripts in parenthesis denote components.

In the case of Hamiltonian systems, we can write

$$\frac{\mathrm{d}}{\mathrm{d}t}z(t) = f(z) = J\nabla H(z),$$

where H is the Hamiltonian. Recall the dynamics are divergence-free, hence the Liouville operator becomes

$$\mathcal{L}_f^*\phi(z) = -\nabla\cdot(f(z)\phi(z)) = -\phi(z)\,(\nabla\cdot f(z)) - f(z)\cdot\nabla\phi(z) = -f(z)\cdot\nabla\phi(z) = -\mathcal{L}_f\phi(z),$$

implying that the operator is skew-adjoint. Hence the Liouvillian for Newton's equations of motion (2.4) $\mathcal{L}_{\mathrm{mc}}^*$ becomes

$$\mathcal{L}_{\mathrm{mc}}^* = -\mathcal{L}_{\mathrm{mc}} = \nabla U(q)\cdot\nabla_p - M^{-1}p\cdot\nabla_q, \tag{2.43}$$

where we describe such constant-energy simulations as *microcanonical*, indicated by the subscript 'mc'.

In order to make sense of numerically compute averages, when comparing to laboratory estimates we will think of our molecular dynamics experiment as being one realization (or sample) from some prescribed probability measure, or ensemble. This ensemble is established by deciding upon a set of characteristic parameters for the virtual laboratory experiment. We assume that the experiment is conducted in equilibrium, so that the overall ensemble density will be stationary with respect to the evolution of the system. These invariant measures for our system correspond to distributions ρ^* satisfying

$$\frac{\partial}{\partial t}\rho^* = \mathcal{L}^*\rho^* = 0, \tag{2.44}$$

where $\mathcal{L}^*$ is the macroevolution operator for our dynamics. In a constant energy simulation, the skew nature of the Liouvillian operator implies that any first integrals of the system satisfies (2.44), in particular any density that is a function of the Hamiltonian will be invariant.

For example, we may choose to fix the number of atoms N in a simulation, as well as defining a box of fixed volume V in $\mathbb{R}^d$, which our system can evolve in. The boundary conditions for this box will also need to be decided upon in order to replicate a sensible experiment, but for the most part we will assume that we use periodic boundary conditions so that the position space is restricted to a torus.

The system's evolution is then governed by prescribed initial energy $H(z) = E_0$. Averages computed with respect to these conditions gives us the microcanonical (or NVE) ensemble, with associated invariant measure

$$\hat{\rho}_{\mathrm{mc}}(z) := \delta\left[H(z) - E_0\right].$$

The spatial average of an observable ϕ in the microcanonical ensemble is then

$$\langle \phi, \rho_{\mathrm{mc}} \rangle = \int_{\Omega} \phi \, \rho_{\mathrm{mc}} \, \mathrm{d}z, \tag{2.45}$$

where ρ_{mc} is the normalized probability distribution

$$\rho_{\mathrm{mc}} = Z_{\mathrm{mc}}^{-1} \, \hat{\rho}_{\mathrm{mc}}, \qquad Z_{\mathrm{mc}} = \int_{\Omega} \hat{\rho}_{\mathrm{mc}}(z) \mathrm{d}z. \tag{2.46}$$

The normalization constant Z_{mc} is often called the microcanonical partition function. The goal of a typical simulation is to compute the average of an observable, with respect to some ensemble. This reduces to the task of evaluating the integral (2.45), where one strategy is to discretize the phase space Ω into a mesh, and approximate the integral by evaluating $\phi(z)\rho_{\mathrm{mc}}(z)$ at each point. The approximation could be improved if necessary by reducing the distance between the meshpoints.

The sting in the tail though comes when considering the enormous dimensionality of the problem, coupled with the sparse nature of ρ_{mc}. For even a modest sized number of particles, creating a grid to span the entire phase space is computationally daunting. For example, taking a naïve approach and considering a small cluster of 100 particles in 3D, the phase space is $\Omega = \mathbb{R}^{600}$. If we only use ten points along each degree of freedom, we still have 10^{600} gridpoints, which is an unworkable amount of data to process or even store ($1GB$ of storage can store around 10^8 double precision numbers). Even worse, most of the meshpoints considered will lie in regions of phase space that contributes little or nothing to the integral defining the average of interest (such parts of the phase space the system will seldom visit - the distribution will be small or vanishing at these points).

For small problems, and for systems where we have sufficient knowledge as to adapt our mesh, this strategy may be workable and practical (indeed we can and will do this to check results for small problems later in this thesis). However, we shall work in the general case for very large N, looking to use a molecular dynamics approach for the calculation of averages such as (2.45).

The primary alternative methodology to MD is a *Monte-Carlo* (MC) approach [4, 42, 97], which moves the atoms at random in accordance with the equilibrium distribution to be sampled. In contrast to MD, no information about the approximate trajectories is required in (or can be generated from using) MC, and hence MD is the preferred tool for understanding any time-dependent quantitative or qualitative behaviour of the system. Hybrid approaches that marry MD and MC have also proven successful in special cases [29], while the flexibility of MC approaches can provide many benefits to specific applications [47, 52].

Molecular dynamics allows us to compute thermodynamically consistent trajectories, with the aim of using these trajectories as a tool to sample the integral (2.45), and calculate averages. This is not as straightforward as it may sound, aside from

how to compute the integral from our numerically discretized trajectory (which is just a sequence of points), the system's dynamics itself may stymie our efforts. One can imagine that a trajectory may not be able to access the entirety of the measure ρ, and hence any information about one part of the integral will be lost, skewing the computed average. The link between the exploration of trajectories in time compared to space is given by the property of ergodicity. In the microcanonical setting, defining the surface of fixed energy $\Omega_E \subset \Omega$, we say a system is ergodic if, for any two sets $A, B \subseteq \Omega_E$ of nonzero Lebesgue measure,

$$\forall t > 0, \quad \Phi_t(A) = A \quad \text{and} \quad \Phi_t(B) = B \qquad \Longrightarrow \qquad A = B,$$

where Φ_t is the system's flow map. Ergodicity of the system means we cannot split the energy surface Ω_E into multiple disconnected invariant sets. Equivalently, this means that the system is ergodic if for any trajectory $z(t)$ with $z(0) \in \Omega_E$, we have

$$\text{(Time average)} \qquad \lim_{T\to\infty} \frac{1}{T} \int_0^T \phi(z(t))\mathrm{d}t = \int_\Omega \phi \, \rho_{\mathrm{mc}} \, \mathrm{d}z \qquad \text{(Space average)}$$

i.e. the average of an observable computed along a single trajectory is equal to the average computed from integrating over the entire space, so the time and space averages are equal. In the general case this requires the trajectory to be able to access the entirety of the measure Ω_E. As an example, consider $(q, p) \in \mathbb{R}^2$, with unit mass and potential energy function given by the triple well potential

$$U(q) = \phi_{\mathrm{tw}}(q) = q^2 \, (q^2 - 1)^2 + \frac{3}{20}(q + 1), \tag{2.47}$$

and consider sampling averages for $E = 0.35$. From Figure 2.6, we can see that the level sets corresponding to $H = E$ are disconnected, and a single trajectory cannot sample the measure appropriately. In such cases, the result of the time average will depend on which invariant set the initial condition is in.

Assuming that have an ergodic dynamics (an ergodic hypothesis), we can approximate the time average by averaging the observable along a discretized trajectory. In general it can be very difficult to prove that a particular system is ergodic, we shall demonstrate some provably ergodic examples in Chapter 3.

2.2.2 The Boltzmann-Gibbs distribution

Returning to our original motivation, we seek to sample the system in a more realistic setting than is provided in the constant-energy framework. We shall present a brief summary of the standard results presented in [4, 42, 72], introducing an alternate 'target distribution' for us to sample the system in a manner more consistent with laboratory experiments.

Consider a system of interest embedded in a much larger system, acting as a heat

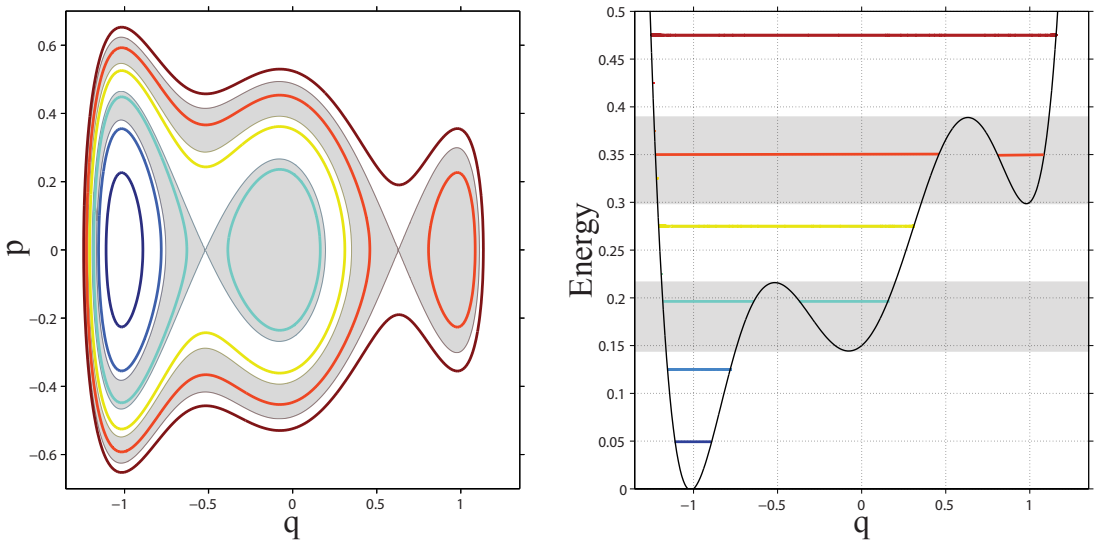


Figure 2.6: Level sets of the energy in phase space (left) are plotted for the system with potential energy given in (2.47), coloured according to their energy level. We also plot the contribution of the potential energy for each level set (right), as a function of the position. As both the kinetic and potential energy are bounded below, we can see in that for some energy levels a high potential energy barrier prevents movement between invariant sets. Level sets in the shaded regions correspond to energy levels E where the measure $\delta[H - E]$ can be partitioned into two disconnected regions, violating ergodicity.

bath. The two systems are assumed to be weakly coupled, so energy is continuously exchanged between them through a dance of jabbing and poking at each other. This energy is assumed to be heat, and therefore any equilibrium logically comes in the form of the systems having matching temperature. The sole transfer between the two is energy, they do not exchange particles or interact through any sort of mechanical process.

We call our system of interest the active system, or part A. The heat bath, or the bulk, is denoted part B. We will assume that the total system has some set of constant parameters defining its evolution, and attempt to "integrate out" over the degrees of freedom in the heat bath in order to recover a distribution for the particles in just the active system.

Let us denote the total energy of the combined system AB as simply E. If our active system has energy E_A then the energy of the heat bath is $E_B = E - E_A$. In line with quantum mechanics, we assume that for a constant number of particles, box volume and energy, there are a discrete number of states that a system can occupy (a 'state' here refers to an eigenfunction of the Hamiltonian operator, but the specifics are not essential). Fixing the box volume for the combined system, we can use our assumption that the two systems do not exchange particles so that the total number of states the combined system can be found in depends solely on the energy of the two

systems. We denote $Z(E_A, E_B)$ the total number of potential states the system can occupy with system A, B having energy E_A, E_B respectively. Combinatorically, we can write

$$Z(E_A, E_B) = Z_A(E_A) \times Z_B(E_B),$$

where $Z_i(E_i)$ denotes the number of states system i can be found in at energy E_i. Taking logarithms and replacing the value for E_B, we have

$$\ln Z(E_A, E - E_A) = \ln Z_A(E_A) + \ln Z_B(E - E_A).$$

The crucial assumption is that each of the states in $Z(E)$ is equally probable, and hence we can find the most likely energy distribution between the systems by maximizing the value of $Z(E_A, E_B)$, or equivalently maximizing $\ln Z(E_A, E - E_A)$. This value of E_A occurs where

$$\frac{\mathrm{d}}{\mathrm{d}E_A} \ln Z(E_A, E - E_A) = 0,$$

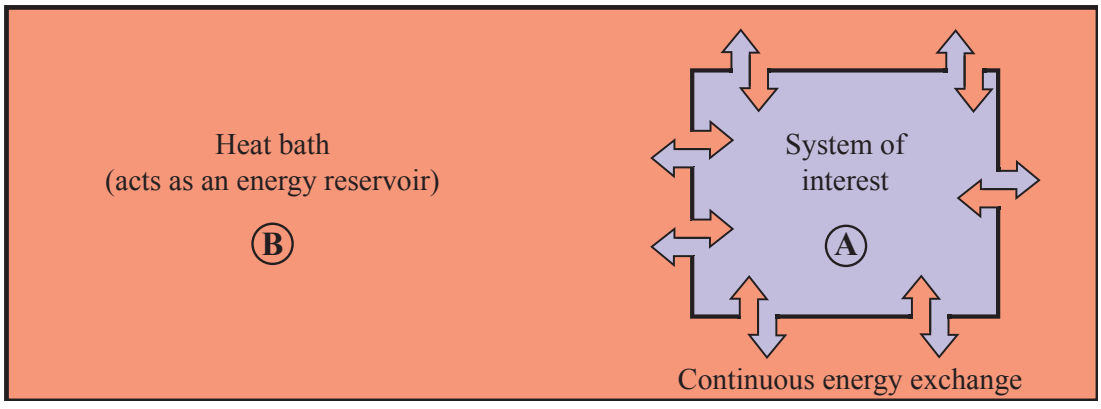
or equivalently where

$$\frac{\mathrm{d}}{\mathrm{d}E_A} \ln Z_A(E_A) = \frac{\mathrm{d}}{\mathrm{d}E_B} \ln Z_B(E_B).$$

We have assumed that our systems are in thermal equilibrium, but if we had not, and instead invested the heat bath with the total system energy, we could imagine that energy would transfer back into system A until the point where this equation is satisfied. We know from the second law of thermodynamics that this equilibrium occurs when the entropy of the system is at its maximum, so we use Boltzmann's famous equation (and in fact, his epitaph) relating the entropy of a system S to the number of eigenstates Z,

$$S \equiv k_B \ln Z,$$

where k_B is Boltzmann's constant $k_B \approx 1.38 \times 10^{-23} J/K$. The temperature of a system



T, is defined thermodynamically as

$$\frac{1}{T} := \frac{\partial S}{\partial E},$$

giving, when the system is in thermal equilibrium,

$$\frac{\mathrm{d}}{\mathrm{d}E_A} \ln Z_A(E_A) = \frac{\mathrm{d}}{\mathrm{d}E_B} \ln Z_B(E_B) = \frac{1}{k_B T} =: \beta. \tag{2.48}$$

We shall refer to the quantity β as the system's inverse or reciprocal temperature.

Consider now the probability P_i of finding the subsystem A in state i, with energy $E_A = e_i$. This is equal to the probability of finding the energy of the heat bath to be $E_B = E - e_i$, and hence

$$P_i = \frac{Z_B(E - e_i)}{\sum_j Z_B(E - e_j)},$$

where the sum in the denominator is over all possible states of the system A.

Expanding the logarithm in a series around E, we have

$$\ln Z_B(E - e_i) = \ln Z_B(E) - \beta e_i + \mathcal{O}(1/E),$$

using (2.48). Assuming the total energy of the system is very large, we truncate and exponentiate to find

$$P_i \propto \exp(-\beta\, e_i),$$

where all the constants are absorbed into the proportionality sign, as we will normalize the distribution anyway. Hence the probability of the system A being in a particular state is proportional to the energy of the state, divided by temperature and exponentiated. This gives the Boltzmann-Gibbs (or canonical) distribution for our system A, sampled at a constant equilibrium temperature T, written (when normalized) as

$$\rho_\beta(z) := Z_\beta^{-1} \exp(-\beta\, H(z)), \qquad Z_\beta := \int_\Omega \exp(-\beta\, H(z))\mathrm{d}z, \tag{2.49}$$

where the parameter $\beta = 1/k_B T$. This ensemble is sometimes called the NVT ensemble, due to the a constant number of particles, constant volume, and a constant (thermodynamical) temperature. As our energy state e_i is replaced by our system Hamiltonian $H(z)$, the energy level of our system should be able to fluctuate as if it were in equilibrium with our heat bath model.

If our Hamiltonian is of the form

$$H(q, p) = p^T M^{-1} p/2 + U(q),$$

for $(q, p) \in \Omega_q \times \Omega_p = \Omega$, as we have considered thus far, then the distribution is decomposable into

$$\rho_\beta(q, p) = \mu_\beta(q) \times \kappa_\beta(p),$$

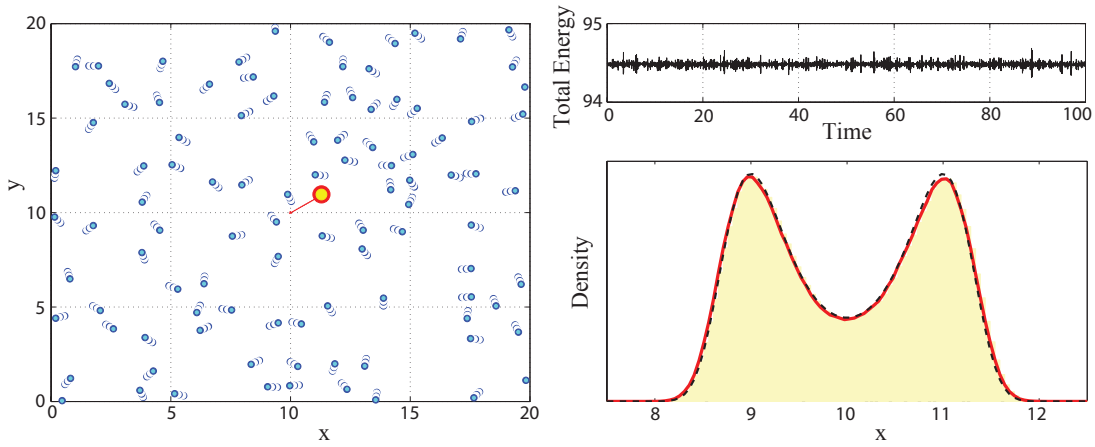


Figure 2.7: We simulate 100 particles in the plane (left) using a periodic box of length $L = 20$, with the Lennard-Jones potential governing all pairwise interactions. The central particle, marked in red, also feels a double-well potential in both coordinates. The total system energy (top right) is conserved, with some variance due to numerical error. The distribution is plotted for the x-coordinate (bottom right) averaged over a single long trajectory. The experimental result (solid red) shows good agreement with the expected canonically distributed result (dashed black).

where

$$\mu_\beta(q) = Z_{\beta,\mu}^{-1} e^{-\beta U(q)}, \qquad Z_{\beta,\mu} := \int_{\Omega_q} e^{-\beta U(q)} \mathrm{d}q,$$

and

$$\kappa_\beta(p) = Z_{\beta,\kappa}^{-1} e^{-\beta p^T M^{-1} p/2}, \qquad Z_{\beta,\kappa} := \int_{\Omega_p} e^{-\beta p^T M^{-1} p/2} \mathrm{d}p = \left(\frac{2\pi}{\beta}\right)^{N/2} (\det M)^{1/2}.$$

The variables q and p are hence probabilistically independent, with the momenta normally distributed. Equilibrium averages that are purely a function of momentum become trivial to compute, and our focus will turn to sampling functions of position with respect to μ_β, or canonical *configurational* sampling.

In order to sample from μ_β, one option would be to simulate the virtual particles we considered in the heat bath, and compute averages along the trajectory for system A. As an example, consider as our system of interest a single particle confined to the plane, and write its instantaneous position $q = (x, y)^T$. If we use periodic boundary conditions with a box length L, we will aim to sample canonically the potential

$$U(x, y) = \phi_{\mathrm{dw}}(x + L/2) + \phi_{\mathrm{dw}}(y + L/2),$$

where recall the double well potential $\phi_{\mathrm{dw}} : \mathbb{R} \to \mathbb{R}$ has minima at ± 1, tethering the particle near to the centre of the box.

In order to simulate the heat bath, we add a multitude of particles around our system of interest, drawing their initial velocities canonically from κ_β, for some prescribed β.

In our analysis we required the total energy of the system to be much larger than the energy of the target system, so we need a very large number of particles in the heat bath to provide a good approximation.

We will use the Lennard Jones potential to simulate all pairwise interaction, in order to provide the necessary jostling of adjacent particles required for the weak coupling (energy transfer) between the target system and heat bath. The Lennard Jones (or $6-12$ potential) is defined as

$$\phi_{\mathrm{lj}}(r) := r^{-12} - r^{-6},$$

where r is the Euclidean distance between any two distinct particles.

We initialise 100 particles on a lattice inside a box of length $L = 20$, with initial temperatures drawn from κ_β using $\beta = 1$. The Verlet $[\![\mathrm{BAB}]\!]$ scheme was used with a step size of $\delta t = 0.01$ to propagate the system in time, sampling until $T = 6 \times 10^4$. We compare the sampled distribution for the x coordinate of the target system with the analytical result of μ_β at $\beta = 1$. The results, shown in Figure 2.7, appear surprisingly good, despite errors from numerical discretization, as well as the finiteness of the heat bath and time interval.

For more general systems, we could estimate the error in sampling due to the finiteness of the heat bath by looking at the sampled distribution of momenta, and comparing it to κ_β. However q and p are independent in the canonical distribution, and so large or small errors in the distribution of one does not imply the same in the other. Of course, this simulation is extremely inefficient. Though the trajectories for the central particle appear to sample canonically, the heat bath particles outnumber the target system particles by 99 to 1. This makes it woefully impractical to scale this method for the simulation of a large cluster of atoms.

What would be preferable is if we could approximate the force terms coming from the bulk, without requiring us to simulate them. It should be clear that the constant-energy dynamics we have used so far (sampling the microcanonical ensemble with constant energy E) will not sample the target system canonically (at least, not ergodically) if we remove the heat bath - such Newtonian trajectories sampling ρ_{mc} cannot access regions of the phase space where $H(z) \neq E$. Regions of the phase space outside the level sets $H(z) = E$ will have nonzero measure in the canonical ensemble.

However, changing perspective to think of trajectories as tools for sampling a distribution, we are free to develop any dynamics we wish that samples the canonical distribution. Even grossly non-physical dynamics are permitted, as our equilibrium formulation does not describe the time evolution of the system [67] (we shall discuss such non-Newtonian dynamics in a later section).

We shall describe any dynamics that sample ρ_β as a thermostat, owing to the 'constant temperature' nature of the canonical ensemble.

2.2.3 The Nosé-Hoover thermostat

We first consider perturbing the vector field for Newtonian dynamics, with an aim to mimic the effect of the heat bath by adding or removing energy from the system. We shall introduce s additional degrees of freedom $\xi = [\xi_{(1)}, \xi_{(2)}, \ldots, \xi_{(s)}]^T \in \mathbb{R}^s$, which will act as an energy governor in the dynamics. The 'artificial' variables ξ have their own invariant distribution, assumed to be independent of the state variables q, p, such that

$$\rho(q, p, \xi) = \rho_\beta(q, p) \times \Xi_\beta(\xi) = \mu_\beta(q) \times \kappa_\beta(p) \times \Xi_\beta(\xi).$$

This distribution allows us to integrate out with respect to the artificial variables ξ and recover canonical sampling. Nosé introduced an extended Hamiltonian dynamics as a perturbation of Newtonian dynamics [89, 90], sampling such a distribution. Hoover later offered a reformulation and time-transformation of Nosé dynamics [55], which gives the Nosé-Hoover thermostat

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} q &= M^{-1} p, \\ \frac{\mathrm{d}}{\mathrm{d}t} p &= -\nabla U(q) - \xi\, p, \\ \frac{\mathrm{d}}{\mathrm{d}t} \xi &= \frac{1}{Q} \left(p^T M^{-1} p - N/\beta \right), \end{aligned} \tag{2.50}$$

for scalar $\xi \in \mathbb{R}$, where N is the dimensionality of q and p, and $Q > 0$ is a mass parameter to be chosen in order to optimize the rate of convergence of averages.

One immediate problem with the dynamics is that some initial conditions do not provide suitable trajectories, for example, consider the initial condition $(q_0, p_0, \xi_0) = (q^*, 0, \xi_0)$, where $\nabla U(q^*) = 0$. As the perturbed force $\xi\, p$ acts along the vector p, if we start at an equilibrium point for Newtonian dynamics with zero initial momentum, then the system remains at that point for all time. Such subsets of initial conditions are called Hoover holes, though as these sets have zero measure (the probability of $p = 0$ is zero) ergodicity does not suffer.

We may hope to prevent choosing such degenerate initial conditions by sampling the initial momenta canonically, though this strategy will not always save us. If initial conditions q_0, p_0 are chosen such that $\nabla U(q_0 + \lambda p_0) \times p_0 = 0$ for all scalar λ, then the system will remain confined to the space $[q_0 + \lambda p_0, \mu p_0, \xi]$ for $\lambda, \mu, \xi \in \mathbb{R}$.

We show that the Nosé-Hoover dynamics preserve the extended canonical distribution, given as

$$\rho_{\mathrm{nh}}(q, p, \xi) = \rho_\beta(q, p) \times \exp(-\beta Q \xi^2/2) \times \sqrt{\frac{Q\beta}{2\pi}}.$$

Recall that the evolution of distribution is given by the Liouville equation

$$\frac{\partial}{\partial t} \rho = \mathcal{L}^* \rho.$$

Hence we require

$$\frac{\partial}{\partial t}\rho_{\text{nh}} = \mathcal{L}^*_{\text{nh}}\,\rho_{\text{nh}} = 0,$$

where $\mathcal{L}^*_{\text{nh}}$ is the characteristic evolution operator given by the dynamics, and computed as the adjoint of the Lie derivative. For the Hamiltonian $H(q,p) = p^T M^{-1} p/2 + U(q)$, the dynamics (2.50) can be split up into

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} q \\ p \\ \xi \end{bmatrix} = \underbrace{J\nabla H(q,p)}_{f_1} + \underbrace{\begin{bmatrix} 0 \\ -\xi p \\ \frac{1}{Q}\left(p^T M^{-1} p - N/\beta\right) \end{bmatrix}}_{f_2}.$$

Using (2.26) we can split the operator similarly, writing $\mathcal{L}^*_{\text{nh}} = \mathcal{L}^*_{f_1} + \mathcal{L}^*_{f_2}$ for corresponding vector fields f_1 and f_2. The operator $\mathcal{L}^*_{f_1} = \mathcal{L}^*_{\text{mc}}$ comes from the Newtonian part of the splitting, and hence it will preserve any function of the Hamiltonian, including ρ_{nh}. It remains to demonstrate that $\mathcal{L}^*_{f_2}\,\rho_{\text{nh}} = 0$, where

$$\mathcal{L}^*_{f_2}\rho = \nabla_p \cdot (\xi\, p\, \rho) - \frac{1}{Q}\left(p^T M^{-1} p - N/\beta\right)\frac{\partial}{\partial \xi}\rho.$$

From its definition, we compute

$$\begin{aligned}
\mathcal{L}^*_{f_2}\rho_{\text{nh}} &= \sqrt{\frac{Q\beta}{2\pi}}\left(\nabla_p \cdot (\xi\, p\, \kappa_\beta(p)) + \beta\xi\left(p^T M^{-1} p - N/\beta\right)\kappa_\beta(p)\right)\mu_\beta(q) e^{-\beta Q \xi^2/2}, \\
&= \left(\xi(\, N - \beta p \cdot M^{-1} p) + \beta\xi\left(p^T M^{-1} p - N/\beta\right)\right)\rho_{\text{nh}}, \\
&= 0,
\end{aligned}$$

and hence the canonical distribution is preserved under the dynamics. Thus, assuming ergodicity, computed trajectories will sample the Boltzmann-Gibbs distribution, as the time average will equal the space average.

As already discussed, proving ergodicity for the dynamics can be a challenging task. The Nosé-Hoover thermostat has been shown to be non-ergodic in the case of $Q \to \infty$, but for more practical choices of the 'thermostat mass' parameter Q, there is strong numerical evidence against ergodicity [66]. The qualitative differences between the canonical distribution and computed distributions for systems of low-dimensionality, or particularly stiff systems can be dramatic. To try and resolve this ergodicity issue, the Nosé-Hoover chains dynamics were introduced [78], which add a multitude of additional artificial degrees of freedom to the system, and couple them together in an effort to more aggressively thermostat the dynamics. The artificial variables are coupled consecutively by index: adding s artificial variables, we denote ξ as a vector $\xi = [\xi_{(1)}, \xi_{(2)}, \ldots, \xi_{(s)}]^T$ and think of each $\xi_{(i)}$ variable as being a link in a harmonic chain. The dynamics is

then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}q &= M^{-1}p, \\
\frac{\mathrm{d}}{\mathrm{d}t}p &= -\nabla U(q) - \xi_{(1)}\, p, \\
\frac{\mathrm{d}}{\mathrm{d}t}\xi_{(1)} &= \frac{1}{Q_{(1)}}\left(p^T M^{-1} p - N/\beta - \xi_{(1)}\xi_{(2)}\right), \\
\frac{\mathrm{d}}{\mathrm{d}t}\xi_{(i)} &= \frac{1}{Q_{(i)}}\left(Q_{(i-1)}\xi_{(i-1)}^2 - 1/\beta - \xi_{(i)}\xi_{(i+1)}\right), \qquad 1 < i < s, \\
\frac{\mathrm{d}}{\mathrm{d}t}\xi_{(s)} &= \frac{1}{Q_{(s)}}\left(Q_{(s-1)}\xi_{(s-1)}^2 - 1/\beta\right),
\end{aligned} \tag{2.51}$$

where we recover the original Nosé-Hoover dynamics as a special case for $s = 1$. The dynamics can be shown to preserve the target distribution ρ_{nhc}, where

$$\rho_{\mathrm{nhc}}(q,p,\xi) = \rho_\beta(q,p) \times \exp\left(-\beta \sum_{i=1}^{s} Q_{(i)}\xi_{(i)}^2/2\right) \times \left(\frac{Q\beta}{2\pi}\right)^{s/2},$$

which will give canonical sampling of the (q,p) space, if the sampling is ergodic. In numerical experiments, ergodicity is shown to be markedly improved when using NHC over NH, particularly in the case of large s.

The choice of the thermostat masses $Q_{(i)}$ are also a central issue in the dynamics, we must choose sensible values for the s-many parameters. Ideally the dynamics should be ergodic for any choice of the masses, but in practice even in simple models we can easily find 'wrong' values in sensible ranges.

We demonstrate this by using the Nosé-Hoover chains thermostat (with $s = 2$) to canonically sample the system with Hamiltonian

$$H(q,p) = q^2/2 + p^2/2,$$

where $(q,p) \in \mathbb{R}^2$, and $\beta = 1$. The left plot in Figure 2.8 demonstrates the tail of a 10^9-step trajectory of points $[q_n, p_n, \xi_{(1),n}] \in \mathbb{R}^3$, computed using thermostat masses $Q_{(1)} = 0.07$ and $Q_{(2)} = 12$, propagated numerically at a small step size $\delta t = 5 \times 10^{-4}$. The parameter values chosen are reasonable for a simulation, but instead of the normally distributed cloud of points we would expect in canonical sampling (discretized points should be weighted according to ρ_{nhc}) we find the trajectory traces out a torus. The small step size used implies that the qualitative behaviour exhibited is not a numerical artifact, but instead a product of resonance effects in the system.

It is clear from the plot that the value of q^2 averaged along the trajectory will not equal the canonical space average of 1. We can see the effect of varying the thermostat masses by considering a 100×100 grid of points in a small region of parameter space $(Q_{(1)}, Q_{(2)}) \in \mathbb{R}^2$. We run an experiment for each parameter set defined by a gridpoint, and plot the resulting error in the time average in the right panel of Figure 2.8. It is clear that there is a well-defined region in the parameter space where the converged

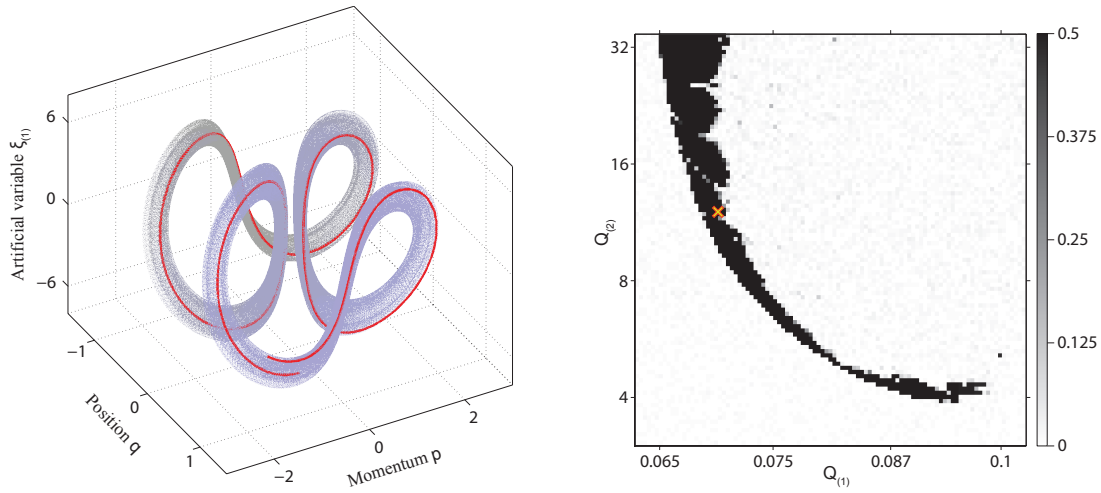


Figure 2.8: We plot points along a single trajectory (left) of the Nosé-Hoover chains dynamics (2.51) for the one-dimensional harmonic oscillator. We use $s = 2$, with thermostat masses $Q_{(1)} = 0.07$ and $Q_{(2)} = 12$ for initial values $q(0) = 1.075$, and $p(0) = \xi_{(1)}(0) = \xi_{(2)}(0) = 0$. A portion of the trajectory is plotted in red. Results were computed using a very small step size of $\delta t = 5 \times 10^{-4}$ to ensure numerical error is not the cause of any sampling errors. The absolute difference between the time and space averages of q^2 is plotted (right) in a 100×100 grid, where each gridpoint corresponding to a numerical experiment with masses given by the horizontal and vertical positions (using the same initial condition as the left plot). Gridpoints are coloured according to the difference between observed averages (along the trajectory) and canonical averages - ergodicity requires the two be commensurate. The yellow cross marks the gridpoint whose experiment corresponds to the left figure.

observed average along trajectories is starkly different to the analytic spatial average.

Though in more complicated systems we would perhaps not expect qualitative errors quite as pronounced as this example, the importance of parameter selection in this method cannot be overstated. By coupling deterministically to the heat bath and incorrectly choosing the masses we run the risk of introducing resonance phenomena that can corrupt averages and warp the trajectories.

The use of the Nosé-Hoover thermostat is still widespread however, as the ergodicity issues demonstrated are rarely so pronounced in large systems where the dynamics are not trapped close to one minima. Additionally, numerical experiments have shown that the thermostat is *gentle* [35], in the sense that computed dynamical quantities (such as correlation functions) have an error inversely proportional to the total system size. We demonstrate a dynamics in Section 3.2 that improves ergodicity without sacrificing this gentle quality.

Chapter 3

Stochastic thermostats

In order to resolve the issues of ergodicity, and generate trajectories sampling only the canonical distribution, we shall move away from deterministic dynamics and consider approximation of the external heat bath using a stochastic force term. We should expect that adding a random term to the dynamics in the correct way could banish ergodicity issues, as the stochasticity could be contrived to give a nonzero probability of transition between any two regions of the phase space.

3.1 Stochastic dynamics

In this section we shall develop some fundamental concepts and tools for stochastic analysis, to allow us to extend many of the machinery we have introduced in the previous chapter to the realm of stochastic differential equations. We present a brief summary of many of the core results in [44, 63, 86, 91]

3.1.1 The Fokker-Planck equation

We begin with the most fundamental term in stochastic calculus: the Wiener process $W(t)$. We can think of $W(t)$ as being a white noise term: a random walk with $W(0) := 0$ and on average (i.e. taking a large number of realizations of $W(t)$) $W(t) \sim \mathcal{N}(0, t)$. Wiener increments between times t and s are independent for non-overlapping intervals, with

$$W(t) - W(s) \sim \mathcal{N}(0, |t - s|) \sim \sqrt{|t - s|}\,\mathcal{N}(0, 1). \tag{3.1}$$

The independence of successive increments of the Wiener process gives rise to a Markovian property, in that its future probable behaviour depends solely on its current state, and not its history. This greatly simplifies the probabilistic analysis.

We shall consider the evolution of stochastic differential equations (SDEs) of the form

$$\mathrm{d}z = f(z)\,\mathrm{d}t + g(z)\,\mathrm{d}W, \tag{3.2}$$

for $z \in \mathbb{R}^N$, $f : \mathbb{R}^N \to \mathbb{R}^N$ and $g : \mathbb{R}^N \to \mathbb{R}^{N\times N}$. The vector $\mathrm{d}W$ represents an

infinitesimal Wiener increment of N random Wiener processes. In some literature, it is instead written as a $\sqrt{\mathrm{d}t}$ due to to the variance given in (3.1). If $g(z) \equiv \sigma$ for constant $\sigma \neq 0$, then we describe the SDE as using additive noise, whereas for non-constant $g(z)$ we have the more complicated multiplicative noise case enabling a wide spectrum of phenomena to be modelled. We shall primarily focus on additive noise in this thesis.

We now seek an analogue to the Lie derivative, governing the evolution of functions with respect to solutions of (3.2). Evidently, due to the stochastic nature of the solutions, we will instead consider the evolution of averages, rather than the evolution of specific trajectories.

For any continuously twice-differentiable function $\phi : \mathbb{R}^{2N} \to \mathbb{R}$, the Itō (or Itō-Doeblin) formula gives its evolution as

$$\mathrm{d}\phi = \nabla\phi \cdot (f(z)\,\mathrm{d}t + g(z)\,\mathrm{d}W) + \frac{1}{2}\left(g(z)g(z)^T\right) : \nabla^2\phi\,\mathrm{d}t, \tag{3.3}$$

where $A : B = \mathrm{trace}(AB^T) = \sum_i \sum_j A_i B_j$ defines the Frobenius product, and recall we denote $\nabla^2\phi = \nabla\nabla^T\phi$ as the Hessian matrix of second derivatives. Taking expectations, we find

$$\mathbb{E}\left[\mathrm{d}\phi\right] = \mathbb{E}\left[\nabla\phi \cdot f(z)\right]\,\mathrm{d}t + \frac{1}{2}\mathbb{E}\left[\left(g(z)g(z)^T\right) : \nabla^2\phi\right]\,\mathrm{d}t,$$

and hence defining an operator $\mathcal{L}$ such that

$$\mathcal{L}\phi = f(z) \cdot \nabla\phi + \frac{1}{2}g(z)g(z)^T : \nabla^2\phi, \tag{3.4}$$

we have the evolution of the expectation of a function ϕ governed by

$$\mathbb{E}\left[\mathrm{d}\phi\right] = \mathbb{E}\left[\mathcal{L}\phi\right], \qquad \text{or} \qquad \mathcal{L}\phi(z) = \lim_{t\to 0}\frac{\mathbb{E}[\phi(z(t))] - \phi(z(0))}{t}.$$

This operator $\mathcal{L}$ is the infinitesimal generator of the process (3.2) serves a similar purpose to Lie derivative in the deterministic setting (in fact, it reduces to the Lie derivative for $g(z) \equiv 0$). Similarly we seek an analogue to the Liouville equation (2.42) for the propagation of a distribution of initial conditions $\rho(z,t)$. For any time t, where ϕ grows at most polynomially fast as $z \to \pm\infty$, its expectation with respect to the distribution $\rho(z,t)$ is well defined:

$$\mathbb{E}\left[\phi(z(t))\right] = \langle\phi(z), \rho(z,t)\rangle = \int_\Omega \phi(z)\rho(z,t)\mathrm{d}z.$$

Differentiating with respect to t, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\phi(z(t))\right] = \int_\Omega \phi(z)\frac{\partial\rho}{\partial t}(z,t)\mathrm{d}z,$$

but as this quantity is given by $\mathbb{E}\left[\mathcal{L}\phi\right] = \langle\mathcal{L}\phi(z), \rho(z,t)\rangle$, taking the adjoint of the

operator $\mathcal{L}$ gives

$$\int_\Omega \phi(z)\frac{\partial \rho}{\partial t}(z,t)\mathrm{d}z = \langle \phi(z), \mathcal{L}^*\rho(z,t)\rangle,$$

where $\mathcal{L}^*$ described the L^2 adjoint mapping of the generator $\mathcal{L}$, easily computed as

$$\mathcal{L}^*\rho(z,t) = -\nabla_z \cdot (f(z)\rho(z,t)) + \frac{1}{2}\nabla_z^2 : g(z)g(z)^T\rho, \tag{3.5}$$

where ∇_z indicates derivatives with respect to the spatial dimensions. Thus, for the SDE (3.2) we have the Fokker-Planck (or forward Kolmogorov) equation

$$\frac{\partial}{\partial t}\rho(z,t) = \mathcal{L}^*\rho(z,t),$$

governing the evolution of distributions of trajectories.

A distribution $\rho(z)$ is described as invariant (or stationary) if $\mathcal{L}^*\rho(z) \equiv 0$. In the special case $g(z) \equiv 0$ the Fokker-Planck equation becomes the Liouville equation (2.42) with $\mathcal{L}^*$ reducing to the Liouvillian. Hence we can view $\mathcal{L}$ and $\mathcal{L}^*$ as appropriate generalizations of deterministic dynamics.

The solution to the SDE (3.2) is found exactly as in the deterministic case; continuing to treat integrals in an Itō sense, integration yields

$$z(t) = z(0) + \int_0^t f(z(s))\mathrm{d}s + \int_0^t g(z(s))\mathrm{d}W(s).$$

In order to evaluate the stochastic integral, we make use of the following standard proposition:

Proposition 3.6 (Itō Isometry)**.** *The stochastic integral $I(t) = \int_0^t b(s)dW(s)$ is itself a normally distributed stochastic process, with mean 0 and variance $\int_0^t b(s)^2 ds$.*

We omit the proof here but point the interested reader towards any standard text on SDEs, such as [44, 63].

Of particular use, and of central importance in this thesis, is the Ornstein-Uhlenbeck (OU) process for $z \in \mathbb{R}^N$, given as

$$\mathrm{d}z = -\gamma z\,\mathrm{d}t + \sigma\,\mathrm{d}W, \tag{3.7}$$

for a positive real constant γ and positive definite matrix $\sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$. Multiplying by $e^{\gamma t}$ we can write

$$\mathrm{d}\left(ze^{\gamma t}\right) = \sigma e^{\gamma t}\,\mathrm{d}W,$$

and hence integrating from 0 to t, we find

$$z(t) = e^{-\gamma t}\,z(0) + \sigma\int_0^t e^{\gamma(s-t)}\,\mathrm{d}W(s).$$

Using Proposition 3.6 to evaluate the integral, we have

$$\int_0^t e^{\gamma(s-t)}\,\mathrm{d}W(s) \sim \mathcal{N}\left(0, c(t)\right),$$

where

$$c(t) = \int_0^t \left(e^{\gamma(s-t)}\right)^2 \mathrm{d}s = \frac{1-e^{-2\gamma t}}{2\gamma}.$$

This allows us to write the exact solution to (3.7) for $t \geq 0$ as

$$z(t) = e^{-\gamma t}\, z(0) + \frac{\sigma}{\sqrt{2\gamma}} \sqrt{1-\exp(-2\gamma t)}\, R_t,$$

where we think of $R_t \sim \mathcal{N}(0,1)$. The Fokker-Planck (forward Kolmogorov) operator for this process $\mathcal{L}_{\mathrm{ou}}^*$ can be computed directly from the definition (3.5) as

$$\mathcal{L}_{\mathrm{ou}}^* \rho = \gamma \nabla \cdot (z\rho) + \frac{1}{2} \nabla \cdot \sigma \nabla \rho. \tag{3.8}$$

If we consider a distribution

$$\nu(z) \propto \exp\left(-z^T X z/2\right),$$

for diagonal positive-definite matrix X, then

$$\mathcal{L}_{\mathrm{ou}}^* \nu(z) \propto \gamma \left(N - z^T X z\right) \nu(z) + \frac{\sigma^2}{2} \left(-NX + z^T X^2 z\right) \nu(z),$$

and hence choosing $X = 2\gamma\sigma^{-2}$ we have

$$\mathcal{L}_{\mathrm{ou}}^* \nu(z) \equiv 0.$$

Notably the choice of $\sigma = \sqrt{2\gamma/\beta}\, M^{1/2}$ provides invariance if $X = \beta\, M^{-1}$. Hence for this choice of σ and arbitrary $\gamma > 0$ we have

$$\mathcal{L}_{\mathrm{ou}}^* \kappa_\beta(z) \equiv 0, \tag{3.9}$$

where recall κ_β defines the normalized canonical momentum distribution function

$$\kappa_\beta(p) = Z_{\beta,\kappa}^{-1}\, e^{-\beta p^T M^{-1} p/2}, \qquad Z_{\beta,\kappa} := \int_{\Omega_p} e^{-\beta p^T M^{-1} p/2} \mathrm{d}p = \left(\frac{2\pi}{\beta}\right)^{N/2} (\det M)^{1/2}.$$

The relationship between σ and γ describes the dynamics satisfying a fluctuation-dissipation law, balancing the $\sigma \mathrm{d}W$ (random fluctuation term) with the friction force γz (dissipation term).

