

# Foundations of Stochastic Thermodynamics

Entropy, Dissipation and Information in Models of Small Systems

Dissertation  
for the award of the degree  
“Doctor rerum naturalium”  
of the Georg–August–Universität Göttingen

in the programme  
International Max Planck Research School  
“Physics of Biological and Complex Systems”  
(IMPRS PBCS)  
of the  
“Göttingen Graduate School for Neurosciences,  
Biophysics, and Molecular Biosciences”  
(GGNB)

submitted by  
**Bernhard Altaner**  
from Ulm

**Göttingen, 2014**

## About this dissertation

### Thesis Committee

Prof. Dr. Jürgen Vollmer, *Supervisor* <sup>1,2</sup>

Prof. Dr. Marc Timme, *Co-supervisor* <sup>3,2</sup>

Prof. Dr. Eberhard Bodenschatz <sup>4,2</sup>

### Thesis referees

Prof. Dr. Jürgen Vollmer <sup>1,2</sup>

Prof. Dr. Stefan Kehrein <sup>5</sup>

### Examination committee

Prof. Dr. Jürgen Vollmer <sup>1,2</sup>

Prof. Dr. Marc Timme <sup>3,2</sup>

Prof. Dr. Stefan Kehrein <sup>5</sup>

Prof. Dr. Stephan Herminghaus <sup>1,2</sup>

Dr. Marco G. Mazza <sup>1</sup>

### Examination date

July, 31st 2014

---

<sup>1</sup>Department “Dynamics of Complex Fluids”, Max Planck Institute for Dynamics and Self-Organization, Göttingen

<sup>2</sup>Institute for Nonlinear Dynamics, Faculty of Physics, Georg-August-Universität Göttingen

<sup>3</sup>Independent Research Group “Network Dynamics”, Max Planck Institute for Dynamics and Self-Organization, Göttingen

<sup>4</sup>Department “Fluid Dynamics, Pattern Formation and Nanobiocomplexity”, Max Planck Institute for Dynamics and Self-Organization, Göttingen

<sup>5</sup>Condensed Matter Theory, Institute for Theoretical Physics, Faculty of Physics, Georg-August-Universität Göttingen

# Foundations of Stochastic Thermodynamics

Entropy, Dissipation and Information in Models of Small Systems

**Bernhard Altaner**

## **Abstract**

Small systems in a thermodynamic medium — like colloids in a suspension or the molecular machinery in living cells — are strongly affected by the thermal fluctuations of their environment. Physicists model such systems by means of stochastic processes. Stochastic Thermodynamics (ST) defines entropy changes and other thermodynamic notions for individual realizations of such processes. It applies to situations far from equilibrium and provides a unified approach to stochastic fluctuation relations. Its predictions have been studied and verified experimentally.

This thesis addresses the theoretical foundations of ST. Its focus is on the following two aspects: (i) The stochastic nature of mesoscopic observations has its origin in the molecular chaos on the microscopic level. Can one derive ST from an underlying reversible deterministic dynamics? Can we interpret ST's notions of entropy and entropy changes in a well-defined information-theoretical framework? (ii) Markovian jump processes on finite state spaces are common models for bio-chemical pathways. How does one quantify and calculate fluctuations of physical observables in such models? What role does the topology of the network of states play? How can we apply our abstract results to the design of models for molecular motors?

The thesis concludes with an outlook on dissipation as information written to unobserved degrees of freedom — a perspective that yields a consistency criterion between dynamical models formulated on various levels of description.

**Göttingen, 2014**



# Contents

<b>1. Introduction</b>	<b>9</b>
1.1. Motivation	9
1.1.1. Stochastic thermodynamics	9
1.1.2. Foundations	11
1.1.3. Entropy, dissipation and information	13
1.1.4. Research questions	16
1.2. The structure of this thesis	17
1.2.1. How to read this thesis	17
1.2.2. Outline	17
1.2.3. Notation, abbreviations and conventions	19
<b>2. Notions of entropy and entropy production</b>	<b>21</b>
2.1. Entropy in classical thermodynamics	22
2.2. Entropy as information or uncertainty	23
2.3. Statistical physics and the distinction between system and medium	25
2.3.1. The second law in statistical physics	26
2.3.2. Entropy changes in statistical physics	27
2.3.3. Hamiltonian dynamics	28
2.4. The models of stochastic thermodynamics	31
2.4.1. Langevin and Fokker–Planck equations	31
2.4.2. Master equations	37
2.4.3. Stochastic fluctuation relations	43
2.5. Effective deterministic models for molecular dynamics	44
2.5.1. Thermostated equations of motion	45
2.5.2. Entropy and dissipation in deterministic dynamics	48
2.5.3. Stroboscopic maps and time-discrete dynamics	51
2.5.4. Reversibility and deterministic fluctuation relations	52
2.6. Measurable dynamical systems	54
2.6.1. Mathematical prerequisites	55
2.6.2. Measurable and topological dynamical systems	56
2.6.3. Topological and measure-theoretic entropy	58
2.7. Summary	60

<b>3. Markovian symbolic dynamics</b>	<b>63</b>
3.1. Symbolic stochastic dynamics	64
3.1.1. Stochastic processes	65
3.1.2. The stochastic process of observed time series	68
3.1.3. Symbolic dynamics	69
3.2. Observables and partitions on phase space	71
3.2.1. Generating partitions	73
3.2.2. Topological and Markov partitions	75
3.3. Markov measures	77
3.3.1. Markov measures for semi-infinite sequences	77
3.3.2. Markov measures for bi-infinite sequences	78
3.3.3. (Natural) Markov measures on phase space	80
3.4. Discussion	82
3.4.1. Connection to ergodic theory	83
3.4.2. Operational interpretation	85
3.5. Summary	88
<b>4. An information-theoretical approach to stochastic thermodynamics</b>	<b>91</b>
4.1. A general information-theoretic framework	92
4.1.1. The measurement process revisited	92
4.1.2. Fine- and coarse-grained entropy	94
4.1.3. The fundamental and derived entropic $\tau$ -chains	97
4.1.4. Temporal variations	100
4.2. Network multibaker maps	102
4.2.1. Formulation of the model	103
4.2.2. Reversibility and further constraints	104
4.2.3. NMBM observables, priors and initial conditions	106
4.2.4. Evolution of the densities	107
4.2.5. The entropic $\tau$ -chains and their variations	108
4.3. Discussion	110
4.3.1. Consistent identification of system and medium entropy	110
4.3.2. Positivity of the variation of the total entropy	111
4.3.3. Foundations of Markovian stochastic thermodynamics	112
4.3.4. Influence of the reference measure	115
4.4. Summary	116
<b>5. The structure of Markov jump processes</b>	<b>119</b>
5.1. Kirchhoff's laws and an electrical analogy	120
5.1.1. Steady states and Kirchhoff's current law	121
5.1.2. Kirchhoff's second law and an electrical analogy	122
5.2. Cycles and trees as the fundamental building blocks of networks	124
5.2.1. Anti-symmetric observables on the edges	124

5.2.2. Algebraic graph theory . . . . .	125
5.2.3. Trees and chords . . . . .	127
5.3. Quantification of fluctuations of physical observables . . . . .	130
5.3.1. Cumulants of random variables . . . . .	130
5.3.2. Asymptotic properties in stochastic processes . . . . .	131
5.3.3. Large deviation theory of Markovian jump processes . . . . .	133
5.3.4. Cycles and fluctuations of physical observables . . . . .	134
5.4. Discussion . . . . .	137
5.4.1. Different cycle decompositions and the flux-cycle transform . . . . .	137
5.4.2. Analogy to field theory . . . . .	138
5.4.3. Relevance for stochastic thermodynamics . . . . .	139
5.5. Summary . . . . .	141
<b>6. Modelling molecular motors</b>	<b>143</b>
6.1. Fluctuations in small systems . . . . .	144
6.1.1. Thermodynamic aspects of molecular motors . . . . .	145
6.1.2. Functional fluctuations . . . . .	146
6.2. Fluctuation-sensitive model reduction . . . . .	147
6.2.1. Heuristic motivation . . . . .	148
6.2.2. Target topologies and coarse-grained observables . . . . .	149
6.2.3. The coarse-graining algorithm . . . . .	150
6.3. Applications to kinesin . . . . .	151
6.3.1. A model for kinesin . . . . .	152
6.3.2. Fluctuations in the coarse-grained model . . . . .	153
6.3.3. Analytical treatment of kinesin's phase diagram . . . . .	155
6.3.4. Simplified models . . . . .	161
6.4. Discussion . . . . .	163
6.4.1. The significance of the SNR . . . . .	163
6.4.2. Consequences for models . . . . .	166
6.5. Summary . . . . .	170
<b>7. Conclusion and outlook</b>	<b>173</b>
7.1. Summary . . . . .	173
7.1.1. Microscopic foundations of stochastic thermodynamics . . . . .	173
7.1.2. Structure and models of stochastic thermodynamics . . . . .	175
7.2. Outlook . . . . .	177
7.2.1. Possible generalizations . . . . .	177
7.2.2. Network multibaker maps as a versatile tool for ergodic theory . . . . .	178
7.2.3. Towards a dynamical picture of local equilibrium . . . . .	178
7.2.4. Deterministic foundations of stochastic fluctuations relations . . . . .	180
7.2.5. Information, complexity and neuroscience . . . . .	181

Contents

7.3. A summarizing (personal) perspective . . . . .	181
7.3.1. On the ubiquity of Markovian statistics . . . . .	182
7.3.2. On information processing systems . . . . .	183
7.3.3. On dissipation . . . . .	185
<b>Acknowledgements</b>	<b>187</b>
<b>A. Mathematical appendix</b>	<b>189</b>
A.1. Construction of the Markov measure . . . . .	189
A.1.1. The basic theorems . . . . .	189
A.1.2. Equivalent definitions of the Markov measures . . . . .	190
A.1.3. Measures on the bi-infinite sequences . . . . .	193
A.2. Sequences of $\tau$ -chains and their variations . . . . .	194
<b>B. Construction of a minimal model for kinesin</b>	<b>197</b>
B.1. State space of the four-state model . . . . .	197
B.2. Choice of parameters . . . . .	199
<b>List of frequently used symbols</b>	<b>201</b>
<b>Bibliography</b>	<b>203</b>



# 1. Introduction

“ Tell them what you are going to say; say it; then tell them what you said. ”

---

Aristotle, *Art of Rhetoric*, 4th century BC

## 1.1. Motivation

Finding an appropriate title for a doctoral thesis is a difficult task. Usually, one starts with a working title. As research progresses and the doctoral candidate’s knowledge deepens, a working title feels increasingly shallow. Often, a good title only emerges when the thesis is almost ready — at a time when it might be impossible to change it any more.

The title of the present thesis is “Foundations of Stochastic Thermodynamics”. Admittedly, such a title sounds rather like the title of a review article than a work of original research. Also the subtitle “Entropy, Dissipation and Information in Effective Models of Small Systems” only slightly specifies the topic of this thesis.

Therefore, as a motivation and introduction to what follows, let us quickly go through the title before we formulate our research question.

### 1.1.1. Stochastic thermodynamics

*Stochastic thermodynamics* (ST) is a modern paradigm for the treatment of small systems in thermodynamic environments [Sei08; Sei12]. In particular, ST studies *non-equilibrium* situations, *i.e.* conditions where a system is actively driven out of equilibrium by some force. Examples include colloids in solution which are driven by external fields [Spe+07; Toy+10; HP11], complex fluids under flow [GO97], actively moving micro-swimmers [Ast97; Rom+12; GC13] as well as small electric devices [Esp+12; Cil+13]. Arguably, the most active field in ST is the study of biologically relevant macro-molecules, ranging from relatively simple molecules like RNA/DNA [Lip+01] to the complex molecular machinery of life [Qia05; LL08; Sei11; BH12].

The above examples show that the mechanisms of driving a system away from equilibrium are as diverse as the systems themselves [CJP10; Sei12]. Experiments on colloids often use optical tweezers, *i.e. external* electrical fields to drive the system. In rheological experiments on soft matter, pressure gradients induce flows. Actively moving particles often carry their own fuel, whereas enzymes and molecular motors reside in a solution of

## 1. Introduction

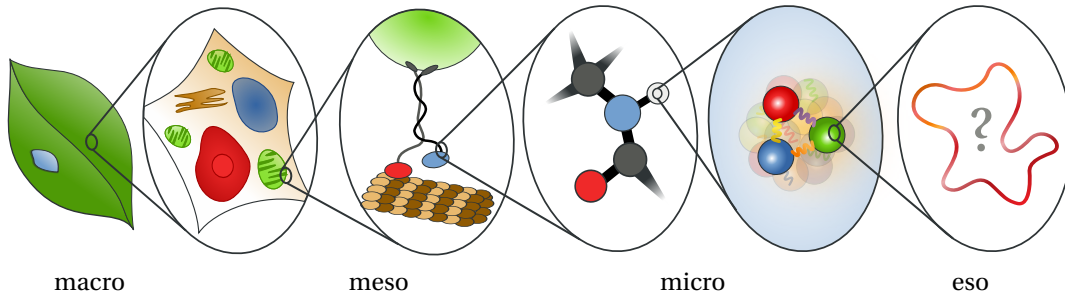
various chemical compounds, which are not in equilibrium with each other. In the latter case, an enzyme's active site acts as a catalyst for otherwise kinetically hindered reactions.

At first sight it seems challenging to capture this variety of systems in one generalized framework. However, for more than one hundred years, *thermodynamics* has been very successful in describing a plethora of different phenomena [GM84]. The key for this success is the abstraction of a thermodynamic *system* and the thermodynamic forces exerted on it by its surrounding *medium*. In this thesis we define a system as the degrees of freedom which are observed in experiments. Hence, the state of a system is defined by the information accessible from a measurement. For the colloid example the state of the system specifies the position of the particle's centre of mass and possibly its velocity and/or rotational degrees of freedom. Similarly, for a complex biological macromolecule one is usually more interested in its tertiary or quaternary structure, *i.e.* its overall geometric shape rather than the position of each atom. Hence, the state of the system may be defined by a set of *coarse-grained* degrees of freedom. All other unresolved degrees of freedom constitute the "medium".

