Production and fluctuation of environmental entropy

Master thesis in physics

by

David Luposchainsky

under supervision of

Prof. Dr. Haye Hinrichsen

at the

University of Würzburg, 2012-13

If you see this, I forgot printing the title page separately.

Contents

1 Introduction

1.1 Overview in German

Die statistische Physik ist ein Teilgebiet der Physik, in welchem mithilfe von wahrscheinlchkeitstheoretischen Methoden Problemstellungen mit typischerweise einer sehr großen Anzahl von Teilnehmern betrachtet werden. Aufgrund der Allgemeinheit dieser Prinzipien sind die Ergebnisse jedoch nicht auf die Physik beschränkt, in welcher die "Teilnehmer" typischerweise einzelne Atome sind: in der Soziologie verhalten sich Meinungen ähnlich wie Spins in einem atomaren Gitter; Zellen enthalten potentiell sehr lange Proteine, die ausgefaltet weit zu groß für sie wären, und der Grund dafür ist der identisch mit dem, warum die Luft in einem Zimmer sich nicht spontan ausschließlich in der rechten Hälfte konzentriert; Flüssigkeiten und Verkehrsmuster ähneln sich in vielen Aspekten. All diese Beispiele haben eine Sache gemein: die grundlegenden Prinzipien dahinter sind sehr einfach, und aus diesem Grund an vielen verschiedenen Stellen in der Natur zu finden.

Die Größe mit der sich diese Arbeit beschäftigt ist die Entropie (welche für das vorige Proteinbeispiel verantwortlich ist). Das Konzept geht zurück auf das 19. Jahrhundert, wo sie eine wichtige Rolle in der klassischen Thermodynamik spielt; wahrscheinlich das bekannteste Auftreten der Entropie heute ist als Bestandteil des "Zweiten Hauptsatzes (der Thermodynamik)", welcher das Phänomen beschreibt, dass in der Natur Dinge viel eher zerstört als kreiert werden, sofern man nicht ein gewisses Maß an Energie aufbringt. In diesem Sinne ist (die Änderung der) Entropie ein Maß für die reversibilität eines Prozesses: mischt man zwei Flüssigkeiten und verrührt sie, wird auch beliebiges "verunrühren" sie nicht wieder trennen.

Es gibt einen weiteren Zugang zur Entropie, eng verwandt mit Informatik und Mathematik. Hier steht die sogenannte Shannon-Entropie für ein Maß des Informationsgehalts eines Systems, oder besser: wie viel unbekannt ist. Man stelle sich einen Sack voll Scrabble-Spielsteinen vor – die Wahrscheinlichkeit zufällig ein paar Buchstaben herauszunehmen, die zusammen ein Wort bilden, ist relativ hoch, und ebenso die Entropie im Sack. Ersetzt man alle Steine nun durch solche die

vornehmlich X, U und F enthalten, ist die Wahrscheinlichkeit ein sinnvolles Wort zu bilden wesentlich geringer, und so auch für die Entropie.

Eine wichtige Erkennis war nun, dass die thermodynamische und die informationstheoretische Entropie die gleiche Größe, nur von verschiedenen Richtungen betrachtet, sind: die Thermodynamik befasst sich mit dem System als Ganzem, die Informatik mit individuellen Teilen auf mikroskopischem Niveau. Die statistische Physik, und damit diese Arbeit, ist genau der Schnittpunkt dieser beiden Disziplinen.

Die Arbeit besteht aus zwei Hälften. In der ersten (Kapitel 2) geht es um ein neues sog. Fluktuationstheorem, welches eine wichtige Eigenschaft zeitlich fluktuierender Größen beschreibt; in diesem Fall handelt es sich um die Wahrscheinlichkeit, dass sich eine bestimmte Form von Entropie erhöht oder um den gleichen Betrag verringert, für die bisher kein solcher Zusammenhang bekannt war. Weitergehende Forschung führt zur Entdeckung einer ganzen neuen Klasse von Fluktuationstheoremen, und eine mögliche Anwendung dieser ist die Untersuchung der Reaktion von Systemen auf plötzliche Temperaturänderungen.

Die zweite Hälfte, Kapitel 3, versucht die vorige Form von Entropie auf ein kontinuierliches System zu übertragen. Dies stellt sich als überraschend schwer heraus, da die naive Übertragung der aus dem diskreten Fall bekannten Gesetzen zu inkorrekten Ergebnissen führt. Es ist dennoch möglich eine plausible Lösung zu finden, die sich zudem recht nah an bekannten physikalischen Intuitionen orientiert, und die Entwicklung dieser Formel ist das Hauptresultat dieses Abschnittes.

1.2 Statistical physics and entropy

Statistical physics is a branch of physics that uses probability theory in order to approximate and solve problems typically involving large amounts of actors. Due to the generality of these principles, the methods and results are far from being confined to the domain of physics, where the “actors” are typically things like single atoms, often revealing surprising similarities between seemingly unrelated fields. To give a few examples: Opinions spread in a fashion that resembles spins in an atomic lattice. Cells contain proteins that, unfolded to full length, would not fit at all and make them explode; a key reason why this never happens also explains why the all the air in the room you’re sitting in does not suddenly move in the bottom right corner simultaneously. Fluids and traffic patterns are very much alike in many respects. All these examples have one thing in common: the basic principles are very simple, and for that reason the emergent phenomena can be found in many places in nature.

The quantity this thesis revolves around is entropy (which accounts for the previous protein example). Its concept dates back all the way to the 19th century,

where it became one of the key quantities in classical thermodynamics; probably the most famous mentioning of entropy nowadays is the then conceived "second law (of thermodynamics)", which in various formulations expresses the phenomenon that things in nature prefer to break rather than be created, unless a considerable amount of energy is invested. In this sense, (the change of) entropy is a measure of reversibility of a process. Mix two fluilds and stir, then no amount of "unstirring" will result in them separating again.

There is another way of approaching entropy though, closely related to computer science and mathematics. Here the so-called Shannon entropy measures how much information is contained in a system, or better: how much about it is unknown. Take a bag of Scrabble pieces: the chance to draw a couple of letters that allow making a word is pretty good, the entropy in the bag is high. Now take a bag of equal size that contains mostly the letters X, U and F. Drawing from this bag is not going to yield any surprise combinations of useful letters; the system is known to be rigged towards three letters, the entropy is low (and incidentially also the chance to create a useful word).

What scientsts then found out is that both the thermodynamic and the information theoretical entropies are the same thing, just viewed from different angles: thermodynamics deals with systems as a whole, information theory with individual parts on a microscopic level. The intersection of these two is precisely where statistical physics, and therefore this thesis, is located.

1.3 Thesis overview

This thesis is divided in two parts, the first one being about environmental entropy in a discrete, the second one in a continuous setting. Although both parts are independent research subjects, they do share certain ideas and definitions. For this reason some issues in chapter 3 are assumed to be known due to being introduced in chapter 2. Both of these chapters will provide their own overviews at the beginning.

Chapter 2 is the result of playing around in a computer simulation, leading to a surprising outcome: when the resulting histogram was weighted using a seemingly arbitrary factor, it turns out to obey a certain symmetry called detailed fluctuation theorem. A detailed fluctuation theorem is an equation of the form

$$p(x) = e^x p(-x) \tag{3.1}$$

that is fulfilled by a distribution p for all values x.

This distribution will be the likelihood to obtain a certain value of one special form of entropy, for which no fluctuation theorems were previously known. Deeper investigation of the phenomenon lead to the discovery of an entire new

class of fluctuation theorems, and as a possible application a model for how sudden temperature changes affect the energy flow from/to the environment of a simple system.

The second topic is discussed in chapter 3, which is concerned with how entropy can be defined in a meaningful way in a continuous setting. This is surprisingly difficult, and naïvely carrying over the definitions from the discrete setting produces incorrect results. To back this up with a simple example, consider the before mentioned definition of (Shannon) entropy for a discrete distribution,

$$S_{\text{discrete}} = -\sum_i p_i \ln p_i \tag{3.2}$$

and make it continuous by scaling down the intervals over which p changes its value. In the limit this will result in an integral of the form

$$S_{\text{cont.}} = -\int_\Omega \mathrm{d}x\, p(x) \ln p(x) \ . \tag{3.3}$$

But there is a problem with this equation: it does not have a lower bound. Intuitively, the δ distribution should have no entropy, as it is the continuous version of a single-valued state. However, inserting it into the equation above results in $S = -\infty$. Take $\Omega = \mathbb{R}$ and a Gaussian with zero mean and standard deviation σ, insert everything in the above equation, then the result will be

$$p_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{3.4}$$

$$\Rightarrow S_{\text{cont.}} = \ln(\sigma) + \text{const.} \tag{3.5}$$

which is logarithmically divergent for $\sigma \to 0$, the same limit that yields the δ distribution for a Gaussian. Another way of looking at it is that a δ centered at a real value μ carries infinite information (hence infinite entropy), namely all the decimals of μ.

This example hints that entropy is not straightforwardly applied to continuous settings. The actual work in chapter 3 will be concerned with how the concept of environmental entropy – in many cases easier to picture as the heat exchange with the environment – can be brought from its easy discrete definition into a setting where the degrees of freedom are continuous.

2 Strong fluctuation theorem for environmental entropy

The contents of this chapter were published in Physical Review E, "Strong fluctuation theorem for nonstationary nonequilibrium systems" [1]. The following is an elaborate description of its contents.

This chapter has two main parts. The first one provides an introduction to fluctuation theorems, Markov processes, entropy, and the terminology used. The secone one then uses these definitions to prove a new general fluctuation theorem, specializes it to appear in a more familiar form, and finally puts it to the test in two numerical examples.

2.1 Preliminaries

2.1.1 Fluctuation theorems

Classical thermodynamics is a theory of averages of certain quantities. For example, temperature is a measure for the *mean* energy per degree of freedom of a gas, or the *mean* cost of energy to increase the *mean* entropy by a certain amount. Unfortunately, this approach requires certain quantities to exist a priori, and does not explain their origins. On the other hand, statistical physics typically starts with all individual constituents at the same time, and tries to explain how certain statistical properties – the mean being just one of them – emerge. In that sense, fluctuation theorems are about **taking away the averages** of classical thermodynamics. One of these averages is the entropy of a system, of which we know that it obeys the second law of thermodynamics, typically written as

$$\mathrm{d}S \geq 0 \,, \tag{1.1}$$

meaning that thermodynamic (macroscopic) entropy is always increasing. In colloquial terms, this describes the phenomenon that systems in nature are automatically driven towards greater uncertainty (a sand castle is a highly ordered

arrangement of particles, and the wind is much more likely of rearranging these particles in a way that makes their position uncertain than to build up a castle by random chance).

However, looking at the microscopic properties of a system, it is indeed possible that entropy spontaneously decreases – when all microstates occur with some nonzero probability, there is also always a nonzero (albeit small) chance of transition to a highly ordered system. Fluctuation theorems provide a characteristic expression for the likelihood for certain statistical events to occur in a system, and have been extensively studied in the past [2–18]. For example in the case of microscopic total entropy changes ΔS, it can be shown that

$$P(\Delta S) = e^{\Delta S} P(-\Delta S) \tag{1.2}$$

which states that the probability of spontaneous entropy increase by ΔS is exponentially more likely than a decrease by the same amount; theorems of the above form will be referred to as **detailed** fluctuation theorems, as the expression yields symmetry information for each point of a distribution individually. In a sense, this is the analogon to the Second Law from the viewpoint of statistical physics, and indeed it can be shown that (1.1) can be recovered from this formula: integrating both sides yields

$$1 = \int_{-\infty}^{\infty} \mathrm{d}(\Delta S) P(\Delta S) = \int_{-\infty}^{\infty} \mathrm{d}(\Delta S) e^{\Delta S} P(-\Delta S) = \left\langle e^{-\Delta S} \right\rangle \tag{1.3}$$

which, using Jensen's inequality $\langle e^x \rangle \geq e^{\langle x \rangle}$, implies

$$\langle \Delta S \rangle \geq 0 \ . \tag{1.4}$$

Averages of this form are known as **integral** fluctuation theorems.

It turns out that this derivation was overly general: in fact, a detailed fluctuation theorem implies arbitrarily many integral fluctuation theorems, since

$$\begin{aligned} B \equiv \langle e^{-\frac{x}{2}} A(x) \rangle &= \int_{-\infty}^{\infty} \mathrm{d}x \, P(x) e^{-\frac{x}{2}} A(x) = \int_{\infty}^{-\infty} \mathrm{d}(-x) P(-x) e^{\frac{x}{2}} A(-x) \\ &= -\int_{-\infty}^{\infty} \mathrm{d}x \, P(x) e^{-\frac{x}{2}} A(x) = -B = 0 \\ &\Rightarrow \langle e^{-\frac{x}{2}} A(x) \rangle = 0 \end{aligned} \tag{1.5}$$

where A is an arbitrary antisymmetric function; the relation (1.3) can be recovered by setting $A(x) = \sinh(x/2) = \frac{1}{2}(e^x - e^{-x})$.

This is just one example of a fluctuation theorem for ΔS; there are many more for other quantities (such as different forms of heat, work, other sub-categories of entropy), as can be seen e.g. in [2], and one of the main results of this chapter will be yet another entire class of fluctuation theorems.

2.1.2 Markov jumping processes

A *jumping process* is a system consisting of a discrete number of states that, governed by certain transition rates or probabilities, spontaneously switches from one state to another repeatedly in a random fashion. A *Markov process* is a special jumping process where the "next" state can only depend on the current state, and has no knowledge about its history. In other words, a Markov process is one in which the system does not have a concept of "momentum" in the colloquial sense.

An example for a jumping process is a gambler in a casino, the state being how much money he has. Each time he places a bet, he changes his state (loses or wins money).

1. A realistic gambler can be seen as a general jumping process, and can make his bets dependent on e.g. whether he lost his last game (which requires to know about the current balance and the one before that), or if he is suspicious kind could alter the winning probabilities in his favour. Note that the concept of "winning streaks" also falls into this category, as "having" one requires knowledge of the previous games/balances.

2. A Markovian gambler would decide how much to bet based solely on how much money he currently has. He would base none of his decisions on his knowledge about previous games.

2.1.3 Terminology

Suppose we have a system Ω of N states, enumerated by j. Time runs from $t = 0$ to $t = T$. At times τ_j, the system transitions from an old state c_{j-1} into a new one, c_j. Grouping multiple successive state changes together, we can describe the history of the system using a trajectory through time/state space with J jumps,

$$\gamma = c_0 \xrightarrow[\tau_1]{} c_1 \xrightarrow[\tau_2]{} \cdots \xrightarrow[\tau_j]{} c_j \xrightarrow[\tau_{j+1}]{} c_{j+1} \xrightarrow[\tau_{j+2}]{} \cdots \xrightarrow[\tau_J]{} c_J \tag{1.6}$$

$$= \{c_{i,\tau_i}\} \tag{1.7}$$

An "element" c_{i,τ_i} of γ can thus be interpreted as "the system is in state c_i starting at time τ_i, until τ_{i+1}". A visualization of the above description can be found in 2.1.

In a typical system, many such paths can be taken when starting from an initial configuration c_0; depending on the dynamics of the system, the jumps can

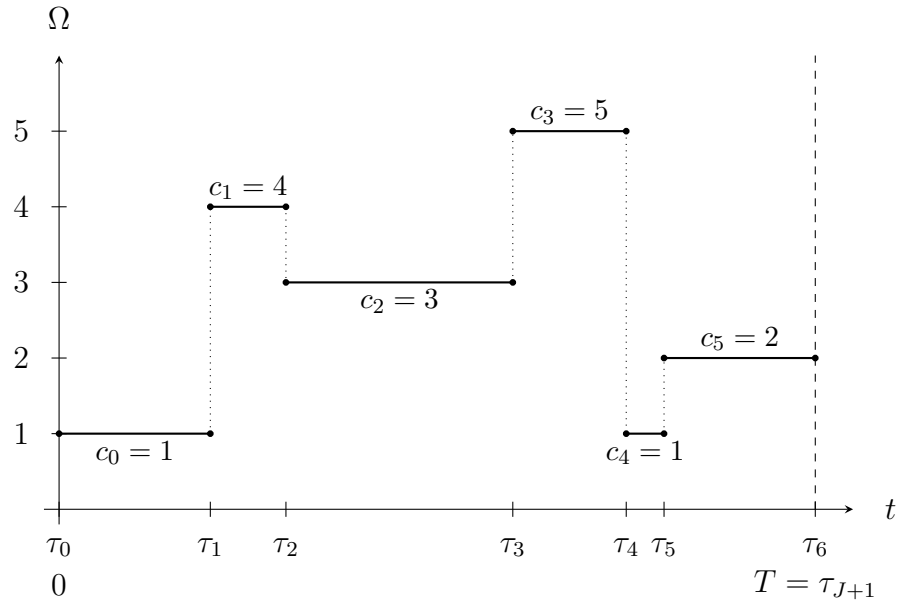

Figure 2.1: Visualization of a discrete jumping process with $J = 5$ jumps. The system, initially in state c_0 at time $t = \tau_0 = 0$, transitions into a new state c_1 at time τ_1, and remains in this state until the next jump occurs. The process then repeats itself, and the resulting trajectory is a representation of the system's history.

occur at different times or to different states. Due to the stochastic nature of the system, trajectories most likely occur with different probabilities; these conditional probabilities will be denoted as $p[\{c_{i,\tau_i}\}|c_0]$, with an explicit mentioning of the starting point n_0.

A Markov process on a set of N states is characterized by a (possibly time-dependent) transition matrix $w_{c\to c'}$, whose entries are the rates at which the system jumps from one configuration c to another one c'. Since later calculations will require ratios of rates it is required that in case of a nonzero jumping rate, the reverse jump also has a nonzero entry. Letting this system run for a certain amount of time will result in such a stochastic trajectory γ through the state space, and these trajectories are what this chapter is about.