If we imagine (3.7) as a process in the momenta p, then the SDE mimics the force term resulting from the barrage of surrounding solvent particles, along with the drag force proportional to its momentum. We can add this SDE term to Hamiltonian dynam-

ics (or equivalently replace the $-\xi p$ term in Nosé-Hoover dynamics) in the momenta, to replicate the forces the system feels from the heat bath. By choosing the magnitude of the fluctuation term to satisfy the fluctuation-dissipation law (which amounts to the correct choice for the matrix σ in the OU process in (3.7)) gives Langevin dynamics

$$\begin{aligned} \mathrm{d}q &= M^{-1}p\,\mathrm{d}t, \\ \mathrm{d}p &= -\nabla U(q)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma/\beta}\,M^{1/2}\,\mathrm{d}W. \end{aligned} \tag{3.10}$$

Due to the linearity of the Fokker-Planck operator in the deterministic part (Liouvillian), we can see that, splitting the vector field (3.10) into

$$\begin{bmatrix} \mathrm{d}q \\ \mathrm{d}p \end{bmatrix} = \underbrace{\begin{bmatrix} M^{-1}p \\ -\nabla U(q) \end{bmatrix}\mathrm{d}t}_{\mathrm{mc}} + \underbrace{\begin{bmatrix} 0 \\ -\gamma p\,\mathrm{d}t + \sqrt{2\gamma/\beta}\,M^{1/2}\,\mathrm{d}W \end{bmatrix}}_{\mathrm{OU}},$$

the Fokker-Planck operator for Langevin dynamics $\mathcal{L}^*_{\mathrm{LD}}$ is given by the sum of the microcanonical Liouvillian term (2.43) and the Ornstein-Uhlenbeck Fokker-Planck operator (3.8) acting in the momentum. The operator is then

$$\begin{aligned} \mathcal{L}^*_{\mathrm{LD}}\,\rho &= \mathcal{L}^*_{\mathrm{mc}}\,\rho + \mathcal{L}^*_{ou}\,\rho, \\ &= \nabla U(q)\cdot\nabla_p\rho - M^{-1}p\cdot\nabla_q\rho + \gamma\nabla_p\cdot(p\rho) + \frac{\gamma}{\beta}\nabla_p\cdot M\nabla_p\rho. \end{aligned} \tag{3.11}$$

We have already seen that $\mathcal{L}^*_{\mathrm{mc}}\,F(H(q,p)) \equiv 0$ for arbitrary F, and it is clear that $\mathcal{L}^*_{\mathrm{ou}}G(q)\kappa_\beta(p) \equiv 0$ for arbitrary G, as the OU process only acts in the momentum. Hence using (3.9) we plainly have

$$\mathcal{L}^*_{\mathrm{LD}}\,\rho_\beta(q,p) \equiv 0, \tag{3.12}$$

where ρ_β is the canonical Boltzmann-Gibbs distribution given in (2.49). Of course this statement alone is not all that meaningful for the purposes of canonical sampling: both Nosé-Hoover dynamics and the constant energy microcanonical dynamics also preserve this distribution. Recall that the preservation of the canonical measure along with ergodicity are necessary for the time average of sampled dynamical trajectories to be commensurate with the spatial average with respect to the weighting ρ_β, the latter being the ultimate goal for our molecular dynamics simulations.

3.1.2 Hörmander's condition

We would expect that, due to the stochastic nature of the dynamics, evolution under Langevin dynamics (4.2) provides a means for the system to access almost any state $z \in \Omega$ (given sufficient smoothness of $U(q)$). The implication is that ergodicity will benefit–how can we separate our phase space into two or more disconnected pieces, if the momentum could take almost any value? We shall describe a condition for

ergodicity by checking that this stochastic perturbation is coupled sufficiently to the system to allow it to access all regions of the phase space Ω.

We have seen that the Fokker-Planck operator (3.5) governing the evolution of the measure for an SDE is a second order partial differential operator with C^∞ coefficients. Any second order differential operator with C^∞ coefficients $\mathcal{L}$ is called hypoelliptic on an open subset $U \subset \Omega$ if, for all distributional solutions ρ satisfying $\mathcal{L}\rho \equiv 0$, ρ is C^∞. A hypoelliptic Fokker-Planck operator implies that a nontrivial C^∞ invariant solution is unique (up to a constant multiple), as if there were multiple nontrivial solutions then we could construct a non-smooth solution violating its hypoellipticity [85, 98]. A sufficient condition for the hypoellipticity of the Fokker-Planck operator of (3.2) comes from Hörmander's condition.

Definition 3.13. *For an open set $U \subset \Omega$, we say the C^∞ vector fields*

$$g_0, g_1, \ldots, g_N, \qquad g_i : U \to \Omega$$

satisfy Hörmander's condition at a point $z \in U$ if the ideal generated by the vector fields at a point z constitutes a basis for the phase space Ω,

$$\Omega \subseteq \mathit{span}\left(g_0, g_1, \ldots g_N, [g_0, g_1], [g_0, g_2], \ldots, [g_N, g_{N=1}], [g_0, [g_0, g_1]], \ldots\right).$$

The Lie bracket for two vectors u and v is given as $[u, v] = v'u - u'v$, where the u' is the Jacobian of the vector field u. For large systems, computing the basis vectors for general $U(q)$ can be a daunting task. The central application of Hörmander's condition is Hörmander's theorem [56, 88].

Theorem 3.14 (Hörmander)**.** *If the SDE is given as* (3.2)*, with the vector field $f(z)$ and each of the columns of the matrix $g(z)$ satisfying Hörmander's condition, then its associated Fokker-Planck evolution operator $\mathcal{L}^*$ is hypoelliptic. Hence the partial differential equation*

$$\mathcal{L}^*\rho \equiv 0$$

has a unique C^∞ solution.

The strategy for proving ergodicity in the stochastic context is straightforward, if repetitive. We aim to show the Lie brackets of combinations of the deterministic vector field $f(z)$ and the columns of $g(z)$ in (3.2) provide a basis for the phase space. For example, we consider Langevin dynamics (3.10) with $q, p \in \mathbb{R}^N$, where, comparing with Definition 3.13, we have

$$g_0 = \begin{bmatrix} M^{-1}p \\ -\nabla U(q) - \gamma p \end{bmatrix}, \qquad g_i = \sqrt{\frac{2\gamma}{\beta}} \begin{bmatrix} 0 \\ M_i^{1/2} \end{bmatrix} \quad 1 \le i \le N,$$

for positive constants γ, β, and positive definite mass matrix M where M_i denotes the i^{th} column of M. The constant vectors g_i for $i > 0$ span half of the phase space on

their own (the momentum space $\mathbb{R}^N$), but we require the whole of $\Omega = \mathbb{R}^{2N}$ to have a basis for ergodicity. We compute for $1 \leq i \leq N$

$$[g_0, g_i] = 0 - g_0' g_i = \sqrt{\frac{2\gamma}{\beta}} \begin{bmatrix} 0 & M^{-1} \\ -\nabla^2 U(q) & -\gamma I_N \end{bmatrix} \begin{bmatrix} 0 \\ M_i^{1/2} \end{bmatrix} = \sqrt{\frac{2\gamma}{\beta}} \begin{bmatrix} M_i^{-1/2} \\ -\gamma M_i^{1/2} \end{bmatrix},$$

and hence the span of the vectors g_i and $[g_0, g_i]$ clearly provide a full basis for the phase space Ω, verifying Hörmander's condition. The Langevin dynamics operator $\mathcal{L}^*_{\mathrm{LD}}$ is therefore hypoelliptic by Theorem 3.14, and hence by (3.12) ρ_β is the unique invariant solution, assuming smoothness of the potential energy function $U(q)$. We expect that Langevin dynamics trajectories will sample the canonical distribution ergodically. What's more, unlike the case of Nosé-Hoover chains in Figure 2.8 we have considerable flexibility in our choice of parameter, as for all values of $\gamma > 0$ Hörmander's condition is satisfied.

3.2 Nosé-Hoover Langevin dynamics

The Nosé-Hoover Langevin dynamics, introduced by Samoletov, Dettmann and Chaplain in [103], couples an Ornstein-Uhlenbeck process to the artificial variable in Nosé-Hoover dynamics (2.50), in an effort to improve the lack of ergodicity demonstrated in Figure 2.8 whilst still sampling canonically. In the deterministic setting, the behaviour of the Nosé-Hoover and Nosé-Hoover Chain thermostats were sensitive to the choice of mass parameters Q, with invariant tori observed in simple models trapping trajectories in regions of the phase space. Adding the OU process to the artificial variable instead of the momentum (as in Langevin dynamics (4.2)) gives a milder perturbation to the dynamics as the particles do not feel a direct random force that leads to fast decorrelation. The result is a thermostat that is both provably ergodic (unlike Nosé-Hoover) and gentle (unlike Langevin dynamics) in that there are smaller random perturbations to the dynamics [70]. Computational benefits may be derived also as the number of random values required is independent of the number of degrees of freedom N (as ξ is a scalar).

The relevant SDEs become

$$\begin{aligned} \mathrm{d}q &= M^{-1} p \, \mathrm{d}t, \\ \mathrm{d}p &= -\nabla U(q) \, \mathrm{d}t - \xi p \, \mathrm{d}t, \\ \mathrm{d}\xi &= Q^{-1}(p^T M^{-1} p - N/\beta) \, \mathrm{d}t - \gamma \xi \mathrm{d}t + \sigma \, \mathrm{d}w. \end{aligned} \tag{3.15}$$

where w is a scalar Wiener process. If a distribution ρ is a stationary solution of the Fokker-Planck equation for these dynamics, then

$$\mathcal{L}^*_{\mathrm{nhl}} \rho = (\mathcal{L}^*_{\mathrm{nh}} + \mathcal{L}^*_{\mathrm{ou}}) \, \rho = 0,$$

where the $\mathcal{L}^*_{\text{ou}}$ operator acts on the artificial variable ξ in this context. It was shown in Section 2.2.3 that the extended canonical distribution

$$\rho_{\text{nh}}(q,p,\xi) = \rho_\beta(q,p) \times \exp(-\beta Q \xi^2/2) \times \sqrt{\frac{Q\beta}{2\pi}},$$

was invariant under Nosé-Hoover dynamics, and hence $\mathcal{L}^*_{\text{nh}}\rho_{\text{nh}} = 0$. It is easily shown that choosing $\sigma = \sqrt{2\gamma/(Q\beta)}$ in (3.15) gives $\mathcal{L}^*_{\text{ou}}\rho_{\text{n}h} = 0$, and hence

$$\mathcal{L}^*_{\text{nhl}}\rho_{\text{nh}} = 0.$$

We summarize the proof of [70], demonstrating the ergodicity of the dynamics in the case of a quadratic potential $U(q) = q^T Bq/2$, by applying Theorem 3.14 with the results in [70]. For simplicity, we can choose $\beta = \gamma = Q = 1$, with $M = I_N$ (this amounts to a change of coordinates). Additionally we select a coordinate system such that B is diagonal, with $B = \text{diag}(\omega_1, \omega_2, \ldots, \omega_N)$ where each ω_i is positive and distinct. The deterministic and stochastic vector fields become

$$g_0 = \begin{bmatrix} p \\ -Bq - \xi p \\ p \cdot p - N - \xi \end{bmatrix}, \qquad \text{and} \qquad g_1 = \begin{bmatrix} 0 \\ 0 \\ \sqrt{2} \end{bmatrix}.$$

We can then compute the Lie brackets

$$X_0 = -\frac{1}{\sqrt{2}}\left[g_1, \xi\, g_1/\sqrt{2} + g_0\right] = \begin{bmatrix} 0 \\ p \\ 0 \end{bmatrix},$$

$$X_1 = g_0 - (p \cdot p - N - \xi)\, g_1/\sqrt{2} - \xi X_0 = \begin{bmatrix} p \\ -Bq \\ 0 \end{bmatrix},$$

$$Y_1 = [X_0, X_1] = \begin{bmatrix} p \\ Bq \\ 0 \end{bmatrix},$$

and define for $k \geq 1$

$$Z_k = \frac{1}{2}\left[Y_k, X_1\right], \qquad Y_{k+1} = -\frac{1}{2}\left[Z_k, X_1\right].$$

It can then be shown by induction that

$$Y_k = \begin{bmatrix} B^{k-1}p \\ B^k q \\ 0 \end{bmatrix}, \qquad Z_k = \begin{bmatrix} B^k q \\ -B^k p \\ 0 \end{bmatrix},$$

where it easily checked that the basic case is satisfied for $k = 1$, and hence assuming

true for k we have

$$Y_{k+1} = -\frac{1}{2}\left[Z_k, X_1\right] = \begin{bmatrix} B^k p \\ B^{k+1} q \\ 0 \end{bmatrix},$$

and similarly

$$Z_{k+1} = \frac{1}{2}\left[Y_{k+1}, X_1\right] = \begin{bmatrix} B^{k+1} q \\ B^{k+1} p \\ 0 \end{bmatrix},$$

as required. Due to the g_1 vector spanning the ξ space, it is evidently enough to show that

$$\operatorname{span}(Y_1, Z_1, Y_2, Z_2, \ldots, Y_N, Z_N) = \mathbb{R}^{2N},$$

which is equivalent to the condition that for any $(x, y) \in \mathbb{R}^{2N}$, we can find coefficients $(a, b) \in \mathbb{R}^{2N}$ such that

$$\sum_{k=1}^{N} (a_k Y_k + b_k Z_k) = (x, y).$$

We can rewrite this condition as

$$\begin{bmatrix} \operatorname{diag}(B^{-1}p) & \operatorname{diag}(q) \\ \operatorname{diag}(q) & -\operatorname{diag}(p) \end{bmatrix} \begin{bmatrix} \mathcal{V} & 0 \\ 0 & \mathcal{V} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}, \tag{3.16}$$

where $\mathcal{V}$ is the $N \times N$ Vandermonde matrix given by $\mathcal{V}_{ij} = \omega_i^j$, with nonzero determinant. Hence we can solve (3.16) as long as its leftmost matrix has nonzero determinant, which requires

$$(q, p) \in U = \left\{(q, p) \,\middle|\, \left(p_i^2/\omega_i + q_i^2\right) \neq 0 \quad \forall\, 1 \leq i \leq N\right\} \subset \mathbb{R}^{2N}.$$

Hence the dynamics (3.15) will be ergodic on the space $U \times \mathbb{R}$. The complement of U in $\mathbb{R}^{2N}$ is precisely the Hoover holes, discussed in Section 2.2.3, which is unsurprising given the perturbation made to the original Nosé-Hoover dynamics was solely through the artificial variable. We could consider adding another OU process to the momenta if we wished to wipe out the Hoover holes, though. as the measure of the inaccessible set will be 0, we consider NHL ergodic on $\mathbb{R}^{2N+1}$ for this potential, where we choose appropriate initial conditions $(q(0), p(0), \xi(0)) \in U \times \mathbb{R}$ for sampling.

The harmonic case (where $U(q)$ quadratic) is one of the most relevant for molecular dynamics as many molecular models contain strong bonds modelled with a linear restoring force, that will prevent the thermostat from thermalizing the system efficiently. For more general potential energy functions it is conjectured that ergodicity is retained, but the more involved calculation for general $U(q)$ makes computation of the iterated Lie brackets challenging. This is contrary to Langevin dynamics, where the strong coupling of the OU process to the momenta made ergodicity much easier to prove using Hörmander's condition.

However, it is precisely the mild stochastic coupling that makes NHL dynamics often a good choice for a thermostatted dynamics. It introduces a relatively small perturbation to the overall dynamics that makes for a more gentle thermostat than Langevin dynamics (4.2) which can corrupt correlation functions for modest choices of friction. If the goal is to calculate system transport constants using Green-Kubo like formulas (such as integrated correlation functions in the case of the diffusion constant) then the NHL thermostat provides ergodic dynamics suitable for computing such quantities. The trade-off is that the system will approach equilibrium more slowly compared to Langevin dynamics, which thermalizes quickly due to the stronger coupling in every degree of freedom.

3.3 Non-Newtonian dynamical sampling

If equilibrium sampling is our goal, then in principle we may choose any ergodic dynamics that samples canonically. We may use any artificial variables we wish, as long as we can recover ρ_β when taking the marginal distribution (i.e. integrating out). Specifically, we seek to sample functions purely of position, then we need not even use momentum at all, as in the case of Brownian (or Smoluchowski) dynamics [32].

Methods of sampling canonically in equilibrium can be devised that use dynamical information to evolve the system, but artificially regulate the temperature at each step, such as velocity rescaling or the Andersen thermostat [5]. Alternately methods can be devised for sampling that are completely dynamics-free, the most popular being Monte Carlo (MC) methods [47, 52, 75].

Our focus however will remain on dynamical thermostat methods, where system trajectories are governed solely by differential equations and require no external or auxiliary devices to maintain sampling. The thermostat methods introduced so far have been built upon the Newtonian dynamics encountered in Chapter 2, where in the case of Langevin and Nosé-Hoover Langevin we have added an Ornstein-Uhlenbeck process (3.7) to the dynamics to improve ergodicity. There is no reason to only perturb Newtonian dynamics if our sole goal is equilibrium sampling, we can choose a dynamics that preserves ρ_β and add OU processes in any degrees of freedom we wish (as long as they are normally distributed) in an effort to sample ergodically (such as GBK schemes introduced in [67]). There are many options for dynamical systems which leave invariant the canonical measure, one may consider arbitrary extensions of the dynamical system by the addition of auxiliary variables, together with interactions which have suitable invariance properties. Potentially these new dynamics could benefit from improved stability (in terms of a discretized step size threshold), efficiency (in terms of rate-of-convergence towards equilibrium) or features amenable to computational exploitation.

To illustrate, we introduce a new GBK method *stochastic line sampling* (SLS) dy-

namics as

$$\begin{bmatrix} \mathrm{d}q \\ \mathrm{d}p \\ \mathrm{d}\xi \end{bmatrix} = \underbrace{\begin{bmatrix} \xi M^{-1} p \\ 0 \\ -Q^{-1}\, p^T M^{-1} \nabla U(q) \end{bmatrix}}_{\text{sls:det}} \mathrm{d}t + \underbrace{\begin{bmatrix} 0 \\ -\gamma p\, \mathrm{d}t + \sqrt{2\gamma/\beta}\, M^{1/2}\, \mathrm{d}W \\ -\lambda \xi\, \mathrm{d}t + \sqrt{2\gamma'/(Q\beta)}\, \mathrm{d}w \end{bmatrix}}_{\text{sls:ou}}, \tag{3.17}$$

for position $q \in \mathbb{R}^N$, momentum $p \in \mathbb{R}^N$ and an artificial variable $\xi \in \mathbb{R}$ with real positive parameters γ, λ and Q, where β is the inverse temperature and there are independent noise processes W (a vector of length N) and w (a scalar Wiener process). Noting that the OU process defining the fluctuations in p is completely decoupled from q and ξ, it is perhaps more natural to think of the dynamics as a random walk on the reduced space (q, ξ), where its direction of travel is generated by an auxiliary OU process.

For small γ the momentum vector varies slowly, and we think of the position q sampling along the line p in phase space, where the scalar artificial variable ξ determines the velocity in the direction p. Thus we refer to these dynamics as line sampling.

In order to show that these dynamics sample the canonical position distribution $\mu_\beta(q)$ ergodically, we shall demonstrate that the unique stationary solution to its Fokker-Planck equation ρ_{sls} is the distribution

$$\rho_{\text{sls}}(q, p, \xi) = \rho_{\text{nh}}(q, p, \xi) = \rho_\beta(q, p) \times \exp(-\beta Q \xi^2/2) \times \sqrt{\frac{Q\beta}{2\pi}}. \tag{3.18}$$

The evolution of measure for the dynamics is governed by its Fokker-Planck operator, computed as the sum of the operators corresponding to the additive splitting given in (3.17),

$$\mathcal{L}^*_{\text{sls}} = \mathcal{L}^*_{\text{sls:det}} + \mathcal{L}^*_{\text{sls:ou}}.$$

It is obvious that, as the sum of distinct Ornstein-Uhlenbeck processes,

$$\mathcal{L}^*_{\text{sls:ou}} f_1(q) \exp(-\beta p^T M^{-1} p/2 - \beta Q \xi^2/2) \equiv 0,$$

for arbitrary $f_1(q)$, and hence $\mathcal{L}^*_{\text{sls:ou}} \rho_{\text{sls}} \equiv 0$. It remains to show that the deterministic part also preserves the prescribed measure. Taking the Lie derivative, we can see that

$$\mathcal{L}^*_{\text{sls:det}} = -\xi M^{-1} p \cdot \nabla_q + Q^{-1} p^T M^{-1} \nabla U(q) \frac{\partial}{\partial \xi},$$

and hence, for arbitrary $f_2(p)$ we have

$$\begin{aligned}
&\mathcal{L}^*_{\text{sls:det}} f_2(p) \exp(-\beta U(q) - \beta Q\xi^2/2) \\
&\qquad = f_2(p) Q^{-1} p^T M^{-1} \nabla U(q) \frac{\partial}{\partial \xi} \exp(-\beta U(q) - \beta Q\xi^2/2) \\
&\qquad\quad - f_2(p) \xi M^{-1} p \cdot \nabla_q \exp(-\beta U(q) - \beta Q\xi^2/2), \\
&\qquad = \left(\beta M^{-1} p \cdot \nabla U(q) - p^T M^{-1} \nabla U(q)\right) \xi \exp(-\beta U(q) - \beta Q\xi^2/2) g(p), \\
&\qquad = 0.
\end{aligned}$$

Therefore $\mathcal{L}^*_{\text{sls:det}} \rho_{\text{sls}} \equiv 0$ and hence ρ_{sls} is preserved under the dynamics (3.17). To show that we sample ergodically we apply Theorem 3.14 with

$$b_0 = \begin{bmatrix} \xi M^{-1} p \\ -\gamma p \\ -Q^{-1}\, p^T M^{-1} \nabla U(q) - \gamma' \xi \end{bmatrix}, \; b_i = \begin{bmatrix} 0 \\ \sqrt{\frac{2\gamma}{\beta}}\, M_i^{1/2} \\ 0 \end{bmatrix}, \; b_{N+1} = \begin{bmatrix} 0 \\ 0 \\ \sqrt{\frac{2\gamma}{Q\beta}} \end{bmatrix},$$

where $1 \le i \le N$, denoting the i^{th} column of M by M_i. Defining

$$c_0 = [b_{N+1}, b_0] + \frac{\lambda}{\sqrt{2\lambda/(Q\beta)}} b_{N+1} = \begin{bmatrix} \sqrt{2\lambda/(Q\beta)} M^{-1} p \\ 0 \\ 0 \end{bmatrix},$$

it is clear that Hörmander's condition is satisfied for all $(q, p, \xi) \in \Omega$, as

$$\text{span}\,(b_1, b_2, \ldots b_N, b_{N+1}, c_1, c_2, \ldots, c_N) = \mathbb{R}^{2N+1} = \Omega,$$

where

$$c_i = [c_0, b_i] = \begin{bmatrix} 2\beta^{-1} \sqrt{\gamma\lambda/(Q)} M_i^{-1/2} \\ 0 \\ 0 \end{bmatrix},$$

for $1 \le i \le N$. Hence the operator $\mathcal{L}^*_{\text{sls}}$ is hypoelliptic, and we expect the stochastic line sampling dynamics to be ergodic with respect to the distribution ρ_{sls}. Due to the statistical independence of each of the state variables, we can integrate out with respect to ξ to recover ergodic canonical sampling with respect to ρ_β.

We compare the trajectories and invariant distributions of the non-Newtonian thermostat (3.17) with the three Newtonian-based thermostats: Langevin dynamics (4.2), Nosé-Hoover dynamics (2.50) and Nosé-Hoover Langevin (3.15). We seek to canonically sample a one-dimensional system with double-well potential energy function given as

$$U(q) = (q^2 - 1)^2.$$

Choosing parameters $\beta = M = \gamma = \lambda = Q = 1$ for the methods, we select arbitrary initial conditions (away from Hoover holes) $q(0) = p(0) = 1$ and $\xi(0) = 0.5$. We choose

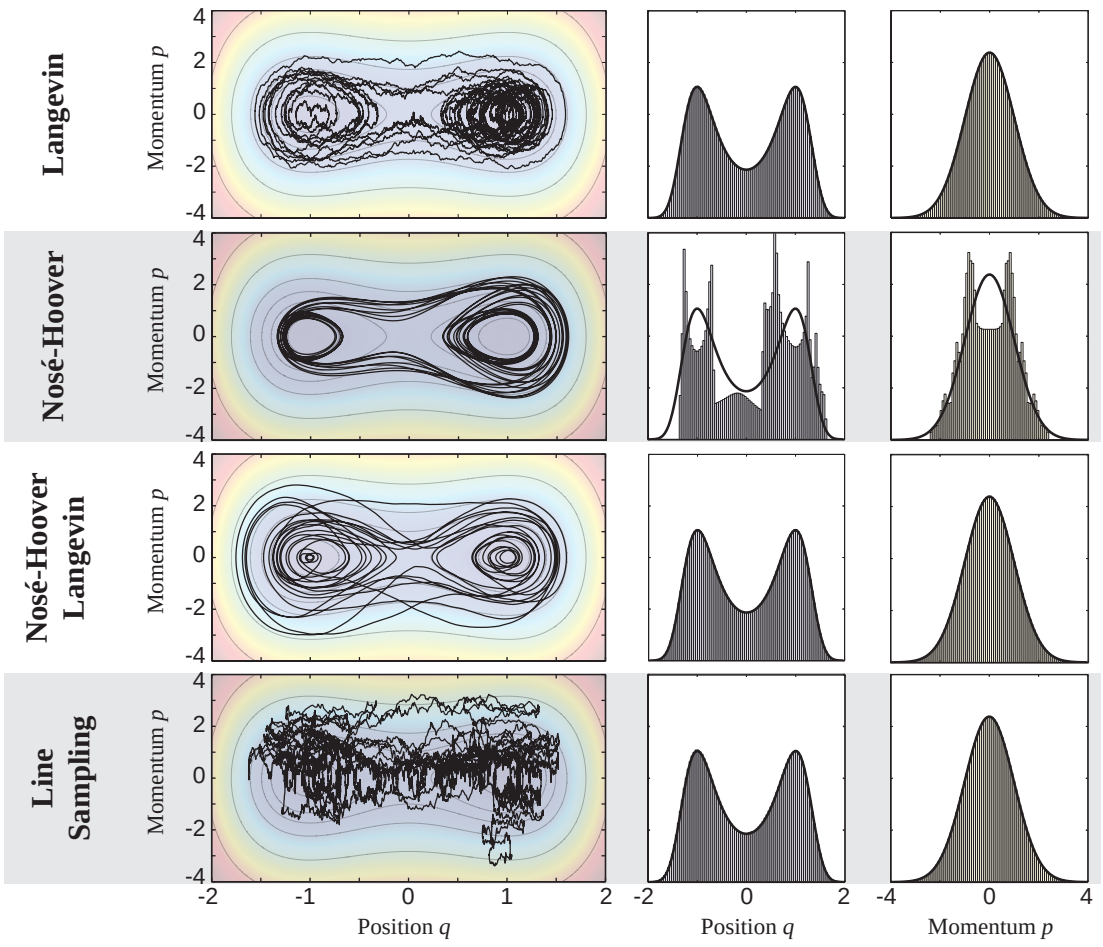


Figure 3.1: We plot trajectories (left) and long-time distributions (right) for a number of thermostatted dynamics. A lack of ergodicity in the case of the deterministic Nosé-Hoover thermostat prevents sampling ergodically, when compared to the exact distributions in bold.

our timestep $\delta t = 0.01$ sufficiently small to eliminate the possibility of numerical errors warping the results.

In Figure 3.1 we give a realization of a trajectory computed using each dynamics, in the $[q, p]$ plane. The Nosé-Hoover and Nosé-Hoover Langevin trajectories are significantly smoother than the others, indicating that these offer much more gentle thermostat techniques to reduce the perturbation to the dynamics. However, the Nosé-Hoover thermostat appears to remain stuck in a periodic cycle preventing exploration of the space. This behaviour persisted using other initial conditions and parameter values, indicating that it is symptomatic of a lack of ergodicity in the system.

By contrast, it is clear that the non-Newtonian thermostat explores the space in a very different way to the Newtonian-based thermostats. However, even in this small model we find that this is not necessarily advantageous, as the space is explored very slowly along the line p (whose direction also changes slowly). It is possible that this

different type of exploration of the phase space could be advantageous in certain models (for example the linear stability threshold for step size may be different), however we find that exploiting this difference may involve tuning the three parameters γ, Q and λ, making it more difficult to work with than Langevin dynamics (which only requires one parameter γ). As such, we think of Stochastic Line Sampling as an illustrative example, rather than as a practical option.

We also plot a histogram of the distributions of the position and momentum in Figure 3.1, computed from a single trajectory of 10^9 discrete points. Overlaying the exact expected result ρ_β, we can see the three stochastic methods perform excellently, whereas the results from using the Nosé-Hoover thermostat reveal extremely poor sampling as a result of the trajectory being trapped.

We would expect that using Nosé-Hoover Chains (2.51) may provide deterministic, ergodic sampling, by using a sufficiently long chain of artificial variables and correctly chosen mass parameters Q_i. However, the choice of these parameter values may prove crucial to the ergodicity of the dynamics, was shown in Figure 2.8. Therefore, we suggest that for equilibrium canonical sampling it is far more efficient to use a stochastic thermostat, in comparison to the ergodicity issues and additional parameter tuning required in the case of deterministic thermostats.

Chapter 4

Splitting methods for Langevin dynamics

In this chapter we provide a framework for the development of suitable numerical methods for Langevin dynamics and rigorous numerical analysis of the sampling bias introduced by the discretization, by considering a splitting (additive decomposition) approach similar to the work presented in Chapter 2. We present the results from work with B. Leimkuhler in [68] along with work completed in collaboration with G. Stoltz in Sections 4.1 and 4.3. Section 4.1 defines the splitting strategy we will consider in the following analysis, as well as presenting a framework for other strategies. Section 4.3 describes the behaviour of the discretization schemes in the case of infinite friction (the overdamped regime of Langevin dynamics), where one method is shown to have a superconvergence property owing to a cancellation at leading order.

Section 4.2 combines work for error estimates in harmonic models, originally presented in [69], with new results computing error estimates using a small nonlinear perturbation. As the harmonic case is often the most relevant for much of molecular dynamics simulation, the work gives insight into many of the results presented in Section 5.3, where we apply the schemes to a solvated biomolecule. Additionally we give some results for the stability of Langevin dynamics for linear systems when using multiple timestepping schemes, that we will directly to the results for the rRESPA scheme in Section 2.1.5.

4.1 Foundations for stochastic splitting

We focus on the weak accuracy properties of a numerical method, which describes the corruption in the law of trajectories in the large time limit. In principle such errors can be eliminated completely through the use of Metropolis acceptance tests [13, 29, 74], though this can destroy the dynamical properties of the system, or become prohibitively costly to implement due to the rejection criteria.

In recent years, there has been widespread interest in multiscale methods for en-

hanced sampling [10, 57, 62, 81] and such methods likely offer the best approach to bridging the timescale gap. We observe that work on enhanced numerical schemes for computing thermodynamically sound trajectories in molecular dynamics remains essential, as it plays a crucial underpinning role in all the enhanced sampling approaches.

We will not make explicit mention of hybrid Monte-Carlo [29, 74] or other schemes based on Metropolis correction [13], although these could be used in conjunction with several of the splitting methods given for Langevin dynamics. The improvement in thermodynamic sampling obtained through the use of more accurate Langevin integrators may, in some cases, provide an alternative to Metropolis-based correction in the practical setting. All methods under discussion require one force evaluation per iteration, and hence have practically the same computational cost.

Methods constructed and analyzed will be based upon splitting the SDE vector field just as for deterministic molecular dynamics, where this type of analysis has also been used for the correction of averages [11, 25]. The approach used here may be compared to other recent works on stochastic numerical methods, and, in particular [1, 12, 15, 82, 87, 108, 111, 130]. Our approach differs in that we seek the explicit quantification of the errors (the perfect sampling bias) introduced in discretization, through the use of the Baker-Campbell-Hausdorff formula. In typical cases which would be relevant for molecular simulation, the error introduced in averages using such methods would be second order in the timestep (i.e. would go to zero quadratically as the step size is reduced).

Our focus on the error arising from discretization is motivated by the extremely long-time simulations that are conducted in many molecular dynamics applications. Of course in practical simulation, many other sources of statistical error exists. Sampling error is generally considered to be the dominant error term in numerical simulation This error comes from two sources: generating a finite number of points (or samples) in our discretized trajectory, and integrating over a finite time interval (as the equality between time and space average is only in the limit of infinite time).

As sampling error is seperate from discretization error, we shall assume that we are able to conduct perfect sampling of the dynamics' native distribution using our numerical method, and hence the only observed errors will be the perfect sampling bias introduced in the numerical discretization. We will discuss other sources of error for particular examples.

4.1.1 Preliminaries

We shall consider a system described by n particles in a d-dimensional setting, with $N = dn$ the total dimensionality of the instantaneous system position and momenta q, p respectively. While the momenta are unrestricted ($p \in \mathbb{R}^N$), we shall consider only the case of compact position space for the following analysis. Periodic boundary conditions shall be assumed, so that $q \in \mathcal{M} = (L\mathbb{T})^N$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the unit torus and $L > 0$ is the box length (this assumption is commensurate with practical MD simulation). We

write the full state space as Ω where $(q,p) = z \in \Omega = \mathcal{M} \times \mathbb{R}^N$.

The system Hamiltonian is

$$H(q,p) = p^T M^{-1} p/2 + U(q),$$

for diagonal mass matrix M, and a C^∞ potential energy function $U : \mathcal{M} \to \mathbb{R}$ (bounded below) describing all particle interactions. The restriction on the smoothness of the potential greatly simplifies the later analysis, but in the molecular dynamics context it can seem far too restrictive. One of the most commonly used potential energy functions is the Lennard-Jones potential

$$U_{\text{lj}}(r) = \frac{\sigma_1}{r^{12}} - \frac{\sigma_2}{r^6},$$

for interatomic separation r and positive constants $\sigma_i > 0$. This potential emulates the strong repulsion felt between two particles, evidently blowing up as two particles come in close proximity. However, the salient features of this potential, such as the steep r^{-12} wall, can be replicated through the use of smooth functions such as Morse or Buckingham potentials. In practice, configurations remain far from singular points, and it is reasonable to expect that identical results are achieved either by restricting the domain or using a smooth cutoff to remove the singularity.

Our goal is to compute integrals with respect to the canonical distribution

$$\begin{aligned} \rho_\beta(q,p) &= Z^{-1} e^{-\beta H(q,p)} = \mu_\beta(q)\kappa_\beta(p), \\ \mu_\beta(q) = Z_\mu^{-1} e^{-\beta U(q)}, &\qquad \kappa_\beta(p) = Z_\kappa^{-1} e^{-\beta p^T M^{-1} p/2}, \end{aligned} \tag{4.1}$$

where $\beta = 1/k_B T$ for Boltzmann's constant k_B and temperature parameter T and

$$Z = \int_\Omega e^{-\beta H(q,p)} \mathrm{d}q\mathrm{d}p, \quad Z_\mu = \int_{\mathcal{M}} e^{-\beta U(q)} \mathrm{d}q, \quad Z_\kappa = \int_{\mathbb{R}^N} e^{-\beta p^T M^{-1} p/2} \mathrm{d}p.$$

The system momenta are Gaussian distributed according to κ_β while the configurational distribution μ_β is significantly more complex. A robust and flexible choice for sampling the canonical distribution ergodically is given by Langevin dynamics,

$$\begin{aligned} \mathrm{d}q &= M^{-1} p \, \mathrm{d}t, \\ \mathrm{d}p &= -\nabla U(q) \, \mathrm{d}t - \gamma p \, \mathrm{d}t + \sqrt{2\gamma/\beta} \, M^{1/2} \, \mathrm{d}W, \end{aligned} \tag{4.2}$$

where $\gamma > 0$ is a (free) scalar parameter modelling friction and $W = W(t)$ denotes a vector of N i.i.d. Wiener random processes (white noise) that approximate the effects of a coupled heat bath bombarding the system under study. The infinitesimal generator for Langevin dynamics, computed from the L^2 adjoint of its Fokker-Planck operator

$\mathcal{L}_{\mathrm{LD}}^*$ (3.11) or directly using Itō's formula in (3.5), is given as

$$\mathcal{L}_{\mathrm{LD}} = M^{-1}p \cdot \nabla_q - \nabla U(q) \cdot \nabla_p - \gamma p \cdot \nabla_p + \frac{\gamma}{\beta}\nabla_p \cdot (M\nabla_p)\,,$$

where for any sufficiently smooth continuous function $\phi(z)$, we have

$$\mathbb{E}\left[\,\phi(z(t)) \,|\, z(0) = z_0\,\right] = \left(e^{t\mathcal{L}_{\mathrm{LD}}}\phi\right)(z_0),$$

for the *evolution operator* $e^{t\mathcal{L}_{\mathrm{LD}}}$. Choosing a Lyapunov function

$$\mathcal{K}_s(q,p) = 1 + \left(p^T p\right)^s,$$

for $s \in \mathbb{N}^*$, gives the existence of solutions to (4.2) for all time (by, for example, [98, Theorem 5.9]). We will make make use of hypocoercivity estimates, which give sharper bounds on the rate of convergence toward the invariant distribution (the spectrum of $\mathcal{L}_{\mathrm{LD}}$). Standard hypocoercivity results [51, 60, 125] demonstrate the existence of constants $K, \lambda > 0$ (whose values depend on the friction γ) such that for any $\gamma, t > 0$ we have

$$\left\|e^{t\mathcal{L}_{\mathrm{LD}}}\right\|_{\mathcal{B}(\mathcal{H}^1)} \leq Ke^{-t\lambda},$$

where $\|\cdot\|_{\mathcal{B}(\mathcal{H}^1)}$ denotes the operator norm on the subspace

$$\mathcal{H}^1 = \left\{ \varphi \in H^1(\rho_\beta) \,\middle|\, \int_\Omega \varphi\,\rho_\beta\,\mathrm{d}z = 0 \right\} \subset H^1(\rho_\beta),$$

endowed with the norm $\|\varphi\|^2_{H^1(\rho_\beta)} = \|\varphi\|^2 + \|\nabla_q\varphi\|^2 + \|\nabla_p\varphi\|^2$. Notably, the inverse operator is bounded, where

$$\left\|\mathcal{L}_{\mathrm{LD}}^{-1}\right\|_{\mathcal{B}(\mathcal{H}^1)} \leq \frac{K}{\lambda}.$$

This bound will be important in the upcoming analysis, as

4.1.2 Decomposition into pieces A, B and C

For general potential energy functions $U(q)$, in order to compute consistent trajectories for the dynamics (4.2) we require the use of an approximation employing a discretization scheme in time (spatial discretization is impractical in the case of a large number of particles). We shall focus on numerical methods built by splitting the Langevin dynamics vector field into additive pieces, and solving each piece exactly in sequence, as if we were building a symplectic integrator for deterministic dynamics.

Splitting methods for Langevin dynamics have been considered in the past [7, 12, 16, 73, 82, 107, 109, 111, 117] but a wide variety of schemes can be constructed by splitting and the rational basis for selecting one scheme over another is seldom presented.

Our approach differs in that we offer a new splitting strategy, and a strict focus on the error of the invariant (numerical) measure being sampled by the discretization

scheme. Many schemes are evaluated on the basis of their temperature control (particularly on the convergence of the average of kinetic energy), while we aim to give a more thorough treatment of the numerical error introduced through discretization.

Just as in Chapter 2, we shall consider the discretized numerical trajectory as sampling a perturbed measure that varies with step size. However, unlike in Chapter 2, we do not construct a perturbed vector field that we associate to our numerical scheme (see [130]). Instead we shall exploit the Baker-Campbell-Hausdorff formula demonstrated for backward error analysis in the deterministic setting (see Section 2.1.3) to approximate the generator of the associated numerical method formed by a composition of splitting pieces. This generator can be viewed as a perturbation of the exact generator for Langevin dynamics introduced in Chapter 3.

However, the increased complexity of Langevin dynamics (4.2) over Newtonian dynamics (2.4) gives us additional options for splitting the vector field. One option is to use a splitting strategy that divides the dynamics into position and momentum components as in [82]. The splitting pieces become

$$\begin{bmatrix} \mathrm{d}q \\ \mathrm{d}p \end{bmatrix} = \underbrace{\begin{bmatrix} M^{-1}p\,\mathrm{d}t \\ 0 \end{bmatrix}}_{A} + \underbrace{\begin{bmatrix} 0 \\ -\nabla U(q)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma/\beta}\,M^{1/2}\,\mathrm{d}W \end{bmatrix}}_{S}, \tag{4.3}$$

denoted A and S. As in Section 2.1.2, we solve the vector fields in sequence to create the method, for example, the stochastic position Verlet method is given by $[\![\mathrm{ASA}]\!]$, where we extend our notation from the first chapter in the obvious way.

The flow map for the S piece is given by

$$\Phi_{S,t}\left(\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}\right) = \begin{bmatrix} q_0 \\ e^{-\gamma t}p_0 + \frac{1-e^{-\gamma t}}{\gamma}\nabla U(q_0) + \beta^{-1/2}\sqrt{1-e^{-2\gamma t}}M^{1/2}R(t), \end{bmatrix},$$

for $R(t) \sim \mathcal{N}(0,t)$ a vector of N random noise processes. In the case of a large friction parameter γ, the influence of the force term on the dynamics will become extremely small, and in the limit of infinite γ methods using this choice of splitting will become inconsistent (the force term is wiped out completely, leaving no information about $U(q)$ in the dynamics). Hence we shall consider coupling the force to the random noise term as evidently not suitable in the case of large friction, although it is desirable to keep the friction and noise terms of the Ornstein-Uhlenbeck process together, as we are able to solve exactly.

An alternative is offered by Owhadi and Bou-Rabee in [12], who suggest a geometric Langevin algorithm (GLA) splitting, decomposing the dynamics into its microcanonical part and Ornstein-Uhlenbeck part

$$\begin{bmatrix} \mathrm{d}q \\ \mathrm{d}p \end{bmatrix} = \underbrace{\begin{bmatrix} M^{-1}p\,\mathrm{d}t \\ -\nabla U(q)\,\mathrm{d}t \end{bmatrix}}_{\mathrm{mc}} + \underbrace{\begin{bmatrix} 0 \\ -\gamma p\,\mathrm{d}t + \sqrt{2\gamma/\beta}\,M^{1/2}\,\mathrm{d}W \end{bmatrix}}_{\mathrm{OU}}. \tag{4.4}$$

Using a general order s symplectic method to evolve the microcanonical dynamics, followed by exactly solving the OU process, the results of [12] demonstrate that for a smooth observable φ we obtain

$$\int_\Omega \varphi \widehat{\rho}_{\delta t} \mathrm{d}z = \int_\Omega \varphi \rho_\beta \mathrm{d}z + \mathcal{O}(\delta t^s).$$

This gives the opportunity to create arbitrary high weak order Langevin dynamics methods by combining the Yoshida-type methods introduced in Section 2.1.4 with the exact OU solve. If we are able to solve the Newtonian part of the system exactly (or in practice to arbitrarily high order through, for example, numerical quadrature) then though the evolution of the distribution through time will be accurate to only first-order, we achieve perfect sampling of averages in the infinite-time limit. Of course, in most cases this is computationally prohibitive.

Utilizing a splitting scheme that isolates all terms involving the friction γ away from the force term means that we would expect this strategy to be more fruitful at evolving the system in the high (or infinite) friction regime. However, examples will be given to show that for some choices of symplectic method we can still fail to sample ρ_β consistently in the case of infinite friction.

We shall provide a framework to analyze and describe a wide class of Langevin dynamics splitting methods, by generalizing the GLA splitting strategy into what we term the ABC splitting:

$$\begin{bmatrix} \mathrm{d}q \\ \mathrm{d}p \end{bmatrix} = \underbrace{\begin{bmatrix} M^{-1} p \, \mathrm{d}t \\ 0 \end{bmatrix}}_{A} + \underbrace{\begin{bmatrix} 0 \\ -\nabla U(q) \, \mathrm{d}t \end{bmatrix}}_{B} + \underbrace{\begin{bmatrix} 0 \\ -\gamma p \, \mathrm{d}t + \sqrt{2\gamma/\beta} \, M^{1/2} \, \mathrm{d}W \end{bmatrix}}_{C}, \tag{4.5}$$

where the A and B pieces of the vector field are the same as was introduced in Section 2.1.2, and the OU process is labelled C for clarity of notation (in contrast to labelling it O in [68, 69]). We can solve each part of the vector field (when taken independently) exactly, with associated flow maps

$$\begin{aligned} \Phi_{A,\delta t}\left(\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}\right) &= \begin{bmatrix} q_0 + \delta t \, M^{-1} p_0 \\ p_0 \end{bmatrix}, \qquad \Phi_{B,\delta t}\left(\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}\right) = \begin{bmatrix} q_0 \\ p_0 - \delta t \, \nabla U(q_0) \end{bmatrix}, \\ \Phi_{C,\delta t}\left(\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}\right) &= \begin{bmatrix} q_0 \\ e^{-\gamma t} p_0 + \beta^{-1/2} \sqrt{1 - e^{-2\gamma t}} M^{1/2} R(t) \end{bmatrix}. \end{aligned} \tag{4.6}$$

As the splitting pieces A and B correspond to those used for deterministic Hamiltonian dynamics, the generators for the evolution of the expectation of observables is given by their Lie derivatives as derived in Section 2.1.2. The generator for the C piece (as seen in Chapter 3) comes from Itō's lemma, where we make explicit the dependence on the

friction by introducing a friction independent operator, giving

$$\mathcal{L}_A\psi = M^{-1}p\cdot\nabla_q\psi, \qquad \mathcal{L}_B\psi = -\nabla U(q)\cdot\nabla_p\psi, \qquad \mathcal{L}_C\psi = \gamma\mathcal{L}_{\bar{C}}\psi,$$
$$\mathcal{L}_{\bar{C}}\psi = -p\cdot\nabla_p\psi + \beta^{-1}\nabla_p\cdot(M\nabla_p)\psi,$$

with the generator of Langevin dynamics given as

$$\mathcal{L}_{\mathrm{LD}} = \mathcal{L}_A + \mathcal{L}_B + \mathcal{L}_C.$$

Its evolution of measure is given by taking its L^2 adjoint, where

$$\mathcal{L}_A^*\psi = -\mathcal{L}_A\psi, \qquad \mathcal{L}_B^*\psi = -\mathcal{L}_B\psi, \qquad \mathcal{L}_C^*\psi = \gamma\mathcal{L}_{\bar{C}}^*\psi,$$
$$\mathcal{L}_{\bar{C}}^*\psi = \nabla\cdot(p\,\psi) + \beta^{-1}\nabla_p\cdot(M\nabla_p)\psi.$$

Our line of attack for the numerical analysis of Langevin dynamics discretization methods is informed by our work in Section 2.1.2 for deterministic dynamics and draws on the work of Talay and Tubaro [116] for expansions of invariant measures in a series of δt. We consider the numerical method to sample from a perturbed problem exactly, where there exists some generator corresponding to its perturbed dynamics. We use a macroscopic approach by working directly with the perturbed generator for the numerical process, as in [27, 79]. The generator for the numerical method $\widehat{\mathcal{L}}_{\delta t}$ is found exactly as in the symplectic case, by computing the product of exponentials in sequential order. Moreover, the evolution of distribution is governed by its adjoint, with

$$\frac{\partial}{\partial t}\rho(q,p,t) = \widehat{\mathcal{L}}_{\delta t}^*\rho(q,p,t),$$

with a unique nontrivial stationary (invariant) distribution

$$\widehat{\mathcal{L}}_{\delta t}^*\,\widehat{\rho}_{\delta t} \equiv 0. \tag{4.7}$$

For an order s method, we write the adjoint of the generator for the numerical method by perturbing the exact operator

$$\widehat{\mathcal{L}}_{\delta t}^* = \mathcal{L}_{\mathrm{LD}}^* + \delta t^s\mathcal{L}_s^* + \delta t^{s+1}\mathcal{L}_{s+1}^* + \dots. \tag{4.8}$$

In light of (4.7) it makes sense to write

$$\widehat{\rho}_{\delta t} = \rho_\beta\,\left(1 + \delta t^s f_s(q,p) + \delta t^{s+1}f_{s+1}(q,p) + \dots\right), \tag{4.9}$$

where the perturbation functions $f(q,p)$ satisfy the PDE (4.7). This implies that the error introduced by a numerical method is itself a stationary average

$$\int_\Omega \phi(q,p)\,\widehat{\rho}_{\delta t}\,\mathrm{d}q\mathrm{d}p - \int_\Omega \phi(q,p)\,\rho_\beta\,\mathrm{d}q\mathrm{d}p = \delta t^s\int_\Omega \phi(q,p)f_s(q,p)\,\rho_\beta\,\mathrm{d}q\mathrm{d}p + \mathcal{O}(\delta t^{s+1}).$$

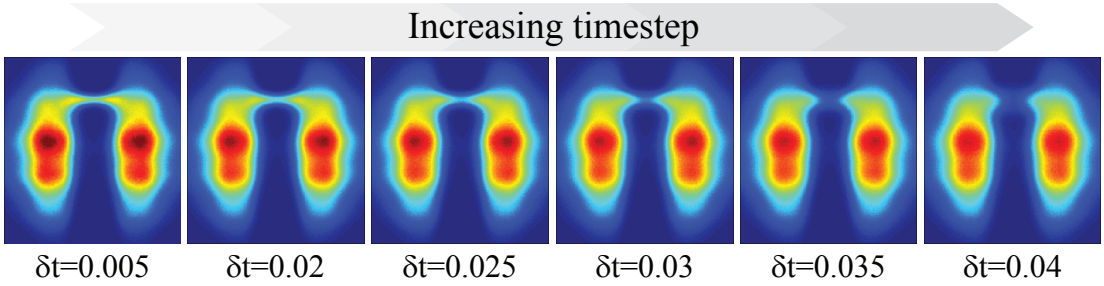


Figure 4.1: Invariant configurational distribution functions for the ⟦ABC⟧ scheme are computed through simulation, using a planar potential energy function with two distinct basins connected by a narrow transition corridor. We plot distributions computed at various step sizes in the plane, where high/low intensity is coloured red/blue respectively. At simulations with a larger timestep we can see a qualitative change in the landscape, with the bridge between the two regions disappearing. Equilibrium averages as well as transport coefficients (such as the diffusion constant) will be perturbed by the corruption of ρ_β.