The effect of driving and drag forces, which are mediated by the medium, are observable thermodynamic *currents*. In addition to these *macroscopic* effects, *small* (sometimes called *mesoscopic*) systems also feel erratic forces. The latter originate in the essentially random motion of the medium's constituents. Usually these effects are collectively summarized as "thermal noise". For small systems thermal noise manifests in fluctuations of physical observables. For large systems the typical energy scales are well above the thermal energy of about  $k_B T \approx 4 \cdot 10^{-11} J$ . Consequently, fluctuations are not relevant and usually negligible on the macroscopic scale. In order to observe these fluctuations experiments require a very high degree of precision. Hence, it is not surprising that the development of the *theoretical framework* of ST in the last twenty year went hand in hand with the refinement of experimental techniques [CJP10].

To account for the apparently random behaviour observed for small systems, the *models* used in ST include fluctuating forces. Thus, the system's *trajectory* is obtained as a random process, rather than given by a deterministic evolution rule. A realization of the fluctuating forces is called the *noise history* of the system. The mathematical framework of *stochastic processes* allows the assignment of probabilities to noise histories. Consequently, one assigns probabilities to fluctuation trajectories and other dynamical observables [VK92; Sek98].

Stochastic thermodynamics obtains its name from its goal to generalize *thermodynamic notions* like heat, work, dissipation and efficiency to this stochastic setting. A big emphasis is put on the *molecular machinery of life*, *i.e.* the *molecular motors* performing work within living cells. The key innovation of modern ST is the definition of entropy changes in the system and its medium for single stochastic trajectories [Sek98; Kur98; LS99; Mae04; Sei05]. In this new approach, one considers both the properties of a single trajectory and of the entire *ensemble*, which specifies the probability of finding the system in a specific state. It was recently realized that this approach leads to a unification of stochastic *fluctuations*



**Figure 1.1.:** Different levels of description. The distinction between the macroscopic, mesoscopic and microscopic levels is not unambiguous. In this work, we make the following distinction: The macroscopic level is described using deterministic, irreversible laws like hydrodynamics. For the mesoscopic level, thermal noise plays a major role and stochastic models are used. The microscopic level refers to any underlying deterministic and reversible description. The esoteric level comprises more fundamental theories which cannot be falsified (yet).

relations [Sei05]. The latter are detailed versions of the *second law of thermodynamics*. They are statements about the probability of finding individual trajectories that yield a decrease rather than an increase of entropy. In fact they are examples of the few exact generally applicable results for thermodynamic systems far from equilibrium [Mae04; Sei05; Sei12].

Besides statements about the entropy, ST also aims to quantify noise-driven fluctuations in other physical observables. Often one is interested in the probability of *rare events* in small systems. For instance, as a result of a fluctuation molecular machines may run in reverse or particles may move against an external field. Note that such events are *not* in contradiction with either the first or the second law of thermodynamics. If a particle moves against an external field, the energy necessary is provided by its medium. However, such a behaviour is atypical, *i.e.* it occurs with a low *probability*. Upon *averaging* over the entire ensemble, we still find that work is dissipated into heat and not the other way round, as guaranteed by the second law.

A well-established mathematical tool for the treatment of rare events is the theory of *large deviations* (*cf.* for instance Ref. [Ell05]). Large-deviations theory has been unified formally in 1966 by Varadhan [Var66]. It formalizes the heuristic ideas of the convergence of probability measures. With its applications in statistical physics in general [Tou09] and ST in particular [AG07; FDP11], large-deviations theory has become a prime example for the application of an abstract mathematical theory in a very interdisciplinary context.

### 1.1.2. Foundations

Stochastic processes and large deviations theory provide the *mathematical foundations* of ST. Consequently, they will play a major role in the present work. However, the “Foundations” appearing in the title of the present thesis also refer to another, more physical, aspect. Stochastic thermodynamics is a framework for the treatment of stochastic behaviour observed in *mesoscopic* systems. In that sense it is an *effective theory* with a

## 1. Introduction

validity for the description on a certain scale of observation. Besides the mathematical foundations, this thesis is mainly concerned with the *microscopic foundations* of ST, *i.e.* the relation of ST to an underlying microscopic dynamics.

Admittedly, the distinction between the macroscopic, mesoscopic and microscopic scale of description is ambiguous. Often typical length scales are used as a distinction. However, there are no definite boundaries between, say, the microscopic and the mesoscopic level. Hence, in the present thesis, we distinguish the scales of description by their model paradigms. More precisely, we call a model or a theory *macroscopic*, if its dynamical equations are deterministic and irreversible, *i.e.* not symmetric upon reversing the direction of time. *Mesoscopic* theories, like ST, are based on stochastic models. In analogy to Hamiltonian mechanics, we say that a system is described by a *microscopic* theory, if it evolves according to time-reversible, deterministic laws, *cf.* Figure 1.1. With this terminology, the microscopic foundations of ST are concerned with a deterministic level of description underlying the stochastic mesoscopic description.

One of the fundamental assumptions of statistical mechanics is the *Markovian postulate* regarding the dynamics of observable states [Pen70]. It states that the system's trajectory is generated by a *memoryless* (so-called Markovian) process.

For ST, the Markovian postulate is understood as a consequence of the assumption of *local equilibrium* (LE) [Sei11]. Local equilibrium is a consistency assumption that relates the statistics of the degrees of freedom of the medium to the statistics of the stochastic terms used in mesoscopic models. More precisely, one assumes that *on the time scale of mesoscopic (or macroscopic) observations*, the distribution of the unobserved degrees of freedom are well-described by *equilibrium* probability densities. Equilibrium distributions are asymptotic distributions, which are encountered in a non-driven system in the long-time limit. They act as *attractors*: Under *equilibrium conditions*, any initial distribution will converge to an equilibrium distribution. In that process, the distribution loses the memory of its past, *i.e.* the memory of its previous interactions with the system.

From this point of view, the Markovian postulate is a prerequisite for LE: The random forces exerted by the medium on the system are assumed to be sampled from an *equilibrium* distribution. As a result, they are uncorrelated with the past of the system or medium.

The *separation of time scales* between the microscopic and mesoscopic levels is also known as an *adiabatic approximation* [VK92]: From the perspective of the medium, the system evolves slowly enough for viewing the medium as being at a (constrained) thermodynamic equilibrium at any time. Assuming an underlying microscopic dynamics in continuous time, the Markovian postulate can only hold in the limit of an *infinite* separation of time scales. Such an infinite separation is itself either an unphysical or uninteresting limit: If the microscopic time scale is finite, the limit implies that *nothing* ever changes on the observable level. On the other hand, if we let the microscopic time scale approach zero we might run into relativistic problems.

Local equilibrium should thus be understood as a useful approximation for *practical*

*purposes* instead than a strict assumption. Additionally, it is desirable to have a proper *dynamical picture* of LE. A major part of this is concerned with the relation between a microscopic deterministic dynamics and a stochastic description of observable, *i.e.* experimentally accessible, states.

Classically, the microscopic-deterministic equations of motion are Hamiltonian. However, modern computer simulations also use non-Hamiltonian, *effective* deterministic-reversible equations of motion. The microscopic character of such an approach is also implicit in the term “Molecular dynamics” (MD), which is often used synonymously with deterministic computer simulations [Hoo83; EM90]. In spite of their name, such models do not treat all molecules of a system individually. For instance, MD is used to model the behaviour of single molecules in solution, without explicitly treating the dynamics of the solvent molecules. Rather, the action of the solvent molecules is reduced to their role as a *heat bath*, *i.e.* the absorption and release of energy from and into the system. Consequently, one speaks of *thermostated* MD.

If microscopic is understood as “from first principles” or “fundamental”, one could (rightfully) argue that effective models like thermostated MD are not microscopic theories. However, in the present work we treat thermostated MD on the same level as Hamilton’s equations of motion. Our argument can be understood with regard to Figure 1.1: If there is no *objective*, physical distinction in the terminology, the distinction must be made elsewhere. The present work is theoretical in its nature. Hence, it is only natural that we use the paradigms for the mathematical modelling to distinguish between different levels of description.

### 1.1.3. Entropy, dissipation and information

Let us now discuss the subtitle “Entropy, Dissipation and Information in Models of Small Systems” of the present thesis. First, note that besides implying a separation of time scales, LE is also a statement about thermodynamic *consistency*. More precisely, the assumption of an equilibrium distribution for the medium allows for a definition of the thermodynamic *entropy* of an observable state. In fact, the term “local” in LE is a remnant of the formulation in its original context, *i.e.* thermodynamic transport theory. The latter is a continuum theory formulated in physical space. In transport theory, LE is the assumption that at any point in space, the fundamental thermodynamic relations are obeyed by density fields for internal energy, entropy, temperature *etc* [GM84].

The notion of entropy first appeared in the work of Clausius [Cla65]. His intuition of entropy was that of energy exchanged with the medium as heat. Building on Carnot’s notion of a reversible process, he arrived at the system’s entropy as a state variable. Reversible processes are infinitely slow. In practice, any real process is *irreversible*.

Upon the completion of an irreversible *cyclic* process, which brings the system back to its original state, the state of the medium has changed. Though some energy might have been converted into the potential energy of a work reservoir (*e.g.* a weight lifted against

## 1. Introduction

gravity), the heat in the medium has increased.<sup>1</sup> Alternatively, we can say the entropy of the medium has increased. This phenomenon is usually referred to as *dissipation*.

With the introduction of statistical mechanics by Gibbs, entropy obtained a *statistical interpretation*. The Gibbs entropy formula

$$S = -k_B \sum_{\omega} p_{\omega} \log p_{\omega}$$

defines entropy with respect to the *probability distribution*  $p_{\omega}$ . In Gibbs' considerations, this probability distribution is interpreted as an ensemble with a *frequentist* interpretation: It specifies the sampling probability of observing a certain state when picking a system from a large number of identical copies.

At the same time, Boltzmann introduced entropy as

$$S = k_B \log \Pi$$

where  $\Pi$  is the number of microscopic states compatible with a given macroscopic state. Using the framework of Hamiltonian mechanics together with the assumption of *ergodicity*, a *microscopical* relation between the two concepts of entropy can be established.

In the first half of the twentieth century, statistical mechanics was mostly discussed following Gibbs' and Boltzmann's lines of thought. Ergodic theory [Hop48; CFS82], which is concerned with probability and the evolution of dynamical systems, was originally perceived within this context. At the same time, scientists started to formalize the notion of deterministic chaos, *i.e.* situations where small changes in the initial state of the system grow exponentially fast with time. Consequently, the ergodic theory for chaotic systems became the major field of study regarding the mathematical foundations of statistical mechanics [Sin72; BC75; BS95; Rue04; Khi13].

In the 1940s, Shannon discovered the importance of Gibbs' formula in his theory of *communication* [Sha48]. More precisely, he found that the entropy formula for probability distributions has all the desired properties of a quantity which characterizes the *uncertainty* of the content of (statistically generated) messages. Nowadays, one refers to the subject founded by Shannon as *information theory*. It constitutes the basis of *all* digital communication, coding and information storage.

Realizing the importance of entropy for applied statistics in general, Jaynes argued that there is no *conceptual difference* which distinguishes entropy in information theory from entropy in statistical mechanics [Jay57]. Based on this premiss, he advocated a view of statistical physics (and science in general) as a theory of logical *statistical inference* [Jay03]. He claims that, if viewed in that way, statistical mechanics can be *logically derived* from the structure of the *underlying fundamental laws* [Jay57]. In that approach, the principle of *maximum entropy* replaces the more technical ergodic requirements demanded by

---

<sup>1</sup>This is also the case for a *heat pump* which uses the energy stored in a work reservoir to cool one heat bath while heating up another. The net heat balance in the medium comprising all reservoirs and heat baths is still positive.

the usual treatment from the perspective of mathematical physics, *cf. e.g.* Ref. [Hop48]. As such it might help us to understand why classical thermodynamic concepts are — perhaps unexpectedly — useful in describing systems whose microscopic dynamics are vastly different from what is usually assumed. An example is provided by the physics of wet granular media as described in Ref. [Her14].

Jaynes’ approach has been both celebrated and rejected by parts of the physics community, partly due to his (physical) interpretation being applied outside of its original context. After all, probability distributions (and thus the corresponding entropies) arise naturally at various levels of and within several different paradigms for the descriptions of physical and mathematical systems, *cf.* also Ref. [FW11]. However, the *thermodynamic* interpretation of the information/entropy associated with an arbitrary probability distribution has to be attempted *cum grano salis*: In order to avoid logical fallacies, it is crucial to carefully review the framework in which these probabilistic notions arise.<sup>2</sup>

In spite of the criticism of Jaynes’ ideas by parts of the physics community, his premiss of a deep conceptual connection between statistical thermodynamics and information theory has been developed further. With the advent of digital computers, Landauer and later Bennett discussed the “thermodynamics of computation” [Lan61; Ben82; Ben03]. Landauer’s principle states that the erasure of an elementary unit of binary information, a *bit*, from a storage medium in a computer comes at the price of at least  $Q = k_B T \log 2$  of dissipated heat [Lan61]. Bennett put this result in the context of the old problem of Maxwell’s or Szilard’s demons [Szi29; Ben03]. He stresses that the *information* that such an imaginary demon *processes* equals the maximal amount of work that can be extracted by the demon. Further thoughts in that direction have recently lead to a general framework of “information thermodynamics” [SU10; Sag12]. Conceptually, a demon can be thought of as a feedback protocol — a point of view that has proven useful for the optimal design of small thermodynamic engines [HP11]. In light of the work discussed above, it should not be surprising that the predictions of information thermodynamics have been confirmed by recent experiments on small systems [Toy+10; Bér+12]. This research as well as other work in the same direction [HBS14] strongly support the information-theoretical perspective on statistical mechanics.

In light of the examples given above, we consider it only natural to look at stochastic thermodynamics from Jaynes’ point of view, *i.e.* as a (dynamical) theory of statistical inference. In fact, one can go a step further and generally understand the statistical mechanics of non-equilibrium situations as the study of models of *information processing systems*. The emphasis on *models* is important; it stresses that information (and thus entropy) needs to be formulated in an operational or descriptive context. At the very end of the present work, we return to these ideas and discuss them in more detail.

---

<sup>2</sup>Examples of common misconceptions of entropy that lead to apparent paradoxes are the “constant entropy paradox” for Hamiltonian dynamics (*cf. e.g.* [Rue99]) and the interpretation of entropy as “disorder”.

## 1. Introduction

### 1.1.4. Research questions

After having motivated the context of this thesis, we formulate its research questions. The work splits into two parts.

