UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics

Final Dissertation

Coarse-graining and entropy production out of equilibrium

Thesis supervisor

Prof. Attilio Stella

Thesis co-supervisor

Dr. Gianluca Teza

Candidate

Riccardo Sanson

Academic Year 2018/2019

Abstract

The problem of establishing how different coarse-graining procedures affect the law of generalized detailed balance and the rate of entropy production is addressed in the context of different models describing systems in stationary states out of equilibrium. The most microscopic descriptions involve Langevin dynamics, while the coarse-grained counterparts amount to Markov jump processes. The study has also the aim to validate and test procedures recently adopted for identifying non-equilibrium in active soft matter systems. The work will involve both analytical and numerical work.

Contents

Contents

Introduction

Let's suppose we are interested in studying the motion of a microscopic particle immersed in a solvent. The most straightforward approach would be to solve the equations of motion for every molecule in the system. This is a sure way to achieve a correct dynamics but, apart from the difficulty of finding a Hamiltonian of the system, it's too computationally expensive, even for modern computers. The alternative is to take a statistical approach, that is to describe the dynamics of the molecules of the solvent using only a few parameters: we then recover the well-known Browian motion, which is expressed matematically by a stochastic differential equation called *Langevin equation*.

Basically what we have done in this example was to reduce the degrees of freedom of the system to facilitate its study. This procedure is a case of *coarse graining*, which stands at the foundations of *Statistical Mechanics*. A coarse-grained description of a system is indeed a premise of every theoretical model, but it's also a powerful tool for simplifying a system, by discarding irrelevant degrees of freedom to highlight some others. It is also an inevitable presence in experiments, where some degrees of freedom are inaccessible due to, e.g., a resolution limit of measurement devices. Even though the coarse-graining plays a role of fundamental importance in statistical physics, its study has been almost confined to thermal equilibrium, with some development in non-equilibrium only in recent years [1,2]. This thesis tries to follow the recent developments in this field, trying to understand the effects of the coarse-graining in particular cases.

In this work we consider systems *out of equilibrium*, for which unfortunately a general theory still doesn't exist. These systems exhibit net current flows between microscopic states, i.e. the transitions rates between any two microstates are not pairwise balanced. This feature is called *broken detail balance* and it's a tell-tale sign of non-equilibrium [3].

There are many areas of application of out of equilibrium physics, and one of the most important of these is living matter. Indeed the typical characteristic of living systems is that they are inherently out of equilibrium [4]. For example a constant consumption and dissipation of energy results in non-equilibrium activity, which lies at the heart of biological functionality: internal activity enables cells to accurately sense and adapt in noisy environments, and for instance it is crucial for high-fidelity DNA transcription and for replication [5–7]. In some cases, non-equilibrium processes also enable subcellular

1

systems to generate forces for internal transport, structural organization and directional motion [8]. Thus, non-equilibrium is an essential characteristic of biological systems and so its study is very important.

Current researches carried out in this area try to find methods to determine the non-equilibrium of living matter through non-invasive, microscopic techniques [9–12]. A recent study [13], conducted on isolated flagella and primary cilia on membranes of living cells, was based on the observation of the steady-state stochastic dynamics of a few mescoscopic degrees of freedom using time-lapse microscopy experiments. The non-equilibrium dynamics of these systems manifests itself as circulating probability currents in a phase space of mesoscopic coordinates, breaking the detail balance condition.

The work done in this thesis fits into this context. In general the intention is to determine the effects of coarse-graining in non-equilibrium. We will consider basically three different procedures of coarse-graining: the first consists on projecting high-dimensional dynamics onto a few preferred degrees of freedom, the second on averaging out dynamics at smaller scales, the third on sampling the time evolution of a system. A question we can ask ourselves is whether these procedures, applied to a condition of non-equilibrium, influence its characteristic of irreversibility. For example, we will see that specific out-of-equilibrium systems may even effectively regain thermodynamic equilibrium at large enough scales, or may simply overlook hidden dissipation from the discarded degrees of freedom.

The approach we use consists of a theretical study, both numerical and analitical, of a model that is at the same time simple and significant, in the sense that it can reproduce the type of results we would see in a typical experiment. More specifically, the model consists in a collection of beads coupled together, and put in a situation of non-equilibrium caused by thermal gradients and/or active forces. The dynamics of the system is then studied in a coarse-grained phase space. The description of the system varies from Langevin equations to Markov jump processes according to the level of wanted resolution. The final objective of the thesis is to provide an exhaustive picture of how a coarse-graining procedure can influence the production of entropy. The solution of this problem can help to establish the correctness of already existing models of non-equilibrium, as well as assist in the analysis of experiments.

The plan for the thesis is the following.

Chapter 1 - Basics of the Model   The «beads-springs» model used for the rest of the thesis is here presented. First, the dynamics of the system was described using both a Langevin and a Fokker-Planck picture. Second, the system was studied from a thermodynamic point of view. In particular, the entropy production rate, present under non-equilibrium conditions, is derived in the steady-state using a stochastical thermodynamics approach. The purpose of these calculations is to give analytical results to be compared to the numerical ones we find after coarse-graining.

Chapter 2 - Coarse Graining A numerical study based on coarse-graining methods is presented, including all the algorithms and methods that are used for finding the results reported in the final chapter. Under a first coarse-graining the dynamics of the system goes from a continuos process to a discrete jumps process. Accordingly, the entropy production rate changes expression and can be calculated under a Markovianity assumption. The steady state dynamics is summarised in plots useful for finding if the system is in equilibrium or not.

Chapter 3 - Results Finally we presented the results obtained from the numerical simulations explained in the previous chapter. The first part deals with the question if it is possible to determine if a system is in equilibrium by looking at the dynamics of particular degrees of freedom using the «probability flux analysis» approach. The second part is more quantitative, and deals with the problem of approaching the correct entropy production rate from the analysis of the coarse-grained dynamics.

Chapter 1

Basics of the Model

We begin this chapter by giving a complete description of the model we will use. Basically, our choice is motivated by the fact that we wanted a relatively simple system to study non-equilibrium. For example, in thermodynamics, the heat transfer between two reservoirs kept at different temperatures is certainly the simplest and probably the most fundamental out-of-equilibrium phenomenon that one can study. Correspondingly we carry this concept to a microscopic level, and study a system of two interacting beads each at a different temperature. Subsequently we increased the complexity of the system by adding more beads (and heat baths) and including active forces.

Most of the ideas in this thesis are illustrated by using a simple stochastic non-equilibrium system with two coupled degrees of freedom. It must be noted that this model doesn't try to mimic any particular natural process. Its strength stands in its simplicity and in the ability to incorporate the essential features for a vast variety of systems. Nevertheless, it is still possible to verify what we theoretically do with a real experiment. For example, our model-system was recently made possible in a laboratory where two viscously coupled particles were trapped with Optical Tweezers and were submitted to an "effective temperature" difference obtained by randomly displacing the position of one of the traps [14,15].

In the first subsection we describe matematically the dynamics of the system. In the second subsection we quantify the non-equilibrium by calculating the entropy production rate. Finally in the third subsection we will add an additional force to describe the active matter, and in the fourth subsection we briefly discuss the effects of anharmonic forces between beads.

1.1 The Equations of Motion

We start this section with a system that consists of two microscopic beads allowed to fluctuate in only one dimension. These beads are connected to each other and to boundary walls by springs with stiffness k as depicted in fig. 1.1 and they are immersed in a solvent

that acts as a viscous fluid with viscosity η. The system is driven out of equilibrium by connecting each bead to local heat baths at temperatures T_1 and T_2 respectively. The relaxation to equilibrium is forbidden by some (undetailed) external constraint which prevents the two thermostats from equilibrating, allowing the system to achieve a nonequilibrium steady state.
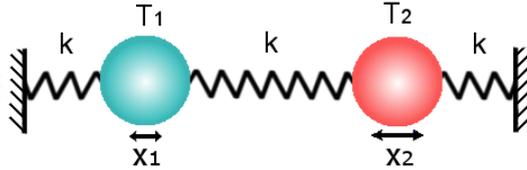

Figure 1.1: Illustration of the model.

To describe the dynamics of the system we have to specify the the forces involved. First, there's the conservative force $-\nabla V$ due to the potential caused by the springs

$$V\left(x_1, x_2\right) = \frac{1}{2}kx_1^2 + \frac{1}{2}kx_2^2 + \frac{1}{2}k\left(x_1 - x_2\right)^2 = k\left(x_1^2 + x_2^2 - x_1 x_2\right), \qquad (1.1)$$

with x_i the displacement of bead i from its equilibrium position. Second, there's the force caused by the collisions with the solvent: since it's clearly not possible to write deterministically this contribution it's necessary to use a statistical approach, hence the forces due to the collisions are treated as a stochastic term. Additionally, we suppose that the two beads are distant enough to anable us to consider the effect of the solvent independent between the two, and we also suppose them to be larger than the molecules of the solvent. Under these hypothesis, the force caused by the collisions acting on the i-th bead can be divided into two contributions: first, an average part $-\zeta\dot{x}_i$ proportional to the velocity and with opposite direction, with ζ called drag coefficient and given, for a spherical particle of radius R, by the Stokes' law $\zeta = 6\pi R\eta$ [16]; second, a random part ξ_i approximately expressed by a Gaussian white noise: in particular the mean and autocorrelation are the following:

$$\langle \xi_i\left(t\right)\rangle = 0, \qquad (1.2)$$

$$\langle \xi_i\left(t\right)\xi_j\left(t'\right)\rangle = 2\zeta k_B T_i \delta\left(t - t'\right)\delta_{ij}, \qquad (1.3)$$

with k_B the Boltzmann constant, $\delta\left(t - t'\right)$ a Dirac's delta to express independence in time, δ_{ij} a Kronecker delta to express independence between beads. The factor in eq. (1.3) is in agreement with the fluctuation-dissipation theorem [17].

The equations of motion of the system are therefore:

$$\begin{cases} \zeta\dot{x}_1 = k\left(-2x_1 + x_2\right) + \xi_1 \\ \zeta\dot{x}_2 = k\left(x_1 - 2x_2\right) + \xi_2 \end{cases} \qquad (1.4)$$

where we have assumed an overdamped regime, i.e. we neglected the inertial part of the equation, using an approximation typical in biophysical system where the Reynolds number is low [18]. This stochastic differential equation is called *overdamped Langevin equation* and stands at the basis of what we will do henceforth. It can be written more compactly in a matrix notation:

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{f}, \tag{1.5}$$

with $\boldsymbol{x} = (x_1, x_2)^T$, $A = \frac{k}{\zeta}\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$ and $\boldsymbol{f} = \zeta^{-1}(\xi_1, \xi_2)^T$ (the T indicates transposition).

Another method, equivalent to the Langevin equation, to describe a stochastic system is to look at the probability distribution $\rho(\boldsymbol{x}, t)$ of finding the system in configuration $\boldsymbol{x}$ at time t. The equation that describes the evolution of $\rho(\boldsymbol{x}, t)$ is called *Fokker-Planck equation*, and can be derived from (1.5) resulting in:

$$\frac{\partial}{\partial t}\rho = -\nabla \cdot (A\boldsymbol{x}\rho - D\nabla\rho) =: -\nabla \cdot \boldsymbol{j}, \tag{1.6}$$

with $\boldsymbol{j}(\boldsymbol{x}, t)$ a probability density current and $D = \zeta^{-1}\begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}$ called diffusion matrix [19].

We can use the Fokker-Plack equation to derive some useful information on the system. In particular we were interested in the steady-state dynamics, so in the following we calculated the steady-state solution $\rho^s(\boldsymbol{x})$ for (1.6). We start using as ansatz a Gaussian distribution:

$$\rho^s(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{\det C}}e^{-\frac{1}{2}\boldsymbol{x}\cdot C^{-1}\boldsymbol{x}}, \tag{1.7}$$

with C a 2×2 symmetric matrix, and with the coefficient determined to impose normalization $\int d^2x\rho^s(\boldsymbol{x}) = 1$. The covariance matrix C can be found inserting the ansatz into the Fokker-Planck equation:

$$0 = \frac{\partial}{\partial t}\rho^s = -\nabla \cdot (A\boldsymbol{x}\rho^s - D\nabla\rho^s)$$

$$\Rightarrow \frac{\partial}{\partial x_i}\left(A_{ij}x_j\rho^s + D_{ij}C_{jk}^{-1}x_k\rho^s\right) = 0$$

$$\Rightarrow -A_{ij}C_{ik}^{-1}x_jx_k - D_{ij}C_{jk}^{-1}C_{il}^{-1}x_lx_k = 0, \tag{1.8}$$

where we used the Einstein notation for the indexes. Relabelling the indexes, noticing that the antisymmetric part of AC^{-1} goes away and remembering that this equation must be

valid for every $\boldsymbol{x}$ gives:

$$-\frac{1}{2}\left(C^{-1}A + AC^{-1}\right) - C^{-1}DC^{-1} = 0$$

$$\Rightarrow AC + CA = -2D. \tag{1.9}$$

Finally, solving manually the two equations system in (1.9) results in the covariance matrix:

$$C = \frac{1}{12k}\begin{pmatrix} 7T_1 + T_2 & 2T_1 + 2T_2 \\ 2T_1 + 2T_2 & T_1 + 7T_2 \end{pmatrix}. \tag{1.10}$$

The Gaussian form for $\rho^s(\boldsymbol{x})$ is due to the linearity of the system and it i preserved even when we add more beads and springs. If the beads would interact with a general anharmonic potential the distribution would be different (see following sections). The probability current density in the steady-state takes the following expression:

$$\boldsymbol{j}^s(\boldsymbol{x}) = \left(A + DC^{-1}\right)\boldsymbol{x}\rho^s(\boldsymbol{x}), \tag{1.11}$$

and it is possible to see that this flux forms a vortex around the center $\boldsymbol{x} = \boldsymbol{0}$.

From $\rho^s(\boldsymbol{x})$ we can find all informations on the stationary system. In particular the average displacementes of the beads is:

$$\langle x_i \rangle = \int d^2x \rho^s(\boldsymbol{x}) x_i = 0, \tag{1.12}$$

by symmetry. Instead the autocorrelation is given by:

$$\langle x_i x_j \rangle = \int d^2x \rho^s(\boldsymbol{x}) x_i x_j =$$
$$= \frac{1}{2\pi\sqrt{\det C}}\frac{\partial}{\partial y_i}\frac{\partial}{\partial y_j}\left[\int d^2x \exp\left(-\frac{1}{2}\boldsymbol{x}\cdot C^{-1}\boldsymbol{x} + \boldsymbol{y}\cdot\boldsymbol{x}\right)\right]_{\boldsymbol{y}=0}, \tag{1.13}$$

and substituting $\boldsymbol{x} = \boldsymbol{z} + C\boldsymbol{y}$ after some semplifications we get:

$$\langle x_i x_j \rangle = \frac{1}{2\pi\sqrt{\det C}}\frac{\partial}{\partial y_i}\frac{\partial}{\partial y_j}\left[\int d^2z \exp\left(-\frac{1}{2}\boldsymbol{z}\cdot C^{-1}\boldsymbol{z} + \frac{1}{2}\boldsymbol{y}\cdot C\boldsymbol{y}\right)\right]_{\boldsymbol{y}=0} =$$
$$= \frac{\partial}{\partial y_i}\frac{\partial}{\partial y_j}\left[\exp\left(\frac{1}{2}\boldsymbol{y}\cdot C\boldsymbol{y}\right)\right]_{\boldsymbol{y}=0} =$$
$$= C_{ij}. \tag{1.14}$$