By solving the PDE (4.7) for the perturbation functions, computed numerical averages can be corrected to higher orders of δt.

The correction of such averages is important for molecular dynamics simulations, where the discretization error at large timesteps can warp the potential energy landscape to such a degree that it qualitatively influences the system behaviour (see Figure 4.1). We write the evolution operator for such numerical methods as $\mathcal{P}_{\delta t}$, where

$$\mathcal{P}_{\delta t}\varphi(z) = \mathbb{E}\left(\varphi\left(z_{k+1}\right)\middle| z_k = z\right).$$

where the subscript on z_k indicates an iteration index. The form of this operator will vary between methods as the generator $\widehat{\mathcal{L}}_{\delta t}$ in (4.7) is method-specific. In order to make clear which method we refer to we shall denote the method as superscript. For example in the case of the method ⟦ABC⟧, we write

$$\mathcal{P}_{\delta t}^{[\![\mathrm{ABC}]\!]} = e^{\delta t \mathcal{L}_A}\, e^{\delta t \mathcal{L}_B}\, e^{\delta t \mathcal{L}_C},$$

corresponding to the evolution operator for the update scheme

$$\begin{aligned} q_{n+1} &= q_n + \delta t\, M^{-1} p_n, \\ p_{n+1/2} &= p_n - \delta t\, \nabla U(q_{n+1}), \\ p_{n+1} &= e^{-\gamma\delta t} p_{n+1/2} + \beta^{-1/2}\sqrt{1 - e^{-2\gamma\delta t}}\, M^{1/2} R_n, \end{aligned} \tag{4.10}$$

where $R_n \sim \mathcal{N}(0,1)$ is a vector of N independent normal random numbers. This method performs the Symplectic Euler method (sE-AB) followed by an OU exact solve, making this method one of the class of GLA methods. Note also, that in the large friction limit the force is again wiped out by the OU process, making this method

inconsistent in this regime. This method is referred to GLA-1 in [12], and is shown to give a first-order error in observable averages.

There are five other first-order methods obtained by composing the three update operators in the ABC splitting so that the resulting method is consistent and only performs each update once (we refer to such methods as elementary first-order schemes). Of these six such schemes, we can split them into two classes

$$\boxed{[\![\mathrm{ABC}]\!], \quad [\![\mathrm{BCA}]\!], \quad [\![\mathrm{CAB}]\!],} \tag{4.11}$$

and

$$\boxed{[\![\mathrm{ACB}]\!], \quad [\![\mathrm{CBA}]\!], \quad [\![\mathrm{BAC}]\!].} \tag{4.12}$$

Iteration of a method in (4.11) or (4.12) can be easily related to another method on the same line, as these schemes share a common sequence of operations. Other first-order methods exist, for example the scheme $[\![\mathrm{ABCA}]\!]$ is first-order and belongs to the class of methods in (4.11).

The precise relation between methods that share a common sequence will be made explicit in Lemma 4.30. It is worth noting that none of these elementary first-order schemes is consistent in the limit $\gamma \to \infty$.

By employing a Strang (symmetric) splitting approach, we can create six second-order schemes which can be divided similarly into three classes of method

$$\underset{\text{(i)}}{\boxed{\begin{array}{c}[\![\mathrm{ABCBA}]\!] \\ \text{and} \\ [\![\mathrm{CBABC}]\!]\end{array}}}, \qquad \underset{\text{(ii)}}{\boxed{\begin{array}{c}[\![\mathrm{ACBCA}]\!] \\ \text{and} \\ [\![\mathrm{BCACB}]\!]\end{array}}}, \qquad \underset{\text{(iii)}}{\boxed{\begin{array}{c}[\![\mathrm{BACAB}]\!] \\ \text{and} \\ [\![\mathrm{CABAC}]\!]\end{array}}}. \tag{4.13}$$

Methods of class (ii) interpose the B step between two C steps, and hence at infinitely large friction it is clear the force will be completely wiped out by the redrawing of p. The two schemes in (ii) must therefore be inconsistent with the dynamics at infinite friction. It should be noted that none of the schemes considered in (4.13) can be classed as GLA-type schemes, which instead involve only a single C step at the end of the deterministic iteration (and as such, are not symmetric).

Other symmetric second-order schemes can be created with the same alphabet, for example our framework allows us to consider arbitrary symmetric schemes such as $[\![\mathrm{ABCACBA}]\!]$, though the complexity involved makes such analysis undesirable for no potential gain. Additionally, these methods could come at a greater computational cost as they may require more than one force evaluation per timestep. We shall consider only the schemes of class (i) and (iii) above, and refer to these as the elementary second-order methods.

For all such schemes, and for functions in $L^\infty_{\mathcal{K}_s}$, where

$$L^\infty_{\mathcal{K}_s} = \left\{ \varphi \text{ measurable} \;\middle|\; \frac{\varphi(q,p)}{\mathcal{K}_s(q,p)} \in L^\infty(\rho_\beta) \right\}$$

endowed with the norm

$$\|\varphi\|_{L^\infty_{\mathcal{K}_s}} = \left\| \frac{\varphi}{\mathcal{K}_s} \right\|_{L^\infty},$$

we have the following important ergodicity result:

Theorem 4.14 (Ergodicity of numerical schemes)**.** *For fixed s^*, there exists a δt^* such that for any positive, finite γ and any $0 < \delta t < \delta t^*$, we have that the evolution operator $\mathcal{P}_{\delta t}$, for any composition scheme given in (4.11-4.13), has a unique invariant probability measure $\widehat{\rho}_{\delta t}$ with finite moments*

$$\int_\Omega \mathcal{K}_s \, \widehat{\rho}_{\delta t} \, dz < \infty, \tag{4.15}$$

for $1 \leq s \leq s^*$. There exists constants $\lambda, K > 0$ such that

$$\left| \mathcal{P}^n_{\delta t} f(q,p) - \int_\Omega f \, \widehat{\rho}_{\delta t} \, dz \right| \leq K \mathcal{K}_s(q,p) e^{-\lambda n} \|f\|_{L^\infty_{\mathcal{K}_s}}, \tag{4.16}$$

for all $n \in N$, all $f \in L^\infty_{\mathcal{K}_s}$ and almost all $z \in \Omega$.

The proof of the theorem follows immediately from [49] under two assumptions, which we show to hold from the definition of the schemes themselves. We will demonstrate proofs in the case of the ⟦BAC⟧ scheme, which has concise update

$$q_{k+1} = q_k + \delta t M^{-1} p_k - \delta t^2 M^{-1} \nabla U(q_k), \tag{4.17}$$

$$p_{k+1} = e^{-\gamma \delta t} \left(p_k - \delta t \nabla U(q_k) \right) + \sqrt{\frac{(1 - e^{-2\gamma \delta t}) M}{\beta}} R_k. \tag{4.18}$$

with R_k a vector of N i.i.d. normal random numbers distributed as $\mathcal{N}(0,1)$.

Proofs given will use the update formula in the case of the ⟦BAC⟧ scheme (other methods follow very similar lines and are omitted).

Lemma 4.19 (Uniform Lyapunov condition)**.** *Consider the evolution operator $\mathcal{P}_{\delta t}$ for any composition scheme given in (4.11-4.13), with a Lyapunov function*

$$\mathcal{K}_s = 1 + (p \cdot p)^s .$$

For any $s^* \in \mathbb{N}^*$, there exists a $\delta t^* > 0$ and constants $C_1, C_2 > 0$ such that for $0 < \delta t < \delta t^*$ and $1 \leq s \leq s^*$

$$\mathcal{P}_{\delta t} \mathcal{K}_s \leq a_{\delta t} \mathcal{K}_s + b_{\delta t}, \qquad 0 \leq a_{\delta t} \leq e^{-C_1 \delta t}, \qquad 0 \leq b_{\delta t} \leq C_2 \delta t. \tag{4.20}$$

Notably, choosing any $T > 0$, we have

$$\mathcal{P}_{\delta t}^{\lceil T/\delta t \rceil}\mathcal{K}_s \leq e^{-C_1 T}\mathcal{K}_s + C_2/C_1. \tag{4.21}$$

Proof. We will initially prove the lemma in the case $s = 1$, before moving to the slightly more involved cases where $s > 1$. Writing $\alpha = \exp(-\gamma\delta t)$ we premultiply (4.18) by its transpose and take expectations to find

$$\begin{aligned}
\mathbb{E}\left(p_{k+1}^T p_{k+1}\middle|\mathcal{F}_k\right) &= \alpha^2 \left(p_k - \delta t \nabla U(q_k)\right)^T \left(p_k - \delta t \nabla U(q_k)\right) + \beta^{-1}\left(1-\alpha^2\right)\mathrm{Tr}\left(M\right), \\
&\leq \alpha^2 p_k^T p_k + 2\delta t \|\nabla U\|\,|p_k| + \delta t^2 \|\nabla U\|^2 + \beta^{-1}\left(1-\alpha^2\right)\mathrm{Tr}\left(M\right), \\
&\leq \left(\alpha^2 + \epsilon^2\delta t\right) p_k^T p_k + \left(\epsilon^{-2}\delta t + \delta t^2\right)\|\nabla U\|^2 + \beta^{-1}\left(1-\alpha^2\right)\mathrm{Tr}\left(M\right),
\end{aligned}$$

where $\epsilon > 0$, by virtue of the relation

$$2\delta t\|\nabla U\|\,|p_k| \leq 2\delta t\|\nabla U\|\,|p_k| + \delta t\left(\epsilon|p_k| - \epsilon^{-1}\|\nabla U\|\right)^2 = \epsilon^2\delta t|p_k|^2 + \epsilon^{-2}\delta t\|\nabla U\|^2.$$

Choosing $\epsilon^2 = \gamma$ gives

$$\alpha^2 + \epsilon^2\delta t = e^{-2\gamma\delta t} + \gamma\delta t \leq e^{-\gamma\delta t/2},$$

for sufficiently small δt. Hence

$$\mathbb{E}\left(p_{k+1}^T p_{k+1}\middle|\mathcal{F}_k\right) \leq a_{\delta t} p_k^T p_k + \widetilde{b}_{\delta t},$$

for

$$a_{\delta t} = e^{-\gamma\delta t/2}, \qquad \text{and} \qquad \widetilde{b}_{\delta t} = \frac{2\delta t}{\gamma}\|\nabla U\|^2 + \frac{4\gamma\delta t}{\beta}\mathrm{Tr}\left(M\right),$$

where δt is sufficiently small. Hence for the Lyapunov function $\mathcal{K}_1 = 1 + p^T p$, we have

$$\mathbb{E}\left(\mathcal{K}_1(q_{k+1}, p_{k+1})\middle|\mathcal{F}_k\right) \leq 1 + a_{\delta t} p_k^T p_k + \widetilde{b}_{\delta t} = a_{\delta t}\mathcal{K}_1(q_k, p_k) + b_{\delta t},$$

with

$$b_{\delta t} = \widetilde{b}_{\delta t} + 1 - a_{\delta t},$$

giving (4.20). The case $s > 1$ follows similarly. We will use the fact that for some $\epsilon > 0$, we can bound powers of $x \in \mathbb{R}$ by considering

$$|x| \leq \frac{1}{\epsilon} \implies |x|^{2s-m} \leq \frac{1}{\epsilon^{2s-m}}, \qquad |x| \geq \frac{1}{\epsilon} \implies |x|^{2s-m} \geq \epsilon^m |x|^{2s},$$

for $0 \leq m \leq 2s$, and hence

$$|x|^{2s-m} \leq \epsilon^m |x|^{2s} + \frac{1}{\epsilon^{2s-m}}. \tag{4.22}$$

Denoting the l^{th} component with a subscript $[l]$ and working componentwise, we take

powers of (4.18) to obtain

$$
\begin{aligned}
p_{[l],k+1}^{2s} &= \left(\alpha \left(p_{[l],k} - \delta t \frac{\partial}{\partial q_{[l]}} U(q_k) \right) + \sqrt{\frac{(1-\alpha^2) m_{[l]}}{\beta}} R_{[l],k} \right)^{2s}, \\
&= \sum_{i=0}^{2s} \begin{pmatrix} 2s \\ i \end{pmatrix} \alpha^i \left(p_{[l],k} - \delta t \frac{\partial}{\partial q_{[l]}} U(q_k) \right)^i \left(\sqrt{\frac{(1-\alpha^2) m_{[l]}}{\beta}} R_{[i],k} \right)^{2s-i}, \\
&\leq \sum_{i=0}^{2s} \sum_{j=0}^{i} \begin{pmatrix} 2s \\ i \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix} \alpha^i \delta t^{i-j} |p_{[l],k}|^j \left\| \frac{\partial}{\partial q_{[l]}} U \right\|^{i-j} \left(\sqrt{\frac{(1-\alpha^2) m_{[l]}}{\beta}} R_{[l],k} \right)^{2s-i}, \\
&\leq \alpha^{2s} |p_{[l],k}|^{2s} + K \sum_{i=0}^{2s-1} \begin{pmatrix} 2s \\ i \end{pmatrix} \alpha^i |p_{[l],k}|^i \left(\sqrt{\frac{(1-\alpha^2) m_{[l]}}{\beta}} R_{[l],k} \right)^{2s-i} \\
&\quad + \alpha \delta t K \sum_{i=1}^{2s} \sum_{j=0}^{i-1} \begin{pmatrix} 2s \\ i \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix} |p_{[l],k}|^j \left(\sqrt{\frac{(1-\alpha^2) m_{[l]}}{\beta}} R_{[l],k} \right)^{2s-i},
\end{aligned}
$$

using $K = \max(1, \left\| \partial / \partial q_{[l]} U \right\|^{2s})$. Taking expectations, only even powers of $R_{[l],k}$ will survive. Hence, bounding $1 - \alpha^2 \leq 4\gamma\delta t$ and applying (4.22), we obtain (4.20) by choosing ϵ sufficiently small. The formula (4.21) follows straightforwardly from iterating (4.20). □

Lemma 4.23 (Uniform minorization condition)**.** *Consider*

$$C = \left\{ (q,p) \in \Omega \,\middle|\, |p| < p_{max} \right\},$$

with fixed $p_{max} > \sqrt{2C_2/C_1}$, where constants C_1, C_2 are chosen as in Lemma 4.19. There exists constants $\delta t^* > 0$ and $a \in (0,1)$ with probability measure ν such that for any bounded measurable non-negative function f, we have

$$\inf_{z \in C} (\mathcal{P}_{\delta t} f(z)) \geq a \int_\Omega f(z) \nu(z)\, dz \tag{4.24}$$

for all $0 < \delta t < \delta t^*$.

Proof. This proof is made significantly easier by virtue of the compactness of the position space $\mathcal{M}$. It is sufficient to prove the lemma in the case of f being the indicator function for Borel sets $A_q \subset \mathcal{M}$ and $A_p \subset \mathbb{R}^N$ with $A = A_q \times A_p \subset \Omega$. The goal is to show that

$$\mathbb{P}\left(z_1 \in A \,\middle|\, z_0 \in C \right) \geq a\nu(A), \tag{4.25}$$

where $z_1 = \mathcal{P}_{\delta t}(z_0)$. Writing (4.17) as

$$p_{k+1} = \widehat{p}_{k+1} + G_{k+1}, \qquad G_{k+1} = \sqrt{\frac{(1 - e^{-2\gamma\delta t}) M}{\beta}} R_k,$$

we have that

$$\mathbb{P}\left(z_1 \in A \,\middle|\, z_0 \in C\right) = \mathbb{P}\left([0, G_1] \in (A_q - q_1) \times (A_p - \widehat{p}_1) \,\middle|\, [q_0, p_0] \in C\right),$$

while by the boundedness of the position space, there exists a constant $E > 0$ such that

$$|q_1| < E, \qquad |\widehat{p}_1| < E.$$

Each component of the variable $G_1 \in \mathbb{R}^N$ is a Gaussian random variable, with

$$\mu = \text{mean}(G_1) = 0, \qquad \sigma^2 = \text{var}(G_1) = \frac{(1 - e^{-2\gamma\delta t})M}{\beta},$$

and hence for any $X \subset \mathcal{M}$ and $Y \subset \mathbb{R}^N$, we have

$$\mathcal{P}([0, G_1] \in X \times Y) = Z^{-1} \int_{X \times Y} \exp(-y^T \sigma^{-2} y/2) \, \mathrm{d}x \, \mathrm{d}y,$$
$$Z = \int_{\Omega} \exp(-y^T \sigma^{-2} y/2) \, \mathrm{d}x \, \mathrm{d}y,$$

for $x \in \mathcal{M}$ and $y \in \mathbb{R}^N$. This gives

$$\mathbb{P}\Big([0, G_1] \in (A_q - q_1) \times (A_p - \widehat{p}_1) \Big| \, [q_0, p_0] \in C\Big)$$
$$= Z^{-1} \int_{(A_q - q_1) \times (A_p - \widehat{p}_1)} \exp(-y^T \sigma^{-2} y/2) \, \mathrm{d}x \, \mathrm{d}y,$$

and thus we define the probability measure

$$\nu(A_q \times A_p) = Z_\nu^{-1} \inf_{\substack{|Q| < E \\ |P| < E}} \int_{(A_q - Q) \times (A_p - P)} \exp(-y^T \sigma^{-2} y) \mathrm{d}x \mathrm{d}y,$$

with appropriate normalization constant Z_ν^{-1}. This gives (4.25). □

The existence and uniqueness of the method's invariant measure $\widehat{\rho}_{\delta t}$ is given by Theorem 1.2 in [49] (where assumption 1 and 2 in [49] are given by Lemmas 4.19 and 4.23 respectively), for the timestep δt chosen sufficiently small. The discrete trajectories generated by $\mathcal{P}_{\delta t}$ will be ergodic in the sense that their infinite time average will sample $\widehat{\rho}_{\delta t}$.

Proof of Theorem 4.14. The equation (4.16) is given by [49]. It remains to show the finiteness of the moments of the distribution in equation (4.15). Averaging (4.20) with respect to the numerical invariant measure gives

$$\int_{\Omega} \mathcal{P}_{\delta t} \, \mathcal{K}_s \widehat{\rho}_{\delta t} \, \mathrm{d}z \leq a_{\delta t} \int_{\Omega} \mathcal{K}_s \, \widehat{\rho}_{\delta t} \, \mathrm{d}z + b_{\delta t},$$

for constants $a_{\delta t}$ and $b_{\delta t}$ as in Lemma 4.19. Using the invariance of the distribution

under $\mathcal{P}_{\delta t}$, we have

$$\int_\Omega \mathcal{P}_{\delta t}\, \mathcal{K}_s \widehat{\rho}_{\delta t}\, \mathrm{d}z = \int_\Omega \mathcal{K}_s \widehat{\rho}_{\delta t}\, \mathrm{d}z,$$

and hence

$$\int_\Omega \mathcal{K}_s \widehat{\rho}_{\delta t}\, \mathrm{d}z \leq \frac{b_{\delta t}}{1 - a_{\delta t}} < \infty,$$

as required. □

The theorem states that the approximation scheme inherits the ergodicity property of Langevin dynamics and that in the infinite-time limit, distributions of initial points propagated using the discretization scheme will converge exponentially fast towards some invariant measure $\widehat{\rho}_{\delta t}$, dependent on the timestep δt. This is very important, as it allows us to bound on the operator itself for functions whose average vanishes with respect to the numerical invariant measure, which in turn allows us to bound the terms appearing in the computed integrals.

As we shall require some control over the derivatives of the observables, we introduce the Sobolev space $\mathcal{W}_s^{i,\infty}$ for $i \in \mathbb{N}^*$, defined recursively as

$$\mathcal{W}_s^{0,\infty} = \mathcal{L}^\infty_{\mathcal{K}_s}, \qquad \mathcal{W}_s^{i,\infty} = \left\{ \varphi \in \mathcal{L}^\infty_{\mathcal{K}_s} \,\middle|\, \nabla\varphi \in \left(\mathcal{W}_s^{i-1,\infty}\right)^{2N} \right\},$$

and define the space of sufficiently smooth functions $\mathcal{S}$ to be such that for any $\varphi \in \mathcal{S}$ and for any $m \in \mathbb{N}$, there exists an $s \in \mathbb{N}$ such that $\varphi \in \mathcal{W}_s^{m,\infty}$. If the goal of the analysis is to simply compute the order of the method, and not find an explicit form for the perturbing error terms, then we do not need such regularity conditions on the derivatives (such as [12]). In contrast, for a method of (weak) order k, the goal of our error analysis will be to find a smooth *correction function* f_k such that

$$\int_\Omega \varphi(z)\, \widehat{\rho}_{\delta t}\, \mathrm{d}z = \int_\Omega \varphi(z)\, \rho_\beta\, \mathrm{d}z + \delta t^k \int_\Omega f_k(z)\, \varphi(z)\, \rho_\beta\, \mathrm{d}z + \delta t^{k+1} r_{k,\delta t}, \tag{4.26}$$

for any sufficiently smooth observable $\varphi(q,p)$ and where the remainder term $r_{k,\delta t}$ is uniformly bounded for sufficiently small step size.

In order to show that the remainder term is uniformly bounded, we shall make use of the fact that the space

$$\widetilde{\mathcal{S}} = \left\{ \varphi \in \mathcal{S} \,\middle|\, \int \varphi \rho_\beta = 0 \right\}, \tag{4.27}$$

is stable under $\mathcal{L}_{\mathrm{LD}}^{-1}$. Such results can easily be obtained for elliptic operators [27], and can be generalized to Langevin dynamics through the results of Talay [115], under appropriate assumptions on the Hamiltonian. Hence, for any sufficiently small δt, we have that for any $\varphi \in \widetilde{\mathcal{S}}$,

$$\mathcal{L}_{\mathrm{LD}}^{-1}\varphi = R \in \widetilde{S},$$

with

$$\left| \int_\Omega R \widehat{\rho}_{\delta t} \mathrm{d}z \right| \leq C < \infty, \tag{4.28}$$

for some $C > 0$.

For a sufficiently smooth function φ, we shall denote the resulting average computed using a composition scheme whose evolution operator has invariant measure $\widehat{\rho}_{\delta t}$, as

$$\langle \varphi \rangle_{\delta t} := \int_\Omega \varphi \, \widehat{\rho}_{\delta t} \, \mathrm{d}z,$$

where $\langle \varphi \rangle_{\delta t}$ is the observed average at step size δt. This is in contrast to the exact average

$$\langle \varphi \rangle := \int_\Omega \varphi \, \rho_\beta \, \mathrm{d}z.$$

The next section will concern finding explicit error estimates $(\langle \varphi \rangle_{\delta t} - \langle \varphi \rangle)$ for particular discretization schemes.

4.1.3 Computing error estimates for general systems

Our goal is to quantify the errors introduced in the averages computed from schemes composed of sequential solves of the vector fields introduced in the ABC splitting (4.5). The evolution operator for a method is computed through a product of semigroups, where each term corresponds to the sequence of updates codifying the method. Just as in the backward error analysis introduced in Section 2.1.2, we use a Lie-Trotter splitting where each individual evolution is exactly integrable.

In the ABC splitting, each update term in a scheme is given by one the flow maps in (4.6), where the overall evolution operator for a method with M stages is

$$\mathcal{P}_{\delta t} = e^{\delta t \widehat{\mathcal{L}}_{\delta t}} = e^{\delta t L_1} \, e^{\delta t L_2} \ldots e^{\delta t L_M},$$

where each L_i is a scalar multiple of one of $\mathcal{L}_A, \mathcal{L}_B$ or $\mathcal{L}_C$. Each elementary evolution semigroup is well-defined, and we can even analytically write down their action:

$$\begin{aligned}
\left(e^{t L_A} \phi\right)(q,p) &= \phi(q + t M^{-1} p, p), \\
\left(e^{t L_B} \phi\right)(q,p) &= \phi(q, p - t \nabla U(q)), \\
\left(e^{t L_C} \phi\right)(q,p) &= \int_{\mathbb{R}^N} \phi\left(q, e^{-\gamma t} p + \sqrt{\frac{1 - e^{-2\gamma t}}{\beta}} M^{1/2} x\right) \frac{e^{-|x|^2/2}}{(2\pi)^{N/2}} \mathrm{d}x.
\end{aligned}$$

The BCH formula (2.30) can be used to compute $\widehat{\mathcal{L}}_{\delta t}$ directly, though at higher orders it becomes cumbersome to work with due to the large number of commutators required to be computed. The resulting expansion for the generator is purely formal, and gives

us no guidance on its convergence. We opt instead to make use of Taylor's theorem

$$\mathcal{P}_{\delta t} = \mathcal{P}_0 + \delta t \left.\frac{\mathrm{d}\mathcal{P}_t}{\mathrm{d}t}\right|_{t=0} + \ldots + \frac{\delta t^n}{n!} \left.\frac{\mathrm{d}^n\mathcal{P}_t}{\mathrm{d}t^n}\right|_{t=0} + \frac{\delta t^{n+1}}{n!}\int_0^1 (1-\theta)^n \left.\frac{\mathrm{d}^{n+1}\mathcal{P}_t}{\mathrm{d}t^{n+1}}\right|_{t=\theta\delta t} \mathrm{d}\theta,$$

with

$$\frac{\mathrm{d}\mathcal{P}_t}{\mathrm{d}t} = L_1\, e^{\delta t L_1}\, e^{\delta t L_2} \ldots e^{\delta t L_M} + e^{\delta t L_1}\, L_2 e^{\delta t L_2} \ldots e^{\delta t L_M} + \ldots + e^{\delta t L_1}\, e^{\delta t L_2} \ldots L_M\, e^{\delta t L_M}.$$

We can see that an intrinsic order is preserved in this sequence through repeated differentiation, so that the operator L_1 will always be to the left of the operator L_M. We introduce the linear ordering operator $\mathbf{T}$ that permutes operators into an order defined by the splitting from left to right, for example

$$\begin{aligned}\mathbf{T}\left[(L_1+L_2+L_3)^2\right] &= \mathbf{T}\left[L_1^2+L_2^2+L_3^2+L_1L_2+L_2L_1+L_1L_3+L_3L_1+L_2L_3+L_3L_2\right],\\ &= L_1^2+L_2^2+L_3^2+2L_1L_2+2L_1L_3+L_2L_3.\end{aligned}$$

Applying this operator simplifies notation to

$$\frac{\mathrm{d}^n\mathcal{P}_t}{\mathrm{d}t^n} = \mathbf{T}\left[(L_1+\ldots+L_M)^n\,\mathcal{P}_t\right],$$

and hence

$$\begin{aligned}\mathcal{P}_{\delta t} = \mathrm{Id} + \delta t\,\mathbf{T}\left[L_1+\ldots+L_M\right] + \ldots + \frac{\delta t^{n-1}}{(n-1)!}\mathbf{T}\left[(L_1+\ldots+L_M)^{n-1}\right]\\ + \frac{\delta t^n}{n!}\mathbf{T}\left[(L_1+\ldots+L_M)^n\right] + \frac{\delta t^{n+1}}{n!}\int_0^1 (1-\theta)^n \mathbf{T}\left[(L_1+\ldots+L_M)^{n+1}\,\mathcal{P}_{\theta\delta t}\right]\mathrm{d}\theta.\end{aligned} \tag{4.29}$$

This formula is very important in our analysis, as it allows us to find an explicit form of the remainder terms in the perturbed series. This gives our analysis a distinct advantage over the formal analysis, using the generator rather than the evolution operator $\mathcal{P}_{\delta t}$, as we can bound the remainder terms.

However, it is often more convenient to compute the terms in the series itself by computing $\widehat{\mathcal{L}}_{\delta t}$ using the BCH formula and matching powers of δt in the expansion

$$\mathcal{P}_{\delta t} = \ \mathrm{Id} + \delta t\widehat{\mathcal{L}}_{\delta t} + \frac{\delta t^2}{2}\widehat{\mathcal{L}}_{\delta t}^2 + \ldots,$$

yielding a more convenient form

$$\mathbf{T}\left[(L_1+\ldots+L_M)^n\right] = (L_1+\ldots+L_M)^n + S_n,$$

where S_n is a sum of commutator terms involving the L_i. This is particularly useful to note, as for our purposes $(L_1+\ldots+L_M)^n = \mathcal{L}_{\mathrm{LD}}^n$, leading to many cancellations when taking products.

The form of the correction function itself is method dependent, and hence before

proceeding with the computation of the correction terms for schemes, we will introduce the following result allowing us to relate algorithms which have a different final or initial term, but the same core iteration (such algorithms were grouped together in (4.13)).

Lemma 4.30 (TU Lemma)**.** *Consider two ergodic numerical schemes for Langevin dynamics, with evolution operators $\mathcal{Q}_{\delta t}$ and $\mathcal{P}_{\delta t}$ and respective invariant measures $\widehat{\rho}_{\delta t,\mathcal{Q}}$ and $\widehat{\rho}_{\delta t,\mathcal{P}}$. If*

$$(\mathcal{Q}_{\delta t})^n = \mathcal{T}_{\delta t}\ (\mathcal{P}_{\delta t})^{n-1}\ \mathcal{U}_{\delta t}, \tag{4.31}$$

for all $n > 0$ for some bounded operators $\mathcal{T}_{\delta t}$ and $\mathcal{U}_{\delta t}$, then

$$\int_\Omega \phi(q,p)\,\widehat{\rho}_{\delta t,\mathcal{Q}}\,dq\,dp = \int_\Omega \left(\mathcal{U}_{\delta t}\phi(q,p)\right)\,\widehat{\rho}_{\delta t,\mathcal{P}}\,dq\,dp, \tag{4.32}$$

for all bounded, measurable functions $\phi(q,p)$.

Proof. The proof follows directly from the assumption (4.31) and ergodicity estimates. Assuming a smooth initial distribution ρ_0,

$$\begin{aligned}\int_\Omega \phi(q,p)\,\widehat{\rho}_{\delta t,\mathcal{Q}}\mathrm{d}q\mathrm{d}p &= \lim_{n\to\infty}\int_\Omega \left[(\mathcal{Q}_{\delta t})^n\,\phi(q,p)\right]\,\rho_0\,\mathrm{d}q\mathrm{d}p,\\ &= \lim_{n\to\infty}\int_\Omega \left[\mathcal{T}_{\delta t}\ (\mathcal{P}_{\delta t})^{n-1}\ (\mathcal{U}_{\delta t}\phi(q,p))\right]\,\rho_0\,\mathrm{d}q\mathrm{d}p.\end{aligned}$$

The lemma then follows by the ergodic property of the method with evolution operator $\mathcal{P}_{\delta t}$, giving the convergence of distribution in the limit. □

The effect of this lemma is a dramatic simplification in the analysis of classes of methods that share a common sequence of updates. For example, letting

$$\mathcal{T}_{\delta t} = e^{\delta t\mathcal{L}_A}\,e^{\delta t\mathcal{L}_B} \qquad \text{and} \qquad \mathcal{U}_{\delta t} = e^{\delta t\mathcal{L}_C},$$

we can relate the methods

$$\mathcal{P}_{\delta t}^{[\![\mathrm{ABC}]\!]} = \mathcal{T}_{\delta t}\,\mathcal{U}_{\delta t} \qquad \text{and} \qquad \mathcal{P}_{\delta t}^{[\![\mathrm{CAB}]\!]} = \mathcal{U}_{\delta t}\,\mathcal{T}_{\delta t},$$

and hence

$$\int_\Omega \varphi(q,p)\,\widehat{\rho}_{\delta t}^{[\![\mathrm{ABC}]\!]}\mathrm{d}q\mathrm{d}p = \int_\Omega \left[e^{\delta t\mathcal{L}_C}\varphi(q,p)\right]\,\widehat{\rho}_{\delta t}^{[\![\mathrm{CAB}]\!]}\mathrm{d}q\mathrm{d}p.$$

From the definition of $\mathcal{L}_C$, we can see that if φ is a function solely of position then the equilibrium averages will be equal under both methods (this can be easily verified by comparison of the corresponding algorithms themselves). Similar such observations can be made for other classes of methods (or for other special cases of observable φ).

We can reduce our workload by considering one scheme in a particular class and reach conclusions about its related methods through the application of Lemma 4.30.

First-order schemes

Our aim is now to proceed with the analysis of the first-order schemes listed in (4.11) and (4.12). We shall approximate the evolution operator by computing itse expansion (in powers of the step size) using the BCH formula, and ignoring high order terms. By expanding the invariant density similarly, we shall solve the resulting PDEs to find the leading order behaviour of the error terms for these schemes.

We shall consider correcting a sufficiently smooth observable $\varphi \in \widetilde{\mathcal{S}}$ defined in (4.27), and hence

$$\langle \varphi \rangle = \int_\Omega \varphi \, \rho_\beta \mathrm{d}z = 0,$$

as this simplifies our analysis. The general case (for functions $\widehat{\varphi}$ where the average is non-zero) can be recovered by using

$$\varphi = \widehat{\varphi} - \int_\Omega \widehat{\varphi} \, \rho_\beta \mathrm{d}z,$$

in the proof. The goal is to find, for a sufficiently small step size δt, the value of R where

$$\langle \varphi \rangle_{\delta t} = \langle \varphi \rangle + \delta t R + \mathcal{O}(\delta t^2), \tag{4.33}$$

with bounded remainder. We may proceed formally to obtain the form of R, however this gives us little guidance on the form and boundedness of the remainder higher-order terms. Our approach differs from the formal analysis presented in [68] as we use Taylor's formula to find the explicit form of the remainder terms in our correction series, however at leading order our results match those that are found formally.

Let us consider one of the elementary first-order schemes of the form $[\![\mathrm{XYZ}]\!]$, with each $X, Y, Z \in \{A, B, C\}$ distinct, and unique invariant distribution written as $\widehat{\rho}_{\delta t} = \widehat{\rho}_{\delta t}^{[\![\mathrm{XYZ}]\!]}$. By definition, the invariant distribution of the method does not change under the flow, and hence for a sufficiently smooth observable ϕ, with $\langle \phi \rangle = 0$, we have

$$\int_\Omega (\mathcal{P}_{\delta t} \phi) \, \widehat{\rho}_{\delta t} \mathrm{d}q\mathrm{d}p = \int_\Omega \phi \widehat{\rho}_{\delta t} \mathrm{d}q\mathrm{d}p, \tag{4.34}$$

where $\mathcal{P}_{\delta t} = \mathcal{P}_{\delta t}^{[\![\mathrm{XYZ}]\!]}$ is the method's evolution operator.

We can compute the generator for the numerical scheme itself as a series in powers of the step size δt using the BCH formula. Formally we have

$$\widehat{\mathcal{L}}_{\delta t} = \mathcal{L}_{\mathrm{LD}} + \frac{\delta t}{2} [\mathcal{L}_Y, \mathcal{L}_Z] + \frac{\delta t}{2} [\mathcal{L}_X, \mathcal{L}_Y + \mathcal{L}_Z] + \mathcal{O}(\delta t^2),$$

though we have no expression for the evaluation of the remainder terms, and the complexity of the series itself makes its general formula difficult to work with. We will choose not to work with $\widehat{\mathcal{L}}_{\delta t}$, and instead use the formula for the evolution operator (4.29), where we are able to give the form of the remainder. Matching terms in the

series expansion, we obtain

$$\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t} = -\mathcal{L}_{\mathrm{LD}} - \frac{\delta t}{2}\left(\mathcal{L}_{\mathrm{LD}}^2 + S_1\right) - \frac{\delta t^2}{2} R_2,$$

where

$$S_1 = [\mathcal{L}_Y, \mathcal{L}_Z] + [\mathcal{L}_X, \mathcal{L}_Y] + [\mathcal{L}_X, \mathcal{L}_Z], \qquad R_2 = \int_0^1 (1-\theta)^2 \mathcal{R}_{\theta\delta t} \mathrm{d}\theta, \tag{4.35}$$

with remainder $\mathcal{R}_t$ a linear combination of terms of the form $\mathcal{L}_X^a e^{t\mathcal{L}_X} \mathcal{L}_Y^b e^{t\mathcal{L}_Y} \mathcal{L}_Z^c e^{t\mathcal{L}_Z}$ with non-negative integer powers $a+b+c=3$. The remainder can be explicitly computed using the ordering operator $\mathbf{T}$, but involves at most 6 derivatives.

Introducing a perturbation function $f_1(q,p)$, we can compute directly

$$\begin{aligned}
&\int_\Omega \left(\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\phi\right)\left[1+\delta t f_1\right]\rho_\beta \mathrm{d}z \\
&\qquad = -\int_\Omega \left(\left[\mathcal{L}_{\mathrm{LD}} + \frac{\delta t}{2}\left(\mathcal{L}_{\mathrm{LD}}^2 + S_1\right) + \frac{\delta t^2}{2} R_2\right]\phi\right)\left[1+\delta t f_1\right]\rho_\beta \mathrm{d}z, \\
&\qquad = -\delta t \int_\Omega \left[\frac{1}{2} S_1\phi + (\mathcal{L}_{\mathrm{LD}}\phi) f_1\right]\rho_\beta \mathrm{d}z \\
&\qquad\quad - \frac{\delta t^2}{2}\int_\Omega \left[\left(R_2\phi\right)\left(1+\delta t f_1\right) + \left(\mathcal{L}_{\mathrm{LD}}^2\phi + S_1\phi\right) f_1\right]\rho_\beta \mathrm{d}z.
\end{aligned} \tag{4.36}$$

Taking the adjoints of the operators in the leading term on the right hand side, we have

$$\int_\Omega \left[\frac{1}{2} S_1\phi + (\mathcal{L}_{\mathrm{LD}}\phi) f_1\right]\rho_\beta \mathrm{d}z = \int_\Omega \phi\left[\frac{1}{2} S_1^* \rho_\beta + \mathcal{L}_{\mathrm{LD}}^* f_1 \rho_\beta\right]\mathrm{d}z,$$

and hence suggests a choice of f_1 such that

$$\mathcal{L}_{\mathrm{LD}}^* f_1 \rho_\beta = -\frac{1}{2} S_1^* \rho_\beta. \tag{4.37}$$

The function f_1 is well-defined as the assumed smoothness of the potential energy function $U(q)$ ensures that ρ_β is sufficiently smooth, and the Fredholm alternative is satisfied:

$$\int_\Omega S_1^* \rho_\beta \mathrm{d}z = 0.$$

The homogeneous solution to (4.37) is $f_1 = \text{constant}$, hence we can choose the constant to ensure the correction function has average 0 with respect to ρ_β, as required.

Plugging in this choice of f_1 into (4.36) leaves only the second order term on the right hand side, giving

$$\begin{aligned}
&\int_\Omega \left(\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\phi\right)\left[1+\delta t f_1\right]\rho_\beta \mathrm{d}z \\
&\qquad = -\frac{\delta t^2}{2}\int_\Omega \left[\left(R_2\phi\right)\left(1+\delta t f_1\right) + \left(\mathcal{L}_{\mathrm{LD}}^2\phi + S_1\phi\right) f_1\right]\rho_\beta \mathrm{d}z.
\end{aligned} \tag{4.38}$$

Comparing this result with our target equation (4.33), we seek to correct the average of φ, which requires us to replace our observable ϕ in (4.38) by

$$\left(\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\right)^{-1}\varphi,$$

but we have no control over the derivatives of this function, making bounding the remainder problematic. However we only require the term at order δt to vanish, so we can approximate the inverse by any function that gives a remainder in higher powers of δt. If

$$\left(\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\right)\mathcal{Q}_{\delta t} = \mathrm{Id}+\delta t^2 Z_2,$$

then we shall call the operator $\mathcal{Q}_{\delta t}$ an approximate inverse. Writing $\mathcal{Q}_{\delta t} = Q_0 + \delta t Q_1$, we have

$$\left(-\mathcal{L}_{\mathrm{LD}} - \frac{\delta t}{2}\left(\mathcal{L}_{\mathrm{LD}}^2 + S_1\right) - \frac{\delta t^2}{2}R_2\right)(Q_0+\delta t Q_1) = \mathrm{Id}+\delta t^2 Z_2,$$

and hence

$$Q_0 = -\mathcal{L}_{\mathrm{LD}}^{-1}, \qquad Q_1 = \frac{1}{2}\left(\mathrm{Id}+\mathcal{L}_{\mathrm{LD}}^{-1}S_1\mathcal{L}_{\mathrm{LD}}^{-1}\right), \tag{4.39}$$

$$Z_2 = -\frac{1}{2}R_2\left(Q_0+\delta t Q_1\right) - \frac{1}{2}\left(\mathcal{L}_{\mathrm{LD}}^2+S_1\right)Q_1. \tag{4.40}$$

Choosing $f_1(q,p)$ to solve the PDE in (4.37), we combine (4.34) with the result of (4.38) to give

$$\begin{aligned}\int_\Omega \frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\phi\widehat{\rho}_{\delta t}\mathrm{d}z - \int_\Omega \frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\phi\left[1+\delta t f_1(q,p)\right]\rho_\beta \mathrm{d}z \\ = \delta t^2\int_\Omega\left[\left(R_2\phi\right)\left(1+\delta t f_1\right)+\frac{1}{2}\left(\mathcal{L}_{\mathrm{LD}}^2\phi+S_1\phi\right)f_1\right]\rho_\beta\mathrm{d}z.\end{aligned}$$

Since the average of φ is 0 with respect to ρ_β, setting $\phi = \mathcal{Q}_{\delta t}\varphi$ is well-defined, and gives an overall formula

$$\int_\Omega \varphi\widehat{\rho}_{\delta t}\mathrm{d}z = \int_\Omega \varphi\left[1+\delta t f_1(q,p)\right]\rho_\beta\mathrm{d}z + \delta t^2 r_{1,\delta t} = \int_\Omega \varphi\rho_\beta\mathrm{d}z + \delta t\int_\Omega \varphi f_1\rho_\beta\mathrm{d}z + \delta t^2 r_{1,\delta t}, \tag{4.41}$$

where the terms in the correction function f_1 and remainder term $r_{1,\delta t}$ will both depend on the friction γ (and as such may become unbounded in the limiting case of infinite friction). However, for fixed friction and δt sufficiently small the remainder term $r_{1,\delta t}$ is uniformly bounded, as S_1 has finitely many derivatives and applying estimates such as (4.28). It should be noted that this formula can be obtained formally (see [68]), but we do not obtain the form of the remainder terms, which can be computed (and bounded) through the approach presented here.