Given such a process, the probability of a path $p[\{c_{i,\tau_i}\}|c_0]$, containing a total of J jumps (and consequently covering $J + 1$ states), can be expressed as a series

of "staying in the same state" and "change state" contributions [2],

$$\begin{aligned} p[\{c_{i,\tau_i}\}|c_0] = \exp\left(-\int_{\tau_0}^{\tau_1} \mathrm{d}\tau\ r_{c_0}(\tau)\right) \\ \times \prod_{j=1}^{J} w_{c_{j-1}\to c_j}(\tau_j) \exp\left(-\int_{\tau_j}^{\tau_{j+1}} \mathrm{d}\tau\ r_{c_j}(\tau)\right) \end{aligned} \tag{1.8}$$

where r_c stands for the escape rate from state c,

$$r_c = \sum_{c' \neq c} w_{c\to c'}\ . \tag{1.9}$$

Here the first line, $\exp\left(-\int_{\tau_0}^{\tau_1} \mathrm{d}\tau\ r_{c_0}(\tau)\right)$, corresponds to the probability for staying in configuration c_0 from time τ_0 to τ_1. This is then followed by a transition to c_1 given by the first factor of the product, namely $w_{\ldots}$, followed by staying in this configuration for times τ_1 to τ_2. The process – jumping, then staying – then repeats itself another $J-1$ times, until all jumps have happened and the trajectory has been run through. The product of the probabilities of the constituents then makes up the overall path probability.

Assuming time-independent rates w, the integrals in (1.8) can trivially be evaluated, and the expression reduces to

$$p[\{c_{i,\tau_i}\}|c_0]\Big|_{w\ \text{const.}} \equiv Q[\gamma] = e^{-(\tau_1-\tau_0)r_{c_0}} \left(\prod_{j=1}^{J} w_{c_{j-1}\to c_j} e^{-(\tau_{j+1}-\tau_j)r_j}\right) \tag{1.10}$$

It will also be necessary to calculate the probability of taking the steps of a path in reverse direction, denoted "†" and illustrated in 2.2, and the above definitions turn out as follows in this case:

$$\begin{aligned} \gamma^\dagger = {} & c_J \underset{T-\tau_J}{\longrightarrow} \cdots \underset{T-\tau_{j+2}}{\longrightarrow} c_{j+1} \underset{T-\tau_{j+1}}{\longrightarrow} c_j \underset{T-\tau_j}{\longrightarrow} \cdots \underset{T-\tau_2}{\longrightarrow} c_1 \underset{T-\tau_1}{\longrightarrow} c_0 \\ = {} & \{c_{i,\tau_i}\}^\dagger \end{aligned}$$

$$\begin{aligned} p^\dagger[\{c_{i,\tau_i}\}^\dagger|c_0^\dagger] &= \exp\left(-\int_{\tau_{J+1}}^{\tau_J} \mathrm{d}(-\tau) r_{c_J}(\tau)\right) \\ &\quad \times \prod_{j=J}^{1} w_{c_j\to c_{j-1}}(\tau_j) \exp\left(-\int_{\tau_j}^{\tau_{j-1}} \mathrm{d}(-\tau) r_{c_{j-1}}(\tau)\right) \\ &= \exp\left(-\int_{\tau_J}^{\tau_{J+1}} \mathrm{d}\tau\ r_{c_J}(\tau)\right) \\ &\quad \times \prod_{j=1}^{J} w_{c_j\to c_{j-1}}(\tau_j) \exp\left(-\int_{\tau_{j-1}}^{\tau_j} \mathrm{d}\tau\ r_{c_{j-1}}(\tau)\right) \end{aligned} \tag{1.11}$$

$$p^\dagger[\{c_{i,\tau_i}\}^\dagger|c_0^\dagger]\Big|_{w\text{ const.}} \equiv Q[\gamma^\dagger] = e^{-(\tau_{J+1}-\tau_J)r_{c_J}}\left(\prod_{j=1}^{J} w_{c_j\to c_{j-1}} e^{-(\tau_j-\tau_{j-1})r_{j-1}}\right) \quad (1.12)$$

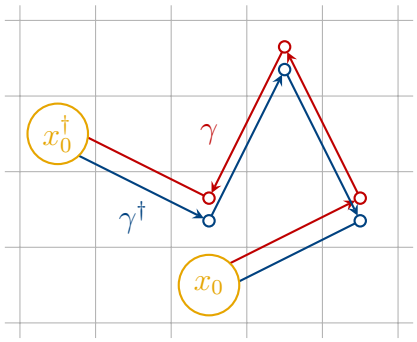

Figure 2.2: Path reversal for randomly moving knight on an infinite chess board. The forward path γ is taken by the knight with probability $p[\gamma|x_0]$. The conjugate path $\gamma^\dagger$ is that same knight making the same jumps, but in opposite order and direction, which happens with probability $p^\dagger[\gamma^\dagger|x_0^\dagger]$.

2.1.4 Entropy in a Markov process

For Markov processes, it is often suitable not to look at the total entropy S_{tot}, but to split it up in two constituents, namely the system entropy S_{sys} and the environmental entropy S_{env}.

The **system entropy** $\mathbf{S}_{\text{sys}}$ is the Shannon entropy of an ensemble. Running a Markov process, starting with some random initial configuration c_0 drawn using some distribution $p_0(t=0)$, then after some time Δt the system will be in a new configuration c_i. Repeating the procedure many times will yield a distribution $p_i(t)$ of the various c_i reached. S_{sys} is now the Shannon entropy associated to that distribution:

$$S_{\text{sys}}(t) = -\sum_i p_i(t)\ln p_i(t) \quad (1.13)$$

The standard interpretation of this relationship is that S_{sys} is a quantity describing how concentrated the system is around certain states at a single point in time, or in other words how much more likely it is to be in some specific state (or set of states) compared to others. For example, if the transition matrix $w_{c\to c'}$ has a bias

to make the system go into a particular subset of configurations, S_{sys}will be lower, while maximum entropy is achieved when all states are reached with the same probability.

On the other hand, the **environmental entropy $\mathbf{S}_{\text{env}}$** takes not the system's internal configuration into account, but characterizes how the current configuration has been reached. Each time the system changes its state, S_{env} increases by a certain amount,

$$\Delta S_{\text{env}}^{c\to c'} = \ln \frac{w_{c\to c'}}{w_{c'\to c}} \tag{1.14}$$

which is known as the Schnakenberg formula, after its discoverer [19]. This provides a characteristic quantity for the reversibility of a process: ΔS_{env} is zero if the jumps and its reverse are equally likely, larger than zero if a transition was made in a more likely state, and lastly smaller than zero if an unlikely jump happens.

A stochastic trajectory γ now is a series of many such jumps, hence the total of the accumulated environmental entropy can be seen as a functional on γ,

$$\Delta S_{\text{env}}[\gamma] = \sum_{i=1}^{J} \ln \frac{w_{c_{i-1}\to c_i}}{w_{c_i\to c_{i-1}}} \tag{1.15}$$

where J is the total number of jumps in γ, and c_i are the configurations it consists of.

2.1.5 Master equation

One way of exploring the properties of a Markov process is to collect statistics via running the same experiment many times. Another approach is starting with the statistics in the first place – instead of beginning with individual processes and measuring their distribution, take the distribution of states and ask how it will evolve. This is what the *Master equation* describes:

$$\dot{p}_c(t) = \underbrace{\sum_{c'} j_{c'\to c}}_{\text{gain}} - \underbrace{\sum_{c'} j_{c\to c'}}_{\text{loss}} \tag{1.16}$$

$$= \sum_{c'} p_{c'}(t) w_{c'\to c} - \sum_{c'} p_c(t) w_{c\to c'} \tag{1.17}$$

The first term describes the system jumping into a state (probability current j towards the state c), while the second one accounts for transitions out of it (current away from c). A stationary Master equation ($\dot{p} \equiv 0$) is characteristic for systems for which macroscopic quantities are constant, as described next.

2.1.6 Detailed balance

Systems obeying statistical dynamics can be classified in multiple categories. On the one end of the spectrum there is total nonequilibrium, where all quantities fluctuate as they please. The opposite scenario is equilibrium, in which all macroscopic properties of the system are constant.

Detailed balance takes an intermediate position between these two: although on a microscopic level the system may fluctuate and change its state, all these fluctuations cancel each other out in their contribution to the macroscopic observables. This leads to a setting where the internal dynamics appear unequilibriated, but viewed from the outside it appears as if the system is in equilibrium.

First definition

Mathematically, the cancellation of the contributions to macroscopic quantities mentioned above translates to the probability current from one state to another being identical to the reverse one,

$$\begin{aligned} j_{c\to c'} &= j_{c'\to c} \\ \Leftrightarrow\ p_c w_{c\to c'} &= p_{c'} w_{c'\to c}\ . \end{aligned} \tag{1.18}$$

This formulation leads directly to a stationary Master equation, which explains how the macroscopic properties of the system are constant. However, it falsely suggests that detailed balance depends on the probability distribution p_c of the system at a point in time, when in fact this is not so, as can be seen in an alternative formulation.

Second definition

A system is in detailed balance iff the product of rates along a closed trajectory $\gamma = \{c_1, \ldots, c_n, c_1\}$ is the same as for the reverse trajectory,

$$w_{c_1\to c_2}\ \cdots\ w_{c_{n-1}\to c_n} w_{c_n\to c_1} = w_{c_1\to c_n} w_{c_n\to c_{n-1}}\ \cdots\ w_{c_2\to c_1}\ . \tag{1.19}$$

This picture, also illustrated in figure 2.3, makes it easier to understand that systems in detailed balance are reversible and therefore produce no entropy in the environment, $\Delta S_{\text{env}}(t) = 0$.

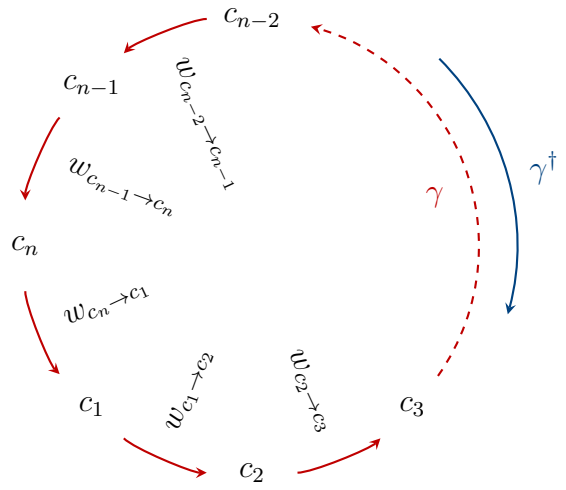

Figure 2.3: Illustration of the second definition of detailed balance. A system obeys detailed balance iff for all closed trajectories γ the product of the rates along them is the same as the product of rates along the corresponding reverse direction $\gamma^\dagger$. In the illustration, $\gamma^\dagger$ would mean that all arrows in the circle and all indices of w are flipped. This definition makes it obvious that detailed balance is a property of the rates of a Markov process only, and does not depend on the probability distribution these rates produce, as (1.18) might suggest.

2.2 Fluctuation theorem for environmental entropy

From now on all rates will be assumed time-independent,

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{c\to c'} = 0 \ , \tag{2.20}$$

for all configurations $c, c' \in \Omega$. This simplifies the equations significantly, at the cost of limiting the system's drive to be constant over time. One important scenario where this is the case is in the study of relaxation processes, where the rates change just before the experiment, within which they will be unchanging; this particular case is described in detail in section 2.2.3.

2.2.1 Preliminary definitions

The main result of this section is the derivation of a new fluctuation theorem for ΔS_{env}. The approach taken will be to first derive an entirely new class of

fluctuation theorems, which will contain the desired result as a special case.

In a given Markov process, each allowed path γ appears with a certain probability $W[\gamma]$. Paired with another functional $F[\gamma]$ along the path, this allows us to express statistical properties of F by integrating over all paths; in particular, the average of F can be written as

$$\langle F \rangle = \int \mathcal{D}\gamma \, W[\gamma] F[\gamma] \ . \tag{2.21}$$

The path integral "$\int \mathcal{D}\gamma$" is shorthand notation for summation over all paths $\gamma = \{c_0, c_1, \ldots, c_n\}$ with N possible configurations per step,

$$\int \mathcal{D}\gamma \, F[\gamma] \equiv \sum_{c_0=1}^{N} \sum_{c_1=1}^{N} \cdots \sum_{c_n=1}^{N} F(c_0, c_1, \ldots c_n) \ . \tag{2.22}$$

Using this formalism, the probability distribution of increasing environmental entropy can, with the help of the δ distribution, be written as a functional integral

$$\begin{aligned} P(\Delta S_{\mathrm{env}} = X) &= \langle \delta(X - \Delta S_{\mathrm{env}}[\gamma]) \rangle \\ &= \int \mathcal{D}\gamma \, W[\gamma] \delta(X - \Delta S_{\mathrm{env}}[\gamma]) \ . \end{aligned} \tag{2.23}$$

For the purpose of the new theorem, the average (2.21) is modified by introducing a weight functional $\chi_{c_0,c_T} > 0$ that only depends on the very first and last configuration along the stochastic trajectory, and satisfies the symmetry relation

$$\chi_{c_0,c_T} p^{\mathrm{init}}_{c_0} = \chi_{c_T,c_0} p^{\mathrm{init}}_{c_T} \tag{2.24}$$

where p^{init}_c is the probability that the whole process is started (initialized) with configuration c. Using this, a weighted average can be defined,

$$\langle F \rangle_\chi \equiv \langle F \, \chi \rangle = \frac{1}{\mathcal{N}} \int \mathcal{D}\gamma \, W[\gamma] F[\gamma] \chi_{c_0,c_T} \tag{2.25}$$

where $\mathcal{N}$ is a normalization constant to take the addition of χ into account,

$$\mathcal{N} = \int \mathcal{D}\gamma \, W[\gamma] \chi_{c_0,c_T} \ . \tag{2.26}$$

Applied to the environmental entropy production ΔS_{env}, this weighted average reads

$$\begin{aligned} \tilde{P}(\Delta S_{\mathrm{env}} = X) &= \langle \delta(X - \Delta S_{\mathrm{env}}[\gamma]) \rangle_\chi \\ &= \frac{1}{\mathcal{N}} \int \mathcal{D}\gamma \, W[\gamma] \delta(X - \Delta S_{\mathrm{env}}[\gamma]) \chi_{c_0,c_T} \end{aligned} \tag{2.27}$$

This allows the main result of this chapter to be stated, namely that the quantity $\tilde{P}$ satisfies a detailed fluctuation theorem,

$$\boxed{\tilde{P}(\Delta S_{\text{env}} = X) = e^X \tilde{P}(\Delta S_{\text{env}} = -X)}\,, \tag{2.28}$$

a relationship that will be proven in the following sections. Note that $\tilde{P}$ is still normalized (and therefore a probability distribution, since $\chi > 0$); explicitly:

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{d}X \tilde{P}(\Delta S_{\text{env}} = X) &= \int_{-\infty}^{\infty} \mathrm{d}X \frac{1}{\mathcal{N}} \int \mathcal{D}\gamma\, W[\gamma]\delta(X - \Delta S_{\text{env}}[\gamma])\chi_{c_0,c_T} \\ &= \frac{1}{\mathcal{N}} \underbrace{\int \mathcal{D}\gamma\, W[\gamma]\chi_{c_0,c_T}}_{=\mathcal{N}\ (2.26)} \underbrace{\int_{-\infty}^{\infty} \mathrm{d}X \delta(X - \Delta S_{\text{env}}[\gamma])}_{=1} \\ &= 1\end{aligned} \tag{2.29}$$

In particular, this allows reducing (2.28) analogous to (1.5) to the integral fluctuation theorem

$$\left\langle e^{-\frac{1}{2}\Delta S_{\text{env}}} A(\Delta S_{\text{env}}) \right\rangle_\chi = 0 \tag{2.30}$$

where again A is an arbitrary antisymmetric function.

2.2.2 Proof of a new class of detailed fluctuation theorems

Path reversal and environmental entropy

While (2.27) provides a statistical summary, an expression relating the actual value of ΔS_{env} to path probabilities will be useful in the proof. (1.15) is such a functional, but is rate- and not path-based; therefore, the following postulate is useful, and will be shown to be equivalent to (1.15). Note that, although not necessary for this thesis, the statement holds even when the rates are time-dependent.

$$\Delta S_{\text{env}} = \ln \frac{p[\{c_{i,\tau_i}\}|c_0]}{p^\dagger[\{c_{i,\tau_i}\}^\dagger|c_0^\dagger]} \tag{2.31}$$

Therefore, for a path with a total of J jumps (hence containing $J+1$ intermediate states),

$$\begin{aligned}
&\ln \frac{p[\{c_{i,\tau_i}\}|c_0]}{p^\dagger[\{c_{i,\tau_i}\}^\dagger|c_0^\dagger]} \\
&= \ln \frac{\exp\left(-\int_{\tau_0}^{\tau_1} \mathrm{d}\tau\; r_{c_0}(\tau)\right) \prod_{j=1}^{J} w_{c_{j-1}\to c_j}(\tau_j) \exp\left(-\int_{\tau_j}^{\tau_{j+1}} \mathrm{d}\tau\; r_{c_j}(\tau)\right)}{\exp\left(-\int_{\tau_{J+1}}^{\tau_J} \mathrm{d}(-\tau) r_{c_J}(\tau)\right) \prod_{j=J}^{1} w_{c_j\to c_{j-1}}(\tau_j) \exp\left(-\int_{\tau_j}^{\tau_{j-1}} \mathrm{d}(-\tau) r_{c_{j-1}}(\tau)\right)} \\
&= \ln \frac{\exp\left(-\int_{\tau_0}^{\tau_1} \mathrm{d}\tau\; r_{c_0}(\tau)\right) \prod_{j=1}^{J} \exp\left(-\int_{\tau_j}^{\tau_{j+1}} \mathrm{d}\tau\; r_{c_j}(\tau)\right)}{\exp\left(-\int_{\tau_J}^{\tau_{J+1}} \mathrm{d}\tau\; r_{c_J}(\tau)\right) \prod_{j=1}^{J} \exp\left(-\int_{\tau_{j-1}}^{\tau_j} \mathrm{d}\tau\; r_{c_{j-1}}(\tau)\right)} + \ln \prod_{j=1}^{J} \frac{w_{c_{j-1}\to c_j}(\tau_j)}{w_{c_j\to c_{j-1}}(\tau_j)} \\
&= \ln \frac{\prod_{j=0}^{J} \exp\left(-\int_{\tau_j}^{\tau_{j+1}} \mathrm{d}\tau\; r_{c_j}(\tau)\right)}{\prod_{j=1}^{J+1} \exp\left(-\int_{\tau_{j-1}}^{\tau_j} \mathrm{d}\tau\; r_{c_{j-1}}(\tau)\right)} + \sum_{j=1}^{J} \ln \frac{w_{c_{j-1}\to c_j}(\tau_j)}{w_{c_j\to c_{j-1}}(\tau_j)} \\
&= \ln \frac{\prod_{j=0}^{J} \exp\left(-\int_{\tau_j}^{\tau_{j+1}} \mathrm{d}\tau\; r_{c_j}(\tau)\right)}{\prod_{j=0}^{J} \exp\left(-\int_{\tau_j}^{\tau_{j+1}} \mathrm{d}\tau\; r_{c_j}(\tau)\right)} + \Delta S_{\mathrm{env}} \\
&= \Delta S_{\mathrm{env}}
\end{aligned}$$

as claimed in (2.31). Specialized to time-independent rates, this then reads

$$\Delta S_{\mathrm{env}}[\gamma] = \ln \frac{Q[\gamma]}{Q[\gamma^\dagger]} \ . \tag{2.32}$$

Generalized statement and proof

Theorem 1 (Modified master fluctuation theorem)
Let $F[\gamma] = -F[\gamma^\dagger]$ be an antisymmetric functional on a path γ (generated by a process with time-independent rates), and g be an arbitrary function acting on it. Then

$$\boxed{\langle g(F[\gamma])\rangle_\chi = \langle e^{-\Delta S_{\mathrm{env}}[\gamma]} g(-F[\gamma])\rangle_\chi} \tag{2.33}$$

Proof. (This largely follows the same ideas as the one for the Master Fluctuation Theorem presented in [2].)