Given that $\rho^s(\boldsymbol{x})$ is a bivariate Gaussian distribution, the first two moments we have calculated fully specify it. In the next figure we represented in a space $x_1 \times x_2$ the distribution $\rho^s(\boldsymbol{x})$ and the two marginal distributions. It's possible to show that these two distributions are Gaussian even when $\rho^s(\boldsymbol{x})$ is asymmetric, i.e. $T_1 \neq T_2$. This is an interesting feature

because it proves that we cannot determine if a system is in equilibrium or not by looking at the distributions: indeed in our example the system exhibits Gaussian fluctuactions both in and out of equilibrium. In the next sections we will also prove, using anharmonic forces, that non-Gaussian fluctuations are possible even in equilibrium.
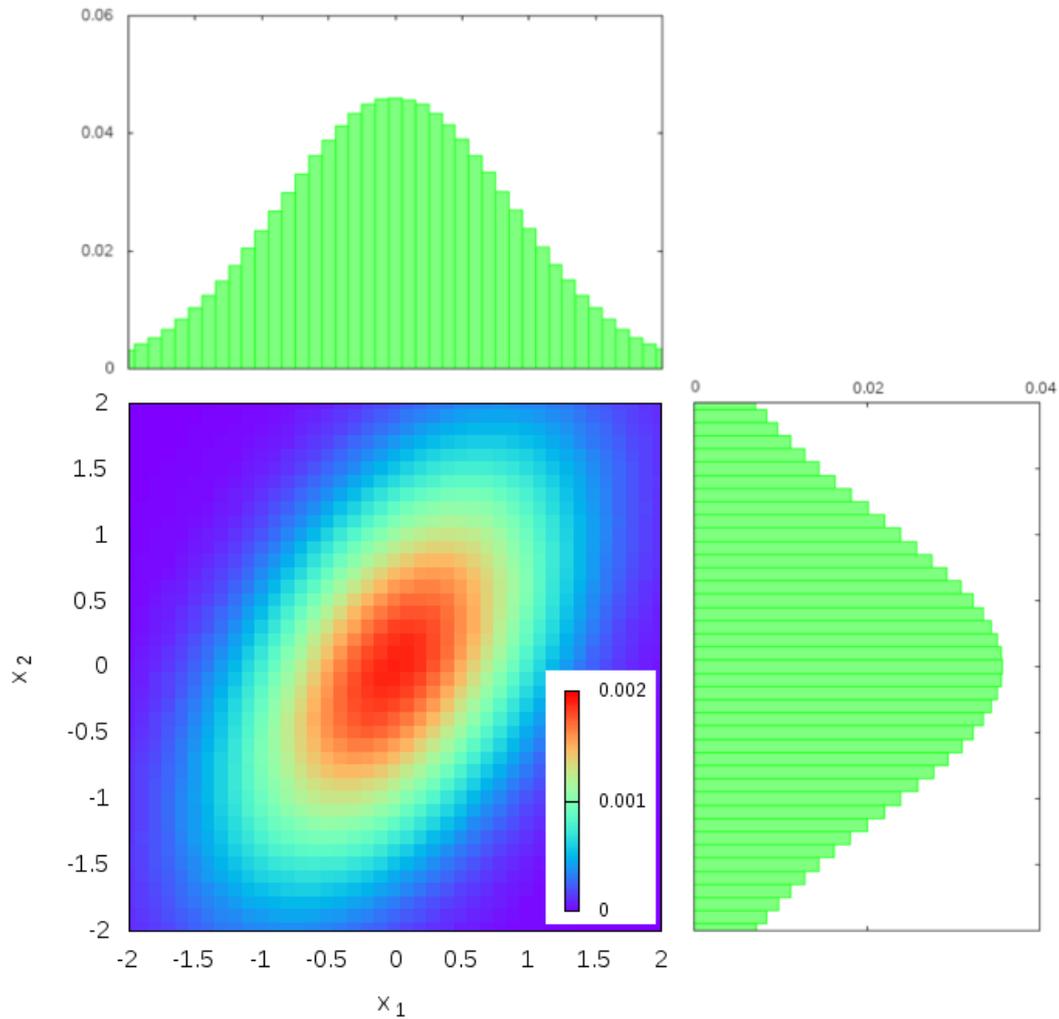

Figure 1.2: The probability density distribution $\rho^s(\boldsymbol{x})$ with its two marginals. Calculated for $T_1 = 1$ and $T_2 = 2$.

A more complex scenario can be obtained by adding a bead to the system. In this case the Langevin equation maintains the same form $\dot{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{f}$ but in three dimensions. An example we used is illustrated in the figure 1.3 where three beads oscillate in one dimension.
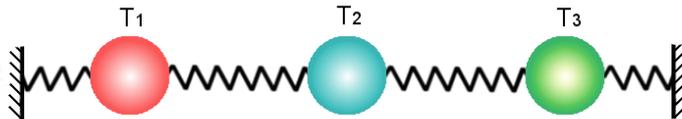
Figure 1.3: Three beads in a row.

The matrix A in this configuration is:

$$A = \frac{k}{\zeta} \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}. \tag{1.15}$$

Another possible configuration is allowing the beads to oscillate in more dimensions. In figure 1.4 the three beads can move in a plane, so every bead has two degrees of freedom. Nevertheless, the same model can be used, and choosing as $\boldsymbol{x} = (x_1, y_1, x_2, y_2, x_3, y_3)$, where y_i is the displacement along the second axis of the ith-bead, we get:

$$A = \frac{k}{4\zeta} \begin{pmatrix} -9 & -\sqrt{3} & 4 & 0 & 1 & \sqrt{3} \\ -\sqrt{3} & -7 & 0 & 0 & \sqrt{3} & 3 \\ 4 & 0 & -9 & \sqrt{3} & 1 & -\sqrt{3} \\ 0 & 0 & \sqrt{3} & -7 & -\sqrt{3} & 3 \\ 1 & \sqrt{3} & 1 & -\sqrt{3} & -2 & 0 \\ \sqrt{3} & 3 & -\sqrt{3} & 3 & 0 & -10 \end{pmatrix}. \tag{1.16}$$

1.2   Entropy Production

There is obviously an interest to quantify to what extent a system is out of equilibrium. Such a quantification could, for example, provide insight into how efficiently molecular motors are able to work together to drive large-scale motions [11].

Qualitatively we can say that in the absence of coupling between the beads, the average rate at which each thermal bath injects energy exactly balances with the rate it absorbs energy due to frictional drag. By coupling the beads, however, there is a net steady-state rate of heat flow from the hot reservoir into the system and out to the cold reservoir. Of course if the reservoirs have the same temperature, the system is in thermal equilibrium.

In this section we calculate the entropy production rate of the system using a theory called stochastic thermodynamics [3, 20]. The starting point is the concept that if a Langevin equation represents the balance of forces on a system, then the Langevin dynam-
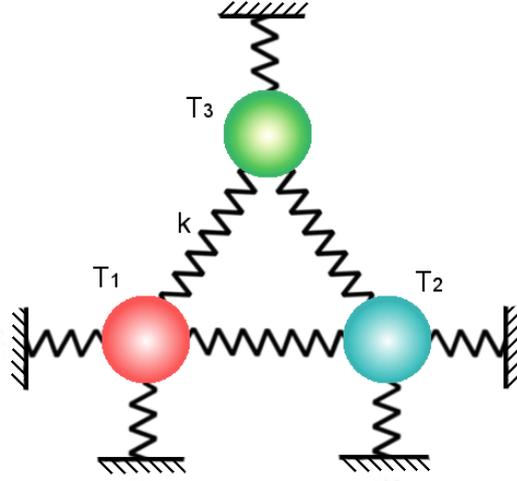
Figure 1.4: Three beads in a plane.

ics conserves the energy of the system plus the surrounding heat bath. This is essentially the first law of thermodynamics

$$đW = dE + đQ \tag{1.17}$$

applyied to an individual trajectory. We use as convention that the heat Q is positive if it's trasferred from the system (the beads) to the environment (the solution), and the work W is positive if it is applied on the system. The energy of the system E is equal to the potential energy V (1.1), because we neglected the kinetic energy by assuming an overdamped motion. Rewriting the Langevin equation we have

$$\mathbf{0} = -\zeta\dot{\boldsymbol{x}} + \boldsymbol{\xi} - \nabla V(\boldsymbol{x}). \tag{1.18}$$

Suppose the system evolves from a configuration $\boldsymbol{x} = (x_1, x_2)$ at time t to $\boldsymbol{x} + d\boldsymbol{x} = (x_1 + dx_1, x_2 + dx_2)$ at time $t + dt$, then multiplying the forces in (1.18) by $-d\boldsymbol{x}$ we get an energy balance equation:

$$0 = -\left(-\zeta\dot{\boldsymbol{x}} + \boldsymbol{\xi}\right) \cdot d\boldsymbol{x} + \nabla V(\boldsymbol{x}) \cdot d\boldsymbol{x} =: dQ + dV, \tag{1.19}$$

where we have identified as dQ the product of $-\left(-\zeta\dot{\boldsymbol{x}} + \boldsymbol{\xi}\right)$, the force exerted by the system to the environment, and $d\boldsymbol{x}$, the displacement of the system. One can write $dQ = dQ_1 + dQ_2$ with:

$$dQ_i = (A\boldsymbol{x})_i \, \dot{x}_i dt, \tag{1.20}$$

the heat absorbed from the bath at temperature T_i. Here and after the product between $\dot{\boldsymbol{x}}$ and a general function of $\boldsymbol{x}$ is evaluated using the Stratonovich rule [3]. In the steady-state

limit, it's possible to show that $dQ_1 = -dQ_2$, in agreement with the fact that $dV = 0$.

Having calculated the heat absorbed from each heat bath, we can now calculate the entropy. It's custumary to divide the total entropy into two contributions. First, the heat dissipated into the environment should obviously be identified with an increase in entropy of the medium S^m such that:

$$dS^m := \frac{1}{T_1} dQ_1 + \frac{1}{T_2} dQ_2. \tag{1.21}$$

Second, there's the Shannon entropy of the system S^{sys} defined as:

$$S^{sys}(\boldsymbol{x}, t) := -k_B \ln \rho(\boldsymbol{x}(t), t), \tag{1.22}$$

where the probability density function $\rho(\boldsymbol{x}(t), t)$ is obtained by solving the Fokker-Planck equation and evaluated along the trajectory $\boldsymbol{x}(t)$.

Then the total entropy production in this time interval is [3, 21]:

$$dS = dS^m + dS^{sys}. \tag{1.23}$$

We proceed to calculate the production rate of entropy by deriving the previous quantities with respect to time. From eq. (1.21) and (1.20) the variation of the entropy of the medium can be written as:

$$\frac{dS^m}{dt} = \left(D^{-1}A\boldsymbol{x}\right) \cdot \dot{\boldsymbol{x}}. \tag{1.24}$$

Instead the variation of the entropy of the system can be calculated using the Fokker-Planck equation (1.6):

$$\begin{aligned}
\frac{dS^{sys}}{dt} &= -\frac{k_B}{\rho(\boldsymbol{x}(t), t)} \left(\frac{\partial \rho(\boldsymbol{x}(t), t)}{\partial t} + \dot{\boldsymbol{x}} \cdot \nabla \rho(\boldsymbol{x}(t), t) \right) \\
&= -\frac{k_B}{\rho(\boldsymbol{x}(t), t)} \frac{\partial \rho(\boldsymbol{x}(t), t)}{\partial t} + \dot{\boldsymbol{x}}(t) \cdot \frac{D^{-1}\boldsymbol{j}(\boldsymbol{x}(t), t)}{\rho(\boldsymbol{x}(t), t)} - \dot{\boldsymbol{x}}(t) \cdot \left(D^{-1}A\boldsymbol{x}(t)\right).
\end{aligned} \tag{1.25}$$

By defining σ the total entropy production rate in the steady-state:

$$\sigma := \left\langle \frac{dS}{dt} \right\rangle = \left\langle \frac{dS^m}{dt} \right\rangle + \left\langle \frac{dS^{sys}}{dt} \right\rangle, \tag{1.26}$$

we notice that the first term in eq. (1.25) vanishes in this limit, and the third term in eq. (1.25) cancels with eq. (1.24), so:

$$\sigma = \left\langle \dot{\boldsymbol{x}}(t) \cdot \frac{D^{-1}\boldsymbol{j}(\boldsymbol{x}(t), t)}{\rho(\boldsymbol{x}(t), t)} \right\rangle. \tag{1.27}$$

The previous average is not a normal average of a quantity involving the position $\boldsymbol{x}(t)$, because it involves also the velocity $\dot{\boldsymbol{x}}$. This complication makes the calculation more delicate. Indeed, to evaluate the average $\langle \dot{\boldsymbol{x}} | \boldsymbol{x}, t \rangle$ conditioned on the position $\boldsymbol{x}$ the Stratonovich

discretization is needed:

$$\langle \dot{\boldsymbol{x}} | \boldsymbol{x}, t \rangle = \lim_{dt \to 0} \frac{\langle \boldsymbol{x}(t+dt) - \boldsymbol{x}(t) | \boldsymbol{x}(t) = \boldsymbol{x} \rangle + \langle \boldsymbol{x}(t) - \boldsymbol{x}(t-dt) | \boldsymbol{x}(t) = \boldsymbol{x} \rangle}{2dt}. \quad (1.28)$$

After some calculations the final result is [3]:

$$\langle \dot{\boldsymbol{x}} | \boldsymbol{x}(t) = \boldsymbol{x} \rangle = A\boldsymbol{x}(t) - D\nabla \log \rho(\boldsymbol{x}, t) = \frac{\boldsymbol{j}(\boldsymbol{x}, t)}{\rho(\boldsymbol{x}, t)}. \quad (1.29)$$

Any subsequent average over position is now trivial leading to:

$$\langle f(\boldsymbol{x}) \dot{\boldsymbol{x}}(t) \rangle = \left\langle f(\boldsymbol{x}) \frac{\boldsymbol{j}(\boldsymbol{x}, t)}{\rho(\boldsymbol{x}, t)} \right\rangle = \int d\boldsymbol{x} f(\boldsymbol{x}) \boldsymbol{j}(\boldsymbol{x}, t). \quad (1.30)$$

So, using the solutions in the steady-state $\rho^s(\boldsymbol{x})$ and $\boldsymbol{j}^s(\boldsymbol{x})$ we found in the previous section, we find from eq. 1.27 and 1.30:

$$\sigma = \int d\boldsymbol{x} \frac{\boldsymbol{j}^s(\boldsymbol{x}) \cdot D^{-1} \boldsymbol{j}^s(\boldsymbol{x})}{\rho^s(\boldsymbol{x})}, \quad (1.31)$$

and performing the integration it results:

$$\begin{aligned}
\sigma &= \zeta \int d\boldsymbol{x} \frac{1}{\rho^s(\boldsymbol{x})} \sum_i T_i^{-1} (j_i^s)^2 \\
&= \zeta \sum_i T_i^{-1} \int d\boldsymbol{x} \left(\sum_j (A + DC^{-1})_{ij} x_j \right)^2 \rho^s(\boldsymbol{x}) \\
&= \zeta \sum_{i,j,k} T_i^{-1} (A + DC^{-1})_{ij} (A + DC^{-1})_{ik} \int d\boldsymbol{x} \rho^s(\boldsymbol{x}) x_j x_k \\
&= \zeta \sum_{i,j,k} T_i^{-1} (A + DC^{-1})_{ij} (A + DC^{-1})_{ik} C_{jk}, \quad (1.32)
\end{aligned}$$

where we have used the expression for D in eq. (1.6) , $\boldsymbol{j}^s$ in eq. (1.11) and the Gaussian integration (1.14). Computing the expression above results in:

$$\sigma = \frac{k}{4\zeta} \frac{(T_1 - T_2)^2}{T_1 T_2}. \quad (1.33)$$

As expected the total entropy production rate is always ≥ 0, vanishes only when $T_1 = T_2$ because the system is in thermodinamical equilibrium, or when $k = 0$ because there's no coupling between the beads. In the graph below I reported how σ varies changing one temperature.
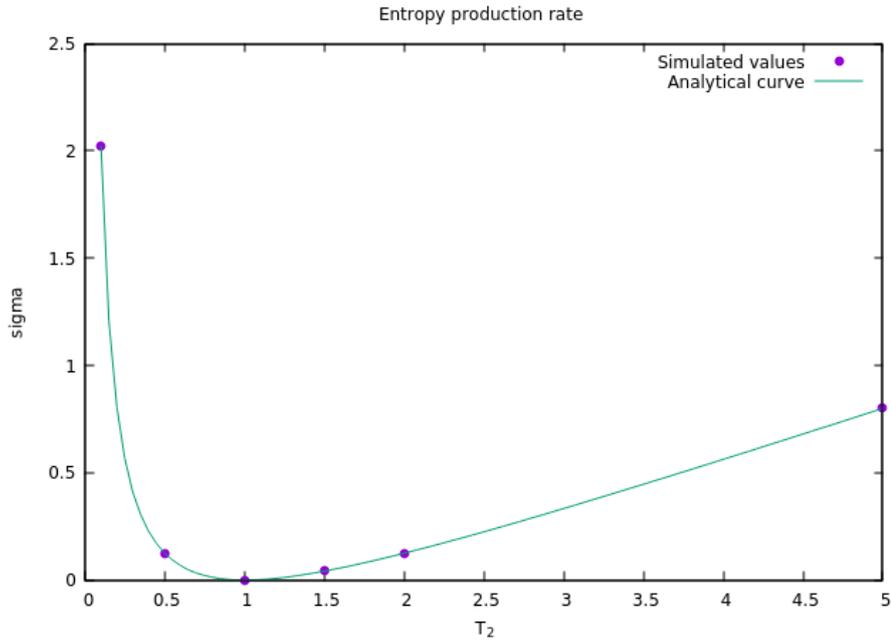
Figure 1.5: Entropy production rate σ as a function of T_2, for fixed $T_1 = 1.0$, $k = 1.0$, $\zeta = 1.0$. The curve is the expression in equation (1.33), the points are the simulated values.