The function $f_1(q,p)$ here is method dependent, determined by the right hand side

of (4.37), where S_1^* is built from the commutators given in (4.35). Using the skew-symmetry of the commutators, and the fact that

$$\left([\mathcal{L}_C^*, \mathcal{L}_A^*] + [\mathcal{L}_C^*, \mathcal{L}_B^*]\right)\rho_\beta = [\mathcal{L}_C^*, \mathcal{L}_A^* + \mathcal{L}_B^*]\,\rho_\beta = 0,$$

we can see that

$$f_1^{[\![\mathrm{ABC}]\!]} = -f_1^{[\![\mathrm{CBA}]\!]} = f_1^{[\![\mathrm{CAB}]\!]} = -f_1^{[\![\mathrm{BAC}]\!]}, \tag{4.42}$$

with $f_1^{[\![\mathrm{ABC}]\!]}$ solving

$$\mathcal{L}_{\mathrm{LD}}^* f_1^{[\![\mathrm{ABC}]\!]}\rho_\beta = \frac{1}{2}\left[\mathcal{L}_A^*, \mathcal{L}_B^*\right]\rho_\beta = \frac{1}{2}\left(\mathcal{L}_A^* + \mathcal{L}_B^*\right)\mathcal{L}_B^*\rho_\beta, \tag{4.43}$$

where

$$\mathcal{L}_B^*\rho_\beta = -\mathcal{L}_A^*\rho_\beta = g(q,p)\rho_\beta, \qquad g(q,p) = -\beta p^T M^{-1}\nabla U(q).$$

The remaining two methods are related through $f_1^{[\![\mathrm{ACB}]\!]} = -f_1^{[\![\mathrm{BCA}]\!]}$, with

$$\begin{aligned}
\mathcal{L}_{\mathrm{LD}}^* f_1^{[\![\mathrm{ACB}]\!]}\rho_\beta &= \frac{1}{2}\left([\mathcal{L}_C^*, \mathcal{L}_B^* - \mathcal{L}_A^*] + [\mathcal{L}_A^*, \mathcal{L}_B^*]\right)\rho_\beta, \\
&= \frac{1}{2}\left([\mathcal{L}_{\mathrm{LD}}^*, \mathcal{L}_B^* - \mathcal{L}_A^*] - [\mathcal{L}_B^* + \mathcal{L}_A^*, \mathcal{L}_B^* - \mathcal{L}_A^*]\right)\rho_\beta + \mathcal{L}_{\mathrm{LD}}^* f_1^{[\![\mathrm{ABC}]\!]}\rho_\beta, \\
&= \mathcal{L}_{\mathrm{LD}}^*\left(g(q,p)\rho_\beta\right) + [\mathcal{L}_B^*, \mathcal{L}_A^*]\,\rho_\beta + \mathcal{L}_{\mathrm{LD}}^* f_1^{[\![\mathrm{ABC}]\!]}\rho_\beta, \\
&= \mathcal{L}_{\mathrm{LD}}^*\left(g(q,p)\rho_\beta\right) - \mathcal{L}_{\mathrm{LD}}^* f_1^{[\![\mathrm{ABC}]\!]}\rho_\beta,
\end{aligned}$$

and hence applying $\mathcal{L}_{\mathrm{LD}}^{-*}$ we can use (4.42) to relate all of the methods through

$$f_1^{[\![\mathrm{ACB}]\!]} = -f_1^{[\![\mathrm{BCA}]\!]} = g(q,p) - f_1^{[\![\mathrm{ABC}]\!]}. \tag{4.44}$$

We may also derive (4.44) by applying Lemma 4.30 with

$$\mathcal{T}_{\delta t} = e^{\delta t\mathcal{L}_B}e^{\delta t\mathcal{L}_C}, \qquad \mathcal{U}_{\delta t} = e^{\delta t\mathcal{L}_A},$$

giving

$$\int_\Omega \varphi\,\widehat{\rho}_{\delta t}^{[\![\mathrm{BCA}]\!]}\mathrm{d}z = \int_\Omega \left(e^{\delta t\mathcal{L}_A}\varphi\right)\widehat{\rho}_{\delta t}^{[\![\mathrm{ABC}]\!]}\mathrm{d}z,$$

and hence

$$\int_\Omega \varphi\,\left[1 + \delta t f_1^{[\![\mathrm{BCA}]\!]}\right]\rho_\beta \mathrm{d}z = \int_\Omega \varphi\,\left[e^{\delta t\mathcal{L}_A^*}\rho_\beta + \delta t e^{\delta t\mathcal{L}_A^*} f_1^{[\![\mathrm{ABC}]\!]}\rho_\beta\right]\mathrm{d}z.$$

Writing $e^{\delta t\mathcal{L}_A^*} = \mathrm{Id} + \delta t\mathcal{L}_A^* + \delta t^2\hat{R}$ and equating powers of the step size yields

$$f_1^{[\![\mathrm{BCA}]\!]}\rho_\beta = f_1^{[\![\mathrm{ABC}]\!]}\rho_\beta + \mathcal{L}_A^*\rho_\beta, \quad \text{and hence} \quad f_1^{[\![\mathrm{BCA}]\!]} = f_1^{[\![\mathrm{ABC}]\!]} - g(q,p).$$

We have seen that the form of the correction function for the $[\![\mathrm{ABC}]\!]$ scheme is defined by the PDE (4.43). If we can solve this, then we can find all of the other first-

order correction functions by exploiting the relations we have derived between them. In order to find the solution, we write

$$\mathcal{L}^*_{\mathrm{LD}} f_1^{[\![\mathrm{ABC}]\!]} \rho_\beta = \frac{1}{2} \left(\mathcal{L}^*_{\mathrm{LD}} - \gamma \mathcal{L}^*_{\bar{C}} \right) g(q,p) \rho_\beta,$$

where recall $\mathcal{L}_C = \gamma \mathcal{L}_{\bar{C}}$. Computing this obfuscating term gives

$$\begin{aligned}
\mathcal{L}^*_{\bar{C}} g(q,p)\rho_\beta &= \nabla_p \cdot (pg\rho_\beta) + \beta^{-1} \nabla_p \cdot (M \nabla_p) g \rho_\beta, \\
&= N g \rho_\beta + \left(p^T + \beta^{-1} \nabla_p^T M \right) \nabla_p g \rho_\beta, \\
&= N g \rho_\beta + \left(p^T + \beta^{-1} \nabla_p^T M \right) \left(-\beta M^{-1} \nabla U(q) + \beta^2 M^{-1} p p^T M^{-1} \nabla U(q) \right) \rho_\beta, \\
&= N g \rho_\beta + \beta \left(\nabla_p \cdot p p^T M^{-1} \nabla U(q) \right) \rho_\beta, \\
&= N g \rho_\beta + (N+1) \beta p^T M^{-1} \nabla U(q) \rho_\beta, \\
&= -g(q,p) \rho_\beta.
\end{aligned}$$

Hence upon applying the relation

$$\mathcal{L}^*_{\mathrm{LD}} \psi(q) \rho_\beta = -p^T M^{-1} \nabla_q \psi(q) \rho_\beta$$

with $\psi(q) = \beta U(q)$, we can write

$$\mathcal{L}^*_{\bar{C}} g(q,p) \rho_\beta = -g(q,p)\rho_\beta = -\beta \mathcal{L}^*_{\mathrm{LD}} U(q) \rho_\beta,$$

and hence

$$\mathcal{L}^*_{\mathrm{LD}} f_1^{[\![\mathrm{ABC}]\!]} \rho_\beta = \frac{1}{2} \mathcal{L}^*_{\mathrm{LD}} g(q,p) \rho_\beta + \frac{1}{2} \gamma \beta \mathcal{L}^*_{\mathrm{LD}} U(q) \rho_\beta,$$

giving an explicit solution to (4.43) as

$$\boxed{f_1^{[\![\mathrm{ABC}]\!]}(q,p) = \frac{\beta}{2} \left(\gamma U(q) - p^T M^{-1} \nabla U(q) \right) - c_1^{[\![\mathrm{ABC}]\!]},} \tag{4.45}$$

for constant

$$c_1^{[\![\mathrm{ABC}]\!]} = \frac{\gamma \beta}{2} \int_\Omega U(q) \rho_\beta \, \mathrm{d}q \, \mathrm{d}p,$$

ensuring the average of $f_1^{[\![\mathrm{ABC}]\!]}$ is zero. The remaining first-order correction functions are given through the relations in (4.42) and (4.44).

Replacing φ by $(\psi - \langle \psi \rangle)$ in (4.41), for sufficiently smooth ψ where $\langle \psi \rangle$ is not necessarily zero, gives

$$\int_\Omega (\psi - \langle \psi \rangle) \, \widehat{\rho}_{\delta t} \mathrm{d}z = \int_\Omega (\psi - \langle \psi \rangle) \, \rho_\beta \mathrm{d}z + \delta t \int_\Omega (\psi - \langle \psi \rangle) \, f_1 \rho_\beta \mathrm{d}z + \delta t^2 r_{1,\delta t}.$$

Hence, adding the constant $\langle \psi \rangle$, we have that there exists a correction function f_1 such

that for any sufficiently smooth ψ, we have

$$\boxed{\int_\Omega \psi \widehat{\rho}_{\delta t} \mathrm{d}z = \int_\Omega \psi \rho_\beta \mathrm{d}z + \delta t \int_\Omega \psi f_1 \rho_\beta \mathrm{d}z + \delta t^2 r_{1,\delta t},} \tag{4.46}$$

for fixed finite friction γ and sufficiently small step size δt. We can rewrite this result as

$$\langle \psi \rangle_{\delta t} = \langle \psi \rangle + \delta t \langle \psi f_1 \rangle + \delta t^2 r_{1,\delta t}.$$

Of course, this formula is not very useful in practice, as in order to evaluate the correction term we need to know $\langle \psi\, f_1 \rangle$ exactly. The evaluation of $f_1(q,p)$ at any instantaneous point is also tricky, as the constant c_1 was introduced in the solution to ensure it has average 0, and its evaluation requires the computation of this average. If we write

$$f_1 = \widetilde{f}_1 - \langle \widetilde{f}_1 \rangle,$$

for some computationally amenable, sufficiently smooth $\widetilde{f}_1$ (for example, the choice of $\widetilde{f}_1 = f_1 + c_1$ gives an explicit formula), then we can apply the result (4.46) recursively, and approximate the correction terms by their observed averages (as any errors are moved to the higher orders in δt). Finally, this gives

$$\langle \psi \rangle_{\delta t} = \langle \psi \rangle + \delta t \left(\left\langle \psi\, \widetilde{f}_1 \right\rangle_{\delta t} - \langle \psi \rangle_{\delta t} \left\langle \widetilde{f}_1 \right\rangle_{\delta t} \right) + \delta t^2 \widetilde{r}_{1,\delta t}, \tag{4.47}$$

where $\widetilde{r}_{1,\delta t}$ is uniformly bounded for sufficiently small δt.

We compare the observed result with the predicted result in (4.46), for the system with $(q,p) \in \mathbb{R}^2$, canonically sampling an uneven double well potential

$$U(q) = (q^2 - 1)^2 + q/2, \tag{4.48}$$

at $\beta = 1$. We shall look at the error in computed averages of a specific observable

$$\varphi(q,p) = p^2 - qU'(q) + 2qp, \qquad \langle \varphi \rangle = 0, \tag{4.49}$$

and compare the estimates provided by (4.46) in the cases of the six first-order schemes, whose correction functions are given by combining (4.45) with (4.42) and (4.44). We run experiments at 32 different values of δt for two different friction values: $\gamma = 1$ and $\gamma = 2$. Each of the six schemes are integrated over a fixed time interval $t \in [0, 10^8]$, using 16 repeat experiments for each average. The results are given in Figure 4.2, where we plot both the observed average, and the prediction of the observed average from our analysis.

In the case of unit friction, we see a first-order trend that matches exactly our expectation: both the magnitude and sign of the error shows good agreement with the theoretical prediction. However, the ⟦ACB⟧ and ⟦BCA⟧ schemes exhibit higher order behaviour in the simulations computed using $\gamma = 2$. As the equilibrium average should

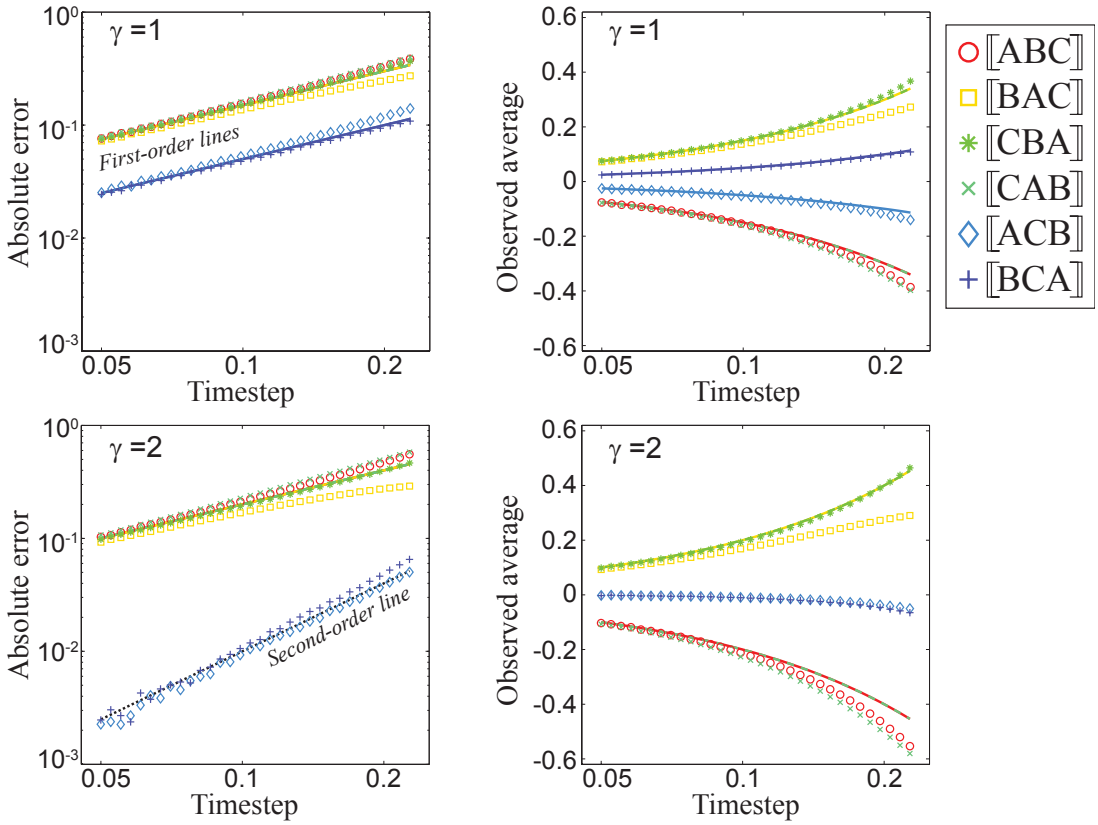


Figure 4.2: We give the error in computed averages of the observable given in (4.49), from simulations using the given timestep and friction for each of the six schemes, with an uneven double-well potential energy function (4.48) at $\beta = 1$. The observed averages (coloured symbols) exactly match the predicted first order error (corresponding coloured lines) at small δt. For $\gamma = 2$, we can see that two of the schemes show higher-order behaviour, which we can explain using our analysis.

be unaffected by the change in friction, this is evidently a phenomenon of the particular numerical schemes employed.

This result is entirely commensurate with our analysis. We compute the correction for ⟦BCA⟧, as the ⟦ACB⟧ and ⟦BCA⟧ schemes have a first order correction that differs only in sign. Thus for general friction $\gamma > 0$, the ⟦BCA⟧ scheme has the first-order correction

$$\begin{aligned}\int_\Omega \psi f_1^{[\![\mathrm{BCA}]\!]} \rho_\beta \mathrm{d}z &= \frac{1}{2}\int_\Omega \left(p^2 - qU'(q) + 2qp\right)\left(pU'(q) + \gamma U(q) - c_1^{[\![\mathrm{BCA}]\!]}\right)\rho_\beta \mathrm{d}q\mathrm{d}p, \\ &= \frac{1}{\beta^2} + \frac{\gamma}{2\beta}\langle U\rangle - \frac{\gamma}{2}\langle qUU'\rangle, \\ &= \frac{(2-\gamma)}{2\beta^2}.\end{aligned}$$

Hence for a critical choice of $\gamma = 2$, the first-order correction term will vanish for the ⟦ACB⟧ and ⟦BCA⟧ schemes, giving (at least) second order errors in observed averages

of the observable (4.49). Additionally, choosing a friction $\gamma = 2 \pm \epsilon$ (for sufficiently small $\epsilon > 0$) shrinks the prefactor of the order δt correction term so that in practice the observed error is dominated by the larger second order term.

For all of the schemes, at the largest tested step sizes we see the emergence of higher-order behaviour influencing the observed averages, as there is a sharp turn away from our prediction. We expect that continuing to find the higher order correction functions would allow us to account for this difference. A more practical approach however, is to consider methods that exhibit higher order behaviour without any correction. We can construct such methods just as in the deterministic case in Section 2.1.4, by considering a symmetric (Strang) splitting scheme.

Second-order schemes

We consider update schemes of the form $[\![\mathrm{XYZYX}]\!]$, with each $X, Y, Z \in \{A, B, C\}$ distinct. The line of attack does not greatly differ from the first-order case, although the additional terms make the computation of the perturbation terms more cumbersome. We can use the symmetric BCH formula (2.35–2.36) to formally write

$$\mathcal{P}_{\delta t}^{[\![\mathrm{XYZYX}]\!]} = e^{\delta t \mathcal{L}_X/2}\, e^{\delta t \mathcal{L}_Y/2}\, e^{\delta t \mathcal{L}_Z}\, e^{\delta t \mathcal{L}_Y/2}\, e^{\delta t \mathcal{L}_X/2} = e^{\delta t \widehat{\mathcal{L}}_{\delta t}},$$

where

$$\widehat{\mathcal{L}}_{\delta t} = \mathcal{L}_X + \mathcal{L}_Y + \mathcal{L}_Z + \delta t^2 S_2 + \mathcal{O}(\delta t^4),$$

with

$$S_2 = \frac{1}{12}\left[\mathcal{L}_Z, [\mathcal{L}_Z, \mathcal{L}_Y]\right] + \frac{1}{12}\left[\mathcal{L}_Y + \mathcal{L}_Z, [\mathcal{L}_Y + \mathcal{L}_Z, \mathcal{L}_X]\right] - \frac{1}{24}\left[\mathcal{L}_Y, [\mathcal{L}_Y, \mathcal{L}_Z]\right] - \frac{1}{24}\left[\mathcal{L}_X, [\mathcal{L}_X, \mathcal{L}_Y + \mathcal{L}_Z]\right]. \tag{4.50}$$

We may then identify terms computed with the BCH formula with those in the Taylor series (4.29), giving

$$\mathcal{P}_{\delta t} = \mathrm{Id} + \delta t\left(\mathcal{L}_{\mathrm{LD}} + \delta t^2 S_2\right) + \frac{\delta t^2}{2}\left(\mathcal{L}_{\mathrm{LD}}^2 + \delta t^2\left(S_2 \mathcal{L}_{\mathrm{LD}} + \mathcal{L}_{\mathrm{LD}} S_2\right)\right) + \frac{\delta t^3}{6}\mathcal{L}_{\mathrm{LD}}^3 + \frac{\delta t^4}{24}\mathcal{L}_{\mathrm{LD}}^4 + \delta t^5 R_{5,\delta t},$$

where we will denote $\mathcal{P}_{\delta t}^{[\![\mathrm{XYZYX}]\!]} = \mathcal{P}_{\delta t}$, and use the property $\mathcal{L}_X + \mathcal{L}_Y + \mathcal{L}_Z = \mathcal{L}_{\mathrm{LD}}$. Rewriting the series gives

$$\frac{\mathrm{Id} - \mathcal{P}_{\delta t}}{\delta t} = -\mathcal{L}_{\mathrm{LD}} - \frac{\delta t}{2}\mathcal{L}_{\mathrm{LD}}^2 - \delta t^2\left(S_2 + \frac{1}{6}\mathcal{L}_{\mathrm{LD}}^3\right) - \delta t^3\left(\frac{1}{24}\mathcal{L}_{\mathrm{LD}}^4 + \frac{1}{2}\left(S_2 \mathcal{L}_{\mathrm{LD}} + \mathcal{L}_{\mathrm{LD}} S_2\right)\right) - \delta t^4 R_{4,\delta t}.$$

We then consider a sufficiently smooth observable $\phi(q,p)$ with $\langle\phi\rangle = 0$, and correction function $f_2(q,p)$ giving

$$\begin{aligned}\int_\Omega \left[\left(\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\right)\phi\right]\left[\left(1+\delta t^2 f_2(q,p)\right)\rho_\beta\right]\mathrm{d}z \\ = -\delta t^2\int_\Omega \phi\left[S_2^*\rho_\beta + \mathcal{L}_{\mathrm{LD}}^* f_2\rho_\beta\right]\mathrm{d}z - \frac{\delta t^3}{2}\int_\Omega \left[\mathcal{L}_{\mathrm{LD}}\phi\right]\left[S_2^*\rho_\beta + \mathcal{L}_{\mathrm{LD}}^* f_2\rho_\beta\right]\mathrm{d}z + \delta t^4\widetilde{R}_{\delta t}.\end{aligned}$$

We now proceed exactly as in the first-order case. We shall choose $f_2(q,p)$ to satisfy

$$\mathcal{L}_{\mathrm{LD}}^* f_2\rho_\beta = -S_2^*\rho_\beta, \tag{4.51}$$

with

$$\int \left(\mathcal{P}_{\delta t}\phi\right)\widehat{\rho}_{\delta t}\mathrm{d}z = \int \phi\widehat{\rho}_{\delta t}\mathrm{d}z,$$

to give

$$\int_\Omega \left[\left(\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\right)\phi\right]\widehat{\rho}_{\delta t}\mathrm{d}z - \int_\Omega \left[\left(\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\right)\phi\right]\left[\left(1+\delta t^2 f_2(q,p)\right)\rho_\beta\right]\mathrm{d}z = -\delta t^4\widetilde{R}_{\delta t}. \tag{4.52}$$

Choosing an approximate inverse $\mathcal{Q}_{\delta t}$ such that

$$\left(\frac{\mathrm{Id}-\mathcal{P}_{\delta t}}{\delta t}\right)\mathcal{Q}_{\delta t} = \mathrm{Id} + \delta t^4 Z_4,$$

and plugging $\phi = \mathcal{Q}_{\delta t}\varphi$ into (4.52) yields

$$\int_\Omega \varphi\widehat{\rho}_{\delta t}\mathrm{d}z = \int_\Omega \varphi\rho_\beta\mathrm{d}z + \delta t^2\int_\Omega \varphi f_2\rho_\beta\mathrm{d}z + \delta t^4 r_{4,\delta t},$$

where only even powers of the step size appear due to the symmetric nature of the considered schemes. Similarly replacing $\varphi = \psi - \langle\psi\rangle$ for suitable ψ gives the correction for second schemes to be

$$\langle\psi\rangle_{\delta t} = \langle\psi\rangle + \delta t^2\langle\psi\, f_2\rangle + \delta t^4 r_{4,\delta t}, \tag{4.53}$$

where $r_{4,\delta t}$ is uniformly bounded for sufficiently small δt and where f_2 is the method-dependent solution to the PDE (4.51). We now desire to solve the PDE in order to find the correction function f_2, so that we may apply this result to obtain higher order averages. However, solving the equation is evidently not as straightforward as in the first-order case, owing to the high order derivatives that appear.

We shall attempt to simplify the PDE to solve as much as possible. From the definition of S_2 in (4.50), it is clearly more convenient to initially consider methods with $\mathcal{L}_X^*\rho_\beta = 0$ and $\left(\mathcal{L}_Y + \mathcal{L}_Z\right)^*\rho_\beta = 0$, so we shall start by considering the scheme

$[\![\mathrm{CBABC}]\!]$, which has

$$S_2^*\rho_\beta = \frac{1}{12}\left[\mathcal{L}_A,[\mathcal{L}_A,\mathcal{L}_B]\right]^*\rho_\beta - \frac{1}{24}\left[\mathcal{L}_B,[\mathcal{L}_B,\mathcal{L}_A]\right]^*\rho_\beta.$$

The terms are computed as

$$\begin{aligned}\left[\mathcal{L}_A,[\mathcal{L}_A,\mathcal{L}_B]\right]^*\rho_\beta &= \left[\mathcal{L}_A^*,[\mathcal{L}_A^*,\mathcal{L}_B^*]\right]\rho_\beta,\\ &= \left(\mathcal{L}_A^*\mathcal{L}_A^*\mathcal{L}_B^* - 2\mathcal{L}_A^*\mathcal{L}_B^*\mathcal{L}_A^* + \mathcal{L}_B^*\mathcal{L}_A^*\mathcal{L}_A^*\right)\rho_\beta,\\ &= \left(\mathcal{L}_A^*\mathcal{L}_A^* + 2\mathcal{L}_A^*\mathcal{L}_B^* - \mathcal{L}_B^*\mathcal{L}_A^*\right)g\rho_\beta,\\ &= \left(\mathcal{L}_A^* + \mathcal{L}_B^*\right)\mathcal{L}_A^* g\rho_\beta + 2\left[\mathcal{L}_A^*,\mathcal{L}_B^*\right]g\rho_\beta.\end{aligned}$$

Similarly,

$$\left[\mathcal{L}_B,[\mathcal{L}_B,\mathcal{L}_A]\right]^*\rho_\beta = -\left(\mathcal{L}_B^* + \mathcal{L}_A^*\right)\mathcal{L}_B^* g\rho_\beta + 2\left[\mathcal{L}_A^*,\mathcal{L}_B^*\right]g\rho_\beta,$$

giving

$$\mathcal{L}_{\mathrm{LD}}^* f_2^{[\![\mathrm{CBABC}]\!]}\rho_\beta = -\frac{1}{12}\left(\mathcal{L}_A^* + \mathcal{L}_B^*\right)\left(\mathcal{L}_A^* + \frac{1}{2}\mathcal{L}_B^*\right)g\rho_\beta - \frac{1}{12}\left[\mathcal{L}_A^*,\mathcal{L}_B^*\right]g\rho_\beta, \tag{4.54}$$

and by computing the related formula for $[\![\mathrm{CABAC}]\!]$ we obtain

$$\mathcal{L}_{\mathrm{LD}}^* f_2^{[\![\mathrm{CABAC}]\!]}\rho_\beta = \frac{1}{12}\left(\mathcal{L}_A^* + \mathcal{L}_B^*\right)\left(\frac{1}{2}\mathcal{L}_A^* + \mathcal{L}_B^*\right)g\rho_\beta - \frac{1}{12}\left[\mathcal{L}_A^*,\mathcal{L}_B^*\right]g\rho_\beta. \tag{4.55}$$

The explicit form of each of the right hand sides is given in (4.92). The equations for the remaining two methods can be computed through the computation of the respective right-hand sides of (4.51) using their corresponding operators S_2, or, alternatively, we may apply Lemma 4.30 with

$$\mathcal{T}_{\delta t} = e^{\delta t\mathcal{L}_A/2}\,e^{\delta t\mathcal{L}_B/2}\,e^{\delta t\mathcal{L}_C/2}, \qquad \mathcal{U}_{\delta t} = e^{\delta t\mathcal{L}_C/2}\,e^{\delta t\mathcal{L}_B/2}\,e^{\delta t\mathcal{L}_A/2},$$

giving

$$\int_\Omega \varphi\widehat{\rho}_{\delta t}^{[\![\mathrm{ABCBA}]\!]}\mathrm{d}z = \int_\Omega \left(\mathcal{U}_{\delta t}\varphi\right)\widehat{\rho}_{\delta t}^{[\![\mathrm{CBABC}]\!]},$$

and hence

$$\int_\Omega \varphi\left[\left(1+\delta t^2 f_2^{[\![\mathrm{ABCBA}]\!]}\right)\rho_\beta\right]\mathrm{d}z = \int_\Omega \varphi\left[\mathcal{U}_{\delta t}^*\rho_\beta + \delta t^2\mathcal{U}_{\delta t}^* f_2^{[\![\mathrm{CBABC}]\!]}\rho_\beta\right]\mathrm{d}z. \tag{4.56}$$

Utilizing the expansion

$$\mathcal{U}_{\delta t}^* = \mathrm{Id} + \frac{\delta t}{2}\mathcal{L}_{\mathrm{LD}}^* + \frac{\delta t^2}{8}\left(\left(\mathcal{L}_{\mathrm{LD}}^*\right)^2 + \left[\mathcal{L}_A^* + \mathcal{L}_B^*,\mathcal{L}_C^*\right] + \left[\mathcal{L}_A^*,\mathcal{L}_B^*\right]\right) + \delta t^3 R_{U,\delta t},$$

and equating powers of the step size in (4.56) gives

$$f_2^{[\![\mathrm{ABCBA}]\!]} = f_2^{[\![\mathrm{CBABC}]\!]} + \frac{1}{8}\left(\mathcal{L}_A^* + \mathcal{L}_B^*\right)g, \tag{4.57}$$

and similarly

$$f_2^{[\![\mathrm{BACAB}]\!]} = f_2^{[\![\mathrm{CABAC}]\!]} - \frac{1}{8}\left(\mathcal{L}_A^* + \mathcal{L}_B^*\right) g, \tag{4.58}$$

where, recall, $g(q,p) = -\beta p \cdot M^{-1}\nabla U(q)$.

Unfortunately, we cannot solve the PDE (4.51) to find explicit solutions for the correction functions for the second order schemes in the general case. However, we will give a strategy in Section 5.1 that permits us to numerically compute the correction terms.

GLA schemes

The first-order GLA schemes reduce to precisely the first order methods considered in previous sections. The second order GLA schemes are of the form $[\![\mathrm{XYXC}]\!]$ for $X, Y \in \{A, B\}$ distinct, where we expect an expansion of

$$\int_\Omega \varphi \widehat{\rho}_{\delta t}^{[\![\mathrm{XYXC}]\!]} = \int_\Omega \varphi \left[\rho_\beta \left(1 + \delta t^2 f_2^{[\![\mathrm{XYXC}]\!]} + \delta t^3 f_3^{[\![\mathrm{XYXC}]\!]}\right)\right] + \delta t^4 r_4. \tag{4.59}$$

Unlike the previously considered second-order schemes, the third order terms do not cancel due to symmetry. Applying Lemma 4.30 with

$$\mathcal{T}_{\delta t} = e^{\delta t \mathcal{L}_X/2} e^{\delta t \mathcal{L}_Y} e^{\delta t \mathcal{L}_X/2} e^{\delta t \mathcal{L}_C/2}, \qquad \mathcal{U}_{\delta t} = e^{\delta t \mathcal{L}_C/2},$$

gives

$$\begin{aligned}
\int_\Omega \varphi \widehat{\rho}_{\delta t}^{[\![\mathrm{XYXC}]\!]} \,\mathrm{d}z &= \int_\Omega \left(\mathcal{U}_{\delta t}\varphi\right) \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \,\mathrm{d}z, \\
&= \int_\Omega \left(\mathcal{U}_{\delta t}\varphi\right) \left[\rho_\beta \left(1 + \delta t^2 f_2^{[\![\mathrm{CXYXC}]\!]}\right)\right] \,\mathrm{d}z + \delta t^4 r_{4,\delta t}, \\
&= \int_\Omega \varphi \rho_\beta \,\mathrm{d}z + \delta t^2 \int_\Omega \left(\mathcal{U}_{\delta t}\varphi\right) f_2^{[\![\mathrm{CXYXC}]\!]} \rho_\beta \,\mathrm{d}z + \delta t^4 r_{4,\delta t}.
\end{aligned}$$

The perturbation term is

$$\begin{aligned}
\int_\Omega \left(\mathcal{U}_{\delta t}\varphi\right) f_2^{[\![\mathrm{CXYXC}]\!]} \rho_\beta \,\mathrm{d}z &= \int_\Omega e^{\delta t \mathcal{L}_C} \varphi \left(f_2^{[\![\mathrm{CXYXC}]\!]} \rho_\beta\right) \,\mathrm{d}z, \\
&= \int_\Omega \varphi \left(e^{\delta t \mathcal{L}_C^*} f_2^{[\![\mathrm{CXYXC}]\!]} \rho_\beta\right) \,\mathrm{d}z, \\
&= \int_\Omega \varphi f_2^{[\![\mathrm{CXYXC}]\!]} \rho_\beta \,\mathrm{d}z + \frac{\delta t}{2} \int_\Omega \varphi \mathcal{L}_C^* f_2^{[\![\mathrm{CXYXC}]\!]} \rho_\beta \,\mathrm{d}z + \delta t^2 \widetilde{r}_{U,\delta t}, \\
&= \int_\Omega \varphi f_2^{[\![\mathrm{CXYXC}]\!]} \rho_\beta \,\mathrm{d}z + \frac{\delta t}{2} \int_\Omega \varphi \rho_\beta \left(\mathcal{L}_C f_2^{[\![\mathrm{CXYXC}]\!]}\right) \,\mathrm{d}z \\
&\quad + \delta t^2 \widetilde{r}_{U,\delta t},
\end{aligned}$$

and hence comparing subsequent powers of δt with those in (4.59) gives

$$f_2^{[\![\mathrm{XYXC}]\!]} = f_2^{[\![\mathrm{CXYXC}]\!]}, \qquad f_3^{[\![\mathrm{XYXC}]\!]} = \frac{1}{2}\mathcal{L}_C f_2^{[\![\mathrm{CXYXC}]\!]}. \tag{4.60}$$

Of course we have

$$\int_\Omega \varphi\rho_\beta \left(\mathcal{L}_C f_2^{[\![\mathrm{CXYXC}]\!]}\right) \mathrm{d}z = \int_\Omega \left(\mathcal{L}_C\varphi\right) \rho_\beta f_2^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z,$$

removing the third order correction for $\varphi = \varphi(q)$. Considering instead the scheme $[\![\mathrm{CXYX}]\!]$ with similar expansion to (4.59), we can use

$$\mathcal{T}_{\delta t} = e^{\delta t\mathcal{L}_C/2} e^{\delta t\mathcal{L}_X/2} e^{\delta t\mathcal{L}_Y} e^{\delta t\mathcal{L}_X/2}, \qquad \mathcal{U}_{\delta t} = e^{\delta t\mathcal{L}_C/2},$$

similarly to find

$$\begin{aligned}
\int_\Omega \varphi\widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \,\mathrm{d}z &= \int_\Omega \left(\mathcal{U}_{\delta t}\varphi\right) \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYX}]\!]} \,\mathrm{d}z, \\
&= \int_\Omega \left(\mathcal{U}_{\delta t}\varphi\right) \left[\rho_\beta \left(1 + \delta t^2 f_2^{[\![\mathrm{CXYX}]\!]} + \delta t^3 f_3^{[\![\mathrm{CXYX}]\!]}\right)\right] \,\mathrm{d}z + \delta t^4 \widehat{r}_{4,\delta t}, \\
&= \int_\Omega \varphi\rho_\beta \,\mathrm{d}z + \delta t^2 \int_\Omega \varphi f_2^{[\![\mathrm{CXYX}]\!]} \rho_\beta \,\mathrm{d}z \\
&\quad + \delta t^3 \int_\Omega \varphi \left(f_3^{[\![\mathrm{CXYX}]\!]} + \frac{1}{2}\mathcal{L}_C f_2^{[\![\mathrm{CXYX}]\!]}\right) \rho_\beta \,\mathrm{d}z + \delta t^4 \widehat{r}_{T,\delta t}.
\end{aligned}$$

Once again equating powers gives correction functions

$$f_2^{[\![\mathrm{CXYX}]\!]} = f_2^{[\![\mathrm{CXYXC}]\!]}, \qquad f_3^{[\![\mathrm{CXYX}]\!]} = -\frac{1}{2}\mathcal{L}_C f_2^{[\![\mathrm{CXYXC}]\!]}. \tag{4.61}$$

4.2 Performance for harmonic systems

The harmonic case is one of the most relevant for molecular dynamics simulations as the timestep is inevitably chosen as close as possible to the timestep stability threshold set by the fastest degrees of freedom in the system: the intramolecular covalent bond stretches [4]. A great deal of the overall behaviour of the system is determined by the harmonic bonds as they imbue the structure of the molecule itself, with the length and frequency of the bonds chosen by chemists with great precision.

4.2.1 Quadratic potential energy functions

We consider a general quadratic potential energy function

$$V(q) = q^T Aq/2 + b^T q + c,$$

for constant matrix positive definite matrix $A \in \mathbb{R}^{N\times N}$, constant vector $b \in \mathbb{R}^N$ and constant scalar $c \in \mathbb{R}$. The Newtonian equations of motion are

$$\mathrm{d}q = M^{-1}p \,\mathrm{d}t, \qquad \mathrm{d}p = (-Aq - b) \,\mathrm{d}t,$$

where upon writing $A = X^{-1}DX$, for positive diagonal D, we can change variables $\hat{Q} = Xq - D^{-1}Xb$ and $\hat{P} = Xp$ to find

$$\mathrm{d}\hat{Q} = M^{-1}\hat{P}\,\mathrm{d}t, \qquad \mathrm{d}\hat{P} = -D\hat{Q}\,\mathrm{d}t.$$

Through a linear change of variables in the system (amounting to a change of basis) we recover a decoupled system of harmonic oscillators. This simplifies greatly analysis, as we only have to consider a model with $q, p \in \mathbb{R}$.

We shall consider sampling a system with Hamiltonian

$$H(q,p) = \omega^2 q^2/2 + p^2/(2m), \tag{4.62}$$

with constant frequency $\omega \in R$ and mass $m > 0$, and investigate the value of the observed averages $\langle \omega^2 q^2 \rangle_{\delta t}$ and $\langle p^2/m \rangle_{\delta t}$, which both have exact value β^{-1} where β is the reciprocal temperature. We make use of the technique for finding the exact invariant averages (without solving for the invariant measure) given in [15].

For any scheme using the ABC splitting (4.5) or the 'AS' splitting (4.3), we may write the update scheme as

$$\begin{bmatrix} q_{n+1} \\ p_{n+1} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} q_n \\ p_n \end{bmatrix} + \begin{bmatrix} Q_n \\ P_n \end{bmatrix},$$

where a, b, c, d are constants depending on the step size δt, and Q_n, P_n are mean zero time-dependent processes independent of the system state $[q, p]$. For our purposes, the $[Q_n, P_n]$ will be normally distributed random numbers with a prescribed mean and variance depending on γ, β and δt.

Taking products of the update equations, we have

$$\begin{aligned}
q_{n+1}^2 &= a^2 q_n^2 + b^2 p_n^2 + Q_n^2 + 2abq_n p_n + 2aQ_n q_n + 2bQ_n p_n, \\
p_{n+1}^2 &= c^2 q_n^2 + d^2 p_n^2 + P_n^2 + 2cdq_n p_n + 2cP_n q_n + 2dP_n p_n, \\
q_{n+1}p_{n+1} &= acq_n^2 + bdp_n^2 + Q_n P_n + (ad+bc)\, q_n p_n + (cq_n + dp_n)\, Q_n + (aq_n + bp_n)\, P_n.
\end{aligned}$$

Then taking expectations in the limit $n \to \infty$ we get

$$\begin{aligned}
\langle q^2 \rangle_{\delta t} &= a^2 \langle q^2 \rangle_{\delta t} + b^2 \langle p^2 \rangle_{\delta t} + \langle Q^2 \rangle_{\delta t} + 2ab \langle qp \rangle_{\delta t}, \\
\langle p^2 \rangle_{\delta t} &= c^2 \langle q^2 \rangle_{\delta t} + d^2 \langle p^2 \rangle_{\delta t} + \langle P^2 \rangle_{\delta t} + 2cd \langle qp \rangle_{\delta t}, \\
\langle qp \rangle_{\delta t} &= ac \langle q^2 \rangle_{\delta t} + bd \langle p^2 \rangle_{\delta t} + (ad+bc) \langle qp \rangle_{\delta t} + \langle QP \rangle_{\delta t},
\end{aligned}$$

where the processes Q and P are statistically independent of the state variables q, p, and $\langle Q \rangle = \langle P \rangle = 0$. Assuming that we know the distributions of Q and P, and hence the values of $\langle Q^2 \rangle_{\delta t}$ and $\langle P^2 \rangle_{\delta t}$, this gives a closed system that can be solved to give the exact values for the stationary averages for the harmonic system with Hamiltonian (4.62) in terms of its update parameters under changes in the frequency, temperature

Scheme	$\langle q^2\rangle_{\delta t}/\langle q^2\rangle$	$\langle p^2\rangle_{\delta t}/\langle p^2\rangle$	$\langle qp\rangle_{\delta t}$
⟦ABC⟧	$[(1+\alpha)/\alpha]\,\Lambda(\alpha\delta t^2)$	$2\Lambda(\alpha\delta t^2)$	$-\delta t\beta^{-1}\Lambda(\alpha\delta t^2)$
⟦CBA⟧	$(1+\alpha)\,\Lambda(\delta t^2)$	$2\Lambda(\delta t^2)$	$\delta t\beta^{-1}\Lambda(\delta t^2)$
⟦CAB⟧	$[(1+\alpha)/\alpha]\,\Lambda(\alpha\delta t^2)$	$1+\dfrac{\delta t^2\omega^2\Lambda(\alpha\delta t^2)}{\alpha m(1+\alpha)}$	$-\delta t\beta^{-1}\Lambda(\alpha\delta t^2)/\alpha$
⟦BAC⟧	$(1+\alpha)\,\Lambda(\delta t^2)$	$1+\dfrac{\delta t^2\omega^2\alpha^2\Lambda(\delta t^2)}{m(1+\alpha)}$	$\delta t\beta^{-1}\alpha\Lambda(\delta t^2)$
⟦ACB⟧	$(1+\alpha)\,\Lambda(\delta t^2)$	$2\Lambda(\delta t^2)$	$-\delta t\beta^{-1}\Lambda(\delta t^2)$
⟦BCA⟧	$[(1+\alpha)/\alpha]\,\Lambda(\alpha\delta t^2)$	$2\Lambda(\alpha\delta t^2)$	$\delta t\beta^{-1}\Lambda(\alpha\delta t^2)$
⟦ABCBA⟧	1	$[1-\delta t^2\omega^2/(4m)]^{-1}$	0
⟦CBABC⟧	$[1-\delta t^2\omega^2/(4m)]^{-1}$	1	0
⟦BACAB⟧	1	$1-\delta t^2\omega^2/(4m)$	0
⟦CABAC⟧	$1-\delta t^2\omega^2/(4m)$	1	0
⟦ASA⟧	$\gamma\delta t(1+\alpha)/(2-2\alpha)$	$2\Lambda(\delta t(\alpha-1)/\gamma)$	0

Table 4.1: The expected long-time computed average using each elementary first and second-order Langevin dynamics method resulting from the ABC splitting as well as the Stochastic Position Verlet scheme ⟦ASA⟧, with quadratic Hamiltonian (4.62) and we use $\Lambda(t) := m(\alpha+1)/(\omega^2 t - 2m[\alpha+1])$ for brevity.

and friction.

As an example, the method coded ⟦ABC⟧ has update constants

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & \delta t/m \\ -\delta t\alpha\omega^2 & \alpha\left(1-\delta t^2\omega^2/m\right) \end{bmatrix}, \qquad \begin{bmatrix} Q_n \\ P_n \end{bmatrix} = \begin{bmatrix} 0 \\ \beta^{-1/2}\sqrt{1-\alpha^2}R_n \end{bmatrix}, \tag{4.63}$$

for $R_n \sim \mathcal{N}(0,1)$ and $\alpha = \exp(-\gamma\delta t)$. We compute the solution for a number of methods, and tabulate the results in Table 4.1.

The first order methods do very poorly, with some cross-correlation occurring between q and p, i.e. $\langle qp\rangle_{\delta t} \neq 0$. The four second-order schemes do much better though, each one giving either exact position and momentum sampling.

The step size limit $\delta t < 2\sqrt{m}/\omega$ is evident for the second order schemes, independent from the friction γ. The threshold for the first order schemes is more complicated, and is a function of the friction (though sampling is inconsistent at the limit of large friction).

Similarly we see that for the stochastic position Verlet (⟦ASA⟧) scheme, also inconsistent for large friction, the error in the average position if of the order $\mathcal{O}(\gamma^2\delta t^2)$. We expect that for small friction, this scheme would give excellent configurational sampling results for harmonic systems.

4.2.2 Perturbed harmonic oscillator

We consider now perturbing slightly the harmonic oscillator potential, with scalar q and using

$$U(q) = q^2/2 + \epsilon q^4/4, \tag{4.64}$$

where $\epsilon > 0$ is a small constant. We cannot write the update schemes in the style of (4.63) as the derivative is no longer linear in q. As a result, the equation for the expectation of $(q_{n+1})^k$ will always involve powers of $(q_n)^j$, for $j > k$, preventing a closed system of equations when we take expectations. For example, if we seek $\langle q^2 \rangle_{\delta t}$, then computing the update q_{n+1}^2 gives terms $q_n^2, \epsilon q_n^4$ and $\epsilon^2 q_n^6$. Taking expectations requires us to compute $\langle q^4 \rangle_{\delta t}$ and $\langle q^6 \rangle_{\delta t}$, which in turn give higher powers.

The terms stymieing the closure of the system have powers of ϵ as prefactors, so if we are willing to solve for the value of $\langle q^2 \rangle_{\delta t}$ up to order ϵ^k, we need only keep terms of sufficient size. Our strategy will to truncate the system of equations as needed : appearing powers of q_n and p_n in the system of equations of order $2j$ will only need to be resolved to order ϵ^{k+1-j}. Odd order terms will not appear due to $\langle R_n^{2j+1} \rangle = 0$, and the statistical independence of R_n from the state variables.

Therefore, choosing some k as a truncation level, in order to resolve $\langle q^2 \rangle_{\delta t}$ to order ϵ^k we only need to keep the powers of q_n and p_n up to order $2k+2$. The $2k+4$-many equations for the state variables of order $2k+2$ will be truncated at order ϵ^0, making them solvable just as in the vanilla harmonic case.