Let

$$W[\gamma] = p_{c_0}^{\mathrm{init}} Q[\gamma] \ . \tag{2.34}$$

(Recall again that p_c^{init} is the probability the experiment is started with initial configuration c.) The left hand side of (2.33) can, using (2.25), be written as

$$\langle g(F[\gamma])\rangle_\chi = \frac{1}{\mathcal{N}} \int \mathcal{D}\gamma\, p_{c_0}^{\mathrm{init}} Q[\gamma] g(F[\gamma]) \chi_{c_0,c_T} \ . \tag{2.35}$$

As the order of summation does not matter ($\mathcal{D}\gamma = \mathcal{D}\gamma^\dagger$), instead of integrating along the trajectory γ, the direction can be reversed by substituting $\gamma \to \gamma^\dagger$ and therefore also swapping c_0 and c_T, giving

$$\langle g(F[\gamma])\rangle_\chi = \frac{1}{\mathcal{N}} \int \mathcal{D}\gamma^\dagger \, p^{\text{init}}_{c_T} Q[\gamma^\dagger] g(F[\gamma^\dagger]) \chi_{c_T,c_0} \ . \tag{2.36}$$

This purely mathematical step is now for the most part undone using (2.32), the antisymmetry of F and the symmetry of χ (2.24),

$$\langle g(F[\gamma])\rangle_\chi = \frac{1}{\mathcal{N}} \int \mathcal{D}\gamma \, p^{\text{init}}_{c_T} e^{-\Delta S_{\text{env}}[\gamma]} Q[\gamma] g(-F[\gamma]) \chi_{c_0,c_T} \ . \tag{2.37}$$

Comparing this result with (2.35) allows the expression on the right hand side as an average as well, resulting the theorem's claim

$$\langle g(F[\gamma])\rangle_\chi = \langle e^{-\Delta S_{\text{env}}[\gamma]} g(-F[\gamma])\rangle_\chi \ . \tag{2.38}$$

Specialization to a theorem about environmental entropy

The theorem just proven is overly general to the point where it is not clear how to apply it. As promised earlier, (2.28) is a special case of theorem 1. First note that environmental entropy is an uneven functional on the trajectory,

$$\Delta S_{\text{env}}[\gamma] = \ln \frac{Q[\gamma]}{Q[\gamma^\dagger]} = -\ln \frac{Q[\gamma^\dagger]}{Q[\gamma]} = -\Delta S_{\text{env}}[\gamma^\dagger] \ . \tag{2.39}$$

Second, choose

$$g(\Delta S_{\text{env}}[\gamma]) = \delta(X - \Delta S_{\text{env}}[\gamma]) \ , \tag{2.40}$$

then (2.28) allows writing (2.33) as

$$\begin{aligned} &\langle \delta(X - \Delta S_{\text{env}}[\gamma])\rangle_\chi = \langle e^{-\Delta S_{\text{env}}[\gamma]} \delta(X + \Delta S_{\text{env}}[\gamma])\rangle_\chi \\ &\Leftrightarrow \tilde{P}(\Delta S_{\text{env}} = X) = e^X \tilde{P}(\Delta S_{\text{env}} = -X) \end{aligned}$$

as claimed.

2.2.3 Applications

One of the challenges of general theorems is fixing the degrees of freedom it provides so that a practically viable result emerges. The Master Fluctuation Theorem [2] does not have this issue as much, because it unifies lots of known fluctuation theorems; in the present case however, the general theorem is proven, but what

it unifies or even describes is not so clear. This section will provide an example of one emergent result: a fluctuation theorem for the energy set free after sudden cooling of an equilibrium system.

Take a classical system that is initially in equlibrium with a heat bath of constant temperature $T_1 = 1/(k_B\beta_1)$. Assuming each configuration $c \in \Omega$ is associated to a specific energy E_c, the distribution of states in this setting is Boltzmannian,

$$p_c^{\text{init}} = \frac{1}{Z(\beta_1)} e^{-\beta_1 E_c} \tag{2.41}$$

with the partition sum $Z(\beta) = \sum_c e^{-\beta E_c}$. Since the system is in equlibrium it also obeys detailed balance (recall that detailed balance is a weaker condition than equilibrium, see section 2.1.6), and consequently the rates can be expressed as follows:

$$p_c w_{c\to c'} = p_{c'} w_{c'\to c} \tag{2.42}$$

$$\Leftrightarrow \frac{w_{c\to c'}}{w_{c'\to c}} = \frac{p_{c'}}{p_c} = e^{-\beta_1(E_{c'}-E_c)} \tag{2.43}$$

At time $t = 0$, the temperature of the bath is now instantaneously changed ("quenched") to β_2, and the system converges to a new stationary equilibrium state. How much energy $\Delta E = E_{c_0} - E_{c_T}$ does this process take out or feed into the system in a given amount of time T? Since the reservoir is not part of the system, changing the temperature leaves the energy associated to a certain state unchanged. The new system will again obey detailed balance. This can be understood by first noting that it is in detailed balance a long time after the quench when it is equilibriated again; but then this means that the rates themselselves obey detailed balance, and since they are constant, they did this for the entire time after the quench, not just asymptotically (cf. the second definition of detailed balance in section 2.1.6 and illustrated in figure 2.3). Therefore, for the new temperature $T_2 = 1/(k_B\beta_2)$,

$$\frac{w_{c\to c'}}{w_{c'\to c}} = e^{-\beta_2(E_{c'}-E_c)} \tag{2.44}$$

According to (1.14), the logarithm of this expression is precisely the increase in environmental entropy such a transition produces, hence

$$\Delta S_{\text{env}}^{c\to c'} = -\beta_2(E_{c'} - E_c) \ . \tag{2.45}$$

Since after the quench $\beta = \beta_2$ is constant, this expression is easily applied jump-wise to (1.15), resulting in the simple functional dependency

$$\Delta S_{\text{env}}[\gamma] = -\beta_2(E_{c_T} - E_{c_0}) \ . \tag{2.46}$$

In order to connect this to the new class of fluctuation theorems obtained in the previous section, define the weights

$$\chi_{c_0,c_T} = \sqrt{\frac{p^{\text{init}}_{c_T}}{p^{\text{init}}_{c_0}}} = e^{-\frac{1}{2}\beta_1(E_{c_T}-E_{c_0})} \equiv e^{-\frac{1}{2}\beta_1\Delta E} \tag{2.47}$$

where p^{init}_c is the probability to obtain the configuration c in the initial distribution of the system (note how they fulfill the symmetry condition (2.24)). In particular, $p^{\text{init}}_{c_T}$ is the probability the system starts out in the final configuration. Inserted into (2.30), writing χ as an explicit factor instead of an index to the average, the result is

$$\left\langle e^{-\frac{1}{2}\Delta S_{\text{env}}} A(\Delta S_{\text{env}})\chi_{c_0,c_T} \right\rangle = 0 \ . \tag{2.48}$$

Using (2.46) this can be expressed in terms of ΔE:

$$\left\langle e^{+\frac{1}{2}\beta_2\Delta E} A(-\beta_2\Delta E)\chi_{c_0,c_T} \right\rangle = 0 \tag{2.49}$$

Upon inserting χ (2.47), rescaling A linearly to $\tilde{A}(\Delta E) = A(-\beta_2\Delta E)$ and $\Delta\beta = \beta_2 - \beta_1$, and absorbing $\mathcal{N}$ by the right hand side's zero, this can be written compactly as

$$\left\langle e^{+\frac{1}{2}\Delta\beta\Delta E} \tilde{A}(\Delta E) \right\rangle = 0 \ . \tag{2.50}$$

With the special choice $\tilde{A}(\Delta E) = \sinh(\frac{1}{2}\Delta\beta\Delta E)$, a final and simple integral fluctuation theorem is obtained:

$$\boxed{\left\langle e^{\Delta\beta\Delta E} \right\rangle = 1} \tag{2.51}$$

This formula provides an insight in the energy fluctuations of a system with constant rates when quenched out of equilibrium, on its way of reaching a new one. Most interestingly it holds for finite time, allowing to observe the relaxation process at an arbitrary point after the quench, and not just when the new equilibrium has been reached. (Note that applying Jensen's inequality to this formula merely yields that the energy in the system increases/decreases on average when the bath is heated/cooled – $\Delta\beta\langle\Delta E\rangle \leq 0$ – which does not provide any new insights.)

2.2.4 Numerical tests

This section describes two numerical simulations to test the theorems previously developed. The first example will pick up right where the last section ended with a concrete model, while the second one makes relatively little assumptions about the dynamics of a system, providing a more general perspective.

Simple growth process

To demonstrate the result of the last section, consider the following growth model, which was originally investgated to explain certain wetting phenomena on an inert substrate [20,21]. On an initially empty d-dimensional lattice with periodic boundary conditions, let $h_i \in \mathbb{N}_0$ be the height of site i. The dynamics now consist of repeatedly depositing particles to or evaporating them from the system at certain rates. However, the heights of adjacent sites are not independent, and subject to certain constraints.

Now consider the simple case of a 1-dimensional ring consisting of N sites ($i + N \equiv i$), subject to the constraint that adjacent sites' heights may only differ by a maximum of one,

$$|h_i - h_{i+1}| \leq 1 \ \forall i \ . \tag{2.52}$$

Particles are deposited at rate q_1 and evaporate at rate 1 (note that rescaling of the rates corresponds to rescaling of time, therefore evaporation can be set to 1 without loss of generality), illustrated in fig. 2.4. For $q_1 < 1$, it is known that the

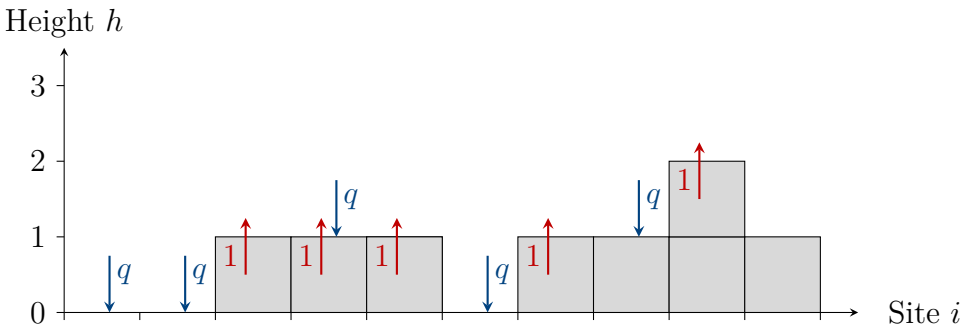

Figure 2.4: Illustration of a simple growth process. Particles are deposited and evaporated at rates q and 1, respectively. However, an event is only allowed if the invariant that adjacent sites can only differ by a maximum of height 1 is maintained; as a result, only the transitions indicated by arrows are permitted.

system behaves stationarily with an exponential probability distribution of the microstates,

$$P(\{h_i\}) \propto q_1^{\sum_j h_j} = e^{-\mu_1 H} \tag{2.53}$$

where $\mu = -\ln q$ can be interpreted as a chemical potential and $H = \sum_{i=1}^{N} h_i$ is the total number of particles in the system. Comparing this with a Boltzmann distribution reveals that μ plays the role of (inverse) temperature, and H is the energy contained in the system.

Making the connection to the previous section, a temperature quench manifests itself here as a "deposition rate quench", e.g. by changing said rate to another value $q_2 < 1$ at $t = \tau_0$, upon which the system will relax into a new equilibrium state. This is precisely what was previously done, and carrying over the analogy just mentioned in the last paragraph, this system should obey the fluctuation theorem

$$\langle e^{\Delta\mu\Delta H} \rangle = 1 \tag{2.54}$$

where $\Delta\mu = -(\ln q_2 - \ln q_1)$ is what previously was an inverse temperature difference.

Figure 2.5 shows how the above average develops in simulations of small systems in a billion trials. Each experiment confirms the fluctuation theorem by clearly converging to 1.

Surprisingly, the quench over the phase transition into a nonequilibrium state ($q_2 > 1$) shows converging behaviour, which is something the theorem does not predict, due to the assumption of having an equilibrium on either (far) end of the quenches. This suggests that the theorem could be more powerful than expected, and could e.g. be useful to find out properties of the critical point $q = 1$ of the phase transition. At this point this remains entirely speculative though, and would be a suitable subject for further investigations.

Nonequilibrium model on a small state space

To test the more general formula (2.28), namely

$$\tilde{P}(\Delta S_{\text{env}} = X) = e^X \tilde{P}(\Delta S_{\text{env}} = -X) \ ,$$

a less specialized setup is necessary. Here, we will only assume the most basic requirements for this theorem, namely having a Markov process with time-independent rates; more explicitly, there will be no assumptions of equilibirum or asymptotic behaviour.

The system consists of a set of states $c_i \in \Omega$ with $|\Omega| = 8$, and therefore a transition matrix $w_{c\to c'}$ with $8^2 - 8 = 56$ nonzero entries ($w_{c\to c} = 0 \ \forall c$). Its elements are randomly chosen (strictly) between 0 and 1 before the experiment (and left fixed throughout each run). In a similar fashion, the initial distribution p^{init} is determined.

A single run now consists of sampling an initial state from p^{init}, and letting it evolve in a Monte-Carlo simulation for some time T (here: 1), corresponding to a certain amount of attempted state transitions (here: 10^5). Each time the system's configuration changes, the generated ΔS_{env} is recorded. As an additional consistency check, the overall distriburion of final states was checked against a

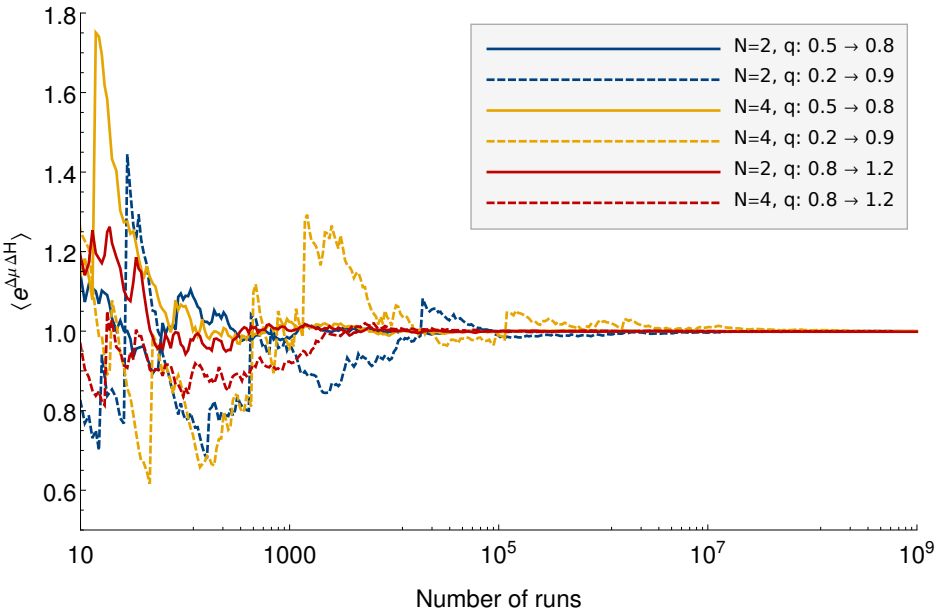

Figure 2.5: Convergence of $\langle e^{\Delta\mu\Delta H}\rangle$ to 1 in the growth model with system sizes of 2 and 4. The initially empty lattice is equilibriated using 50 time steps to reach the initial configuration, before the final state is obtained using three additional steps. Each run records one value $e^{\Delta\mu\Delta H}$, which is then averaged over many individual runs. The blue and yellow graphs show the development of said average over time in an ordinary quench, and confirm the fluctuation theorem's claim. The red line corresponds to changing the rates so that the system is brought beyond the phase transition ($q > 1$); in this scenario, there is no relaxation into a new stationary state, as the interface grows arbitrarily high over time. Remarkably, the fluctuation theorem seems to hold even in this case, despite the assumption of asymptotic equilibrium in its derivation.

numerical solution of the Mater equation

$$\dot{p}_c(t) = \underbrace{\sum_{c'} w_{c'\to c} p_{c'}(t)}_{\text{jump into } c} - \underbrace{\sum_{c'} w_{c\to c'} p_c(t)}_{\text{jump out of } c} \quad \forall c \ , \tag{2.55}$$

which describes the distribution of states on a statistical level directly, whereas the main simulation is based on individual runs, which are then later used for statistics.

A simulation now consists of a decent amount (10^7) individual runs, the result of which is then binned and put in a histogram. As can be seein in figures 2.6 and 2.7, the theorem is reproducible in simulation.