1.3   Active Matter

The non-equilibrium caused by thermal gradients we considered so far does not cover all possible manifestations of non-equilibrium. Therefore we decided to include a new form of non-equilibrium caused by active matter. Active matter is composed of large numbers of active "agents", each of which consumes energy in order to move or to exert mechanical forces. Due to the energy consumption, these systems are intrinsically out of thermal equilibrium [4]. We decided to study this effect by adding to each bead an active force $f_{A,i}(t)$ modelled as a zero-average random telegraph process of amplitude f_0:

$$\langle f_{A,i}(t) \rangle = 0, \tag{1.34}$$

$$\langle f_{A,i}(t) f_{A,j}(t') \rangle = \frac{f_0^2}{4} e^{-|t-t'|/\tau_A} \delta_{ij}. \tag{1.35}$$

The characteristic time constant τ_A can be expressed as $\tau_A^{-1} = \tau_{on}^{-1} + \tau_{off}^{-1}$ (in our case with $\tau_{on} = \tau_{off}$), where τ_{on} (τ_{off}) stands as the mean time to switch on (off) the active motor. Although this is a simple model for the dynamics of motor-generated forces, it's able to describe correctly some important properties, for instance the power spectral density of the position [4,9].
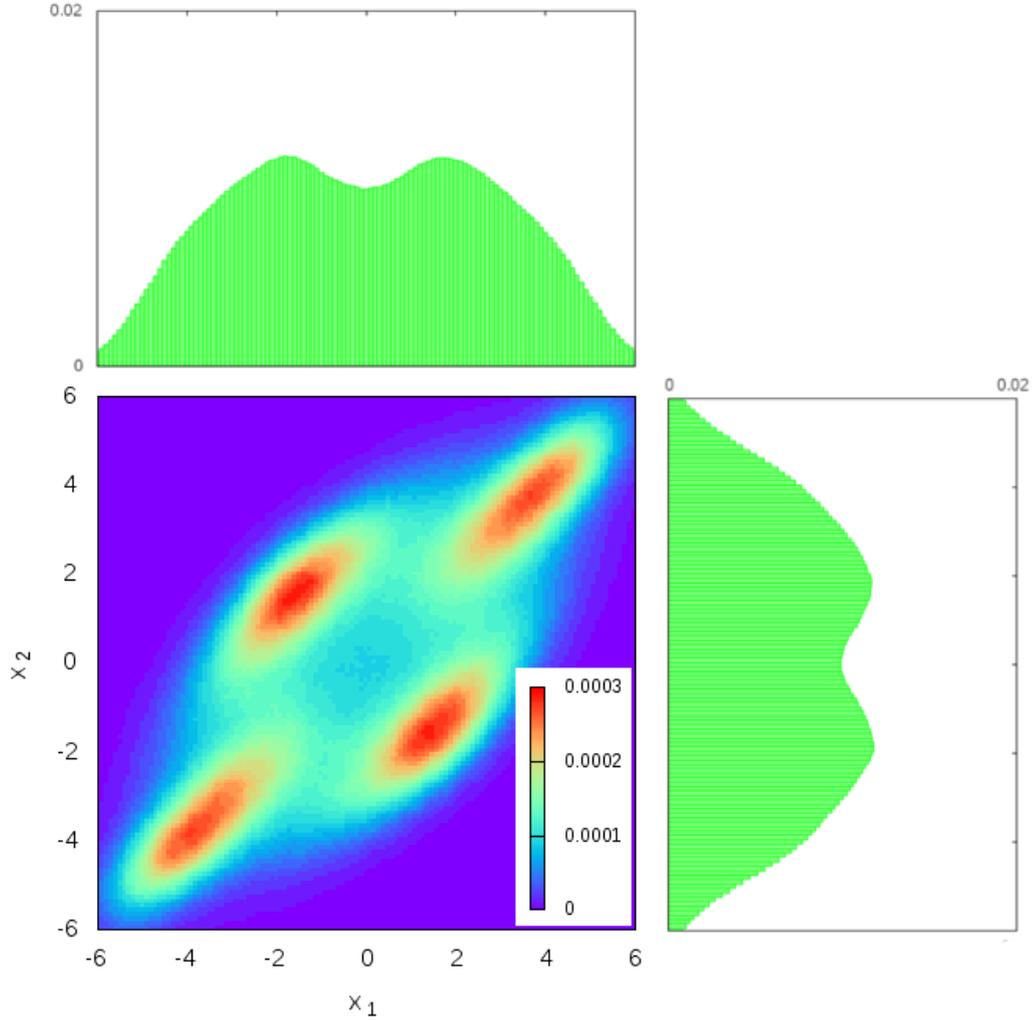
Figure 1.6: The probability density distribution $\rho^s(\boldsymbol{x})$ with its two marginals. Calculated for $T_1 = T_2 = 1$ and $\tau_A = 1$.

1.4 General Coupling

In this section we change for the moment the interaction force between the beads to show what happens to the dynamics of the two-beads system of fig. 1.1 in presence of general anharmonic forces. In particular we chose as interacting potential the following:

$$\tilde{V}(r) = \epsilon \left(-\frac{1}{r^2} + \frac{1}{r^4} \right),$$

with $r = x_2 - x_1 + \sqrt{2}$ (this value ensures that the force is null at $\boldsymbol{x} = \boldsymbol{0}$). The Langevin equation is therefore:

$$\begin{cases} \zeta \dot{x}_1 = -kx_1 - \frac{d\tilde{V}}{dr} + \xi_1 \\ \zeta \dot{x}_2 = -kx_2 + \frac{d\tilde{V}}{dr} + \xi_2 \end{cases} \tag{1.36}$$

No analytical result can be obtained for the distributions we calculated in the harmonic case. Nevertheless solving numerically the dynamics in the next figure we can illustrate a possible example for $T_1 = T_2$. It is possible to show that the two marginal distributions are not Gaussian, so we proved the assertion we made before, that we cannot conclude that a system is out of equilibrium if its fluctactions are not Gaussian.
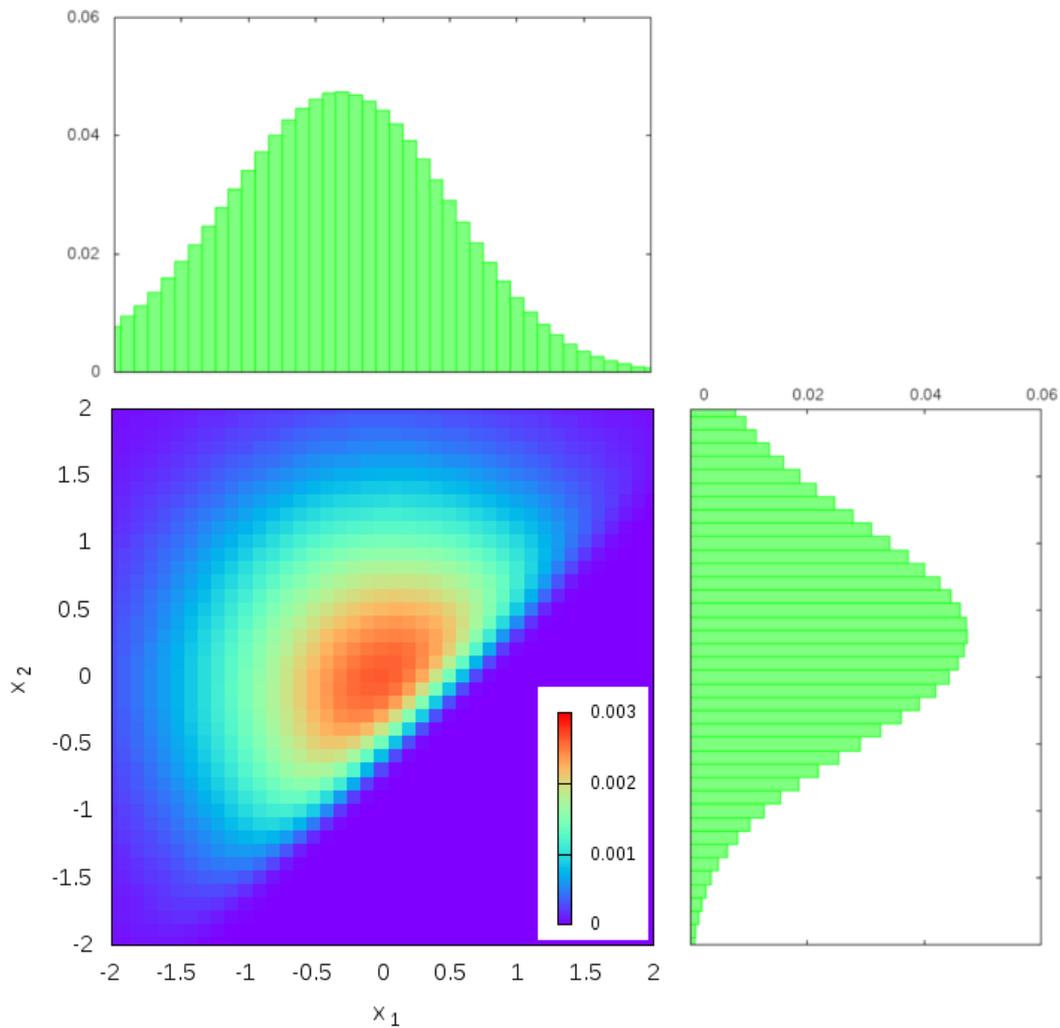

Figura 1.7: The probability density distribution $\rho^s(\boldsymbol{x})$ with its two marginals. Calculated for $T_1 = T_2 = 1$.

Chapter 2

Coarse Graining

In the previous chapter we delineated a model where we were able to study the steady-state dynamics for the simple case of linear-coupled beads in a non-equilibrium situation. In the event of only thermal gradients and no active forces, we even found an analytical formula for the entropy production rate σ. This means that knowing the parameters of the model, we can determine how far the system is from equilibrium by calculating σ.

Considering for the moment a real experiment, if we knew a valid model of the system and we were able to measure all necessary parameters, we could probably calculate σ using a (numerical) approach based on stochastic thermodynamics like we did before. However, this approach is problematic, in the sense that even a good model of the system is always a «coarse-grained» approximation of the system, raising the question of how trustworthy our results would be. Besides, any measurement of the system is limited by the finite resolution of the instruments. So, we may not observe all the required degrees of freedom. These are all reasons for introducing a coarse-graining element in our model. The coarse-graining, in our analysis, is not only a method, but also an object of study. The intention was to establish coarse-grained procedures useful to analyse data and, at the same, consistent with stochastic thermodynamics.

We turn the discussion again about how to tackle an experiment. Most of the developed experimental methods used so far to probe a system for non-equilibrium are called «invasive», and are, for instance, performed by perturbing the system and looking at its response [4]. Such approaches are not ideal for investigating delicate microscopic systems, because they can alter unexpectedly its dynamics. Ideally, one would like to avoid the technical and conceptual difficulties of invasive protocols to probe for non-equilibrium behaviour. This raised the question if it's possible to measure a system's non-equilibrium behaviour simply by looking at it [13]. In the previous section we showed that by looking at a single degree of freedom we couldn't distinguish an equilibrium system from a non-equilibrium one. Recently, though, it was developed a method that indeed uses conventional video microscopy data to detect broken detailed balance, and therefore non-equilibrium behaviour [13]. We will start this chapter by rewriting the methods introduced

by the authors in this reference, and applying them to our model.

Since the purpose of this thesis is not solving simple analytical problem, but rather studying methods that can apply to general systems, we must abandon the view we have developed in the previous chapter. From now on, we have to pretend we don't know the microscopic details of our system, for instance the temperature of our beads. Indeed, we have to assume that we have performed a «non-invasive» experiment, i.e. by only microscopy measurements. The questions we must try to answer are then: is the system in equilibrium? If not, what is its entropy production rate σ? Naturally, when we have these answers, we can compare them with the analytical results we already found.

In this thesis we used a total of three different coarse-graining techniques. These are introduced basically one at each section, and in summary are the following:

1. a discretization of the configurational space (for instance $x_1 \times x_2$), with which we translate the problem from a continous space to a discrete set of states; this coarse-graining mimics a finite resolution of the experimental apparatus; in the end, by doing a «probability flux analysis», it enables to see if the system is in equilibrium (sect. 2.1);

2. a projection of degrees of freedom, which permits to explore the effect of overlooking parts of the system (sect. 2.2);

3. a sampling of the trajectory, with which we studied the effect of memory in the jumping dynamics (sect. 2.3).

2.1 Subdivision of the Configurational Space

In this section we will describe a coarse-grained method which consists of discretizing the continous degrees of freedom of the system, or in other words by partitioning the configurational space into discrete states. This method was used as a natural approach toward modeling the kinetics of chaotic systems. For example, it was used for examining chaotic dynamical systems, with the purpose of studying deterministic chaos through Markov processes [22]. Another reason for using this discretized representation of phase space was to be able to obtain informative results on experimental data with limited statistics [23]. More recently it was used to analyse experimental data of some biological system, in particular a flagellum and cell's cilia [13]. The dynamics of such a system was captured by conventional video microscopy. To quantify this measured stochastic dynamics, they parameterized the configuration of the system by decomposing the shape of the flagellum into normal modes. In their analysis, the mode amplitudes represented time-dependent generalized coordinates of the system, and were the equivalent of our beads displacements.

We decided to start following the same approach to coarse-graining that was used in this research. They used this method to detect non-equilibrium, but without quantifying the entropy production rate or examining the effects of coarse-graining.

In this section, we are going to describe the basis and methodology that can be used to infer broken detailed balance, which is going to be our main test for non-equilibrium. We considered for the moment the system with only two beads. The dynamics was described by the Langevin equation we have written in the previous chapter. The evolutions of the beads displacements $\boldsymbol{x}(t) = (x_1(t), x_2(t))$ were found numerically by integrating this equation using Euler's algorithm, i.e. by discretizing the time in intervals of duration Δt:

$$\begin{cases} x_1(t + \Delta t) = x_1(t) + \frac{k}{\zeta}(-2x_1(t) + x_2(t))\Delta t + \sqrt{\frac{2k_B T_1 \Delta t}{\zeta}}\mathcal{N}(0,1) \\ x_2(t + \Delta t) = x_2(t) + \frac{k}{\zeta}(x_1(t) - 2x_2(t))\Delta t + \sqrt{\frac{2k_B T_2 \Delta t}{\zeta}}\mathcal{N}(0,1) \end{cases} \tag{2.1}$$

where $\mathcal{N}(0,1)$ is a value sampled from a normal distribution with zero mean and unit variance. The previous equations are a very good approximation of the dynamics if Δt is much smaller than the characteristic time of the system, which in this case is ζ/k. For the majority of the simulations we have done, we used unit values for the parameters, i.e. $k = 1$, $\zeta = 1$ and $k_B = 1$; for the temperatures typical values were $T = 1 \div 2$; for Δt we used values 10^{-3} or smaller. The system of equations (2.1) applies in the absence of active forces. In the presence of these forces we need to add a term $f_{A,i}(t)$ to the right-hand side (with i index of the bead), given by the following process:

$$\begin{cases} f_{A,i}(0) = \frac{f_0}{2\zeta}\Delta t, \\ f_{A,i}(t + \Delta t) = \begin{cases} f_{A,i}(t) & \text{if } u_i \geq \frac{\Delta t}{2\tau_a} \\ -f_{A,i}(t) & \text{elsewise} \end{cases} \end{cases} \tag{2.2}$$

with u_i a random number picked uniformly in the interval $(0,1)$ at each time. Indeed, we can notice that the previous system results in a telegraphic process with the properties listed in the previous chapter, so basically a force independent between beads, always with the same absolute value, that switches between positive and negative sign with characteristic time τ_A. We must keep in mind also that we must use a $\Delta t \ll \tau_A$.

A typical solution to eq. (2.1) is illustrated in fig. 2.1. In the following we are going to work always in the steady-state limit of the system. To be sure of that, in the simulations we made the system evolve for a certain time (much longer than τ) before keeping track of the dynamics.

