Lectures notes on Advanced Computational Statistical Physics

Lectures given by Ralf Everaers, ENS-Lyon

Contents

Introduction

In the 19th century, classical mechanics, rooted in Newton's laws, dominated physics. Pierre-Simon Laplace famously articulated the deterministic worldview: given the initial conditions of a system, its future could be perfectly predicted through precise mathematical equations. This perspective treated the universe like a clockwork machine, where every event followed from the initial state.

However, as the study of thermodynamics and many-particle systems advanced, the limits of this purely deterministic approach became clear. Statistical physics emerged to address these complexities, particularly through the work of James Clerk Maxwell and Ludwig Boltzmann. Their pioneering contributions, such as Maxwell's velocity distribution in gases and Boltzmann's statistical interpretation of entropy, introduced probabilistic methods to understand the behavior of large ensembles of particles.

In this new framework, the precise motion of individual particles became less important; instead, statistical averages and distributions described macroscopic properties like temperature and pressure. While Laplace envisioned a universe governed by strict determinism, statistical physics embraced the unpredictability inherent in large systems, marking a profound shift in understanding.

This shift continued to resonate into the 20th century, influencing the work of physicists like Philip W. Anderson. Anderson famously argued that "more is different," suggesting that the behavior of complex systems cannot be fully understood by analyzing individual components alone. This echoes the insights of 19th-century statistical physics, where collective behavior emerged from many interacting parts, challenging the reductionist views of classical mechanics.

In summary, while classical mechanics remained essential for describing deterministic systems, the development of statistical physics in the 19th century introduced a probabilistic approach that transformed our understanding of many-body systems and laid the groundwork for modern physics.

I. Computational Statistical Physics

Computational methods allow us to simulate these complex systems directly, providing detailed insights into their behavior. Using modern computing power, scientists can model the interactions of millions, or even billions, of particles, making it possible to observe emergent phenomena such as phase transitions, critical behavior, and chaotic dynamics. This computational approach helps us overcome the "many-body" problem, where the sheer number of interactions in a system defies exact solutions.

One reason computational statistical physics is so powerful is that it can handle systems that are analytically intractable. For instance, systems with strong correlations between particles or those far from equilibrium, which are difficult to study using traditional methods, can be simulated using Monte Carlo techniques, molecular dynamics, and other algorithms.

Additionally, computational approaches enable the study of phenomena at different scales—from the atomic scale, where quantum effects dominate, to macroscopic scales governed by classical statistical physics. This versatility allows for a deeper understanding of both microscopic mechanisms and their macroscopic consequences, bridging the gap between theoretical models and real-world systems.

In essence, computational statistical physics allows us to explore systems that are too complex for exact solutions, providing a practical and powerful way to study emergent behavior, phase transitions, and non-equilibrium systems. By leveraging the power of computers (which is increasing exponentially since the second half of the 19th century 1.1), it opens up new frontiers for understanding the vast complexity of the physical world, which neither classical mechanics nor early analytical statistical methods could fully address.
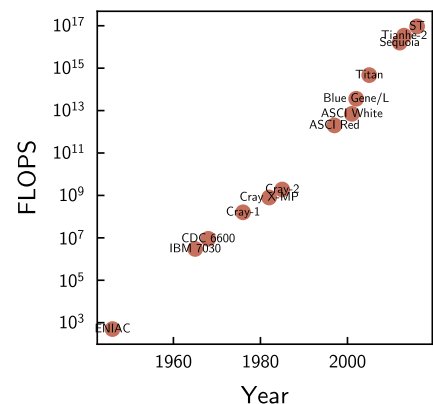

Figure 1.1 – Number of Floating points operations evolution for the biggest supercomputer

II. Invariant Measures and Ergodicity

In statistical physics, the concept of an invariant measure plays a crucial role in understanding the long-term behavior of dynamical systems, and the symetry of the system (Noether's theorem). One of the main goals of computational statistical physics is to find algorithms preserving the invariant measure of the systems we are studying. For example we will see that the Euler algorithm is not a good choice for simulating Hamiltonian systems, as it does not conserve the energy of the system, as opposed to the Verlet algorithm, which is based on clever considerations of the symplectic[1]. Ergodicity is another important concept in statistical physics, which ensures that the system explores the entire phase space in the long run. Which results in one of the most important hypothesis in computational statistical physics, the ergodic hypothesis, which states that the time average of a system is equal to its ensemble average.

Molecular Dynamics

I. Hamiltonian Dynamics

A. Hamiltonian Formalism

The Legendre transformation of the Lagrangian allows us to define the Hamiltonian of a system, which is a function of the generalized coordinates q_i and the generalized momenta p_i of the system (see CITE). The Hamiltonian is defined as :

$$H \overset{\text{def}}{=} \frac{\mathbf{p}^T \mathbf{M}^{-1} \mathbf{p}}{2} + U(q)$$

with $U(q)$ the potential energy of the system and $\mathbf{M}$ the mass matrix of the system. The equations of motions can then be written using the Euler-Lagrange equations :

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q} \end{cases}$$

This can be written in a more compact matrix form :

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = \mathbf{J} \nabla H(\mathbf{p}, \mathbf{q})$$

with

$$\mathbf{J} = \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix}$$

These type of non linear equations can only be ensured locally, but since the systems are Hamiltonian we can unearth the global existence and uniqueness of the solutions. Indeed we just need to show that using the energy constraints of the system the solutions are bounded for all time. For the impulsion this is trivial considering a potential global minimum :

$$\frac{\mathbf{p}^T \mathbf{M}^{-1} \mathbf{p}}{2} \leq E_0 - U_{min}$$

For the position we just need an assumption on the level sets of the potential energy.

$$\Sigma_\alpha = \mathbf{q} | U(\mathbf{q} = \alpha)$$

If these levels are bounded then the solutions are bounded for all time. This is not the case for all potentials, this is why for solving this kind of equations we will add sometimes confining potentials to the system.

B. Map Flow

Since the Hamiltonian equations are solvable, it seems natural to define a map flow $\mathcal{F}$ such that for an initial condition z_0 and a considered point z_t we have :

$$\mathbf{z}_t = \mathcal{F}_t(\mathbf{z}_0)$$

This flow map is obviously invertible and Hamiltonian conservative. The key of numerical integration is then to approximate the true flow map of the system by the numerical flow map $\mathcal{T}$ such that the physical properties of the system are conserved over time.

Considering the linear system for the differentiation :

$$\frac{d\mathbf{z}}{dt} = f(\mathbf{z}) = A\mathbf{z} \qquad (2.1)$$

One way to solve this equation is to use the matrix exponential :

$$\mathbf{z}_t = \exp(At)\mathbf{z}_0$$

Then it appears that the flow map is given by :

$$\mathcal{F}_t(\mathbf{z}) = \exp(tA)\mathbf{z} \qquad (2.2)$$

C. Symplectic Form

a) Volume Preservation

One of the most important properties of the flow map is the preservation of the volume in the phase space. Indeed, for a Hamiltonian system we have thanks to Liouville's theorem the volume of a given sets of solution governed by 2.1 is preserved over time if

$$\nabla \cdot f = 0 \,.$$

It is easy to show that for a Hamiltonian system :

$$\nabla \cdot f = \nabla \cdot \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} \nabla H = 0$$

Due to the $\mathcal{C}^2$ property of the Hamiltonian. The main consequence of this property is that the flow map is volume preserving (Liouville Theorem), which has to be an important concern while building an integration algorithm to approximate the flow map.
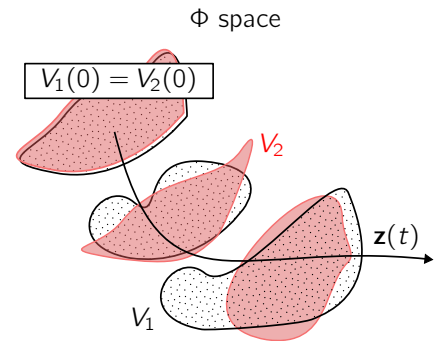
Φ space

$V_1(0) = V_2(0)$

V_2

$\mathbf{z}(t)$

V_1

Figure 2.1 – Volume conservation for the true Hamiltonian in black and the approximated Hamiltonian in red

b) Symplectic Property

The symplectic property of the flow map is another important characteristic of the Hamiltonian system. Indeed, the symplectic property of the flow map is defined as :

$$\mathcal{F}_t^T J \mathcal{F}_t = J \qquad (2.3)$$

With $J = \begin{pmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{pmatrix}$ the symplectic matrix. This property leads to the conservation of the volume of the phase space, indeed one can easily show using (2.3) that we have for the Jacobian associated with the map : $|\mathcal{F}'_t|^2 = 1$.

One interesting point of symplectic maps is that they form a group (they preserved the composition) which can be used to build other more complex symplectic maps. For example to increase the order of symplectic integrators $\mathcal{G}_h$ approximating the flow map $\mathcal{F}_h$, we can compose this latter with an other interesting symplectic integrator, this final integrator will then be symplectic and volume conservative. We can also split the integrator in two symplectic part if they are independant which is the case for Hamiltonian flow map in a pure potential $V(\mathbf{q})$. This is called the splitting method [1], and will be the object of the next discussion.