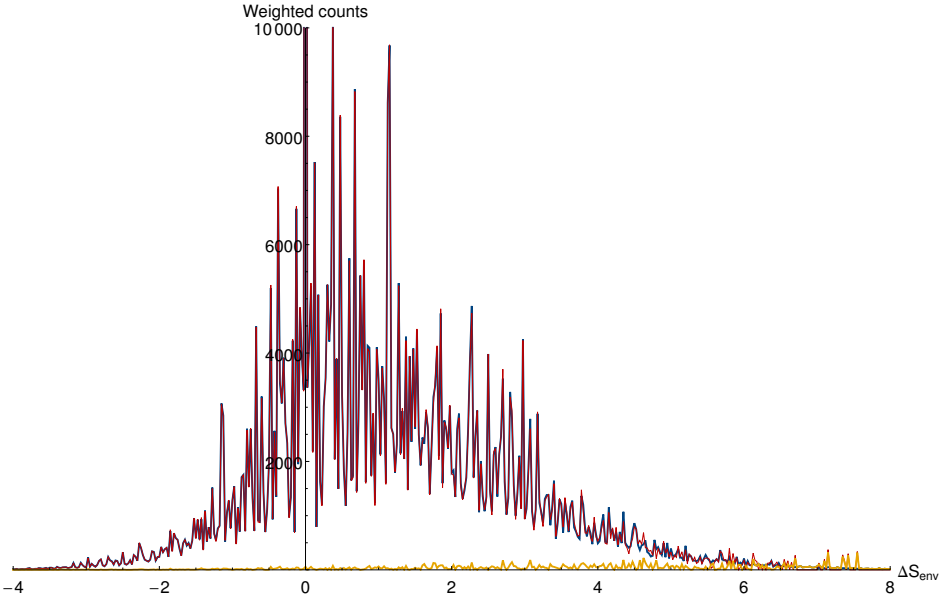

Figure 2.6: Linear histogram of a simulation of the fluctuation theorem for ΔS_{env}, obtained by binning the results of 10^7 individual runs. Shown in blue and red are the actual histogram and that histogram with the symmetry relation $f(x) = e^x f(-x)$ (corresponding to the fluctuation theorem) applied; the yellow line is the (absolute) deviation between these two representations. It can be seen that both histograms match well, with only a few barely visible red peaks overshooting, which is visualized by the yellow graph more clearly. The overlap is especially good in the region around $\Delta S_{\text{env}} = 0$ (most notably *at* 0, where two peaks about $5 \cdot 10^4$ units high meet), whereas in the large deviation regime towards the edges of the plot show rare events for which the statistics are not as good, but still well within acceptable bounds.

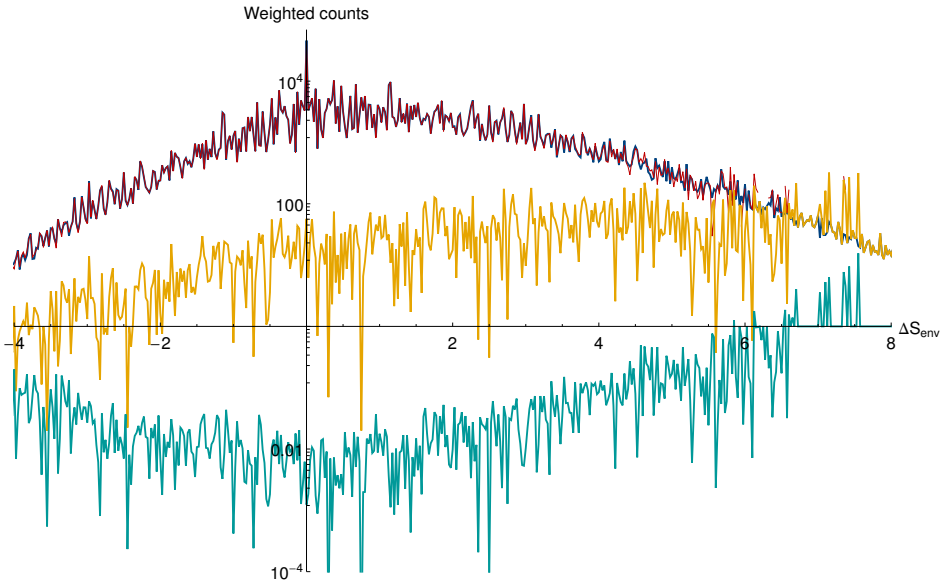

Figure 2.7: Same histogram as figure 2.6, but with logarithmic y-axis. In addition to the graphs for the histogram (blue), the symmetrized version (red, again just barely visible) and the absolute deviation (yellow), this also shows the relative deviation in cyan. This demonstrates again that both graphs match fairly well, particularly in around $\Delta S_{\text{env}} = 0$, where most events happen. Leaving the middle region the overlap becomes worse as $|\Delta S_{\text{env}}|$ increases, which can be explained by the fact that these events are relatively rare, hence the statistics aren't as good (compare e.g. about 100 weighted counts at $\Delta S_{\text{env}} = 7$ vs. 10^4 around $\Delta S_{\text{env}} = 1$).

2.3 Conclusion

One of the main results of this chapter is a set of fluctuation theorems for continuous time Markov processes. An important consequence – and the other main result – of this is a detailed fluctuation theorem for for the environmental entropy production ΔS_{env}, which holds even in finite time, and not just asymptotically. This suggests that it is suitable to study relaxation processes from arbitrary initial conditions into equilibrium, as well as nonequilibrium seady states (NESS).

One application of this result is to investigate a system that is initially in equilibrium, before it is suddenly heated or cooled to a different tempreature. Properties of the energy fluctuations of this system as it approaches the new equilibrium at the new temperature are captured by the mentioned fluctuation theorem. Even more, the simulation suggests that the theorem holds even when the target tem-

perature is in a nonequilibrium region – something assumed not to be the case in its derivation – suggesting that it may be suitable for studying phase transitions. This claim should be handled with care until the phenomenon is better understood though.

3 Entropy production in continuous phase space systems

The contents of this chapter are again an in-depth description of a publication, this time in the Journal of Statistical Physics, under the same title [22].

An overview is as follows: The first section will talk about how to transfer some of the ideas used in the previous chapter to a continuous setting, and what problems arise on the way there. Afterwards, in section 3.2, a brief overview over stochastic differential equations (SDEs) is given, before the main model of interest is introduced in section 3.3. Section 3.4, the chapter's main part, will then develop a new entropy definition, and compare it with a previously known one in the process.

The most important result will be that the production of environmental entropy is in a certain sense ambiguous, allowing multiple different definitions that all agree in their end results.

3.1 From discrete to continuous systems

The basis of chapter 2 was the assumption of a discrete system $\Omega = \{c_1, c_2, \ldots\}$. On the other hand, many processes occurring in nature are not discrete. This chapter describes an attempt to transfer the concept of environmental entropy used before to a continuous phase space system [23–26] governed by Hamiltonian equations of motion.

Recall that a Markov process is fully described by an initial configuration p^{init} and a set of transition rates w, which provide a way of getting from one state to the next. Conceptually, this can be continualized to the case where the configuration is the location of a particle, leading to a system described by

$$\dot{q}(t) = f(q,t) + g(q,t)\xi(t) \ , \tag{1.1}$$

which can be understood in two different ways. Coming from a Markov process, f corresponds to the bias of the transition rates to go towards a certain state, and

g accounts for how the noise depends on location. On the other hand, seen from a neutral standpoint, f is of course a force term, and $g\xi$ is the randomness of the system due to a heat bath represented by Gaussian noise $\xi(t)$, for example.

Most notably, the above equation does not include any form of momentum. This can be seen as a residue of the Markov property, which was previously informally described as "momentum-less"; in a continuous differential equation (DE), this description is much more meaningful, because very high damping gets rid of any momentum right away. For this reason, equations of the form (1.1) are called *overdamped*, and could for example describe the motion of a bacteria in water (which happens to be quite thick for small organisms [27]). Many of the quantities associated with Markov processes have clear corresponding quantities in the overdamped case, in particular the notion of a "jump" from one state to the other has a reverse, allowing the definition of environmental entropy to carry over quite directly [3],

$$\Delta S_{\text{env}} = \ln \frac{p[\gamma|q_0]}{p^\dagger[\gamma^\dagger|q_0^\dagger]} \tag{1.2}$$

where $p[\gamma]$ is the probability of taking the forward path γ, and $p^\dagger[\gamma^\dagger]$ is the likelihood of walking that same path in reverse direction. The $\dagger$ ("dagger") or *conjugation* operation encodes this reversal in the obvious way: for a process running from time $t = 0$ to T,

$$\begin{aligned} q^\dagger(t) &= q(T-t) \\ q_0^\dagger(t) &= q_T \\ t &\xrightarrow{\dagger} T - t \end{aligned} \tag{1.3}$$

where the last line is important if the coefficients in (1.1) have an explicit time dependency. Note that the $\dagger$ operation is *not reversing time*, it only stands for walking the same path in reverse directions, probably better described as "making the same decisions, but in reverse order". Figure 3.1 also illustrates this process.

Of course not all systems in nature are overdamped in the continuous case, just like not every discrete system has the Markov property. These non-overdamped – *underdamped* – systems take the full Hamiltonian phase space into account, and the Hamiltonian equations of motion take the form

$$\begin{aligned} \dot{q}(t) &= f_q(q,p,t) + g_q(q,p,t)\xi_q(t) \\ \dot{p}(t) &= f_p(q,p,t) + g_p(q,p,t)\xi_p(t) \end{aligned} \tag{1.4}$$

As will be discussed later, it is nontrivial to introduce the concept of environmental entropy production in this scenario. The key reason to this is that a path cannot be reversed in a straight-forward way (i.e. the "$\dagger$" operation is not that easy to

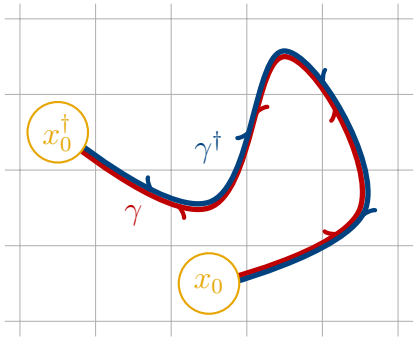

Figure 3.1: The knight from figure 2.2, but now in a continuous setting, where the configuration is $(x, y) \in \mathbb{R}^2$. The forward path γ is taken with probability $p[\gamma|x_0]$, and the backward/conjugate path $\gamma^\dagger$ occurs with probability $p^\dagger[\gamma^\dagger|x_0^\dagger]$. In both cases, the paths can be constructed so time runs forward as the particle follows the trajectory.

find), and the main result of this chapter is a new way this path reversal can be defined, and what the consequences are.

To explain this, consider simply carrying over the † operation from the above (overdamped) case. The stochastic trajectory, now through (q, p) space, will be transformed according to

$$\begin{aligned} \gamma &= \Big\{\big(q(t), p(t)\big)\Big\} \\ \xrightarrow{\dagger} \quad \gamma^{\dagger\text{naïve}} &= \Big\{\big(q^{\dagger\text{naïve}}(t), p^{\dagger\text{naïve}}(t)\big)\Big\} = \Big\{\big(q(T-t), p(T-t)\big)\Big\} \end{aligned} \tag{1.5}$$

but it turns out that the velocity $\dot{q}$ has a different sign as momentum p in this definition, leading to an unphysical path.

3.2 Stochastic differential equations

There are multiple equivalent approaches to stochastic differential equations (SDEs), i.e. differential equations with a probabilistic or "noise" term. There are two important distinctions to make between the three commonly used ones.

The goal of this section is not providing an introduction to SDEs, but to illustrate different formalisms briefly, and mention their key concepts. A more rigorous yet practical treatment of stochastic calculus can be found in e.g. [28], which also served as a main source for what follows.

3.2.1 Mathematics vs. physics

A stochastic differential equation in mathematics usually takes the form

$$\mathrm{d}X_t = f(X_t, t)\mathrm{d}t + g(X_t, t)\mathrm{d}W_t \tag{2.6}$$

describing how a *microscopic* quantity X evolves under the influence of a deterministic drive f and a stochastic contribution $g\,\mathrm{d}W_t$, where $\mathrm{d}W_t$ is the Wiener measure accounting for the stochasticity. The Weiner process is the fundamental building block in SDEs; W_t can be interpreted as the position of a random walk at time t, and $\mathrm{d}W_t$ is a "small" increase of this random walk as time advances a little. The solution of such a SDE will be a probability distribution for the value of X. Informally – but maybe more intuitive – the above equation is "divided by $\mathrm{d}t$" to yield

$$\dot{X}(t) = f(X(t), t) + g(X(t), t)\frac{\mathrm{d}W_t}{\mathrm{d}t}\ , \tag{2.7}$$

called a Langevin equation, and $\mathrm{d}W_t(t)/\mathrm{d}t$ can be interpreted as Gaussian noise $\xi(t)$,

$$\dot{X}(t) = f(X(t), t) + g(X(t), t)\xi(t)\ . \tag{2.8}$$

This is the form often used in physics. In this setting, the noise is usually defined by its statistical properties, most commonly

$$\langle \xi(t) \rangle = 0 \tag{2.9}$$

$$\langle \xi(t)\xi(t') \rangle \propto \delta(t - t') \tag{2.10}$$

and can therefore be pictured as a quantity fluctuating around 0 so wildly that there are no correlations between any given different points in time. The advantage of this approach of course is that ξ can be treated as an ordinary function with certain special properties.

3.2.2 Itō vs. Stratonovich

There is a surprising phenomenon when integrating SDEs numerically. The naïve approach would take the current state of the system $X(t)$, and extrapolate its new value using the SDE (using a suitably generated random number $\mathcal{R}$ for the noise) a timespan $\mathrm{d}t$ later, using

$$X(t + \mathrm{d}t) = f(X(t))\mathrm{d}t + g(X(t))\mathcal{R}\ . \tag{2.11}$$

However, this leaves one question open: why should the functions be evaluated at the starting point of the interval $[t, t + \mathrm{d}t]$, and not for example at $t_{\mathrm{eval}} = t + \mathrm{d}t/2$?

And indeed it turns out that the result depends on the choice of this evaluation point. In other words, there is an ambiguity in solving SDEs: all of the schemata

$$X(t + \mathrm{d}t) = f(X(t + a\,\mathrm{d}t))\mathrm{d}t + g(X(t + a\,\mathrm{d}t))\mathcal{R} \tag{2.12}$$

$$a \in [0, 1] \tag{2.13}$$

are theoretically valid ways of defining the integration of (2.6). Which value of a to choose depends on a number of factors, a couple of which are:

- *Modelling.* What is the nature of the noise? In finance, the randomness of the market is assumed to be a result of the current state and appears random due to the fact people are unpredictable. This is what the Itō method models by setting $a = 0$. On the other hand, for example in physics or biology, the noise often originates from a second, underlying and inaccessible process, such as temperature-related fluctuations, and setting $a = 1/2$ (= Stratonovich) accounts for this independence of noise and system by not preferring either side of the timestep over the other.

- *Pragmatism.* Some formulas are easier in certain formulations. A good example is the chain rule, which is called just that in the Stratonovich case, whereas its analogon has the telling name "Itō's lemma" [28].

In the context of these issues, it is important to mention that both approaches are equivalent and can be converted into each other via

$$\int_0^T f(W_t, t) \circ \mathrm{d}W_t = \int_0^T f(W_t, t)\,\mathrm{d}W_t + \frac{1}{2}\int_0^T \partial_x f(W_t, t)\,\mathrm{d}t\ , \tag{2.14}$$

where the convention is that "$\circ\,\mathrm{d}W_t$" stands for Stratonovich integration, and anything without special syntax is Itō. The spatial derivative ∂_x is to be read as with respect to f's first argument (which appears with explicit mention of the Wiener process in the integrands). As can be seen above, both approaches are identical if the integrand f is spatially constant; otherwise, the nonvanishing integral will be referred to as a *drift term*.

3.2.3 Micro vs. macro

SDEs like eq. (2.6) assume a single quantity X, track its behaviour over time, and accumulate this in a final probability distribution for its possible values. This is a microscopic approach, effectively equivalent to repeating the same experiment with identical initial conditions (i.e. particle in the same place) many times, recording each individual result, and putting that in a histogram.

However, individual partcles are rarely accessible[1], and the starting point of the experiment is a probability distribution in the first place. This scenario is described by the Fokker-Planck (FP) equation, which can be interpreted as a generalized Heat equation:

$$\begin{aligned} \partial_t P(x,t) &= -\partial_x(f(x,t)P(x,t)) + \partial_x^2(D(x,t)P(x,t)) \\ &= -\partial_x j(x,t) \end{aligned} \tag{2.15}$$

Like the Heat equation, the FP equation describes a flow in terms of a probability current, just that this current can be more complex. Nevertheless, it still describes the evolution of an entire distribution over time, and not of one of its constituents. It can be shown that the FP equation is an alternative and equivalent formulation of the SDE

$$\mathrm{d}X_t = f(X_t,t)\mathrm{d}t + \sqrt{2D(X_t,t)}\,\mathrm{d}W_t \ . \tag{2.16}$$

when integrated according to the Itō calculus [28].

3.3 Underdamped particle with additive noise

The treatment of the most general case of Hamiltonian motion, as stated for a single particle in (1.4), turns out to be very complicated. For this reason, the model used here is a special case of said equations. That is not to say it is impractical though; many of the basic systems of classical mechanics can be described by them:

$$\boxed{\begin{aligned} \dot{q} &= p \\ \dot{p} &= -V'(q) - \mu(p)p + \Gamma(p)\xi(t) \end{aligned}} \tag{3.17}$$

This models a single one-dimentional particle of unit mass in full phase space (q,p) subject to a conservative force field $-V'(q)$, momentum-dependent friction $\mu(p)$ and multiplicative noise $\Gamma(p)\xi(t)$, where $\xi(t)$ is Gaussian noise, implicitly defined via

$$\begin{aligned} \langle \xi(t) \rangle &= 0 \\ \langle \xi(t)\xi(t') \rangle &= 2D\delta(t-t') \end{aligned} \tag{3.18}$$

and D is a diffusive coefficient.

The main question this chapter is trying to answer is what mechanism leads to the production of environmental entropy in a system that follows the above

[1] For a case where this *is* possible, consider biophysical models of individual cells, or the Brownian motion of small particles in liquid [27].

dynamics. In the sense of classical thermodynamics, environmental entropy can be thought of as the flow of energy in the form of heat in and out of a system (held at constant temperature). In the specific case here, the noise ξ adds energy to the system (decreasing the entropy in the environment due to the resulting cooling), while the friction takes out heat (increasing the entropy in the environment). Equipped with this simple knowledge of classical thermodynamics, it is already possible to formulate a definition for environmental entropy for a special case of the present model, namely if the particle is in equilibrium with a heat bath of temperature T,

$$T\mathrm{d}S_{\mathrm{env}} = -\mathrm{d}Q \tag{3.19}$$

where Q is the heat contained in the system, and S_{env} is the environment's entropy. Since heat flowing from the environment into the system increases its entropy by $\mathrm{d}S_{\mathrm{env}}$ while decreasing the heat of the environment by $\mathrm{d}Q$, a negative sign appears in the equation. Due to conservation of energy (first law of thermodynamics in the absence of mechanical work), the heat flow is precisely the energy flow between the systems, $\mathrm{d}Q = \mathrm{d}E$, and E is simply the particle's energy,

$$E = \frac{p^2}{2} + V(q) \tag{3.20}$$

allowing the entropy production $\mathrm{d}S_{\mathrm{env}}$ to be written as

$$\begin{aligned} \mathrm{d}S_{\mathrm{env}}(q,p) &= -\beta\,\mathrm{d}\left(\frac{p^2}{2} + V(q)\right) \\ \Rightarrow \mathrm{d}S_{\mathrm{env}}(q,p) &= -\beta p\mathrm{d}p - \beta V'(q)\mathrm{d}q\ . \end{aligned} \tag{3.21}$$

Reproducing this result from a microscopic approach however is nontrivial, and poses a strong consistency check for any theory trying to explain environmental entropy production in a similar setting. It will later turn out that under certain circumstances, some *correct* theories fail to do so, yet yield correct physical results after another calculation. The reason and significance of this will be discussed at that point, where the terminology and formalism to deal with location-dependent "thermodynamic" quantities, like $\mathrm{d}S_{\mathrm{env}}(q,p)$ above, has been developed.