**Microscopic foundations** Within the framework of having a microscopic-deterministic and a coarse-grained, mesoscopic-stochastic level of description, we formulate two questions:

- What are the implications of the *Markovian postulate* on the mesoscopic level of description for the microscopic dynamics?
- Can, and if yes how, stochastic thermodynamics be obtained in an *information-theoretical framework*?

Both questions point towards a dynamical or information-theoretical picture of local equilibrium. Hence, in our investigations we will point out when certain physical assumptions appear as logical-probabilistic consistency relations between different models.

**Mathematical foundations** In the second part of the present thesis, we deal with the mathematical foundations of ST formulated on discrete state spaces. The network of states, which we use to describe a mesoscopic system, is represented as a graph. Using concepts from graph theory and the theory of large deviations we address the following questions:

- What is the general structure of discrete ST and how can we use it in order to *characterize fluctuations* of physical observables?
- How can we use such concepts in order to compare different *mesoscopic models* for real physical systems with each other?

In the context of the first question, we see how the results of Kirchhoff on electrical circuits reappear in the present setting. More precisely, we discuss the importance of *cycles* for small systems driven away from equilibrium. As a solution to the second question we propose to consider the statistics of *dissipation*, which we interpret as information written to unobservable degrees of freedom. We illustrate our results using models for a system which plays a huge role for the function of living cells: The molecular motor *kinesin*.



## 1.2. The structure of this thesis

### 1.2.1. How to read this thesis

The initial quote, in some form or another, is usually attributed to Aristotle and his teachings on rhetorics. Admittedly, I have never studied the alleged “Master of Rhetorics” himself nor heard him speak. Thus, I cannot say whether the quote is original. However, it seems equally good advice for both writing a thesis and for giving an oral presentation. I mention the advice at this point, because it may serve as a guide on how to *read* the present work.

In the spirit of Aristotle’s suggestion, the multiple hierarchical levels of this thesis also show some amount of intended redundancy. On the highest level, the outline presented in the next subsection will tell the reader what and what not to expect from the story told by this thesis. Similarly, the discussion in the final chapter comes back to the general picture presented here.

The central Chapters 2–6 are written in the same spirit. Each chapter starts with an initial quote followed by a short introduction in order to give an idea of “What is this about?”. After the introduction, a presentation of the methods and results precedes a detailed discussion of the latter. Finally, we give a short summary and motivate the connection to the contents of the subsequent chapter.

### 1.2.2. Outline

Chapter 2 reviews different notions of entropy and entropy changes as they occur in different physical and mathematical settings. Consequently, that chapter should be considered as an extended introduction, providing the necessary mathematical and physical terminology needed in what follows. In particular, we focus on entropy and dissipation in both stochastic and deterministic models of complex systems in thermodynamic environments.

The main part of the thesis is divided into two parts. The first part starts with Chapter 3, which revisits the above-mentioned Markovian postulate. More precisely, we make explicit the requirements on dynamics, observables and ensembles such that the Markovian postulate holds. For this formal treatment, we introduce an abstract framework for the process of recording mesoscopic time series on a system evolving according to deterministic microscopic laws. Eventually, the mathematical results are put into the context of ergodic theory and we equip them with *operational* interpretations.

In Chapter 4 we attempt an information-theoretical interpretation of the framework introduced in Chapter 3. However, we will *not* make use of the Markovian postulate or the concept of local equilibrium. Instead we try to base our argument purely on information-theoretical aspects. In order to make our considerations more transparent in examples, we introduce a versatile, yet analytically tractable, microscopic model dynamics. We will see that the Markovian postulate holds rigorously for that model, and ST emerges as

## 1. Introduction

an information-theoretical interpretation. Based on this central result, we conjecture a general mechanism for the emergence of ST from an underlying microscopic dynamics.

The second part of the thesis starts with Chapter 5, where we deal with the mathematical theory of Markovian dynamics on a finite state space. Finiteness ensures that the topology induced by the stochastic dynamics on state space can be represented as a graph. Viewing the graph as an electrical circuit, we present an electro-dynamical analogy of ST. The rationale behind this analogy are algebraic-topological considerations, pioneered already in the nineteenth century by Kirchhoff. In analogy to Kirchhoff's "mesh" or "circuit law", we see how *cycles* play a fundamental role in non-equilibrium situations. This in turn gives an intuition of the intimate connection between cycles and the thermodynamic (macroscopic) forces that drive the system.

Building on the electro-dynamical analogy, we investigate the structure of Markovian jump processes from the theory of algebraic topology. We establish an analytical way to quantify fluctuations in these processes, *i.e.* any behaviour that deviates from ensemble expectations. Our results stress that the topology of the network is extremely important: Fluctuations of *any* physical observable are shown to depend only on the fluctuation statistics of currents associated with a set of fundamental cycles.

Chapter 6 is concerned with fluctuations in models of ST. This is particularly relevant for models of the molecular machinery of living cells. In the light of evolution it is not surprising that there are many cases where fluctuations are important for the function of an organism.

We explicitly discuss the design and structure of chemo-mechanical models using the molecular motor *kinesin* as an example. As a main result, we present a fluctuation-sensitive model reduction procedure and investigate its heuristic motivation from the topological perspective established in Chapter 5.

In addition, we demonstrate how minimal models can be designed in a systematic way. With our methods we give a detailed account of kinesin's phase diagram, which is spanned by chemical and mechanical driving forces. In contrast to previous characterizations using approximations or numerics, our results are completely analytic. Moreover, we find that the fluctuation statistics found in our simplified models agree very well with the prediction of a more complex model known in the literature. The relative mismatches amount to only few percent in the majority of the phase diagram — for values ranging over twenty logarithmic decades. Finally, we show how our method unifies previous approaches to the exact calculation of dynamic properties of molecular machines, like drift and diffusion.

Chapter 7 provides a summary and an outlook on interesting future research. We finish with a personal perspective on non-equilibrium thermodynamics as the study of information processing devices.

### 1.2.3. Notation, abbreviations and conventions

A Ph.D. thesis is always composed of work which has been obtained over an extended period of time. During that time, preliminary results are being generalized and new definitions or formulations are constantly being created at the expense of older ones. Consequently, it is fair to say that the general notation has evolved quite a bit during both research for and the formulation of a thesis.

In the optimal case, this evolution leads to a consistent presentation of the results. As in so many cases, this optimum is hardly ever reached. The current thesis is no exception to that rule. Still, the reader might benefit from the following remarks.

**Language** We tried to use British English as a convention throughout the entire thesis. Abbreviations are usually introduced in the context where they first appear. The most commonly used ones are: stochastic thermodynamics (ST), local equilibrium (LE), [non-equilibrium] molecular dynamics ([NE]MD), subshift of finite type (SFT), network multibaker map (NMBM), [scaled] cumulant-generating function ([S]CGF) and adenosine triphosphate (ATP).

**Mathematical notation** In the present work, “log” denotes the *natural* logarithm. The natural numbers  $\mathbb{N} = (0, 1, \dots)$  always include zero as the neutral element of addition. Ensemble averages  $\langle \cdot \rangle_t$  are denoted by chevrons and a subscript indicates that the probability density reflects an ensemble at time  $t$ . *Time series*  $\underline{\omega}$  and orbits  $\underline{x}$  are discrete or continuous successions of values and exhibit an under-bar to distinguish them from values  $\omega_t$  or  $x_t$  at a specific point in time. Time series  $\underline{\omega}^{(\tau)}$  of finite run length  $\tau$  are equipped with a superscript. Similarly, averages  $\langle \cdot \rangle_t^{(\tau)}$  which are taken over an ensemble of trajectories that start at time  $t$  and extend until time  $t + \tau$  carry both decorators. The *time average*  $\overline{\varphi}_t^{(\tau)}$  of an observable  $\varphi$  along a single trajectory  $\underline{\omega}^{(\tau)}$  is denoted with an over-bar. Generally, a single point in time is indicated by a subscript  $t$  while a run length is indicated by a superscript  $(\tau)$ . In addition, a list of *frequently used symbols* can be found at the end of the present work.

**Figures** All of the sketches were designed using the free software *Inkscape*. Contour plots were rendered using *Mathematica*<sup>TM</sup>.

**Copyright** This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>.





## 2. Notions of entropy and entropy production

“ You should call it entropy, for two reasons. In the first place, your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage. ”

---

J. v. Neumann to C. E. Shannon, 1940–1941

### What is this about?

The introductory quote (or slightly different formulations thereof) has its origin in a conversation between John von Neumann and Claude E. Shannon attributed to a period of time between autumn 1940 and spring 1941 [TM71]. At that time, Shannon was working on his post-doctoral studies at the Institute for Advanced Study in Princeton, New Jersey, where von Neumann was one of the faculty members. Previous to the conversation Shannon had realized the importance of the expression

$$-\sum_i p_i \log p_i$$

for his statistical formulation of signal transmission (*cf.* Section 2.2). He thought about calling it “uncertainty” rather than “information”, because he was concerned that the latter term is already overly used and might be misleading. The quote above is Neumann’s alleged answer to Shannon when he was asked about the naming issue.

The present chapter picks up on the second part of the quote which regards the nature and meaning of entropy. More precisely, we present different notions of entropy and entropy production that arise in different branches of physics and mathematics. A main goal of this thesis is to outline and discuss connections between these notions. The review character of this chapter sets the stage for the original results presented in Chapters 3–6.

In the present chapter we introduce the notation and terminology for the rest of this work. In contrast to von Neumann’s suggestion, we aim to disentangle the different

## 2. Notions of entropy and entropy production

meanings of entropy. If we are successful in that task, the reader of this thesis should know exactly what entropy *is* — at least from the perspective of the following investigations.

This chapter is structured as follows: In Section 2.1 we review entropy in the classical thermodynamics of the mid-nineteenth century. After that, Section 2.2 reviews Shannon's and related notions of entropy as *uncertainty* or information of data. In Section 2.3 we use the latter notion to define the entropy of a *system* as the uncertainty in its observed configurations. Consequently, we assign the entropy of a system's environment (which we will refer to as its *medium*) to the (dynamics of) unobservable degrees of freedom. Section 2.4 makes the distinction explicit for stochastic models and introduces the basic idea of *stochastic thermodynamics*. In Section 2.5 we investigate this distinction in the context of deterministic models of complex systems in thermodynamic environments. Finally, Section 2.6 returns to mathematical notions of entropy (production), which characterize the *complexity* of abstract dynamical systems.

### 2.1. Entropy in classical thermodynamics

In classical thermodynamics, the *variation of the entropy of a thermodynamic system* is defined by the relation

$$\Delta S^{\text{sys}} := - \int \frac{\delta Q_{\text{rev}}^{\text{med}}}{T}.$$

In this definition,  $T$  is the thermodynamic temperature and  $Q_{\text{rev}}^{\text{med}}$  is the (integrated) heat flow into<sup>1</sup> the *medium* for a so-called *reversible* process. A reversible process is *defined* to be a sequence of changes to the system's state, such that the integral on the right-hand side depends only on the initial and final state of the system. For a *cyclic* process, the system state is the same both at the beginning and at the end of the process. Hence, irrespective of its specific nature, a reversible cyclic process (in particular, a *Carnot process*) obeys:

$$- \oint \frac{\delta Q_{\text{rev}}^{\text{med}}}{T} = \Delta S^{\text{sys}} = 0.$$

This path-independence ensures that the *entropy of the system*  $S^{\text{sys}}$  is well-defined and obeys the differential relationship  $T dS^{\text{sys}} = \delta Q_{\text{rev}}^{\text{med}}$  for such reversible processes. The Clausius inequality states that *any* cyclic process obeys [Cla54]

$$- \oint \frac{\delta Q^{\text{med}}}{T} \leq 0.$$

This is one of the many formulations of the *second law of thermodynamics*. Note that this equation does not imply that there has been no heat exchange with the medium. Rather, it states that the integrated *ratio* of a heat flux and a (generally varying) temperature

---

<sup>1</sup>Note that we define the heat flow from the perspective of the medium rather of the system. Hence, our sign convention differs from Clausius' classical work [Cla54].

vanishes. Combining a reversible with an irreversible process yields

$$\Delta S^{\text{sys}} \geq - \int \frac{\delta Q^{\text{med}}}{T} =: -\Delta S^{\text{med}},$$

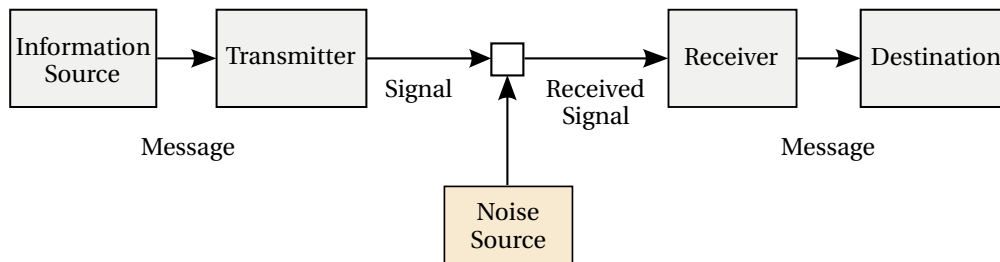
where the right-hand side *defines* the *entropy variation in the medium*. With that, we arrive at a formulation of the second law, where heat  $Q^{\text{med}}$  and temperature  $T$  do not appear explicitly any more:

$$\Delta S^{\text{tot}} := \Delta S^{\text{sys}} + \Delta S^{\text{med}} \geq 0. \quad (2.1)$$

This is the famous formulation of the second law that states that the *total entropy* of a system together with its environment never decreases.<sup>2</sup>

## 2.2. Entropy as information or uncertainty

Information theory is the branch of mathematics that deals with the quantification of information. It was developed in 1948 by C.E. Shannon as a theoretical framework for the processing of electrical signals. At that time Shannon was working at Bell labs, and his seminal work “A Mathematical Theory of Communication” appeared in the *Bell Labs Technical Journal* [Sha48]. The main goal of the paper was to lay out the central elements of communication and to formalize them mathematically (*cf.* figure 2.1).



**Figure 2.1.:** The elements of communication according to Shannon's original paper [Sha48].

Information theory is a framework developed to make *quantitative* statements about the information content of *messages*. In information theory, a message  $\omega$  is a string of *letters*  $\omega \in \Omega$  composed from a finite *alphabet*  $\Omega$ . More precisely, information theory is concerned with the *probability* of a certain letter appearing in a message. One can rephrase that statement as follows: Information theory deals with strings of letters which are generated by a random source. In that regard it can make statements about uncertainty, redundancy and encoding of messages. However, it does not refer to *qualitative* properties such as their meaning or their relevance.