---

[1]Symplectic geometry has its origins in the Hamiltonian formulation of classical mechanics where the phase space of certain classical systems takes on the structure of a symplectic manifold

Andrea Combette

D. Error Analysis for Hamiltonian splitting

a) Lie Derivatives and Poisson Brackets

Lie Derivatives

In the case of a non-Linear Hamiltonian system, the flow map can be approximated by the Lie derivative of the Hamiltonian, we would like to find again the convenient results :

$$\mathcal{F}_t = \exp(tA)\mathbf{z}$$

Let's consider a functional Φ of the phase space, the Lie derivative of Φ is defined as :

$$\mathcal{L}_f \Phi = \nabla \Phi \cdot f$$

This is a generalization of the directional derivative to the phase space. Expanding the Φ map in a Taylor series we have :

$$\Phi(\mathbf{z}(t)) = \sum_i \frac{t^i}{i!} (\mathcal{L}_f^i \Phi)(\Phi(\mathbf{0}))$$
$$= \exp(t\mathcal{L}_f)(\Phi(\mathbf{0}))$$

Applying this along the trajectory gives us the following map :

$$\mathcal{F}_t(\mathbf{z}) = \exp(t\mathcal{L}_f)(\mathbf{z})$$

Poisson Brackets

A common notation introduced in Hamiltonian mechanics is the Poisson bracket, which is defined as :

$$\{g_1, g_2\} = \sum_{i=1}^N \left(\frac{\partial g_1}{\partial q_i} \frac{\partial g_2}{\partial p_i} - \frac{\partial g_2}{\partial q_i} \frac{\partial g_1}{\partial p_i} \right)$$
$$= \nabla g_1^T \mathbf{J} \nabla g_2$$

Considering a smoooth scalar value function F of the phase space, we can show that the Lie derivative of F is given by the Poisson bracket of F and the Hamiltonian :

$$\mathcal{L}_{\mathbf{J}\nabla H} F = \mathcal{L}_H F = \{F, H\} \qquad (2.4)$$

Poisson brackets can be related to the Lie derivating noticing that for every real valued function f:

$$[\mathcal{L}_{H_1}, \mathcal{L}_{H_2}] f = \mathcal{L}_{\{H_1, H_2\}} f$$

b) Error Analysis for non Commuting Hamiltonian

The main idea behinds this study is to consider integration method as a splitting of the Hamiltonian into several parts oftenly independant of one of the two system of coordinates $(\mathbf{q}, \mathbf{p})$, as we will see further. The poisson brackets are linear with respect to the Hamiltonian (2.4), which allows us to write the following considering $H = H_1 + H_2$,

$$\mathcal{L}_H = \mathcal{L}_{H_1} + \mathcal{L}_{H_2}$$

The flow map of the system is then defined as : $\mathcal{F}_t(\mathbf{z}) = e^{t()\mathcal{L}_{H_1} + \mathcal{L}_{H_2}}\mathbf{z}$

The splitting method has the following flow map :

$$\mathcal{G}_t h = e^{h\mathcal{L}_{H_1}} e^{h\mathcal{L}_{H_2}}$$

Which is not the same as the true flow map. To evaluate the error done considering the splitted Hamiltonian, we can use the Baker-Campbell-Hausdorff formula and the correspondance between the Lie derivative and the Poisson bracket unearthing :

$$e^{h\mathcal{L}_{H_1}} e^{h\mathcal{L}_{H_2}} = e^{h\mathcal{L}_{\tilde{H}_h}}$$

With $\tilde{H}_h$ the so called *shadow Hamiltonian* [1] :

$$\tilde{H}_h = H_1 + H_2 + \frac{h}{2}\{H_1, H_2\}$$
$$+ \frac{h^2}{12}(\{H_1\{H_1, H_2\}\} - \{H_2\{H_1, H_2\}\})$$
$$+ \dots$$
$$(2.5)$$

From this we can easily understand that as long as the two split Hamiltonian do not commute, we have at least a linear error (one example is the sympletic Euler Method). One way to do that, is to find a split with commutating Hamiltonian, or to split H into three Hamiltonian such that that the linear term vanishes when we develop the expansion (such as the Velvet Velovity & position method). The main idea behind this development was to show that the flow map $\mathcal{F}_t$ we approximate using the other $\mathcal{G}_t$ stands in the splitting method for an exact solution to an other (but similar) Hamiltonian system.

To put in a nutshell the splitting method allows us to build sympletic integrator using Hamiltonian transformations. Not all interesting integrators have sympletic structures, but they should all have volume conservation properties at least on long time scales.

II. Time Integration

Here we will restrict our study to simple linear system. To perform time integration, we discretize time into small intervals δt . At each time step $t_{n+1} = n\delta t$, we approximate the change in the system using matrix iteration.

$$\begin{pmatrix} x(n\delta t) \\ y(n\delta t) \end{pmatrix} = \mathcal{T}^n(\delta t) \begin{pmatrix} x(0) \\ y(0) \end{pmatrix}$$

The goal is then to find the correct matrix such that iterating over time will not change the invariance of the system. To fit the previous notations the flow map can be defined as :

$$\mathcal{F}_{\delta t}(\mathbf{z}) = T(\delta t)\mathbf{z}$$

with $\mathcal{F}_{\delta t}$ a volume conservatuve flow map.

A. Application to the Harmonic Oscillator

The Harmonic oscillator is a simple system with a known solution, despite its simplicity it unearthed so key concerns for time integration algorithms (stability of the integrator, pseudo-conservation of the energy, etc.).

$$\ddot{x} - \omega^2 x = 0$$

a) Exact Propagator

The exact time evolution of the system for a time translation of δt is given by the following propagator :

$$\begin{pmatrix} x(\delta t) \\ \dot{x}(\delta t) \end{pmatrix} = \mathcal{T}(\delta t) \begin{pmatrix} x(0) \\ \dot{x}(0) \end{pmatrix}$$

With

$$\mathcal{T}(\delta t) = \begin{pmatrix} \cos(wt) & \frac{1}{w}\sin(wt) \\ w\sin(wt) & \cos(wt) \end{pmatrix}$$

In the next analysis we will use a dimensionless time $\tau = \omega t$, a dimensionless position $\xi = \sqrt{\frac{k}{k_B T}}x$ and a dimensionless impulsion $\Pi = \sqrt{\frac{1}{mk_b T}}p$, which leads to the following time Propagator :

$$\mathcal{F}_{\delta \tau} = \mathcal{T}(\delta \tau) = \begin{pmatrix} \cos(\delta\tau) & \sin(\delta\tau) \\ -\sin(\delta\tau) & \cos(\delta\tau) \end{pmatrix}$$
$$= \exp(\mathbf{J}\delta\tau)$$

Here we related as did previously the flow map to the derivative operation for our system. For the Hamiltonian we recover :

$$H = \frac{\mathcal{H}}{k_B T} = \frac{1}{2}(\Pi^2 + \xi^2)$$

Andrea Combette

Then let's consider a microstate of the system $|\tau\rangle = \begin{pmatrix} \xi(\tau) \\ \Pi(\tau) \end{pmatrix}$ The time evolution of the system is given by the following equation :

$$|\tau + \delta\tau\rangle = \mathcal{T}(\delta\tau)\,|\tau\rangle$$

With the unitary propagation matrix

$$\mathcal{T}(\delta\tau) = \begin{pmatrix} \cos(\delta\tau) & \frac{1}{\omega}\sin(\delta\tau) \\ -\omega\sin(\delta\tau) & \cos(\delta\tau) \end{pmatrix}$$

Solving the characteristic equation $\|\mathcal{T} - \lambda I\| = 0$, we unearth the eigenvalues.

$$\lambda_\pm = \exp(\pm i\delta\tau),$$

Considering the two **orthogonal** eigenvectors $|\pm\rangle$, we easily show that :

$$|n\delta t\rangle = \mathcal{T}(\delta t)^n\,|0\rangle$$
$$= a_+\lambda_+^n\,|+\rangle + a_-\lambda_-^n\,|-\rangle$$

Considering the initial decomposition

$$|0\rangle = a_+\,|+\rangle + a_-\,|-\rangle$$

Then we can easily show that :

$$\frac{E}{k_b T} = \frac{1}{2}\langle n\delta\tau | n\delta\tau\rangle$$
$$= \frac{1}{2}\left(|a_+|^2 + |a_-|^2\right)\mathcal{T}(\delta t)^n$$
$$= \frac{1}{2}\langle 0|0\rangle$$

Which shows that the energy is conserved over time. We can also show an interesting result about the conserved quantities : Every conserved property of the system is proportional to the Hamiltonian.

$$\langle 0|A|0\rangle = \langle \tau|A|\tau\rangle$$
$$\Rightarrow A \equiv \mathcal{T}^\dagger A\mathcal{T}$$
$$\Rightarrow A \propto H$$

b) Euler Algorithm & Propagator

The Euler algorithm is the simplest algorithm to perform time integration, based on the first order approximation in Taylor expansion. The explicit non symplectic Euler algorithm is given by the following flow map :

$$\mathcal{G}_{\delta\tau}^{\text{euler}} = I - S\delta\tau$$

With S verifying

$$\frac{d\mathbf{z}}{d\tau} = S\mathbf{z}$$

One can remark that for volume conservation the map should verify the following property :

$$\det(I - S\delta\tau) = 0 \qquad (2.6)$$

Which is generally not the case in physical system were the eigen values of S are in many cases, pure imaginary. If we write the propagation matrix for the Euler algorithm, we obtain the following :

$$\mathcal{T}(\delta\tau) = \begin{pmatrix} 1 & \delta\tau \\ -\omega^2\delta\tau & 1 \end{pmatrix}$$

Which is not unitary, and not time reversible. We can show that the energy of the system is not conserved over time, through the same reasoning as before. Indeed, the two eigenvalues of the matrix are :

$$\lambda_\pm = 1 \pm i\delta\tau$$

Which leads to divergence of the energy with the following expression :

$$H = \frac{1}{2}\langle 0|0\rangle\left(1 + \delta\tau^2\right)^n$$
$$= \frac{1}{2}\langle 0|0\rangle\exp\left(n\ln\left(1 + \delta\tau^2\right)\right)$$

This divergence of the system can be seen in fig 2.3. This lack of stability is closely related to the relation (2.6), but stability analysis is a more complex subject, which will not be discussed here (see [1] for more information).

c) Verlet Algorithm & Propagator

The Verlet algorithm is a second-order algorithm, which is based on the symplectic structure of the phase space and directly derived from the splitting method. The point is to split the Hamiltonian into two parts, one depending only on the position and the other only on the impulsion.

$$H = H_q + H_p$$

With $H_q = U(q)$ and $H(p) = \frac{p^T M^{-1} p}{2}$. If we just do that it is obviously from the shadow Expansion of the Hamiltonian (2.5) an order one method (the Euler symplectic method). For the velvet algorithm we want that the order one term vanishes, which leads to the following splits:

$$H = H_1 + H_2 + H_3$$

With the following possibility :

$$\begin{cases} H_1 = H_3 = \frac{1}{4}p^T M^{-1}p & | & H_2 = U(q) \\ H_1 = H_3 = \frac{1}{2}U(q) & | & H_2 = \frac{1}{2}p^T M^{-1}p \end{cases}$$