We already explained in the previous chapter that analysing separately x_1 and x_2, we are unable to tell if the system is in equilibrium. Indeed the time evolution in the previous figures are similar to what we are expecting from an equilibrium model. Therefore to gain more information on the system we must keep track of both the degrees of freedom simultaneously. To do so, we traced a «trajectory» of the system in a configurational space $\boldsymbol{x} = (x_1, x_2)$, the same one we used before to plot the joint probability distribution. In fig. 2.2 we plotted an illustrative example of trajectory in this space. The trajectory is clearly a sequence of points $\boldsymbol{x}(t_i)$ with times $t_i = i\Delta t$.
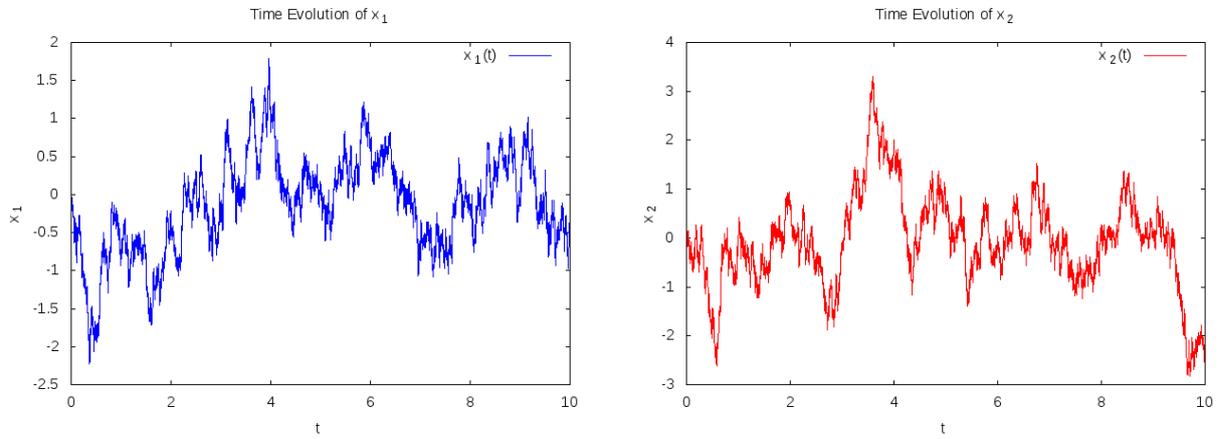
Figure 2.1: Time evolutions of the beads' displacement in the steady-state. Temperatures $T_1 = 1$ and $T_2 = 2$.
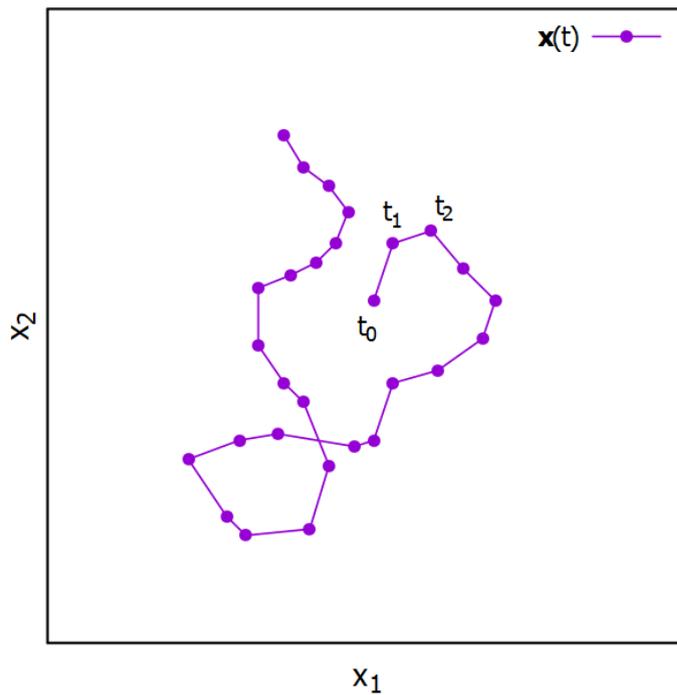

Figure 2.2: Illustrative example of trajectory in the configurational space. Equation (2.1) is solved starting at a point at time t_0, then at every Δt the state of the system is updated and saved in the trajectory.

At this point we introduced the coarse-grained procedure: this configurational space was subdivided into a grid of equally sized, squared cells, each of which represented a discrete state a the system. Such a discrete state encompassed a continuous set of microstates, each of which belonged to a unique, discrete state. In the following figure we showed the new space, which was divided into cells of size $\Delta x \times \Delta x$.
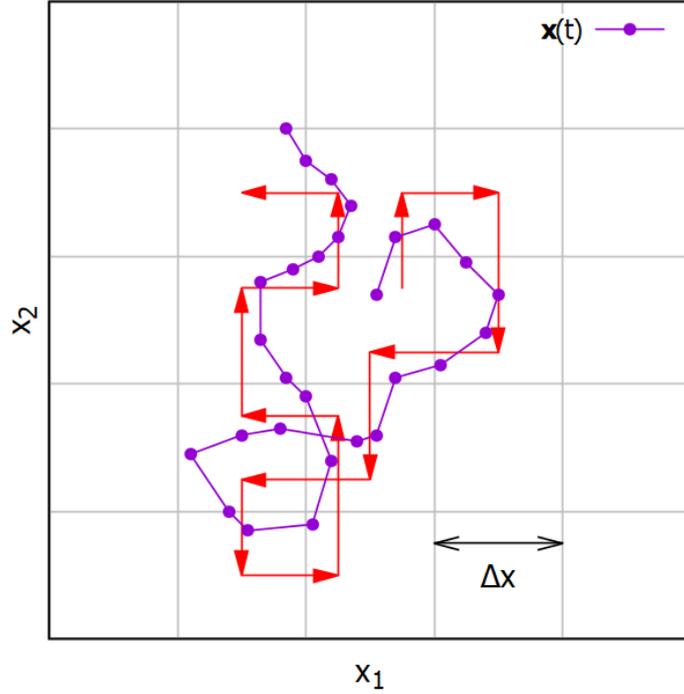

Figura 2.3: Coarse-graining of the configurational space. The trajectory is replaced by jumps between cells (red arrows).

The Langevin trajectory was then associated to a sequence of jumps in this grid. Basically, indicating every cell of the system by the couple of indeces $\boldsymbol{n} = (n_1, n_2) \in \mathbb{Z}^2$, the state of the system $\boldsymbol{x}(t)$ was replaced by $\boldsymbol{n}(t)$ using the cell in which belonged the system at that time. For example, if $\boldsymbol{x}(t)$ moved inside the same cell for a certain time, $\boldsymbol{n}(t)$ didn't change during that time. If instead $\boldsymbol{x}(t)$ leaved its cell, typically the new cell was one at the sides of the starting one. However, in a small fraction of cases, some transitions could go from one cell to a non-adjacent cell in a single timestep due to the limited time resolution: for these cases, we performed a linear interpolation of the measured trajectory to capture all transitions between adjacent cells. So, in our approach we considered only the possibility of jumping between neighbouring states, i.e., using a graph theory terminology, every vertex of our graph had degree 4. Additionally we assumed that our graph was ergodic, i.e., every state could be reached by every other state in a finite time, and the transition rates were reversible, i.e. if a certain jump was possible then also the reverse jump was possible.

Because the description of the system passed from a continuos to discrete view, the equations of motion changed as well, from a Langevin equation to a Master equation. The Master equation is more appropriately the correspondent of the Fokker-Planck equation, and it was the following:

$$\frac{d}{dt} p\left(\boldsymbol{n}, t\right) = \sum_{\boldsymbol{n}' \neq \boldsymbol{n}} \left(p\left(\boldsymbol{n}', t\right) w\left(\boldsymbol{n}'; \boldsymbol{n}\right) - p\left(\boldsymbol{n}, t\right) w\left(\boldsymbol{n}; \boldsymbol{n}'\right) \right),\tag{2.3}$$

where $p\left(n_1, n_2, t\right)$ was the probability of being inside the cell and $w\left(n_1, n_2; n_1', n_2'\right)$ the rates of jumping from cell (n_1, n_2) to cell (n_1', n_2'). This rates presumed that the dynamics had a «memoryless» or «Markovian» property: in the following sections we will see that this hypothesis was not correct, but for the moment we are going to ignore this problem. Under the Markovianity assumption and steady-state conditions, the system was entirely describeded by associating at every cell of the space the steady-state solution of eq. (2.3), which we indicated $p^s\left(n_1, n_2\right)$, and the rates $w\left(n_1, n_2; n_1', n_2'\right)$. Because we considered only the possibility of jumping to an adjacent cell we had $w\left(n_1, n_2; n_1', n_2'\right) \neq 0$ only if $n_1' = n_1 \pm 1$ and $n_2' = n_2 \pm 1$.

An analytical determination of p^s and w was not possible, so in this context the simulated trajectory came useful. We could indeed calculate these quantities numerically after having captured a trajectory long enough to have repeatedly spanned over the cells. The probability $p^s\left(n_1, n_2\right)$ was estimated as:

$$p^s\left(n_1, n_2\right) = \frac{t\left(n_1, n_2\right)}{t_{tot}},\tag{2.4}$$

where $t\left(n_1, n_2\right)$ was the accumulated time spent in the state (n_1, n_2) and t_{tot} is the total time of the trajectory simulated. Similarly the rates were estimated as:

$$w\left(n_1, n_2; n_1', n_2'\right) = \frac{N\left(n_1, n_2; n_1', n_2'\right)}{t_{tot} \cdot p\left(n_1, n_2\right)},\tag{2.5}$$

with $N\left(n_1, n_2; n_1', n_2'\right)$ the total number of recorded jumps from (n_1, n_2) to (n_1', n_2'). A useful quantity to consider was the probability flux $j\left(n_1, n_2; n_1', n_2'\right)$, which could be defined from the Master equation as:

$$j\left(\boldsymbol{n}; \boldsymbol{n}'\right) := -p\left(\boldsymbol{n}', t\right) w\left(\boldsymbol{n}'; \boldsymbol{n}\right) + p\left(\boldsymbol{n}, t\right) w\left(\boldsymbol{n}; \boldsymbol{n}'\right).\tag{2.6}$$

From this quantity we defined a vector $\boldsymbol{j}\left(n_1, n_2\right)$, which represented the net flux of probability that exited from a cell, in agreement with eq. (2.5):

$$\boldsymbol{j}\left(n_1, n_2\right) := \frac{1}{t_{tot}} \begin{pmatrix} N\left(n_1, n_2; n_1 + 1, n_2\right) - N\left(n_1 + 1, n_2; n_1, n_2\right) \\ N\left(n_1, n_2; n_1, n_2 + 1\right) - N\left(n_1, n_2 + 1; n_1, n_2\right) \end{pmatrix},\tag{2.7}$$

In the following image we illustrated the quantities introduced so far.
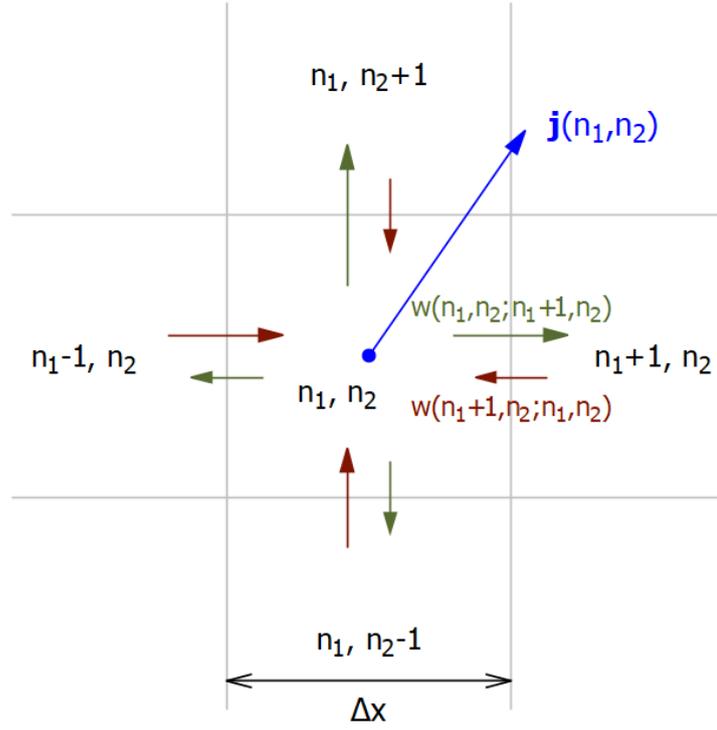
Figure 2.4: Illustration of the dynamics near a cell. A cell is indicated by one index for each dimension, i.e. $\boldsymbol{n} = (n_1, n_2)$. At every side it has an outward $w(\boldsymbol{n}; \boldsymbol{n}')$ and inward $w(\boldsymbol{n}'; \boldsymbol{n})$ rate. Finally every cell has a flux vector $\boldsymbol{j}$ attached, which represents the direction and intensity of the net probability flux.

With the coarse-graining procedure we have defined in the previous pages we developed a tool to identify non-equilibrium. Indeed using the fact that when a system reaches thermodynamical equilibrium, not only it becomes stationary in time, but also it requires that transition rates between any two states to be pairwise balanced, we have:

$$j\left(\boldsymbol{n}; \boldsymbol{n}'\right) = 0 \quad \forall \boldsymbol{n}, \boldsymbol{n}' \iff \text{equilibrium.}$$

The previous statement is called detailed balance principle, and its connection to entropy and non-equilibrium was first discovered by Boltzmann [24]. Therefore, to show if the system was in equilibrium or not, we needed only to calculate $\boldsymbol{j}(\boldsymbol{n})$ for every cell, using the expression in (2.7), and see if the fluxes were all zero. For example, in the next figure we plotted a «flux map», i.e. the flux vector $\boldsymbol{j}(\boldsymbol{n})$ at every cell. It's possible to see without further analysis that the system was out of equilibrium because these fluxes were clearly present and formed a vortex around the center.
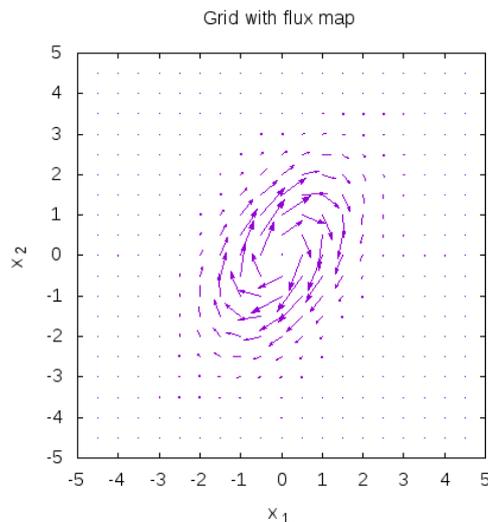
Figure 2.5: Example of probability flux map. For every cell we calculated its flux $\boldsymbol{j}$ that we plotted here as an arrow. A non-zero vorticity of the fluxes indicates broken detailed balance, so this system was out of equilibrium (indeed we used $T_1 = 1$ and $T_2 = 2$).

The discussion we have done so far allows only to discern equilibrium from non-equilibrium. What remains now is to quantify the non-equilibrium. We are going to do that in next part of this section, by calculating the total entropy production rate, in analogy to what we have done in the previous chapter. Now it's not clear if the entropy we are going to calculate will differ from the previous quantity, because when we introduced a coarse-graining we basically gathered different microstates into discrete states, leading to a loss of information and a change in the amount of disorder of the system. The effects of the coarse-graining procedure are illustrated in chapter 3.