Considering powers up to ϵ^k, we can take expectation of the resulting update equations to obtain a closed linear system of $k(k+2)$ equations to solve. As sampling the momentum in this context is not interesting (it is trivial to sample from its Gaussian distribution), we instead look solely at the averages of $\langle q^2 \rangle_{\delta t}$. We find that choosing $k = 2$ is sufficient to explicitly resolve the leading order behaviour of the system, and we give the differences between observed numerical averages and exact averages below, where the superscript indicates the considered method:

$$\begin{aligned}
\left\langle q^2 \right\rangle_{\delta t}^{[\![\mathrm{ABCBA}]\!]} - \left\langle q^2 \right\rangle &= \frac{3\delta t^2 \epsilon \left[48\epsilon - 4 + 3\gamma^2(9\epsilon - 1)\right]}{12\gamma^2 + 16} + \mathcal{O}(\epsilon \delta t^4 + \epsilon^3 \delta t^2), \\
\left\langle q^2 \right\rangle_{\delta t}^{[\![\mathrm{BACAB}]\!]} - \left\langle q^2 \right\rangle &= -\frac{9\delta t^2 \epsilon^2}{6\gamma^2 + 8} - \frac{3\delta t^4 \epsilon^2 \left[6\gamma^4 + 19\gamma^2 + 16\right]}{8\left(3\gamma^2 + 4\right)^2} + \mathcal{O}(\epsilon^2 \delta t^6 + \epsilon^3 \delta t^2), \\
\left\langle q^2 \right\rangle_{\delta t}^{[\![\mathrm{CBABC}]\!]} - \left\langle q^2 \right\rangle &= \delta t^2 \left(\frac{1}{4} + \frac{\epsilon \left[144\epsilon + 81\epsilon\gamma^2 - 9\gamma^2 - 12\right]}{12\gamma^2 + 16} \right) + \mathcal{O}(\delta t^4 + \epsilon^3 \delta t^2), \\
\left\langle q^2 \right\rangle_{\delta t}^{[\![\mathrm{CABAC}]\!]} - \left\langle q^2 \right\rangle &= -\delta t^2 \left(\frac{1}{4} + \frac{18\epsilon^2}{12\gamma^2 + 16} \right) + \mathcal{O}(\delta t^4 + \epsilon^3 \delta t^2).
\end{aligned} \tag{4.65}$$

The behaviour of the methods can thus be summarized as

$$\begin{aligned}
\left\langle q^2 \right\rangle_{\delta t}^{[\![\text{ABCBA}]\!]} - \left\langle q^2 \right\rangle &= \mathcal{O}\left(\epsilon \delta t^2\right), \\
\left\langle q^2 \right\rangle_{\delta t}^{[\![\text{BACAB}]\!]} - \left\langle q^2 \right\rangle &= \mathcal{O}\left(\frac{\epsilon^2 \delta t^2}{1+\gamma^2} + \epsilon^2 \delta t^4\right), \\
\left\langle q^2 \right\rangle_{\delta t}^{[\![\text{CBABC}]\!]} - \left\langle q^2 \right\rangle &= \mathcal{O}\left(\delta t^2\right), \\
\left\langle q^2 \right\rangle_{\delta t}^{[\![\text{CABAC}]\!]} - \left\langle q^2 \right\rangle &= \mathcal{O}\left(\delta t^2\right).
\end{aligned} \tag{4.66}$$

We would expect that both the ⟦ABCBA⟧ and ⟦BACAB⟧ schemes will have a leading error term proportional to ϵ, given that it was shown in Table 4.1 that in the pure harmonic case (at $\epsilon = 0$) these methods give perfect sampling. The twofold surprising result though, is that the leading error of the ⟦BACAB⟧ scheme has behaviour proportional to ϵ^2, as well as the leading order term being annihilated for large γ.

It would seem that by ordering the ABC update terms favourably in an algorithm we are able to achieve significant sampling improvements on systems that are "closely harmonic", at no extra cost (in terms of force evaluations).

To illustrate this numerically, we use each of the four given second order methods to compute the absolute error in $\left\langle q^2 \right\rangle_{\delta t}$ using the perturbed harmonic potential (4.64) at a range of timesteps $\delta t \in [0.1, 2]$ and nonlinearity parameters $\epsilon \in [0.0001, 1]$. We choose $\beta = \gamma = 1$ for all experiments, and run the schemes using each parameterset for a fixed time window of length $T = 10^8$. The results are shown in Figure 4.3, and agree with the results of (4.66).

The ⟦BACAB⟧ scheme performs extremely well in this example, even in regimes of large ϵ and step size where our formal analysis is no longer valid. Though Table 4.1 demonstrates that both the ⟦ABCBA⟧ and ⟦BACAB⟧ schemes give exact results at $\epsilon = 0$, there remains qualitative difference between the two, owing to the order ϵ^2 leading term in the latter scheme. However, even at large values of the nonlinearity such as $\epsilon = 0.1$, we can achieve sampling of a comparable quality to the ⟦CBABC⟧ scheme by using the ⟦BACAB⟧ scheme at timesteps an order of magnitude larger.

There is also an evident kink in the error for the plot of the ⟦CBABC⟧ scheme: for fixed δt and increasing ϵ the error decreases (most easily seen by tracing the yellow contour of constant error). This is explained by considering the leading term in (4.65), where the prefactor for the δt^2 term decreases in size as ϵ increases, giving rise to the observed behaviour for fixed δt. In contrast, the prefactor to the second order term in the ⟦CABAC⟧ scheme increases at order ϵ^2, leading to an behaviour in averages that appears indifferent to ϵ.

Finally, we comment that the stability of the ⟦BACAB⟧ and ⟦CABAC⟧ class of methods is greatly improved over the alternative for small ϵ. It is unsurprising of course that these methods share stability properties as they clearly share the same internal structure when iterated. Indeed, simulations were successfully completed for all $\epsilon \leq 0.003$ at the maximum tested timestep for these methods, whereas ⟦ABCBA⟧

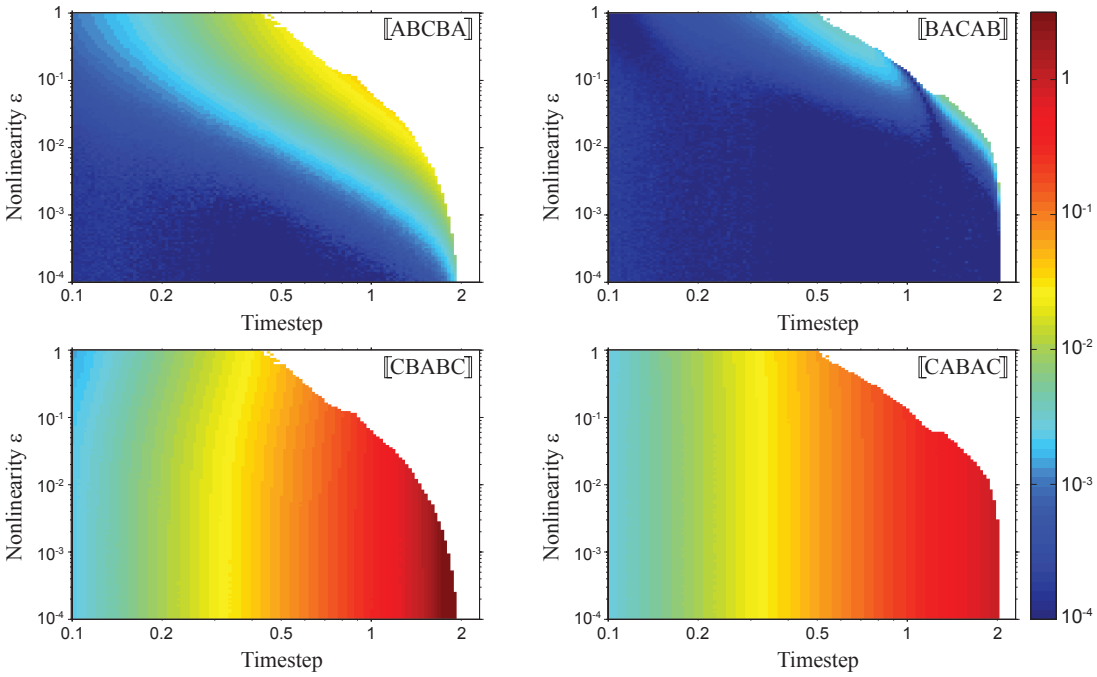


Figure 4.3: A 128×128 grid showing the error in $\langle q^2 \rangle_{\delta t}$ computed with one of the four elementary second order methods, using step size δt (horizontal) and potential energy (4.64) at a fixed value of ϵ (vertical). Pixels are coloured by the absolute error in the observed average, with white pixels denoting instability at that parameterset.

and ⟦CBABC⟧ required a significant reduction of ∼15% before the methods were stable. Even at the lowest ϵ tested, the ⟦ABCBA⟧ and ⟦CBABC⟧ schemes were not stable at the maximum δt used.

4.2.3 Stability for systems with multiple scales

Recall the (r-RESPA) scheme, designed to propagate systems which have a fast highly oscillatory motion, as well as a slower motion. Splitting the potential energy function into fast and slow components in the deterministic setting such as equation (2.39) proved unsuccessful, as we encountered linear resonance phenomena at step sizes that were integer multiples of half a period [9]. In systems where the timescale gap is very large, the r-RESPA scheme does not give significant improvement, with the stability threshold scaling identically to the the Symplectic Euler method [72]. Our hope would be that using Langevin dynamics to add stochasticity would disrupt resonance effects and permit larger usable timesteps. While some considerable efficiency improvements have been demonstrated using multiple timestepping methods for Langevin dynamics with constraints [30], we shall investigate the simpler fast/slow splitting demonstrated in Section 2.1.5. We would expect that a full treatment of the weak order for such constraint algorithms could be achieved in a similar manner to the Langevin splitting algorithms given in previous sections (for example see [73, 118]).

We consider propagating a system with Hamiltonian

$$H(q,p) = p^2/2 + U(q), \qquad U(q) = U_F(q) + U_S(q)$$

where $q, p \in \mathbb{R}$ and U_S and U_F denote the slow and fast components of the potential energy respectively. For the purposes of analysis, we shall model these with harmonic potentials, with

$$U_S(q) = q^2/2, \qquad U_F(q) = q^2/(2\epsilon^2), \tag{4.67}$$

where $\epsilon > 0$ is a given small constant. The question we shall be interested in is whether Langevin dynamics gives us any extra stability for methods that do not dynamics into fast and slow terms, such as those using the general ABC splitting. In the case of the first order methods, we can see that any stability relationship between the friction and step size are not at all, useful as the methods are not consistent for general potentials in the case of large friction. For the remaining second-order methods using the ABC splitting, examining Table 4.1 with $m = 1$ and $\omega^2 = (1 + \epsilon^{-2})$, we can see all of the methods are stable only for

$$\delta t \leq \frac{2}{\sqrt{1+\epsilon^{-2}}},$$

indicated by a divergence or sign change in the averages. This is the same stability condition for the deterministic case (2.40), suggesting that the stochasticity has not bought us extra stability. We can instead split Langevin dynamics into

$$\begin{bmatrix} \mathrm{d}q \\ \mathrm{d}p \end{bmatrix} = \underbrace{\begin{bmatrix} p\,\mathrm{d}t \\ -q/\epsilon^2\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma/\beta}\,\mathrm{d}W \end{bmatrix}}_{\text{Fast}} + \underbrace{\begin{bmatrix} 0 \\ -q\,\mathrm{d}t \end{bmatrix}}_{\text{Slow}}, \tag{4.68}$$

and proceed as in the case of the rRESPA scheme. As before, we consider the case where we are able to solve the fast part exactly, though in practice a much smaller timestep is used as the fast parts are considered cheap, making the error in fast dynamics negligible in analysis anyway. Writing

$$F = \begin{bmatrix} 0 & 1 \\ -\epsilon^{-2} & -\gamma \end{bmatrix},$$

we can write down the flow map for the solution to the fast dynamics as

$$\Phi_{\text{Fast},t}\left(\begin{bmatrix} q \\ p \end{bmatrix}\right) = \exp(Ft)\begin{bmatrix} q \\ p \end{bmatrix} + B_t R_t,$$

where R_t is a column vector of two i.i.d. random normal numbers distributed as $\mathcal{N}(0,1)$, and

$$B_t B_t^T = \frac{2\gamma}{\beta}\int_0^t \exp(F(s-t))\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\exp(F^T(s-t))\mathrm{d}s.$$

Consider a method taking one step of the slow system, followed by a step of the fast

system, using fixed step size $\delta t > 0$. If we denote $z = [q, p]^T$ as the system state, then we can write this scheme as

$$z_{n+1} = Xz_n + YR_n,$$

for

$$X = \begin{bmatrix} 1 & 0 \\ -\delta t & 1 \end{bmatrix} \exp(\delta t A), \qquad Y = \begin{bmatrix} 1 & 0 \\ -\delta t & 1 \end{bmatrix} B_{\delta t}.$$

The stability of the method therefore depends solely on the spectral radius of the matrix X. Denoting $\omega = \sqrt{\gamma^2 - 4/\epsilon^2}$, then we compute

$$\mathrm{Det}(X) = e^{-\gamma \delta t}, \qquad \mathrm{Tr}(X) = \frac{\omega - \delta t}{\omega} e^{(\omega-\gamma)\delta t/2} + \frac{\omega + \delta t}{\omega} e^{(-\omega-\gamma)\delta t/2},$$

giving a stability condition

$$\left| \frac{\omega - \delta t}{\omega} e^{(\omega-\gamma)\delta t/2} + \frac{\omega + \delta t}{\omega} e^{(-\omega-\gamma)\delta t/2} \right| \leq 1 + e^{-\gamma \delta t}.$$

Multiplying by $e^{\gamma \delta t/2}$ gives

$$\left| \cosh\left(\frac{\omega \delta t}{2}\right) + \frac{\delta t}{\omega} \sinh\left(-\frac{\omega \delta t}{2}\right) \right| \leq \cosh\left(\frac{\gamma \delta t}{2}\right),$$

which reduces to the deterministic stability condition (2.41) in the limit $\gamma = 0$. For $\gamma > 0$, we can simplify to

$$\left| \cosh\left(\frac{\omega \delta t}{2} - \mathrm{atanh}\left(\frac{\delta t}{\omega}\right)\right) \sqrt{1 - \frac{\delta t^2}{\omega^2}} \right| \leq \cosh\left(\frac{\gamma \delta t}{2}\right).$$

We consider the case with $\gamma \ll \epsilon^{-1}$, allowing us to write $\omega = i\widehat{\omega}$ for $\widehat{\omega} \in \mathbb{R}$, and

$$\left| \cos\left(\frac{\widehat{\omega} \delta t}{2} - \mathrm{atanh}\left(\frac{\delta t}{\widehat{\omega}}\right)\right) \sqrt{1 + \frac{\delta t^2}{\widehat{\omega}^2}} \right| \leq \cosh\left(\frac{\gamma \delta t}{2}\right).$$

The term on the left hand side of the inequality is oscillatory in δt, while the term on the right hand side uniformly grows with step size. This implies that for any suitable $\gamma > 0$, we can find small pocket regions of the step size where the left hand side is sufficiently minimized in order to give stability. However, just as in the deterministic case, these island stability regions are difficult to predict the location of in a more general system.

The oscillatory cosine term is bounded between ± 1, implying that we have stability for all step sizes provided the condition

$$\sqrt{1 + \frac{\delta t^2}{\widehat{\omega}^2}} \leq \cosh\left(\frac{\gamma \delta t}{2}\right), \tag{4.69}$$

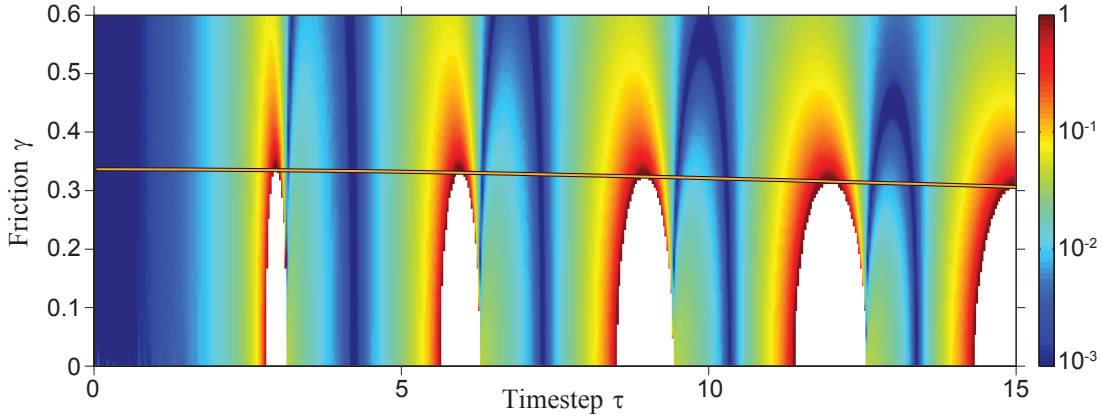


Figure 4.4: We repeat the experiment given in Figure 2.5 using Langevin dynamics, with potential energy function (4.67) choosing $\epsilon = 1/3$. Each pixel represents a simulation undertaken with parameters $(\delta t, \gamma)$, with the colour of a pixel indicating the absolute error in $\langle q^2 \rangle_{\delta t}$, with white pixels indicating instability. The plotted line gives the boundary to the region in parameter space where (4.69) is satisfied.

holds. We expect there to exist stability regions for smaller values of the friction where this condition is violated, but these will be sandwiched between unstable regions coming from the oscillatory term. We have guaranteed stability where (4.69) holds, which we can simplify to

$$\frac{\delta t}{\sqrt{4\epsilon^{-2} - \gamma^2}} \leq \sinh\left(\frac{\gamma \delta t}{2}\right).$$

For a fixed value of the friction, if we increase the step size the right hand side will grow exponentially compared to the linear growth on the left hand side, so it would seem that for large δt we actually need a smaller friction in order to to be stable. Hence the minimum required friction for stability can be found by considering the case of small timestep. This gives a bound

$$\frac{1}{\sqrt{4\epsilon^{-2} - \gamma^2}} \leq \frac{\gamma}{2},$$

which gives a stability condition for all step sizes δt as long as

$$\gamma \geq \frac{\sqrt{2 - 2\sqrt{1 - \epsilon^4}}}{\epsilon} \geq \sqrt{2}\epsilon.$$

Therefore we expect that the minimum friction value required for stability for all timesteps $\delta t > 0$, scales with the value of ϵ. Unlike the deterministic case, we are able to remove linear resonance effects with sufficiently large friction, where γ is an arbitrary parameter in our dynamics governing the strength of the noise.

We illustrate this numerically in Figure 4.4, by applying the splitting scheme (4.68) using various timesteps and frictions. Each gridpoint marks an independent simulation of 10^8 steps with the given parameter values, where the plotted line marks the boundary of the region where (4.69) is satisfied. We simulate using $\epsilon = 1/3$, making our results

for small friction commensurate with the deterministic case presented in Figure 2.5, which is the limiting case of zero friction.

The island regions of instability vanish as expected for large values of friction, while choosing $\gamma = \sqrt{2}\epsilon \approx 0.471$ in this example is plainly sufficient to give stability for all visible step sizes.

4.3 The overdamped regime

We will consider the behaviour of the elementary Langevin splitting methods discussed in the overdamped (infinite friction) limit of Langevin dynamics. This involves a resample of momenta from its equilibrium distribution κ_β every timestep, removing memory effects such as inertia. The movement of the system through space becomes purely diffusive, acting under Brownian dynamics, and evolving with respect to the governing SDE

$$\mathrm{d}q = -M^{-1}\nabla U(q) + \sqrt{2/\beta}M^{1/2}\mathrm{d}W, \tag{4.70}$$

with generator $\mathcal{L}_{\mathrm{ovd}}$ given through Itō's lemma as

$$\mathcal{L}_{\mathrm{ovd}} := -M^{-1}\nabla U(q)\cdot\nabla_q + \beta^{-1}\nabla_q\cdot(M\nabla_q).$$

Solutions to the so-called overdamped limit of Langevin dynamics, resulting from taking $\gamma \to \infty$, can be shown to converge to solutions to (4.70) with matching initial condition [41, 51]. As such, the dynamics is ergodic with respect to the configurational distribution μ_β.

The limit of infinite friction is known as the *Kramers to Smoluchowski* limit [51], where we refer to (4.70) as the Smoluchowski equation. In this limit, the Ornstein-Uhlenbeck update (the C step in the ABC splitting) becomes a resampling of momenta from its Gaussian distribution, with its evolution $\exp(\delta t\mathcal{L}_C)$ replaced by the operator π, where

$$(\pi\phi)(q) = \int_{\mathbb{R}^N}\phi(q,p)\rho_\beta\mathrm{d}p \tag{4.71}$$

for an observable $\phi(q,p)$. Note that as we would expect, step size (and indeed time itself) plays no role in this resampling step. Our interest in this section will be on the errors in schemes at the overdamped limit, and the transition between finite and infinite friction.

4.3.1 Error estimates in the limit of infinite friction

As we have solved the OU process exactly in the ABC splitting, we can simply plug $\gamma = \infty$ into the algorithms for any of the methods to derive an algorithm for integrating (4.70). We can rewrite the schemes compactly in one-line form, for example the

⟦CBABC⟧ scheme has the limiting method

$$q_{n+1} = q_n - \frac{\delta t^2}{2} M^{-1} \nabla U(q_n) + \frac{\delta t}{\sqrt{\beta}} R_n, \tag{4.72}$$

for $R_n \sim \mathcal{N}(0,1)$ a vector of N i.i.d normal random numbers. This is evidently the Euler-Maruyama scheme integrating (4.70), using an effective timestep $h = \delta t^2/2$ corresponding to a diffusive timescale [63]. In contrast, the limiting method for the ⟦BACAB⟧ scheme is

$$q_{n+1} = q_n - \frac{\delta t^2}{2} M^{-1} \nabla U(q_n) + \frac{\delta t}{2\sqrt{\beta}} \left(R_{n+1} + R_n \right), \tag{4.73}$$

where curiously (q_n) is no longer a Markov chain, due to the the correlation between successive steps (the noise process has a 'one-step memory'). We can recover the Markovian property by extending the space as necessary.

We may expect that the overdamped limit may serve to benefit some methods. For example, the results in equation (4.66) show that for large friction, when using the perturbed harmonic potential, the leading term of the error for the ⟦BACAB⟧ scheme is pushed to $\mathcal{O}(\delta t^4)$, despite it being a second order method. Of course, this could be a special case for the chosen observable or system.

Other methods are clearly unusable in this limit. The Stochastic Position Verlet method ⟦ASA⟧, or any method using the splitting (4.3) coupling the force together with the Ornstein-Uhlenbeck process cannot be used at large or infinite friction, as the drag term dominates the force term resulting in a random walk. It is not enough to separate the OU process from the force though: all of the first-order schemes considered in Section 4.1.3 are inconsistent as γ becomes large. For example, consider the limiting method of the scheme ⟦ABC⟧

$$q_{n+1} = q_n + \frac{\delta t}{\sqrt{\beta}} M^{1/2} R_n,$$

which is plainly not consistent for any step size as no information about $U(q)$ is used in the update. Any initial distribution will spread out evenly across the space in a random walk, evidently making the invariant distribution the uniform distribution in Ω.

Some second order schemes with exact OU solves are also inconsistent. For example the ⟦ACBCA⟧ and ⟦BCACB⟧ schemes both sandwich an A step between two C steps, which redraw the momentum before the position is updated: wiping out all force information.

Our focus in this section will be to examine the behaviour of averages of observables with respect to the remaining four elementary second-order discretization methods presented in Section 4.1.3, in the case of overdamped Langevin dynamics corresponding to infinite friction. If we could solve (4.51) for the perturbation functions f_2 then this would be trivial, as we would know exactly its behaviour for general potential energy

functions. Similarly in a formal sense, we expect the results of Table 4.1 and (4.66) to be consistent with our analysis in the limit of $\gamma \to \infty$.

We shall solely consider averages of functions of position in this section, since in the limiting case of infinite friction we can think of the system state wholly defined by instantaneous position. In order to simplify the analysis, we shall take $M = I_N$ Recall we denote by μ_β the canonical configurational distribution of position

$$\mu_\beta(q) = Z_{\beta,\mu}^{-1} \exp(-\beta U(q)), \qquad Z_{\beta,\mu}^{-1} = \int_{\mathcal{M}} \exp(-\beta U(q)) \mathrm{d}q,$$

where β is reciprocal temperature as usual. Numerical averages of sufficiently smooth observables $\phi = \phi(q)$ are taken with respect to the distribution

$$\langle \phi \rangle_{\delta t} = \int_{\mathcal{M}} \phi(q) \widehat{\mu}_{\delta t} \mathrm{d}q,$$

where we can think of $\widehat{\mu}_{\delta t}$ representing a method's corrupted configurational invariant distribution.

Let us consider the evolution operator for the overdamped methods, denoted $\mathcal{P}_{\infty,\delta t}$ and where the C operation is replaced by $\widetilde{C}$ in our notation to indicate that we are specifically considering the infinite friction case.

In the overdamped limit, a Langevin dynamics method of the form $[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]$ for $X, Y \in \{A, B\}$ distinct has evolution operator

$$\mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]} = \pi \, e^{\delta t \mathcal{L}_X/2} \, e^{\delta t \mathcal{L}_Y} \, e^{\delta t \mathcal{L}_X/2} \, \pi = \pi \, \mathcal{P}_{\delta t}^{[\![\mathrm{XYX}]\!]} \, \pi,$$

where $\mathcal{P}_{\delta t}^{[\![\mathrm{XYX}]\!]}$ is the evolution operator for the deterministic numerical method $[\![\mathrm{XYX}]\!]$, which corresponds to either position or velocity Verlet. Using (4.29), we can write

$$\mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]} = \pi + \delta t \pi \mathcal{L}_{\mathrm{mc}} \pi + \frac{\delta t^2}{2} \pi \mathcal{L}_{\mathrm{mc}}^2 \pi + \frac{\delta t^3}{6} \pi S_3 \pi + \frac{\delta t^4}{24} \pi S_4 \pi + \frac{\delta t^5}{120} \pi S_5 \pi + \delta t^6 \pi R_{6,\delta t} \pi, \tag{4.74}$$

using $\mathcal{L}_X + \mathcal{L}_Y = \mathcal{L}_A + \mathcal{L}_B = \mathcal{L}_{\mathrm{mc}}$. The S_j terms can be computed using the BCH formula, or explicitly through $S_j = \mathbf{T}\left[(L_1 + L_2 + L_3)^j\right]$, where $L_1 = L_3 = \mathcal{L}_X$ and $L_2 = \mathcal{L}_Y$. It is easily verified that $S_1 = \mathcal{L}_{\mathrm{mc}}$ and $S_2 = \mathcal{L}_{\mathrm{mc}}^2$ as they appear in (4.74).

In order to efficiently compute the right hand side of (4.74), we can make use of a number of observations about the interplay of the different operators. The action of the $\mathcal{L}_A$ and $\mathcal{L}_B$ operators respectively increases or decreases the order of the momentum terms by 1, while the action of the π operator averages the momentum terms with respect to a Gaussian weighting, leaving solely a function of position. Given that the odd moments of a Gaussian distribution are zero, we can exploit a parity argument in order to simplify terms. From their definitions, we can immediately see that for all $n > 0$

$$\pi f(q) = f(q), \qquad \pi^n = \pi, \qquad \mathcal{L}_B^n \pi = 0.$$

A consideration of the order of the momentum terms gives that for any $i, j, k \in \mathbb{N}$ where $i + j + k$ is odd,

$$\pi \mathcal{L}_B^i \mathcal{L}_A^j \mathcal{L}_B^k \pi = \pi \mathcal{L}_A^i \mathcal{L}_B^j \mathcal{L}_A^k \pi = 0.$$

Simple computation also yields

$$\pi \mathcal{L}_B \mathcal{L}_A \pi = -\nabla U \cdot \nabla_q \pi, \qquad \pi \mathcal{L}_A^2 \pi = \beta^{-1} \Delta_q \pi.$$

In order to simplify (4.74), we note also that terms $\pi S_n \pi$ will always be of the form $\pi \mathcal{L}_X^i \mathcal{L}_Y^j \mathcal{L}_X^k \pi$, where $i + j + k = n$ and $i, j, k \in \mathbb{N}$. Hence for any odd n, the $\pi S_n \pi$ term will be zero.

Applying these relations gives

$$\begin{aligned} \pi S_1 \pi &= \pi S_3 \pi = \pi S_5 \pi = 0, \\ \pi S_2 \pi &= \pi (\mathcal{L}_A + \mathcal{L}_B)^2 \pi = \pi (\mathcal{L}_A^2 + \mathcal{L}_B \mathcal{L}_A) \pi = \pi \mathcal{L}_{\text{ovd}} \pi. \end{aligned}$$

The term $\pi S_4 \pi$ requires a little more work to obtain. Using the symmetric BCH formula (2.35)

$$\mathcal{P}_{\delta t}^{[\![\text{XYX}]\!]} = \text{Id} \ + \delta t \widehat{\mathcal{L}} + \frac{\delta t^2}{2} \widehat{\mathcal{L}}^2 + \frac{\delta t^3}{6} \widehat{\mathcal{L}}^3 + \frac{\delta t^4}{24} \widehat{\mathcal{L}}^4 + \mathcal{O}(\delta t^5),$$

with

$$\widehat{\mathcal{L}} = \mathcal{L}_X + \mathcal{L}_Y + \frac{\delta t^2}{24} \left(2[\mathcal{L}_Y, [\mathcal{L}_Y, \mathcal{L}_X]] - [\mathcal{L}_X, [\mathcal{L}_X, \mathcal{L}_Y]] \right) + \mathcal{O}(\delta t^4).$$

Equating the powers of δt^4 gives

$$S_4 = \mathcal{L}_{\text{mc}}^4 + \frac{1}{2} \left(\mathcal{L}_{\text{mc}} [2\mathcal{L}_Y + \mathcal{L}_X, [\mathcal{L}_Y, \mathcal{L}_X]] + [2\mathcal{L}_Y + \mathcal{L}_X, [\mathcal{L}_Y, \mathcal{L}_X]] \mathcal{L}_{\text{mc}} \right).$$

We can simplify greatly by considering $\pi S_4 \pi$, as many of the terms will cancel. The only surviving terms are a linear combination of the operators

$$\begin{aligned} \pi \mathcal{L}_A^4 \pi \varphi &= 3\pi \mathcal{L}_A^2 \pi \mathcal{L}_A^2 \pi \varphi = 3\beta^{-2} \Delta^2 \varphi, \\ \pi \mathcal{L}_B \mathcal{L}_A^3 \pi \varphi &= 3\pi \mathcal{L}_B \mathcal{L}_A \pi \mathcal{L}_A^2 \pi \varphi = -3\beta \nabla U \cdot \nabla \left(\Delta \varphi \right), \\ \pi \mathcal{L}_A^2 \mathcal{L}_B \mathcal{L}_A \pi \varphi &= \pi \mathcal{L}_A^2 \pi \mathcal{L}_B \mathcal{L}_A \pi \varphi = -\beta^{-1} \left(2\nabla^2 U : \nabla^2 \varphi + \nabla U \cdot \nabla (\Delta \varphi) + \nabla (\Delta U) \cdot \nabla \varphi \right), \\ \pi \mathcal{L}_B \mathcal{L}_A \mathcal{L}_B \mathcal{L}_A \pi \varphi &= \pi \mathcal{L}_B \mathcal{L}_A \pi \mathcal{L}_B \mathcal{L}_A \pi \varphi = \nabla U \cdot \left(\nabla^2 \varphi \right) \nabla U + \nabla U \cdot \left(\nabla^2 U \right) \nabla \varphi, \\ \pi \mathcal{L}_A \mathcal{L}_B \mathcal{L}_A^2 \pi \varphi &= -2\beta^{-1} \left(\nabla^2 U : \nabla^2 \varphi + \nabla U \cdot \nabla (\Delta \varphi) \right), \\ \pi \mathcal{L}_B^2 \mathcal{L}_A^2 \pi \varphi &= 2 \nabla U \cdot \left(\nabla^2 \varphi \right) \nabla U. \end{aligned} \tag{4.75}$$

Here, as $\varphi = \varphi(q)$, we have denoted $\nabla = \nabla_q$ and $\Delta = \Delta_q = \nabla_q \cdot \nabla_q$. Using these terms,

it is easy to compute

$$\begin{aligned}
\pi\mathcal{L}_{\mathrm{mc}}^4\pi &= \pi\left(\mathcal{L}_A+\mathcal{L}_B\right)^2\left(\mathcal{L}_A+\mathcal{L}_B\right)^2\pi,\\
&= \pi\left(\mathcal{L}_A+\mathcal{L}_B\right)^2(\mathcal{L}_A^2+\mathcal{L}_B\mathcal{L}_A)\,\pi,\\
&= \pi\left(\mathcal{L}_A^4+\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A^2+\mathcal{L}_B\mathcal{L}_A^3+\mathcal{L}_B^2\mathcal{L}_A^2+\mathcal{L}_A^2\mathcal{L}_B\mathcal{L}_A+\mathcal{L}_B\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A\right)\pi,\\
&= 3\pi\left(\mathcal{L}_A^2+\mathcal{L}_B\mathcal{L}_A\right)\pi\left(\mathcal{L}_A^2+\mathcal{L}_B\mathcal{L}_A\right)\pi\\
&\quad+\pi\left(\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A^2+\mathcal{L}_B^2\mathcal{L}_A^2-2\mathcal{L}_A^2\mathcal{L}_B\mathcal{L}_A-2\mathcal{L}_B\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A\right)\pi,\\
&= 3\pi\mathcal{L}_{\mathrm{ovd}}^2\pi+\pi\left(\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A^2+\mathcal{L}_B^2\mathcal{L}_A^2-2\mathcal{L}_A^2\mathcal{L}_B\mathcal{L}_A-2\mathcal{L}_B\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A\right)\pi.
\end{aligned}$$

Then denoting the commutator term as

$$\Psi(\mathcal{L}_X,\mathcal{L}_Y)=\left(\mathcal{L}_X+\mathcal{L}_Y\right)\left[2\mathcal{L}_Y+\mathcal{L}_X,[\mathcal{L}_Y,\mathcal{L}_X]\right]+\left[2\mathcal{L}_Y+\mathcal{L}_X,[\mathcal{L}_Y,\mathcal{L}_X]\right]\left(\mathcal{L}_X+\mathcal{L}_Y\right),$$

we can write

$$\pi\Psi(\mathcal{L}_A,\mathcal{L}_B)\pi=\pi\left(\mathcal{L}_B^2\mathcal{L}_A^2+\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A^2+\mathcal{L}_A^2\mathcal{L}_B\mathcal{L}_A-2\mathcal{L}_B\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A-\mathcal{L}_B\mathcal{L}_A^3\right)\pi,$$

and

$$\pi\Psi(\mathcal{L}_B,\mathcal{L}_A)\pi=\pi\left(\mathcal{L}_B^2\mathcal{L}_A^2-2\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A^2-2\mathcal{L}_A^2\mathcal{L}_B\mathcal{L}_A-2\mathcal{L}_B\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A+2\mathcal{L}_B\mathcal{L}_A^3\right)\pi.$$

Hence we can write

$$\pi S_4\pi=3\pi\left(\mathcal{L}_{\mathrm{ovd}}^2+D\right)\pi,$$

where D is the method dependent operator

$$\begin{aligned}
D^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]} &= \frac{1}{2}\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A^2+\frac{1}{2}\mathcal{L}_B^2\mathcal{L}_A^2-\frac{1}{2}\mathcal{L}_A^2\mathcal{L}_B\mathcal{L}_A-\mathcal{L}_B\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A-\frac{1}{6}\mathcal{L}_B\mathcal{L}_A^3,\\
D^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} &= \frac{1}{2}\mathcal{L}_B^2\mathcal{L}_A^2-\mathcal{L}_A^2\mathcal{L}_B\mathcal{L}_A-\mathcal{L}_B\mathcal{L}_A\mathcal{L}_B\mathcal{L}_A+\frac{1}{3}\mathcal{L}_B\mathcal{L}_A^3.
\end{aligned}$$

It makes sense to denote $h:=\delta t^2/2$ as the Brownian timestep, giving

$$\mathcal{P}_{\infty,\delta t}=\pi+h\pi\mathcal{L}_{\mathrm{ovd}}\pi+\frac{h^2}{2}\pi\left(\mathcal{L}_{\mathrm{ovd}}^2+D\right)\pi+h^3R_{2,\infty,h}.$$

We can then proceed exactly as in the case of the first-order methods in Section 4.1.3, writing

$$\begin{aligned}
&\int_{\mathcal{M}}\left[\frac{\pi-\mathcal{P}_{\infty,\delta t}}{h}\phi\right]\left[\mu_\beta\left(1+hf_{2,\infty}\right)\right]\mathrm{d}q\\
&\qquad=\int_{\mathcal{M}}\left[\left(-\mathcal{L}_{\mathrm{ovd}}-\frac{h}{2}\left(\mathcal{L}_{\mathrm{ovd}}^2+D\right)-h^2R_{2,\infty,h}\right)\phi\right]\left[\mu_\beta\left(1+hf_{2,\infty}\right)\right]\mathrm{d}q\\
&\qquad=-h\int_{\mathcal{M}}\left(\mathcal{L}_{\mathrm{ovd}}\phi\left[\mu_\beta f_{2,\infty}\right]+\frac{1}{2}\left[D\phi\right]\mu_\beta\right)\mathrm{d}q+h^2\widetilde{R}_{2,\infty,h}.
\end{aligned}$$

By taking adjoints in the leading term, it is clear that $f_{2,\infty}$ should satisfy

$$\mathcal{L}_{\text{ovd}}^{*}\mu_\beta f_{2,\infty} = -\frac{1}{2}D^{*}\mu_\beta, \tag{4.76}$$

where D is method dependent with finitely many derivatives. Using

$$\int_{\mathcal{M}} \left(\mathcal{P}_{\infty,\delta t}\phi\right)\widehat{\mu}_{\delta t}\mathrm{d}q = \int_{\mathcal{M}} \phi\widehat{\mu}_{\delta t}\mathrm{d}q,$$

we can write

$$\begin{aligned}\int_{\mathcal{M}} \left[\frac{\pi - \mathcal{P}_{\infty,\delta t}}{h}\phi\right]\widehat{\mu}_{\delta t}\mathrm{d}q - \int_{\mathcal{M}} \left[\frac{\pi - \mathcal{P}_{\infty,\delta t}}{h}\phi\right]\left[\mu_\beta\left(1 + hf_{2,\infty}\right)\right]\mathrm{d}q \\ = h\int_{\mathcal{M}} \left(\mathcal{L}_{\text{ovd}}\phi\left[\mu_\beta f_{2,\infty}\right] + \frac{1}{2}\left[D\phi\right]\mu_\beta\right)\mathrm{d}q - h^2\widetilde{R}_{2,\infty,h},\end{aligned}$$

and hence choosing $f_{2,\infty}$ to solve (4.76) we have

$$\int_{\mathcal{M}} \left[\frac{\pi - \mathcal{P}_{\infty,\delta t}}{h}\phi\right]\widehat{\mu}_{\delta t}\mathrm{d}q = \int_{\mathcal{M}} \left[\frac{\pi - \mathcal{P}_{\infty,\delta t}}{h}\phi\right]\left[\mu_\beta\left(1 + hf_{2,\infty}\right)\right]\mathrm{d}q - h^2\widetilde{R}_{2,\infty,h},$$

Using the result of (4.39), we choose the approximate inverse

$$\mathcal{Q}_h = -\mathcal{L}_{\text{ovd}} + \frac{h}{2}\left(\text{Id } + \mathcal{L}_{\text{ovd}}^{-1}D\mathcal{L}_{\text{ovd}}^{-1}\right), \qquad \left(\frac{\pi - \mathcal{P}_{\infty,\delta t}}{h}\right)\mathcal{Q}_h = \text{Id } + h^2 Z_2,$$

and set $\phi = \mathcal{Q}_h\varphi$ to give

$$\boxed{\int_{\mathcal{M}} \varphi\widehat{\mu}_{\delta t}\mathrm{d}q = \int_{\mathcal{M}} \varphi\mu_\beta\mathrm{d}q + h\int_{\mathcal{M}} \varphi f_{2,\infty}\mu_\beta\mathrm{d}q + h^2 r_{2,\infty},}$$

for method dependent correction function $f_{2,\infty}$, and where $r_{2,\infty}$ in uniformly bounded by virtue of D having finitely many derivatives and the set $\mathcal{S} \cap \ker(\pi)$ being stable with respect to $\mathcal{L}_{\text{ovd}}^{-1}$ (by arguments similar to (4.28)). The simplification produced by moving to the overdamped limit (i.e. eliminating momentum) allows us to find explicit solutions to (4.76) and give formulas for $f_{2,\infty}$. Letting

$$\psi_1 = \beta^2\left|\nabla U\right|^2 - \beta\Delta U, \qquad \Delta\mu_\beta = \psi_1\mu_\beta,$$

we can compute

$$\begin{aligned}D^{[\![\widetilde{\text{C}}\text{ABA}\widetilde{\text{C}}]\!]} &= \frac{1}{2\beta}\nabla(\Delta U)\cdot\nabla - (\nabla U)\cdot(\nabla^2 U)\nabla, \\ &= \nabla\left(\frac{1}{2\beta}\Delta U - \frac{1}{2}\left|\nabla U\right|^2\right)\cdot\nabla, \\ &= -\frac{1}{2\beta^2}\nabla\psi_1\cdot\nabla.\end{aligned}$$

Taking the adjoint and operating on μ_β gives

$$\begin{aligned}\left(D^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]}\right)^* \mu_\beta &= \frac{1}{2\beta^2}\nabla\cdot(\mu_\beta\,\nabla\psi_1)\,,\\ &= \frac{1}{2\beta^2}\left(\nabla\mu_\beta\cdot\nabla\psi_1+\mu_\beta\,\Delta\psi_1\right),\\ &= \frac{1}{2\beta^2}\,\mu_\beta\left(-\beta\nabla U\cdot\nabla+\Delta\right)\psi_1,\\ &= \frac{1}{2\beta}\,\mu_\beta\mathcal{L}_{\mathrm{ovd}}\psi_1,\\ &= \frac{1}{2\beta}\mathcal{L}^*_{\mathrm{ovd}}\psi_1\mu_\beta.\end{aligned}$$

This gives the solution

$$f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]} = -\frac{1}{4\beta}\psi_1 = \frac{1}{4}\left(\Delta U-\beta\,|\nabla U|^2\right), \qquad f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]}\mu_\beta = -\frac{1}{4\beta}\Delta\mu_\beta. \tag{4.77}$$

From the definition of ψ_1, we can also write

$$\int_{\mathcal{M}}\varphi\, f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]}\mu_\beta\mathrm{d}q = -\frac{1}{4\beta}\int_{\mathcal{M}}\varphi\,\Delta\mu_\beta\mathrm{d}q = -\frac{1}{4\beta}\int_{\mathcal{M}}\Delta\varphi\mu_\beta\mathrm{d}q,$$

giving a more succinct correction

$$\int_{\mathcal{M}}\varphi\widehat{\mu}_{\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]}\mathrm{d}q = \int_{\mathcal{M}}\varphi\mu_\beta\mathrm{d}q - \frac{h}{4\beta}\int_{\mathcal{M}}\Delta\varphi\mu_\beta\mathrm{d}q + h^2\hat{R}_\infty.$$

Similarly

$$D^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} = \frac{2}{\beta}\nabla^2U:\nabla^2+\frac{1}{\beta}\nabla\Delta U\cdot\nabla-\nabla U\cdot\left(\nabla^2U\right)\nabla,$$

and defining

$$\psi_2 = \beta\Delta U-\frac{\beta^2}{2}\,|\nabla U|^2\,,$$

we can compute the adjoint operator as

$$\left(D^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]}\right)^*\mu_\beta = \frac{1}{\beta^2}\nabla\cdot(\mu_\beta\nabla\psi_2) = \frac{1}{\beta}\mathcal{L}^*_{\mathrm{ovd}}\psi_2\mu_\beta.$$

Hence we obtain the solution

$$f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} = c_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]}-\frac{1}{2\beta}\psi_2 = c_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]}-\frac{1}{2}\left(\Delta U-\frac{\beta}{2}\,|\nabla U|^2\right),$$

where the constant $c_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]}$ is chosen so that $\left\langle f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]}\right\rangle = 0$ as required. As $\left\langle f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]}\right\rangle = 0$, it is easy to see that

$$c_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} = \frac{1}{2\beta}\langle\psi_2\rangle = \left\langle\frac{\Delta U}{4}\right\rangle.$$

The correction can be written as

$$\int_{\mathcal{M}} \varphi \widehat{\mu}_{\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} \mathrm{d}q = \int_{\mathcal{M}} \varphi \mu_\beta \mathrm{d}q + \frac{h}{4\beta} \int_{\mathcal{M}} \left(\Delta\varphi + \beta\left(1-\varphi\right)\Delta U\right) \mu_\beta \mathrm{d}q + h^2 \hat{R}_\infty.$$

The remaining limiting solutions can be computed using Lemma 4.30. We set

$$\mathcal{U}_{\infty,\delta t} = \pi e^{\delta t \mathcal{L}_B/2} e^{\delta t \mathcal{L}_A/2}, \quad \mathcal{T}_{\infty,\delta t} = e^{\delta t \mathcal{L}_A/2} e^{\delta t \mathcal{L}_B/2} \pi$$

so that

$$\mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} = \mathcal{U}_{\infty,\delta t}\mathcal{T}_{\infty,\delta t}, \qquad \mathcal{P}_{\infty,\delta t}^{[\![\mathrm{AB}\widetilde{\mathrm{C}}\mathrm{BA}]\!]} = \mathcal{T}_{\infty,\delta t}\mathcal{U}_{\infty,\delta t}.$$

Using the BCH formula,

$$\mathcal{U}_{\infty,\delta t} = \pi\left(\mathrm{Id} + \frac{\delta t}{2}\left(\mathcal{L}_A + \mathcal{L}_B\right) + \frac{\delta t^2}{8} G_2 + \frac{\delta t^3}{4} 8 G_3 + \delta t^4 R_{U,4}\right), \tag{4.78}$$

where

$$\pi\left(\mathcal{L}_A + \mathcal{L}_B\right)\pi = 0, \qquad \pi G_2 \pi = \pi\left(\left(\mathcal{L}_A + \mathcal{L}_B\right)^2 + \left[\mathcal{L}_B, \mathcal{L}_A\right]\right)\pi = \mathcal{L}_{\mathrm{ovd}} + \pi\mathcal{L}_B\mathcal{L}_A\pi,$$

using the above rules. The G_3 term will be a linear combination of terms $\mathcal{L}_B^i \mathcal{L}_A^j$, where $i, j \in \mathbb{N}$ and $i + j = 3$. Hence, by a parity argument $\pi G_3 \pi = 0$. Then applying Lemma 4.30, with $\varphi = \pi\varphi$, we have

$$\begin{aligned}\int_{\mathcal{M}} \varphi \widehat{\mu}_{\delta t}^{[\![\mathrm{AB}\widetilde{\mathrm{C}}\mathrm{BA}]\!]} \mathrm{d}q &= \int_{\mathcal{M}} \left(\mathcal{U}_{\infty,\delta t}\varphi\right) \widehat{\mu}_{\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} \mathrm{d}q \\ &= \int_{\mathcal{M}} \left(\mathcal{U}_{\infty,\delta t}\varphi\right) \left[\mu_\beta (1 + h f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]})\right] \mathrm{d}q + h^2 \widehat{r}_{2,h},\end{aligned}$$

and equating powers of h we have

$$f_{2,\infty}^{[\![\mathrm{AB}\widetilde{\mathrm{C}}\mathrm{BA}]\!]} \mu_\beta = f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} \mu_\beta + \frac{1}{4}\left(\pi G_2 \pi\right)^* \mu_\beta.$$

Computing

$$\left(\pi\mathcal{L}_B\mathcal{L}_A\pi\right)^* \mu_\beta = \mu_\beta\left(\Delta U - \beta\left|U\right|^2\right),$$

gives

$$f_{2,\infty}^{[\![\mathrm{AB}\widetilde{\mathrm{C}}\mathrm{BA}]\!]} = f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} + \frac{1}{4}\left(\Delta U - \beta\left|U\right|^2\right) = \left\langle \frac{\Delta U}{4} \right\rangle - \frac{\Delta U}{4}.$$

Alternately, setting

$$\mathcal{U}_{\infty,\delta t} = \pi e^{\delta t \mathcal{L}_A/2} e^{\delta t \mathcal{L}_B/2}, \quad \mathcal{T}_{\infty,\delta t} = e^{\delta t \mathcal{L}_B/2} e^{\delta t \mathcal{L}_A/2} \pi,$$

gives

$$\mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]} = \mathcal{U}_{\infty,\delta t}\mathcal{T}_{\infty,\delta t}, \qquad \mathcal{P}_{\infty,\delta t}^{[\![\mathrm{BA}\widetilde{\mathrm{C}}\mathrm{AB}]\!]} = \mathcal{T}_{\infty,\delta t}\mathcal{U}_{\infty,\delta t}.$$

We can then proceed by exchanging the $\mathcal{L}_A$ and $\mathcal{L}_B$ symbols in G_2, to give

$$\begin{aligned} f_{2,\infty}^{[\![\mathrm{BA}\widetilde{\mathrm{C}}\mathrm{AB}]\!]}\mu_\beta &= f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]}\mu_\beta + \frac{1}{4}\left(\pi\left((\mathcal{L}_A+\mathcal{L}_B)^2 + [\mathcal{L}_A,\mathcal{L}_B]\right)\pi\right)^*\mu_\beta, \\ &= f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]}\mu_\beta + \frac{1}{4}\left(\pi\mathcal{L}_A^2\pi\right)^*\mu_\beta, \\ &= f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]}\mu_\beta + \frac{1}{4\beta}\Delta\mu_\beta, \\ &= 0, \end{aligned} \tag{4.79}$$

by (4.77). This is perhaps a surprising result, as it implies that the scheme (4.73) is second-order in h, or equivalently fourth-order in δt, for any potential energy function $U(q)$ and any smooth observable φ. Schemes have been previously been constructed for Langevin dynamics with order > 2 [86, 114], but these require multiple evaluations of the force; for this reason they are not normally viewed as competitive alternatives for molecular sampling [65]. By contrast, the ⟦BACAB⟧ scheme (like all Langevin splitting schemes considered in this thesis) requires only one force-evaluation per timestep.