One can show that the first order term vanishes in the shadow Hamiltonian, which defined a second order method. The first split is called the **position Verlet algorithm** , and the second the **velocity Verlet algorithm** . Let's focus on the velocity algorithm, this split is equivalent to the following canonical change of variable :

$$\hat{\mathbf{P}} = \mathbf{P} - \frac{h}{2}\nabla U(\mathbf{q}) \qquad (2.7)$$

$$\mathbf{Q} = \mathbf{q} + \frac{h}{2}\mathbf{p}^T M^{-1}\mathbf{p} \qquad (2.8)$$

$$\mathbf{P} = \hat{\mathbf{P}} - \frac{h}{2}\nabla U(\mathbf{Q}) \qquad (2.9)$$

This consist in a first drift $\frac{h}{2}\nabla U(\mathbf{q}$, following by a kick $\frac{h}{2}\mathbf{p}^T M^{-1}\mathbf{p}$ and a final kick. This respect the hamiltonian srtucture of the system and gives the following propagation matrix :

$$\mathcal{T}(\delta\tau) = \begin{pmatrix} 1 - \frac{\delta\tau^2}{2} & \delta\tau \\ -\delta\tau\left(1 - \frac{1}{4}\delta\tau^2\right) & 1 - \frac{\delta\tau^2}{2} \end{pmatrix}$$

Here we find an unitary matrix, which conserves a pseudo energy, but not the true Hamiltonian. This non conservation of the Hamiltonian is related to non orthogonality of eigenspaces, however we can still study the so called conserved pseudo energy and show that :

$$\langle 0|A|0\rangle = \langle \tau|A|\tau\rangle$$
$$\Rightarrow A \equiv \mathcal{T}^\dagger A\mathcal{T}$$
$$\Rightarrow A \propto \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{(1-\delta\tau^2/2)^2} \end{pmatrix} \xrightarrow{\delta\tau\to 0} \propto H$$

Hence, this time step dependent pseudo-energy converged to the wanted conserved value. This can be related to the discussion on the *shadow* Hamiltonian, were we saw that the Hamiltonian describing the system in these kind of integration was changing, due to commutator terms. Hence, the true energy is not conserved but the energy corresponding to the exact description of the integration system is definitely conserved.
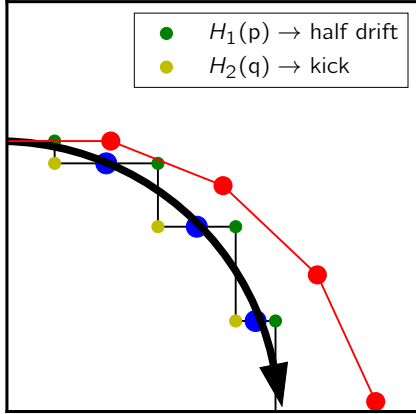
Andrea Combette

Figure 2.2 – Here we compared the construction of the flow map for the euler and the velvet algorithm. We can see that for great time steps the Euler algorithm diverges from the true flow map, while the velvet algorithm stays close to the true flow map.

This is quite relevant to notice that for periodic system the euler algorithm is definitely not stable due to the non conservation of the volume. We can also see that as a consequence of taking the tangent for each time steps of the true flow map and will leads to exponential divergence from the trajectory in the phase space. This is why the velvet algorithm is a better choice for time integration of Hamiltonian systems.
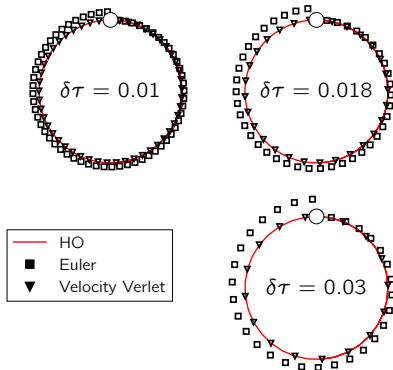

Figure 2.3 – To have a more general view of the efficiency of the integration scheme using a simple Harmonic oscillator we plotted the phase space trajectories for the Euler and the Velvet Algorithm, for two different timestep values

Constant Temperature Dynamics

In molecular dynamics simulations, maintaining a constant temperature is essential for studying temperature-dependent properties and sampling the right thermodynamic ensemble. Vari-

ous thermostats have been developed to control the temperature of the system, each with its own advantages and limitations. In this chapter, we will explore some common thermostats used in constant-temperature molecular dynamics simulations and discuss their properties.

I. Different Ensembles

In statistical physics, an ensemble is a large collection of hypothetical copies of a system, each representing a possible state that the system can be in. Ensembles are used to describe the macroscopic properties of systems based on the statistical behavior of their microscopic components. The concept of ensembles is fundamental in statistical mechanics because it allows the connection between microscopic interactions (the behavior of individual particles) and macroscopic observable (such as temperature, pressure, and magnetization).

A. Micro-canonical Ensemble

The micro-canonical set is the simplest ensemble in statistical mechanics, where the system is isolated and has a fixed energy. It is described using a constant number of particles N, volume V, and energy E. In an isolated system the fundamental postulates states that the system will explore all possible microstates with the same probability. Hence we can define the density of micro-states for the micro-canonical ensemble as :

$$\rho_{eq} = \frac{1}{\Omega}$$

Where Ω is the number of microstates of the system with energy E, some key properties can be derived from this definition, such as the entropy of perfect gas, thermodynamic equilibrium pressure and chemical potential properties, but we will assume that known for the reader. The Micro-canonical ensemble is the one used when one simulate a physical system due to conservation laws regarding the Energy, however it is much more convenient to work with constant Temperature, for this we can introduce the canonical ensemble.
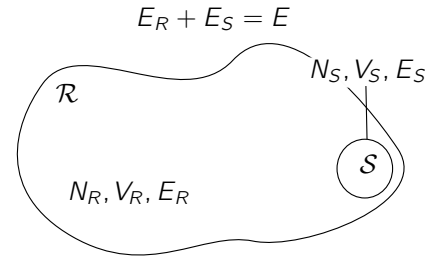
II. Canonical Ensemble

Figure 3.1 – Canonical ensemble schematic representation

In the Canonical ensemble we consider that our system is in contact with a biog reservoir, thus this two system composed a bigger isolated system. Then since the system $\mathcal{S} + \mathcal{R}$ is a microcanonical ensemble we get the following :

$$p_{\mathcal{S}} = \frac{\Omega_{\mathcal{R}}(E_{\mathcal{R}} = E - E_{\mathcal{S}})}{\Omega_{tot}}$$

Considering the reservoir at equilibrium we get :

$$\Omega_{\mathcal{R}}(E_{\mathcal{R}} = E - E_{\mathcal{S}}) = \exp\left[\frac{S_R}{k_B}(E - E_S)\right]$$

then if we develop the entropy around the mean energy of the system we get :

$$p_{\mathcal{S}} = \frac{1}{Z}\exp\left(-\frac{E_S}{k_B T_0}\right)$$

This stands for the Boltzmann-Gibbs distribution with Z defining the partition function of the canonical ensemble and verifying the following relation :

$$Z = \sum_s \exp(-E_s/k_B T_0)$$

.

A simple formulation of the partition function can be derived counting the degeneracy of the Energy level $\omega = e^{S(e)/k_B}$. This will lead to the following expression :

$$Z = \int \exp(-\beta(E - T_0 S(E)))dE \tag{3.1}$$

This motivates the definition of the free energy $F = E - TS$, and developing this latter around its minimum at the second order in the into grand gives the following :

$$Z = e^{-\beta(E^* - T_0 S(E^*))}\sqrt{\frac{2\pi k_B}{-\frac{\partial^2 S(E^*)}{\partial E^2}}} \tag{3.2}$$

Andrea Combette

This can be used to prove that the free energy is minimal at equilibrium in the thermodynamic limit. But here we will show that this imply a gaussian distribution of the energy. Indeed using the previous calculation is leads to :

$$p(E) \propto e^{\frac{(E-E^*)^2}{2k_B} \frac{\partial^2 S(E^*)}{\partial E^2}}$$

The study of the second moment then unearth :

$$\langle \Delta E^2 \rangle = -k_B \frac{\partial^2 S(E^*)}{\partial E^2} = k_B T_0^2 C_V \tag{3.3}$$

With $C_V = \frac{\partial E}{\partial T}_{(eq,N,V)}$ the heat capacity at fixed volume. This implied that the considering this latter at constant in a small range of temperature the energy distribution is a gaussian whose width is proportional to T^2.

III. Thermostats

This huge reservoir we used in the canonical ensemble is called a thermostat and is really convenient to dissipate external constraints apply on our system $\mathcal{S}$ while conserving a velocity in rough correspondence with temperature. This reservoir has to be able to exert forces and absorb energy from the system, this interaction can be defined in several ways from deterministic to stochastic method.

A. Stochastic thermostat Thermostat

a) Andersen Thermostat

The Andersen thermostat consists in refreshing the velocities of particles with a given probability at each timestep which $P = \nu \delta t$ with the collision rate ν. The refresh process is done by drawing a new velocity from the Boltzmann distribution at the given temperature, this can be done from scrambling or for every particles. This method is really simple and do not introduce concerns about ergodicity. However, despite the equation of motion being unchanged too high collision rate can dominate the underlying dynamics, slowing down the speed of exploring the phase space, and breaking the dynamics. Indeed it will leads to loss of memory and exponential decay of autocorrelation functions, while it is not the case in many MD systems.