3.3.1 Fokker-Planck equation

A key quantity later will be the propagator of the Fokker-Planck equation, which can be seen as the continuous analogon of the Master quation (1.16). For this reason, (3.17) has to be transformed into Fokker-Planck form. This will be done following the notation of [24] and defining phase space coordinates $\boldsymbol{x} = (q,p)$, allowing it to be written as

$$\mathrm{d}x_i = A_i(\boldsymbol{x},t)\mathrm{d}t + B_i(\boldsymbol{x},t)\mathrm{d}W_i \tag{3.22}$$

where, as introduced in section 3.2.1, $\mathrm{d}W_i$ can be interpreted as the random walk generated by the noise $\xi(t)$.

The FP equation introduced in 3.2.3 was just concerned with a single dimension. In the present two-dimensional space (with coordinates $\boldsymbol{x}$), that equation has to be modified to

$$\partial_t P(\boldsymbol{x},t) = -\sum_{i=q,p} \partial_{x_i}(f_i(\boldsymbol{x},t)P(\boldsymbol{x},t)) + \partial_{x_i}^2(D_i(\boldsymbol{x},t)P(\boldsymbol{x},t)) \tag{3.23}$$

$$= -\sum_{i=q,p} \partial_{x_i} j_i(\boldsymbol{x},t) \tag{3.24}$$

taking into account the individual fluctuations each component of $\boldsymbol{x}$ may be subject to. The coefficients are related to the above Langevin equation via

$$f_i(\boldsymbol{x},t) = A_i(\boldsymbol{x},t) \tag{3.25}$$

$$D_i(\boldsymbol{x},t) = \frac{1}{2}B_i(\boldsymbol{x},t)^2 \tag{3.26}$$

when integrating accoring to Itō scheme, done so for pragmatic reasons (there will be plenty of probably more interesting ambiguities later). The FP equation corresponding to (3.17) there turns out to be

$$\begin{aligned} \partial_t P(q,p,t) &= \mathcal{L}\,P(q,p,t) \\ \mathcal{L} &= \mu(p) + p\mu'(p) + \Gamma(p)\Gamma''(p) + \Gamma'(p)^2 \\ &\quad + \Big(p\mu(p) + 2\Gamma(p)\Gamma'(p) + V'(q)\Big)\partial_p \\ &\quad + \frac{1}{2}\Gamma(p)^2\partial_p^2 - p\partial_q\ . \end{aligned} \tag{3.27}$$

3.3.2 Detailed balance in continuous systems

The idea of detailed balance (section 2.1.6) is also applicable to continuous systems, where as mentioned the Fokker-Planck equation (2.15) takes over for the Master equation (1.16). Analogous, a stationary FP equation means all macroscopic state variables are constant; in the present context it is of particular interest that no heat or environmental entropy is exchanged with the bath.

This can probably be better understood than in the discrete scenario even, where the explanation was quite abstract ("probability currents cancel out"). Eqns. (3.17) contain a friction term responsible for heating up the environment as the system moves, and also a noise term transferring energy (also in the form of heat) into the system. If these two contributions cancel out just right, then all that's left is a Hamiltonian system of a free particle in a potential, which is of course reversible and does not produce any heat.

Being in detailed balance introduces a connection between the friction terms $\mu(p)$ and the noise coefficient $\Gamma(p)$ in (3.17). This relation can be obtained explicitly by inserting a Boltzmann distribution (which is the equilibrium distribution of a system in thermal equilibrium with a heat bath) into the FP equation (3.23) and setting it equal zero, namely

$$P_{\text{Boltz}}(q,p) = \frac{1}{Z} e^{-\beta E(q,p)} = \frac{1}{Z} \exp\Big(-\beta\Big(\tfrac{p^2}{2} + V(q)\Big)\Big) \tag{3.28}$$

The result is a first-order differential equation for $\mu(p)$,

$$\begin{aligned} 0 = \frac{1}{2}\beta^2 p^2 \Gamma(p)^2 - \beta p^2 \mu(p) - 2\beta p \Gamma(p)\Gamma'(p) - \frac{1}{2}\beta\Gamma(p)^2 \\ + \Gamma(p)\Gamma''(p) + \Gamma'(p)^2 + p\mu'(p) + \mu(p) \ . \end{aligned} \tag{3.29}$$

Thanks to computer algebra systems[2], the solution of this equation is obtained quite easily, it is

$$\mu(p) = \frac{1}{2}\beta\Gamma(p)^2 - \frac{1}{p}\Gamma(p)\Gamma'(p) + C\frac{1}{p}e^{\frac{\beta p^2}{2}} \tag{3.30}$$

where C is an arbitrary integration constant. However, since the last term is exponentially divergent for $p \to \infty$, it unphysical: the mean acceleration of a system in equilibrium would be infinite. Therefore $C = 0$ is the only suitable choice, and what remains is

$$\mu(p) = \frac{1}{2}\beta\Gamma(p)^2 - \frac{1}{p}\Gamma(p)\Gamma'(p) \ . \tag{3.31}$$

This generalizes the usual condition for detailed balance $\mu = \frac{1}{2}\beta\Gamma^2$ (also known as the Einstein relation) for momenum-independent ("additive") noise and linear friction.

3.3.3 Short-time propagator

General case

Like in the discrete case, a key ingredient to defining entropy production in a continuous setting is the notion of a path probability. This probability is given by the propagator or Green's function of the FP equation. It will be denoted $G(\boldsymbol{x}'|\boldsymbol{x}; \Delta t)$, and can be read as the probability that a particle starting at phase space coordinates $\boldsymbol{x} = (q,p)$ is found at a new point $\boldsymbol{x}' = (q',p')$ at Δt time later.

[2]Mathematica 9.0.0.0

In the present work however, it will be sufficient to take the *short-time* proagator into account, i.e. Δt is very small.

The short-time propagator $G(\boldsymbol{x}'|\boldsymbol{x};\, \mathrm{d}t)$ solves the FP equation to leading order in $\mathrm{d}t$, and is therefore not uniquely defined:

1. The short-time propagator need not be Gaussian, as long as the ("arbitrary-time") macroscopic propagator is reproduced by it. According to the central limit theorem, any distribution with the correct first and second moments is sufficient. (Intuitively, this can be understood by the fact that many functions have identical first-order power series coefficients, but their actual global shape is influenced by the higher-order terms. In the present scenario, these terms are neglected, leading to said ambiguity.)

2. It is not clear where the fields contained in the FP equation (3.23), $f_i(\boldsymbol{x}, t)$ and $D_i(\boldsymbol{x}, t)$, have to be evaluated: at $\boldsymbol{x}$, $\boldsymbol{x}'$, between, somewhere in the neighbourhood? In general, an evaluation point at any $\boldsymbol{r} = \varphi(\boldsymbol{x}, \boldsymbol{x}')$ with bijective φ will do [29]. Conceptually, this is similar to the Itō-Stratonovich dilemma introduced in the conext of stochastic differential equations in section 3.2.2, but it is important to emphasize that this is a *new, independent* ambiguity in addition to the previous one (which was resolved to integration according to Itō scheme to obtain the FP equation from the SDE).

To get a hold of the infinitely many short-time propagators, assume the evaluation point is located linearly between the two end points,

$$\boldsymbol{r} = (1-a)\boldsymbol{x} + a\boldsymbol{x}' \qquad a \in [0,1]\ , \tag{3.32}$$

as done in [24]. Using [29, 30] this leads to the following expression for the propagator:

$$\begin{aligned} G_a(\boldsymbol{x}'|\boldsymbol{x};\, \mathrm{d}t) = \prod_i \frac{1}{\sqrt{2\pi D_i \mathrm{d}t}} \\ \exp\left(-\frac{(\mathrm{d}x_i - A_i \mathrm{d}t + 2aD_i' \mathrm{d}t)^2}{4D_i \mathrm{d}t} - aA_i' \mathrm{d}t + a^2 D_i'' \mathrm{d}t \right) \end{aligned} \tag{3.33}$$

with the short-hands

$$\begin{aligned} A_i &= A_i(\boldsymbol{r}, t) & A_i' &= \partial_{r_i} A_i(\boldsymbol{r}, t) & \\ D_i &= D_i(\boldsymbol{r}, t) & D_i' &= \partial_{r_i} D_i(\boldsymbol{r}, t) & D_i'' &= \partial_{r_i}^2 D_i(\boldsymbol{r}, t)\ . \end{aligned}$$

Applied to the model

The system in question, (3.17), consists of two coupled differential equations, but only one of them is stochastic, giving $D_1 \equiv 0$. However, the propagator (3.33) requires dividing by this quantity so it is not well-defined. For this reason, (3.17) is modified to add a "very small" (new, independent from ξ_p) noise term to the location, resulting in the system

$$\begin{aligned} \dot{q}(t) &= p + \varepsilon\xi_q(t) \\ \dot{p}(t) &= -V'(q) - \mu(p)p + \Gamma(p)\xi_p(t) \ ; \end{aligned} \tag{3.34}$$

the coefficient ε will later fall away on its own, justifying this rather ad-hoc addition. The resulting propagator for this model then reads

$$\begin{aligned} {}_\varepsilon G_a(q', p'|q, p;\, \mathrm{d}t) = \frac{1}{2\pi\varepsilon\Gamma(r)\mathrm{d}t} \exp\Big(& a^2(\Gamma(r)\Gamma''(r) + \Gamma'(r)^2)\mathrm{d}t \\ & - \frac{\Big(V'(q + a\,\mathrm{d}q)\mathrm{d}t + 2a\Gamma(r)\Gamma'(r)\mathrm{d}t + \mathrm{d}p + r\mu(r)\mathrm{d}t\Big)^2}{2\Gamma(r)^2\mathrm{d}t} \\ & + a\Big(r\mu'(r) + \mu(r)\Big)\mathrm{d}t - \frac{(\mathrm{d}q - r\mathrm{d}t)^2}{2\varepsilon^2\mathrm{d}t}\Big) \end{aligned} \tag{3.35}$$

with

$$\mathrm{d}q = q' - q \qquad \mathrm{d}p = p' - p \qquad r = (1-a)p + ap' \ .$$

As a consistency check, consider the limit $\varepsilon \to 0$. According to the differential equations above, this system should be deterministic in the position coordinate. And indeed the terms involving ε in (3.35) are of Gaussian shape, just like in the usual derivation of the δ distribution from a narrow Gaussian. More explicitly, the limit reads

$$G_a(q', p'|q, p;\, \mathrm{d}t) = \frac{\delta(\mathrm{d}q - r\mathrm{d}t)}{\sqrt{2\pi\mathrm{d}t}\Gamma(r)} \exp(\cdots) \tag{3.36}$$

where the δ accounts for the deterministic location, and the exponent "..." is independent of the location q.

3.4 Differential entropy production

The focus of the main part of this chapter is the definition of differential environmental entropy $\mathrm{d}S_{\mathrm{env}}$. The term *differential* will be used in multiple similar but different contexts; it can always be understood as some form of Taylor expansion of

the finite entropy increase ΔS_{env} in $\mathrm{d}t$, $\mathrm{d}q$ and $\mathrm{d}p$ to lowest order. Which meaning is appropriate will be specified explicitly in ambiguous places.

The outline of what follows is this. First, an overview of Spinney and Ford's approach to defining $\mathrm{d}S_{\text{env}}$ is given, which leads to the question what should be "the state" of an SDE. Inspired by this question, a new way of defining $\mathrm{d}S_{\text{env}}$ will be proposed, and frequently compared to the other approach in the process of finding the corresponding formulas.

3.4.1 Spinney and Ford's entropy definition

In order to solve the issue of naïve path reversal mentioned in section 3.1, Spinney and Ford (SF) [24] choose to reverse time itself ($t \to -t$); more precisely, they take the time parity of all quantities into account, and mirror only those changing sign under time reversal. They define "†" as [25]

$$\begin{aligned} \gamma &= \Big\{\big(q(t), p(t)\big)\Big\} \\ \xrightarrow{\dagger} \quad \gamma^{\dagger_{\text{SF}}} &= \Big\{\big(q^{\dagger_{\text{SF}}}(t), p^{\dagger_{\text{SF}}}(t)\big)\Big\} = \Big\{\big(\varepsilon q(T-t), \varepsilon p(T-t)\big)\Big\} \end{aligned} \tag{4.37}$$

where ε is a parity operator, which is $+1$ for even quantities such as q and -1 for odd ones such as p under time reversal. Using this, they are able to use the entropy definition

$$\Delta S_{\text{env}} = \ln \frac{p[\gamma|q_0]}{p^{\dagger_{\text{SF}}}[\gamma^{\dagger_{\text{SF}}}|q_0^{\dagger_{\text{SF}}}]} \tag{4.38}$$

to derive an expression for the differential environmental entropy production of arbitrary underdamped stochastic systems in full phase space [24], given by

$$\mathrm{d}S_{\text{env}}^{\text{SF}} = \ln \frac{G_a(\boldsymbol{x}'|x;\, \mathrm{d}t)}{G_b(\boldsymbol{x}'^{\dagger}|\boldsymbol{x}^{\dagger};\, \mathrm{d}t)} \;; \tag{4.39}$$

recall that the propagator $G_a(\boldsymbol{x}'|x;\, \mathrm{d}t)$ corresponds to the probability to go from $\boldsymbol{x} = (q, p)$ to $\boldsymbol{x}'$ in the short time $\mathrm{d}t$, plus an ambiguity parameter a. In the case of the underdamped particle (3.34) this then reads

$$\mathrm{d}S_{\text{env}}^{\text{SF}} = \ln \frac{{}_{\varepsilon}G_a(q', p'|q, p;\, \mathrm{d}t)}{{}_{\varepsilon}G_b(q, -p|q', -p';\, \mathrm{d}t)} \;. \tag{4.40}$$

A graphical representation of the paths constructed here can be found in the left part of figure 3.2 (page 45).

In order to make this equation well-behaved in the limit of the ad-hoc parameter $\varepsilon \to 0$, both parts of the fraction have to peak at the same location; taking a step

away from mathematics, this can be understood as the δ distributions in (3.36) cancelling. This leads to the condition

$$(q' - q) - (p + a(p' - p))\mathrm{d}t = -\Big((q - q') - (-p' + b(-p + p'))\mathrm{d}t\Big) \tag{4.41}$$

which is solved by $b = 1 - a$, and therefore eliminates one of the propagator ambiguity parameters in (4.40). SF obtain the same relation after a lengthy derivation in their appendix, and choose $a = 0$ in the main text without further comment [24]. Direct calculation of the entropy production results in the expression

$$\begin{aligned} \mathrm{d}S^{\mathrm{SF}}_{\mathrm{env}} = \frac{1}{\Gamma^2}\Big(&-\mu\Gamma^2 - \Gamma^3\Gamma'' + \Gamma^2\Gamma'^2 - \Gamma^2 p\mu' \\ &+ 2\mu\Gamma p\Gamma' - 2\mu p V' - 2\Gamma\Gamma' V'\Big)\mathrm{d}t \\ &+ \frac{1}{\Gamma^2}(-2\Gamma\Gamma' - 2\mu p)\mathrm{d}p + \mathcal{O}(\mathrm{d}t^2) + \mathcal{O}(\mathrm{d}p^3) \end{aligned} \tag{4.42}$$

where for the sake of brevity the explicit arguments of $\Gamma(p)$, $\mu(p)$ and $V(q)$ have been dropped. Note that the series expansion is to first order in $\mathrm{d}t$ and *second* order in $\mathrm{d}p$.

This formula has some surprising properties. Most notably, in the case of a particle subject to linear friction ($\mu(p) = \text{const.}$) in detailed balance ($\Gamma = \sqrt{2T\mu}$) it reduces to

$$\mathrm{d}S^{\mathrm{SF}}_{\mathrm{env}} = -\beta p\mathrm{d}p - \beta V'(q)\mathrm{d}q - \mu\mathrm{d}t \tag{4.43}$$

which has an additional term $-\mu\mathrm{d}t$ compared to the textbook example (3.21), indefinitely producing (negative) environmental entropy regardless of the state of the system – take for example a resting free particle, $p = 0, V = \text{const.}$ to make this obvious – when there should be no such production. This implausibility was in fact one of the motivations behind investigating the topic of continuous entropy further, although this specific issue was resolved upon further investigation (section 3.4.3).

SF's approach has some further conceptually unsatisfactory properties though, first and foremost the $\dagger$ operation is non-local in phase space: it maps $p \to -p$, which may be in a completely different part of phase space where the Hamiltonian flow is not necessarily identical to the mirror image of the original flow. Furthermore, the start and end points of the normal trajectory are different from the ones for the conjugate process, which seems to be inconsistent with the idea behind the discrete environmental entropy production (1.14).

One last thing to mention here due to its later relevance is that the calculation to obtain (4.42) can be done in the general case $a = 1 - b$ instead of setting $a = 0$

from the beginning on. The result is quite long and was therefore therefore moved appendix A.1, but for $a = b = \frac{1}{2}$ the resulting entropy reads

$$\begin{aligned} \mathrm{d}S_{\mathrm{env}}^{\mathrm{SF}_{1/2}} = &\frac{1}{\Gamma^2}\Big(-2p\mu V' - 2\Gamma\Gamma' V'\Big)\mathrm{d}t + \frac{1}{\Gamma^2}\Big(-2p\mu - 2\Gamma\Gamma'\Big)\mathrm{d}p \\ &+ \frac{1}{\Gamma^2}\Big(-p\mu' + 2p\mu\tfrac{\Gamma'}{\Gamma} - \mu - \Gamma\Gamma'' + \Gamma'^2\Big)\mathrm{d}p^2 \\ &+ \mathcal{O}(\mathrm{d}t^2) + \mathcal{O}(\mathrm{d}p^3) \end{aligned} \tag{4.44}$$

which is of course different from the $a = 0$ result. Not very meaningful on its own, this term will reappear again in a different approach to entropy; the significance of this will become apparent later.