In summary, we give the correction terms for all four of the limiting schemes as

$$\begin{aligned} f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{ABA}\widetilde{\mathrm{C}}]\!]} &= \frac{1}{4}\left(\Delta U - \beta\left|\nabla U\right|^2\right), & f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{BAB}\widetilde{\mathrm{C}}]\!]} &= c - \frac{1}{2}\left(\Delta U - \frac{\beta}{2}\left|\nabla U\right|^2\right), \\ f_{2,\infty}^{[\![\mathrm{BA}\widetilde{\mathrm{C}}\mathrm{AB}]\!]} &= 0, & f_{2,\infty}^{[\![\mathrm{AB}\widetilde{\mathrm{C}}\mathrm{BA}]\!]} &= c - \Delta U/4, \end{aligned} \tag{4.80}$$

where $c = \langle \Delta U/4 \rangle$. It can be verified that the above correction terms match those given for the perturbed harmonic oscillator in (4.65) for $q \in \mathbb{R}$ and using the observable $\varphi(q) = q^2$ with potential energy function

$$U(q) = q^2/2 + \epsilon q^4/4,$$

when taking the limit $\gamma \to \infty$, and truncating at $\mathcal{O}(\epsilon^3)$.

4.3.2 Large and finite friction

in this section we consider the behaviour of the composition schemes as we approach the limit of infinite friction. We shall make use of the following lemma that allows us to bound the inverse of the Langevin dynamics operator for large friction, which gives us a step towards bounding remainder terms of our solution in this limit.

Lemma 4.81. *For any $\varphi \in \mathcal{H}^1_\perp$, where*

$$\mathcal{H}^1_\perp = \left\{ \varphi \in \mathcal{H}^1 \;\middle|\; (\pi\varphi)(q) \text{ is constant} \right\} \subset \mathcal{H}^1,$$

there exists a constant $K > 0$ such that for any $\gamma \geq 1$

$$\|\mathcal{L}_{LD}^{-1}\varphi\|_{H^1(\rho_\beta)} \leq K\|f\|_{H^1(\rho_\beta)}.$$

Proof. Defining

$$\mathcal{L}_1 = \mathcal{L}_A + \mathcal{L}_B + \mathcal{L}_{\bar{C}},$$

using $\gamma\mathcal{L}_{\bar{C}} = \mathcal{L}_C$ we have

$$\mathcal{L}_{\mathrm{LD}} = \mathcal{L}_1 + (\gamma - 1)\mathcal{L}_{\bar{C}}.$$

We will make use of a result in [51], showing that for all smooth functions $u \in \mathcal{H}^1$ there exists a $\theta > 0$ such that

$$-\langle\langle u, \mathcal{L}_1 u\rangle\rangle \geq \theta\langle\langle u, u\rangle\rangle,$$

where the bilinear form is defined as

$$\langle\langle u, v\rangle\rangle = a\langle u, v\rangle + b\langle\nabla_q u, \nabla_q v\rangle + b\langle\nabla_p u, \nabla_p v\rangle - \langle\nabla_p u, \nabla_q v\rangle - \langle\nabla_q u, \nabla_p v\rangle,$$

for appropriate constants $a \gg b \gg 1$. The norm defined by this bilinear form is equivalent to the norm defined on $H^1(\rho_\beta)$.

Thus we have

$$-\langle\langle u, \mathcal{L}_{\mathrm{LD}} u\rangle\rangle \geq \theta\langle\langle u, u\rangle\rangle - (\gamma - 1)\langle\langle u, \mathcal{L}_{\bar{C}} u\rangle\rangle. \tag{4.82}$$

Denoting the i^{th} component of p as p_i, can write

$$\mathcal{L}_{\bar{C}} = -p \cdot \nabla_p + \beta^{-1}\Delta p = -\frac{1}{\beta}\sum_{i=1}^{N} \partial_{p_i}^* \partial_{p_i}, \tag{4.83}$$

where $\partial_{p_i} = \partial/\partial p_i$, and $\partial_{p_i}^* = -\partial_{p_i} + \beta p_i$ is the adjoint map of ∂_{p_i} in $L^2(\rho_\beta)$. Hence direct computation yields

$$\begin{aligned}
\langle\langle u, \partial_{p_i}^* \partial_{p_i} u\rangle\rangle &= a\|\partial_{p_i} u\|^2 + b\|\nabla_q \partial_{p_i} u\|^2 + b\|\nabla_p \partial_i u\|^2 + b\beta\|\partial_{p_i} u\|^2 \\
&\quad - 2\langle\langle\nabla_q \partial_{p_i} u, \nabla_p \partial_{p_i} u\rangle\rangle - \beta\langle\langle\partial_{q_i} u, \partial_{p_i} u\rangle\rangle, \\
&\geq \left(a + \beta\left(b - \frac{1}{2}\right)\right)\|\partial_{p_i} u\|^2 + (b-1)\|\nabla_q \partial_{p_i} u\|^2 + (b-1)\|\nabla_p \partial_i u\|^2 \\
&\quad - \frac{\beta}{2}\|\partial_{q_i} u\|^2,
\end{aligned}$$

using the identities

$$\begin{aligned}
-2\langle\langle\nabla_q \partial_{p_i} u, \nabla_p \partial_{p_i} u\rangle\rangle &= \|(\nabla_q - \nabla_p)\partial_{p_i} u\|^2 - \|\nabla_q \partial_{p_i} u\|^2 - \|\nabla_p \partial_{p_i} u\|^2, \\
-\beta\langle\langle\partial_{q_i} u, \partial_{p_i} u\rangle\rangle &= \frac{\beta}{2}\|(\partial_{q_i} - \partial_{p_i})u\|^2 - \frac{\beta}{2}\|\partial_{q_i} u\|^2 - \frac{\beta}{2}\|\partial_{p_i} u\|^2.
\end{aligned}$$

We can bound the remaining negative term as the Gaussian momentum measure $\kappa_\beta(p)$ satisfies a Poincaré inequality, and hence there exists a constant $K' > 0$ such that for all $i = 1, \ldots, N$ we have

$$\|\partial_{q_i} u\| \leq K'\|\nabla_p \partial_{q_i} u\|,$$

by virtue of

$$\pi\partial_{q_i} u = \partial_{q_i}\pi u = 0,$$

for all $u \in \mathcal{H}^1_\perp$. Thus, for appropriate choices of a and b, we have

$$\langle\langle u, \partial^*_{p_i}\partial_{p_i} u\rangle\rangle > 0,$$

and hence combining this result with (4.82) and (4.83), we can deduce there exists some $K > 0$ such that

$$\forall \gamma \geq 1, \quad \|u\|_{H^1(\rho_\beta)} \leq K\|\mathcal{L}_{\mathrm{LD}}\|_{H^1(\rho_\beta)},$$

for any $u \in \mathcal{H}^1_\perp \cap H^2(\rho_\beta)$. Taking the inverse gives us the desired result. □

For sufficiently large finite friction γ, we would expect the behaviour of computed averages to approach the limiting behaviour of the overdamped scheme. In order to make this behaviour precise, we give a bound on the difference between the underdamped and overdamped Ornstein-Uhlenbeck evolution operators.

Lemma 4.84. *For any fixed $S \in \mathbb{N}^*$, there exists constants $K, \alpha > 0$ such that for all $1 \leq s \leq S$ and $\delta t > 0$*

$$\left\| e^{\delta t \mathcal{L}_C} - \pi \right\|_{\mathcal{B}(\mathcal{L}^\infty_{\mathcal{K}_s})} \leq K e^{-\alpha\gamma\delta t}. \tag{4.85}$$

Proof. The result follows by, for instance, an application of [98, Theorem 8.7] when considering the Ornstein-Uhlenbeck process with unit friction and mass $M = \mathrm{Id}$

$$\mathrm{d}p = -p\,\mathrm{d}t + \beta^{-1}\,\mathrm{d}W,$$

with generator

$$\mathcal{L}_{\bar{C}} = -p \cdot \nabla_p + \beta^{-1}\Delta_p.$$

The action of the operator on the Lyapunov function $\mathcal{K}_s = 1 + (p \cdot p)^s$ is such that for

$$\mathcal{L}_{\bar{C}}\mathcal{K}_s = -2s\,(p \cdot p)^s + \beta^{-1}\left(2Ns\,(p \cdot p)^{s-1} + 4s(s-1)\,(p \cdot p)^{s-1}\right) \leq -\mathcal{K}_s + \alpha_s,$$

where $\alpha_s \geq 0$ is an appropriate constant. Thus we have constants $K, \lambda > 0$ such that for any $t > 0$

$$\left| e^{t\mathcal{L}_{\bar{C}}}\psi(p) - \int_{\mathbb{R}^N} \psi(p)\kappa_\beta(p)\,\mathrm{d}p \right| \leq K e^{-\lambda t}\mathcal{K}_s(p)||\psi||_{L^\infty_{\mathcal{K}_s}},$$

giving exponential convergence towards its average. Plugging in $t = \gamma\delta t$ and noting that $\mathcal{L}_C = \gamma\mathcal{L}_{\bar{C}}$ gives (4.85) by applying the bound to any function $\varphi(q,p) \in L^\infty_{\mathcal{K}_s}$ and taking the supremum with respect to q. □

We can apply this lemma to bound the differences between the evolution of the Ornstein-Uhlenbeck process in a scheme at finite and infinite friction. Consider the

over and underdamped evolution operators of a method $[\![\mathrm{CXYXC}]\!]$, for $X, Y \in \{A, B\}$ distinct. Writing

$$\mathcal{P}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} - \mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]} = \left(e^{\delta t \mathcal{L}_C/2} - \pi\right) \mathcal{P}_{\delta t}^{[\![\mathrm{XYX}]\!]} \pi + e^{\delta t \mathcal{L}_C/2} \mathcal{P}_{\delta t}^{[\![\mathrm{XYX}]\!]} \left(e^{\delta t \mathcal{L}_C/2} - \pi\right), \tag{4.86}$$

allows us to bound their difference. This implies that for a smooth function solely of position $\phi(q)$, we have

$$\int_\Omega \left(\mathrm{Id} - \mathcal{P}_{\delta t}^{[\![\mathrm{CXYXC}]\!]}\right) \phi \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z = \int_\Omega \left(\mathrm{Id} - \mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]}\right) \phi \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z + r_1 = 0,$$

where the remainder

$$r_1 = \int_\Omega \left(\mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]} - \mathcal{P}_{\delta t}^{[\![\mathrm{CXYXC}]\!]}\right) \phi \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z,$$

can be bounded through an application of Lemma 4.84 with (4.86), for δt sufficiently small. Choosing a function $f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]}$ such that (4.76) is satisfied, we have

$$\int_\Omega \left(\mathrm{Id} - \mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]}\right) \phi \left[\rho_\beta \left(1 + \frac{\delta t^2}{2} f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]}\right)\right] \mathrm{d}z = \delta t^4 r_{2,\infty},$$

by virtue of $\rho_\beta = \kappa_\beta(p) \times \mu_\beta(q)$. Hence

$$\begin{aligned}\int_\Omega \left(\mathrm{Id} - \mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]}\right) \phi \left[\rho_\beta \left(1 + \frac{\delta t^2}{2} f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]}\right)\right] \mathrm{d}z \\ - \int_\Omega \left(\mathrm{Id} - \mathcal{P}_{\infty,\delta t}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]}\right) \phi \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z = r_1 + \delta t^4 r_{2,\infty}.\end{aligned}$$

Through the appropriate approximate inverse $\mathcal{Q}_{\delta t}$, and replacing $\phi = -\mathcal{Q}_{\delta t}\varphi$ as in the underdamped case, we can then write

$$\int_\Omega \varphi \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z = \int_\Omega \varphi \rho_\beta \mathrm{d}z + \frac{\delta t^2}{2} \int_\Omega \varphi f_{2,\infty}^{[\![\widetilde{\mathrm{C}}\mathrm{XYX}\widetilde{\mathrm{C}}]\!]} \rho_\beta \mathrm{d}z + \delta t^4 \widehat{r}_{2,\infty} + r_1, \tag{4.87}$$

where the r_1 can be bounded by terms exponentially small in $\gamma \delta t$, while $\widehat{r}_{2,\infty}$ is uniformly bounded for sufficiently small step size. Hence for sufficiently large frictions, we expect the leading δt^2 error term for the underdamped schemes of the form $[\![\mathrm{CXYXC}]\!]$, to be the same as the leading error term in their corresponding overdamped schemes.

It remains to consider the two schemes with the central Ornstein-Uhlenbeck update, where we used the TU lemma to compute the $f_{2,\infty}$ terms. Let us consider the $[\![\mathrm{YXCXY}]\!]$ scheme, again with $X, Y \in \{A, B\}$ distinct. We choose

$$\mathcal{T}_{\delta t} = e^{\delta t \mathcal{L}_Y/2} e^{\delta t \mathcal{L}_X/2} e^{\delta t \mathcal{L}_C/2}, \quad \mathcal{U}_{\delta t} = e^{\delta t \mathcal{L}_C/2} e^{\delta t \mathcal{L}_X/2} e^{\delta t \mathcal{L}_Y/2}, \quad \mathcal{U}_{\infty,\delta t} = \pi e^{\delta t \mathcal{L}_X/2} e^{\delta t \mathcal{L}_Y/2},$$

then by application of lemma 4.30, we have

$$\begin{aligned}\int_\Omega \varphi \widehat{\rho}_{\delta t}^{[\![\mathrm{YXCXY}]\!]} \mathrm{d}z &= \int_\Omega \left(\mathcal{U}_{\delta t}\varphi\right) \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z, \\ &= \int_\Omega \left(\mathcal{U}_{\infty,\delta t}\varphi\right) \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z + \int_\Omega \left(\left[\mathcal{U}_{\delta t} - \mathcal{U}_{\infty,\delta t}\right]\varphi\right) \widehat{\rho}_{\delta t}^{[\![\mathrm{CXYXC}]\!]} \mathrm{d}z.\end{aligned}$$

The second integral on the right hand side can be bounded by an exponentially small term through Lemma 4.84, while the first term is precisely of the form (4.87). Expanding the operator as in (4.78), we find an identical result to the overdamped case. Thus for $\gamma \geq 1$ and sufficiently small δt, for any of the symmetric second order schemes that are consistent at large friction, for sufficiently smooth observable $\varphi(q)$ we have

$$\int_\Omega \varphi(q)\widehat{\rho}_{\delta t} = \int_\Omega \varphi(q)\rho_\beta + h\int_\Omega \varphi(q) f_{2,\infty}(q)\rho_\beta + R,$$

where $f_{2,\infty}$ is the corresponding overdamped correction given in (4.80), with $h = \delta t^2/2$ and

$$|R| \leq K_1 h^2 + K_2 e^{-\widehat{\lambda}\gamma\delta t},$$

for appropriate constants $K_1, K_2, \widehat{\lambda} > 0$. Thus we have the same result for schemes with a central Ornstein-Uhlenbeck update, and schemes with two Ornstein-Uhlenbeck steps book-ending the method.

This implies that for sufficiently large $\gamma\delta t$, the ⟦BACAB⟧ scheme (still in the underdamped regime) gives an effective fourth-order result for averages of any configurational observables. Note that the restriction on φ to be solely a function of position is of little consequence, as we have already described that sampling the momentum is trivial and irrelevant in practical tasks. However, the price paid for this superconvergence property may be too steep: the increasing the friction to make $\gamma\delta t$ sufficiently large may have an adverse effect on the rate of exploration through phase space. As a consequence, it may not always be optimal (when sampling with Langevin dynamics) to blindly increase γ to improve the accuracy.

Let us illustrate the convergence of the correction function from f_2 to $f_{2,\infty}$ as γ increases. We generalize the form of the correction function's governing PDE for each method in (4.51) by considering u as the solution to the Poisson equation

$$\mathcal{L}_{\mathrm{LD}} u(q,p) = W(q,p), \tag{4.88}$$

where $W(q,p)$ is independent of the friction γ. In (4.51), the form of $W(q,p)$ is determined by the method we consider and the solution u is the second order correction function. We make the ansatz that we can write

$$u = \gamma u_{-1} + u_0 + \frac{1}{\gamma}u_1 + \frac{1}{\gamma^2}u_2 + \dots$$

and seek to rigorously ground this assumption. From (4.88), we have

$$(\mathcal{L}_A + \mathcal{L}_B)u + \gamma \mathcal{L}_{\bar{C}} u = W(q,p),$$

and plugging in our ansatz and equating powers of the friction gives

$$\begin{aligned}
\mathcal{L}_{\bar{C}} u_{-1} &= 0, \\
(\mathcal{L}_A + \mathcal{L}_B)u_{-1} + \mathcal{L}_{\bar{C}} u_0 &= 0, \\
(\mathcal{L}_A + \mathcal{L}_B)u_0 + \mathcal{L}_{\bar{C}} u_1 &= W, \\
(\mathcal{L}_A + \mathcal{L}_B)u_i + \mathcal{L}_{\bar{C}} u_{i+1} &= 0, \qquad \text{for } i \in \mathbb{N}^*.
\end{aligned}$$

The first equation implies $u_{-1} = u_{-1}(q)$, and hence $\mathcal{L}_B u_{-1} = 0$. The second equation then becomes

$$\mathcal{L}_{\bar{C}} u_0 = -\mathcal{L}_A u_{-1} = -p \cdot \nabla_q u_{-1},$$

with general solution

$$u_0 = p \cdot \nabla_q u_{-1} + \widetilde{u}_0(q),$$

for some function of position $\widetilde{u}_0(q)$. Plugging this solution into the third equation gives the formula

$$\mathcal{L}_{\bar{C}} u_1 = W - (\mathcal{L}_A + \mathcal{L}_B)u_0 = W + \nabla U(q) \cdot \nabla_q u_{-1} - p^T(\nabla^2 u_{-1})p - p \cdot \nabla_q \widetilde{u}_0, \quad (4.89)$$

which has solution only when the right hand side has average 0 with respect to the distribution κ_β, or equivalently when it is in the kernel of π. We therefore need $\pi W = \pi(\mathcal{L}_A + \mathcal{L}_B)u_0$, and hence

$$\pi W = \frac{1}{\beta}\Delta u_{-1} - \nabla U(q) \cdot \nabla_q u_{-1} = \mathcal{L}_{\text{ovd}} u_{-1},$$

giving $u_{-1} = \mathcal{L}_{\text{ovd}}^{-1} \pi W$. Substituting this into (4.89) gives

$$\mathcal{L}_{\bar{C}} u_1 = W - \pi W + \frac{1}{\beta}\Delta u_{-1} - p^T(\nabla \nabla^T u_{-1})p - p \cdot \nabla_q \widetilde{u}_0,$$

and using the identity

$$\mathcal{L}_{\bar{C}}\left(p \cdot Xp + p \cdot Y\right) = -p \cdot (X + X^T)p + \frac{2}{\beta}\text{Tr}(X) - p \cdot Y,$$

for matrix X and vector Y, we choose $X = X^T = \nabla^2 u_{-1}/2$ and $Y = \nabla_q \widetilde{u}_0$ to give

$$\mathcal{L}_{\bar{C}} u_1 = W - \pi W + \mathcal{L}_{\bar{C}}(p^T(\nabla^2 u_{-1})p/2 + p \cdot \nabla_q \widetilde{u}_0),$$

and hence

$$u_1 = \mathcal{L}_{\bar{C}}^{-1}(W - \pi W) + p^T(\nabla^2 u_{-1})p/2 + p \cdot \nabla_q \widetilde{u}_0 + \widetilde{u}_1(q),$$

for some arbitrary function of position $\widetilde{u}_1$. The next equation gives

$$\mathcal{L}_{\bar{C}} u_2 = -(\mathcal{L}_A + \mathcal{L}_B) u_1,$$

with solvability condition $\pi(\mathcal{L}_A + \mathcal{L}_B)u_1 = 0$, which amounts to the choice of

$$\pi(\mathcal{L}_A + \mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1}(W - \pi W) = -\pi(\mathcal{L}_A + \mathcal{L}_B)p \cdot \nabla_q \widetilde{u}_0 = -\mathcal{L}_{\mathrm{ovd}}\widetilde{u}_0,$$

giving $\widetilde{u}_0 = -\mathcal{L}_{\mathrm{ovd}}^{-1}\pi(\mathcal{L}_A + \mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1}(W - \pi W)$. Hence we have that

$$\mathcal{L}_{\mathrm{LD}}\left(u - \gamma u_{-1} - u_0 - \frac{1}{\gamma}u_1\right) = -\frac{1}{\gamma}(\mathcal{L}_A + \mathcal{L}_B)u_1.$$

Using Lemma 4.81, for $\gamma \geq 1$ there exists a constant $K > 0$ such that for any $v \in \mathcal{H}_\perp^1$ we have

$$\left\|\mathcal{L}_{\mathrm{LD}}^{-1} v\right\|_{H^1(\rho_\beta)} \leq K \|v\|_{H^1(\rho_\beta)}. \tag{4.90}$$

Setting $v = -(\mathcal{L}_A + \mathcal{L}_B)u_1/\gamma$, we obtain

$$\left\| u - \gamma u_{-1} - u_0 - \frac{1}{\gamma}u_1 \right\|_{H^1(\rho_\beta)} \leq \frac{K}{\gamma} \left\|(\mathcal{L}_A + \mathcal{L}_B)u_1\right\|_{H^1(\rho_\beta)}$$

and hence there exists some constant $\widetilde{K} > 0$ such that for any $\gamma \geq 1$ we have

$$\|u - \gamma u_{-1} - u_0\|_{H^1(\rho_\beta)} \leq \frac{\widetilde{K}}{\gamma},$$

for our choices of u_{-1} and u_0, giving

$$\left\|\mathcal{L}_{\mathrm{LD}}^{-1} W - \gamma \mathcal{L}_{\mathrm{ovd}}^{-1}\pi W - p \cdot \nabla_q \mathcal{L}_{\mathrm{ovd}}^{-1}\pi W - \mathcal{L}_{\mathrm{ovd}}^{-1}\pi(\mathcal{L}_A + \mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1}(W - \pi W)\right\|_{H^1(\rho_\beta)} \leq \frac{\widetilde{K}}{\gamma}. \tag{4.91}$$

Returning to methods denoted by ⟦CXYXC⟧, using (4.51) we can make use of the identity $\mathcal{L}_{\mathrm{LD}}^* f_2 \rho_\beta = -\rho_\beta \mathcal{L}_{\mathrm{LD}} f_2$ to find

$$\mathcal{L}_{\mathrm{LD}} f_2 = w,$$

where $w\rho_\beta = S_2^* \rho_\beta$, such that w is independent of the friction γ and computed from the right hand sides of (4.54–4.55). By the invariance $(\mathcal{L}_A^* + \mathcal{L}_B^*)\rho_\beta = 0$, we have

$$\begin{aligned}
[\mathcal{L}_A^*, \mathcal{L}_B^*]\, g\rho_\beta &= (\mathcal{L}_A^* + \mathcal{L}_B^*)(g^2\rho_\beta) - g\rho_\beta(\mathcal{L}_A^* + \mathcal{L}_B^*)g = g\rho_\beta(\mathcal{L}_A^* + \mathcal{L}_B^*)g, \\
\left(\mathcal{L}_A^* + \frac{1}{2}\mathcal{L}_B^*\right) g\rho_\beta &= \rho_\beta \left(\mathcal{L}_A^* + \frac{1}{2}\mathcal{L}_B^*\right) g - \frac{1}{2}g^2\rho_\beta, \\
\left(\frac{1}{2}\mathcal{L}_A^* + \mathcal{L}_B^*\right) g\rho_\beta &= \rho_\beta \left(\frac{1}{2}\mathcal{L}_A^* + \mathcal{L}_B^*\right) g + \frac{1}{2}g^2\rho_\beta,
\end{aligned}$$

where recall $g(q,p)\rho_\beta = \mathcal{L}_B^*\rho_\beta = -\beta p\cdot\nabla U\rho_\beta$, and hence

$$(\mathcal{L}_A^*+\mathcal{L}_B^*)\left(\mathcal{L}_A^*+\frac{1}{2}\mathcal{L}_B^*\right)g\rho_\beta + [\mathcal{L}_A^*,\mathcal{L}_B^*]\,g\rho_\beta = \rho_\beta\,(\mathcal{L}_A^*+\mathcal{L}_B^*)\left(\mathcal{L}_A^*+\frac{1}{2}\mathcal{L}_B^*\right)g,$$
$$(\mathcal{L}_A^*+\mathcal{L}_B^*)\left(\frac{1}{2}\mathcal{L}_A^*+\mathcal{L}_B^*\right)g\rho_\beta - [\mathcal{L}_A^*,\mathcal{L}_B^*]\,g\rho_\beta = \rho_\beta\,(\mathcal{L}_A^*+\mathcal{L}_B^*)\left(\frac{1}{2}\mathcal{L}_A^*+\mathcal{L}_B^*\right)g.$$

We can compute

$$\left(\mathcal{L}_A^*+\frac{1}{2}\mathcal{L}_B^*\right)g = \left(-p\cdot\nabla_q+\frac{1}{2}\nabla U\cdot\nabla_p\right)g = \beta\left(p^T\left(\nabla^2U\right)p-\frac{1}{2}|\nabla U|^2\right),$$

and hence

$$(\mathcal{L}_A^*+\mathcal{L}_B^*)\left(\mathcal{L}_A^*+\frac{1}{2}\mathcal{L}_B^*\right)g = \beta\left(3p^T\left(\nabla^2U\right)\nabla U-\nabla^3U:p\otimes p\otimes p\right),$$

similarly

$$(\mathcal{L}_A^*+\mathcal{L}_B^*)\left(\frac{1}{2}\mathcal{L}_A^*+\mathcal{L}_B^*\right)g = \beta\left(3p^T\left(\nabla^2U\right)\nabla U-\frac{1}{2}\nabla^3U:p\otimes p\otimes p\right).$$

Comparing with (4.54–4.55), we have

$$\begin{aligned}\mathcal{L}_{\mathrm{LD}}f_2^{[\![\mathrm{CBABC}]\!]} &= w^{[\![\mathrm{CBABC}]\!]} = \frac{\beta}{12}\left(\nabla^3U:p\otimes p\otimes p-3p^T\left(\nabla^2U\right)\nabla U\right),\\ \mathcal{L}_{\mathrm{LD}}f_2^{[\![\mathrm{CABAC}]\!]} &= w^{[\![\mathrm{CABAC}]\!]} = \frac{\beta}{12}\left(3p^T\left(\nabla^2U\right)\nabla U-\frac{1}{2}\nabla^3U:p\otimes p\otimes p\right).\end{aligned}\tag{4.92}$$

Noting that $\pi w = 0$ in both cases, we can make use of the bound (4.91) to show that there exists some constant $K>0$ such that for $\gamma\geq 1$

$$\left\|f_2-\mathcal{L}_{\mathrm{ovd}}^{-1}\pi(\mathcal{L}_A+\mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1}w\right\|_H \leq \frac{K}{\gamma}.\tag{4.93}$$

Using the formulas

$$\mathcal{L}_{\bar{C}}\left(\nabla^3U:p\otimes p\otimes p\right) = -3\nabla^3U:p\otimes p\otimes p+\frac{6}{\beta}p^T\nabla_q\Delta_qU,$$
$$\mathcal{L}_{\bar{C}}\left(p^T\nabla_q\Delta_qU\right) = -p^T\nabla_q\Delta_qU,\qquad \mathcal{L}_{\bar{C}}\left(p^T\left(\nabla^2U\right)\nabla U\right) = -p^T\left(\nabla^2U\right)\nabla U,$$

it is clear that

$$\begin{aligned}w^{[\![\mathrm{CBABC}]\!]} &= \frac{\beta}{12}\mathcal{L}_{\bar{C}}\left(-\frac{1}{3}\nabla^3U:p\otimes p\otimes p-\frac{2}{\beta}p^T\nabla_q\Delta_qU+3p^T\left(\nabla^2U\right)\nabla U\right),\\ w^{[\![\mathrm{CABAC}]\!]} &= \frac{\beta}{12}\mathcal{L}_{\bar{C}}\left(\frac{1}{6}\nabla^3U:p\otimes p\otimes p+\frac{1}{\beta}p^T\nabla_q\Delta_qU-3p^T\left(\nabla^2U\right)\nabla U\right).\end{aligned}$$

Considering the $[\![\text{CBABC}]\!]$ scheme, we can use the formulas

$$\nabla^3 U : p \otimes p \otimes p = \mathcal{L}_A^3 U = \mathcal{L}_A^3 \pi U, \tag{4.94}$$

$$p^T \nabla_q \Delta_q U = \mathcal{L}_A \Delta_q U = \mathcal{L}_A \pi \Delta_q U, \qquad p^T \left(\nabla^2 U\right) \nabla U = \frac{1}{2} \mathcal{L}_A |\nabla U|^2 = \frac{1}{2} \mathcal{L}_A \pi |\nabla U|^2, \tag{4.95}$$

to write

$$\begin{aligned}
\pi(\mathcal{L}_A + \mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1} w^{[\![\text{CBABC}]\!]} &= \frac{\beta}{12} \pi(\mathcal{L}_A + \mathcal{L}_B) \left(3p^T \left(\nabla^2 U\right) \nabla U - \frac{1}{3} \nabla^3 U : p \otimes p \otimes p \right. \\
&\qquad \left. - \frac{2}{\beta} p^T \nabla_q \Delta_q U \right), \\
&= \frac{\beta}{12} \pi(\mathcal{L}_A + \mathcal{L}_B) \left(\frac{3}{2} \mathcal{L}_A \pi |\nabla U|^2 - \frac{1}{3} \mathcal{L}_A^3 \pi U - \frac{2}{\beta} \mathcal{L}_A \pi \Delta_q U \right), \\
&= \frac{\beta}{12} \pi \left(\frac{3}{2} \left(\mathcal{L}_A^2 + \mathcal{L}_B \mathcal{L}_A\right) \pi |\nabla U|^2 - \frac{2}{\beta} \left(\mathcal{L}_A^2 + \mathcal{L}_B \mathcal{L}_A\right) \pi \Delta_q U \right. \\
&\qquad \left. - \frac{1}{3} \mathcal{L}_B \mathcal{L}_A^3 \pi U - \frac{1}{3} \mathcal{L}_A^4 \pi U \right).
\end{aligned}$$

Recall the formulas in (4.75), giving

$$\pi \left(\mathcal{L}_A^2 + \mathcal{L}_B \mathcal{L}_A\right) \pi = \pi \mathcal{L}_{\text{ovd}} \pi,$$

$$\pi \mathcal{L}_A^4 \pi = 3\pi \mathcal{L}_A^2 \pi \mathcal{L}_A^2 \pi, \qquad \pi \mathcal{L}_B \mathcal{L}_A^3 \pi = 3\pi \mathcal{L}_B \mathcal{L}_A \pi \mathcal{L}_A^2 \pi.$$

Using $\pi \mathcal{L}_A^2 \pi U = \Delta_q U / \beta$, we have

$$\pi(\mathcal{L}_A + \mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1} w^{[\![\text{CBABC}]\!]} = \frac{\beta}{12} \left(\frac{3}{2} \mathcal{L}_{\text{ovd}} |\nabla U|^2 - \frac{3}{\beta} \mathcal{L}_{\text{ovd}} \Delta_q U \right),$$

and hence

$$\mathcal{L}_{\text{ovd}}^{-1} \pi(\mathcal{L}_A + \mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1} w^{[\![\text{CBABC}]\!]} = \frac{1}{8} \left(c + \beta |\nabla U|^2 - 2\Delta_q U \right), \tag{4.96}$$

where the constant c is chosen to ensure the right hand side has average 0. Similarly for the $[\![\text{CABAC}]\!]$ scheme,

$$\begin{aligned}
\pi(\mathcal{L}_A + \mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1} w^{[\![\text{CABAC}]\!]} = \frac{\beta}{12} \pi \left(\frac{1}{6} \left(\mathcal{L}_A^4 + \mathcal{L}_B \mathcal{L}_A^3\right) \pi U + \frac{1}{\beta} \left(\mathcal{L}_A^2 + \mathcal{L}_B \mathcal{L}_A\right) \pi \Delta_q U \right. \\
\left. - \frac{3}{2} \left(\mathcal{L}_A^2 + \mathcal{L}_B \mathcal{L}_A\right) |\nabla U|^2 \right),
\end{aligned}$$

giving the result

$$\mathcal{L}_{\text{ovd}}^{-1} \pi(\mathcal{L}_A + \mathcal{L}_B)\mathcal{L}_{\bar{C}}^{-1} w^{[\![\text{CABAC}]\!]} = \frac{1}{8} \left(\Delta_q U - |\nabla U|^2 \right). \tag{4.97}$$

For the other second order schemes, we can perform a similar procedure to find equiv-

alent results. Alternately, we make use of (4.57-4.58) with

$$(\mathcal{L}_A^* + \mathcal{L}_B^*)\, g = \beta \left(p^T \left(\nabla^2 U\right) p - |\nabla U|^2\right),$$

to yield the relations

$$f_2^{[\![\mathrm{CBABC}]\!]} = f_2^{[\![\mathrm{ABCBA}]\!]} - \frac{\beta}{8}\left(p^T \left(\nabla^2 U\right) p - |\nabla U|^2\right),$$
$$f_2^{[\![\mathrm{CABAC}]\!]} = f_2^{[\![\mathrm{BACAB}]\!]} + \frac{\beta}{8}\left(p^T \left(\nabla^2 U\right) p - |\nabla U|^2\right).$$

Combining this result with the results (4.96) and (4.97), we can use (4.93) to write

$$\left\| f_2^{[\![\mathrm{CBABC}]\!]} - \frac{1}{8}\left(c + \beta|\nabla U|^2 - 2\Delta_q U\right)\right\|_H \leq \frac{K}{\gamma}, \tag{4.98}$$

$$\left\| f_2^{[\![\mathrm{CABAC}]\!]} - \frac{1}{8}\left(\Delta_q U - \beta|\nabla U|^2\right)\right\|_H \leq \frac{K}{\gamma}, \tag{4.99}$$

$$\left\| f_2^{[\![\mathrm{ABCBA}]\!]} - \frac{1}{8}\left(c + \beta p^T \left(\nabla^2 U\right) p - 2\Delta_q U\right)\right\|_H \leq \frac{K}{\gamma}, \tag{4.100}$$

$$\left\| f_2^{[\![\mathrm{BACAB}]\!]} - \frac{1}{8}\left(\Delta_q U - \beta p^T \left(\nabla^2 U\right) p\right)\right\|_H \leq \frac{K}{\gamma}, \tag{4.101}$$

where $c = \langle \Delta U \rangle$. By taking averages of the limiting functions with respect to the canonical momentum distribution κ_β, it is evident that in the limit of large friction the underdamped correction functions converge to their corresponding overdamped corrections, given in (4.80) (where $h = \delta t^2/2$). The rate of this convergence is at least $1/\gamma$, although in the case of the $[\![\mathrm{BACAB}]\!]$ scheme for the perturbed harmonic oscillator, it is observed in (4.66) to converge at a rate $1/\gamma^2$. We would expect that by continuing to find more terms in our ansatz, we could solve to find the leading order perturbation for this special case.

4.3.3 Numerical experiment

We consider the one-dimensional system with the uneven double-well potential

$$U(q) = (q^2 - 1)^2 + q/2,$$

and consider using each of the four second order schemes to compute the average of $\varphi = q^2 - \langle q^2 \rangle$. We use $\beta = 1$ with friction parameters $\gamma = 0.5$ and $\gamma = 5$, verifying the results of (4.98). For each timestep, we run 128 experiments of 10^9 total iterations, using each of the four schemes. We then plot the error and magnitude of the computed averages for each method in Figure 4.5.

We also plot the expected limiting results for the overdamped case as solid lines, except in the case of the $[\![\mathrm{BACAB}]\!]$ scheme where we expect higher order behaviour. We demonstrate the approximate convergence of the $[\![\mathrm{BACAB}]\!]$ scheme using a fourth-order

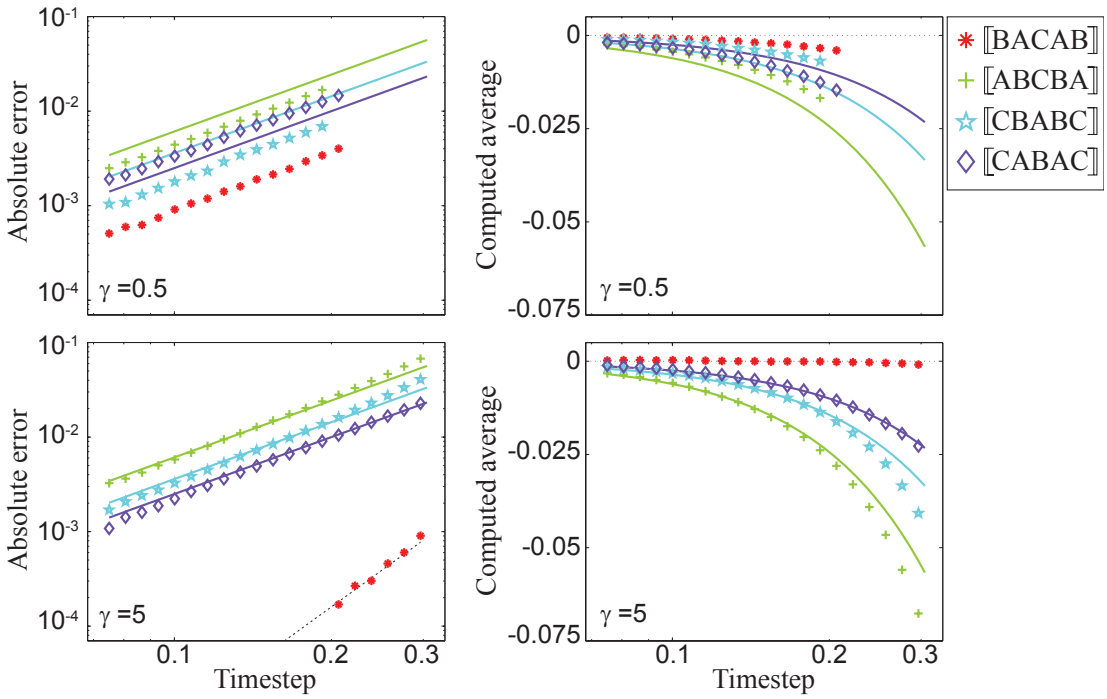


Figure 4.5: We give the error in computed averages of $q^2 - \langle q^2 \rangle$ at $\beta = 1$ for the uneven double well potential (4.48), from simulations using the given timestep and friction for each of the second order schemes. We plot the observed averages (coloured symbols) with the results expected from our analysis of the overdamped limit (corresponding coloured lines). All methods give a second-order trend at low friction, but at high friction the ⟦BACAB⟧ method demonstrates a superconvergence property. As expected, as friction is increased the errors in the averages converge towards the errors given for their overdamped versions. We plot a fourth-order line (black,dotted) to estimate the behaviour of the ⟦BACAB⟧ method.

guideline (black, dashed) in Figure 4.5.

As we expect, for small friction the schemes all give second-order errors. Increasing γ has a dramatic effect on the ⟦BACAB⟧ scheme, where despite the relatively large number of computations we are unable to demonstrate any trend in its errors for timesteps below 0.2, due to sampling error dominating the results. The other schemes show reasonable convergence towards their respective limiting curves, where we would expect that increasing the friction would reduce the deviation further.

Chapter 5

Applications in molecular sampling

This chapter will focus on practical applications for the framework and numerical discretization schemes we have developed in previous sections. We begin with describing two strategies for evaluating the error term during simulation.

Section 5.1 summarizes work in collaboration with G. Stoltz on numerical estimation of the perfect sampling bias error on-the-fly during simulation. Using a linear response identity, we are able to evaluate the leading error term as the integral of a correlation function.

Section 5.2 details new work on perturbing the Langevin dynamics vector field in order to target higher-order sampling for one specific observable, by manipulating the PDE governing the leading error term to ensure that the correction function is orthogonal to the desired observable, with respect to the canonical weighting ρ_β.

We conclude the chapter by giving detailed results of a molecular simulation comparing the four second-order schemes developed with four schemes in current use in MD packages. This work combines new simulation results with those presented in collaboration with B. Leimkuhler in [69].

5.1 Error estimation using linear response techniques

The results of Chapter 4 show that by using an order s numerical method, we can find a suitable correction function f_s such that for a sufficiently smooth observable $\varphi(q,p)$ we have

$$\int_\Omega \varphi \hat{\rho}_{\delta t} \mathrm{d}z = \int_\Omega \varphi \rho_\beta \mathrm{d}z + \delta t^s \int_\Omega f_s \varphi \rho_\beta \mathrm{d}z + \delta t^{s+1} r_{s,\delta t}, \tag{5.1}$$

for uniformly bounded $r_{s,\delta t}$, and where f_s is a solution to the Poisson equation

$$\mathcal{L}_{\mathrm{LD}}^* f_s \rho_\beta = w_s \rho_\beta, \tag{5.2}$$

for method dependent function $w_s(q,p)$. We are able to solve for f_s in only select cases due to the complexity of the Langvin dynamics operator $\mathcal{L}_{\mathrm{LD}}$, such as the first-order schemes in (4.43) or the overdamped case in (4.80). However, in the general case it is very challenging to solve (5.2) and difficult to compute a numerical solution (using finite difference methods, for example) due to the extremely high dimensionality in which the PDE is posed. Thus we are unable to correct the second order methods at arbitrary, finite friction, as we cannot solve their corresponding partial differential equations (4.51).

Instead, in order to compute the correction value we will work with the right hand side of (5.2), which is easily found using the BCH formula and commutator expansion. The correction can then be written as an integrated correlation function by virtue of the linear response identity

$$\mathcal{L}_{\mathrm{LD}}^{-1} = -\int_0^\infty e^{t\mathcal{L}_{\mathrm{LD}}}\mathrm{d}t,$$

as an operator on $\mathcal{H}^1$ (see [50, 51, 93] for precise justfication of this equality). For any value of τ, we have

$$\int_\Omega \left(e^{\tau\mathcal{L}_{\mathrm{LD}}}\phi\right)\psi\rho_\beta \mathrm{d}z = \mathbb{E}\left[\phi(z(\tau))\,\psi(z(0))\right],$$

where the expectation is taken with respect to all initial conditions $z(0)\in\Omega$ distributed canonically according to ρ_β. Combining these formulas, we can express the correction as a double integral in space and time, and evaluate the integral using simulation data. By switching the order of integration, the correction term becomes

$$\begin{aligned}\int_\Omega f_s\varphi\rho_\beta \mathrm{d}z = \int_\Omega \left(\mathcal{L}_{\mathrm{LD}}^{-*}w_s\rho_\beta\right)\varphi \mathrm{d}z &= -\int_0^\infty\int_\Omega \left(e^{t\mathcal{L}_{\mathrm{LD}}}\varphi\right)w_s\rho_\beta \mathrm{d}t\,\mathrm{d}z \\ &= -\int_0^\infty \mathbb{E}\left[\varphi(z(t))\,w_s(z(0))\right]\mathrm{d}t.\end{aligned} \tag{5.3}$$

In order to calculate this correction value accurately, we shall consider appropriate methods for computing the integral to high order, using a Riemann sum with a perturbed observable function much in the style of backward error analysis and Section 2.1.2.

Let us consider two sufficiently smooth observables $\phi,\psi\in\widetilde{\mathcal{S}}$, so

$$\langle\phi\rangle = \langle\psi\rangle = 0.$$

We seek a perturbed observable $\phi_{\delta t}$ that allows us to approximate a correlation function to order $\alpha\in\mathbb{N}^*$, such that

$$\int_0^\infty \mathbb{E}\left[\phi(z(t))\,\psi(z(0))\right]\mathrm{d}t = \delta t\sum_{k=0}^\infty \mathbb{E}_{\delta t}\left(\phi_{\delta t}(z_k)\psi(z_0)\right) + \mathcal{O}(\delta t^\alpha),$$

where $\mathbb{E}_{\delta t}$ denotes that the expectation is taken with respect to all initial conditions $z_0 \in \Omega$, distributed according to $\widehat{\rho}_{\delta t}$. The (z_k) are the discretized points from the trajectory, where the subscript indicates iteration number, such that $z_0 = z(0)$.