In order to tackle this issue Andersen proposed $\nu \propto C_T \rho^{\frac{2}{3}} N^{-\frac{2}{3}}$ as a typical collision rate such that it should preserve the dynamics in most cases.

b) Langevin Thermostat

Changing the dynamics of the system

The Langevin thermostat is a stochastic thermostat that introduces a friction term and a random force to the modified equations of motion. This thermostat mimics the effect of solvent particles collisions on the particles of the system (like the **Anderson** thermostat). The random force is a gaussian force whose amplitude is chosen independently from the dynamics but only on the Temperature of the system and the friction coefficient, so the more we damp the system the more we have to excitate it, this ensure that even tough we increase the damping effect the exploration of the phase space will not be too slow despite we broke the dynamics (as for the Andersen thermostat), on the other hand if we decrease the damping the exploration of the canonical distribution of energy will not be ensured. The position Langevin equatiions in an harmonic potential in stationary state are given by :

$$\begin{cases} 0 = \xi v(t) - kx(t) + f(t) \\ \langle f(t)f(t') \rangle = A\delta(t-t') \\ \langle f(t) \rangle = 0 \end{cases}$$

With the characteristic time $\tau = \xi/k$ and the friction coefficient ξ. The solution can be written as following :

$$x(t) = x_0 e^{-t/\tau} + \tau^{-1} \int_0^t e^{-(t-t')/\tau} f(t')dt'$$

Amplitude of the noise

And one can easily assert what we previously saw regarding the independence of the random force with the dynamics.

$$\langle x(t)^2 \rangle = A\xi^{-2} \int_{-\infty}^t e^{-2(t-t')/\tau}$$

$$= A\xi^{-2} \frac{\tau}{2} = \frac{k_B T}{k}$$

We then got the following : $A = 2\xi k_B T$, this is a pretty interesting results since the noise amplitude is directly proportional to the friction constant (the more we damped the more we excite).

Mean Square Displacement

At very short time the t the mean square displacement (**msd**) $\langle \delta r^2 \rangle$ is only given by inertial effects (expending at the first order) and using the equi-partition theorem on the Kinetic energy,

$$\langle \delta r^2 \rangle = d \frac{k_B T}{m} t^2$$

For longer timescales we have to study the effect of the friction term :

$$\langle (x_t - x_0) \rangle^2 = 2 \left(\langle x^2 \rangle - \langle x_t x_0 \rangle \right),$$
$$= 2 \frac{k_B T}{\tau} \left(1 - e^{-t/\tau} \right)$$

this is done using the equi-partition theorem and :

$$\langle x_t x_0 \rangle = 2 \frac{k_B T}{\tau} e^{-t/\tau}$$

For time $t \ll \tau$ this leads to a diffusive process recovering the Enstein relation with the diffusion constant : $D = \frac{k_B T}{\xi}$,

$$\langle (x_t - x_0) \rangle^2 = 2Dt$$

The τ constant is most of the time several order of magnitude above the characteristic time of the dynamics [2]. So we will generally considerate only the inertial and the diffusive regime

$$D = \frac{1}{2} \lim_{t \to \infty} \frac{d\delta r^2}{dt} \tag{3.4}$$

$$D = \frac{\delta r^2}{2t} \tag{3.5}$$

One can note that the memory loss using the Langevin thermostat will be continuously impacted by the white noise, we used for sampling the canonical ensemble, leading to corrupted autocorrelation functions. To deal with that one can implement n memory term in the noise [1, 2], which has to be chosen carefully to not break the dynamics of the system. In addition this stochasticity break the continuous behavior of the system, and leads to discontinuity in average variables at least for andersen method.

B. Deterministic Thermostat

Studying the stochastic thermostats leads obviously to find and tweaks all the parameters such that we can preserve the physical behavior of the system, one can argue that we can use deterministic method to do the same sampling of the canonical ensemble, however these kind of methods generally lacks of ergodicity, and hardly reproduce the typical variation of the Kinetic Energy in the canonical ensemble [CITE].

a) Velocity Rescaling

The simplest way to obtain a constant temperature dynamics is to rescale the velocity by a factor :

$$v' = \sqrt{\frac{K}{\mathcal{K}}} v \tag{3.6}$$

With $K, \mathcal{K}$ respectively the total Kinetic energy and the desired Kinetic energy derived from the targeted temperature. Evans and Morriss (1983b) have shown that it leads to the following ensemble distribution $p(\mathbf{r}, \mathbf{q})$:

$$\delta(K(\mathbf{p}) - K_0)\delta(\mathbf{P} - \mathbf{P}_0)\exp\left(\frac{-V(\mathbf{r})}{k_B T}\right)$$
(3.7)

With $\mathbf{P}, K$ respectively the total momentum and kinetic energy. For the conformation probability distribution it is clearly the canonical one. For the momentum one, it does not fit the canonical distribution. Which can lead to smaller oscillation in kinetic energy than the ones in the canonical ensemble. One can propose to add a specific timescale to relax the normalization which defines the Berendsen thermostat, where the velocity is this time slowing rescaled to the desired value :

$$\mathbf{v}' = \mathbf{v}\sqrt{1 + \frac{\delta t}{\tau}\left(\frac{K}{\mathcal{K}} - 1\right)} \qquad (3.8)$$

With τ the relaxation time, this method is really simple and can be used for equilibration, however it is not time reversible, and one can also show [1] that it leads for vanishing τ to microcanonical ensemble and isokinetic for diverging τ. Rescaling the velocity seems to be a too strong constraint to sample correctly the canonical ensemble. On average, this will lead to the same results as canonical sampling since every ensemble is, in the thermodynamic limit, equivalents, however for fluctuations dependent observable this will exhibit wrong results. These concerns were of strong interest in the 80's, leading to several (partial) solution.

One way to solve this is to introduce a random velocity rescaling to mimic the fluctuations in the canonical ensemble, this can be done using the simple Velocity Rescaling (3.6) or the Berendsen rescaling (3.8). To do so we just use a random pick for the Kinetic Energy drawn from the canonical distribution.

$$v' = \sqrt{\frac{\bar{K}}{\mathcal{K}}}, \qquad \bar{K} \in$$

This is called **Canonical Velocity rescaling** (CVR), and as shown by (Bussi), under the assumption of ergodicity this sample correctly the canonical ensemble. One can argue that the dynamics can be altered, because of this stochasticity.

An other totally deterministic approach is the Nose-Hoover formulation of the Dynamics.

b) Nose-Hoover Thermostat

The main idea here is to extend the system using two a new coordinate s and a momentum p_s representing the reservoir, the energy is then allowed to flow between the system and the reservoir. This leads to the following Hamiltonian of the total system :

$$\mathcal{H} = H(\mathbf{p}, \mathbf{r}) + \frac{p_s^2}{2Q} + gk_B T$$

With Q the thermal inertia of the reservoir, g the number of degree of freedom of the system. The choices of these parameter must me carefully done in order to ensure close coupling between thermostating and dyncamics. The ensemble distribution is then :

$$\exp\left(\frac{-H(\mathbf{q},\mathbf{p})}{k_B T}\right)\exp\left(\frac{-p_s^2}{2Qk_B T}\right)$$

This beautifully stands for the canonical distribution in $\mathbf{q},\mathbf{p}$. However Nosehimself found that this approach was difficultly exploring the phase for very simple system such as harmonic oscillator. Indeed, Nothing to drive the drive the fluctuations of p_s [CITE] and fluctuations of the thermostating velocites are necessary to ensure the ergodicity of the system, since it will fill difficulty the phase space (coupling between the system and the reservoir). One possible solution is nose-hoover thermostatic chains, in order to drive the thermostat velocities with other reservoirs, this stands with the **CVR** as a perfect solution but Bussi argued we lost the beauty and the simplicity of the theory multiplying the thermostatic chain.
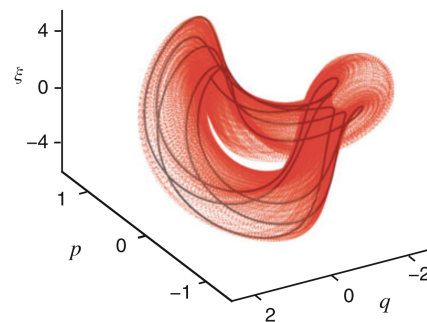

Figure 3.2 – length 2 Nose Hoover Chain applied on harmoic oscillator, trajectories are cought in an attractor, we do not recover the cannonical ensemble distribution (increase the number of chains [CITE], from [CITE]

C. Comparing Thermostats

Several studies can be done to compare the thermostat,the probability dis-

tribution function of the energy to see if it matches the canonical one, with high level of fluctuation. Regarding the dynamics we have to see if the velocity auto-correlation function is not dominated by thermostating effect, leading to corrupted transport coefficient calculation.

a) Energy Distribution

Studying the impact of thermostating on the energy distribution is a simple way to see if you correctly sample the wanted ensemble without breaking the ergodicity or the dynamics of the system.

Using (3.3) we unearthed that the typical fluctuations of the energy in the canonical ensemble should be quite important, proportional to T^2. To ckeck which thermostat was able to mimic these fluctuations, we simulated a simple **LJ**-fluids with 32 particles equilibrated by different thermostats. The first step of the study was to calculate C_V for this system, to do so we use the **CVR**, and showed that the heat capacity was remaining constant in the range of study.
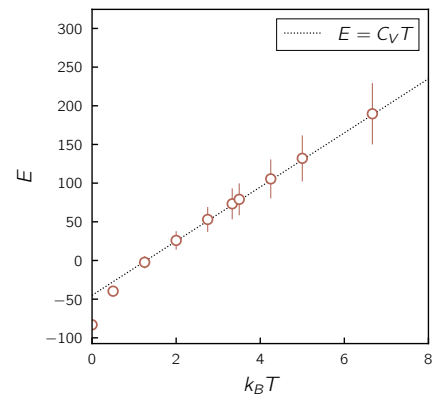

Figure 3.3 – Energy over Temperature, this plot unearths a value of Heat capacity at fixed volume : $C_V = 3.5e^1 \pm 1$

Knowing that we can estimate the canonical distribution of the energy at second order using our previous study (3.3).