In the following we will motivate information theory in the original setting of a discrete

<sup>2</sup> We refrain from using a statement referring to the *universe*, as we do not divert into a discussion of entropy and information in cosmology. The interested reader is referred to Ref. [Bek03] for the general idea and to Ref. [Bou02] and the references therein for a detailed treatment.

## 2. Notions of entropy and entropy production

random variable  $X$  taking values on a finite set  $\Omega = \{1, 2, \dots, N\}$ . We denote the probability to find a letter  $\omega$  by  $p_\omega^X$ . The probabilities of all possible letters are summarized in the stochastic vector  $\mathbf{p}^X = (p_\omega^X)_{\omega \in \Omega}$ . Entropy is a scalar quantity that characterizes the average *uncertainty* of a letter (or more abstractly, an *event*) to occur. Hence, entropy quantifies the amount of additional information obtained by *observing* a letter in a message generated by a source solely characterized by  $\mathbf{p}^X$ .

The requirements on such an entropy have been formalized mathematically in the so-called Khinchin axioms [Khi57]:

**2.1 Definition (Shannon entropy)** Let  $X$  be a random variable taking values  $\omega \in \Omega$  on a finite set  $\Omega = \{1, 2, \dots, N\}$  with a probability distribution  $\mathbf{p}^X := (p_\omega^X)_{\omega \in \Omega}$ . Then, we call a scalar function  $H[X]$  the *entropy* (or *uncertainty* or *Shannon information*) of  $X$  if it obeys the following axioms:

1.  $H[X]$  depends only on  $\mathbf{p}^X$ , *i.e.* the enumeration of its entries must not matter.
2.  $H[X]$  takes its maximum value for the uniform distribution.
3. Let  $Y$  be a random variable taking values  $y$  on a larger set  $\Omega^Y = \{1, 2, \dots, M\} \supset \Omega^X$  such that its distribution  $\mathbf{p}^Y$  obeys  $p_\omega^Y = p_\omega^X$  for all  $\omega$  in  $\Omega^X$ . Then,  $H[X] = H[Y]$ .
4. For any two random variables  $X$  and  $Y$  with values in  $\Omega^X$  and  $\Omega^Y$ , respectively, we have

$$H[X, Y] = H[X] + \sum_{\omega \in \Omega^X} p_\omega^X H[Y|X = \omega],$$

where  $H[X, Y]$  is the entropy of the joint distribution for the tuple  $(X, Y)$  and  $H[Y|X = \omega]$  is the entropy of the distribution of  $Y$  conditioned on  $X = \omega$ .

It can be shown [Khi57] that the only functional  $H[\cdot]$  form satisfying these axioms is

$$H[X] = H[\mathbf{p}^X] = - \sum_{\omega \in \Omega} [p_\omega \log_b p_\omega], \quad (2.2)$$

where  $\log_b$  denotes the logarithm with respect to base  $b$ . The dependence on the base can also be understood as choosing the *unit* of entropy. For instance, if  $b = 2$  the unit of entropy is called a *bit*. In statistical mechanics, often the natural logarithm is used and entropy is measured in units of the Boltzmann constant  $k_B$ . In the remainder of this thesis we will use the natural logarithm and set  $k_B \equiv 1$ .

To see that this definition of entropy appropriately captures the notion of the uncertainty of  $X$ , let us take a closer look at the first three axioms: The first one says that  $H[X]$  must be independent of the specific nature or enumeration of the events  $\omega \in \Omega$ , *i.e.*  $H[(\frac{1}{3}, \frac{2}{3})] = H[(\frac{2}{3}, \frac{1}{3})]$ . Hence, entropy is well-defined for *any* random variable and we can compare arbitrary random variables with each other. This certainly is a useful thing to demand of a *generally applicable* concept of uncertainty. The second axiom specifies that



### 2.3. Statistical physics and the distinction between system and medium

entropy should be maximal if no event is more probable than any other, in agreement with the informal meaning of uncertainty. The third axiom states that adding zero-probability events to the possible values of a random variable does not change its uncertainty.

Finally, the fourth axiom specifies the additivity of uncertainty. More precisely, it says that the uncertainty of conditional events averages to the uncertainty of the joint distribution. Indeed, this axiom is necessary in order to obtain equation (2.2). However, relaxing or dropping this axiom gives rise to a whole class of generalized entropies, with applications in contexts where a weaker form of additivity is sufficient or desired [Rén61; Tsa88; BS95].

Because a discrete probability vector has entries in the interval  $[0, 1]$ , the entropy (2.2) is always positive. This is not true for the *differential entropy* of a probability density  $\varrho: \Gamma \rightarrow [0, \infty)$  on a continuous space  $\Gamma$ :

$$H[\varrho] := - \int_{\Gamma} \varrho(x) \log \varrho(x) dx \quad (2.3)$$

As the integral is a generalized sum, we will usually use the differential notion of entropy, even if  $\varrho$  is actually a probability distribution  $\mathbf{p}$  on a discrete space. Despite the fact that the expression (2.3) can take negative values (and hence without the direct interpretation as “uncertainty”), the differential entropy is readily used in physics, especially in statistical mechanics.

Another important quantity is the *relative entropy* or *Kullback–Leibler divergence*. For two probability distributions  $\varrho$  and  $\varrho'$  on a state space  $\Gamma$  such that  $\varrho' = 0$  implies  $\varrho = 0$ , it is defined as

$$D_{\text{KL}}[\varrho \parallel \varrho'] := \int_{\Gamma} \varrho(x) \log \frac{\varrho(x)}{\varrho'(x)} dx. \quad (2.4)$$

By using the concavity of the logarithm, it is straightforward to show that  $D_{\text{KL}} \geq 0$  in general and that  $D_{\text{KL}} = 0$  implies measure-theoretic equality of the distributions.

Another quantity we encounter in this work is the cross-entropy of two distributions. It is a measure for the error one makes if a distribution  $\varrho'$  is assumed for a random variable with real distribution  $\varrho$ :

$$H^{\times}[\varrho; \varrho'] := - \int_{\Gamma} \varrho(x) \log \varrho'(x) dx = H[\varrho] + D_{\text{KL}}[\varrho \parallel \varrho'], \quad (2.5)$$

where the second equality requires that  $D_{\text{KL}}$  is defined.

## 2.3. Statistical physics and the distinction between system and medium

In this section we review the fundamental aspects of statistical physics we will need in the remainder of this work. Classical statistical physics has been developed in order

## 2. Notions of entropy and entropy production

to provide a microscopic background for thermodynamics. It is based on Hamiltonian dynamics, which is a deterministic evolution rule for microscopic states. A microscopic state contains information about the degrees of freedom of all the particles that make up a macroscopic system. The number of such degrees of freedom is very large. Thus, computing the dynamics of individual configurations is cumbersome. Moreover, for several reasons which we will analyse in more detail later, such calculations are also not effective in order to obtain physical statements. Hence, rather than focussing on individual microscopic configurations, statistical physics makes probabilistic statements. For instance, it features a statistical derivation of the second law of thermodynamics (2.1).

### 2.3.1. The second law in statistical physics

In classical thermodynamics, the second law is a *macroscopic* statement about macroscopic states. Similarly, the fundamental equations of *thermodynamic transport theory* are continuity equations for macroscopically defined quantities [GM84]. In both cases, matter is treated as a continuum and one neglects the existence of individual atoms or molecules. At macroscopic scales, the granularity of matter is not visible and the continuum approximation is sufficient. For smaller systems, however, *fluctuations* due to finite particle numbers play a role. For electrical systems, this effect is referred to as *shot noise* [BB00].

In classical statistical physics, one relies on the notion of a *thermodynamic limit*, where the number of particles goes to infinity. In this limit, fluctuations are negligible. In contrast, modern statistical physics does not necessarily assume this limit. Consequently, fluctuations in non-macroscopic systems become relevant and should be included in the theory. Modern generalizations of the second law are thus detailed probabilistic statements, rather than statements about (macroscopic) averages. However, consistency requires that the second law of thermodynamics as formulated in (2.1) must emerge in the macroscopic limit.

The recent years have seen a multitude of such generalizations of the second law for different (non-thermodynamic) models of complex systems. Amongst the most famous of such statements are the inequalities of C. Jarzynski [Jar97] and G. Crooks [Cro99]. Even more recently, these relations have been understood as being consequences of the so-called fluctuation relations for finite systems in thermodynamic environments [Mae04; Sei12]. Moreover, they have been tested and verified numerically and experimentally [CJP10].

For the formulation of fluctuation relations, one defines entropy changes associated with the *system* and its surrounding *medium*, similar to equation (2.1). While this distinction is quite clear for macroscopic thermodynamic systems like engines or refrigerators, for small systems it becomes more subtle. In this work, we identify the system with the set of *observed* degrees of freedom. Consequently, the medium contains the *unobserved* degrees of freedom.

### 2.3. Statistical physics and the distinction between system and medium

This distinction based on observability has the advantage that there is no need for a *spatial* separation of the system and the medium. This is already an implicit feature of any hydrodynamic theory. For instance, in the Navier–Stokes equation, viscosity acts as a *transport coefficient* for an energy flow from observable hydrodynamic to unobservable internal degrees of freedom.

Other examples are systems in chemical environments. In particular, we are interested in biological macromolecules which are often surrounded by different chemical compounds. In biology, a macromolecular system often acts as a *catalyst* which enables (or at least strongly accelerates) reactions between the chemical species. If such a catalytic reaction additionally triggers an (observable) conformational change on the level of the system itself, one also speaks of *molecular motors*. In these examples, the medium is composed of the molecules of the solvent and the solutes as well as unobservable microscopic degrees of freedom of the macromolecule. Even in a well-mixed environment, the solute concentrations need not be in equilibrium with each other. Hence, the medium provides a *heat bath* as well as different *chemical reservoirs*, which are not spatially separated.

Although a distinction between system and environment based on *observability* seems useful, it comes at the price of subjectivity: Observability is always an operational, and thus a subjective quality, which is determined by the choice or capability of an observer performing *measurements* on the system. One goal of this thesis is to shed light on physical implications of that type of subjectivity.

#### 2.3.2. Entropy changes in statistical physics

Keeping the issue of subjectivity discussed in the last subsection in mind, we look for definitions of the entropy changes  $\Delta S^{\text{sys}}$  and  $\Delta S^{\text{med}}$  in modern statistical physics. We begin with some general considerations here and then explicitly define these quantities for modern model paradigms. In particular, we will look at stochastic (jump) processes and molecular dynamics simulations in sections 2.4 and 2.5, respectively.

A concept common to all models in statistical physics is the notion of an *ensemble*. An ensemble specifies the probability of picking a system at a certain microscopic state from a large number of copies of a system. Mathematically, ensembles are probability densities<sup>3</sup>  $\rho^{\text{sys}}: X \rightarrow [0, \infty)$  defined on the *state space*  $X$  of a *model*. The *system's entropy*  $S^{\text{sys}}$  is defined to be the (differential) entropy of the distribution  $\rho^{\text{sys}}$  of the observed degrees of freedom

$$S^{\text{sys}} := H[\rho^{\text{sys}}] \equiv - \int_X \rho^{\text{sys}} \log \rho^{\text{sys}} dx. \quad (2.6)$$

Subjectivity also enters into purely theoretical considerations of mathematical *models* for physical systems, even without the reference to a measurement: It manifests in the degrees of freedom we choose to make up the state space of a model. A *dynamical* model

---

<sup>3</sup>In this section, we only consider situations where such a density exists. We do not yet discuss the measure-theoretic formulation. For an account of the latter cf. Chapter 3 or Ref [Alt p].

## 2. Notions of entropy and entropy production

specifies an evolution rule on the state space. Consequently, the dynamics prescribes an evolution operator  $U_t^{(\tau)}: \rho_t \mapsto \rho_{t+\tau}$  for the ensemble  $\rho_t$ .<sup>4</sup> Hence, the system's entropy becomes a time-dependent quantity  $S_t^{\text{sys}} := H[\rho_t^{\text{sys}}]$ . The *temporal variation* of the system entropy in the interval  $[t, t + \tau]$  is defined as

$$\Delta_t^{(\tau)} S^{\text{sys}} := S_{t+\tau}^{\text{sys}} - S_t^{\text{sys}}.$$

As for classical thermodynamics, the entropy change in the medium is related to the *irreversibility* of a process. Let us denote the evolution operator of a suitable *reversed process* by  $RU_t^{(\tau)}$ . Often, the term or operator responsible for the *temporal variation of the entropy in the medium* has the form

$$\Delta_t^{(\tau)} S^{\text{med}} \sim \int \log \left( \frac{U_t^{(\tau)}}{RU_{t+\tau}^{(\tau)}} \right) \rho_t dx. \quad (2.7)$$

Various examples of this relation can be found in [Mae04].

In Sections 2.4 and 2.5 we will be more concrete and give the expressions for  $\Delta_t^{(\tau)} S^{\text{sys}}$  and  $\Delta_t^{(\tau)} S^{\text{med}}$  for some common models of complex systems in thermodynamic environments. Beforehand, we revisit the microscopic theory of *isolated systems*, namely Hamiltonian dynamics.

### 2.3.3. Hamiltonian dynamics

Classical statistical mechanics is formulated based on *Hamiltonian dynamics* [Gib48; CS98; KHi13]. In Hamiltonian dynamics, a point  $x = (\vec{q}, \vec{p}) \in \Gamma$  fully represents the state of a system. The dynamics is deterministic, *i.e.* the state  $x_t$  after some time  $t$  is fully determined by the *initial condition*  $x_0$ . The *phase space*  $\Gamma$  of Hamiltonian dynamics is the state space of an *isolated system*.<sup>5</sup> The degrees of freedom  $x$  split into the (generalized) coordinates  $\vec{q}$  and (generalized) momenta  $\vec{p}$  of all  $N$  particles that constitute the system. For brevity, here and in the following we use the notation  $\vec{q} = \{\vec{q}_k\}_{k=1}^N$ ,  $\vec{p} = \{\vec{p}_k\}_{k=1}^N$  where no ambiguity can arise.

The *Hamiltonian*<sup>6</sup>

$$\mathcal{H}(x) = V(\vec{q}) + \frac{\vec{p}^2}{2m}. \quad (2.8a)$$

is the dynamic variable that represents the total energy  $E$  of the system. It determines the

<sup>4</sup> The evolution operator for deterministic dynamics is often called the Frobenius–Perron operator, whereas for stochastic systems it is often called the Smoluchowski or Fokker–Planck operator.