The entropy of the system can be defined as anology to the Langevin case:

$$S^{sys}(t) := -\log p(\boldsymbol{n}(t), t), \qquad (2.8)$$

with $p(\boldsymbol{n}(t), t)$ the solution to the master equation (2.3) applied to a single trajectory $\boldsymbol{n}(t)$. The entropy of the medium instead can be derived considering the time reversal conterpart of our trajectory. Let's call γ the trajectory formed by the succession of occupied states (cells) $\boldsymbol{n}_0, \boldsymbol{n}_1, \boldsymbol{n}_2, \ldots$, and $\bar{\gamma}$ the reversed trajectory. Then it's possible to show that the ratio of the probabilities of the direct trajectory γ and the reversed counterpart $\bar{\gamma}$ is [25]:

$$\frac{\mathbb{P}(\gamma)}{\mathbb{P}(\bar{\gamma})} = \prod_i \frac{w(\boldsymbol{n}_i; \boldsymbol{n}_{i+1})}{w(\boldsymbol{n}_{i+1}, \boldsymbol{n}_i)}. \qquad (2.9)$$

This ratio can be associated to the variation of the entropy of the medium during the trajectory [3]:

$$\Delta S^m = \log \frac{\mathbb{P}(\gamma)}{\mathbb{P}(\bar{\gamma})}, \qquad (2.10)$$

so we can identify an entropy production due to a single transition $\boldsymbol{n}_i \to \boldsymbol{n}_{i+1}$ as:

$$dS^m = \log \frac{w\left(\boldsymbol{n}_i; \boldsymbol{n}_{i+1}\right)}{w\left(\boldsymbol{n}_{i+1}, \boldsymbol{n}_i\right)}. \tag{2.11}$$

The total entropy production was therefore given by the sum of the two term we just found:

$$\Delta S^{tot} = \Delta S^{sys} + \Delta S^m. \tag{2.12}$$

In the steady-state the contribution at the total entropy production rate due to the system is zero, so we have:

$$\sigma = \left\langle \dot{S}^{tot} \right\rangle = \left\langle \dot{S}^m \right\rangle = \sum_{\boldsymbol{n}, \boldsymbol{n}'} p^s\left(\boldsymbol{n}\right) w\left(\boldsymbol{n}, \boldsymbol{n}'\right) \log \frac{w\left(\boldsymbol{n}, \boldsymbol{n}'\right)}{w\left(\boldsymbol{n}', \boldsymbol{n}\right)}. \tag{2.13}$$

It's possible to prove that σ can be written also with the following formula:

$$\sigma = \sum_{\boldsymbol{n}, \boldsymbol{n}'} p^s\left(\boldsymbol{n}\right) w\left(\boldsymbol{n}, \boldsymbol{n}'\right) \log \frac{p^s\left(\boldsymbol{n}\right) w\left(\boldsymbol{n}, \boldsymbol{n}'\right)}{p^s\left(\boldsymbol{n}'\right) w\left(\boldsymbol{n}', \boldsymbol{n}\right)}. \tag{2.14}$$

This equation shows clearly that the breaking of detail balance is the cause of the entropy production. To conclude, in this section we showed how to determine if the system satisfies detail balance, and so equilibrium, by plotting the probability entropy fluxes. Then with eq. (2.13) we are able to quantify how far drom equilibrium is the system.

2.2   Projecting Degrees of Freedom

It is important to note that for a system in steady-state dynamics, broken detailed balance is direct evidence of non-equilibrium, but showing that a system obeys detailed balance in a subspace of coordinates is insufficient to prove equilibrium. Indeed, even for systems out of equilibrium, broken detailed balance is not necessarily apparent at the supramolecular scale [26]. In this section we worked with a system with more than two degrees of freedom, and we showed exactly what happens when one or more of these are ignored.

Let's suppose, without loss of generality, to have a system of three beads in a row, like the example we showed in chapter one. Then we could perform the same coarse-graining procedure we have done for the two-dimensional case. The only differences are that the configurational space is three-dimensional, the cells are cubic boxes with six adjacent boxes, etc. Basically the same results we achieved in the previous section are still valid, in particular the results for the entropy production rate σ (eq. (2.13)) and the fact that the system is in equilibrium if and only if the three-dimensional probability fluxes are all zero. In other words, if we were able to observe the stochastic motion of all beads in the system, we could measure the full probability current and extract information about the complete

non-equilibrium dynamics of the system. However, we wanted to study what happens by only tracking two of the three beads in the chain. This analysis tries to describe a typical experiment where only a small subset of the degrees of freedom can be tracked.

When we completely forget one of the beads, we have only two degrees of freedom and we recover the situation of the previous section. Moreover, we can repeat the analysis three times, in each case ignoring a different bead. In the end, we have three different flux maps and three different σ's. The flux maps can all be thought as a projection of the three-dimensional fluxes onto the planes perpendicular to the direction of the ignored degree of freedom, so we decided to represent these fluxes in figures such as the following example.
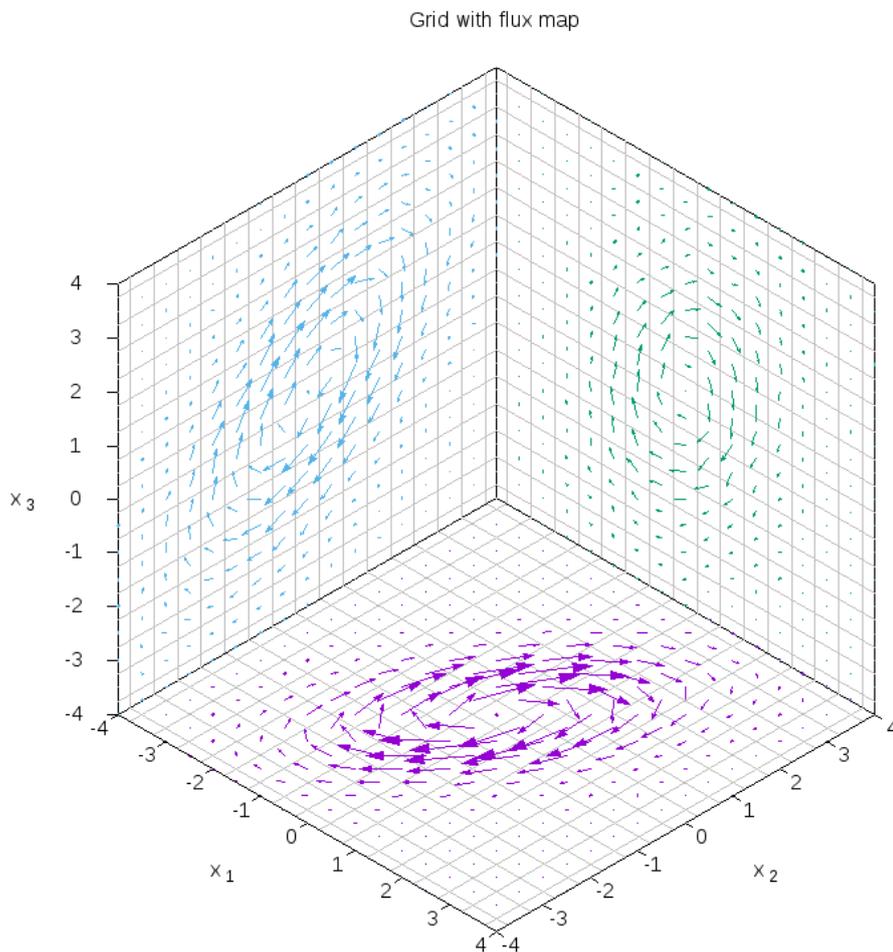
Figura 2.6: Projection of the 3-dimensional fluxes into the Cartesian plans. In the plane $x_1 \times x_2$ we have the flux map in the case we don't track x_3, etc for the other planes. Temperatures: $T_1 = 1$, $T_2 = 2$ and $T_3 = 3$.

Let's calculate instead the entropy production rate σ^i when we ignore the i-bead in

the chain. For example, assuming the elimination of the third bead (the rightmost one), we can find, after integrating x_3 from the total Fokker-Planck equation, that the system is described by a multivariate steady-state probability density $\tilde{\rho}^s(\tilde{\boldsymbol{x}})$ and probability flux $\tilde{\boldsymbol{j}}^s(\tilde{\boldsymbol{x}})$, with $\tilde{\boldsymbol{x}} = (x_1, x_2)$ given by:

$$\tilde{\rho}^s(\tilde{\boldsymbol{x}}) = \frac{1}{2\pi\sqrt{\det\tilde{C}}}e^{-\frac{1}{2}\tilde{\boldsymbol{x}}\cdot\tilde{C}^{-1}\tilde{\boldsymbol{x}}},$$

$$\tilde{\boldsymbol{j}}^s(\tilde{\boldsymbol{x}}) = \tilde{A}\tilde{\boldsymbol{x}}\tilde{\rho}^s - \tilde{D}\tilde{\nabla}\tilde{\rho}^s,$$

with the matrixes given by:

$$\tilde{A} = \left(\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right) + \left(\begin{array}{c} A_{13} \\ A_{23} \end{array}\right)\left(\begin{array}{cc} C_{31} & C_{32} \end{array}\right)\left(\begin{array}{cc} C_{11}^{-1} & C_{12}^{-1} \\ C_{21}^{-1} & C_{22}^{-1} \end{array}\right), \qquad (2.15)$$

$$\tilde{D} = \frac{1}{\zeta}\left(\begin{array}{cc} T_1 & 0 \\ 0 & T_2 \end{array}\right), \qquad (2.16)$$

which can be thought as an effective interaction matrix $\tilde{A} \in \mathbb{R}^{2\times 2}$ and effective diffusion matrix $\tilde{D} \in \mathbb{R}^{2\times 2}$ for the reduced system composed by the first two beads. Finally it's possible to prove the following inequality:

$$\sigma - \sigma^3 = \int d\boldsymbol{x}\frac{\boldsymbol{j}^s(\boldsymbol{x})\cdot D^{-1}\boldsymbol{j}^s(\boldsymbol{x})}{\rho^s(\boldsymbol{x})} - \int d\tilde{\boldsymbol{x}}\frac{\tilde{\boldsymbol{j}}^s(\tilde{\boldsymbol{x}})\cdot\tilde{D}^{-1}\tilde{\boldsymbol{j}}^s(\tilde{\boldsymbol{x}})}{\tilde{\rho}^s(\tilde{\boldsymbol{x}})} \geq 0,$$

after expliciting all the integrands [12]. This means that $\sigma^i \leq \sigma$ (for the general bead) with σ the entropy production calculated without losing degrees of freedom. This result is somewhat obvious, because by losing degrees of freedom we lose also information on the system and so entropy. In the extreme case of keeping track of only one bead, the entropy production is zero, as we showed that this sub-system alone is undistinguishable from equilibrium.

In the next chapter we show some interesting cases where by eliminating a bead the manifestation of the non-equilibrium properties changed drastically.

2.3 Sampling the data

In this section we studied the effects of the Markovianity assumption on the Master equation dynamics. In particular we are going to introduce a coarse-graining method with the intention to make more valid this memoryless assumption. In the next chapter we proved in detail that the dynamic is indeed non-Markovian, using two methods: (i) showing that the time of permanence in a cell is not exponential; (ii) showing that the jumping rates depend directly from the memory.

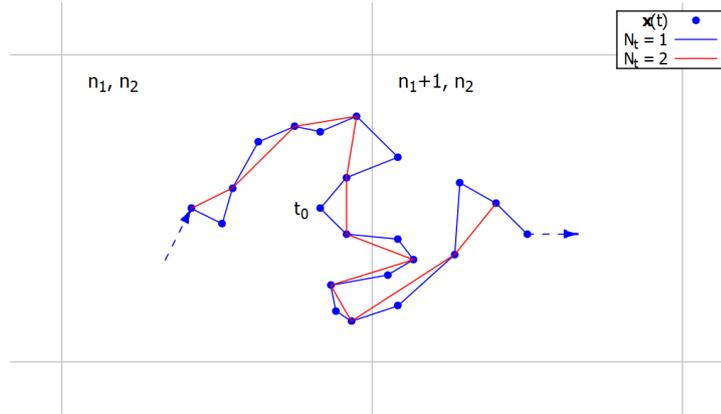First of all, we need to introduce a new set of states and rates that keep track of some

Figure 2.7: The blue points are $\boldsymbol{x}\left(t_i\right)$, the simulated evolution of the system, and the blue line is the line connecting these points. In red the trajectory using $N_t = 2$, i.e. by jumping avery two points.

memory, and we decided to consider only the last cell occupied different from the present one. Therefore, from now on a state of the system at time t is described by the three following indeces: n_1 and n_2 which express the position of the cell at time t, and

$$\alpha \in \{"up","down","right","left"\},$$

which stands for the position of the previous cell relative to the present one. To better explain the concepts in this section we illustrated an example in figure 2.7, plotting in blue the simulated trajectory of the system. In this example, at point labelled by t_0, the state is given by $(n_1, n_2, \alpha) = (n_1, n_2, "right")$. The probability of being in the state (n_1, n_2, α) is $P(n_1, n_2, \alpha)$, and it was estimated it by dividing the number of times the trajectory occupied this state by the total points in the trajectory, as it was done in the previous analysis. The total probability of a cell is thus

$$P(n_1, n_2) = \sum_{\alpha} P(n_1, n_2, \alpha).$$

In the example, assuming that previously the trajectory was on the left and in no other time them trajectory passed again in this cell, we have $P(n_1, n_2, "right") = 6/N$ with N the total length of the trajectory. Finally, we considered the rate of going from state (n_1, n_2, α) to state (n_1', n_2', α') using the notation $w(n_1, n_2, \alpha, \beta)$, where

$$\beta \in \{"up","down","right","left"\}$$

is the possible direction of the jump, e.g. β is the opposite of α' in this notation. In the example, using the available points $w(n_1 + 1, n_2, "left", "left") = 2/(9 \cdot \Delta t)$.

Now that we have completely defined a method that takes into account a memory

effect on the dynamic, we can study the effect of the memory and how to reduce it. If the dynamic were Markovian we would have $w(n_1, n_2, \alpha, \beta) = w(n_1, n_2, \alpha', \beta)$ for $\alpha \neq \alpha'$, but this was not generally the case. As a way to measure the importance of this effect of memory, we decide to consider the following figure of merit:

$$f := \sum_{n_1, n_2} \sum_{\beta} P(n_1, n_2) \frac{sd_\alpha [w(n_1, n_2, \alpha, \beta)]}{mean_\alpha [w(n_1, n_2, \alpha, \beta)]}, \qquad (2.17)$$

where sd_α and $mean_\alpha$ are the standard deviation (the square root of the variance) and the average of the rates varying the index α. The second factor of f is a measure of how much different are the rates with respect to the memory. The first factor is the probability of the cell and is introduced to weight more the states that are more important in the dynamic. For a Markovian process $f = 0$ because the rates are equal with respect to α (such that $sd_\alpha(w) = 0$). It's possible to say then that a system with lower f than another is «more Markovian». So we can modify our system trying to minimize f, with the effect of reducing the (unwanted) importance of memory.

To do so, we introduce a «coarse-grain on the time» by keeping track of only some positions along the trajectory. In particular, if we have a trajectory $\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, ...$, with $\boldsymbol{x}_n$ calculated at time $t_n = n\Delta t$, we can always reduce to a new trajectory that is $\boldsymbol{x}_0, \boldsymbol{x}_{N_t}, \boldsymbol{x}_{2N_t}, ...$, i.e. using only a subset of points of the initial trajectory. In figure 2.7 I illustrated the case with $N_t = 2$. Clearly, by changing N_t, every quantity described previously can vary. Indeed, for instance, the rates can change drastically, as its clear by looking at the figure. In the next chapter we show the results of entropy and memory-effects when the figure of merit is minimized. Since the dynamics with the best N_t verifies more closely the assumption of Markovianity that is needed in the calculations we have done previously, we expect that the values of σ are closer to what we have calculated analytically. This will prove that an additional coarse-grain is needed to find results compatible with stochastic thermodynamics.

Chapter 3

Results

In this final chapter we are going to show some significative examples of the methods we introduced previously. In the first section we worked with the Probability Flux Analysis we described in sections 2.1-2.2. Basically we chose various cases where the effect of coarse-graining was more evident, to show how it affects the properties of the system. Depending on the effect we wanted to study, we chose a particular model among those introduced in chapter 1, fixed the parameters (i.e. the temperatures) and performed the simulation and coarse-graining procedure of chapter 2. The purpose of this section was to test if the probability flux analysis could be a method to probe the system for non-equilibrium behaviour.

In the section 3.2 we showed the results regarding the property of Markovianity of the coarse-grained trajectory and the sampling of data procedure we discussed in section 2.3. After a first part where we reported the tests to prove the non-Markovianity of the system, we applied our sampling procedure to some interisting cases. The purpose of this last section was to test if using this method it is possible to recover a quantitative correct value of the entropy production rate.

3.1   Probability Flux Analysis

In the last chapter we already showed what we called a «flux map», i.e. a plot of the configurational space ($x_1 \times x_2$ for a two-beads system) divided in cells with an arrow, representing the flux of probability exiting from the cell, attached to each one. We made the claim that it was possible, from looking at the plot, to decide if the system was in equilibrium or not. In principle we can say that if there are non-zero fluxes, the system breaks detail balance and so it is out of equilibrium. Vice versa, if all the fluxes are zero the system is in equilibrium. In this section we put this claim to the test.