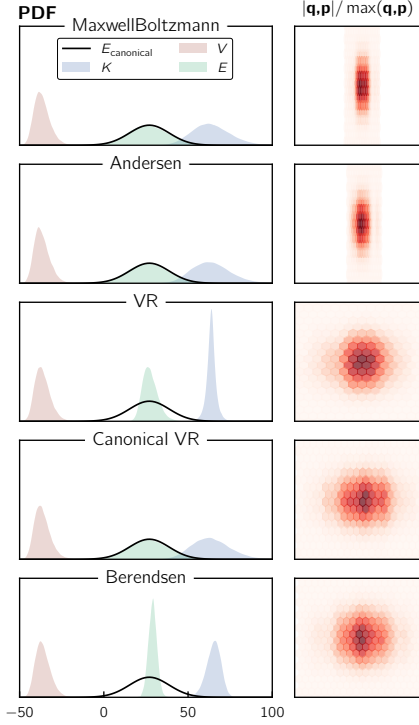$$p(E) \propto e^{\frac{(E-E^*)^2}{2k_B T_0^2 C_V}}$$

Andrea Combette

Figure 3.4 – On the left we plots PDFs of potential V, kinetic K and total energy E for several thermostats. On the right we plot the normalized phase space hexbin histogram (in the sens of explored domain size

As intended the **Maxwell Boltzmann, Andersen, CVR** sample correctly the canonical distribution of energy, which appears normal considering their stochastic behavior. For the **VR, Berendsen** thermostat their non-canonical properties are clearly demonstrated with very small fluctuations of the Kinetic energy. However, one can notice on the fig 3.4 that the phase space exploration is inefficient for the **Maxwell-Boltzmann, Andersen** thermostat, with a very restricted exploration of the conformation space, this is mainly due to the fact that we broke the system dynamic, thermostating the velocities too often for this kind of thermostat. This issue does not appeared in the three other thermostats, where the phase space exploration is efficient.

To put it in a nutshell thermostating is a subtle problem where we have to find the right equilibrium between sampling correctly the ensemble we want and the phase space in a reasonable time without breaking the dynamics, and ensuring ergodicity.

D. Calculating Transport Coefficients

To check if the thermostats preserve the physics of the system, we can study auto-correlations functions. Indeed referring to (Haile and Gupta, 1983), suitable thermostats should preserve the auto-correlations function of the velocity. It is simply a way of cheeking if the characteristic times of the system are unchanged. In addition to that a whole theory is dedicated to recover transport coefficient from auto-correlation functions (Green-Kubo relations). Here we will focus on the diffusion constant, which can be calculated from the velocity auto-correlation function (VACF).

a) Long time tails

In our simulated **LJ**-system the auto-correlation function of the system should decay slowly at large times, this is called long time tails phenomenon and is universal in this kind of fluid system. The first apparition of this result in the scientific litterature was back in the 70's witj the work of Alder and Wainwright simulating hard sphere.

Indeed they unearthed a long time tailed of the form $\sim t^{-\frac{d}{2}}$ contradicting the believed exponential decay of auto-correlation function (If we suppose that the dynamic follows the Langevin equation it is quite straightforward that it is the case). This dependancy was studied afterward by Pomeau and Résibois using long times hydrodynamics arguments.

The main idea is rely on the local equilibrium implied by the presence of multiple particle fig.3.5.b) these latter will interact between each other, leading to a constant velocity in a small volume Ω_t around the particle of interest, if you consider n particle interacting in this small volume then :

$$v(\tau) = v(0)/n\Omega_t$$

This volume will spread with time fig.3.5.c) dominated by hydrodynamic long time viscosity effect :

$$\Omega_t = (\nu t)^{\frac{d}{2}}$$

with d the dimension of the system. This leads to the following expression for the velocity auto-correlation function :

$$v(\tau) \sim v(0) \left(\frac{1}{\nu t}\right)^{\frac{d}{2}} \quad (3.9)$$
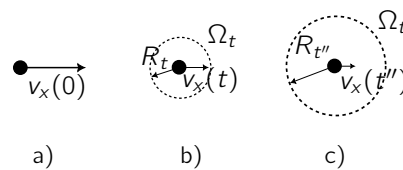

a)   b)   c)

Figure 3.5 – Schematics picturing the slow decay of auto-correlation functions, redraw from (Pomeau, Resibois [3])

This dissipative spreading of the small volume Ω_t can be pictured by the following reasoning : considering a particle interacting in a repulsive way with others particles, this interaction will create an eddy behind the particle that will ensure that the particle continue in its original direction during a sufficiently long time.

Memory loss

This slow memory loss is a key feature of the system, and can be used to check if the dynamics of the system is preserved by the thermostat. Indeed if the thermostat is too strong, the long time tail will be altered.
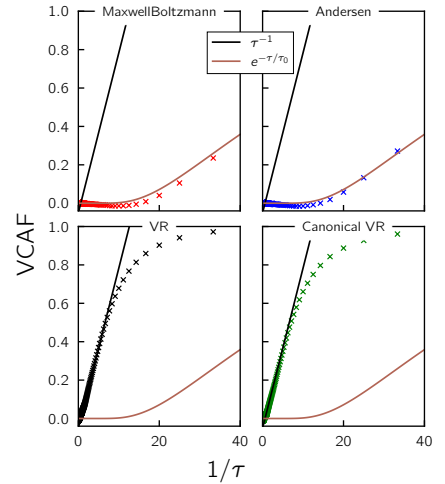

Figure 3.6 – Normalized Velocity auto-correlation function

Dependance of the density

To fall in this regime the system has to be dense enough as highlighted by [4]. Indeed the long time tail is a collective effect, and will be more pronounced in dense system, where the particle are more likely to interact with each other and the vortex formation may not be sustained. To ckeck this behavior we simulated the same 2d system. Since the regime we studies should be dominated by collisions, we must know the time between the collision, the Enskog kinetic theory predicts :

$$\tau_E = \frac{1}{4\rho\sigma^2 g(\sigma)} \sqrt{\frac{m}{\pi k_B T}} \quad (3.10)$$

With $g(\sigma)$ the radial distribution function at the contact distance σ, this will

Andrea Combette

be the characteristic length of the **LJ** potential in our case. For hard spheres it gives the following [CITE] :

$$g(\sigma) = \frac{1 - \pi\sigma^3\rho/12}{(1 - \pi\sigma^3\rho/6)^3}$$

We then decided to normalize the time by this time, and plot the VACF for different densities. Hence the density dependancy should not play a role anymore if it has only an effect on the collision time.
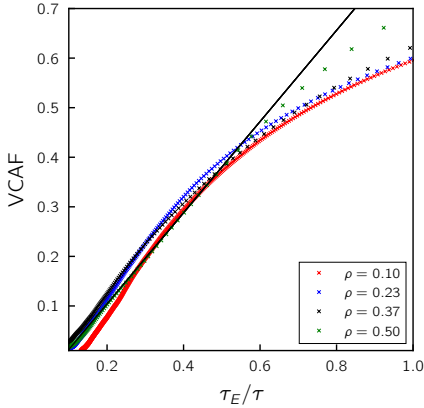

Figure 3.7 – Normalized Velocity auto-correlation function over the approximated frequency of collision.

The density dependancy is not visible anymore except for very low or high collision frequency, and the **VACF** arounf the order of magnitude of the collision time is as expected comparable to $\sim \tau^{-1}$

b) Calculation of Diffusion Constant

The calculation of the velocity auto-correlation function is of great interest since it can be used to calculate the diffusion constant of the system, indeed we have :

$$D = \lim_{t\to\infty} \frac{1}{2t}\langle \Delta r^2(t)\rangle$$
$$= \lim_{t\to\infty} \frac{1}{2t}\int_0^t \int_0^t dt'' dt' \langle v(t''-t')v(0)\rangle$$
$$= \lim_{t_1\to\infty} \frac{1}{2t}\int_0^t dt' \int_0^t dt'' \langle v(t'')v(0)\rangle$$
$$= \lim_{t\to\infty} \frac{1}{2t}\int_{-t}^t C(\tau)d\tau \int_{\max(-\tau,\tau)}^{\min(2t-\tau,2t+\tau)} \frac{dv}{2}$$
$$= \lim_{t\to\infty} \int_0^t C(\tau)(1-|\frac{\tau}{t}|)d\tau$$
$$= \int_0^\infty C(\tau)d\tau$$

We lead this integration explecitely since it is rarely done in the literature, we used the following change of variable :

$$\tau = t'' - t', \quad v = t'' + t'$$

The final expression is only valid when $C(\tau)$ decays faster than $\frac{1}{\tau}$ otherwise we got a logarithmic growth of the integral over time, this one of the numerically challenging issue when we focus on long time decay of the VACF, since we need to push the integration further, and run longer simulation. In 2d simulations we might even not be able to integrate because of the logarithmic growth of the integral. This way of calculation is commonly use but one should remember this integration subtility. With this latter we now have three way of calculating the Diffusion coefficient, one can ask which one is the most reliable depending on the situation. M.P.Allen and D.J.Tidesley already discussed about it [2], and drew the following conclusions. For exponential decay of the **VCAF**, it might be more accurate to compute gradient of the **msd** rather than dividing it by the time. So we will generally prefer (3.4) than (3.5). Regarding the calculation of the diffusion constant using the auto-correlation integration, we just need to integrate over the whole time to not miss long-time contributions.
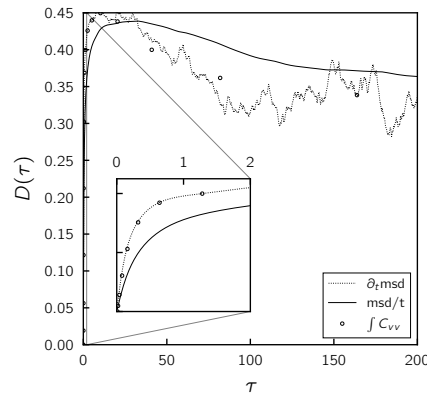

Figure 3.8 – Diffusion coefficient over time using the previous described methods

Metropolis Algorithm

I. Monte Carlo Integration

In Monte Carlo integration, we estimate an integral

$$I = \int_\Omega f(x)\,d\mu(x)$$

by averaging the function values at a set of randomly sampled points. If we sample N independent and identically distributed (i.i.d.) points $x_1, x_2, \ldots, x_N$ uniformly over the domain Ω, the Monte

Carlo estimate of the integral is given by:

$$I = V(\Omega)\langle f(x)\rangle_\Omega$$
$$\hat{I}_N = \frac{V(\Omega)}{N}\sum_{i=1}^N f(x_i).$$

The Monte Carlo estimate is unbiased, meaning that the expected value of the estimate is equal to the true integral. But one can remark that the convergence is quite slow, indeed for the uniform sampling we evoked before it leads to the following variance of the estimator :

$$\sigma^2(\hat{I}_N) = \frac{V(\Omega)^2}{N}\sigma^2(f(x)) \qquad (4.1)$$

the error can then be estimated bu the standard deviation σ which converges as $O(\frac{V}{\sqrt{N}})$. This is quite promising since the error convergence rate is independent of the dimensionality which beats the "*curse of dimensionality*". Indeed for a simple deterministic 2nd order Simson rule the convergence will in a space of dimension d : $O(N^{2/d})$, due to the number of evaluation we do. However one could not be satisfied by this poor convergence rate of $\frac{1}{2}$. Some solutions can be proposed as non-uniform sampling or important sampling.