<sup>5</sup> Closed and open systems can be obtained by considering only subsets of the phase space as the system, whereas the rest is identified with the medium.

<sup>6</sup> In this notation, the term  $\vec{p}^2/2m$  (2.8a) is short for  $\sum_{k=1}^N \frac{\vec{p}_k^2}{2m_k}$  including the (possibly different) masses  $m_k$ .

equations of motion

$$\dot{\vec{q}} = \nabla_{\vec{p}} \mathcal{H} = \frac{\vec{p}}{m}, \quad (2.8b)$$

$$\dot{\vec{p}} = -\nabla_{\vec{q}} \mathcal{H} = -\nabla_{\vec{q}} V(\vec{q}). \quad (2.8c)$$

In the above equations (2.8),  $\dot{x} := \frac{dx}{dt}$  denotes the total derivative with respect to time and  $\nabla_{\{x\}}(\cdot)$  denotes the vector gradient (also denoted  $\text{grad}_{\{x\}}(\cdot)$ ) of a scalar function with respect to the set of coordinates  $\{x\}$ .

The first term in the Hamiltonian,  $V(\vec{q})$ , is a potential that gives rise to (conservative) forces  $F^{\text{cons}}(\vec{q}) := -\nabla_{\vec{q}} V(\vec{q})$ . The second term denotes the total kinetic energy. Moreover, the Hamiltonian is a constant of motion, *i.e.*  $\mathcal{H}(x_t) = \mathcal{H}(x_0) = E$  does not change over time. Hence, energy is conserved as we would expect it from an isolated system.

Hamiltonian dynamics are a standard example of deterministic-chaotic systems. Its equations are usually non-linear and high dimensional, and thus generically show a sensitive dependence on initial conditions: The distance of infinitesimally separated points in phase space  $\delta x$  shows an (initial) exponential growth with time. In contrast, detailed information on microscopic initial conditions is never accurately available for real systems. Hence, it must be specified in a probabilistic way — which lead to the notion of Gibbs' statistical ensembles.<sup>7</sup> Moreover, Gibbs was the first to write down the functional form of the (differential) entropy (2.2) associated with a phase-space ensemble  $\rho_t$ .

A probability density for a dynamics  $\partial_t x = f(x)$  satisfies a continuity equation, because probability is conserved. Henceforth, we denote the partial derivative with respect to time  $t$  by  $\partial_t$  and the divergence of a vector field  $f$  by  $\text{div}(f) \equiv \nabla \cdot f$ . The continuity equation then reads:

$$\begin{aligned} \partial_t \rho_t &= -\text{div}(f(x) \rho_t) \\ &= -\rho_t \text{div}(f(x)) - f(x) \cdot \text{grad}(\rho_t). \end{aligned} \quad (2.9)$$

Rearranging this equation, we find for the total derivative of the probability density:

$$\frac{d\rho_t}{dt} = \partial_t \rho_t + f(x) \cdot \text{grad}(\rho_t) = -\rho_t \text{div}(f(x)).$$

Note that this equation can be rewritten as

$$\frac{d(-\log \rho_t)}{dt} = \text{div}(f) =: \Lambda, \quad (2.10)$$

where  $\Lambda(x)$  is called the *phase space expansion rate*. For Hamiltonian dynamics, phase space volume is conserved, *i.e.* the expansion rate identically vanishes:

$$\Lambda \equiv \text{div}(f) := \frac{d(\partial_t q)}{dq} + \frac{d(\partial_t p)}{dp} = \frac{d^2 \mathcal{H}}{dq dp} - \frac{d^2 \mathcal{H}}{dp dq} = 0 \quad (2.11)$$

<sup>7</sup> A collection of Gibbs' pioneering work can be found in Ref. [Gib48]

## 2. Notions of entropy and entropy production

and thus

$$\frac{d\rho_t}{dt} = 0. \quad (2.12)$$

This statement, usually known as the “Liouville theorem”, was first written down by Gibbs in the context of his statistical ensembles. He soon realized that conservation of phase space volumes implies the conservation of the entropy:

$$\frac{dH[\rho_t]}{dt} = 0. \quad (2.13)$$

This fact is often referred to as the “paradox of constant entropy” in Hamiltonian systems, as it seems to be in contradiction with observations. However, this problem is remedied if one accepts that one never has access to the microscopic density. All that we can hope for in real observations is to find a distribution for some coarser, *effective* degrees of freedom. Indeed, the apparent paradox is resolved if one adapts our initial point of view, in which the system consists of observable and thus operationally accessible degrees of freedom, *cf.* Ref. [Pen70; Rue99].

In the following, we reserve the term “Gibbs entropy”  $S^G$  for the entropy obtained by a *maximum entropy principle*. More precisely, we say that  $\rho \equiv \rho(\{(a_i, \varphi_i)\})$  is *compatible* with the macroscopic constraints  $\{(a_i, \varphi_i)\}$ , if for the observables  $\{\varphi_i: \Gamma \rightarrow \mathbb{R}\}$  one has

$$\langle \varphi_i \rangle := \int_{\Gamma} \varphi_i \rho \, dx = a_i, \quad \forall i. \quad (2.14)$$

In that case, the Gibbs entropy specified by  $\{(a_i, \varphi_i)\}$  is defined as

$$S^G := \sup_{\rho'} H[\rho'(\{(a_i, \varphi_i)\})], \quad (2.15)$$

where the supremum is taken with respect to all compatible ensembles  $\rho'(\{(a_i, \varphi_i)\})$ . Often the supremum is given by a unique ensemble  $\rho^G(\{(a_i, \varphi_i)\})$ , which we will call the *Gibbs ensemble* or *Gibbs distribution*.

Hamilton’s equations of motion are appealing because they constitute a microscopic theory derived from first principles. However, besides the paradox of constant entropy they suffer another huge practical problem: For macroscopic physical systems the number of particles,  $N \sim 10^{23}$ , is very large and makes computations hard. The problem also does not vanish if we consider much smaller, *mesoscopic*<sup>8</sup> systems. Such systems are usually immersed in some solvent (*e.g.* water) and Hamiltonian dynamics requires us to treat this environment explicitly.

Thus, treating meso- or macroscopic systems in thermodynamic environments with the microscopic equations of motion (2.8) is a challenging task. Even with state-of-the-art

<sup>8</sup> Usually, for the *mesoscopic range* one considers typical molecular scales (less than 10 nm) and typical macroscopic scales (larger than 10  $\mu\text{m}$ ) as lower and upper boundaries, respectively. Because of this wide range, we prefer to define the term with respect to the modelling paradigm, *cf.* Sec. 1.1.

supercomputers, simulations of no more than a few ( $10^4$  to  $10^6$ ) particles on small time scales ( $10^2$  ns to  $10^4$  ns) are possible. Hence, developing and applying effective dynamical models with fewer degrees of freedom is a major subject of modern physics. In the next two sections, we review modelling paradigms for systems in thermodynamic environments. We start with models based on stochastic processes, which have their origins already in the beginning of the twentieth century. After that, we focus on deterministic models used in modern molecular dynamics simulations.

## 2.4. The models of stochastic thermodynamics

The first stochastic models were introduced as a theoretical framework to study the phenomenon of *Brownian motion* in systems at or close to equilibrium. Brownian motion provides an archetypal example of the dynamics of systems in thermodynamic environments. As we will see shortly, already the study of the thermodynamic aspects of such a simple system yields important physical results. The most famous one is the so-called Einstein relation which connects microscopic fluctuations to macroscopic dissipation.

*Stochastic thermodynamics* is the area of statistical physics that seeks such relations for increasingly complex systems in non-equilibrium environments [Sei08]. Already in the middle of 20th century, stochastic models were formulated for a variety of (non-equilibrium) phenomena in many disciplines of science [VK92]. They all have in common that the statistically random forces on the system exerted by the environment are modelled using stochastic terms. For small systems like biomolecules in solution, these forces lead to notable *fluctuations* in the system's dynamics.

Hill and Schnakenberg pioneered a thermodynamic interpretation of non-equilibrium steady states of master equations [Hil77; Sch76]. In particular, they proposed a general relation between abstract notions of entropy production for stochastic processes and thermodynamic dissipation. These early considerations were based on the temporal evolution of an ensemble as specified by the master equation. More recently, authors started to discuss notions of entropy and entropy production for individual *realizations* of stochastic processes [Kur98; LS99]. This idea led to the unification of a variety of fundamental non-equilibrium *fluctuation relations* (FR) concerning the probability distributions of heat, work and entropy production [Mae04; Sei05]. Here, we only briefly discuss stochastic FR in Section 2.4.3. For a review on the general theory, we refer the reader to Ref. [Sei12].

### 2.4.1. Langevin and Fokker–Planck equations

The first stochastic models were introduced in the beginning of the twentieth century by Einstein [Ein05], Langevin [Lan08] and Smoluchowski [Smo15]. Their goal was to model the diffusion of a relatively heavy *tracer particle* surrounded by a large number of much lighter particles. One usually refers to the tracer particle as performing *Brownian motion* in its fluid environment. Today we know that every fluid, though it might appear

## 2. Notions of entropy and entropy production

as continuous, is made out of particles. Further, we understand Brownian motion as the result of the irregular forces that the lighter particles exert on the tracer. Hence, Brownian motion can be understood as a kind of shot noise, *i.e.* an erratic behaviour that has its origin in the granularity of matter. However, at the end of the 18th century the atomistic view had not been generally accepted. Einstein emphasized that the success of the theory of Brownian motion gives an estimation of Avogadro's number and thus confirms the existence of molecules [Ein05].

### Brownian motion

We start by illustrating the ideas of stochastic models in the framework of Brownian motion. The mathematics are essentially the same for more general situations. For a comprehensive review of stochastic thermodynamics, we direct the reader to Ref. [Sei12].

Consider a particle with position  $q$  and velocity  $\dot{q}$  in a fluid environment. The particle is subject to conservative forces  $F^{\text{cons}} = -\partial_q V$  and a (Stokes) drag force  $F^{\text{drag}} = -\zeta \dot{q}$ , where  $\zeta$  denotes a phenomenological drag coefficient. Further, we consider a microscopic noise term  $\xi$  to model the collisions of the tracer with the fluid molecules.

In the *overdamped limit* one assumes that accelerations are immediately damped away by the environment. Hence, the macroscopic forces balance, *i.e.*  $F^{\text{cons}} + F^{\text{drag}} = 0$  and thus  $\dot{q} = F^{\text{cons}}/\zeta$ . To this *macroscopic* equation of motion we add the noise  $\xi$  to obtain the *overdamped Langevin equation*:

$$\dot{q} = \frac{-\partial_q V}{\zeta} + \xi \quad (2.16)$$

A common assumption (which we will adopt here) is that  $\xi$  obeys the statistics of *white noise*. White noise is uncorrelated with zero mean and variance  $2D$ . More precisely, the averages realization of the stochastic force at time  $t$  obey

$$\langle \xi(t) \rangle_t = 0, \quad \langle \xi(t)\xi(0) \rangle_t = 2D\delta(t), \quad (2.17)$$

where  $\delta(t)$  denotes the Dirac  $\delta$ -distribution.

For Langevin dynamics, the average of an observable  $\varphi: \Gamma \rightarrow \mathbb{R}$  can be written as an integral over a probability density  $\varrho_t$ :

$$\langle \varphi \rangle_t := \int_{\Gamma} \varphi \varrho_t dx.$$

The density  $\varrho_t$  specifies a time-dependent ensemble. For white noise, its evolution is governed by the *Smoluchowski equation* [Smo15]:

$$\partial_t \varrho_t = -\partial_q j_t. \quad (2.18a)$$

Probability conservation is guaranteed by this equation, as the right-hand side amounts



to the divergence of the *instantaneous probability current*

$$j_t := \frac{-\partial_q V}{\zeta} \varrho_t - D \partial_q \varrho_t. \quad (2.18b)$$

The first contribution to the probability current is associated with the macroscopic force balance. It thus expresses the macroscopic *drift*. The second term is an undirected diffusive current which is determined by the strength of the noise  $D$ . For dilute systems, the probability density  $\varrho$  can also be understood as a particle density. If the current  $j_t$  in Equations (2.18) is interpreted in that way, then  $D$  is called a *diffusion constant*.

Equilibrium is defined as a steady state ( $\partial_t \varrho_t = 0$ ) where probability currents vanish:

$$j_t \equiv 0. \quad (2.19)$$

In that case one also says that the system obeys *detailed balance*. For the current in Equation (2.18b), the equilibrium condition (2.19) yields

$$0 = - \left( \frac{\partial_q V}{\zeta} + D \partial_q \right) \varrho_t.$$

Consistency with statistical mechanics requires that the equilibrium probability density amounts to a Boltzmann-distribution<sup>9</sup>, i.e.  $\varrho(q) \propto \exp \frac{-V(q)}{T}$ . Hence, we get

$$D = \frac{T}{\zeta}, \quad (2.20)$$

where  $T$  is the temperature of the (isothermal) system. In the context of Brownian motion one usually uses a Stokes drag constant  $\zeta = 6\pi\eta R$ , where  $R$  denotes the radius of the particle and  $\eta$  is the dynamic viscosity of the fluid. In that case, (2.20) is the so-called Smoluchowski–Einstein fluctuation-dissipation relation (FDR)

$$D = \frac{T}{6\pi\eta R}. \quad (2.21)$$

It thus relates the erratic motion of the tracer particle in equilibrium (diffusion) to the *linear response* of the system to an externally applied force (drag).

A general connection between equilibrium fluctuations and the response to externally applied (small) forces is the *fluctuation-dissipation theorem* [CW51]. For systems close to equilibrium, this theorem implies a linear response, which results in the purely exponential decay of fluctuations.

Another example of a linear response result close to equilibrium are the Onsager relations [Ons31]. They are statements about the *thermodynamic current*  $J_\alpha$  induced by a (small) *thermodynamic force* or *affinity*  $A_\alpha$ . The index  $\alpha$  distinguishes between the different driving mechanisms, because there may be multiple forces acting on the same system. The driving forces are either external forces (like electric fields) or spatial gradients of

---

<sup>9</sup>Note that  $k_B \equiv 1$ .