More precisely the expression «non-zero» must be used carefully. Indeed, the finite length of experimental or computer-simulated trajectories limits the accuracy with which we can estimate fluxes in configurational space. This is because a short trajectory does not

travel across the cells enough times to generate a good statistics of jumps from which we can calculate the fluxes. Therefore, it is important to determine if the estimated currents are statistically signicant from zero. The method we use consists in repeating the simulation more times in order to have a large set of different and independent trajectories. Let's call $\boldsymbol{j}_i(\boldsymbol{n}) = (j_{1,i}(\boldsymbol{n}), j_{2,i}(\boldsymbol{n}))$ the vector flux calculated from the i-th generated trajectory for cell $\boldsymbol{n}$. Then from this set of values one can study the statistics of the fluxes: we decide to use as global flux $\bar{\boldsymbol{j}}(\boldsymbol{n})$ the average of the set, and as a test to measure the discrepancy from zero the value:

$$\sqrt{\left(\frac{\text{mean}(j_{1,i})}{\text{sd}(j_{1,i})}\right)^2 + \left(\frac{\text{mean}(j_{2,i})}{\text{sd}(j_{2,i})}\right)^2}, \tag{3.1}$$

also called *compatibility* with zero [27]. The compatibility measures how many «sigmas» of difference there are between the average of a set and zero. If the compatibility has a low value, i.e. less than 2, we can conclude that the set of fluxes is not statistically different from zero. To conclude, we make a flux map using for each cell an arrow with length $\bar{\boldsymbol{j}}(\boldsymbol{n})$ and with a color that indicates its compatibility.

We start our analysis using the simplest model of two beads with no active forces, and we expect to see fluxes statistically different from zero only when $T_1 \neq T_2$. Some parameters of the system are not important in this analysis so we decided to fix them throughout all the simulations to the following values: $k = 1$, $\zeta = 1$, $k_B = 1$. The timestep of the simulation was typically $\Delta t = 10^{-3 \div -4}$ and the total length of the trajectory was $10^{9 \div 10}\Delta t$. The timestep value was chosen with the intention of making the average jump smaller than the width of the cell $\Delta x \sim 0.1$.
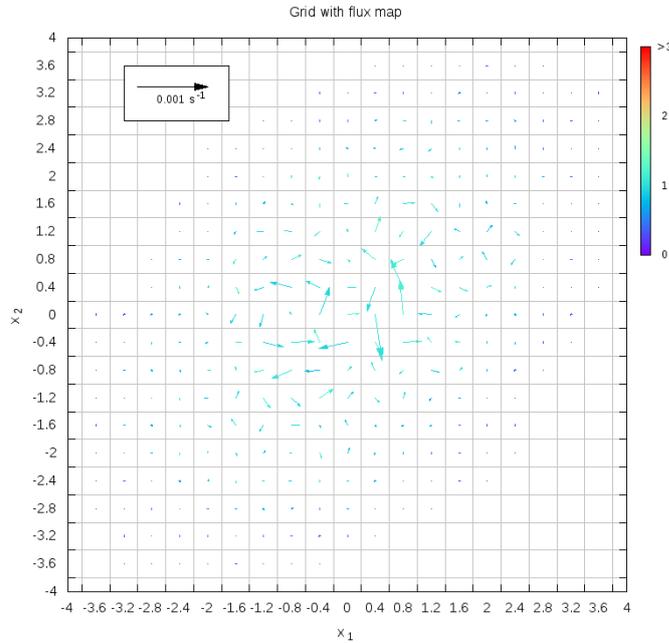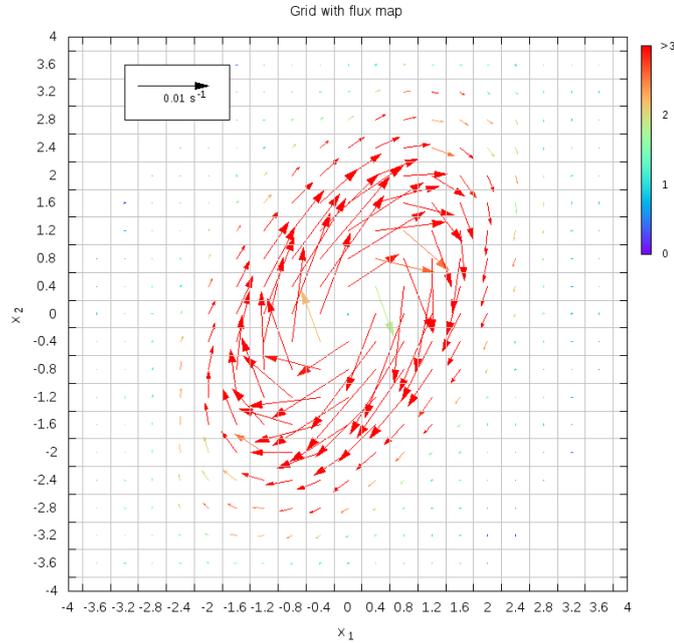

Figura 3.1: Two beads with no active force and $T_1 = T_2$. Grid size $\Delta x = 0.4$. The color of the arrows represents the compatibility (eq. (3.1)).

The trajectory length instead was chosen long enough as to provide a good statistic. In the two plots here we illustrate two exemplary cases: in the first (fig. 3.1) the temperatures are $T_1 = T_2 = 1$, in the second (fig. 3.2) $T_1 = 1$ and $T_2 = 2$. We can clearly notice how the plots prove the assertion we made before: in the case of equilibrium the fluxes are nearly zero, and are also pointing at random directions, reinforcing the idea that there is no net flow of probability; in the case of non-equilibrium all the fluxes are significantly different from zero, and they make a pattern in the configurational space. As shown in the legend of the figures, the absolute value of the fluxes in the first case was smaller (by nearly an order of magnitude) than the second case. The particular value of the equilibrium case is connected to the trajectory length, in particular when the trajectory increases the flux diminishes, because the statistic improves and we approach the correct value of zero fluxes.

Figura 3.2: Two beads with no active force and $T_1 = 1$, $T_2 = 2$. Grid size $\Delta x = 0.4$. The color of the arrows represents the compatibility (eq. (3.1)).

We reported below two additional cases regarding the two-beads system. In the first (fig. 3.3) we modified the spring constants and drag friction coefficients to obtain an asymmetric system. From a simulation with $T_1 = T_2$ it results an equilibrium flux map, therefore proving that the result obtained before was not a conclusion of symmetry, and that this method works also for asymmetric systems. In the next case (fig. 3.4) we used a symmetric system with the two beads at the same temperature but adding an active force to both. The result shows that an active force is a source of non-equilibrium behaviour. The breaking of detailed balance in this case does not arise from energy exchange between the beads, but rather from how stochastic motor forces induce position correlations.
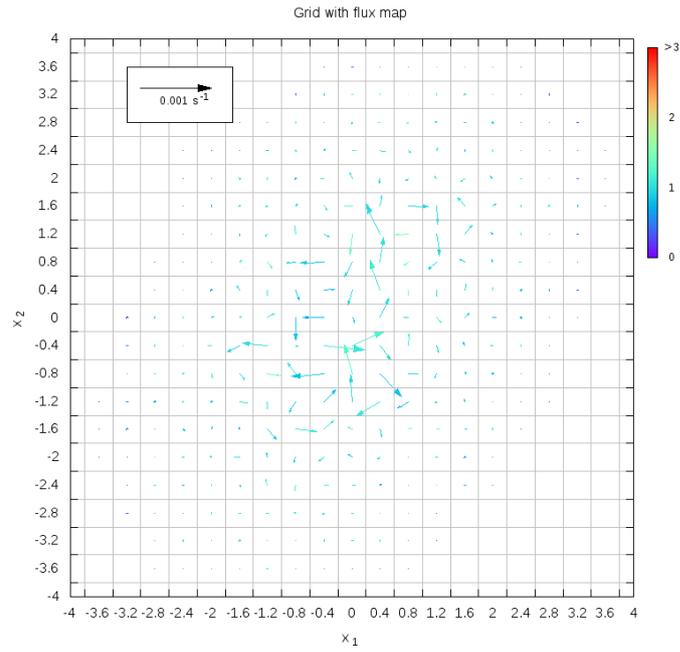
Figura 3.3: Asymmetric two beads system with no active force $T_1 = T_2 = 1$. The spring constants were in order from left to right 2, 1 and 1, and the drag coefficients $\zeta_1 = 1$ and $\zeta_2 = 2$. Grid size $\Delta x = 0.4$. The color of the arrows represents the compatibility (eq. (3.1)).
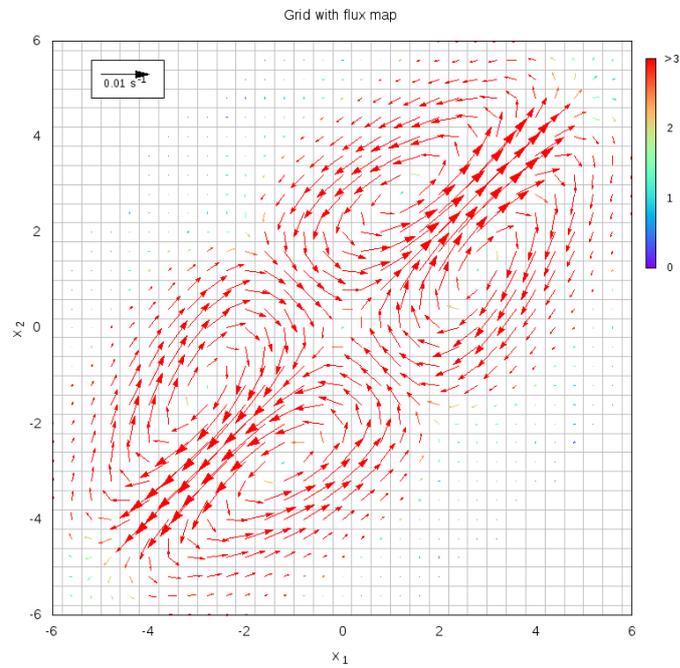

Figura 3.4: Two beads with active force $(f_0 = 10$ and $\tau_A = 10)$ to both beads and $T_1 = T_2 = 1$. Grid size $\Delta x = 0.4$. The color of the arrows represents the compatibility (eq. (3.1)).

In this last part of the section we reported some results related to the system with three beads and the projection in the configurational space. Technically the simulation and analysis of a three-dimensional (or more) system is the same as a two-dimensional one, and the same conclusions as before can be obtained. Anyhow, as already explained in the previous chapter, in this thesis we tried to study the effect of neglecting degrees of freedom (DoF) by keeping track of only a subset of these. We performed this analysis by projecting the three-dimensional fluxes into a plane perpendicular to the DoF eliminated. The three possible choices of these planes can all be represented in the same plot as shown in the previous chapter (fig. 2.6). Therefore from these plots it is possible to determine if a particular DoF is essential to spot a non-equilibrium behaviour or not.
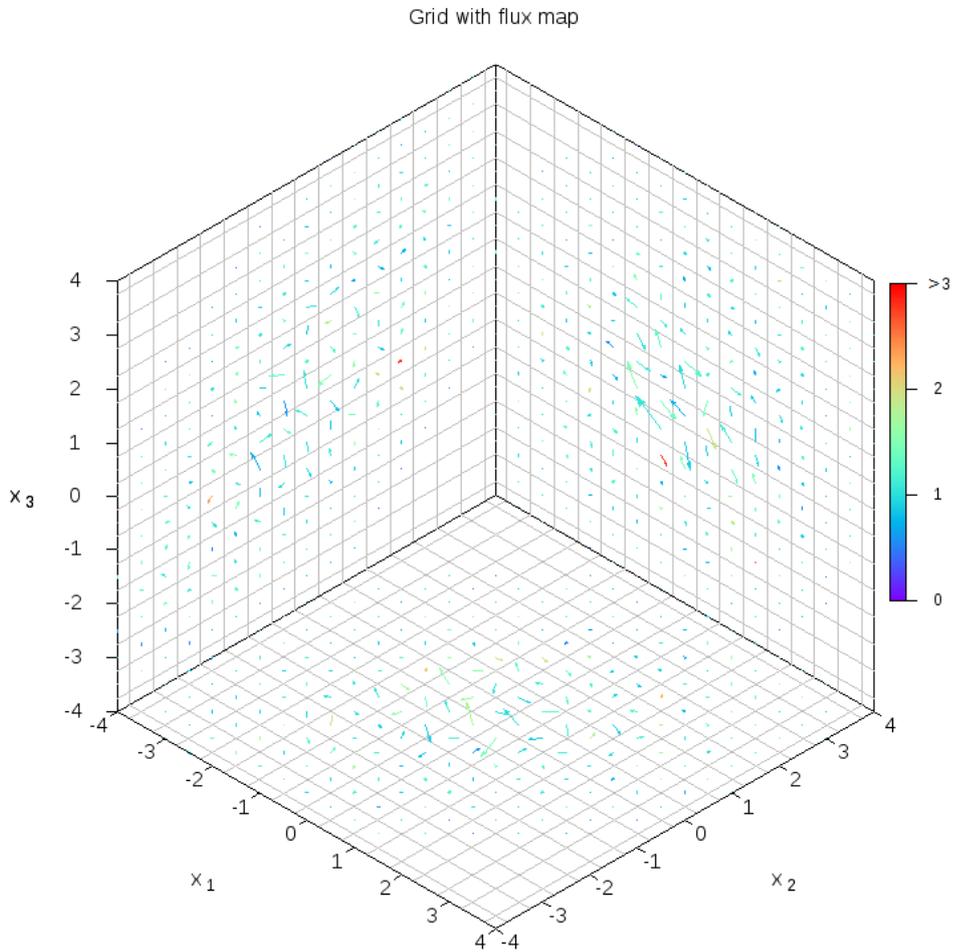

Figura 3.5: Three beads in a line system. No active forces and $T_1 = T_2 = T_3 = 1$. Grid size $\Delta x = 0.5$. The color of the arrows represents the compatibility (eq. (3.1)).

We decide to use as first system the three beads on a line with no active forces (fig. 1.3). The irrelevant physical parameters were put to the same values as the two-bead system,

so the important values are only the three temperatures. If these have all the same value, then the system is clearly in equilibrium, and, accordingly, all the three planes show no significant fluxes (fig 3.5). Instead, if one bead has a different temperature from the other two, then there are two possible cases depending on whether the bead is one at the ends of the chain (fig. 3.6) or the one in the middle (fig. 3.7).
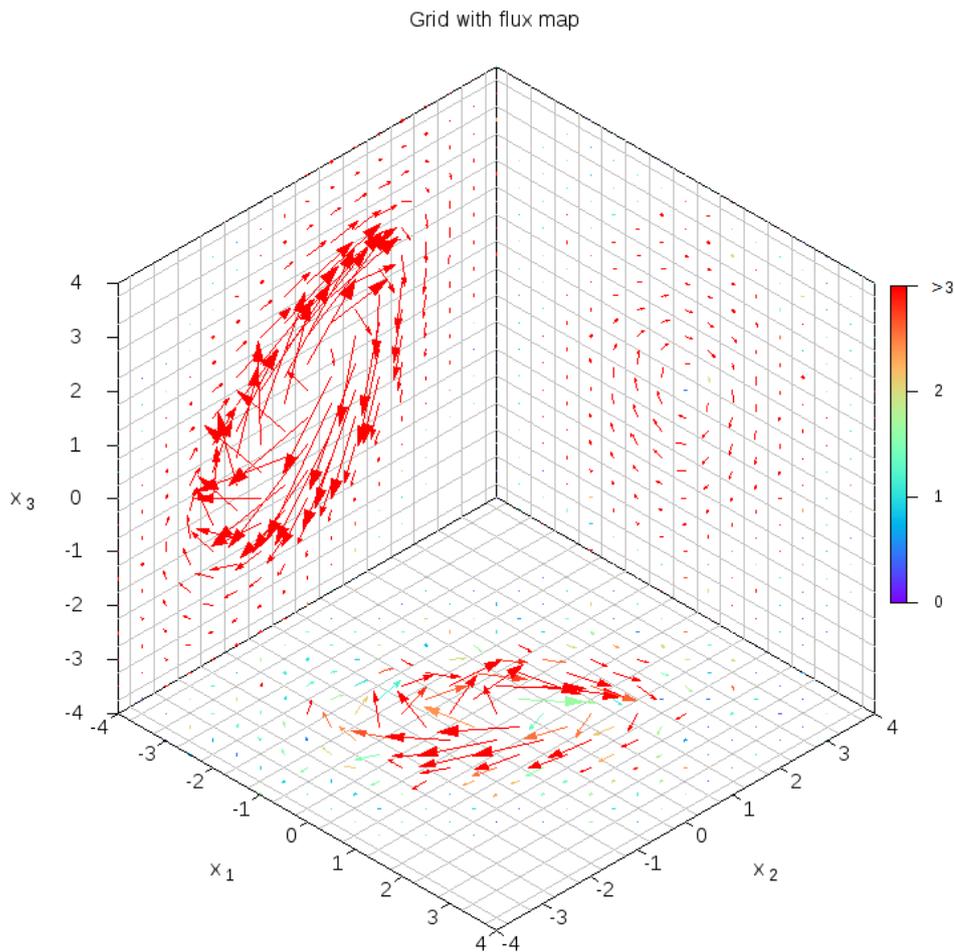