II. Importance Sampling in Monte Carlo Simulations

Importance sampling is a technique used in Monte Carlo simulations to improve the efficiency of sampling by reducing the variance of the estimator. The goal is to focus computational effort on regions of the state space that contribute most significantly to the density of states. However this sampling cannot be done without care and should respect some properties of the density of states such as its constancy at equilibrium.

A. Concept of Detailed Balance and Equilibrium

In statistical mechanics, **detailed balance** is a condition that ensures a system will reach and maintain equilibrium. It requires that, for any two states ν and μ, the rate of transitions from ν to μ must equal the rate of transitions from μ to ν at equilibrium. Mathematically, this condition is written as:

$$P(\nu)W(\nu \to \mu) = P(\mu)W(\mu \to \nu) \qquad (4.2)$$

Andrea Combette

This ensures that at equilibrium the probability distribution remains unchanged over time. This detailed balance properties, appears simple but restrict set of possible scheme Monte Carlo simulation. Indeed if we consider the following scheme for a Physics system, where we choose to do a move in the phase space dependind on the implied energy variation ΔE :

$$\begin{cases} \text{If } \Delta E < 0, \text{ accept the move} \\ \text{If } \Delta E > 0, \text{ reject the move} \end{cases}$$

Since the scheme does not allow upward energy fluctuations (where $\Delta E > 0$), the condition:

$$P(\nu)W(\nu \to \mu) = P(\mu)W(\mu \to \nu)$$

is not satisfied for all transitions. Here we understand that thermal fluctuations are crucial for the system to explore the entire configuration space properly.

B. Metropolis Algorithm

To ensure that the Monte Carlo algorithm preserves detailed balance, transitions that increase energy must be allowed with appropriate probabilities most of the case it is given by the canonical ratio of the two state probabilities p_μ, p_ν). This is typically done using the **Metropolis algorithm**. Indeed, The Metropolis criterion ensures that the detailed balance condition is satisfied because:

$$\frac{W(\nu \to \mu)}{W(\mu \to \nu)} = \frac{P(\mu)}{P(\nu)} = \exp\left(-\frac{\Delta E}{k_B T}\right),$$

where ΔE is the energy difference between states ν and μ. This allows the system to transition both to higher and lower energy states, maintaining the equilibrium distribution.

III. Ising Model

The Ising model is a fundamental model in statistical physics used to study phase transitions and critical phenomena.

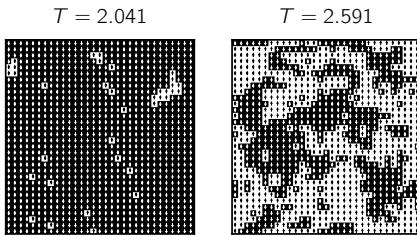| $T = 2.041$	$T = 2.591$

Figure 4.1 – Two configurations of the same random walker (respectively for $T < T_C$ and $T > T_C$, with $T_C \approx 2.269$ using a $L = 32$ size periodic ising model

It consists of discrete spins arranged on a lattice, where each spin can point up or down and interacts only with its neighbors. Ising model is a really simple model that unearth some interesting physical behavior while studying phase transitions.

A. Theoritical Background

The constant coupling ising model is defined by its Hamiltonian:

$$\mathcal{H} = -J\sum_{i,j} s_i s_j - B\sum_i s_i$$

Where s_i are the spins laying on a d dimensional grid, J is the coupling constant, and B is the external magnetic field. This model has been solve exactly for dimension 1 and only without external magnetic field in dimension 2. Some simplifications can be made considering **Mean Field Approximation**, but these approximation leads to misleading behavior near the phase transitions with wrong critical exponent. Here we can define two important quantities, the spontaneous magnetization and the coupling term M, Q :

$$M = \sum_i s_i$$
$$Q = \sum_{i,j} s_i s_j$$

Generally we study the reduced quantity m, q with respectively : $m = \frac{M}{N}$ and $q = \frac{Q}{N}$, we will also define the susceptibility $\chi = \frac{\partial m}{\partial B}$.

Phase Transitions

For the next discussion we will focus on thermodynamic limit, the system at stake will then be infinite. A phase transition occurs when a given quantity called the order parameter change from a zero value to a non zero value (for ferromagnetic system the order parameter is the spontaneous magnetization). Depending on the behavior of the system at the critical point one can define several type of phase transition.

For first order transition first derivatives of the free energy are discontinuous at the transition temperature. For the second order first derivatives are continuous. It will lead to metastable states or hysteresis phenomena in the first case and critical phenomena in the second case, with divergence of correlation length and time. In our case a 2nd order transition can be observed while sweeping the temperatures, and 1st order while seeping the external magnetic field for fixed temperature

Applying these definitions to our Ising model, we can unearth this two order of transition.

Magnetization discontinuity for first order transition

For sufficiently low temperature, bellow the critical temperatures. One can find a first order transition for $B = 0$ while sweeping the external magnetic field B. The system will undergo an abrupt change from a positive or negative magnetization depending on the starting state, for example if we start from a positive magnetization state we will go from $M_{sp}(B \to 0^+)$ to $-M_{sp}(B \to 0^-)$. We can illustrate that using the simple mean field approximation , we recall that this leads to the following expression :

$$m = \tanh\left(\frac{4Jm + B}{k_B T}\right) \quad (4.3)$$
$$\frac{k_B T x - B}{4J} = \tanh(x)$$

With $x = \frac{4Jm + B}{k_B T}$. This implicit equation can be solved graphically, for T bellow critical temperature ($T_c = \frac{4J}{k_B}$, which is not the write value for the 2d ising model). For now we will limit ourselves to the following two case ($m < 0$, $B < 0$) and ($m > 0$, $B > 0$).
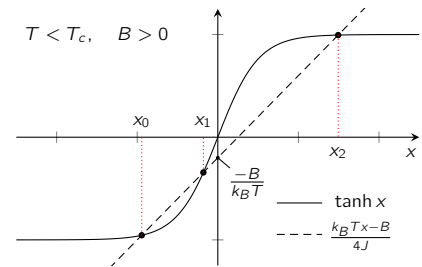

Figure 4.2 – Graphical resolution of the mean field equation (4.3) for $T < T_c$ and $B > 0$

If we limit to the case ($m > 0$, $B > 0$), only the x_2 solution is feasible since we got the following requirements $x > 0$, one can also suggest graphically that the x_2 value has an infimum independent from the magnetization and only dependent of the temperature since it's a solution of an implicit equation independent of m $x^+ > 0$. This will lead in the limit of vanishing magnetic field starting from positive magnetization

$$m^+ = x^+ k_B T$$

Andrea Combette

From the same reasoning we got similarly

$$m^- = -x^+ k_B T$$

With this simple modeling we explain how the ising model can exhibit discontinuities in the order parameter and hence a first order transition, another cleaner proof can be done using fluctuation dissipation theorem [5], exhibiting divergent behavior of the susceptibility at the transition point.

$$\chi = \frac{N}{k_B T}(\langle m^2 \rangle - \langle m \rangle^2) \propto N \quad (4.4)$$

Despite its usefulness this model present some limitations for this order of transiting, the mean field approximation consists in considering small fluctuations at the transition, this locality leads to meta stable states, that cannot exist in the 2d infinite ising model, since it is always at equilibrium. This metastables states arised from the x_0, x_1 solution and scales with the Coupling Constant of the system (see fig.4.3).
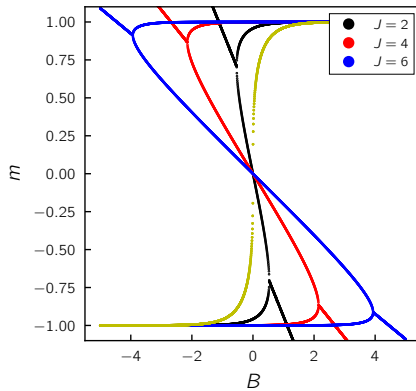
Figure 4.3 – Magnetization Isotherm for several coupling constants, solved using approximation of the implicit equation

Here we see the characteristic hysteresis behavior of the first first order transition for the mean field aplsing model, and results from the locality of the approximation, highly correlated structure are then more resistant to change of the external magnetic field.

Critical behavior for Second Order Transition

For second order transition, one can show that thermodynamic properties scales as power low of the reduced distance to the critical temperature $\epsilon = |1 - \frac{T}{T_c}|$. In the Ising model, Onsager unearthed that a second order transition is obtained for $B = 0$, $\epsilon \to 0$. In this case

we can express with this kind of power laws the order parameter m, the heat capacity C_V and the magnetic susceptibility χ, which respectively characterized how the system react to a change of Temperature and External Magnetic Field. The exponents of theses laws are called critical exponents and are usually constants for $T > T_C$ and $T < T_C$ (fore the amplitude of the power law this is not the case [6]). for the magnetization it is more complicated in the ising case, since it should vanish infinitely fast at 0, when $T > T_c$. For example in the Ising model we got the following scaling laws when $\epsilon \to 0$ [7].

$$m \sim \epsilon^\beta \quad , \quad \beta = \frac{1}{8}$$
$$C_V \sim \epsilon^{-\alpha} \quad , \quad \alpha = 0$$
$$\xi \sim \epsilon^{-\nu} \quad , \quad \nu = 1$$
$$\chi \sim \epsilon^{-\gamma} \quad , \quad \gamma = \frac{7}{4}$$

With m the magnetization, C_V the heat capacity, ξ the correlation length and χ the magnetic susceptibility. These critical exponents are universal, meaning that they are independent of the microscopic details of the system.

a) Finite Size Effect

Since we cannot simulate an infinite system, we have to restrict to finite system where the thermodynamic limit is not ensured anymore, for this study we will not consider boundary effect that can play an important role at the critical point [6].