## 2. Notions of entropy and entropy production

intrinsic thermodynamic variables (like temperature or chemical potential). With the matrix of transport coefficients  $\mathbb{L}_{\alpha\beta}$ , the *Onsager relations*

$$A_\alpha = \sum_{\beta} \mathbb{L}_{\alpha\beta} J_\beta \quad (2.22)$$

provide a prime example of a linear-response relation. In most cases we also have *reciprocity*, which means that the Onsager coefficients are symmetric, *i.e.*  $\mathbb{L}_{\alpha\beta} = \mathbb{L}_{\beta\alpha}$ .

Above, we have derived the Smoluchowski-Einstein FDR, from the *thermodynamic consistency argument*, namely the assumption of a Boltzmann distribution. In general, linear response theory close to equilibrium follows from a more general thermodynamic consistency assumption called *local equilibrium*. We will discuss local equilibrium in more detail below.

### Entropies for the system and the medium

In order to identify entropies and entropy changes in the system and the medium we follow Seifert's work [Sei05; Sei12]. In agreement with the general prescription (2.6), the system's entropy is the differential entropy of the ensemble:

$$S_t^{\text{sys}} := - \int_{\Gamma} \rho_t \log \rho_t \, dq.$$

The instantaneous entropy change of the system is its time-derivative

$$\delta_t S^{\text{sys}} := \partial_t S_t^{\text{sys}}.$$

Denoting the change in the medium by  $\delta_t S^{\text{med}}$  and the total change by  $\delta_t S^{\text{tot}}$ , it splits into two contributions:

$$\delta_t S^{\text{sys}} = \delta_t S^{\text{tot}} - \delta_t S^{\text{med}}. \quad (2.23a)$$

With the introduction of the velocity distribution,  $v_t := \frac{j_t}{\rho_t}$  one finds that [Sek98; Sei12]

$$\delta_t S^{\text{med}} = - \int_{\Gamma} \frac{v_t F^{\text{cons}}}{T} \rho_t \, dq, \quad (2.23b)$$

$$\delta_t S^{\text{tot}} = \int_{\Gamma} \frac{j_t^2}{D \rho_t} \, dq = \frac{\langle v_t^2 \rangle_t}{D} \geq 0. \quad (2.23c)$$

The thermodynamic interpretation is straightforward: In the overdamped limit, any work performed in a potential  $V$  is immediately dissipated. The ensemble average of the instantaneous dissipated heat  $\delta_t Q^{\text{med}}$  is thus the associated power  $\delta_t Q^{\text{med}} = v_t F^{\text{cons}}$ . Under isothermal conditions, the entropy change in the medium is the heat  $Q$  divided by temperature  $T$ . The total entropy is always positive and can be written in a form which is well-known from transport theory [GM84].

In this interpretation, the relations (2.23) yield a differential form of the second law (2.1):

$$\delta_t S^{\text{tot}} = \delta_t S^{\text{sys}} + \delta_t S^{\text{med}} \geq 0.$$

### Underdamped motion and generalizations

The Langevin equation is easily formulated for more general situations. In fact, the original Langevin equation was formulated as an underdamped equation [Lan08]. In that case, the macroscopic equation is Newton's second law  $\dot{p} = F^{\text{tot}} = F^{\text{drag}} + F^{\text{cons}}$ , where  $p = m\dot{q}$  is the momentum of the particle with mass  $m$ . Again, by adding a noise term to model the irregular microscopic forces we obtain:

$$\dot{q} = \frac{p}{m}, \quad (2.24a)$$

$$\dot{p} = -\partial_q V - \frac{\zeta}{m} p + \xi. \quad (2.24b)$$

Here, the strength of the noise fulfils a different fluctuation-dissipation relation, which can be found from demanding a Maxwell–Boltzmann distribution for the momenta. Further generalizations consider multiple interacting particles in more spatial dimensions. Because the evolution equation for the probability density retains the form of a linear advection-diffusion equation similar to Eq. (2.18), one can at least formally solve it.

In practice, one is often interested in observable *collective degrees of freedom*, like the hydrodynamic modes of a continuous density. In order to obtain equations for these *effective* degrees of freedom, one applies approximations at some point which turn the high-dimensional linear equation for the probability distribution into a lower dimensional form. Unfortunately, the linear form of the evolution equation is usually lost [Zwa61; Mor65].

However, there are situations where a linear description of the evolution equation is still appropriate. In that case, one can formulate a generalized version of the Langevin equation also for the collective degrees of freedom. Common examples of such collective degrees of freedom are *reaction coordinates* in biochemical systems or *order parameters* in the physics of condensed matter [CL00].

The most general formulation of a classical Langevin equation for an arbitrary set of degrees of freedom  $\omega$  reads [Sei12]:

$$\partial_t \omega = \mathbb{M} \left( -\nabla_\omega V(\omega) + F^{\text{diss}}(\omega) \right) + \xi. \quad (2.25)$$

The mobility tensor  $\mathbb{M}$  summarizes the appropriate phenomenological transport coefficients. If  $\mathbb{M} = \zeta^{-1} \mathbb{1}$  has scalar form, then it is the inverse of the drag coefficient  $\zeta$ .

Dissipative forces  $F^{\text{diss}}$  may include drag and other non-conservative forces. Again, the microscopic noise term has white noise statistics and needs to be connected with

## 2. Notions of entropy and entropy production

macroscopic transport properties. More precisely, the noise correlations obey

$$\langle \xi_i(t) \xi_j(0) \rangle_t = 2T \mathbb{M}_{ij} \delta(t), \quad (2.26)$$

where  $\xi_i(t)$  denotes the component of the noise associated with the  $i$ th component of  $\omega$  and the numbers  $\mathbb{M}_{ij}$  are the entries of the (positive semi-definite) mobility tensor.

White noise ensures a linear evolution equation for the probability densities  $\rho_t(\omega)$ . The resulting partial differential equation is called the *Fokker–Planck equation*:

$$\begin{aligned} \partial_t \rho_t(\omega) &= -\nabla_\omega \cdot j_t(\omega) \\ &:= -\nabla_\omega \cdot \left( \mathbb{M} \left( -\nabla_\omega V(\omega) + F^{\text{diss}}(\omega) \right) \rho_t(\omega) - T \mathbb{M} \nabla_\omega \rho_t(\omega) \right). \end{aligned} \quad (2.27)$$

In (experimental) applications, often time-dependent (*e.g.* oscillatory) forces are used to probe the response of the system and hence determine the transport coefficients  $\mathbb{M}_{ij}$  [CL00]. Stochastic thermodynamics with explicitly time-dependent forces is thoroughly reviewed in Ref. [Sei12].

### Local equilibrium

One of the crucial assumptions in stochastic thermodynamics is *local equilibrium* (LE). In order to appreciate its meaning in the present context, consider the following situation: An experimenter takes measurements on a many-particle system. The possible measurement outcomes are the values  $\omega \in \Omega$  of an *observable*  $M: \Gamma \rightarrow \Omega$  characterizing the measurement process. Let us further assume that the dynamics on the level of the collective variables  $\omega = M(x)$  are modelled by a generalized Langevin equation (2.25). On that level of description, one ignores the hidden structure and dynamics of the microstates  $x \in \Gamma$ .

In order to connect the stochastic dynamics with statistical physics, one (implicitly) assumes a distribution  $\rho_\omega(x)$  for those microstates  $x$  which yield a certain measurement result  $M(x) = \omega$ . The energy associated with a state  $\omega$  is thus an *internal energy* obtained as the conditioned expectation value of a (Hamiltonian) energy function  $\mathcal{H}(x)$ . Besides the average internal energy, the distribution  $\rho_\omega$  also specifies an internal entropy. The forces  $F^{\text{cons}} = -\nabla_\omega V(\omega)$  must therefore include *entropic forces* as well. The latter arise from the fact that different values of  $\omega$  might be compatible with a different number of microscopic states.

In order to make the notion of compatibility precise, we formulate an ansatz for the constrained microscopic distribution  $\rho_\omega(x)$ . The *internal* entropy is then given as  $S_\omega^{\text{int}} = H[\rho_\omega]$ . Local equilibrium specifies the value of this entropy, by making assumptions about the distribution  $\rho_\omega(x)$ . Similar to the constraints on the strength of the noise in the formulation of Langevin equations, local equilibrium is a consistency assumption: It demands consistency with the treatment of microstates in the statistical physics of equilibrium systems. Further, it ensures that quantities like the potential  $V(\omega)$  of a collective variable  $\omega$  are well defined. Note that here “local” refers to the association with a certain

value  $\omega$  rather than a spatial localization.

One of the necessary requirements for local equilibrium is the existence of a separation of time scales between effective mesoscopic degrees of freedom  $\omega$  and the microscopic degrees of freedom  $x$ : The dynamics of the collective variables are assumed to evolve on a typical time scale  $\tau_{\text{mes}}$  which is much larger than the microscopic time scale  $\tau_{\text{mic}}$ . Consequently, one expects that the distribution  $\rho_\omega$  of the microstates  $x$  has effectively relaxed to an equilibrium distribution. More precisely,  $\rho_\omega = \rho^G(\omega, M)$  is assumed to be the Gibbs distribution compatible with the mesoscopic value  $\omega$  and any further thermodynamic constraints. In particular, the internal entropy of a state  $\omega$  is assumed to have the value of the corresponding Gibbs entropy. Then, the potential  $V(q)$  should be understood as *thermodynamic potential* like a *free energy*.

Note that the separation of time scales implicitly enters the Langevin equation (2.25) through the assumption of white noise (2.26). For the example of Brownian motion, the  $\delta$ -correlations are approximations to the real collision statistics, which have a small (but finite) relaxation time  $\tau_{\text{mic}}$ . We assume that the temporal resolution for the observation of the motion of the heavy colloid is much larger than a microscopic time scale  $\tau_{\text{mic}}$ . The time scale  $\tau_{\text{mic}}$  also determines the decay time of microscopic–mesoscopic correlations.

In the context of bio-chemical systems, collective variables are usually chosen to reflect experimentally accessible observations. Such variables may describe the configuration and chemical composition of macromolecules. In that context, they are also known as reaction coordinates. As an example, consider the configurational changes associated with the folding of proteins. They occur on time scales  $\tau_{\text{conf}} \approx 10^{-4} \text{ s}$  to  $10^{-4} \text{ s}$ , which is much larger than the time scales of the microscopic constituents and the solvent.

This separation of time scales and the resulting quick loss of correlations, is called the *Markov property*. It means that the future of a state only depends on its current state. Neither the system nor its environment keeps memory of the system's evolution in the past. This is explicitly visible in the Fokker–Planck equation (2.27): Both the conservative and the dissipative forces only depend on the instantaneous value of  $\omega$ . The Fokker–Planck (or Smoluchowski) equation is a certain time- and space-continuous form of a *master equation*. It is appropriate if  $\omega \in \Omega$  takes continuous values. However, often it is enough to consider a discrete space of observations  $\Omega$ .

### 2.4.2. Master equations

The master equation is the discrete-time version of the Fokker–Planck equation. We have already seen that the Fokker–Planck equation specifies a Markovian (*i.e.* memoryless) evolution on a continuous phase space. In this section, we consider time-homogeneous Markov processes on a finite state space  $\Omega = \{1, 2, \dots, N\}$ . The ensemble at time  $t$  is specified by a probability vector  $\mathbf{p}_t$ .

## 2. Notions of entropy and entropy production

The *master equation* for its evolution in continuous time reads

$$\partial_t \mathbf{p}_t = \mathbf{p}_t \mathbb{W}, \quad (2.28)$$

where  $\mathbb{W}$  is a *rate matrix* containing the transition rates  $w_{\omega'}^{\omega}$ . For  $\omega \neq \omega'$  they obey  $w_{\omega'}^{\omega} \geq 0$ . The diagonal entries of  $\mathbb{W}$  amount to  $w_{\omega}^{\omega} := -\sum_{\omega' \neq \omega} w_{\omega'}^{\omega}$ , because probability needs to be conserved.

In Chapters 3 and 4, we are more interested in the time-discrete case of *Markov chains*. The master equation for Markov chains reads:

$$\mathbf{p}_{t+1} = \mathbf{p}_t \mathbb{W}, \quad (2.29)$$

where the entries  $\mathbb{W}$  obey  $0 \leq w_{\omega'}^{\omega} \leq 1$  and  $\sum_{\omega'} w_{\omega'}^{\omega} = 1$ . A matrix that satisfies these conditions is called a *stochastic* or *transition matrix*.

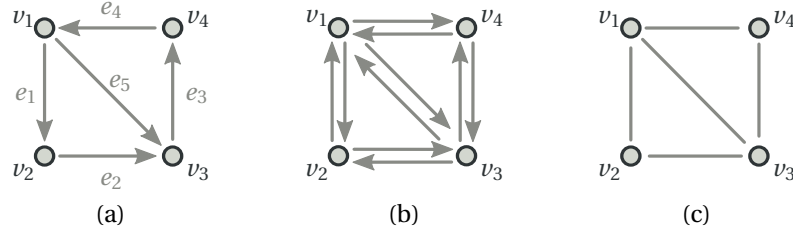
For both continuous and discrete-time Markov chains we define the *probability flux*  $\phi$  from state  $\omega$  to state  $\omega'$  as

$$\phi_{\omega'}^{\omega}(t) := p_t(\omega) w_{\omega'}^{\omega}, \quad \omega \neq \omega'. \quad (2.30)$$

The difference of the forward and backward flux is the *probability current*  $J$  with entries

$$J_{\omega'}^{\omega}(t) := \phi_{\omega'}^{\omega}(t) - \phi_{\omega}^{\omega'}(t). \quad (2.31)$$

### The graph of the network of states



**Figure 2.2.:** The network of states as a graph. a) A directed graph  $G^d$  consisting of four vertices  $v_i$  and five edges  $e_i$ . b) The directed graph  $G^d$  of a dynamically reversible Markovian jump process. c) By identifying the forward and backward edges  $(\omega, \omega')$  and  $(\omega', \omega)$  in b) with the unordered pair  $\{\omega, \omega'\}$  we obtain an undirected graph  $G^u$ . Note that the directed graph in a) is an oriented version of  $G^u$ , where we pick an arbitrary orientation for each undirected edge.

For both continuous and discrete time dynamics on finite state spaces, we have a visual representation of the *network of states* as a *graph*  $G$ . A directed graph  $G^d = (V, E)$  is a finite collection of *vertices*  $V$  and a set of tuples  $E \subset V \times V$ , which are called *edges*. In our case, we choose  $V = \Omega$  and demand that  $(\omega, \omega') \in E$  if and only if  $w_{\omega'}^{\omega} > 0$ . Its adjacency matrix  $\mathbb{A}$  is a Boolean matrix (*i.e.* a matrix consisting of zeros and ones) with an entry  $a_{\omega'}^{\omega} = 1$  whenever  $(\omega, \omega') \in E$ .