Figura 3.6: Three beads in a line system. No active forces and $T_1 = T_2 = 1$ and $T_3 = 2$. Grid size $\Delta x = 0.5$. The color of the arrows represents the compatibility (eq. (3.1)).
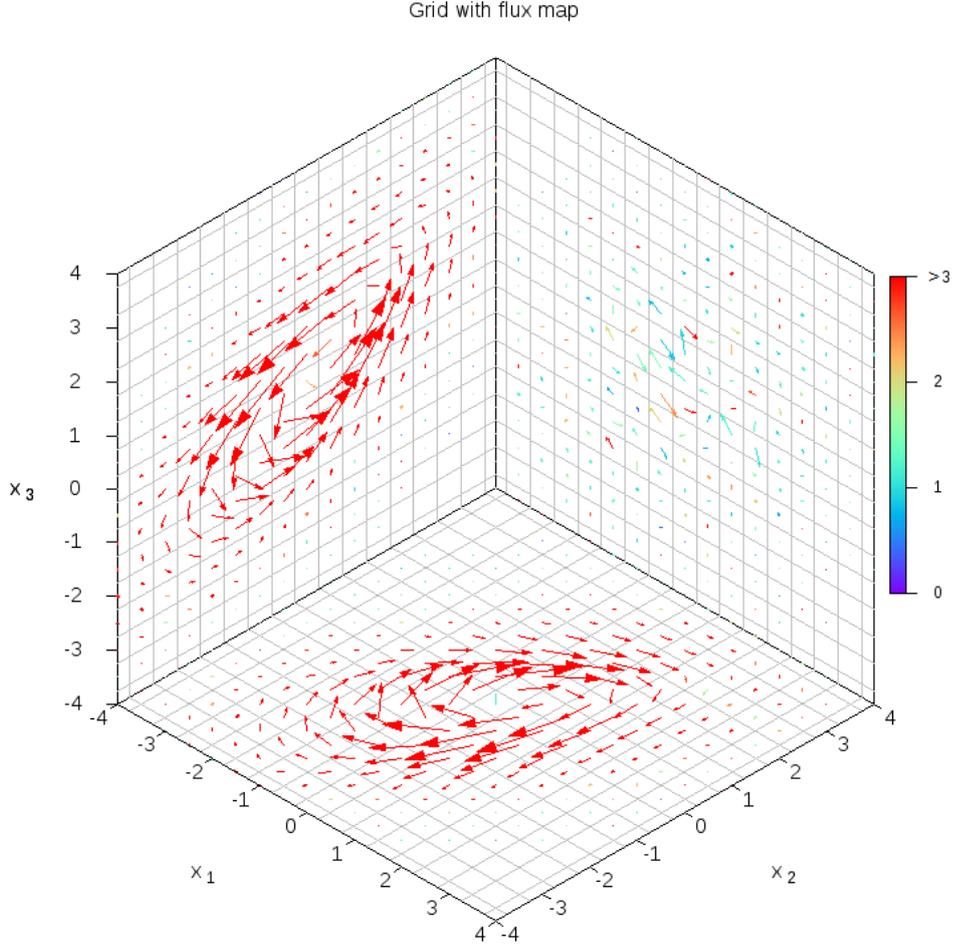
Figura 3.7: Three beads in a line system. No active forces and $T_1 = T_3 = 1$ and $T_2 = 2$. Grid size $\Delta x = 0.5$. The color of the arrows represents the compatibility (eq. (3.1))

In the first case (fig. 3.6) the flux map shows that the system's non-equilibrium behaviour is noticeable from each plane: even by looking only at the 1st and 2nd beads, which are at the same temperature, we can conclude that the system is out of equilibrium. In the second case the different bead is the one in the center (the 2nd): the system is out of equilibrium but that is evident only by looking at planes $x_1 \times x_2$ and $x_2 \times x_3$. Indeed the plot shows that the fluxes computed by tracking only the beads at the extremities aren't statistically different from zero, and moreover have random directions. This behaviour is caused by the symmetry of the system and it does not appear in presence of asymmetries. Nevertheless this is not a case that we can overlook, and therefore we have to conclude that when we have not traced all the DoF a flux map that shows an equilibrium bahaviour can possibly conceal a non-equilibrium system. Instead, it is obvious that if the flux map shows a non-equilibrium behaviour, then the only possible conclusion is that the system is truly out of equilibrium.

In the next figures we reported more cases, these times removing the thermal gradient $(T_1 = T_2 = T_3)$ and adding active forces to some beads. Figure 3.8 and 3.9 are the equivalents respectively of fig. 3.6 and 3.7. They show that the analyses and conclusions we have reached so far using a thermal gradient as source of non-equilibrium are generally the same as when we use active forces, therefore indicating that our approach to the study of non-equilibrium is suitable to very general systems. In figure 3.10 we used instead an asymmetric system, showing that the reason we couldn't detect non-equilibrium in the previous case was indeed due to the particular symmetry.
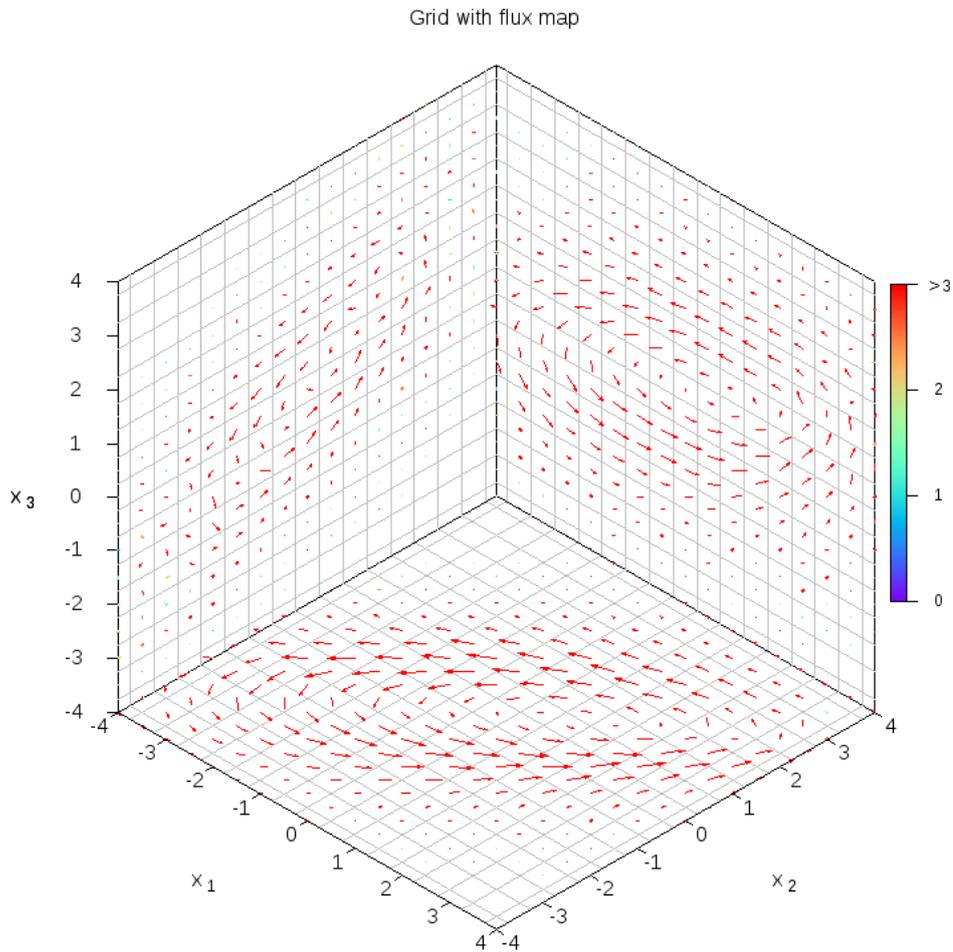

Figura 3.8: Three beads in a line system. $T_1 = T_2 = T_3 = 1$ and active force ($f_0 = 10$ and $\tau_A = 10$) on the 1st bead. Grid size $\Delta x = 0.5$. Same flux scale on all planes. The color of the arrows represents the compatibility (eq. (3.1)).
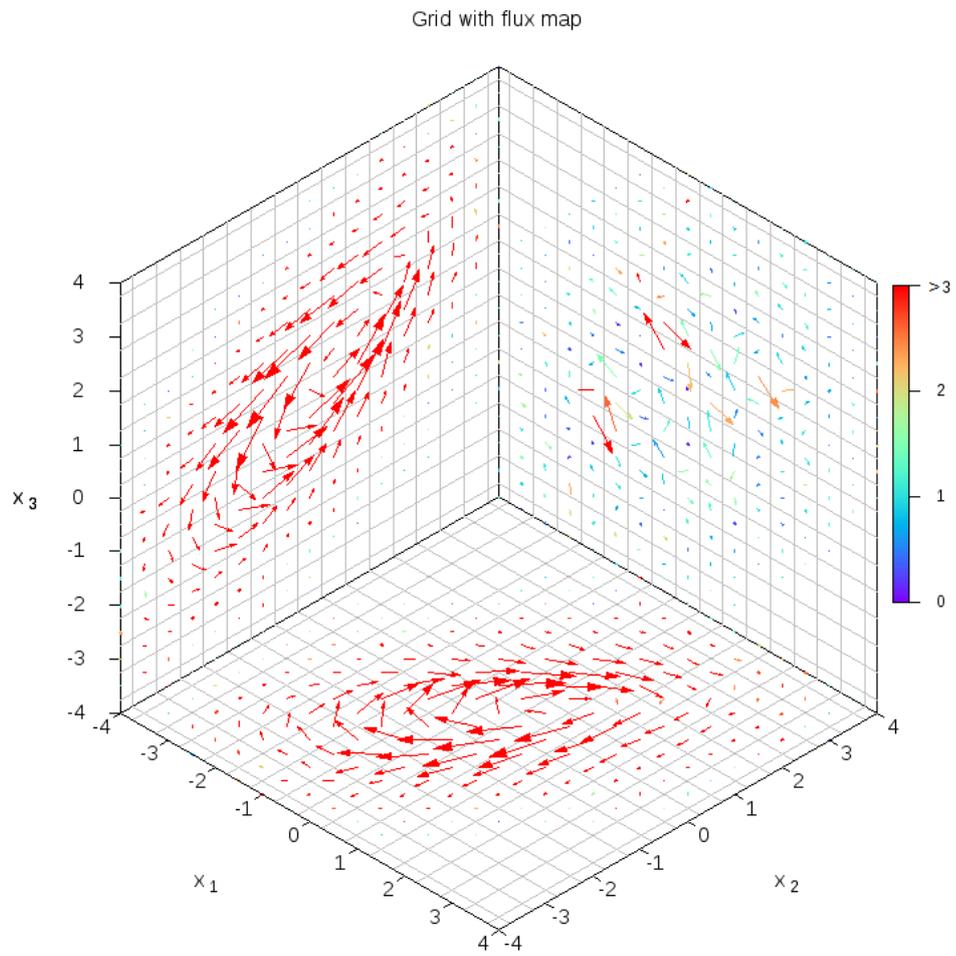
Figura 3.9: Three beads in a line system. $T_1 = T_2 = T_3 = 1$ and active force ($f_0 = 5$ and $\tau_A = 1$) on the 2nd bead. Grid size $\Delta x = 0.5$. Flux scale of plane $x_1 \times x_3$ is $5\times$ the others. The color of the arrows represents the compatibility (eq. (3.1))
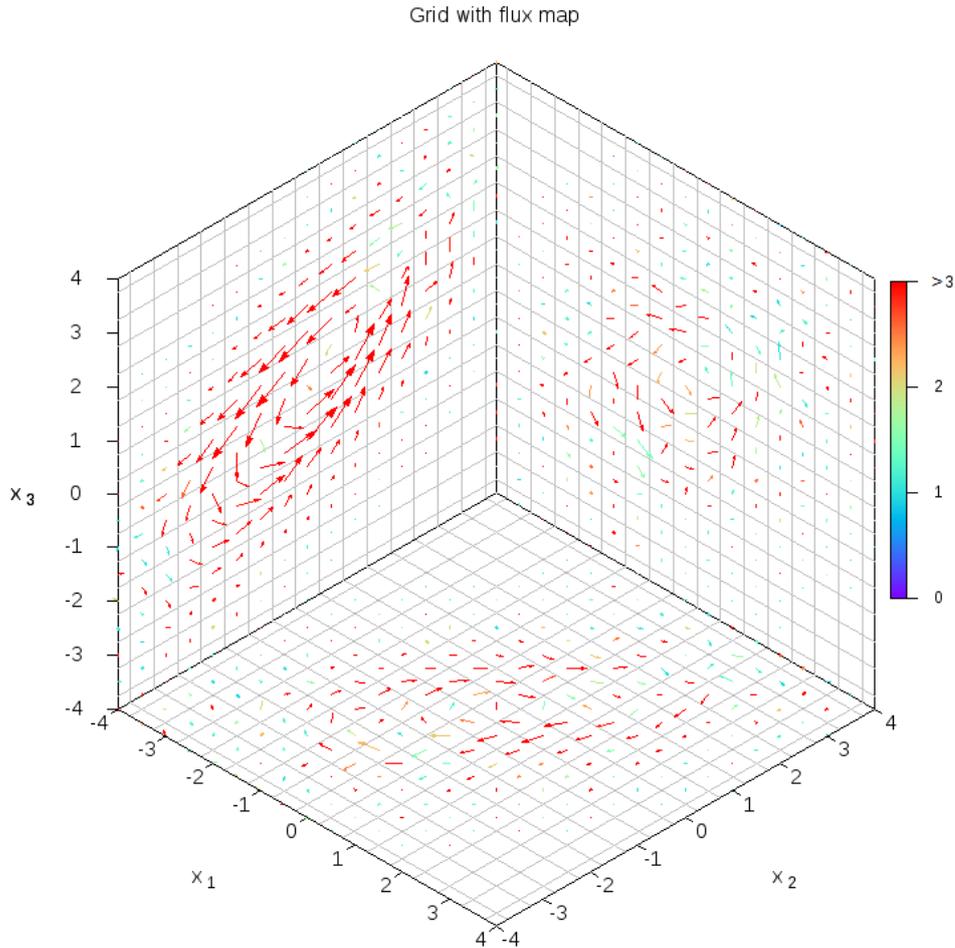
Figura 3.10: Asymmetric three beads in a line system. Three different drag coefficients: $\zeta_1 = 0.5$, $\zeta_2 = 1$ and $\zeta_3 = 2$. $T_1 = T_2 = T_3 = 1$ and active force ($f_0 = 5$ and $\tau_A = 10$) on the 2nd bead. Grid size $\Delta x = 0.5$. Same flux scale on all planes. The color of the arrows represents the compatibility (eq. (3.1)).

3.2 Entropy and Markovianity

In this section we are going to show the results regarding the entropy production rate and the coarse-graining method that consists on sampling the data. It is possible to notice that without this method the entropy production rate σ calculated from the discretized trajectory is different from the theoretical one σ^{th} calculated using stochastic thermodynamics. Our claim is that the error stands in the assumption of Markovianity used to calculate σ, and in the first part of this section we are going to prove that the jumps' dynamic is indeed non-Markovian. Next, we will use the sampling method described in section 2.3 to increase the level of Markovianity, and we will find that this procedure

increases the accuracy of σ as well.

To prove that the jumps' dynamic is non-Markovian we use two methods:

1. by simply looking at the rates $w\,(n_1, n_2, \alpha, \beta)$ we can notice that there is a dependence on the index α which represents the previous cell occupied: this is a direct evidence of a memory property; this is the reason why the values of the figure of merit f reported in the following part are greater than zero;

2. by looking at the time of permanence t in a cell: in the Markov case it should have an exponential distribution $\propto \exp\left(-t/\tau\right)$, whereas in our case we find a different distribution for short times; in the next figure we reported an example.
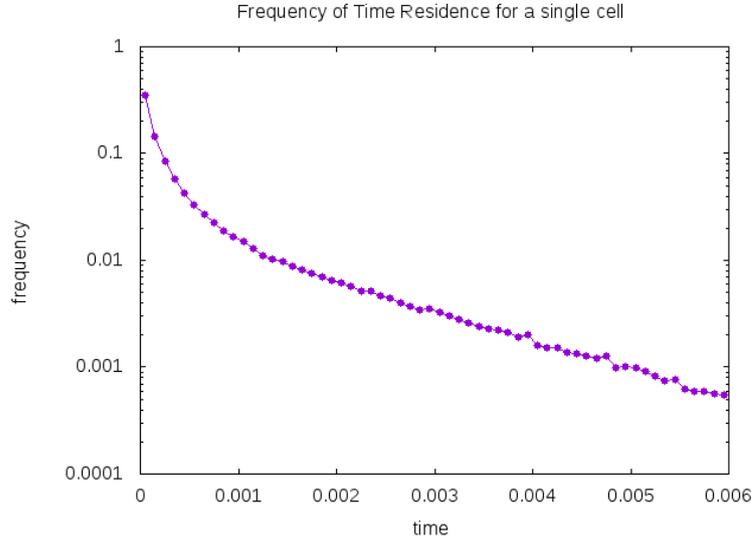

Figura 3.11: Distribution of the time permanence in a cell (notice the log-scale in the y-axis).