Assume we have a numerical scheme with evolution operator $\mathcal{P}_{\delta t}$ and invariant distribution $\widehat{\rho}_{\delta t}$ such that for sufficiently smooth φ and sufficiently small δt

$$\int \varphi \rho_\beta \mathrm{d}z = \int \varphi \widehat{\rho}_{\delta t} + \mathcal{O}(\delta t^\alpha),$$

and

$$-\left(\frac{\mathrm{Id} - \mathcal{P}_{\delta t}}{\delta t}\right) = \mathcal{L}_{\mathrm{LD}} + \delta t S_1 + \ldots + \delta t^{\alpha-1} S_{\alpha-1} + \delta t^\alpha R_\alpha,$$

where R_α uses finitely many derivatives. Assuming we have sufficiently small step size, we introduce the invariant numerical distribution through $\langle \phi \rangle = \left\langle \mathcal{L}_{\mathrm{LD}}^{-1} \phi \right\rangle = 0$, where

$$\begin{aligned}\int_\Omega \left(\mathcal{L}_{\mathrm{LD}}^{-1}\phi\right) \psi \rho_\beta \mathrm{d}z &= \int_\Omega \left(\mathcal{L}_{\mathrm{LD}}^{-1}\phi\right) \left(\psi - \langle \psi \rangle_{\delta t}\right) \rho_\beta \mathrm{d}z, \\ &= \int_\Omega \left(\mathcal{L}_{\mathrm{LD}}^{-1}\phi\right) \left(\psi - \langle \psi \rangle_{\delta t}\right) \widehat{\rho}_{\delta t} \mathrm{d}z + \mathcal{O}(\delta t^\alpha).\end{aligned}$$

Then using

$$\left(\delta t \sum_{k=0}^{\infty} \mathcal{P}_{\delta t}^k\right)\left(\frac{\mathrm{Id} - \mathcal{P}_{\delta t}}{\delta t}\right) = \mathrm{Id},$$

we have

$$-\mathcal{L}_{\mathrm{LD}}^{-1} = \left(\delta t \sum_{k=0}^{\infty} \mathcal{P}_{\delta t}^k\right)\left(\mathrm{Id} + \delta t S_1 \mathcal{L}_{\mathrm{LD}}^{-1} + \ldots + \delta t^{\alpha-1} S_{\alpha-1}\mathcal{L}_{\mathrm{LD}}^{-1} + \delta t^\alpha R_\alpha \mathcal{L}_{\mathrm{LD}}^{-1}\right),$$

allowing us to write

$$\int_\Omega \left(\mathcal{L}_{\mathrm{LD}}^{-1}\phi\right)\left(\psi - \langle \psi \rangle_{\delta t}\right)\widehat{\rho}_{\delta t}\mathrm{d}z = -\delta t \int_\Omega \left(\sum_{k=0}^{\infty} \mathcal{P}_{\delta t}^k \widehat{\phi}_{\delta t}\right)\left(\psi - \langle \psi \rangle_{\delta t}\right)\widehat{\rho}_{\delta t}\mathrm{d}z + \mathcal{O}(\delta t^\alpha), \quad (5.4)$$

for

$$\widehat{\phi}_{\delta t} = \left(\mathrm{Id} + \delta t S_1 \mathcal{L}_{\mathrm{LD}}^{-1} + \ldots + \delta t^{\alpha-1} S_{\alpha-1}\mathcal{L}_{\mathrm{LD}}^{-1}\right)\phi.$$

Making use of the identity

$$\begin{aligned}\int_\Omega \left(\mathcal{P}_{\delta t}^n \phi\right)\left(\psi - \langle \psi \rangle_{\delta t}\right)\widehat{\rho}_{\delta t}\mathrm{d}z &= \int_\Omega \left(\mathcal{P}_{\delta t}^n \left(\phi - \langle \phi \rangle_{\delta t}\right)\right)\left(\psi - \langle \psi \rangle_{\delta t}\right)\widehat{\rho}_{\delta t}\mathrm{d}z, \\ &= \int_\Omega \left(\mathcal{P}_{\delta t}^n \left(\phi - \langle \phi \rangle_{\delta t}\right)\right)\psi\widehat{\rho}_{\delta t}\mathrm{d}z,\end{aligned}$$

we find

$$\int_\Omega \left(\sum_{k=0}^{\infty} \mathcal{P}_{\delta t}^k \widehat{\phi}_{\delta t}\right)\left(\psi - \langle \psi \rangle_{\delta t}\right)\widehat{\rho}_{\delta t}\mathrm{d}z = \int_\Omega \left(\sum_{k=0}^{\infty} \mathcal{P}_{\delta t}^k \left(\widehat{\phi}_{\delta t} - \langle \widehat{\phi}_{\delta t} \rangle_{\delta t}\right)\right)\psi\widehat{\rho}_{\delta t}\mathrm{d}z,$$

and hence we can combine the formulas to give

$$\begin{aligned}\int_0^\infty \mathbb{E}\left[\phi(z(t))\,\psi(z(0))\right]\mathrm{d}t &= \int_\Omega \int_0^\infty \left(e^{t\mathcal{L}_{\mathrm{LD}}}\phi\right)\,\psi\rho_\beta \mathrm{d}t\mathrm{d}z,\\ &= -\int_\Omega \left(\mathcal{L}_{\mathrm{LD}}^{-1}\phi\right)\,\psi\rho_\beta \mathrm{d}z,\\ &= -\int_\Omega \left(\mathcal{L}_{\mathrm{LD}}^{-1}\phi\right)\left(\psi - \langle\psi\rangle_{\delta t}\right)\widehat{\rho}_{\delta t}\mathrm{d}z + \mathcal{O}(\delta t^\alpha)\\ &= \delta t\int_\Omega \left(\sum_{k=0}^\infty \mathcal{P}_{\delta t}^k\left(\widehat{\phi}_{\delta t} - \left\langle\widehat{\phi}_{\delta t}\right\rangle_{\delta t}\right)\right)\psi\widehat{\rho}_{\delta t}\mathrm{d}z + \mathcal{O}(\delta t^\alpha).\end{aligned}$$

Thus, choosing $\phi_{\delta t} = \widehat{\phi}_{\delta t} - \left\langle\widehat{\phi}_{\delta t}\right\rangle_{\delta t}$, where

$$\widehat{\phi}_{\delta t} = \phi + \delta t S_1\mathcal{L}_{\mathrm{LD}}^{-1}\phi + \ldots + \delta t^{\alpha-1}S_{\alpha-1}\mathcal{L}_{\mathrm{LD}}^{-1}\phi,$$

we have

$$\int_0^\infty \mathbb{E}\left[\phi(z(t))\,\psi(z(0))\right]\mathrm{d}t = \delta t\sum_{k=0}^\infty \mathbb{E}_{\delta t}\left[\phi_{\delta t}(z_k)\psi(z_0)\right] + \delta t^\alpha\widetilde{r}_\alpha, \tag{5.5}$$

for sufficiently small δt, and sufficiently smooth functions ϕ and ψ with uniformly bounded $\widetilde{r}_\alpha$.

If we choose $\alpha = 1$, then $\phi_{\delta t} = \phi_{\delta t}^{[1]} = \phi - \langle\phi\rangle_{\delta t}$, reducing (5.5) to a Riemann sum. In the more relevant case of symmetric second order methods, we have $\alpha = 2$ and $S_1 = \mathcal{L}_{\mathrm{LD}}^2/2$. In this case, we choose $\phi_{\delta t} = \phi_{\delta t}^{[2]}$ where

$$\phi_{\delta t}^{[2]} = \phi - \langle\phi\rangle_{\delta t} + \delta t\left(\mathcal{L}_{\mathrm{LD}}\phi - \langle\mathcal{L}_{\mathrm{LD}}\phi\rangle_{\delta t}\right)/2, \tag{5.6}$$

and hence using (5.4) we have

$$\sum_{k=0}^\infty \mathbb{E}_{\delta t}\left[\phi_{\delta t}^{[2]}(z_k)\psi(z_0)\right] = \sum_{k=0}^\infty \mathbb{E}_{\delta t}\left[\phi_{\delta t}^{[1]}(z_k)\psi(z_0)\right] - \frac{1}{2}\int\left(\mathcal{L}_{\mathrm{LD}}^{-1}\left(\mathcal{L}_{\mathrm{LD}}\phi - \langle\mathcal{L}_{\mathrm{LD}}\phi\rangle_{\delta t}\right)\right)\psi\widehat{\rho}_{\delta t}\mathrm{d}z.$$

Of course,

$$\frac{1}{2}\int\left(\mathcal{L}_{\mathrm{LD}}^{-1}\left(\mathcal{L}_{\mathrm{LD}}\phi - \langle\mathcal{L}_{\mathrm{LD}}\phi\rangle_{\delta t}\right)\right)\psi\widehat{\rho}_{\delta t}\mathrm{d}z = \frac{1}{2}\langle(\phi - \langle\phi\rangle_{\delta t})\psi\rangle_{\delta t} + \mathcal{O}(\delta t^2),$$

giving

$$\sum_{k=0}^\infty \mathbb{E}_{\delta t}\left[\phi_{\delta t}^{[2]}(z_k)\psi(z_0)\right] = \frac{1}{2}\langle\phi_{\delta t}^{[1]}\psi\rangle_{\delta t} + \sum_{k=1}^\infty \mathbb{E}_{\delta t}\left[\phi_{\delta t}^{[1]}(z_k)\psi(z_0)\right] + \mathcal{O}(\delta t^2).$$

This is precisely the 'trapezoidal rule', although we would not expect identical results in comparison to the trapezoidal rule as the remainder terms will differ, leading to different behaviour at higher orders. For schemes of higher order than 2 (for example

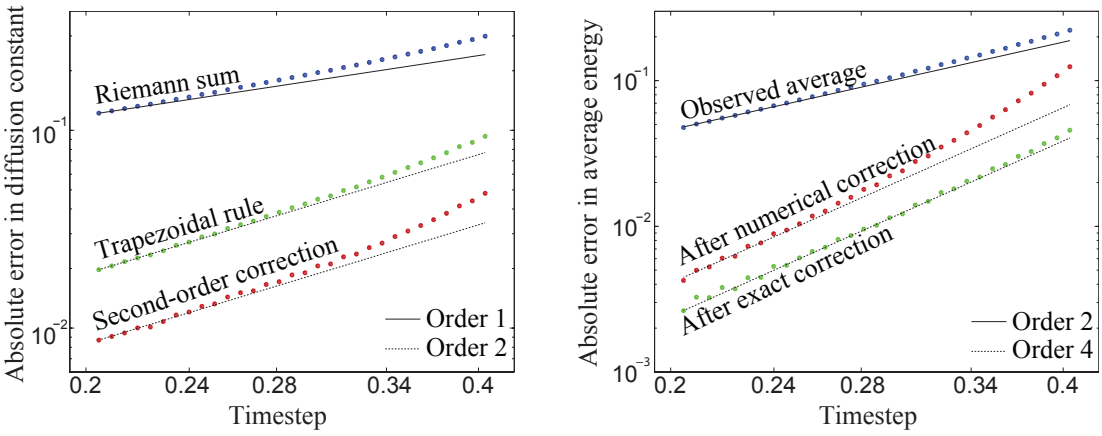


Figure 5.1: Left: The error in the value of the integrated velocity autocorrelation function is compared at a number of time steps when computed using a Riemann sum (blue) or the correction term provided in (5.5) (red). The result from computing the integral using the trapezoidal rule (green) is also shown. Right: The error in the computed average of total energy is plotted (blue), with the correction term computed using the same step size (green) demonstrating the practical application of the method. We can test the validity of (5.7) in principle by computing the correction more accurately at a smaller time step in a separate simulation, this result is plotted in green. All results are computed using the scheme ⟦CBABC⟧ with $\beta = \gamma = 1$.

a Yoshida-type scheme) we would expect to be able to continue with this procedure to evaluate the required expectation functions at equally high order.

Returning to (5.3), we can approximate the correction function by choosing $\phi = \varphi$ and $\psi = w_s$ in (5.5), to yield

$$\int_\Omega f_s \varphi \rho_\beta \mathrm{d}z = -\delta t \sum_{k=0}^{\infty} \mathbb{E}_{\delta t} \left[\varphi_{\delta t}(z_k) w_s(z_0) \right] + \delta t^s \widetilde{r}_s, \tag{5.7}$$

for appropriate $\varphi_{\delta t}$. Thus instead of solving (4.51) to find explicit solutions for the correction functions f_2, we can compute the correction on-the-fly during simulation using this method. This allows us to correct an average, or gauge the quality of the discretized sampling.

The function $w_s(q, p)$, which comes from the right hand side of (4.51), may itself be computationally expensive to evaluate at every timestep, so this method may not be practical in all cases (although f_s may also be expensive).

We illustrate numerically the results of (5.5) and (5.7). We consider a small planar model $q = (x, y) \in \mathcal{M} = (2\pi\mathbb{T})^2$, with the potential energy function

$$U(q) = 2\cos(2x) + \cos(y).$$

We fix temperature $\beta = 1$, with $\gamma = 1$ and $M = \mathrm{Id}$, and use the ⟦CBABC⟧ scheme

to evolve the dynamics. We aim to correct the average energy $\left\langle p_x^2/2 + p_y^2/2 + U(q) \right\rangle$ using (5.1) to yield a fourth-order average, by evaluating the correction term through (5.7), where the w_s function was computed in (4.92) as

$$w_2^{[\![\mathrm{CBABC}]\!]}(q,p) = \frac{\beta}{12}\left(\nabla^3 U : p \otimes p \otimes p - 3p^T\left(\nabla^2 U\right)\nabla U\right).$$

We compare the observed and corrected averages to the exact average, where the exact average can be computed to arbitrary numerical precision (due to the simplicity of the model) using standard solvers.

Additionally we seek to compute the diffusion constant in the x-direction D_x, as the integral of the velocity autocorrelation function

$$D_x = \frac{1}{2}\int \mathbb{E}\left[p_x(t)p_x(0)\right] \mathrm{d}t,$$

to order $\alpha = 2$. As well as using the trapezoidal rule, we use the perturbed Riemann sum given in (5.5), where we set $\phi(q,p) = p_x$ and $\psi(q,p) = p_x/4$. The perturbed observable we use is given by (5.6) as

$$\begin{aligned}\phi_{\delta t}(q,p) &= \phi(q,p) + \frac{\delta t}{2}\mathcal{L}_{\mathrm{LD}}\phi(q,p) - \langle \phi(q,p) + \frac{\delta t}{2}\mathcal{L}_{\mathrm{LD}}\phi(q,p)\rangle_{\delta t},\\ &= \left(1 - \frac{\gamma\delta t}{2}\right)p_x - \frac{\delta t}{2}F_x(q) - \left\langle \left(1 - \frac{\gamma\delta t}{2}\right)p_x - \frac{\delta t}{2}F_x(q)\right\rangle_{\delta t},\end{aligned}$$

where $F_x(q)$ is the force in the x-direction. The computed value for D_x is compared to a baseline value computed using a much smaller step size ($\delta t = 0.05$).

For both sets of results, we run 10^3 repeat simulations for a fixed time interval of 2×10^8, where $\delta t \in [0.2, 0.4]$. The required Riemann sums were computed during simulation by maintaining a history of discretized points. This history was kept at fixed time length 20, at which point ϕ and ψ were suitably decorrelated, minimizing the expected error. We accept that truncating the length of the Riemann sums in time would lead to errors, however this is both a necessity of the hardware, as well as a computational concern: due to the relative simplicity of the model, the computational bottleneck in simulation was maintaining the history for computing the correction estimates.

We give the results for the simulations in Figure 5.1. The left plot demonstrates that we are able to compute the velocity autocorrelation to second order accuracy, with the result improving on the trapezoidal rule. The difference between the two comes from the presence of the additional higher order terms when using $\phi_{\delta t}$, whose interplay may help or hinder the sampling. We do not expect this method to consistently outperform the trapezoidal rule for more general systems.

The right hand plot in Figure 5.1 shows that the observed error in the average energy is second-order, as expected. We can apply (5.7) to correct the average, giving a fourth-order relation (marked in red). However at higher values of δt, we find that the numerical estimation of the correction introduces large errors itself, that come close

to complementing the observed average, introducing more error rather than correcting it. We can remove this behaviour by computing the correction in a separate simulation, using (5.7) with a much smaller step size ($\delta t = 0.1$), and then use this value to correct the average. This result (green) is shown to be far more robust, even at larger step sizes.

5.2 Modified Langevin dynamics equation

We consider using the framework presented in Chapter 4 to analyze the weak order of schemes integrating a perturbed version of the Langevin dynamics equation, with a view to increase the weak order of a desired method with respect to ρ_β. While introducing modified equations to increase the order of symplectic methods is well-studied [18, 48, 99], we are motivated by recent results for modified SDE equations [1, 108, 130] and attempt to optimize the efficiency of a scheme by considering only one class of perturbation to the standard Langevin dynamics. We shall proceed under the assumption that the step size is sufficiently small and the perturbation is sufficiently smooth such that the evolution operator of the method (and its derivatives) remain bounded.

For a fixed step size $\delta t > 0$, we will consider perturbing Langevin dynamics by a smooth function $\nu(q,p) \in \mathbb{R}^N$, with a new splitting strategy

$$\begin{bmatrix} \mathrm{d}q \\ \mathrm{d}p \end{bmatrix} = \underbrace{\begin{bmatrix} M^{-1}p\,\mathrm{d}t \\ 0 \end{bmatrix}}_{A} + \underbrace{\begin{bmatrix} 0 \\ -\nabla U(q) + \delta t^2 \nu(q,p)\,\mathrm{d}t \end{bmatrix}}_{\widehat{B}} + \underbrace{\begin{bmatrix} 0 \\ -\gamma p\,\mathrm{d}t + \sqrt{2\gamma/\beta}\,M^{1/2}\,\mathrm{d}W \end{bmatrix}}_{C}. \tag{5.8}$$

The reason we perturb the B piece and not the A or C pieces is for reasons of computation–we anticipate that the function ν will require some system information in the form derivatives of the potential energy function. As the B step requires a force evaluation anyway (and hence calculation of all interatomic distances), it will be most efficient to evaluate higher derivatives of $U(q)$ at the same time (see [76] for examples with Hessian-vector products). However, we accept that some additional computational cost will be required.

We cannot necessarily solve the $\widehat{B}$ piece exactly, though as q remains constant we expect that a high-accuracy solution is attainable using implicit methods or numerical quadrature. As in the case of multiple time stepping in Section 4.2, we assume that the evolution is computed to sufficient accuracy that we are able to use the exact evolution operator in our analysis.

Proceeding as in Section 4.1.2, we define our method through a composition of the corresponding evolution operators, where we have the L^2 generator for the $\widehat{B}$ vector field

$$\mathcal{L}_{\widehat{B}} = \mathcal{L}_B + \delta t^2 \nu(q,p) \cdot \nabla_p, \qquad \mathcal{L}^*_{\widehat{B}} = \mathcal{L}^*_B - \delta t^2 \nabla_p \cdot (\nu(q,p)\,\cdot\,).$$

We shall assume that for sufficiently small step size δt, the evolution operator and its derivatives can be bounded in such a way that we retain sufficient control over any remainder terms in the Taylor expansion of the method operator (4.29). As an example, let us set $M = I_N$ consider the scheme $[\![\mathrm{CA\widehat{B}AC}]\!]$, where

$$\mathcal{P}_{\delta t}^{[\![\mathrm{CA\widehat{B}AC}]\!]} = e^{\delta t \mathcal{L}_C/2}\, e^{\delta t \mathcal{L}_A/2}\, e^{\delta t \mathcal{L}_{\widehat{B}}}\, e^{\delta t \mathcal{L}_C/2}\, e^{\delta t \mathcal{L}_A/2} = e^{\delta t \widehat{\mathcal{L}}_{\delta t}},$$

with

$$\widehat{\mathcal{L}}_{\delta t} = \mathcal{L}_A + \mathcal{L}_{\widehat{B}} + \mathcal{L}_C + \delta t^2 S_2^{[\![\mathrm{CABAC}]\!]} + \mathcal{O}(\delta t^4) = \mathcal{L}_{\mathrm{LD}} + \delta t^2 S_2^{[\![\mathrm{CA\widehat{B}AC}]\!]} + \mathcal{O}(\delta t^4),$$

where we have introduced

$$S_2^{[\![\mathrm{CA\widehat{B}AC}]\!]} = S_2^{[\![\mathrm{CABAC}]\!]} + \nu(q,p)\nabla_p.$$

Proceeding as in Section 4.1.3 using (4.51), we obtain

$$\mathcal{L}_{\mathrm{LD}}^* f_2^{[\![\mathrm{CA\widehat{B}AC}]\!]} \rho_\beta = -\left(S_2^{[\![\mathrm{CA\widehat{B}AC}]\!]}\right)^* \rho_\beta = -\left(S_2^{[\![\mathrm{CABAC}]\!]}\right)^* \rho_\beta + \nabla_p \cdot (\nu\, \rho_\beta)\,,$$

and hence

$$\mathcal{L}_{\mathrm{LD}} f_2^{[\![\mathrm{CA\widehat{B}AC}]\!]} = w^{[\![\mathrm{CABAC}]\!]} + \nabla_p \cdot \nu(q,p) - \beta p \cdot \nu(q,p). \tag{5.9}$$

This gives us the ability to edit the right-hand side of (4.51) as we see fit. Recall that for any sufficiently smooth observable φ with $\langle \varphi \rangle = 0$, the correction is of the form

$$\int_\Omega \varphi\, f_2\, \rho_\beta \mathrm{d}q\mathrm{d}p.$$

Hence if we seek high order sampling for only one particular observable, we need not necessarily rely on a method that gives higher order for all observables (which may be computationally prohibitive). If we can choose a correction function f_2 (such that $\langle f_2 \rangle = 0$) orthogonal to φ with respect to the weighting ρ_β, then we obtain higher order sampling if we are able to tailor the right-hand side of (5.9) (through judicious choice of ν) such that f_2 is a solution.

The resulting method is not a higher-order method, but instead it is tuned for a particular observable (or class of observables) to give higher order. As an example, if we seek higher order sampling for any function solely of position $\varphi(q)$, then choosing a suitable pair of functions ν and f_2, such that f_2 is a solution of (5.9) and $\pi f_2 = 0$, will give high order sampling.

The practical benefit of this approach is that by choosing an appropriate splitting scheme (our choice of $[\![\mathrm{CABAC}]\!]$ is motivated by the simplicity of $w^{[\![\mathrm{CABAC}]\!]}$ and its independence with respect to γ), we can search for any such suitable f_2 that makes $\nu(q,p)$ a computationally amenable function.

The approach also allows for more general corrections, if the computational weight

of ν is not a central issue. Completely annihilating the right hand side of (5.9) will give a higher order scheme for all observables. For example, recall from (4.92) that

$$w^{[\![\mathrm{CABAC}]\!]} = \frac{\beta}{12}\left(3p\cdot\left(\nabla^2 U\right)\nabla U - \frac{1}{2}\nabla^3 U : p\otimes p\otimes p\right).$$

Direct computation yields that for

$$\nu_1(q,p) = \nabla_q\cdot p^T\left(\nabla^2 U\right)p, \quad \nu_2(q,p) = \nabla_q^T\left(\nabla^2 U\right), \quad \nu_3(q,p) = \left(\nabla^2 U\right)\nabla U,$$

we have

$$\begin{aligned}
\nabla_p\cdot\nu_1 - \beta p\cdot\nu_1 &= 2\nabla_q\cdot\left(\nabla^2 U\right)p - \beta\nabla^3 U : p\otimes p\otimes p,\\
\nabla_p\cdot\nu_2 - \beta p\cdot\nu_2 &= -\beta\nabla_q\cdot\left(\nabla^2 U\right)p,\\
\nabla_p\cdot\nu_3 - \beta p\cdot\nu_3 &= -\beta p\cdot\left(\nabla^2 U\right)\nabla U.
\end{aligned}$$

Hence, choosing

$$\nu(q,p) = \frac{1}{4}\nu_3(q,p) - \frac{1}{24}\nu_1(q,p) - \frac{1}{12\beta}\nu_2(q,p),$$

we have

$$\mathcal{L}_{\mathrm{LD}} f_2^{[\![\mathrm{CA\widehat{B}AC}]\!]} = w^{[\![\mathrm{CABAC}]\!]} + \nabla_p\cdot\nu(q,p) - \beta p\cdot\nu(q,p) = 0,$$

giving higher order weak sampling for any suitable $\varphi(q,p)$, as $f_2^{[\![\mathrm{CA\widehat{B}AC}]\!]} = 0$.

We test this method on the uneven double well potential for $(q,p)\in\mathbb{R}^2$, where

$$U(q) = \left(q^2-1\right)^2 + q/2.$$

We will correct the average of

$$\varphi(q) = q^2 - \langle q^2\rangle, \tag{5.10}$$

at $\beta=\gamma=1$. We compare the results of the $[\![\mathrm{CABAC}]\!]$ and $[\![\mathrm{CA\widehat{B}AC}]\!]$ schemes in Figure 5.2, where each datapoint is an average from 128 runs of 10^8 iterations.

While we do indeed achieve fourth-order sampling from this scheme, comparing the results of the $[\![\mathrm{CA\widehat{B}AC}]\!]$ scheme with results in Figure 4.5 we see that at the larger tested step sizes of $\delta t = 0.2$ the error computed using the $[\![\mathrm{BACAB}]\!]$ scheme is approximately two orders of magnitude smaller than the $[\![\mathrm{CA\widehat{B}AC}]\!]$ scheme (which has a nontrivial increase in computational cost due to the required computation of $\nabla^3 U$).

5.3 Application of Langevin dynamics schemes to configurational sampling

In this section we shall compare the elementary second-order methods developed in Chapter 4 with popular Langevin dynamics methods in the literature that are derived

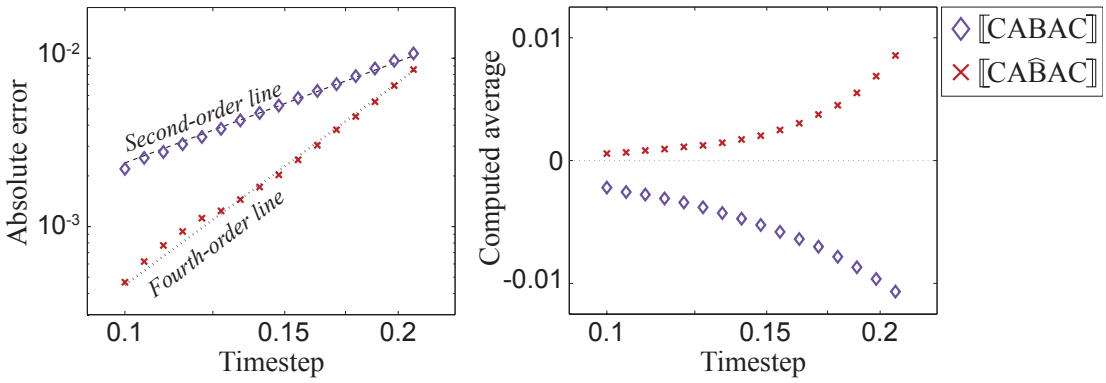


Figure 5.2: We verify that the modification to the Langevin dynamics method gives higher-order sampling for the observable given in (5.10), in the case of the ⟦CAB̂AC⟧ scheme.

through different approaches than the ABC splitting (4.5). All methods we consider will be second order and require one force evaluation per timestep: hence we consider them all to have the same computational cost as this is the dominant computational bottleneck in large systems. In smaller systems, the complexity of the algorithm (such as repeated state updates or generating multiple random numbers) will become more of a factor, though we will not consider this here.

Numerical methods using Metropolis conditioning, or any form of accept/reject branching will not be used, although these techniques complement trajectory-based algorithms nicely. Our focus will remain on methods that produce trajectory data that is consistent with Langevin dynamics. The metric we shall use to grade methods will be the computed error in the configurational sampling resulting from taking the average along the trajectory. This gives an indication of the corruption to the invariant measure.

As well as the four elementary methods

$$\llbracket \mathrm{CBABC} \rrbracket, \qquad \llbracket \mathrm{CABAC} \rrbracket, \qquad \llbracket \mathrm{ABCBA} \rrbracket, \qquad \llbracket \mathrm{BACAB} \rrbracket, \tag{5.11}$$

we shall test four additional numerical schemes from the literature by implementing them in the NAMD (Not Another Molecular Dynamics) package [94].

In comparison to alternate MD packages such as GROMACS (Groningen MAchine for Chemical Simulation) [122] , NAMD offers a lightweight version of the package (while retaining much of its functionality) called NAMD Lite, designed specifically for the purpose of testing new algorithms [54]. This allows utilization of the NAMD force routines, enabling testing using forcefields applied to standard problems in the literature. Procedures such as load balancing (for parallel simulation) or file handling are inherited from the NAMD source, making the implementation of the four elementary methods given in (5.11) very simple, owing to the modular nature of our splitting

approach. The velocity Verlet scheme ⟦BAB⟧ is standard in almost all MD codes, making the ⟦CBABC⟧ and ⟦BACAB⟧ schemes particularly easily to implement.

As many packages do not offer a choice of algorithms for the numerical integration of Langevin dynamics, for our implementation we have adapted the existing workflow and routines (such as random number generation) utilized in the software. In collaboration with D. Hardy, we have added the ⟦BACAB⟧ algorithm to the standard NAMD package.

We list the four additional methods that we will compare the schemes in (5.11) against:

Brünger, Brooks and Karplus (BBK)

Derived in [14], this extremely popular method is the common benchmark for many MD codes. Packages such as NAMD use this scheme as standard for the generation of Langevin dynamics trajectories.

Stochastic Position Verlet (SPV or ⟦ASA⟧)

The stochastic versions of the position and velocity Verlet schemes were proposed in [82] by Melchionna, using the splitting strategy given in (4.3) splitting the momentum from the position. At large friction, the force becomes dominated by the noise term (leading to inconsistency), while at low friction we expect ⟦ASA⟧ and ⟦ABCBA⟧ to have very similar behaviour.

Van Gunsteren and Berendsen (VGB)

Used as standard in packages such as GROMACS [122], this method was proposed in [123] and later revised in a more compact form in [124]. It is not derived from an additive decomposition of the vector field, but instead from local consideration of the trajectory itself.

Langevin Impulse (LI)

The Langevin impulse scheme was suggested by Skeel and Izaguirre in [111]. The version of the algorithm we consider is for sampling functions solely of position.

While the BBK and SPV methods are very simple to implement, we found that the LI and VGB schemes required significant time investment to ensure they were working correctly, due to the high complexity of the algorithms (all of the schemes are summarized in the Appendix). Some popular schemes reduce to algorithms we have already considered, these are given below:

Bussi-Parrinello

Proposed by Bussi and Parrinello [16], this scheme can be rewritten as exactly the ⟦CBABC⟧ scheme. Similarly, the scheme given in [109] is equivalent to ⟦CBABC⟧ using a reparameterized timestep.

Reverse leapfrog

The reverse leapfrog scheme [15] of Burrage and Lythe was derived to sample positions of harmonic potentials exactly by design, using a Runge-Kutta partioning strategy. Its update scheme (using unit masses) is

$$\begin{aligned}
p_{n+1/4} &= p_n - \delta t \nabla U(q_n)/2, \\
p_{n+2/4} &= \left(p_{n+1/4} + \frac{1}{2}\sqrt{2\gamma\delta t/\beta} R_n \right) (1 + \gamma\delta t/2)^{-1}, \\
q_{n+1} &= q_n + \delta t p_{n+2/4}, \\
p_{n+3/4} &= (1 - \gamma\delta t/2)\, p_{n+2/4} + \frac{1}{2}\sqrt{2\gamma\delta t/\beta} R_n, \\
p_{n+1} &= p_{n+3/4} - \delta t \nabla U(q_{n+1})/2.
\end{aligned}$$

For some fixed constant $\eta > 0$, plugging in

$$\gamma = \frac{4}{\delta t}\left(\left(1 + e^{-\eta\delta t}\right)^{-1} - \frac{1}{2} \right),$$

reduces the reverse leapfrog method to precisely the ⟦BACAB⟧ update scheme, using friction η. Correspondingly, using infinite friction in the symmetric ABC schemes is equivalent to setting $\gamma = 2/\delta t$ in the reverse leapfrog method, giving the overdamped method in (4.73). Noting that $\gamma = \eta + \mathcal{O}(\delta t^2 \eta^3)$, we expect much of the same behaviour as the ⟦BACAB⟧ method for low friction, though the reverse leapfrog scheme is plainly not suitable at large frictions.

GLA methods

Geometric Langevin Methods were introduced in [12], using splitting strategy given in equation (4.4). The Newtonian part is separated from the Ornstein-Uhlenbeck process, with the pieces integrated in sequence. The correction functions for the second order GLA methods (using the Verlet schemes for the Newtonian part) were shown in (4.60-4.61) to be identical to the symmetric methods ⟦CBABC⟧ and ⟦CABAC⟧ up to order δt^4 for configurational sampling. In fact, as the A step is the only operation to affect the position, we expect all of the methods denoted by

$$⟦\text{BCBA}⟧, \quad ⟦\text{CBAB}⟧, \quad ⟦\text{CBABC}⟧, \quad ⟦\text{BABC}⟧, \quad ⟦\text{ABCB}⟧,$$

to give identical averages when sampling functions of position.

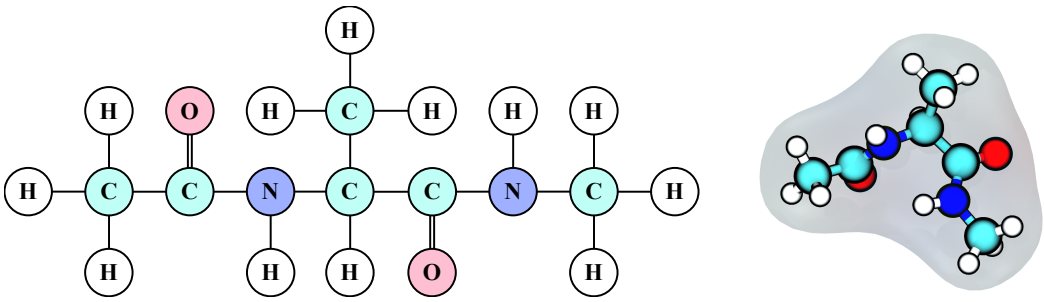


Figure 5.3: The structure of the alanine dipeptide protein is shown (left) with a visualization of the molecule in simulation (right).

As our focus will be on configurational sampling we will not specifically include comparisons with GLA schemes as the observed averages in experiments will be exactly as the either the ⟦CBABC⟧ or ⟦CABAC⟧ scheme, depending on the flavour of Verlet chosen as the second-order integrator.

5.3.1 A description of the alanine dipeptide molecule

We study the alanine dipeptide molecule, a classic test case for molecular dynamics. We implement each of the eight considered methods into the NAMD lite package [54], and observe the effect of discretization error (if any) on computed configurational averages. The CHARM22 force field was used to parameterize the potential energy functions suitably for simulation.

The alanine dipeptide model has 22 atoms with $d = 3$, where each atom is bonded to at least one other atom through a harmonic potential. Additionally there are 41 dihedral potential terms confining its shape, however only two of these dihedral terms are internal, making extensive sampling of its dihedral states computationally painless.

One of the most important features of a numerical method for ergodic dynamics (such as Langevin dynamics) is its preservation of the theoretical global phase space exploration rate. The spectral properties of the operator $\mathcal{L}^*_{\mathrm{LD}}$ guarantee that we will explore the entire phase space, while the relatively small perturbations to the operator induced by numerical discretization do not significantly alter the rate of search. Ultimately, pushing the timestep up is the only way to breach timescale gaps, although this comes at the cost of corruption to the long-time averages.

The self-diffusion coefficient gives a metric quantifying the diffusion rate. It is often used as a way to compare the rate of phase space exploration between methods, and typically calculated using the integral of the velocity auto-correlation function. However, arbitrary methods can be constructed to artificially scale the velocity auto-correlation function, hence giving inaccurate diffusion constants.

Indeed the calculation of a system's average temperature using the average of kinetic

energy (the equipartition theorem) by

$$Nk_BT = \left\langle p^T M^{-1} p \right\rangle, \tag{5.12}$$

gives a similar problem: methods can be created that give perfect sampling of this quantity but give comparatively poor sampling of configurations. For example, consider the results in Table 4.1 using the ⟦CBABC⟧ scheme with the harmonic oscillator. We see that we achieve exact results for (5.12) for any suitable step size δt, whereas configurational sampling suffers. Conversely for the ⟦ABCBA⟧ scheme, we have perfect configurational sampling, but poor results for the average temperature when computed using the momenta.

In the overdamped case, where momentum (if one chooses to use it) is resampled at each step, it is clearly not appropriate to use any averages of momentum as they are completely decoupled from the dynamics. Many alternative functions, including functions of q only, can be obtained whose averages are proportional to the system temperature [59, 64, 100]. For an arbitrary vector field $B : \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ we have

$$k_BT = \frac{\langle \nabla H \cdot B \rangle}{\langle \nabla \cdot B \rangle}, \tag{5.13}$$

where H is the system Hamiltonian and B is chosen such that the above numerator and denominator are nonzero and finite. Choosing $B = [0, p]^T$ gives exactly (5.12), while choosing $B = [q, 0]^T$ gives Clausius' virial theorem

$$Nk_BT = -\left\langle q \cdot \nabla U(q) \right\rangle, \tag{5.14}$$

which we shall refer to as the *configurational temperature*. Of course there are many other possibilities, a measure of configurational temperature is given in [17, 28] as

$$k_BT = \frac{\left\langle |\nabla U|^2 \right\rangle}{\langle \Delta U \rangle},$$

by taking $B = [\nabla U(q), 0]^T$ in (5.13). However, due to the extra computation required by evaluating the Laplacian at each step, as well as the relatively unpredictable denominator at large δt, we will use the virial theorem (5.14) to compute the configurational temperature of the system.

Were one able to solve the dynamics exactly, the kinetic and configurational temperatures would of course be commensurate. However, using a numerical method in the large-timestep regime we instead sample expectations with respect to the perturbed density $\hat{\rho}_{\delta t}$, that may introduce discrepancies between configurational and kinetic temperatures. A configuration-based temperature calculation is normally more useful and relevant for assessing the quality of configurational sampling methods.

In a similar way, the speed of exploration of the space should not be determined solely from functions of momentum, but should use actual barrier crossing rates or times

to reach some target region of phase space as a metronome to gauge the exploration rate.

5.3.2 Numerical results for vacuum simulation

Our initial test involves running the alanine dipeptide model without boundary conditions or solvent. This approach allows us to run a huge number of simulations in order to get a flavour for the performance of the algorithms.

We simulate the alanine dipeptide molecule at 300K for a fixed time interval of 2.5ns, using each of the eight second-order schemes. Multiple runs using different step sizes and friction constants are used to observe how different simulation parameters affect the computed averages. A region of the parameter space (where $(\delta t, \gamma) \in \mathbb{R}^2$) was divided into a 50×50 grid, with each point on the grid corresponding to a $(\delta t, \gamma)$ parameter set for a simulation. Values for the friction range from 0.01/ps to 100/ps while step size runs from 1fs to 3.29fs, with both parameter ranges increased by logarithmic increments. This parameter spectrum is considered broad enough to contain values that would realistically be used in computation, but was not so wide as to yield a large number of unsuitable parameter sets (for example, using an unstable step size) leading to wasted computation. All the schemes were unstable for the maximum step size tested, ensuring we gain some information about the relative stability of the schemes.

The results of the 20,000 total independent simulations are given in Figure 5.4, where we colour points on the 50×50 grid of parameter sets to indicate the results from that respective simulation. The colouring gives the relative error in the computed simulation's average, where the "exact" comparison value is taken from averaging ten 2.5ns runs using the ⟦BACAB⟧ scheme at $\delta t = 0.25$fs, where it is expected that discretization error is not significant.

We give errors of four averages–the average total potential energy and average total bond energy, as well as the error in the average temperature computed using kinetic data (5.12) or configuration data (5.14). As our focus is on configurational sampling, we include the kinetic temperature data to verify any correlation between errors in sampling q and p. Indeed, the results are similar to those of the harmonic oscillator in Table 4.1, with the ⟦BACAB⟧ and ⟦ABCBA⟧ schemes giving excellent results for the bond energy (which is quadratic in position), while the ⟦CABAC⟧ and ⟦CBABC⟧ schemes give superior results for the kinetic temperature (which is quadratic in momentum). Evidently judging one by the other is not an optimal strategy, and at worst would be misleading.

Simulating on such a short timescale, we would perhaps expect to see a very "noisy" result: high variances due to the sampling error vastly outweighing the discretization error. But in fact the discretization error dominates in most of the results and is plainly observable at step sizes significantly below the stability threshold. The results for the average total potential energy for the ⟦BACAB⟧, ⟦ABCBA⟧ and SPV schemes do show high variance due to sampling error, though we would expect this to decrease given

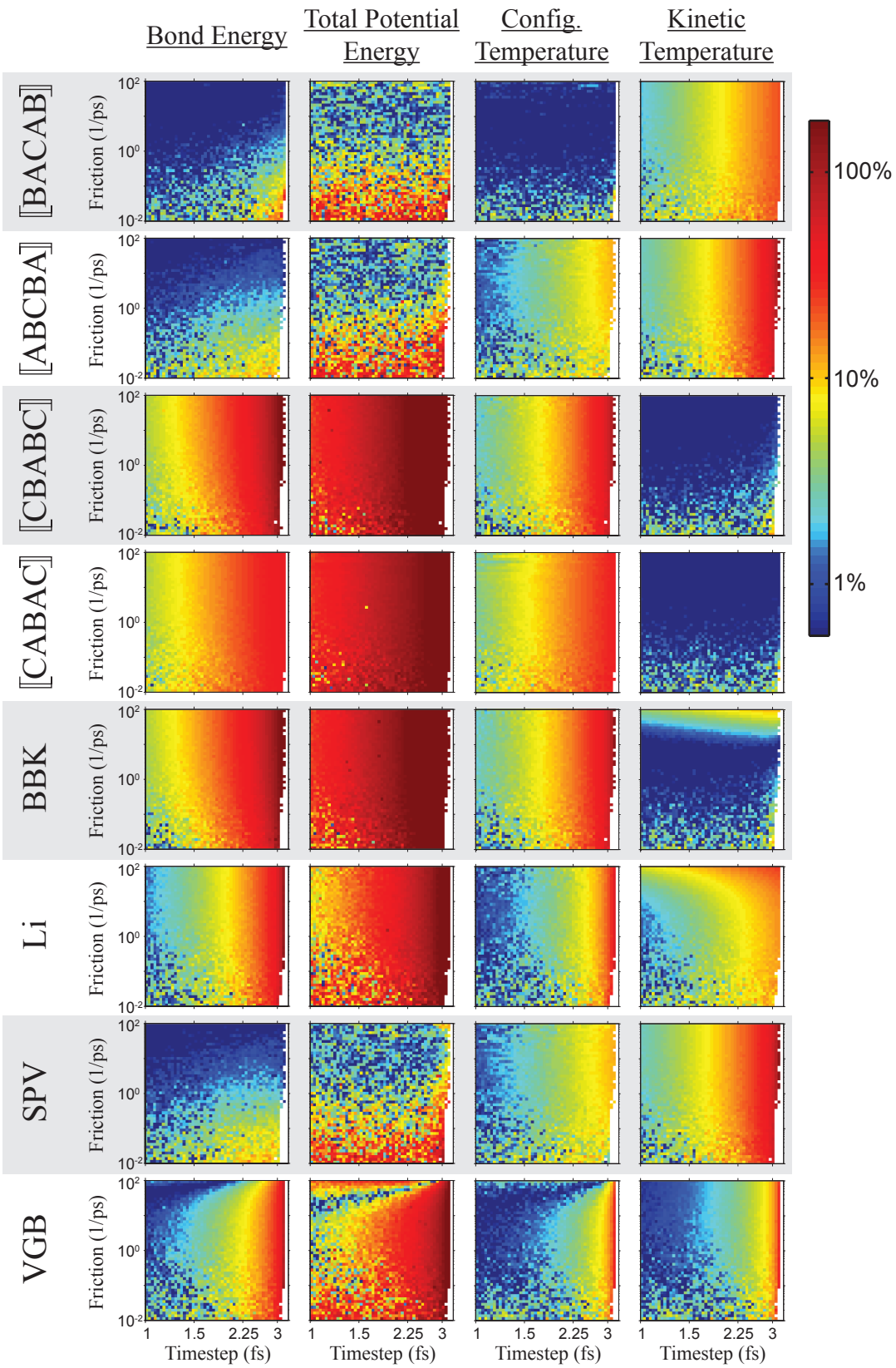


Figure 5.4: We simulate an alanine dipeptide molecule at 300K in a vacuum for 2.5ns using each of the eight schemes. Pixels colours denote the relative errors of computed averages in simulations with the prescribed timestep and friction parameters. White pixels denote instability.

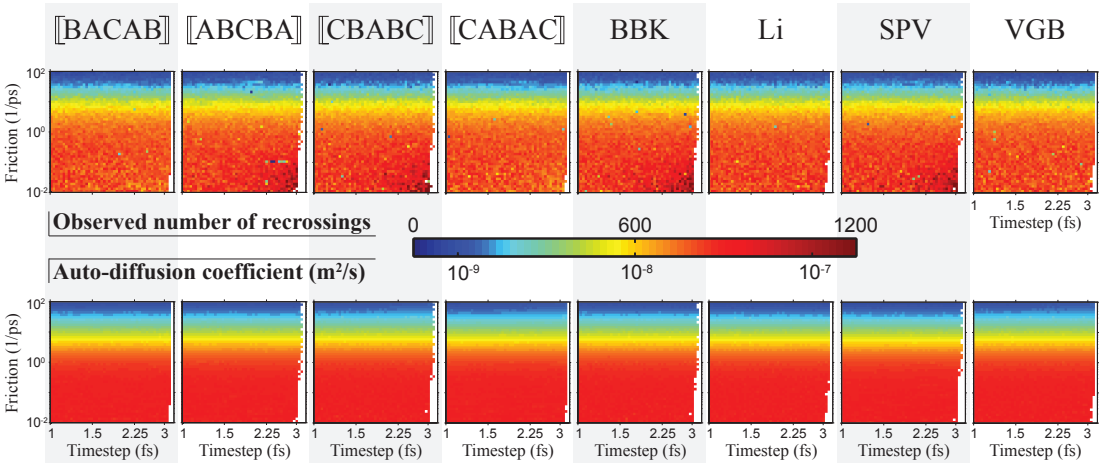


Figure 5.5: The rate of exploration through the phase space for the alanine dipeptide molecule in a vacuum is estimated by counting the number of orientation switches the particle makes (top) and computing the auto-diffusion coefficient using a Green-Kubo relation (bottom). We achieve similar results for all schemes tested.

additional repeat experiments. Alternatively we could use an interpolation scheme, as neighbouring gridpoints use have only small differences in their parameter sets. We elect to leave the noise as it is however, in favour of showing a 'clean' unfiltered result.