Finiteness of a system lead to important change in the critical behavior, indeed the thermodynamic properties are smoothed and do not diverge near the transition point [6], and the order parameter change continuously during the transition, this might lead to difficulty distinguishing 1st and 2nd order transition. For second order transition in finite system the correlation length cannot diverge (as in first order transition) because it is obviously limited by the system size L.

For first order transition the discontinuity of the magnetization is smeared out by the finite susceptibility, leading to a continuous variation of the magnetization while sweeping B contradicting our previous study.

First Order Transition in Finite Size System

Here we will consider the most interesting case for finite size i.e $T < T_c$. In

these conditions the system will undergo a smooth change of magnetization, and one could unearth some metastable state due to the locality implied by the finite size. Indeed even if we let the system equilibrate it will be trapped in these local minima of free energy due to predominance of surface effect. In addition to that the susceptibility is no more spiking for equilibrium state (as we saw previously) [5]. Indeed due, to finitness the susceptibility is bounded and evolves proportionally to the system size (eq.4.4), explaining the smooth change of the magnetization (see fig.4.4).
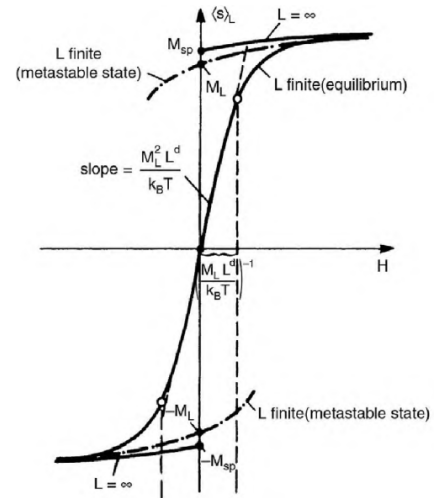

Figure 4.4 – Variation of the magnetization in a finite Ising model, M_{sp} denotes the magnetization at for the infinite phase transition, M_L for the finite one, from (Landau, David P. and Binder, Kurt, 2014) [6]

.

The system will then jumped back and forth between the two minimum of the free energy, leading to a unevenly weighted bimodal distribution of the magnetization (see fig.4.5). The existence of metastable states results in hysteresis loop depending on the starting configuration. Hence we can see that determining the critical point can be really hard knowing this hysteresis phenomenons, since it is constantly shifted by the existence of metastable states whose weight is dependent on the system size and the simulations parameters.

This bahavior had been studied in the fig.4.5. Where the characteristic bimodal distribution appeared, the width size dependancy shows correct agreement with [6], with narrower distribution for higher L. The size dependant slope is also clearly visible, denoting finite susceptibility at the transition point $\propto N$.

Andrea Combette

In our study the hysteresis behavior has only been found for $T < T_C$ as opposed to [8] where they used sinusoidal driving. The number of sweeps we used for Monte carlo simulations does not seem to play a great role contradicting at first sight the results of [8], but here they used logarihmic valued of the sweeping parameter, this can lead to very fast exploration of the phase space but also to issues in the equilibration of the walkers. The simulation temperature has interesting impacts on the hysteresis behavior. Indeed, increasing the temperatures seems to decrease the hysteresis effect, while reducing the population in the metastable states.
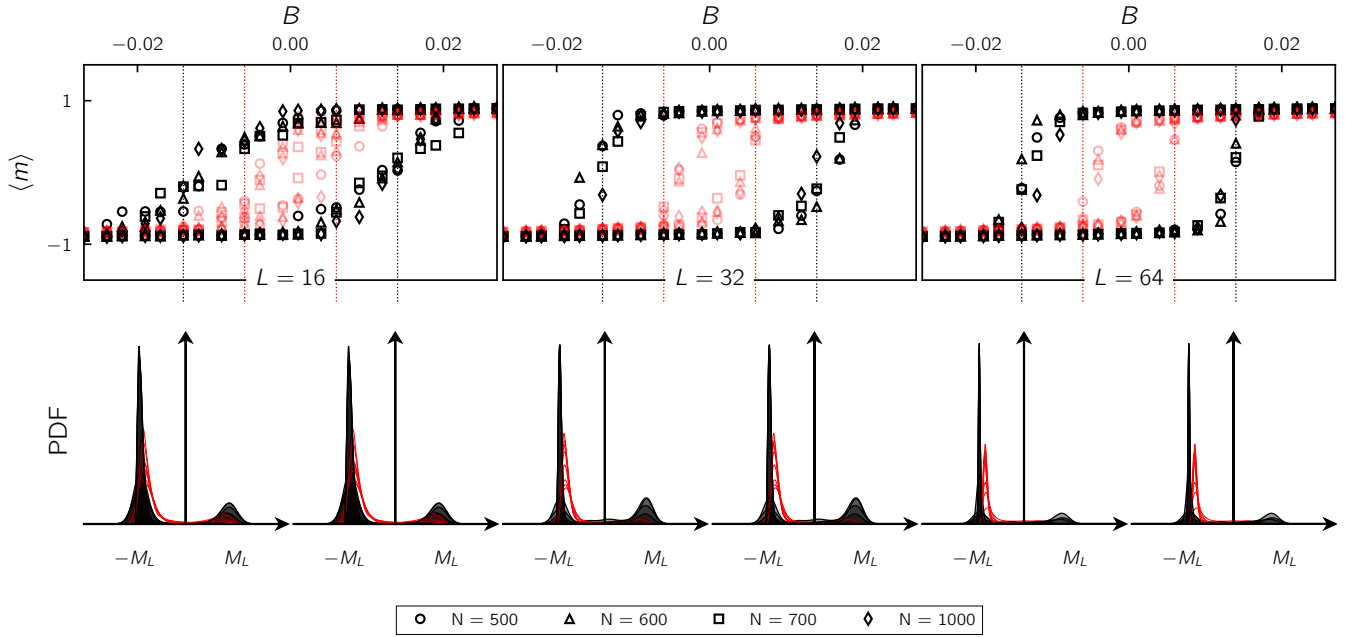

Figure 4.5 – On the top row, we plot the magnetization evolution while seeping the magnetic field B for different system size. The bottom row shows the corresponding histogram of the magnetization for $T = 2$ and several Monte Carlo sweeps (N). Red transparent dots stands for simulations done near the critical point $T = 2.269$, and black solid dot correspond to $T = 2 < T_C$. To get this clear metastability we started with negative B for all-down-state, and with positive B for all-up-state, then we increased and decreased respectively the magnetic field until we reach the opposite conformation. The sweeping frequency study has been motivated by [8] where it is shown that the magnetic sweeeping rate has a strong impact on the hysteresis below and above T_C

In first order transition oner can show [6] that the size of the system only play a role in the width and in the ampllitude of the distriubution with the same contribution a L^d factor. One can then understand that smart rescaling can leads to correct infinte behavior of the simulation.

In 2nd order phase transition the thermodynamics are governed by critical power laws, so we expect this finite scaling to change.

Second Order Transition in Finite Size System

It might seem more difficult to extract inifinte behavior for this kind of power laws. However one is able quite straightforwardly to extract infinite behavior from finite size simulations considering finite size scaling of the free energy [6] (this express the dependancy of the free energy over the length of the system in a simple way). Differentiation of this latter yields the following scaling laws for the order parameter and the susceptibility.

$$m_L = L^{-\beta/\nu} \mathcal{M}_0(\epsilon L^{\frac{1}{\nu}}) \tag{4.5}$$

$$\chi_L = L^{\gamma/\nu} \chi_0(\epsilon L^{\frac{1}{\nu}}) \tag{4.6}$$

Where $\mathcal{M}_0$, and χ_0 are scaling functions of the infinite system. Then we can easily propose to rescale the finite thermodynamics properties by the corresponding length dependencies, and to use the rescaled relative distance to the critical temperature $\epsilon L^{\frac{1}{\nu}}$. For example here we studied the rescaled susceptibility and magnetization for different system size (see fig.4.6).
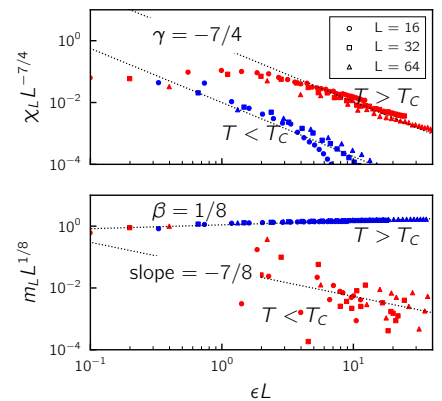

Figure 4.6 – Scaling laws for respectively the susceptibility and the magnetization for different system size. The simulations have been done using 10 random walkers with 500 checkboards sweep.

Andrea Combette

First the curves collapse on a single one, proving the validity of these scaling laws over the finiteness of the system. In addition to that we recover quite precisely the infinite scaling laws for the susceptibility and the magnetization. With the same slope for heating or cooling down the system for χ. For the magnetization it seems to exhibit a $\frac{7}{8}$ critical exponents already discussed in [6].

b) critical slowing down

As we saw previously the correlation length is diverging when we reach T_C, long wavelength phenomena are then dominant. In this situation one can understand that local spin flip will not lead to correct sampling of the phase space since it does not allow to perturb the highly correlated system. For a constant observable A, one can show (Hohenberg and Halperin, 1977) that the spatial components $\tilde{A}(q)$ of the Fourier transform $\tilde{A}$ follows the following relation :

$$\tau_{AA} = (D_{AA}q^2)^{-1}$$

With D_{AA} a transport coefficient. For example in spin exchange trial Move (Kawasaki), the concentration for each spin values is maintained constant. Hence for due to divergence of the correlation length the time to decorrelate the system will diverge as well, leading to a critical slowing down. To tackle this issues on can propose other trial move than just a local flip of one spin. The idea is then to flip larger domains of the system while preserving the flux of random walker (4.2). This can be done using percolation method (Wolff cluster algorithm) to determine the spins to flip in a single move, in this case we can choose the percolation rate in order to always flip the selected cluster while preserving the flux, which leads to extremely fast exploration of the phase space. Checkboards methods can also be used to apply basic flip trial on all non neighboring spins without breaking the simulations.