3.4.2 What is "the state" of a stochastic system in continuous phase space?

The entropy definition, even in the continuous case, is based on the idea that environmental entropy is produced when the system changes from one state to another. SF's approach implicitly considers such a state to be the current location in phase space. As shown just above, this notion leads to strange results, illustrating an important difference between discrete and continuous settings: while randomness is the only source of dynamics in discrete models, continuous phase space systems consist of two parts: random *and deterministic*. When a continuous system cℌanges "its state" in the previous sense, i.e. moves to a different point in phase space, this is not necessarily due to the noise, but could just be a consequence of deterministic processes. For example the "SDE" of the Hamiltonian model (3.17) with $\mu \equiv \Gamma \equiv 0$, i.e. only the deterministic part of the system, seems to change state, but as it moves deterministically on the Hamiltonian flow it of course produces no entropy.

This realization leads to the question whether "the state" of a stochastic system is maybe something different, and not just the current location in phase space, and as will be shown in the following, an alternative definition can be made that solves these issues.

3.4.3 Hamiltonian orbit based entropy definition

Reversible processes do not produce entropy, and any Hamiltonian process is reversible. A new definition of environmental entropy should satisfy this constraint, and therefore yield $\Delta S_{\mathrm{env}} = 0$ if a process follows a single Hamiltonian orbit; quite naturally, the idea that "a state switch" is a switch of such orbits, leading from one Hamiltonian process to another, lends itself as the basis for a new entropy

definition, which will be developed in a moment. This approach will also have the side effect of eliminating the need to distinguish between odd and even variables, and yield a somewhat different expression for ΔS_{env}. However, it will turn out that the observables predicted by the new model lead to the *same* expressions as SF's approach, indicating that there is a certain ambiguity in the way ΔS_{env} cab be defined. The only necessary condition on the dynamics of the system for the following approach is that the equations of motion can be separated in Hamiltonian and non-Hamiltonian parts, so that the Hamiltonian orbits (i.e. the flow of the Hamiltonian part) can be identified.

More detailed, the new definition of the forward and backward paths is based on two Hamiltonian orbits σ and σ'. The forward path γ is given in the same way as in SF's model, but the construction of the conjugate path – denoted as $\tilde{\gamma}$ instead of $\gamma^\dagger$ to distinguish it from SF's approach – is different, as illustrated in figure 3.2 (page 45). The algorithm for the construction is as follows:

1. The start point of the process is a phase space coordinate $\boldsymbol{x}$, located on a Hamiltonian orbit σ.

2. The system evolves according to the full dynamics for a timespan $\mathrm{d}t$, reaching a point $\boldsymbol{x}'$ after following a trajectory γ. $\boldsymbol{x}'$ is located on some other Hamiltonian orbit σ'.

3. The weight of this path γ is given by the propagator $G_a(\boldsymbol{x}'|\boldsymbol{x};\, t, \mathrm{d}t)$.

4. The starting point of the conjugate process $\tilde{\boldsymbol{x}}'$ is reached by tracing back the Hamiltonian orbit through $\boldsymbol{x}'$, σ', for a timespan $\mathrm{d}t$.

5. Similarly, the end point of the conjugate process $\tilde{\boldsymbol{x}}$ is where the system would have gone from $\boldsymbol{x}$, had it only followed the Hamiltonian orbit σ for a time $\mathrm{d}t$.

6. The weight of the conjugate process is $G_b(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{x}}';\, t, \mathrm{d}t)$.

This can be interpreted as calculating the weight of a jump between the orbits σ and σ', and enables the new definition of entropy production in the familiar terms of (1.15):

$$\boxed{\Delta S^{\text{HF}}_{\text{env}}(\sigma'|\sigma) = \ln \frac{P_{\sigma\to\sigma'}}{P_{\sigma'\to\sigma}}} \tag{4.45}$$

$$\boxed{\mathrm{d}S^{\text{HF}}_{\text{env}}(\boldsymbol{x}'|\boldsymbol{x};\, \mathrm{d}t) = \ln \frac{G_a(\boldsymbol{x}'|\boldsymbol{x};\, t, \mathrm{d}t)}{G_b(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{x}}';\, t, \mathrm{d}t)}} \tag{4.46}$$

where *HF* is short for *Hamiltonian Flow*, as opposed to SF for the previous model. Note that the propagator ambiguities a and b are independent at this point, and

will be dealt with later when the entropy definition is applied to a concrete model. This definition will also frequently be referred to as "Flow entropy".

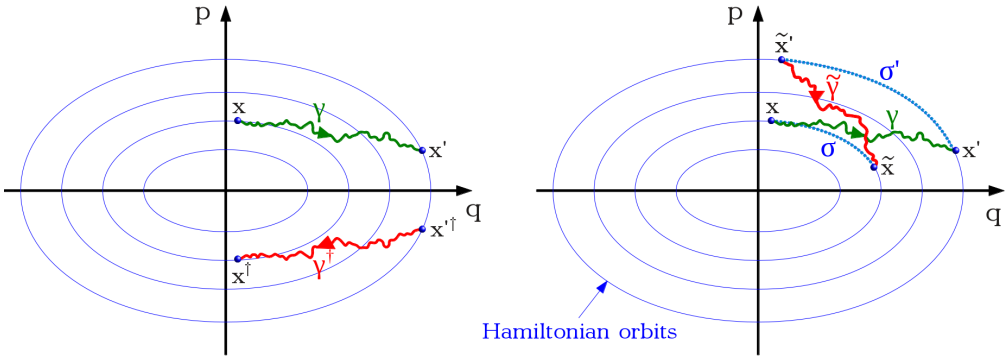

Figure 3.2: Path construction in the two models. The left side shows SF's approach, which mirrors the forward trajectory γ into a different part of phase space (yielding $\gamma^\dagger$), and calculates the ratio of the weights of these two paths. On the right hand side is the flow-based approach, which does not mirror momenta; instead, it traces the Hamiltonian flow in order to reach the start and end points of the conjugate process $\tilde{\gamma}$. (Image taken from [22].)

3.4.4 Differences between the models

In order emphasize the differences between the models recall figure 3.2, where the two path constructions (SF and flow) are pictured alongside each other. Important consequences compared to SF's model are as follows:

- The HF entropy definition is local in phase space: nothing is mirrored into a different part of phase space, and the entire quotient is constrained to an arbitrarily small part of phase space for sufficiently small timespans.
- Inside the HF entropy definition, distinction between odd and even variables is necessary, again because of the absence of time or momentum reversal.
- No differential entropy is produced for "transitioning" onto the same orbit again in the HF case, whereas for SF this is not necessarily so.

The second point is particularly important, as it contains what is probably the crucial conceptual difference between HF and SF. For HF, the splitting is done *a priori* with only the equations of motion known, in particular before entropy is even mentioned. On the other hand, in Spinney/Ford's case, reversal of dynamics

is hard-wired into the entropy definition itself. In other words, **the HF approach factors out the reversal of dynamics**.

In order to obtain the conjugate propagator to the original forward process in the HF model, the Hamiltonian part of the equations of motion needs to be known, which allows tracing the Hamiltonian flow back along the deterministic trajectory to reach the starting position of the reverse process, and similarly for the end position. In other words, the equations of motions must be partitioned in two parts, one for the Hamiltonian flow and one for the non-Hamiltonian processes such as dissipation and friction.

The principle behind the algorithm for the separation is inspired by the definition of the Riemann tensor: the system is run in a process that should return to the origin in the end; if it does not, the discrepancy is characteristic for a certain property of that system. In case of the Riemann tensor that discrepancy is the curvature of space, in phase space system it accounts for the non-Hamiltonian parts of the dynamics.

Formally, the most general equations of motion read, with $x = q$ or p:

$$\dot{x} = f_x(q,p) + \Gamma_x(q,p)\xi_x(t) \tag{4.47}$$

This can be split up in Hamiltonian ($\dot{x}_H$) and non-Hamiltonian ($\dot{x}_\Delta$) parts:

$$\dot{x} = \dot{x}_H + \dot{x}_\Delta \tag{4.48}$$

To obtain the summands individually, the following algorithm is used:

1. Calculate $x(t + \mathrm{d}t/2)$.
2. Reverse the system's dynamics by substituting $p \to -p$.
3. Evolve the for another $\mathrm{d}t/2$, starting at the previously reached end point.
4. The end position is $x(t + \mathrm{d}t)$, from which $x_\Delta = x(t + \mathrm{d}t) - x(t)$ can be determined.

For example, this procedure applied to the underdamped particle discussed in the present work

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -V'(q) - \mu(p)p + \Gamma(p)\xi(t) \end{aligned}$$

results in

$$\begin{aligned} \dot{q}_H &= p & \dot{q}_\Delta &= 0 \\ \dot{p}_H &= -V'(q) & \dot{p}_\Delta &= -\mu(p)p + \Gamma(p)\xi(t)\ . \end{aligned}$$

This approach is quite general and can be applied to more complicated systems, in which the separation may not be as clear.

3.4.5 Flow-based entropy production of the model

This new model can now be applied to the one-dimensional underdamped particle introduced before in section 3.3. Here, the forward process starts at $\boldsymbol{x} = (q, p)$ and ends at $\boldsymbol{x}' = (q', p')$. The conjugate process starts at $\tilde{\boldsymbol{x}}' = (q' - p'\mathrm{d}t, p' - f(q')\mathrm{d}t)$, which as described before is where tracing back the Hamiltonian orbit through (q', p') for a timespan $\mathrm{d}t$ reaches; similarly, $\tilde{\boldsymbol{x}} = (q + p\mathrm{d}t, p + f(q)\mathrm{d}t)$. The entropy thus reads

$$\mathrm{d}S^{\mathrm{HF}}_{\mathrm{env}}(\boldsymbol{x}'|\boldsymbol{x};\,\mathrm{d}t) = \lim_{\varepsilon\to 0} \frac{{}_{\varepsilon}G_a(q', p'|q, p;\,\mathrm{d}t)}{{}_{\varepsilon}G_b(q + p\mathrm{d}t, p + f(q)\mathrm{d}t)|q' - p'\mathrm{d}t, p' - f(q')\mathrm{d}t;\,\mathrm{d}t)} \tag{4.49}$$

where the propagator has no explicit t dependence, which has therefore been dropped in the notation. The same argument that has been made in the SF case in (3.36) – namely that in the limit $\varepsilon \to 0$ the location has to behave deterministically, giving a condition on the ambiguity parameters a and b – can be used here to yield

$$\begin{aligned}
&(q' - q) - \Big(p + a(p' - p)\Big)\mathrm{d}t \\
&= -\Big((q - q') + (p + p')\mathrm{d}t \\
&\qquad - \Big(p' - f(q')\mathrm{d}t + b(p - p' + (f(q) + f(q'))\mathrm{d}t)\Big)\Big)
\end{aligned} \tag{4.50}$$

which holds iff $a = b = \frac{1}{2}$, as can be shown by expanding the force $f(q')$ around q and neglecting terms of $\mathcal{O}(\mathrm{d}t^3)$. This makes a lot of sense, as it means that the fields are evaluated around the center of each trajectory, which to leading order in time is the same for both processes. This result is much stronger than what is obtained in SF's case, where $a = b$ is imposed, without giving them a specific value.

Like in the SF chapter before, the entropy production (4.49) can be calculated in this case, yielding

$$\begin{aligned}
\mathrm{d}S^{\mathrm{HF}}_{\mathrm{env}} = &\frac{1}{\Gamma^2}\Big(-2p\mu V' - 2\Gamma\Gamma' V'\Big)\mathrm{d}t + \frac{1}{\Gamma^2}\Big(-2p\mu - 2\Gamma\Gamma'\Big)\mathrm{d}p \\
&+ \frac{1}{\Gamma^2}\Big(-p\mu' + 2p\mu\frac{\Gamma'}{\Gamma} - \mu - \Gamma\Gamma'' + \Gamma'^2\Big)\mathrm{d}p^2 \\
&+ \mathcal{O}(\mathrm{d}t^2) + \mathcal{O}(\mathrm{d}p^3)
\end{aligned} \tag{4.51}$$

where once again the explicit dependencies of the coefficients have been dropped for brevity. This result shows some important differences compared to SF's result (4.42):

- No entropy is produced to leading order in $\mathrm{d}t$ along the deterministic trajectory $\mathrm{d}p = -V'(q)\mathrm{d}t$. This is of course a direct result of the construction, which specifically addressed the issue that only switching orbits should yield an entropy contribution.
- While the SF result does not include a term proportional to $\mathrm{d}p^2$, the expression here does.
- Applied to a particle in a system with linear friction and additive noise in detailed balance, the expression reduces to the expected textbook result (3.21).
- The result coincides with the result SF would have gotten, had they not chosen $a = \frac{1}{2}$ instead of 0 in their model – (4.44) is identical to (4.51)! This brings up the question once again as to why they made this specific choice, but as mentioned earlier, this is done without further comment in [24].

3.5 Towards observable quantities

The previously defined quantity $\mathrm{d}S_{\mathrm{env}}(\boldsymbol{x}'|\boldsymbol{x};\,\mathrm{d}t)$ is a two-point function, yielding the expected entropy production for the transition $\boldsymbol{x} \to \boldsymbol{x}'$. However, this quantity is not easily accessible by experiment or simulation: while the starting point $\boldsymbol{x}$ can be set, $\boldsymbol{x}'$ is much less accessible due to it being statistically distributed over arbitrary phase space regions. The following sections will successively integrate out the unknowns, in order to yield a testable result.

3.5.1 Local entropy production

The *local entropy production rate* integrates out the final phase space coordinate:

$$\mathrm{d}S_{\mathrm{env,\,loc}}(\boldsymbol{x};\,\mathrm{d}t) = \int_\Omega \mathrm{d}\boldsymbol{x}' G_c(\boldsymbol{x}'|\boldsymbol{x};\,\mathrm{d}t)\mathrm{d}S_{\mathrm{env}}(\boldsymbol{x}'|\boldsymbol{x};\,\mathrm{d}t) + \mathcal{O}(\mathrm{d}t^2) \tag{5.52}$$

This describes the expected change in environmental entropy in a short timespan $\mathrm{d}t$, given an initial location $\boldsymbol{x}$ in phase space Ω. Note that the propagator to encode the transition likelihood contains yet another ambiguity parameter in addition to a and b implicitly contained in $\mathrm{d}S_{\mathrm{env}}$.

Translated to the present model of a particle in two-dimensional phase space, the equation reads

$$\begin{aligned}\mathrm{d}S_{\mathrm{env,\,loc}}(q,p;\,\mathrm{d}t) = &\int_{-\infty}^{\infty}\mathrm{d}q' \int_{-\infty}^{\infty}\mathrm{d}p'\; G_c(q',p'|q,p;\,\mathrm{d}t)\mathrm{d}S_{\mathrm{env}}(q',p'|q,p;\,\mathrm{d}t) \\ &+ \mathcal{O}(\mathrm{d}t^2)\;.\end{aligned} \tag{5.53}$$

To calculate its value, first expand the ordinary differential entropy production in terms of differences of its arguments,

$$\mathrm{d}S_{\mathrm{env}}(q', p'|q, p;\, \mathrm{d}t) = \sum_{i,j=0}^{\infty} \sigma_{ij}(q, p, \mathrm{d}t)(q' - q)^i (p' - p)^j + \mathcal{O}(\mathrm{d}t^2) \tag{5.54}$$

so that σ_{ij} are the Taylor coefficients of $\mathrm{d}S_{\mathrm{env}}$ in terms of q and p. Inserting this expansion in (5.53) yields

$$\begin{aligned} \mathrm{d}S_{\mathrm{env,\,loc}}(q, p;\, \mathrm{d}t) = \sum_{i,j=0}^{\infty} \sigma_{ij}(q, p, \mathrm{d}t) \\ \times \int_{-\infty}^{\infty} \mathrm{d}q' \int_{-\infty}^{\infty} \mathrm{d}p'\, G_c(q', p'|q, p;\, \mathrm{d}t)(q' - q)^i (p' - p)^j \\ + \mathcal{O}(\mathrm{d}t^2) \end{aligned} \tag{5.55}$$

where the second line can be identified as the (i, j)-th (q, p) propagator moments respectively,

$$M_{ij,c}(q, p, \mathrm{d}t) = \int_{-\infty}^{\infty} \mathrm{d}q' \int_{-\infty}^{\infty} \mathrm{d}p'\, G_c(q', p'|q, p;\, \mathrm{d}t)(q' - q)^i (p' - p)^j \tag{5.56}$$

which allows rewriting $\mathrm{d}S_{\mathrm{env,\,loc}}$ as

$$\mathrm{d}S_{\mathrm{env,\,loc}}(q, p;\, \mathrm{d}t) = \sum_{i,j=0}^{\infty} \sigma_{ij}(q, p, \mathrm{d}t) M_{ij,c}(q, p, \mathrm{d}t) + \mathcal{O}(\mathrm{d}t^2) \tag{5.57}$$

where the expression is split up in two parts that easily combine into the local entropy production: σ_{ij} is obtained via the expansions of the known propagator quotients from the previous section using (5.54), and the rest of the moments $M_{ij,c}$.

In the case of the present model of the underdamped particle (3.34), the propagator moments can be calculated. Due to the length of the intermediate results, this calculation can be found in appendix A.2. The result obtained there is

$$M_{00;c} = 1 + \mathcal{O}(\mathrm{d}t) \tag{5.58}$$

$$M_{01;c} = (-p\gamma(p) - V'(q))\mathrm{d}t + \mathcal{O}(\mathrm{d}t^2) \tag{5.59}$$

$$M_{02;c} = \Gamma(p)^2 \mathrm{d}t + \mathcal{O}(\mathrm{d}t^2) \tag{5.60}$$

$$M_{ij;c} = \mathcal{O}(\mathrm{d}t^{i+1}) \quad (i, j > 0) \tag{5.61}$$

These expressions are noteworthy in a couple of ways:

1. The ambiguity parameter c does not appear in terms of leading order, indicating that the propagator moments are independent of said ambiguity at least in the present special case, where the fields are evaluated linearly between start and end points. (Recall that the most general version, introduced in section 3.3.3, allows evaluation at arbitrary points $\varphi(\boldsymbol{x}, \boldsymbol{x}')$, whereas here $(1-c)\boldsymbol{x} + c\boldsymbol{x}'$ has been chosen.)