A graph can be easily drawn, *cf.* figure 2.2a. Vertices  $\omega \in V$  are represented by points and an arrow pointing from  $\omega$  to  $\omega'$  is drawn if  $(\omega, \omega') \in E$ . The visual representation of the

state space as a graph provides an intuition for *paths* and *cycles*. A directed path  $\gamma^{(\tau)}$  of length  $\tau$  is a sequence of successive states  $\gamma^{(\tau)} = (\omega_0, \omega_1, \dots, \omega_{\tau-1}, \omega_\tau)$  that obey

$$(\omega_{i-1}, \omega_i) \in E, \quad 0 < i \leq \tau.$$

Further, we require that no *internal state*  $\omega_i$  with  $0 < i < \tau$  appears multiple times in a path  $\gamma^{(\tau)}$ . A cycle  $\alpha^{(\tau)}$  of length  $\tau$  is a closed path where  $\omega_0 = \omega_\tau$ .

A graph is said to be *connected*, if it does not consist of multiple disconnected parts. We say that it is *strongly connected*, if there exists a directed path between any two vertices. If a graph is strongly connected, its adjacency matrix is called *irreducible*. We use the same terminology to refer to the transition matrix  $\mathbb{W}$ , which yields an adjacency matrix  $\mathbb{A} = |\text{sgn } \mathbb{W}|$ . A Markov chain or a Markovian jump process on a finite state space with an irreducible adjacency matrix is also called *ergodic*.

For continuous-time Markov processes, ergodicity implies the existence of a *unique* invariant distribution  $\mathbf{p}_\infty$ . For time-discrete Markov chains uniqueness of  $\mathbf{p}_\infty$  additionally requires that the transition matrix  $\mathbb{W}$  is *aperiodic*. Aperiodicity means that returns to any state are possible at arbitrary, but sufficiently large time-differences  $t$ . Formally, we demand that there is a  $\tau \in \mathbb{N}$  such that for all  $t > \tau$  we have

$$(\mathbb{W}^t)_{\omega, \omega} > 0, \quad \forall \omega \in \Omega, \quad (2.32)$$

where  $(\mathbb{W}^t)_{\omega, \omega}$  denotes the diagonal entries of the  $t$ th power of  $\mathbb{W}$ . One also says that an irreducible and aperiodic Markov chain is *mixing*.

For the rest of this section we assume *dynamical reversibility*, i.e.  $w_{\omega'}^\omega > 0 \Leftrightarrow w_\omega^{\omega'} > 0$ . This means, that the network of states is a simply connected, bi-directional graph. In that case, we can also draw an undirected graph  $G^u = (V, E^u)$ , where  $E^u$  contains sets of pairs  $\{\omega, \omega'\}$  rather than tuples (cf. figure 2.2c). The physical motivation of dynamic reversibility has its origin in the reversibility of the microscopic (Hamiltonian) dynamics: Assume that the value  $\omega$  of a collective dynamical variable does not depend on the momenta of an underlying microscopic system. Now consider a microscopic trajectory that takes the system from state  $\omega$  to  $\omega'$ . By flipping the momenta, we reverse the direction of time and hence obtain a microscopic trajectory that takes the system from  $\omega'$  to  $\omega$ .

### Stochastic thermodynamics for master equations

For dynamically reversible Markov processes, we define the logarithmic ratio

$$\mathcal{E}_{\omega'}^\omega := \log \frac{w_{\omega'}^\omega}{w_\omega^{\omega'}}. \quad (2.33)$$

for both continuous- and discrete-time dynamics.

Because of an analogy with electrical networks (which we will discuss in Chapter 5), we refer to  $\mathcal{E}_{\omega'}^\omega$  as the *motance* of a transition  $\omega \rightarrow \omega'$ . The integrated motance  $\mathcal{E}_\alpha$  of a

## 2. Notions of entropy and entropy production

cycle  $\alpha = \alpha^{(\tau)}$  is defined as

$$\mathcal{E}_\alpha := \sum_{i=1}^{\tau} \mathcal{E}_{\omega_i}^{\omega_{i-1}}. \quad (2.34)$$

By analogy to vector calculus we say that the motance is *conservative*, if it vanishes along any cycle:

$$\mathcal{E}_\alpha = 0, \forall \alpha.$$

Then, the integrated motance depends only on the initial and final state of any path  $\gamma^{(\tau)}$ . As a consequence one can write the motance between two states as the difference  $\Delta U_{\omega'}^\omega$  of a potential  $U_\omega$  defined on the states. A formal treatment of this analogy will be the topic of Chapter 5. An equivalent condition for a conservative motance is that the so-called *Kolmogorov cycle criterion* holds for every closed path  $\underline{\omega}^{(\tau)}$  with  $\omega_0 = \omega_\tau$ :

$$\prod_{i=1}^{\tau} w_{\omega_i}^{\omega_{i-1}} = \prod_{i=1}^{\tau} w_{\omega_{i-1}}^{\omega_i}. \quad (2.35)$$

For the moment, let us focus on the steady state where we have an invariant probability distribution  $\mathbf{p}_t = \mathbf{p}_\infty$  with  $\partial_t \mathbf{p}_\infty = 0$ . In that case, fluxes  $\phi_{\omega'}^\omega$  and currents  $J_{\omega'}^\omega$  are time-independent. For the steady state, we define the *edge affinity*

$$A_{\omega'}^\omega := \log \frac{\phi_{\omega'}^\omega}{\phi_{\omega}^{\omega'}} \equiv \log \frac{p_\infty(\omega) w_{\omega'}^\omega}{p_\infty(\omega') w_\omega^{\omega'}} \quad (2.36)$$

as the logarithmic ratio of forward and backward fluxes in the steady state. Because the steady state probabilities cancel in the logarithmic ratio, the affinity  $A_\alpha$  of a cycle of length  $\tau$  corresponds to the motance of that cycle:

$$\begin{aligned} A_\alpha &:= \sum_{t=1}^{\tau} \log \frac{p_\infty(\omega) w_{\omega'}^\omega}{p_\infty(\omega') w_\omega^{\omega'}} \\ &= \sum_{t=1}^{\tau} \log \frac{w_{\omega'}^\omega}{w_\omega^{\omega'}} = \mathcal{E}_\alpha \end{aligned} \quad (2.37)$$

The affinities and motances of cycles and edges play an important role in the early formulations of stochastic thermodynamics by T. Hill [Hil77] and J. Schnakenberg [Sch76]. In analogy to the discussion of the Fokker–Planck equation, we say that the system is in *equilibrium* if all steady state currents vanish identically:

$$J_{\omega'}^\omega = 0, \forall \omega, \omega'.$$

The detailed balance condition can be written using the fluxes:

$$\phi_{\omega'}^\omega = \phi_\omega^{\omega'}.$$



In a steady state, the master equation can be rewritten as

$$\sum_{\omega'} J_{\omega'}^{\omega} = 0.$$

We discuss an interpretation of this equation in more detail in Chapter 5. Whether a system will exhibit an equilibrium or a non-equilibrium steady state (NESS), depends on the elements of  $\mathbb{W}$ . The Kolmogorov criterion (2.35) provides a necessary and sufficient condition for a system to relax to an equilibrium steady state [Sch76].

Initially, let us discuss continuous-time master equations. Hill was the first one to attempt a thermodynamic interpretation of the master-equation framework. He noticed the importance of cycles for non-equilibrium steady states [Hil77] which motivated Schnakenberg's network theory [Sch76]. These authors also realized that the cycle affinities  $A_{\alpha}$  obtain the values of (linear combinations) of the thermodynamic forces acting on the system. Moreover, cyclic currents  $J_{\alpha}$  are the analogues of the thermodynamic currents for these systems. Schnakenberg also found that cycle currents and affinities close to equilibrium obey

$$A_{\alpha} = \sum_{\beta} \mathbb{L}_{\alpha\beta} J_{\beta}, \quad (2.38)$$

with positive symmetric coefficients  $\mathbb{L}_{\alpha\beta} = \mathbb{L}_{\beta\alpha}$ . Hence, Eq. (2.38) is another manifestation of Onsager's linear response relations (2.22).

This in turn motivates the identification of the bilinear expression

$$\delta S_{\infty}^{\text{tot}} = \sum_{\alpha} A_{\alpha} J_{\alpha} \quad (2.39)$$

as the total entropy production in the steady state. The sum in Eq. (2.39) runs over a set of *fundamental cycles* defined in Schnakenberg's network theory. We will discuss the abstract algebraical features of master equations in more detail in Chapter 5. For now, we just mention that we can formulate Eq. (2.39) also using edge currents and edge affinities:

$$\delta S_{\infty}^{\text{tot}} = \frac{1}{2} \sum_{\omega, \omega'} A_{\omega'}^{\omega} J_{\omega'}^{\omega} \geq 0 \quad (2.40)$$

The factor  $\frac{1}{2}$  is needed to avoid double-counting of edges. Positivity immediately follows from the fact that  $\text{sgn}(A_{\omega'}^{\omega}) = \text{sgn}(J_{\omega'}^{\omega})$ , cf. Equations (2.36) and (2.31).

Schnakenberg's considerations for the steady state can be generalized to the transient case. We canonically identify the system's entropy as

$$S_t^{\text{sys}} := - \sum_{\omega \in \Omega} p_t(\omega) \log p_t(\omega). \quad (2.41)$$

## 2. Notions of entropy and entropy production

With the master equation (2.28), the time-derivative of the system's entropy reads

$$\partial_t S_t^{\text{sys}} = \sum_{\omega, \omega'} p_t(\omega) w_{\omega'}^\omega \log \frac{p_t(\omega)}{p_t(\omega')} \quad (2.42a)$$

It can be split into two contributions  $\partial_t S_t^{\text{sys}} = \delta S_t^{\text{tot}} - \delta S_t^{\text{med}}$ :

$$\delta S_t^{\text{med}} := \sum_{\omega, \omega'} p_t(\omega) w_{\omega'}^\omega \log \frac{w_{\omega'}^\omega}{w_{\omega'}^{\omega'}}, \quad (2.42b)$$

$$\begin{aligned} \delta S_t^{\text{tot}} &:= \sum_{\omega, \omega'} p_t(\omega) w_{\omega'}^\omega \log \frac{p_t(\omega) w_{\omega'}^\omega}{p_t(\omega') w_{\omega'}^{\omega'}} \\ &= \frac{1}{2} \sum_{\omega, \omega'} \left[ \left( \phi_{\omega'}^\omega(t) - \phi_{\omega'}^{\omega'}(t) \right) \log \frac{\phi_{\omega'}^\omega(t)}{\phi_{\omega'}^{\omega'}(t)} \right] \geq 0. \end{aligned} \quad (2.42c)$$

Equation (2.42c) is the transient version of (2.40), and hence we identify  $\delta S_t^{\text{tot}}$  with the transient generalization of the total (instantaneous) entropy variation.<sup>10</sup>

From the second line of the equality, it is immediately clear that this quantity is positive for all  $t \geq 0$ . Positivity is also evident because  $\sum_{\omega, \omega'} [p_t(\omega) w_{\omega'}^\omega] = \sum_{\omega, \omega'} [p_t(\omega') w_{\omega'}^{\omega'}] < \infty$ . Then, the term  $\delta S_t^{\text{tot}}$  is a Kullback–Leibler divergence of two normalized distributions and hence always positive.

The right-hand side of Equation (2.42b) is the average of the motance  $\mathcal{E}$  over all jumps  $\omega \rightarrow \omega'$  that appear at time  $t$ . It involves the logarithmic ratio of forward and backward transitions, and thus provides an example of the relation (2.7). Consequently, one identifies  $\delta S_t^{\text{med}}$  as the instantaneous variation of the entropy in the medium [Sei05].

### The time-discrete case

In the last subsection we have introduced the instantaneous change of the system's and the medium's entropy for continuous-time Markov processes. We will also need the expressions for the discrete-time case. Using the discrete-time master equation (3.2), we split the variation

$$\Delta S_t^{\text{sys}} = S_t^{\text{sys}} - S_{t-1}^{\text{sys}} \quad (2.43a)$$

of the system entropy (2.41) into  $\Delta S_t^{\text{sys}} = \Delta S_t^{\text{tot}} - \Delta S_t^{\text{med}}$  using the expressions

$$\Delta S_t^{\text{med}} := \sum_{\omega, \omega'} p_{t-1}(\omega) w_{\omega'}^\omega \log \frac{w_{\omega'}^\omega}{w_{\omega'}^{\omega'}}, \quad (2.43b)$$

$$\Delta S_t^{\text{tot}} := \sum_{\omega, \omega'} p_{t-1}(\omega) w_{\omega'}^\omega \log \frac{p_{t-1}(\omega) w_{\omega'}^\omega}{p_t(\omega') w_{\omega'}^{\omega'}}. \quad (2.43c)$$

<sup>10</sup> Even without the thermodynamic considerations above, the expressions (2.42) also arise naturally when treating non-equilibrium stochastic thermodynamics as a *gauge theory* based on information-theoretical considerations [Pol12]. More precisely, the gauge corresponds to the above-mentioned choice of a reference measure when defining entropies. Then, the expression for the (instantaneous) *total* entropy production is the simplest non-trivial gauge invariant term associated with changes in entropy.

In the continuous-time limit, the expressions (2.43) yield their continuous-time analogues (2.42). Again, the motance of a transition determines the temporal variation of the entropy in the medium  $\Delta S_t^{\text{med}}$ . Note that we cannot write the variation  $\Delta S_t^{\text{tot}}$  of the total entropy as the sum over positive terms. However, note that independently of time  $t$  we have  $\sum_{\omega, \omega'} [p_t(\omega) w_{\omega'}^{\omega}] = \sum_{\omega, \omega'} [p_{t-1}(\omega') w_{\omega}^{\omega'}] = 1$ . Hence, expression (2.43c) fulfils the properties of a Kullback-Leibler divergence and is thus always positive.