IV. Thermodynamic Integration

Thermodynamic integration is a method used to compute the free energy. The major issue is the impossibility of sampling all the phase space in a simulations. Indeed the free energy is derived using the partition function of the ensemble

$$F = -k_B T \log Z$$

The partition function is a sum over all the possible states of the system, and is then impossible to compute exactly.

However, the exact value of the free energy is not necessary, since we are often interested in the relative free energy between two thermodynamic states. We can then use several method to directly or iteratively calculate this differences.

A. Kirkwood Coupling Parameter

The Kirkwood coupling parameter method provides a simple way to connect two states using a linear interpolation parameter λ between their Hamiltonians:

$$H(\lambda) = (1 - \lambda)H_0 + \lambda H_1,$$

where H_0 and H_1 are the Hamiltonian of the initial and final states. The derivative of $H(\lambda)$ with respect to λ is given by:

$$\frac{\partial H(\lambda)}{\partial \lambda} = H_1 - H_0.$$

The Free energy has interesting properties related to the derivative of the Hamiltonian, indeed supposing the Hamiltonian can be written as a function of the coupling parameter λ, we can express

$$\frac{dF}{d\lambda} = \langle \frac{dH(\lambda)}{d\lambda} \rangle_\lambda$$

Thus, the free energy difference between the two states $|1\rangle$, $|2\rangle$ is given by:

$$\Delta F = \int_0^1 \langle H_1 - H_0 \rangle_\lambda \, d\lambda.$$

a) Mechanical and Thermodynamic Observable

One can notice that we link the derivative of a global thermodynamic observable depending over the whole phase space to the derivative of a measurable mechanical observable H defined for each specific state. This is the general the idea of the thermodynamic integration, linking the global thermodynamic properties to the local mechanical properties.

b) Ising Model case study

In the Ising Model defined previsouly one can easily verify that :

$$\begin{cases} \frac{\partial F}{\partial B} = \langle M \rangle \\ \frac{\partial F}{\partial k} = -\langle Q \rangle \end{cases}$$

Hence at first order for a small change of the magnetic field B and the coupling constant k we can easily compute the free energy difference between two states (h_i, k_i), (h_j, k_j) by expending F around these two points.

$$\Delta F = \frac{h_j - h_i}{2}(\langle M \rangle_i + \langle M \rangle_j)$$
$$+ \frac{k_j - k_i}{2}(\langle Q \rangle_i + \langle Q \rangle_j)$$

This concludes the basic Kirkwood scheme for our model, in a practical way it is useful to use neighboring states for this type of integration, so the expansion will not be too coarse. Hence, for computing the free energy difference between two far states one have to find a path of neighboring states between this two states (this is called **Thermodynamic integration**. It will lead to the following formulation of the free energy difference.

$$\Delta F = \sum_{i=1}^N \frac{h_{i+1} - h_i}{2}(\langle M \rangle_i + \langle M \rangle_{i+1})$$
$$+ \frac{k_{i+1} - k_i}{2}(\langle Q \rangle_i + \langle Q \rangle_{i+1}) \tag{4.7}$$

B. Exponential Averaging

Exponential averaging, or the Bennett acceptance ratio method, is an alternative approach that improves the efficiency of free energy calculations. It uses an exponential weighting of energy differences to minimize the impact of poorly sampled regions. From dividing the two associated partition functions one can show that, the free energy difference can be calculated as:

$$\Delta F = \frac{1}{\beta} \log \langle \exp\left(\beta(H_1 - H_0)\right)\rangle_1,$$

Hence using thermodynamic integration we can easily compute the free energy for a given path.

$$\Delta F = \sum_{i=1}^N \log \langle \exp\left((H_i - H_{i-1})\right)\rangle_i \tag{4.8}$$

This method is more genrally more precise than the Kirkwood coupling parameter since it does not use the 1st order approximation of the free energy.

V. Multi Histogram Analysis

The goal here is to compute the density of states $\Omega(M, Q)$ for the 2d ising model, in order to do so we will use the multi histogram analysis. This method is based on the following idea, we will divide the phase space (M, Q) in N^2 bins, and compute the probability of finding the system in each of these bins. The density of states for a given group of walker ν is then given by the following relation in the canonical ensemble :

$$\Omega_\nu(M, Q) = P_\nu(M, Q) e^{(\beta(E_\nu(M,Q) - F_\nu))} \tag{4.9}$$

Where $P_\nu(M, Q) = \langle \frac{N_\nu(M,Q)}{N_\nu} \rangle$ is the probability of finding the system in the

Andrea Combette

bin (M, Q) for the group of walker ν. From this we can derive all the thermodynamic quantity of the system for one group of walker ν, i.e one pair of value (h, k). One can already tell that sampling the entire phase space (M, Q) using only one pair of value h, k is unfeasiblem hence we need to find a way combining several group of walker, and weight their contribution to the final estimation Ω_{est} of the densoty of states.

This is generally done using the **WHAM**, weighted histogram analysis : Where the estimate is computed using the following relation :

$$\Omega_{est}(M, Q) = \frac{\sum_\nu N_\nu \Omega_\nu(M, Q)}{\sum_\nu N_\nu} \quad (4.10)$$
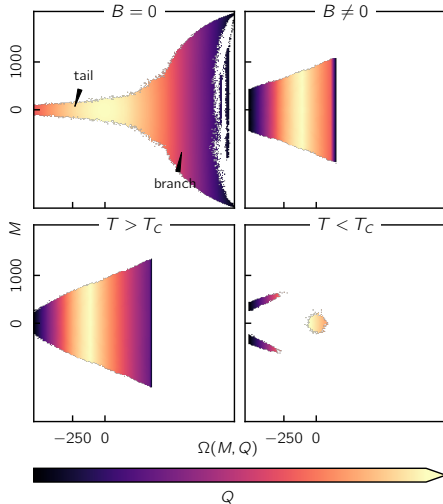

Figure 4.7 – Here we computed the density of state estimation using **WHAM** for 4 different typical cases on a $L = 32$ size Ising model. To obtain this estimation we compiled all the datasets used for our previous studies including several starting configuration (`all-spin-up`, `all-spin-down`, `random`), different number of random walkers, and different random walk length. The Intermediate computation of the free energy used Thermodynamic Integration and Exponential averaging over all the data collected

However the exponential term in $\Omega_\nu(M, Q)$ can be a source of numerical instability, hence it is often replaced by the following relation :

$$\Omega_{est}(M, Q) \approx \frac{\sum_\nu N_\nu(M, Q)}{\sum_\nu N_\nu e^{\beta(F_\nu - E_\nu(M,Q))}}$$

In the figure 4.7 we can see the estimation of the density of states for the 2d ising model, it appears that the density of states for the zeros external field is close from the litterature one. Intersting things can be notice here, in the the case of $T < T_C$ the two branch of the magnetization valid in low magnetic external field vanish, this is probably due to the anti-ferromagnetic value of the coupling constant, decreasing other weights. One can verify that we recover the two branches (metastable state) when we compute the density only with positive value of the coupling constant.

For $T > T_C$ one can see that external magnetic field will broaden the distribution in the "tail" of the density of states,which seems totally correct since magnetized state are more probable in non vanishing magnetic field, here we also lose the branches of the magnetization which is physically correct (we can verify that removing the anti-ferrmoagnetic contribution).

This conlude our study on the 2d ising model and more globally on this introduction to computationnal statistical physics. Some concerns should be raise about free energy calculation since we absolutely need to do not have large unexplored domain of the phase space, to compute it. In order to do so one can design clever method to avoid being trapped in some region of the phase space wuile simulating (paralell tempering, cluster algorithm, multi-canonical sampling . . .), but we will not discuss this here, and let the reader refer to the litterature for more information.

Bibliography

1. Leimkuhler, B. & Matthews, C. *Molecular Dynamics: With Deterministic and Stochastic Numerical Methods* English. ISBN: 978-3-319-16374-1 (Springer, May 2015).

2. Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids* ISBN: 9780198803195. https://doi.org/10.1093/oso/9780198803195.001.0001 (Oxford University Press, June 2017).

3. Pomeau, Y. & Résibois, P. Time dependent correlation functions and mode-mode coupling theories. *Physics Reports* **19,** 63–139. ISSN: 0370-1573. https://www.sciencedirect.com/science/article/pii/0370157375900198 (1975).

4. Ghosh, K. & Krishnamurthy, C. V. Molecular dynamics of partially confined Lennard-Jones gases: Velocity autocorrelation function, mean squared displacement, and collective excitations. *Physical Review E* **98.** ISSN: 2470-0053. http://dx.doi.org/10.1103/PhysRevE.98.052115 (Nov. 2018).

5. Cai, W. *ME346A Introduction to Statistical Mechanics* 2011.

6. Landau, D. P. & Binder, K. *A Guide to Monte Carlo Simulations in Statistical Physics* 4th ed. (Cambridge University Press, 2014).

7. Stanley, H. Scaling, Universality and Renormalization: Three Pillars of Modern Critical Phenomena. *Rev. Mod. Phys.* **71,** S358 (Mar. 1999).

8. Fan, Z., Jinxiu, Z. & Xiao, L. Scaling of hysteresis in the Ising model and cell-dynamical systems in a linearly varying external field. *Phys. Rev. E* **52,** 1399–1402. https://link.aps.org/doi/10.1103/PhysRevE.52.1399 (2 Aug. 1995).

9. Lydéric Bocquet, M. B. *Lecture notes on Statistical Physics* 2023.

Andrea Combette