2. None of the nonzero $\mathrm{d}q$ moments matter; they all contribute to higher-order in $\mathrm{d}t$ terms only.

3. The ad-hoc quantity ε of the equations of motion (3.34) vanishes on its own in the calculation, once again justifying its previous introduction.

Now $\mathrm{d}S_{\mathrm{env,\,loc}}$ can be calculated for the Spinney-Ford case with $a = 0$, the Hamiltonian Flow model and therefore also SFwith $a = \frac{1}{2}$, and remarkably the results are all identical, namely

$$\begin{aligned}\mathrm{d}S_{\mathrm{env,\,loc}}(q,p,\mathrm{d}t) = \Big(&-\mu(p) - p\mu'(p) + \frac{2p^2\mu(p)^2}{\Gamma(p)^2} + \frac{4p\mu(p)\Gamma'(p)}{\Gamma(p)} \\ &+ \Gamma'(p)^2 - \Gamma(p)\Gamma''(p)\Big)\,\mathrm{d}t + \mathcal{O}(\mathrm{d}t^2)\ .\end{aligned} \tag{5.62}$$

The terms proportional to $\mathrm{d}p^2$ in (4.51) compensate the other differences compared to (4.42) during integration (i.e. multiplication with the propagator moments) by moving them to higher orders of $\mathrm{d}t$, explaining how different differential entropies can yield the same local entropy production.

3.5.2 Global entropy production

The *global entropy production* once again averages over all positions, to give an expectation value for the short-time environmental entropy production in an ensemble of many (non-interacting) particles. It is obtained from the local entropy production developed just above via

$$\begin{aligned}\langle \mathrm{d}S_{\mathrm{env}}\rangle(t) &= \mathrm{d}t \int_\Omega \mathrm{d}\boldsymbol{x} P(\boldsymbol{x},t)\mathrm{d}S_{\mathrm{env,\,loc}}(\boldsymbol{x},\mathrm{d}t) \\ &= \mathrm{d}t \int_{-\infty}^{\infty}\mathrm{d}q \int_{-\infty}^{\infty}\mathrm{d}p\, P(q,p,t)\,\mathrm{d}S_{\mathrm{env,\,loc}}(q,p,\mathrm{d}t)\end{aligned} \tag{5.63}$$

where P is the probability density for being in a certain point of phase space.

For example, take a system in thermal equilibrium. Here, $\mu(p)$ and $\Gamma(p)$ are related via the generalized Einstein relation (3.31), and $P(q,p)$ is Boltzmannian.

The differential entropy productions then turn out to be

$$\mathrm{d}S_{\mathrm{env}}^{\mathrm{SF}} = -\beta p(\mathrm{d}p + V'(q)\mathrm{d}t) - \frac{\beta}{2}\Gamma(p)^2 + \mathcal{O}(\mathrm{d}t^2) + \mathcal{O}(\mathrm{d}p^3) \tag{5.64}$$

$$\mathrm{d}S_{\mathrm{env}}^{\mathrm{HF}} = -\beta p(\mathrm{d}p + V'(q)\mathrm{d}t) + \mathcal{O}(\mathrm{d}t^2) + \mathcal{O}(\mathrm{d}p^2) \tag{5.65}$$

which also showcases the previously claimed fact that SF's result does not vanish along the deterministic trajectory $\mathrm{d}p = -V'(q)\mathrm{d}t$, while the HF result does. Maybe more importantly even, SF's result does not reproduce the result previously obtained using only classical thermodynamics in (3.21). However, as claimed earlier, the resulting local entropy production is identical for both results,

$$\mathrm{d}S_{\mathrm{env,\,loc}}(q,p,\mathrm{d}t) = \left(-\tfrac{1}{2}\beta\Gamma(p)^2 + \tfrac{1}{2}\beta^2 p^2\Gamma(p)^2 - \beta p\Gamma(p)\Gamma'(p)\right)\,\mathrm{d}t + \mathcal{O}(\mathrm{d}t^2)\ . \tag{5.66}$$

By integrating this together with the Boltzmann distribution

$$P(q,p) = \frac{1}{Z}e^{-\beta\left(\frac{p^2}{2}-V(q)\right)} \tag{5.67}$$

in (5.63), the global entropy production can be obtained. As the explicit calculation in appendix A.3 shows, the result is

$$\langle \mathrm{d}S_{\mathrm{env}}\rangle(t) = 0\ \forall t \tag{5.68}$$

which is what a system obeying detailed balance in thermal equilibrium should obey. This is in agreement with classical thermodynamics again, and the necessity for the previous calculation also hints at why the expression for $\mathrm{d}S_{\mathrm{env}}$ in (3.21) is not entirely correct: although it assumes *some* properties equilibrium provides, it is a "thermodynamic" quantity in terms of q and p, while classical thermodynamics only deals with averages in which single appearances of location and momentum are meaningless.

3.6 Conclusion

Although the initial goal of the presented subject was finding a better definition of entropy production in continuous settings, it turned out that both the new (flow based) and the old (SF, [24]) approach yield identical results both for the local and the global environmental entropy production in the scenarios presented above. This leads to the conclusion that (two-point) **short-time environmental entropy production is not a uniquely defined quantity**, and should be regarded as a mathematical intermediate step. However, the flow-based approach seems to be closer to physical intuition, as it does not require a time reversal

(which is particularly difficult to map to the real world), its intermediate results vanish along the deterministic trajectory of the system, the distinction between odd and even variables is not necessary, and the propagator ambiguities fall away during the calculation without making further assumptions (which is important to mention because SF and HF models are identical in their two-point entropy production, given the *arbitrary* choice of $a = \frac{1}{2}$ in the former model).

It is worth noting that the case of a more general system than (3.17) has not been investiated, and it is unknown whether both path constructions still agree in this scenario. Furthermore, the SDE of the system has been integrated according to the Itō scheme in order to obtain the FP equation; the consequences of choosing Stratonovich or an arbitrary parameter value for the stochastic integration are also subject to further research.

A Appendix

A.1 General expressions for environmental entropy production

The general expressions obtained for SF and HF are different; despite this fact, they lead to identical results in the investigated scenarios. Here are the full expressions, with the ambiguities a and b left as free and independent parameters and differing terms boxed, for $\mathrm{d}S_{\mathrm{env}}$:

Spinney-Ford:

$$
\begin{aligned}
\mathrm{d}S_{\mathrm{env}}^{\mathrm{SF}} = a\Bigg(&-\frac{\mathrm{d}p^2 p\mu'(p)}{\Gamma(p)^2}+\frac{2\mathrm{d}p^2 p\mu(p)\Gamma'(p)}{\Gamma(p)^3}-\frac{\mathrm{d}p^2\mu(p)}{\Gamma(p)^2}+\frac{2\mathrm{d}p^2\Gamma'(p)V'(q)}{\Gamma(p)^3}-\frac{3\mathrm{d}p\Gamma'(p)}{\Gamma(p)}\\
&+\mathrm{d}tp\mu'(p)-\frac{2\mathrm{d}tp\mu(p)\Gamma'(p)}{\Gamma(p)}+\mathrm{d}t\mu(p)-\frac{2\mathrm{d}t\Gamma'(p)V'(q)}{\Gamma(p)}\Bigg)\\
+b\Bigg(&\frac{\mathrm{d}p^2 p\mu'(p)}{\Gamma(p)^2}-\frac{2\mathrm{d}p^2 p\mu(p)\Gamma'(p)}{\Gamma(p)^3}+\frac{\mathrm{d}p^2\mu(p)}{\Gamma(p)^2}-\frac{3\mathrm{d}p^2\Gamma''(p)}{\Gamma(p)}+\frac{3\mathrm{d}p^2\Gamma'(p)^2}{\Gamma(p)^2}\\
&\boxed{+\frac{2\mathrm{d}p^2\Gamma'(p)V'(q)}{\Gamma(p)^3}}-\frac{3\mathrm{d}p\Gamma'(p)}{\Gamma(p)}-\mathrm{d}tp\mu'(p)+\frac{2\mathrm{d}tp\mu(p)\Gamma'(p)}{\Gamma(p)}\\
&-\mathrm{d}t\mu(p)\boxed{-\frac{2\mathrm{d}t\Gamma'(p)V'(q)}{\Gamma(p)}}\Bigg)\\
+(a^2-b^2)\Bigg(&-\frac{5\mathrm{d}p^2\Gamma''(p)}{2\Gamma(p)}+\frac{5\mathrm{d}p^2\Gamma'(p)^2}{2\Gamma(p)^2}+\mathrm{d}t\Gamma(p)\Gamma''(p)-\mathrm{d}t\Gamma'(p)^2\Bigg)\\
&-\frac{\mathrm{d}p^2 p\mu'(p)}{\Gamma(p)^2}+\frac{2\mathrm{d}p^2 p\mu(p)\Gamma'(p)}{\Gamma(p)^3}-\frac{\mathrm{d}p^2\mu(p)}{\Gamma(p)^2}+\frac{\mathrm{d}p^2\Gamma''(p)}{2\Gamma(p)}\\
&-\frac{\mathrm{d}p^2\Gamma'(p)^2}{2\Gamma(p)^2}\boxed{-\frac{2\mathrm{d}p^2\Gamma'(p)V'(q)}{\Gamma(p)^3}}-\frac{2\mathrm{d}ppp\mu(p)}{\Gamma(p)^2}+\frac{\mathrm{d}p\Gamma'(p)}{\Gamma(p)}-\frac{2\mathrm{d}tp\mu(p)V'(q)}{\Gamma(p)^2}
\end{aligned}
$$

Flow model:

$$
\begin{aligned}
\mathrm{d}S_{\mathrm{env}}^{\mathrm{HF}} = a\Big(&-\frac{\mathrm{d}p^2 p\mu'(p)}{\Gamma(p)^2} + \frac{2\mathrm{d}p^2 p\mu(p)\Gamma'(p)}{\Gamma(p)^3} - \frac{\mathrm{d}p^2\mu(p)}{\Gamma(p)^2} + \frac{2\mathrm{d}p^2\Gamma'(p)V'(q)}{\Gamma(p)^3} - \frac{3\mathrm{d}p\Gamma'(p)}{\Gamma(p)} \\
&+ \mathrm{d}t p\mu'(p) - \frac{2\mathrm{d}t p\mu(p)\Gamma'(p)}{\Gamma(p)} + \mathrm{d}t\mu(p) - \frac{2\mathrm{d}t\Gamma'(p)V'(q)}{\Gamma(p)}\Big) \\
+ b\Big(&\frac{\mathrm{d}p^2 p\mu'(p)}{\Gamma(p)^2} - \frac{2\mathrm{d}p^2 p\mu(p)\Gamma'(p)}{\Gamma(p)^3} + \frac{\mathrm{d}p^2\mu(p)}{\Gamma(p)^2} - \frac{3\mathrm{d}p^2\Gamma''(p)}{\Gamma(p)} + \frac{3\mathrm{d}p^2\Gamma'(p)^2}{\Gamma(p)^2} \\
&\boxed{+\frac{4\mathrm{d}p^2\Gamma'(p)V'(q)}{\Gamma(p)^3}} - \frac{3\mathrm{d}p\Gamma'(p)}{\Gamma(p)} - \mathrm{d}t p\mu'(p) + \frac{2\mathrm{d}t p\mu(p)\Gamma'(p)}{\Gamma(p)} \\
&- \mathrm{d}t\mu(p)\boxed{-\frac{4\mathrm{d}t\Gamma'(p)V'(q)}{\Gamma(p)}}\Big) \\
+ (a^2 - b^2)\Big(&-\frac{5\mathrm{d}p^2\Gamma''(p)}{2\Gamma(p)} + \frac{5\mathrm{d}p^2\Gamma'(p)^2}{2\Gamma(p)^2} + \mathrm{d}t\Gamma(p)\Gamma''(p) - \mathrm{d}t\Gamma'(p)^2\Big) \\
&- \frac{\mathrm{d}p^2 p\mu'(p)}{\Gamma(p)^2} + \frac{2\mathrm{d}p^2 p\mu(p)\Gamma'(p)}{\Gamma(p)^3} - \frac{\mathrm{d}p^2\mu(p)}{\Gamma(p)^2} + \frac{\mathrm{d}p^2\Gamma''(p)}{2\Gamma(p)} \\
&- \frac{\mathrm{d}p^2\Gamma'(p)^2}{2\Gamma(p)^2}\boxed{-\frac{3\mathrm{d}p^2\Gamma'(p)V'(q)}{\Gamma(p)^3}} - \frac{2\mathrm{d}p p\mu(p)}{\Gamma(p)^2} + \frac{\mathrm{d}p\Gamma'(p)}{\Gamma(p)} - \frac{2\mathrm{d}t p\mu(p)V'(q)}{\Gamma(p)^2} \\
&\boxed{+\frac{\mathrm{d}t\Gamma'(p)V'(q)}{\Gamma(p)}}
\end{aligned}
$$

A.2 Calculation of Fokker-Planck propagator moments

The Fokker-Planck (FP) propagator has the moments

$$
M_{ij;c} = \int_{-\infty}^{\infty} \mathrm{d}(\mathrm{d}q) \int_{-\infty}^{\infty} \mathrm{d}(\mathrm{d}p) G_c(q + \mathrm{d}q, p + \mathrm{d}p | q, p; \mathrm{d}t) \mathrm{d}q^i \mathrm{d}p^j \ . \tag{A.1}
$$

The following describes how to calculate these up to relevant orders. For the system discussed in this thesis (3.34) [22], G_c takes the form

$$
\begin{aligned}
&G_c(q,p,\mathrm{d}q,\mathrm{d}p,\mathrm{d}t) = \\
&\frac{1}{2\pi \mathrm{d}t\varepsilon\Gamma(c\mathrm{d}p+p)}\ \exp\Bigg(\frac{1}{2\mathrm{d}t} \\
&\quad\times\Bigg(-\frac{(\mathrm{d}q-\mathrm{d}t(c\mathrm{d}p+p))^2}{\varepsilon^2} \\
&\qquad+\frac{(\mathrm{d}t(c\mathrm{d}p+p)\gamma(c\mathrm{d}p+p)+2c\mathrm{d}t\Gamma(c\mathrm{d}p+p)\Gamma'(c\mathrm{d}p+p)+\mathrm{d}p+\mathrm{d}tv'(c\mathrm{d}q+q))^2}{\Gamma(c\mathrm{d}p+p)^2} \\
&\qquad+2\mathrm{d}t^2c\left(-(c\mathrm{d}p+p)\gamma'(c\mathrm{d}p+p)-\gamma(c\mathrm{d}p+p)\right) \\
&\qquad-2\mathrm{d}t^2c^2\left(\Gamma(c\mathrm{d}p+p)\Gamma''(c\mathrm{d}p+p)+\Gamma'(c\mathrm{d}p+p)^2\right)\Bigg)\Bigg)
\end{aligned}
\tag{A.2}
$$

A.2.1 i-th dq moment

Calculation of the i-th $\mathrm{d}q$ moment is a straightforward application of the standard method of steepest descent

$$
\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty}\mathrm{d}q f(q)e^{\frac{1}{\varepsilon}g(q)} = \lim_{\varepsilon\to 0} f(q_{\max})e^{\varepsilon g(q_{\max})}\sqrt{\frac{2\pi}{-\varepsilon g''(q_{\max})}} \tag{A.3}
$$

where $q_{\max}$ denotes the extremum of $g(q)$.

Here the "small" quantity is ε, which appears only in the prefactor and the first summand in the exponential in (A.2); $q_{\max}$ can be seen to be $\mathrm{d}t(c\,\mathrm{d}p+p)$ from the propagator's exponential term. Steepest descent then, using $\tilde{p}_c = c\,\mathrm{d}p + p$ as a shorthand, yields

$$
\begin{aligned}
&\frac{\mathrm{d}t^{i-\frac{1}{2}}\tilde{p}_c^{\,i}}{\sqrt{2\pi}\Gamma(\tilde{p}_c)}\exp\Bigg(\\
&\qquad\mathrm{d}tc\left(\tilde{p}_c\gamma'(\tilde{p}_c)+\gamma(\tilde{p}_c)\right)+\mathrm{d}tc^2\left(\Gamma(\tilde{p}_c)\Gamma''(\tilde{p}_c)+\Gamma'(\tilde{p}_c)^2\right) \\
&\qquad-\frac{(\mathrm{d}t\tilde{p}_c\gamma(\tilde{p}_c)+2\mathrm{d}tc\Gamma(\tilde{p}_c)\Gamma'(\tilde{p}_c)+\mathrm{d}tv'(c\tilde{p}_c\mathrm{d}t+q)+\mathrm{d}p)^2}{2\mathrm{d}t\Gamma(\tilde{p}_c)^2}\Bigg)
\end{aligned}
\tag{A.4}
$$

as the intermediate result for (A.1) after the $\mathrm{d}(\mathrm{d}q)$ integration. (The ε fall away naturally without even taking the limit $\varepsilon \to 0$.)

A.2.2 j-th dp moment

For the $\mathrm{d}p$ moments the approach is similar, but a little more complicated:

1. The standard formula for steepest descent assumes that the exponential's prefactor is independent of the "small" quantity ($\mathrm{d}t$ in this case). This is not the case for the previously obtained result after $\mathrm{d}(\mathrm{d}q)$ integration.

2. Terms quickly grow with increasing order. While it must be made sure that no significant terms are omitted, calculating "too much" and dropping unnecessary terms afterwards introduces a lot of work and potential oversights. For that reason one should very carefully keep track of the orders of the terms to be calculated; the criteria for this turn out to be quite different in each order.