We now apply the sampling method to the case of two beads at different temperatures and no active forces. We can choose for example the case of $T_1 = 1$ and $T_2 = 2$, and all the other parameters set as 1. The theoretical entropy production rate is given from eq. (1.33) and is $\sigma^{th} = \frac{k}{4\zeta}\frac{(T_1 - T_2)^2}{T_1 T_2} = 0.125$. Changing the sampling parameter N_t (the number of Δt-jumps to skip) we obtaine different trajectories from which we can calculate the Figure of Merit f (eq. (2.17)) and the σ (eq. (2.14)). The numerical results are reported in the tables below. These values are also represented in the near figures, where on the horizontal axis there is N_t, on the left vertical axis the figure of merit f, and on the right vertical axis the entropy production rate σ. The horizontal dotted line is the true value of the entropy production rate σ^{th}. The vertical line has a N_t-value which correspond to the minumum of f.

N_t	f	σ
1	4.794	0.0173
20	1.728	0.0622
30	1.382	0.0752
40	1.186	0.0860
50	1.077	0.0955
60	1.022	0.1040
70	1.000	0.1116
75	0.996	0.1153
80	0.997	0.1187
85	1.000	0.1220
90	1.005	0.1252
100	1.021	0.1313
110	1.041	0.1372
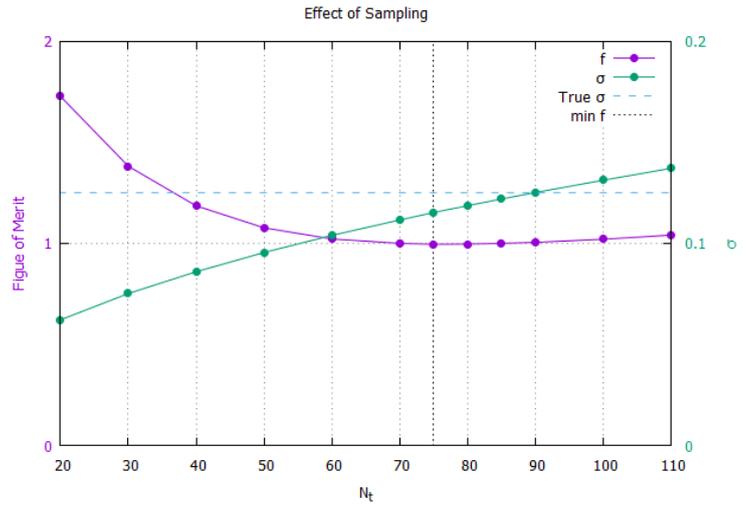
Figure 3.12: Temperatures $T_1 = 1$ and $T_2 = 2$. Intervals $\Delta t = 10^{-4}$ and $\Delta x = 0.2$.

N_t	f	σ
10	1.186	0.1001
15	1.020	0.1189
17	1.000	0.1254
19	0.993	0.1317
20	0.993	0.1346
25	1.016	0.1484
30	1.056	0.1606
40	1.152	0.1817
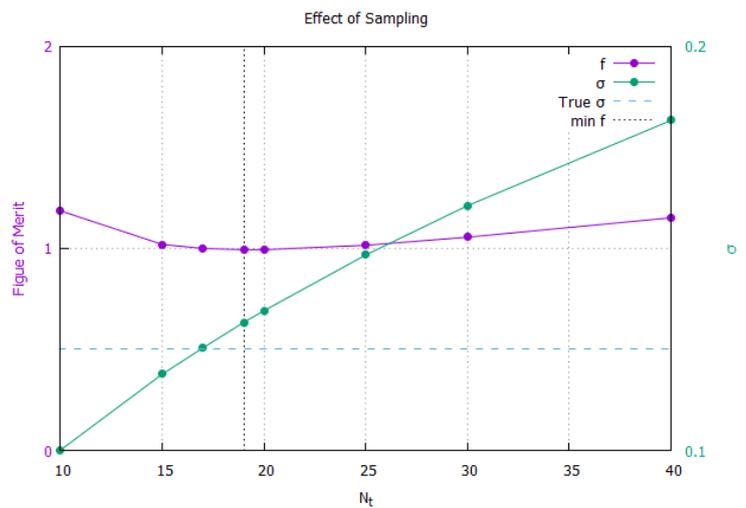
Figure 3.13: Temperatures $T_1 = 1$ and $T_2 = 2$. Intervals $\Delta t = 10^{-4}$ and $\Delta x = 0.1$.

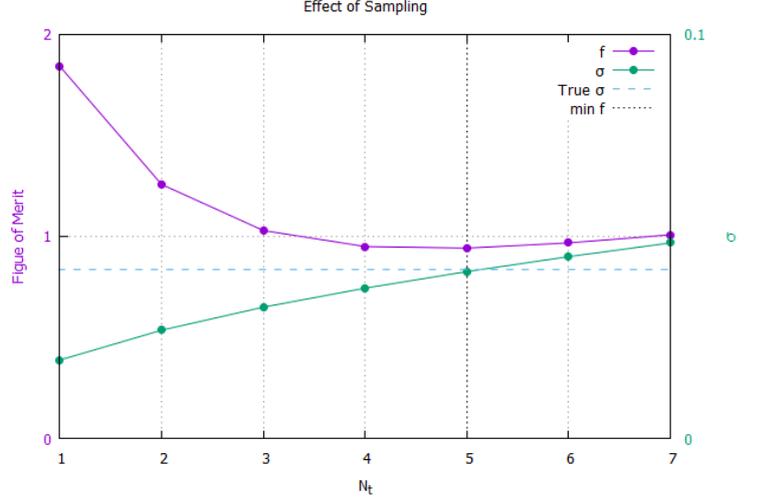| N_t	f	σ
1	1.842	0.0194
2	1.257	0.0268
3	1.029	0.0325
4	0.950	0.0372
5	0.942	0.0413
6	0.968	0.0450
7	1.008	0.0484

Figure 3.14: Temperatures $T_1 = 2$ and $T_2 = 3$. Intervals $\Delta t = 10^{-3}$ and $\Delta x = 0.2$.

From the examples shown it is possible to notice that generally the σ computed from the jumps' trajectory differ significantly from the correct value. In particular for small jumps σ is lower than σ^{th}, and increases monotonically increasing N_t. The figure of merit instead shows an absolute minimum at a particular N_t. This represents the best value we can use to minimize the memory effects of the dynamics. The important conclusion we can deduce from these examples is that using this N_t, the σ computed from the simulation is very close from σ^{th}. This is the proof that a good part of the discrepancy of the entropy production rate is due to non-Markovian effects, indeed by reducing these effects (by minimizing f) the value of σ approaches the correct value.

In the previous calculations we always used a small values for Δx (0.2 for fig. 3.12 and 3.14, 0.1 for fig. 3.13), such that the coarse-graining of the configurational space was not «drastic». We can ask ourselves what happens if we increase Δx: naturally we expect a greater discrepancy between the countinous trajectory and the discretized one. In particular, since a cell will gather more microstates and eliminate dynamics at increasingly larger scales, we expect that the entropy production rate will be smaller than the theoretical one. This is indeed the case, as it is possible to notice from the example reported in fig. 3.15, where using $\Delta x = 0.8$ the computed entropy was $\sigma = 0.0745$, compared to $\sigma^{th} = 0.125$. This argument shows the importance of the resolution of an instrument in the measurement of the non-equilibrium properties of a system.

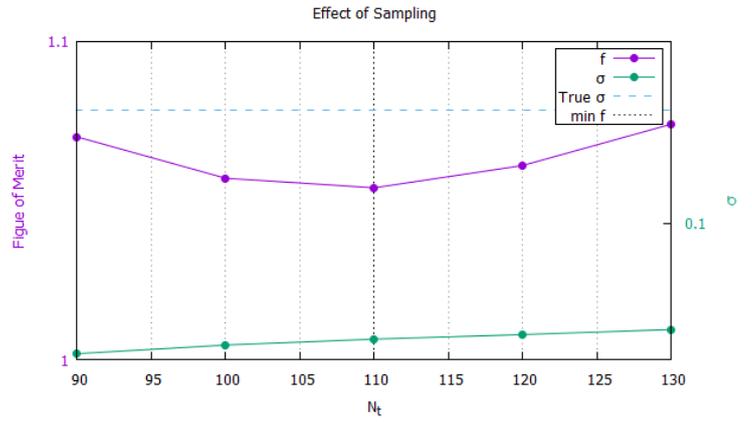| N_t	f	σ
90	1.070	0.0713
100	1.057	0.0732
110	1.054	0.0745
120	1.061	0.0755
130	1.074	0.0766

Figure 3.15: Temperatures $T_1 = 1$ and $T_2 = 2$. Intervals $\Delta t = 10^{-3}$ and $\Delta x = 0.8$.

Conclusions

In this thesis we have discussed the problem of establishing how different coarse-graining procedures affect the dynamics of a non-equilibrium system. In particular we worked with numerical simulations of models with few degrees of freedom in order to be able to compare simulations with analytical results. We presume however that the general results found in this thesis are also valid for more generic and realistic systems.

The first question we tried to answer was if it is possible to determine if a system is out of equilibrium by only "looking" at it, as we were trying to simulate a "non-invasive" experiment. It was easily proved that recording just one degree of freedom is not sufficient to draw conclusions, because even out of equilibrium the time series of random displacements were individually still indistinguishable from those found in equilibrium dynamics. In particular, displacements can maintain a Gaussian distribution even in a non-equilibrium situation. Instead, by looking at multiple degrees of freedom it is possible, in the majority of cases, to test for violations of detailed balance using the probability flux analysis, as it was done in ref. [13]. However, we found that in some cases, when there was a particular simmetry and we neglected some degree of freedom, this method infers equilibrium even when there is not. These results, if carried out in further research, could be useful in shedding light on the contingent effects of discarding degrees of freedom, which can be thought as a coarse-graining method present in most models and experiments.

The final result we obtained concerns a more quantitative characterization of non-equilibrium. We tried to quantify how much a system is out of equilibrium by means of the entropy production rate, a quantity easily calculated from the microscopic physical parameters. However, we were interested in a method to calculate this quantity by only monitoring the configurations of the system, like in a microscopy experiment, i.e. without knowing the parameters of the model. An expression was found only in the assumption of Markovian dynamics, which though conflicts with the discretization of the space in cells. A solution that solves this problem was found by adding an additional coarse-graining that samples the time series. At the end, it was possible to find an estimate of the entropy production rate comparable with the true value. This proves that our analysis can be used to find quantitative results regarding non-equilibrium in an experiment where the only accessible quantities are the trajectories of the degrees of freedom.

Bibliography

[1] Joel L. Lebowitz and Herbert Spohn. A Gallavotti-Cohen Type Symmetry in the Large Deviation Functional for Stochastic Dynamics. *Journal of Statistical Physics*, 95(1/2):333–365, 1999. arXiv: cond-mat/9811220.

[2] A Puglisi, S Pigolotti, L Rondoni, and A Vulpiani. Entropy production and coarse graining in Markov processes. *J. Stat. Mech.*, 2010(05):P05015, May 2010.

[3] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.*, 75(12):126001, December 2012.

[4] F. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz. Broken detailed balance and non-equilibrium dynamics in living systems. *Rep. Prog. Phys.*, 81(6):066601, June 2018. arXiv: 1710.03456.

[5] Ganhui Lan, Pablo Sartori, Silke Neumann, Victor Sourjik, and Yuhai Tu. The energy–speed–accuracy trade-off in sensory adaptation. *Nature Phys*, 8(5):422–428, May 2012.

[6] P. Mehta and D. J. Schwab. Energetic costs of cellular computation. *Proceedings of the National Academy of Sciences*, 109(44):17978–17982, October 2012.

[7] J. J. Hopfield. Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity. *Proceedings of the National Academy of Sciences*, 71(10):4135–4139, October 1974.

[8] M E Cates. Diffusive transport without detailed balance in motile bacteria: does microbiology need statistical physics? *Rep. Prog. Phys.*, 75(4):042601, April 2012.

[9] J. Gladrow, N. Fakhri, F. C. MacKintosh, C. F. Schmidt, and C. P. Broedersz. Broken Detailed Balance of Filament Dynamics in Active Networks. *Phys. Rev. Lett.*, 116(24):248301, June 2016.

[10] Fabian Knoch and Thomas Speck. Cycle representatives for the coarse-graining of systems driven into a non-equilibrium steady state. *New J. Phys.*, 17(11):115004, November 2015. arXiv: 1507.04247.

[11] Junang Li, Jordan M. Horowitz, Todd R. Gingrich, and Nikta Fakhri. Quantifying dissipation using fluctuating currents. *Nat Commun*, 10(1):1666, December 2019. arXiv: 1809.02118.

[12] Federica Mura, Grzegorz Gradziuk, and Chase P. Broedersz. Non-equilibrium scaling behaviour in driven soft biological assemblies. *Phys. Rev. Lett.*, 121(3):038002, July 2018. arXiv: 1803.02797.

[13] Christopher Battle, Chase P. Broedersz, Nikta Fakhri, Veikko F. Geyer, Jonathon Howard, Christoph F. Schmidt, and Fred C. MacKintosh. Broken detailed balance at mesoscopic scales in active biological systems. *Science*, 352(6285):604–607, April 2016.

[14] Massimiliano Esposito. Stochastic thermodynamics under coarse graining. *Phys. Rev. E*, 85(4):041125, April 2012.

[15] Hyun-Myung Chun and Jae Dong Noh. Hidden entropy production by fast variables. *Phys. Rev. E*, 91(5):052128, May 2015.

[16] L. D. Landau, E. M. Lifshit?s?, and L. D. Landau. *Fluid mechanics*. Number v. 6 in Course of theoretical physics. Pergamon Press, Oxford, England ; New York, 2nd ed., 2nd english ed., rev edition, 1987.

[17] L. D Landau, E. M Lifshit?ıs, and L. P Pitaevskiĭ. *Statistical physics. Part 1 Part 1.* 2009. OCLC: 1107194189.

[18] Chris H. Wiggins, D. Riveline, A. Ott, and Raymond E. Goldstein. Trapping and Wiggling: Elastohydrodynamics of Driven Microfilaments. *Biophysical Journal*, 74(2):1043–1060, February 1998.

[19] Hans Christian Öttinger. *Stochastic processes in polymeric fluids: tools and examples for developing simulation algorithms.* Springer, Berlin ; New York, 1996.

[20] Ken Sekimoto. Langevin Equation and Thermodynamics. *Prog. Theor. Phys. Suppl.*, 130:17–27, 1998.

[21] Udo Seifert. Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem. *Phys. Rev. Lett.*, 95(4):040602, July 2005.

[22] G. Nicolis and C. Nicolis. Chaotic dynamics, Markovian coarse-graining and information. *Physica A: Statistical Mechanics and its Applications*, 163(1):215–231, February 1990.

[23] S Krumscheid, M Pradas, G A Pavliotis, and S Kalliadasis. Data-driven coarse graining in action: Modelling and prediction of complex systems. page 9.

[24] Ludwig Boltzmann. Further Studies on the Thermal Equilibrium of Gas Molecules. In *History of Modern Physical Sciences*, volume 1, pages 262–349. PUBLISHED BY IMPERIAL COLLEGE PRESS AND DISTRIBUTED BY WORLD SCIENTIFIC PUBLISHING CO., July 2003.

[25] Saar Rahav and Christopher Jarzynski. Fluctuation relations and coarse-graining. *J. Stat. Mech.*, 2007(09):P09012–P09012, September 2007. arXiv: 0708.2437.

[26] D. A. Egolf. Equilibrium Regained: From Nonequilibrium Chaos to Statistical Mechanics. *Science*, 287(5450):101–104, January 2000.

[27] Maurizio Loreti. *Teoria degli Errori e Fondamenti di Statistica*. 2006.