White pixels in each grid in Figure 5.4 represent parameters where the method is unstable, showing that in general there is a small stability threshold increase for the large-friction case, though there is no significant increase in this threshold between the methods. Recall that the threshold is set by the fastest degrees of freedom in this system, which in this case are the $N - H$ bond stretches. In the case of the harmonic oscillator, all of the consistent second order methods tested had the same step size limit equivalent to the Verlet step size threshold. Therefore we would not expect any significant differences in stability where the fastest motion is harmonic.

Of course a step size can be stable but unusable if we desire the error in averages below some tolerance. For example, if we require at most a 10% error in the computed average potential energy, then the BBK scheme is evidently unusable at any step size that we have considered, regardless of the time interval sampled over.

One salient feature of the results of Figure 5.4, is that for the ⟦BACAB⟧ scheme there is consistently less than a 1% error in the computed configurational temperature (for moderate friction) across all step sizes - even the largest stable timesteps tested. The relative errors obtained were so small that no discernible trend (with step size) can be shown, (due to the sampling error dominating below this value), whereas the other schemes tested show smooth error growth consistent with second-order schemes.

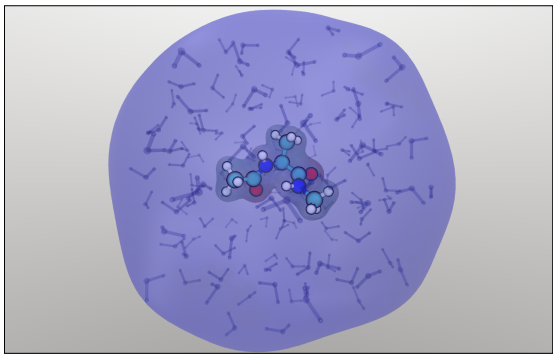
Self-diffusion coefficients are calculated from integrating the computed velocity autocorrelation function, where a history is kept of the velocities for 1ps. The values plotted in Figure 5.5 show that changing the step size within the indicated range us-

ing any of the schemes has only a very slight effect on the diffusion coefficient, while increasing the friction can dramatically reduce it. Examining the graphs, we settle on $\gamma = 1/\text{ps}$ as the largest value of γ for which the diffusion coefficient is unperturbed for all the schemes.

Additionally we compute a physical measure to compare the passage of time. We count the number of recrossings the molecule makes, which gives the number of times the orientation of the ϕ and ψ dihedral angles switch between their two states. For a fixed value of friction, for all methods this rate of recrossing remains constant, but is severely decreased for large values of γ above $1/\text{ps}$.

It is interesting that larger damping parameters do not substantially improve numerical stability for any of the methods, except in an extreme case for the VGB method ($\gamma \approx 100/\text{ps}$), where the diffusion constant is drastically reduced). From our analysis in Section 4.2 and Section 4.3, we would expect that for large friction the ⟦BACAB⟧ scheme should give a superconvergence property. It is likely that we do not see this effect due to the dominance of sampling error preventing any reliable statistics at sufficient accuracy.

5.3.3 Numerical results from simulations using solvent



In order to provide a more realistic comparison, we immerse the alanine dipeptide molecule in a sphere of TIP3P water (10Å radius, total system contains 424 atoms) and equilibrate for 1ns at 300K to generate an initial configuration. Simulations were completed using each scheme considered in the unsolvated case, using a 10Å cutoff for the electrostatic and van der Waals potentials. The value of friction was fixed at $1/\text{ps}$ in all simulations, given the optimal performance of this parameter in Figure 5.5. We test a range of timesteps, starting from $\delta t = 1\text{fs}$, with subsequent simulations increasing the step size by 5%, until reaching a step size where all of the methods fail. This was found after 22 such increments in the timestep, giving a maximum stable step size of approximately 2.8fs for this model. Each simulation runs for 5ns at 300K using spherical (harmonically restrained) boundary conditions to maintain the shape of the molecule.

In order to have a good baseline 'exact' result to compare our results, we average 10 comparable simulations using a 0.5fs step size. The baseline simulations were computed using the ⟦BACAB⟧ scheme, though we expect that using a discretization parameter so small means that sampling error should dominate over discretization error and hence the exact result should be effectively independent of scheme.

We plot the results of the growth in discretization error in Figure 5.6 for the 176 independent simulations. As this model is considerably larger than the vacuum model

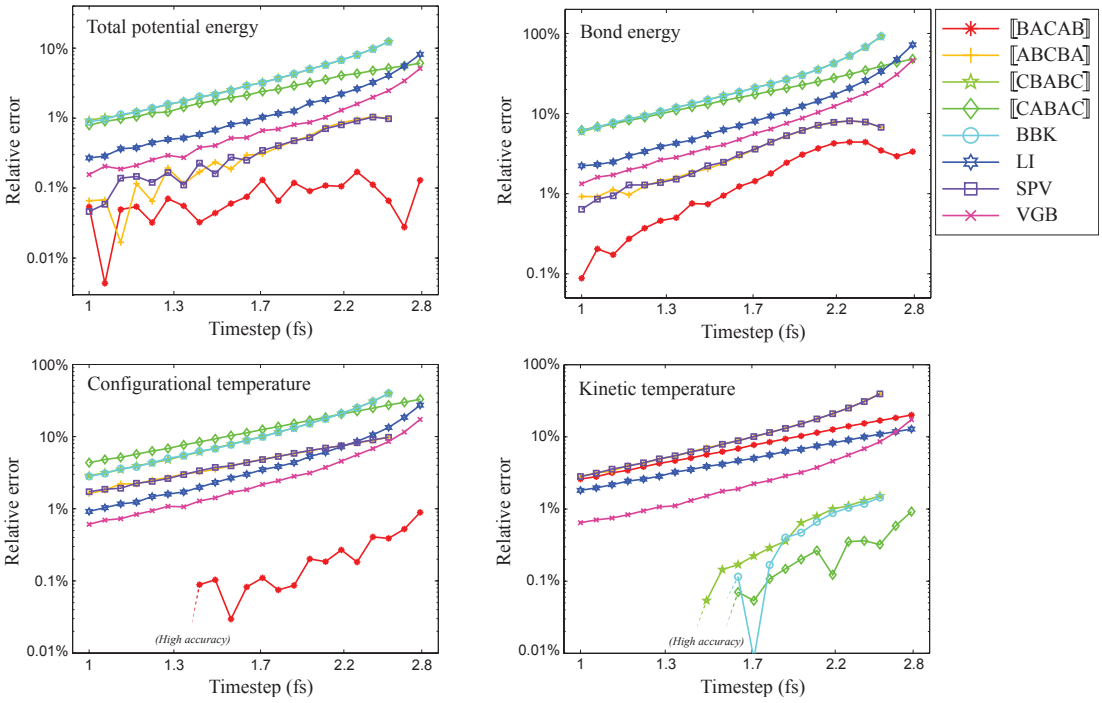


Figure 5.6: We plot the relative error of observed averages computed using each scheme for the solvated alanine dipeptide model with $\gamma = 1/\text{ps}$ and at $300K$. Errors are computed against a baseline solution averaged from ten 5ns simulations using the ⟦BACAB⟧ scheme at $\delta t = 0.5\text{fs}$. We omit some results at the smaller step sizes for the average temperature results due to the high accuracy of some schemes giving results with high variance.

considered in Section 5.3.2, it becomes prohibitively expensive to perform repeat experiments to drive down the variance. Averaging techniques such as importance sampling [75] or block averaging [74] could be employed to give greater accuracy in the results.

Instead, we recognize that the errors we record provide only an approximation of the overall behaviour. In the case of the average temperature results, some schemes give extremely accurate results of around 0.1% error below the step size of 2fs. At smaller step sizes, the sampling error becomes too large for any trend to be approximated.

Noticeably the ⟦BACAB⟧ scheme provides an order of magnitude improvement in the configurational temperature results at all step sizes tested. However, the results for its average kinetic temperature are rather poor, especially compared to its twin method ⟦CABAC⟧ produced from shuffling the terms around. This foregrounds the potentially misleading assumption of equivalence between the two types of sampling, as configurational sampling is of primary concern.

The surprising downward trend of the error using the ⟦BACAB⟧ scheme could be indicative of higher-order terms dominating in the error expansion, showing that our asymptotic approach does not give definitive answers about behaviour in the large step size regime. The analytic results obtained for are understood only for $\delta t \to 0$.

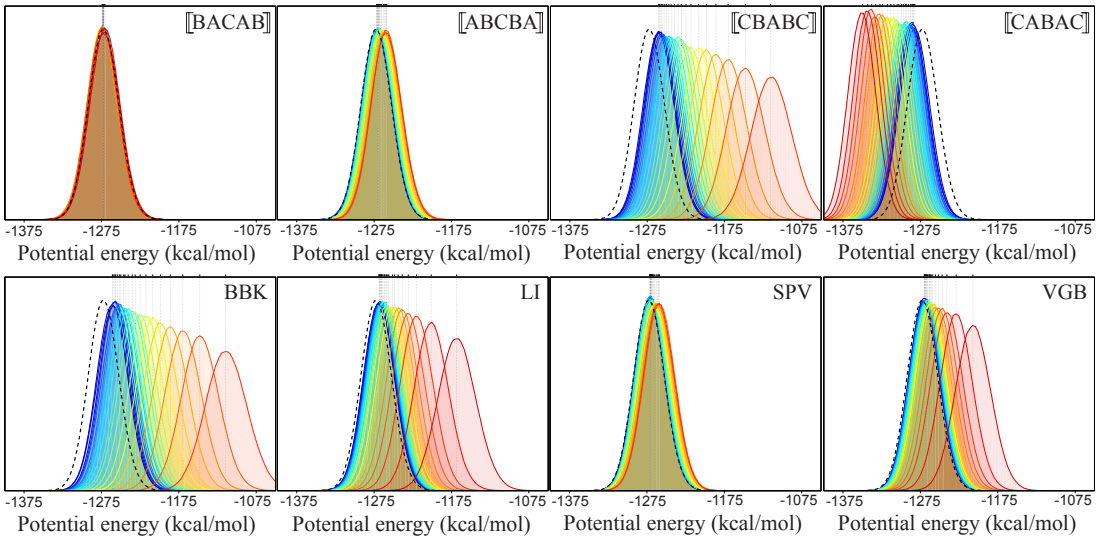


Figure 5.7: We plot the normalized potential energy distributions for simulations of solvated alanine dipeptide using each scheme, coloured for increasing step size (1fs in blue to 2.8fs in red) with successive lines indicating a 5% increase in the step size. Vertical dotted lines indicate the corresponding progression of the average total potential energy.

The greater restriction on the step size compared to the vacuum case (where schemes were stable above 3fs) is due to the introduction of the higher frequency O-H bonds. The slight differences in stability between some of the schemes is surprising as we have seen that in models where the step size threshold is governed by the harmonic terms all of the schemes have commensurate stability regions for fixed friction. It is likely that for multiple repeats, or a longer integration time, similar stability properties would be observed.

We additionally compute a histogram of the distribution of potential energy for each simulation, by dividing the interval $[-1500, -1000]$ (kcal/mol) into 150 bins and normalizing the resulting density. We compare the computed distributions with a baseline distribution computed using the same 'exact' results from Figure 5.6. We plot all of the resulting histograms in Figure 5.7, colouring the results according to the step size used to obtain them (increasing step size is marked from blue to red), overlaying the exact distribution in black.

The ⟦ABCBA⟧ and SPV schemes perform very well, with only a noticeable difference from the exact distribution as the step size approaches the stability limit. Some schemes, such as BBK or ⟦CABAC⟧ perform very poorly, with even the smallest step sizes tested giving a visible difference between the distributions. This indicates that the system is able to access much higher energy states than it would do in exact sampling. At the highest step sizes the exact and observed distributions have very little overlap, suggesting a complete failure to sample the system.

The distributions generated from trajectories computed using the ⟦BACAB⟧ scheme

deviate by only a very small amount, even at the highest step sizes tested. Comparing the energy breakdown, primarily the average bond and dihedral energy terms are responsible for the errors in average potential energy. From the results of the perturbed harmonic analysis in equation (4.66), the annihilation of the leading term in the error series (at order ϵ) appears to lead to the behaviour of the ⟦BACAB⟧ scheme. The parallel success of the ⟦ABCBA⟧ and SPV schemes (which share a similar update strategy, especially at small friction) support this observation, as such schemes also give excellent configurational sampling properties in (4.66).

The success of the ⟦BACAB⟧ scheme suggests that large improvements could be made to any methods that use Langevin dynamics as a tool for trajectory generation. Improving a method's efficiency (e.g. allowing the use of a larger basic timestep in simulation) thus has a knock-on effect on the efficiency of all the methods that rely on such trajectories. While relative improvements of a few percentages in efficiency can already warrant a minor change in software implementation, our analysis points to a more dramatic (even *qualitative*) difference among various methods leading to prospects for much greater efficiencies by selecting a suitable method.

The contribution of error coming from the restraining boundary condition energy was extremely small, suggesting that the properties of the bulk water in the model are responsible for the differences in efficiency seen here. Hence we would expect the observed corruption of averages to be generalizable to any simulations involving other boundary conditions, or other simulations involving water. Note that we would not expect similar results had we looked at the momentum distribution, but instead this behaviour is unique to configurational sampling.

Summary and conclusion

The research given in this thesis is devoted to explicit computation of the errors introduced through numerical discretization of Langevin dynamics, with a focus on the computed trajectories used to sample canonical configurational averages of a molecular dynamics system..

We have analyzed and tested four new second-order Langevin dynamics discretization schemes that require one force evaluation per timestep, based upon a new (ABC) splitting strategy for the SDEs. In a molecular dynamics simulation of a solvated protein, one of the new schemes was found to significantly outperform all other methods in tests involving configurational sampling. At large step size, the discretization error was found to dominate the random sampling error for typical integrators found in standard MD packages.

This has a potentially large impact on many disciplines that rely on molecular dynamics simulation. If simulations can significantly increase the timestep without compromising on the sampling quality, then either the required wall-clock simulation time is dramatically decreased, or it enables a longer the integration interval to sample over.

We make rigorous the error estimates given in [68], and make precise statements on the relations between errors in the observed averages of different schemes, solving to find the leading correction terms in the case of the first order methods. In the case of the second order methods, we use a linear response technique to numerically estimate the required correction. Alternatively we discuss methods of modifying the Langevin dynamics equations to provide higher-order sampling tuned to specific observables.

Schemes for the overdamped (infinite friction) limit of Langevin dynamics (or Brownian dynamics) are studied as the limiting cases of methods splitting the Ornstein-Uhlenbeck process up from the Newtonian dynamics. We solve to find the correction term for these overdamped schemes, and prove that one scheme has a superconvergence property in this limit. The limiting method of this scheme gives a fourth-order error in observed averages, while requiring only one force evaluation per timestep.

There are a wide variety of future directions for this work. The framework we use can be naturally extended to provide explicit weak error estimates for SDEs other than Langevin dynamics (such as Nosé-Hoover Langevin dynamics (3.15) or line sampling (3.17)), assuming similar hypocoercivity results hold allowing us to bound the remain-

der terms. Potentially this would allow estimation of weak sampling errors between schemes that integrate different dynamics, although we would not necessarily expect to find the cancellations at higher order that we demonstrate in some Langevin dynamics schemes.

The results of perturbing the Langevin equation in Section 5.2 suggest that the framework presented could be extended to include schemes that compute averages of non-equilibrium systems, which amounts to perturbing the force in a similar manner. This could allow greater fidelity in the computation as static quantities such as the diffusion coefficient.

To overcome some numerical stability issues in simulation, holonomic constraints are sometimes used in molecular dynamics to fix bond lengths or angles between nuclei. It is not obvious whether discretization schemes for Langevin dynamics with constraints that use a splitting analagous to those presented in this thesis inherit the same properties as discretization methods without constraints (e.g. exact sampling of harmonic potentials). This is particularly relevant for the method with the superconvergence property.

We conclude by stating that the ⟦BACAB⟧ scheme, and its limiting method in (4.73), demonstrate excellent configurational sampling properties in both experiment and analysis. The scheme shows great potential to significantly improve the weak accuracy of many algorithms employing Langevin dynamics to sample molecular dynamics systems.

Bibliography

[1] A. Abdulle, D. Cohen, G. Vilmart, and K. Zygalakis. High weak order methods for stochastic differential equations based on modified equations. *SIAM Journal on Scientific Computing*, 34(3):A1800–A1823, 2012.

[2] B. J. Alder and T. E. Wainwright. Phase transition for a hard sphere system. *The Journal of Chemical Physics*, 27(5):1208–1209, 1957.

[3] B. J. Alder and T. E. Wainwright. Studies in molecular dynamics. I. General method. *The Journal of Chemical Physics*, 31(2):459–466, 1959.

[4] P. Allen and D. Tildesley. *Computer Simulation of Liquids*. Oxford Science Publications. Clarendon Press, 1989.

[5] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384–2393, 1980.

[6] V. Arnold. *Mathematical methods of classical mechanics*. Graduate texts in mathematics. Springer-Verlag, 1978.

[7] M. Athénes. A path-sampling scheme for computing thermodynamic properties of a many-body system in a generalized ensemble. *The European Physical Journal B - Condensed Matter and Complex Systems*, 38:651–663, 2004.

[8] E. Barth, B. Leimkuhler, and C. Sweet. Approach to thermal equilibrium in biomolecular simulation. In B. Leimkuhler, C. Chipot, R. Elber, A. Laaksonen, A. Mark, T. Schlick, C. Schtte, and R. Skeel, editors, *New Algorithms for Macromolecular Simulation*, volume 49 of *Lecture Notes in Computational Science and Engineering*, pages 125–140. Springer Berlin Heidelberg, 2006.

[9] J. J. Biesiadecki and R. D. Skeel. Dangers of multiple time step methods. *Journal of Computational Physics*, 109:318–328, Dec. 1993.

[10] P. Bolhuis, D. Chandler, C. Dellago, and P. Geissler. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annual review of physical chemistry*, 53(1):291–318, 2002.

[11] S. D. Bond and B. J. Leimkuhler. Molecular dynamics and the accuracy of numerically computed averages. *Acta Numerica*, 16:1–65, 4 2007.

[12] N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM Journal on Numerical Analysis*, 48(1):278–297, 2010.

[13] N. Bou-Rabee and E. Vanden-Eijnden. Pathwise accuracy and ergodicity of metropolized integrators for SDEs. *Communications on Pure and Applied Mathematics*, 63(5):655–696, 2010.

[14] A. Brünger, C. L. Brooks, and M. Karplus. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chemical Physics Letters*, 105(5):495–500, 1984.

[15] K. Burrage and G. Lythe. Accurate stationary densities with partitioned numerical methods for stochastic differential equations. *SIAM J. Numer. Anal.*, 47(3):1601–1618, Apr. 2009.

[16] G. Bussi and M. Parrinello. Accurate sampling using Langevin dynamics. *Phys. Rev. E*, 75:056707, May 2007.

[17] B. D. Butler, G. Ayton, O. G. Jepps, and D. J. Evans. Configurational temperature: Verification of monte carlo simulations. *The Journal of Chemical Physics*, 109(16):6519–6522, 1998.

[18] M. Calvo and J. Sanz-Serna. The development of variable-step symplectic integrators, with application to the two-body problem. *SIAM Journal on Scientific Computing*, 14(4):936–952, 1993.

[19] J. Candy and W. Rozmus. A symplectic integration algorithm for separable Hamiltonian functions. *Journal of Computational Physics*, 92(1):230 – 256, 1991.

[20] M. Ceriotti, G. Bussi, and M. Parrinello. Colored-noise thermostats á la carte. *Journal of Chemical Theory and Computation*, 6(4):1170–1180, 2010.

[21] P. J. Channell. Symplectic integration algorithms. Technical Report AT–6:ATN 83–9, Los Alamos National Laboratory, 1983.

[22] P. J. Channell and C. Scovel. Symplectic integration of Hamiltonian systems. *Nonlinearity*, 3(2):231, 1990.

[23] T. W. Clark, J. A. McCammon, and L. R. Scott. Parallel molecular dynamics. In *Proceedings of the Fifth SIAM Conference on Parallel Processing for Scientific Computing*, pages 338–344, Philadelphia, PA, USA, 1992. Society for Industrial and Applied Mathematics.

[24] V. Daggett. Long timescale simulations. *Current Opinion in Structural Biology*, 10(2):160 – 164, 2000.

[25] R. L. Davidchack. Discretization errors in molecular dynamics simulations with deterministic and stochastic thermostats. *Journal of Computational Physics*, 229(24):9323–9346, 2010.

[26] R. de Vogelaere. Methods of integration which preserve the contact transformation property of the Hamiltonian equations. Technical Report 4, Department of Mathematics, University of Notre Dame, 1956.

[27] A. Debussche and E. Faou. Weak backward error analysis for SDEs. *SIAM Journal on Numerical Analysis*, 50(3):1735–1752, 2012.

[28] J. Delhommelle and D. J. Evans. Correspondence between configurational temperature and molecular kinetic temperature thermostats. *The Journal of Chemical Physics*, 117(13):6016–6021, 2002.

[29] S. Duane, A. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid monte carlo. *Physics Letters B*, 195(2):216–222, 1987.

[30] P. Eastman and S. Doniach. Multiple time step diffusive Langevin dynamics for proteins. *Proteins: Structure, Function, and Bioinformatics*, 30(3):215–227, 1998.

[31] D. L. Ermak and H. Buckholz. Numerical integration of the Langevin equation: Monte carlo simulation. *Journal of Computational Physics*, 35(2):169–182, 1980.

[32] D. L. Ermak and J. A. McCammon. Brownian dynamics with hydrodynamic interactions. *The Journal of Chemical Physics*, 69:1352–1360, 1978.

[33] J. J. Erpenbeck and W. W. Wood. Molecular-dynamics calculations of the velocity autocorrelation function: Hard-sphere results. *Phys. Rev. A*, 32:412–422, Jul 1985.

[34] H. Esmaeilzadeh, E. Blem, R. S. Amant, K. Sankaralingam, and D. Burger. Power challenges may end the multicore era. *Commun. ACM*, 56(2):93–102, Feb. 2013.

[35] D. J. Evans and B. L. Holian. The Nosé–Hoover thermostat. *The Journal of Chemical Physics*, 83(8):4069–4074, 1985.

[36] K. Feng. On difference schemes and symplectic geometry. In *Proceedings of the 1984 Beijing Symposium on Differential Geometry and Differential Equations*, pages 42–58, Beijing, 1985. Science Press.

[37] K. Feng. Difference schemes for Hamiltonian formalism and symplectic geometry. volume 4, pages 279–289. Science Press, Utrecht, The Netherlands, The Netherlands, July 1986.

[38] K. Feng. Symplectic geometry and numerical methods in fluid dynamics. In F. Zhuang and Y. Zhu, editors, *Tenth International Conference on Numerical Methods in Fluid Dynamics*, volume 264 of *Lecture Notes in Physics*, pages 1–7. Springer Berlin Heidelberg, 1986.

[39] E. Forest and R. D. Ruth. Fourth-order symplectic integration. *Physica D: Nonlinear Phenomena*, 43(1):105 – 117, 1990.

[40] R. F. Fox. Long-time tails and diffusion. *Phys. Rev. A*, 27:3216–3233, Jun 1983.

[41] M. Freidlin. Some remarks on the Smoluchowski-Kramers approximation. *Journal of Statistical Physics*, 117(3-4):617–634, 2004.

[42] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Computational science series. Elsevier Science, 2001.

[43] B. Garca-Archilla, J. Sanz-Serna, and R. Skeel. Long-time-step methods for oscillatory differential equations. *SIAM Journal on Scientific Computing*, 20(3):930–963, 1998.

[44] C. Gardiner. *Handbook of stochastic methods for physics, chemistry, and the natural sciences*. Springer series in synergetics. Springer, 1994.

[45] T. C. Germann and K. Kadau. Trillion-atom molecular dynamics becomes a reality. *International Journal of Modern Physics C*, 19(09):1315–1319, 2008.

[46] J. B. Gibson, A. N. Goland, M. Milgram, and G. H. Vineyard. Dynamics of radiation damage. *Phys. Rev.*, 120:1229–1253, Nov 1960.

[47] N. Go. Theoretical studies of protein folding. *Annual Review of Biophysics and Bioengineering*, 12(1):183–210, 1983. PMID: 6347038.

[48] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics. Springer, second edition, 2006.

[49] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for markov chains. In R. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progress in Probability*, pages 109–117. Springer Basel, 2011.

[50] M. Hairer and G. Pavliotis. Periodic homogenization for hypoelliptic diffusions. *Journal of Statistical Physics*, 117(1-2):261–279, 2004.

[51] M. Hairer and G. Pavliotis. From ballistic to diffusive behavior in periodic potentials. *Journal of Statistical Physics*, 131(1):175–202, 2008.

[52] U. H. Hansmann and Y. Okamoto. New monte carlo algorithms for protein folding. *Current Opinion in Structural Biology*, 9(2):177 – 183, 1999.

[53] T. Hansson, C. Oostenbrink, and W. van Gunsteren. Molecular dynamics simulations. *Current Opinion in Structural Biology*, 12(2):190 – 196, 2002.

[54] D. J. Hardy. NAMD-Lite. University of Illinois at Urbana-Champaign, `http://www.ks.uiuc.edu/Development/MDTools/namdlite/`, 2007.

[55] W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, Mar 1985.

[56] L. Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119:147–171, 1967.

[57] J. Izaguirre, C. Sweet, and V. Pande. Multiscale dynamics of macromolecules using normal mode Langevin. In *Pac Symp Biocomput*, volume 15, pages 240–251, 2010.

[58] J. A. Izaguirre, Q. Ma, T. Matthey, J. Willcock, T. Slabach, B. Moore, and G. Viamontes. Overcoming instabilities in Verlet-I/r-RESPA with the mollified impulse method. In T. Schlick and H. Gan, editors, *Computational Methods for Macromolecules: Challenges and Applications*, volume 24 of *Lecture Notes in Computational Science and Engineering*, pages 146–174. Springer Berlin Heidelberg, 2002.

[59] O. G. Jepps, G. Ayton, and D. J. Evans. Microscopic expressions for the thermodynamic temperature. *Phys. Rev. E*, 62:4757–4763, 2000.

[60] R. Joubaud and G. Stoltz. Nonequilibrium shear viscosity computations with Langevin dynamics. *Multiscale Modeling & Simulation*, 10(1):191–216, 2012.

[61] K. Kadau, T. C. Germann, and P. S. Lomdahl. Large-scale molecular-dynamics simulation of 19 billion particles. *International Journal of Modern Physics C*, 15(01):193–201, 2004.

[62] S. Kirmizialtin and R. Elber. Revisiting and computing reaction coordinates with directional milestoning. *The Journal of Physical Chemistry A*, 115(23):6137–6148, 2011.

[63] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of Mathematics. Springer, 1992.

[64] L. D. Landau and E. M. Lifshitz. *Statistical Physics*. Nakura, 1952.

[65] L. Larini, R. Mannella, and D. Leporini. Langevin stabilization of molecular-dynamics simulations of polymers by means of quasisymplectic algorithms. *The Journal of Chemical Physics*, 126(10):104101, 2007.

[66] F. Legoll, M. Luskin, and R. Moeckel. Non-ergodicity of the Nosé-Hoover thermostatted harmonic oscillator. *Archive for Rational Mechanics and Analysis*, 184(3):449–463, 2007.

[67] B. Leimkuhler. Generalized Bulgac-Kusnezov methods for sampling of the Gibbs-Boltzmann measure. *Phys. Rev. E*, 81:026703, Feb 2010.

[68] B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Applied Mathematics Research eXpress*, 2013(1):34–56, 2013.

[69] B. Leimkuhler and C. Matthews. Robust and efficient configurational molecular sampling via Langevin dynamics. *The Journal of Chemical Physics*, 138(17):174102, 2013.

[70] B. Leimkuhler, E. Noorizadeh, and F. Theil. A gentle stochastic thermostat for molecular dynamics. *Journal of Statistical Physics*, 135(2):261–277, 2009.

[71] B. Leimkuhler and S. Reich. A reversible averaging integrator for multiple time-scale dynamics. *Journal of Computational Physics*, 171(1):95 – 114, 2001.

[72] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2005.

[73] T. Lelièvre, M. Rousset, and G. Stoltz. Langevin dynamics with constraints and computation of free energy differences. *Mathematics of Computation*, 81(280):2071, 2012.

[74] T. Lelièvre, G. Stoltz, and M. Rousset. *Free Energy Computations: A Mathematical Perspective*. Imperial College Press, 2010.

[75] J. S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer, corrected edition, Jan. 2008.

[76] M. López-Marcos, J. Sanz-Serna, and R. Skeel. Explicit symplectic integrators using hessian–vector products. *SIAM Journal on Scientific Computing*, 18(1):223–238, 1997.

[77] Q. Ma, J. Izaguirre, and R. Skeel. Verlet-I/r-RESPA/Impulse is limited by nonlinear instabilities. *SIAM Journal on Scientific Computing*, 24(6):1951–1973, 2003.

[78] G. J. Martyna, M. L. Klein, and M. Tuckerman. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *The Journal of Chemical Physics*, 97(4):2635–2643, 1992.

[79] J. Mattingly, A. Stuart, and M. Tretyakov. Convergence of numerical time-averaging and stationary measures via poisson equations. *SIAM Journal on Numerical Analysis*, 48(2):552–577, 2010.

[80] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stoch. Proc. Appl.*, 101(2):185–232, 2002.

[81] E. Meerbach, E. Dittmer, I. Horenko, and C. Schütte. Multiscale modelling in molecular dynamics: Biomolecular conformations as metastable states. *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*, pages 495–517, 2006.

[82] S. Melchionna. Design of quasisymplectic propagators for Langevin dynamics. *The Journal of Chemical Physics*, 127(4):044108, 2007.

[83] C. R. Menyuk. Some properties of the discrete Hamiltonian method. *Physica D: Nonlinear Phenomena*, 11(12):109 – 129, 1984.

[84] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[85] S. S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer, London, 1993.

[86] G. Milstein and M. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computation. Springer, 2004.

[87] G. N. Milstein and M. V. Tretyakov. Quasisymplectic methods for Langevin-type equations. *IMA Journal of Numerical Analysis*, 23(4):593–626, 2003.

[88] F. Nier and B. Helffer. *Hypoelliptic Estimates and Spectral Theory for Fokker-Planck Operators and Witten Laplacians*. Springer, 2005.

[89] S. Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2):255–268, 1984.

[90] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–519, 1984.

[91] B. Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Hochschultext / Universitext. Springer, 2003.

[92] D. Okunbor and R. Murty. Parallel molecular dynamics using force decomposition. In P. Deuflhard, J. Hermans, B. Leimkuhler, A. Mark, S. Reich, and R. Skeel,

editors, *Computational Molecular Dynamics: Challenges, Methods, Ideas*, volume 4 of *Lecture Notes in Computational Science and Engineering*, pages 483–494. Springer Berlin Heidelberg, 1999.

[93] M. Ottobre and G. A. Pavliotis. Asymptotic analysis for the generalized Langevin equation. *Nonlinearity*, 24(5):1629, 2011.

[94] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.

[95] S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1 – 19, 1995.

[96] A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev.*, 136:A405–A411, Oct 1964.

[97] D. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2004.

[98] L. Rey-Bellet. Ergodic properties of markov processes. In S. Attal, A. Joye, and C.-A. Pillet, editors, *Open Quantum Systems II*, volume 1881 of *Lecture Notes in Mathematics*, pages 1–39. Springer, 2006.

[99] G. Rowlands. A numerical algorithm for Hamiltonian systems. *Journal of Computational Physics*, 97(1):235 – 239, 1991.

[100] H. H. Rugh. Dynamical approach to temperature. *Phys. Rev. Lett.*, 78:772–774, Feb 1997.

[101] R. D. Ruth. A canonical integration technique. *Nuclear Science, IEEE Transactions on*, 30(4):2669–2671, 1983.

[102] J.-P. Ryckaert, G. Ciccotti, and H. J. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3):327 – 341, 1977.

[103] A. Samoletov, C. Dettmann, and M. Chaplain. Thermostats for slow configurational modes. *Journal of Statistical Physics*, 128(6):1321–1336, 2007.

[104] J. M. Sanz-Serna and P. Calvo. *Numerical Hamiltonian problems*. Applied mathematics and mathematical computation. Chapman & Hall, 1994.

[105] T. Schlick. Time-trimming tricks for dynamic simulations: splitting force updates to reduce computational work. *Structure (London, England : 1993)*, 9(4), Apr. 2001.

[106] T. Schlick. *Molecular Modeling and Simulation: An Interdisciplinary Guide*. Interdisciplinary applied mathematics. Springer Science+Business Media, LLC, 2010.

[107] T. Shardlow. Splitting for dissipative particle dynamics. *SIAM Journal on Scientific Computing*, 24(4):1267–1282, 2003.

[108] T. Shardlow. Modified equations for stochastic differential equations. *BIT Numerical Mathematics*, 46(1):111–125, 2006.

[109] D. A. Sivak, J. D. Chodera, and G. E. Crooks. Using nonequilibrium fluctuation theorems to understand and correct errors in equilibrium and nonequilibrium simulations of discrete Langevin dynamics. *Phys. Rev. X*, 3:011007, Jan 2013.

[110] R. Skeel. What makes molecular dynamics work? *SIAM Journal on Scientific Computing*, 31(2):1363–1378, 2009.

[111] R. D. Skeel and J. A. Izaguirre. An impulse integrator for Langevin dynamics. *Molecular Physics*, 100(24):3885–3891, 2002.

[112] E. Smargiassi and P. A. Madden. Free-energy calculations in solids from first-principles molecular dynamics: Vacancy formation in sodium. *Phys. Rev. B*, 51:117–128, Jan 1995.

[113] Y. Suris. Some properties of methods for the numerical integration of systems of the form x = f(x). *USSR Computational Mathematics and Mathematical Physics*, 27(5):149 – 156, 1987.

[114] D. Talay. *Simulation and Numerical Analysis of Stochastic Differential Systems: A Review*. Rapports de recherche. Institut National de Recherche en Informatique et en Automatique, 1990.

[115] D. Talay. Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Proc. Rel. Fields*, 8:163–198, 2002.

[116] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4):483–509, 1990.

[117] F. Thalmann and J. Farago. Trotter derivation of algorithms for Brownian and dissipative particle dynamics. *The Journal of Chemical Physics*, 127(12):124109, 2007.

[118] K. P. Travis and C. Braga. Configurational temperature control for atomic and molecular systems. *The Journal of Chemical Physics*, 128(1):014111, 2008.

[119] M. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics*, 97(3):1990–2001, 1992.

[120] M. E. Tuckerman, B. J. Berne, and G. J. Martyna. Molecular dynamics algorithm for multiple time scales: Systems with long range forces. *The Journal of Chemical Physics*, 94(10):6811–6815, 1991.

[121] M. E. Tuckerman and M. Parrinello. Integrating the Car–Parrinello equations. I. Basic integration techniques. *The Journal of Chemical Physics*, 101(2):1302–1315, 1994.

[122] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen. Gromacs: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16):1701–1718, 2005.

[123] W. van Gunsteren and H. Berendsen. Algorithms for Brownian dynamics. *Molecular Physics*, 45(3):637–647, 1982.

[124] W. van Gunsteren and H. Berendsen. A leap-frog algorithm for stochastic dynamics. *Molecular Simulation*, 1(3):173–185, 1988.

[125] C. Villani. *Hypocoercivity*. American Mathematical Soc.

[126] H. Weyl. *The Classical Groups: Their Invariants and Representations*. Princeton landmarks in mathematics and physics / Princeton landmarks in mathematics and physics. Princeton University Press, 1997.

[127] W. W. Wood and J. J. Erpenbeck. Molecular dynamics and monte carlo calculations in statistical mechanics. *Annual Review of Physical Chemistry*, 27(1):319–348, 1976.

[128] W. W. Wood and F. R. Parker. Monte carlo equation of state of molecules interacting with the Lennard-Jones potential. I. A supercritical isotherm at about twice the critical temperature. *The Journal of Chemical Physics*, 27(3):720–733, 1957.

[129] H. Yoshida. Construction of higher order symplectic integrators. *Physics Letters A*, 150(57):262 – 268, 1990.

[130] K. Zygalakis. On the existence and the applications of modified equations for stochastic differential equations. *SIAM Journal on Scientific Computing*, 33(1):102–130, 2011.

Appendix A

Numerical Methods

We give the numerical methods used in this thesis, assuming timestep δt, friction constant γ with diagonal mass matrix M and position and momentum vectors $q, p \in \mathbb{R}^N$ respectively. The force is $F(q) := -\nabla U(q)$, where we use temperature T and k_B is Boltzmann's constant. R_n is an N-vector of independent, identically distributed normal random numbers with zero mean and unit variance. Where $J > 1$ random numbers are required per degree of freedom, multiple independent (uncorrelated) random vectors are denoted $R_n^{(j)}, j = 1, \ldots, J$ in the schemes.

$$-- \quad [\![\text{BACAB}]\!] \quad --$$

Note: This scheme is available in recent versions of NAMD (after Jan 2013) by including options '*langevin on*' and '*langevinBAOAB on*' in the input parameter file.

$$\begin{aligned}
p_{n+1/3} &= p_n + \frac{\delta t}{2} F(q_n), \\
q_{n+1/2} &= q_n + \frac{\delta t}{2} M^{-1} p_{n+1/3}, \\
p_{n+2/3} &= e^{-\gamma \delta t} p_{n+1/3} + \sqrt{k_B T \left(1 - e^{-2\gamma \delta t}\right)} M^{1/2} R_n, \\
q_{n+1} &= q_{n+1/2} + \frac{\delta t}{2} M^{-1} p_{n+2/3}, \\
p_{n+1} &= p_{n+2/3} + \frac{\delta t}{2} F(q_{n+1})
\end{aligned}$$

$$-- \quad [\![\text{ABCBA}]\!] \quad --$$

$$\begin{aligned}
q_{n+1/2} &= q_n + \frac{\delta t}{2} M^{-1} p_n, \\
p_{n+1/3} &= p_n + \frac{\delta t}{2} F(q_{n+1/2}), \\
p_{n+2/3} &= e^{-\gamma \delta t} p_{n+1/3} + \sqrt{k_B T \left(1 - e^{-2\gamma \delta t}\right)} M^{1/2} R_n, \\
p_{n+1} &= p_{n+2/3} + \frac{\delta t}{2} F(q_{n+1/2}), \\
q_{n+1} &= q_{n+1/2} + \frac{\delta t}{2} M^{-1} p_{n+1}
\end{aligned}$$

$$-- \quad [\![\text{CBABC}]\!] \quad --$$

$$\begin{aligned}
p_{n+1/4} &= e^{-\gamma\delta t/2} p_n + \sqrt{k_B T\,(1 - e^{-\gamma\delta t})}\, M^{1/2} R_n^{(1)}, \\
p_{n+2/4} &= p_{n+1/4} + \frac{\delta t}{2} F(q_n), \\
q_{n+1} &= q_n + \delta t M^{-1} p_{n+2/4}, \\
p_{n+3/4} &= p_{n+2/4} + \frac{\delta t}{2} F(q_{n+1}), \\
p_{n+1} &= e^{-\gamma\delta t/2} p_{n+3/4} + \sqrt{k_B T\,(1 - e^{-\gamma\delta t})}\, M^{1/2} R_n^{(2)}
\end{aligned}$$

$$-- \quad [\![\text{CABAC}]\!] \quad --$$

$$\begin{aligned}
p_{n+1/3} &= e^{-\gamma\delta t/2} p_n + \sqrt{k_B T\,(1 - e^{-\gamma\delta t})}\, M^{1/2} R_n^{(1)}, \\
q_{n+1/2} &= q_n + \frac{\delta t}{2} M^{-1} p_{n+1/3}, \\
p_{n+2/3} &= p_{n+1/3} + \delta t F(q_{n+1/2}), \\
q_{n+1} &= q_{n+1/2} + \frac{\delta t}{2} M^{-1} p_{n+2/3}, \\
p_{n+1} &= e^{-\gamma\delta t/2} p_{n+2/3} + \sqrt{k_B T\,(1 - e^{-\gamma\delta t})}\, M^{1/2} R_n^{(2)}
\end{aligned}$$

$$-- \quad \text{Van Gunsteren/Berendsen (VGB)} \quad --$$

We must initialize the vector $X \in \mathbb{R}^N$,

$$X_1 = \kappa_4 M^{-1/2} R_0^{(3)},$$

and then iterate

$$\begin{aligned}
V_{n+1} &= \kappa_1 M^{-1/2} R_n^{(1)}, \\
\hat{V}_{n+1} &= \kappa_2 X_n + \kappa_3 M^{-1/2} R_n^{(2)}, \\
p_{n+1} &= e^{-\gamma\delta t} p_n + \frac{1 - e^{-\gamma\delta t}}{\gamma} F(q_n) + M\left(V_{n+1} - e^{-\gamma\delta t}\hat{V}_{n+1}\right), \\
X_{n+1} &= \kappa_4 M^{-1/2} R_n^{(3)}, \\
\hat{X}_{n+1} &= \kappa_5 V_{n+1} + \kappa_6 M^{-1/2} R_n^{(4)}, \\
q_{n+1} &= \frac{e^{\gamma\delta t/2} - e^{-\gamma\delta t/2}}{\gamma} M^{-1} p_{n+1} + X_{n+1} - \hat{X}_{n+1},
\end{aligned}$$

where we use vectors $X, \hat{X}, V, \hat{V} \in \mathbb{R}^N$. Constants $\kappa_i \in \mathbb{R}$ are given as

$$\begin{aligned}
\kappa_1 &= \sqrt{k_B T \left(1 - e^{-\gamma \delta t}\right)}, \\
\kappa_2 &= \frac{2\gamma - \gamma e^{\gamma \delta t/2} - \gamma e^{-\gamma \delta t/2}}{\gamma \delta t - 3 + e^{-\gamma \delta t}\left(4 e^{\gamma \delta t/2} - 1\right)}, \\
\kappa_3 &= \sqrt{k_B T} \sqrt{\frac{\gamma \delta t \left(e^{\gamma \delta t} - 1\right) - 4\left(e^{\gamma \delta t/2} - 1\right)^2}{\gamma \delta t - 3 + e^{-\gamma \delta t}\left(4 e^{\gamma \delta t/2} - 1\right)}}, \\
\kappa_4 &= \gamma^{-1} \sqrt{k_B T} \sqrt{\gamma \delta t - 3 + e^{-\gamma \delta t}\left(4 e^{\gamma \delta t/2} - 1\right)}, \\
\kappa_5 &= \gamma^{-1} \left(\frac{2 - e^{\gamma \delta t} - e^{-\gamma \delta t}}{e^{-2\gamma \delta t} - 1}\right), \\
\kappa_6 &= \gamma^{-1} \sqrt{k_B T} \sqrt{\frac{\gamma \delta t \left(e^{-\gamma \delta t} - 1\right) + 4\left(e^{-\gamma \delta t/2} - 1\right)^2}{e^{-\gamma \delta t} - 1}}.
\end{aligned}$$

$$-- \quad \text{Stochastic Position Verlet (SPV)} \quad --$$

$$\begin{aligned}
q_{n+1/2} &= q_n + \frac{\delta t}{2} M^{-1} p_n, \\
p_{n+1} &= e^{-\gamma \delta t} p_n + \frac{1 - e^{-\gamma \delta t}}{\gamma} F(q_{n+1/2}) + \sqrt{k_B T \left(1 - e^{-2\gamma \delta t}\right)} M^{1/2} R_n, \\
q_{n+1} &= q_{n+1/2} + \frac{\delta t}{2} M^{-1} p_{n+1}
\end{aligned}$$

$$-- \quad \text{Langevin Impulse (LI)} \quad --$$

We use the algorithm designed for configurational sampling; a correction term is given in [111] to improve the sampling of momenta, though this has no effect on configurational averages. We must initialize the vector $Z \in \mathbb{R}^N$,

$$Z_1 = M^{1/2}\left(\alpha_0 R_0 + \hat{\alpha} R_1\right).$$

and then iterate

$$\begin{aligned}
p_{n+1/4} &= p_n + \omega \delta t F(q_n), \\
p_{n+2/4} &= e^{-\gamma \delta t/2}\left(p_{n+1/4} + \omega Z_n\right), \\
q_{n+1} &= q_n + \frac{1 - e^{-\gamma \delta t}}{\gamma e^{-\gamma \delta t/2}} M^{-1} p_{n+2/4}, \\
Z_{n+1} &= M^{1/2}\left(\alpha R_n + \hat{\alpha} R_{n+1}\right), \\
p_{n+3/4} &= e^{-\gamma \delta t/2} p_{n+2/4} + \hat{\omega} Z_{n+1}, \\
p_{n+1} &= p_{n+3/4} + \hat{\omega} F(q_{n+1}),
\end{aligned}$$

where

$$\omega = \frac{e^{-\gamma\delta t} + \gamma\delta t - 1}{\gamma\delta t\,(1 - e^{-\gamma\delta t})}, \qquad \hat{\omega} = 1 - \omega,$$
$$a = k_B T \left(2\omega^2\gamma\delta t + \omega - \hat{\omega}\right),$$
$$b = k_B T \left(2\omega\hat{\omega}\gamma\delta t + \hat{\omega} - \omega\right),$$
$$c = k_B T \left(2\hat{\omega}^2\gamma\delta t + \omega - \hat{\omega}\right),$$
$$\alpha = 2^{-1/2}\sqrt{c + a + \sqrt{(c+a)^2 - 4b^2}},$$
$$\hat{\alpha} = 2^{-1/2}\sqrt{c + a - \sqrt{(c+a)^2 - 4b^2}},$$
$$\alpha_0 = \sqrt{\alpha^2 - c}.$$

−− Brünger/Brooks/Karplus (BBK) −−

$$p_{n+1/2} = \left(1 - \frac{\gamma\delta t}{2}\right) p_n + \frac{\delta t}{2} F(q_n) + \frac{1}{2}\sqrt{2\gamma k_B T \delta t}\, M^{1/2} R_n,$$
$$q_{n+1} = q_n + \delta t M^{-1} p_{n+1/2},$$
$$p_{n+1} = \left(1 + \frac{\gamma\delta t}{2}\right)^{-1} \left(p_{n+1/2} + \frac{\delta t}{2} F(q_{n+1}) + \frac{1}{2}\sqrt{2\gamma k_B T \delta t}\, M^{1/2} R_{n+1}\right).$$