The expression of interest is

$$I_k(\mathrm{d}t) = \int_{-\infty}^{\infty} \mathrm{d}(\mathrm{d}p)\, \mathrm{d}p^k f_{\mathrm{d}t}(\mathrm{d}p)\, e^{\frac{1}{\mathrm{d}t} g(\mathrm{d}p)} \tag{A.5}$$

where the integrand "$f e^{g/\mathrm{d}t}$" is given by (A.4), which as you recall is the already integrated part of (A.1); note the potential $\mathrm{d}t$ dependency of f denoted by an index. By substitution $\mathrm{d}p \to \mathrm{d}p'\sqrt{\mathrm{d}t}$ and following that Taylor expansion of f and g around $\mathrm{d}p'\sqrt{\mathrm{d}t} = \underline{\mathrm{d}p'}\sqrt{\mathrm{d}t}$, this can be transformed into

$$\begin{aligned} I_k(\mathrm{d}t) = \sqrt{\mathrm{d}t}^{k+1} \exp\left(\frac{1}{\mathrm{d}t} g\left(\underline{\mathrm{d}p'}\sqrt{\mathrm{d}t}\right)\right) \int_{-\infty}^{\infty} \mathrm{d}(\mathrm{d}p')\mathrm{d}p'^k \\ \times \exp\left(\frac{\mathrm{d}p'^2}{2} g''(\underline{\mathrm{d}p'}\sqrt{\mathrm{d}t})\right) \exp\left(\frac{\mathrm{d}p'^3\sqrt{\mathrm{d}t}}{2} g'''(\underline{\mathrm{d}p'}\sqrt{\mathrm{d}t})\right) \left(1 + \mathcal{O}(\mathrm{d}p'^4\mathrm{d}t^1)\right) \\ \times \left(f_{\mathrm{d}t}(\underline{\mathrm{d}p'}\sqrt{\mathrm{d}t}) + f'_{\mathrm{d}t}(\underline{\mathrm{d}p'}\sqrt{\mathrm{d}t})\mathrm{d}p'\sqrt{\mathrm{d}t} + f''_{\mathrm{d}t}(\underline{\mathrm{d}p'}\sqrt{\mathrm{d}t})\frac{\mathrm{d}p^2\mathrm{d}t}{2} + \mathcal{O}(\mathrm{d}p'^3\mathrm{d}t^{3/2})\right) \end{aligned} \tag{A.6}$$

This intermediate result resembles a corresponding expression in the method of steepest descent: suppose f does not have an implicit $\mathrm{d}t$ dependence (in the present notation: does not depend on the index), then the integral would simply be Gaussian in the limit $\mathrm{d}t \to 0$, standard formula (A.3) would be recovered. Compared to the $\mathrm{d}q$ case, $\mathrm{d}p_{\max}$ could not be easier, as it turns out to be 0.

The following integral will prove handy later ($a < 0$):

$$J_n = \int_{-\infty}^{\infty} \mathrm{d}x\, x^n e^{\frac{ax^2}{2}} = \begin{cases} \sqrt{\frac{2^{n-1}}{-a^{n+1}}}\Gamma\left(\frac{n+1}{2}\right) & n \text{ even} \\ 0 & n \text{ odd} \end{cases} \tag{A.7}$$

(derivation: substitute the exponent by a linear expression, leading to the usual Γ integral representation.) The first couple of values are

$$J_0 = \sqrt{\frac{2\pi}{-a}} \qquad J_2 = \sqrt{\frac{2\pi}{-a^3}} \qquad J_4 = \sqrt{\frac{18\pi}{-a^5}}\ . \tag{A.8}$$

Zeroth dp moment

The zeroth moment appears in the context of σ_{00}, which is accompanied by $\mathrm{d}t$ implicitly (due to the shape of the propagator quotients' expansions, namely (4.42)/SF and (4.51)/HF) in (5.57). Therefore, in order to obtain an overall result of up to $\mathcal{O}(\mathrm{d}t^1)$, it is sufficient to calculate only $\mathcal{O}(\mathrm{d}t^0)$ here.

$$\begin{aligned} I_0(\mathrm{d}t) &= \sqrt{\mathrm{d}t}\exp\left(\frac{1}{\mathrm{d}t}g(0)\right)\int_{-\infty}^{\infty}\mathrm{d}(\mathrm{d}p')\exp\left(\frac{\mathrm{d}p'^2}{2}g''(0)\right)\left(1+\mathcal{O}(\mathrm{d}p'^3\mathrm{d}t^{1/2})\right) \\ &\quad\times\left(f(0)+\cancel{f'(0)\mathrm{d}p'\sqrt{\mathrm{d}t}}+f''(0)\frac{\mathrm{d}p'^2\mathrm{d}t}{2}+\mathcal{O}(\mathrm{d}p'^3\mathrm{d}t^{3/2})\right) \\ &= f(0)e^{\frac{g(0)}{\mathrm{d}t}}\sqrt{\frac{2\pi\mathrm{d}t}{-g''(0)}}+\mathcal{O}(\mathrm{d}t^1) \end{aligned}$$

Inserting the expressions for f and g (obtained by comparing (A.5) with (A.4)) and using (A.7) to calculate/cancel out the integrals, then results in

$$M_{i0,c} = \mathrm{d}t^i p^i + \mathcal{O}(\mathrm{d}t^{i+1})\ \forall c \tag{A.9}$$

First dp moment

Unlike the zeroth order, the $\mathrm{d}p^1$ moment does not have an implicit $\mathrm{d}t$ prefactor in the entropy formula; expansion therefore has to be up to $\mathcal{O}(\mathrm{d}t^1)$ here. The surprising property of this moment is that the Gaussian exponential is *not* the only exponential contributing to the result up to the desired order: the term $\exp(\propto \mathrm{d}p'^3)$ needs to be considered as well, otherwise the result additionally contains an incorrect c-dependent expression. Using (A.7) again allows us to calculate the integrals, leading to

$$\begin{aligned} I_1(\mathrm{d}t) &= \mathrm{d}t\exp\left(\frac{1}{\mathrm{d}t}g(0)\right)\int_{-\infty}^{\infty}\mathrm{d}(\mathrm{d}p')\mathrm{d}p' \\ &\quad\times\exp\left(\frac{\mathrm{d}p'^2}{2}g''(0)\right)\exp\left(\frac{\mathrm{d}p'^3\sqrt{\mathrm{d}t}}{2}g'''(0)\right)\left(1+\mathcal{O}(\mathrm{d}p'^4\mathrm{d}t^1)\right) \\ &\quad\times\left(f(0)+f'(0)\mathrm{d}p'\sqrt{\mathrm{d}t}+\mathcal{O}(\mathrm{d}p'^2\mathrm{d}t^1)\right) \\ &= f'(0)e^{\frac{g(0)}{\mathrm{d}t}}\sqrt{\frac{2\pi\mathrm{d}t^3}{-g''(0)^3}}+\frac{f(0)g'''(0)}{6}e^{\frac{g(0)}{\mathrm{d}t}}\sqrt{\frac{3^2 2\pi\mathrm{d}t^3}{-g''(0)^5}}+\mathcal{O}(\mathrm{d}t^2) \end{aligned}$$

so the moment turns out to be

$$M_{i1,c} = -\mathrm{d}t^{i+1}\left(p^{i+1}\gamma(p)+p^i V'(q)-p^{i-1}ci\Gamma(p)^2\right)+\mathcal{O}(\mathrm{d}t^2)\ . \tag{A.10}$$

The only case relevant to entropy production is $i = 0$, as it is the only one containing a term of order $\mathrm{d}t^1$. This also gets rid of the odd term involving $c \cdot i$.

Second dp moment

The highest moment is, maybe surprisingly, the easiest one of the three required ones: all the complicated terms are of order higher than $\mathcal{O}(\mathrm{d}t^1)$, and the only significant expression left over is a single square root term:

$$\begin{aligned} I_2(\mathrm{d}t) &= \mathrm{d}t^{3/2} \exp\left(\frac{1}{\mathrm{d}t} g(0)\right) \int_{-\infty}^{\infty} \mathrm{d}(\mathrm{d}p') \exp\left(\frac{\mathrm{d}p'^2}{2} g''(0)\right) \left(1 + \mathcal{O}(\mathrm{d}p'^3 \mathrm{d}t^{1/2})\right) \\ &\quad \times \left(f(0) + \cancel{f'(0) \mathrm{d}p' \sqrt{\mathrm{d}t}} + f''(0) \frac{\mathrm{d}p'^2 \mathrm{d}t}{2} + \mathcal{O}(\mathrm{d}p'^3 \mathrm{d}t^{3/2})\right) \\ &= f(0) e^{\frac{g(0)}{\mathrm{d}t}} \sqrt{\frac{2\pi \mathrm{d}t^3}{-g''(0)^3}} + \mathcal{O}(\mathrm{d}t^2) \end{aligned}$$

and one last time (A.7) is used for integration,

$$M_{i2,c} = \mathrm{d}t^{i+1} p^i \Gamma(p)^2 + \mathcal{O}(\mathrm{d}t^2) \ . \tag{A.11}$$

Higher-order dp moments

Since terms of higher than $\mathrm{d}t^1$ aren't interesting for short-time entropy production, and keeping in mind that $f_{\mathrm{d}t}$ involves a factor of $\mathrm{d}t^{-1/2}$, it can be seen that for $k > 2$ (A.6) will always yield a result of negligible order. Therefore, the $\mathrm{d}p$ moments of power greater than 2 do not appear in short-time entropy and need not be considered, as it is the case for $\mathrm{d}q$ moments of order greater than 0.

Summary of the moments

Taking the previous sections together, it can be seen that the moments are as follows:

$$\begin{aligned} M_{00;c} &= 1 + \mathcal{O}(\mathrm{d}t) \\ M_{01;c} &= (-p\gamma(p) - V'(q))\mathrm{d}t + \mathcal{O}(\mathrm{d}t^2) \\ M_{02;c} &= \Gamma(p)^2 \mathrm{d}t + \mathcal{O}(\mathrm{d}t^2) \\ M_{ij;c} &= \mathcal{O}(\mathrm{d}t^{i+1}) \quad \text{else} \end{aligned}$$

as claimed in the main text.

A.3 Calculation of the global entropy production

To calculate the global entropy production in section 3.5.2, the integral in question is

$$\langle \mathrm{d}S_{\mathrm{env}} \rangle(t) = \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}q \int_{-\infty}^{\infty} \mathrm{d}p \, P(q,p,t) \, \mathrm{d}S_{\mathrm{env,\,loc}}(q,p,\mathrm{d}t) \tag{A.12}$$

with

$$P(q,p,t) = \tfrac{1}{Z} e^{-\beta\left(\frac{p^2}{2} - V(q)\right)} \tag{A.13}$$

$$\mathrm{d}S_{\mathrm{env,\,loc}}(q,p,\mathrm{d}t) = \left(-\tfrac{1}{2}\beta\Gamma(p)^2 + \tfrac{1}{2}\beta^2 p^2 \Gamma(p)^2 - \beta p \Gamma(p)\Gamma'(p)\right) \mathrm{d}t \tag{A.14}$$

where terms of higher than linear order have been dropped. Noting that

$$-\tfrac{1}{2}\beta\Gamma(p)^2 + \tfrac{1}{2}\beta^2 p^2 \Gamma(p)^2 - \beta p \Gamma(p)\Gamma'(p) \tag{A.15}$$

$$= \tfrac{1}{2}\beta^2 p^2 \Gamma(p)^2 - \frac{\beta}{2}\left(\Gamma^2 + p\Gamma(p)\Gamma'(p)\right) \tag{A.16}$$

$$= \underbrace{\tfrac{1}{2}\beta^2 p^2 \Gamma(p)^2}_{\equiv A} - \underbrace{\frac{\beta}{2}\frac{\mathrm{d}}{\mathrm{d}p}\left(p\Gamma(p)^2\right)}_{\equiv B} \tag{A.17}$$

the integral can be divided up in two contributions over A and B; then

$$\begin{aligned} I_B &= \frac{\beta}{2Z} \int_{-\infty}^{\infty} \mathrm{d}p e^{-\beta\left(\frac{p^2}{2} - V(q)\right)} \frac{\mathrm{d}}{\mathrm{d}p}\left(p\Gamma(p)^2\right) \\ &= \frac{\beta}{2Z}\left(\left[-\beta p e^{-\beta\left(\frac{p^2}{2} - V(q)\right)}\right]_{p=-\infty}^{\infty} - \int_{-\infty}^{\infty} \mathrm{d}p \, (-\beta p) e^{-\beta\left(\frac{p^2}{2} - V(q)\right)} p\Gamma(p)^2\right) \\ &= \frac{\beta^2}{2Z} \int_{-\infty}^{\infty} \mathrm{d}p e^{-\beta\left(\frac{p^2}{2} - V(q)\right)} p^2 \Gamma^2 \\ &= I_A \end{aligned} \tag{A.18}$$

so both terms cancel out (before the $\mathrm{d}q$ integration is even necessary), resulting in

$$\langle \mathrm{d}S_{\mathrm{env}} \rangle(t) = 0 \;\forall t \tag{A.19}$$

as claimed in (5.68).

A.4 Abbreviations

DE	Differential equation
FP	Fokker-Planck
HF	Hamiltonian-Flow-based model for environmental entropy production
SDE	Stochastic Differential equation
SF	Spinney and Ford, authors of [24]

Thanks

This thesis would not have been possible without the help of others. I would like to thank Prof. Dr. Wolfgang Kinzel for making me part of the group of theoretical physics 3, making the research possible in the first place.

My advisor Prof. Dr. Haye Hinrichsen provided the topic along with many important discoveries/ideas (most notably it was him who found the histogram symmetry that lead to the research in chapter 2, and also finding the continuous version of entropy having implausible results – the motivation behind section 3), cross-checking results, constructively judging methods.

Prof. Dr. Joel Lebowitz' seminar over my one year stay at Rutgers University was the first research-level contact with the area of statistical physics, and is what made me consider taking this path.

Last but not least I was part of a group of people that will be missed; in the age of internet crawlers they shall be named explicitly (ordered by office location): Dr. Jörg Reichardt, Steffen Zeeb, Peter Janotta, Dr. Otti d'Huys, Prof. Dr. Georg Reents, Marco Winkler, Barbara Schlögl and Peter Reisenauer.

Bibliography

[1] D. Luposchainsky, A. Barato, H. Hinrichsen
Phys. Rev. E **87**, 042108 (2013)
Strong fluctuation theorem for nonstationary nonequilibrium systems

[2] U. Seifert
Rep. Prog. Phys. **75**, 126001 (2012)
Stochastic thermodynamics, fluctuation theorems and molecular machines

[3] U. Seifert
Phys. Rev. Lett. **95**, 040602 (2005)
Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem

[4] G. Crooks
Phys. Rev. E **60**, 2721-2726 (1999)
Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences

[5] M. Esposito
Phys. Rev. Lett. **104**, 090601 (2010)
Three Detailed Fluctuation Theorems

[6] D. Evans, E. Cohen, G. Morriss
Phys. Rev. Lett. **71**, 2401-2404 (1993)
Probability of second law violations in shearing steady states

[7] G. Gallavotti, E. Cohen
Phys. Rev. Lett. **74**, 2694-2697 (1995)
Dynamical Ensembles in Nonequilibrium Statistical Mechanics

[8] D. Evans, D. Searles
Phys. Rev. E **50**, 1645-1648 (1994)
Equilibrium microstates which generate second law violating steady states

[9] J. Kurchan
J. Phys. A: Math. Gen. **31** 3719 (1998)
Fluctuation theorem for stochastic dynamics

[10] J. Kurchan
J. Stat. Mech. (2007) P07005
Non-equilibrium work relations

[11] J. Lebowitz, H. Spohn
J. Stat. Phys. **95**, 333-365 (1999)
A Gallavotti-Cohen-Type Symmetry in the Large Deviation Functional for Stochastic Dynamics

[12] C. Jarzynski
Phys. Rev. Lett. **78**, 2690-2693 (1997)
Nonequilibrium Equality for Free Energy Differences

[13] G. Crooks
Phys. Rev. E **60**, 2721-2726 (1999)
Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences

[14] G. Crooks
Phys. Rev. E **61**, 2361-2366 (2000)
Path-ensemble averages in systems driven far from equilibrium

[15] C. Maes
J. Stat. Phys. **95**, 367-392 (1999)
The Fluctuation Theorem as a Gibbs Property

[16] T. Hatano, S. Sasa
Phys. Rev. Lett. **86**, 3463-3466 (2001)
Steady-State Thermodynamics of Langevin Systems

[17] D. Andrieux, P. Gaspard
J. Stat. Phys. **127**, 107-131 (2007)
Fluctuation Theorem for Currents and Schnakenberg Network Theory

[18] R. Harris, G. Schütz
J. Stat. Mech. (2007) P07020
Fluctuation theorems for stochastic dynamics

[19] J. Schnakenberg
Rev. Mod. Phys **48**, 571-585 (1976)

Network theory of microscopic and macroscopic behavior of master equation systems

[20] J. Leeuwen, H. Hilhorst
Physica A **107**, 319-329 (1981)
Pinning of a rough interface by an external potential

[21] H. Hinrichsen, R. Livi, D. Mukamel, A. Politi
Phys. Rev. E **68**, 041606 (2003)
Wetting under nonequilibrium conditions

[22] D. Luposchainsky and H. Hinrichsen
J. Stat. Phys. **153**, 828-841 (2013)
Entropy production in continuous phase space systems

[23] H. Lee, C. Kwon, H. Park
arXiv:1209.0543 [cond-mat.stat-mech]
Fluctuation theorems in general stochastic processes with odd-parity variables

[24] R. Spinney and I. Ford
Phys. Rev. E **85**, 051113 (2012)
Entropy production in full phase space for continuous stochastic dynamics

[25] R. Spinney and I. Ford
Phys. Rev. Lett. **108**, 170603 (2012)
Nonequilibrium Thermodynamics of Stochastic Systems with Odd and Even Variables

[26] R. Spinney and I. Ford
Fluctuation Relations: A Pedagogical Overview, in *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*
Editors: R. Klages, W. Just and C. Jarzynski
Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany (2013)

[27] A. Sengupta
Summer term 2012 at Rutgers/USA
Lecture notes on theoretical biophysics

[28] L. Evans
Lecture notes on mathematics
An Introduction to Stochastic Differential Equations
`http://math.berkeley.edu/ evans/SDE.course.pdf`
Retrieved December 2012 (version 1.2)

[29] C. Wissel
Z. Phys. B **35**, 185 (1979)
Manifolds of equivalent path integral solutions of the Fokker-Planck equation

[30] J. Donoso, J. Salgado, M. Soler
J. Phys. A: Math. Gen. **32** (1999) 3681-3695
Short-time propagators for nonlinear Fokker-Planck equations

Eidesstattliche Erklärung

Ich bin der alleinige Autor des Inhalts dieser Arbeit. Zur ihrer Erstellung wurden keine außer der angegebenen Quellen und Hilfsmittel benutzt. Die Arbeit wurde weder bisher noch gleichzeitig einer anderen Prüfungsbehörde zur Erlangung eines akademischen Grades vorgelegt.

David Luposchainsky
Würzburg, 14. 11. 2013