### 2.4.3. Stochastic fluctuation relations

In this section we briefly discuss the notion of fluctuation relations (FR) in stochastic thermodynamics. In their most common form [Sei05], the FR compares the probability  $\mathbb{P}$  of finding a certain value  $a$  of the total entropy change  $\delta^{(\tau)} s_t^{\text{tot}}$  in the interval  $[t, t + \tau]$  with the probability of observing the negative value  $-a$ :

$$\frac{\mathbb{P}[\delta^{(\tau)} s_t^{\text{tot}} = a]}{\mathbb{P}[\delta^{(\tau)} s_t^{\text{tot}} = -a]} = \exp a. \quad (2.44)$$

First results in this direction are Kurchan's [Kur98] and Crooks' [Cro99] relations for Langevin and Fokker–Planck equations. Lebowitz and Spohn formulated a related result for the master equation approach [LS99]. A general framework was suggested by Maes [Mae04] and Seifert [Sei05].

Note that (2.44) is indeed a *detailed* rather than an *integral* formulation of the second law [Sei05]: Rather than talking about averages, it is a statement about probabilities. Because entropy (or heat) is an extensive quantity, for macroscopic systems and macroscopic time scales  $\tau$  the mean value of the total entropy production becomes very large. Hence, for macroscopic systems, Eq. (2.44) states that observing a negative value of the dissipation is not *impossible*, but extremely *improbable*: In principle, heat can be extracted from a single reservoir to perform work on the system. However, the FR states that the probability of such an event is extremely unlikely on macroscopic scales.

All of the fluctuation relations consider single realizations (so-called *noise histories*) of the stochastic dynamics. More precisely, the random variable  $\delta^{(\tau)} s_t^{\text{tot}}$  for the total entropy production depends on a stochastic *trajectory*  $\underline{\omega}^{(\tau)}$ . Consequently, the probability  $\mathbb{P}$  in equation (2.44) is obtained by marginalizing the probability of noise-histories in an interval  $[t, t + \tau]$ . In this thesis, we focus on the trajectory-dependent entropic random variables in the framework of time-discrete Markov chains. For models in continuous time we refer the reader to the review article [Sei12].

In discrete time, the stochastic trajectory  $\underline{\omega}^{(\tau)}$  is a *time series*

$$\underline{\omega}^{(\tau)} = (\omega_0, \omega_1, \dots, \omega_{\tau}) \in \Omega^{\tau+1}. \quad (2.45)$$

For a Markov chain, the probability of seeing a finite sequence of states  $\underline{\omega}^{(\tau)}$  at time  $t$  is

## 2. Notions of entropy and entropy production

given as

$$\mathbb{P}_t[\underline{\omega}^{(\tau)}] = p_t(\omega_0) \prod_{t=1}^{\tau} w_{\omega_t}^{\omega_{t-1}}. \quad (2.46)$$

The probability of a time series of length one is the joint transition probability or probability flux (2.30).

Let  $\varphi_{t_0}^{(1)}(\omega, \omega')$  be a (possibly time-dependent) function. We define the *jump average* as

$$\langle\langle \varphi \rangle\rangle_{t_0}^{(1)} := \sum_{\omega, \omega'} \phi_{\omega'}^{\omega}(t_0) \varphi_{t_0}^{(1)}(\omega, \omega'), \quad (2.47)$$

where  $\phi_{\omega'}^{\omega}(t_0)$  denotes the flux at time  $t_0$ .

Because of the Markovian character of the dynamics, the expressions for the entropy changes depend only on the final transition  $\omega_{\tau-1} \rightarrow \omega_{\tau}$  of  $\underline{\omega}$ . They read

$$\delta s_t^{\text{sys}}[\underline{\omega}^{(\tau)}] := \log \frac{p_{t+\tau-1}(\omega_{\tau-1})}{p_{t+\tau}(\omega_{\tau})}, \quad (2.48a)$$

$$\delta s_t^{\text{med}}[\underline{\omega}^{(\tau)}] := \log \frac{w_{\omega_{\tau}}^{\omega_{\tau-1}}}{w_{\omega_{\tau-1}}^{\omega_{\tau}}}, \quad (2.48b)$$

$$\delta s_t^{\text{tot}}[\underline{\omega}^{(\tau)}] := \log \frac{p_{t+\tau-1}(\omega_{\tau-1}) w_{\omega_{\tau}}^{\omega_{\tau-1}}}{p_{t+\tau}(\omega_{\tau}) w_{\omega_{\tau-1}}^{\omega_{\tau}}}. \quad (2.48c)$$

It is easy to see that for any of these quantities it holds that

$$\sum_{\underline{\omega}^{(\tau)}} \mathbb{P}[\underline{\omega}^{(\tau)}] \delta s_t[\underline{\omega}^{(\tau)}] = \langle\langle \delta s \rangle\rangle_{t+\tau-1}^{(1)} = \Delta S_{t+\tau}$$

where  $\delta s$  and  $\Delta S$  stand for the expressions in Equations (2.48) and (2.43), respectively.

The above definitions can be generalized to the case of random variables that depend on the whole trajectory rather than only on its two last states. In mathematics, such random variables are also known as  $\tau$ -chains. We will use this more general notion in Chapter 4. Further, Appendix A.2 treats  $\tau$ -chains in a more formal way and proves some general results.

## 2.5. Effective deterministic models for molecular dynamics

In the previous section, we have introduced stochastic processes as models for complex systems. We have seen that it is not necessary to model every degree of freedom explicitly. Rather, we can restrict our models to observable (collective) degrees of freedom. The interaction of the system with the medium was modelled using both fluctuating microscopic forces and phenomenological macroscopic coefficients. For the former we assumed stochastic white noise and arrived at a stochastic differential equation.

For the purposes of computer simulations it may be desirable to have a deterministic

description instead.<sup>11</sup> In that case, one speaks of *molecular dynamics* (MD) simulations. The most common form of MD simply uses Hamiltonian dynamics.

Because Hamiltonian dynamics are consistent only with isolated systems, the “environment” of a non-isolated subsystems has to be modelled explicitly. In that case, the large number of particles often render Hamiltonian dynamics unattractive for the purpose of simulations. To deal with that problem, deterministic modifications to the equations of motion have been proposed. Such equations are called *thermostated* equations of motion, because the equations of motion contain the dynamics of an effective “heat bath” or *thermostat*.

### 2.5.1. Thermostated equations of motion

In thermostated MD, one introduces *auxiliary* degrees of freedom which are equipped with their own dynamics instead of stochastic noise. Throughout this section, we follow the systematic approach to deterministic thermostats presented in Ref. [SDC07]. We mostly focus on the thermostated equations that are used for equilibrium molecular dynamics (EMD).

The construction of EMD equations resembles the formulation of a Langevin equation: One starts with a phenomenological equation of motion for the observable degrees of freedom under consideration. In general, such equations include a mobility/drag term to model the dissipative effects of the environment. Then, instead of adding a stochastic noise term, one chooses at least one of the following methods: (i) Promoting the drag coefficient to a dynamical variable and specify an evolution rule for the latter or (ii) adding a deterministic “noise” term which mimics the fluctuations caused by the environment.

The new dynamical variables introduced in that approach are the auxiliary degrees of freedom we mentioned above. As in the stochastic case, the choice of these terms has to be consistent with the thermodynamic properties of the bath. Again, this is achieved by demanding that the stationary distribution  $\rho_\infty$  of the physical degrees of freedom becomes a Gibbs distribution  $\rho^G$  under equilibrium conditions.

#### Nosé–Hoover thermostats

As an example for the construction of thermostated equations of motion, we review Nosé–Hoover scheme. Starting with Newton’s equation for the macroscopic forces, we obtain a system of equations similar to (2.16). However, we do not add a noise term but rather promote the drag  $\zeta$  to a dynamical variable  $\tilde{\zeta}(t)$  whose evolution is specified by a

---

<sup>11</sup> After all, pseudo-random number generators as they are used in stochastic simulations are deterministic algorithms.

## 2. Notions of entropy and entropy production

function  $g(\vec{q}, \vec{p})$ :

$$\dot{\vec{q}} = \frac{\vec{p}}{m}, \quad (2.49a)$$

$$\dot{\vec{p}} = -\nabla_{\vec{q}}V(\vec{q}) - \tilde{\zeta} \frac{\vec{p}}{m}, \quad (2.49b)$$

$$\dot{\tilde{\zeta}} = g(\vec{q}, \vec{p}). \quad (2.49c)$$

In order to find a thermodynamically consistent form of  $g$ , we refer to the steady state. We know that in equilibrium *equipartition* holds for the momenta, *i.e.*

$$\left\langle \sum \frac{\vec{p}^2}{m} \right\rangle_{\infty} = dNT, \quad (2.50)$$

where  $d$  is the dimension of physical space. This motivates a choice of  $g$  that accelerates the system if there is too little kinetic energy in the degrees of freedom and decelerates it otherwise. Hence, one chooses

$$g(\vec{q}, \vec{p}) = \frac{m}{Q} \left( \frac{\vec{p}^2}{m} - dNT \right). \quad (2.51)$$

where  $Q = dNT\tau_{\vec{p}}^2$  is a constant related to the time scale  $\tau_{\vec{p}}$  of the relaxation of the momenta. This choice of  $g$  leads to the following stationary distribution [SDC07]:

$$\rho^{\infty} \propto \exp\left(-\frac{1}{T} \left( V(\vec{q}) + \frac{\vec{p}^2}{2m} + F(\tilde{\zeta}) \right)\right), \quad (2.52)$$

where

$$F(\tilde{\zeta}) = \frac{Q}{2m^2} \tilde{\zeta}^2. \quad (2.53)$$

Momenta and coordinates are distributed according to a Maxwell–Boltzmann distribution, while the statistics of  $\tilde{\zeta}$  are determined by the “potential”  $F(\tilde{\zeta})$ . The purely quadratic form of  $F$  is generic for the following two reasons: Firstly, it leads to an expectation value  $\langle \tilde{\zeta} \rangle_{\infty} = 0$  ensuring that the system is neither accelerated nor decelerated under equilibrium conditions. Secondly,  $\tilde{\zeta}$  is used to model a force that is the outcome of many quasi-independent microscopic contributions. Hence, the central limit theorem states that the distribution should be approximately Gaussian. Note that  $Q$  is the only free parameter. The discussion of the Einstein FDR (2.21) relates the variance of the noise  $2D$  to the phenomenological drag constant  $\zeta$ . If  $\zeta$  is given, this amounts to  $Q \stackrel{\dagger}{=} dNT \frac{m^2}{\zeta^2}$ . Note that the drag defines a relaxation time scale  $\tau_p := \frac{m}{\zeta}$ . Hence, without the reference to a phenomenological drag  $\zeta$  we can understand  $Q$  as defining the time scale  $\tau_p$  via  $Q = dNT\tau_p^2$ .

Note that there are other motivations of the Nosé–Hoover equations of motion. For instance, one can start with the stationary distribution (2.52) and *infer*  $g = g(\vec{q}, \vec{p}, \tilde{\zeta})$  by consistency. For more details we refer the reader to Refs. [SDC07] and [JR10].

### Configurational thermostats

Configurational thermostats are the deterministic analogue to the overdamped Langevin equation (2.16). It is assumed that the momenta have relaxed to their equilibrium values and the dynamics can be described by the equations

$$\dot{\vec{q}} = \tilde{\mu} \nabla_{\vec{q}} V, \quad (2.54a)$$

$$\dot{\tilde{\mu}} = \frac{1}{Q_\mu} \sum \left[ (\nabla_{\vec{q}} V)^2 - T (\nabla_{\vec{q}})^2 V \right]. \quad (2.54b)$$

Note that if one sets  $\tilde{\mu} = \tilde{\zeta}^{-1}$ , equation (2.54a) looks like the macroscopic part of (2.16). As above, the form of the dynamics of  $\tilde{\mu}$  is found by demanding a canonical form for the stationary distribution of the coordinates. Whereas Eq. (2.51) ensured equipartition (2.50), Eq. (2.54b) is consistent with Rugh's notion of a configurational temperature [Rug97]. As above, the constant  $Q_\mu$  can be used to set a relaxation time scale  $\tau_q$  for the positions  $\vec{q}$ .

The problem with Eqs. (2.54) is that the dynamics are not ergodic: Mechanical equilibria  $\nabla_{\vec{q}} V = 0$  act as attracting fixed points where the system comes to rest. To restore ergodicity, further modifications of the equations of motion are required. In one variant of the so-called SDC scheme (after Samoilev, Dettmann and Chaplain [SDC07]) sampling of phase space is enhanced by introducing another auxiliary dynamical variable  $\vec{\xi}$ . This auxiliary variable is used to perturb the dynamics around mechanical equilibria, similar to what the stochastic noise term would do in the Langevin equation. With the addition of  $\vec{\xi}$  the equations of motion are

$$\dot{\vec{q}} = \tilde{\mu} \nabla_{\vec{q}} V + \vec{\xi} \quad (2.55)$$

with the dynamics of  $\tilde{\mu}$  as in Eq. (2.54b) and

$$\dot{\vec{\xi}} = h(\vec{\xi}, \vec{q}). \quad (2.56)$$

Consistency with the Boltzmann distribution requires that  $h$  yields a dynamics that satisfies  $\vec{\xi} \cdot \nabla_{\vec{q}} V = 0$ . There are essentially three different possibilities to satisfy this condition. They correspond to differently constrained fluctuations around mechanical equilibria, which are described in detail in [SDC07]. Moreover, further generalizations can be found there and in Ref. [SDC10]. The latter work focuses on a variant of the general scheme, the so-called Braga–Travis (BT) thermostat [BT05]. So far it is not known if or under which conditions the dynamics created by the SDC thermostating scheme are ergodic. However, numerical simulations using deterministic SDC schemes indicate ergodicity at least in some variants [SDC07].

### Non-equilibrium molecular dynamics

The schemes presented here are useful for thermostated equilibrium MD, *i.e.* the modelling of non-isolated but closed systems. However, there is also the need to model driven (open) non-equilibrium situations, which is the subject of *non-equilibrium molecular*

## 2. Notions of entropy and entropy production

*dynamics* (NEMD). To achieve non-equilibrium conditions one adds additional terms to the thermostated equations of motion. For instance, one can couple the momentum equation to non-conservative forces, which constantly accelerates the system. The thermostat then removes (*i.e.* dissipates) the energy added in that way. We will not go into the details of non-equilibrium molecular dynamics, but refer the reader to the literature, *e.g.* Refs. [Hoo83; EM90; ES02; JR10].

### 2.5.2. Entropy and dissipation in deterministic dynamics

Like for the case of stochastic dynamics, we would like to identify the entropy  $H[\rho_t]$  of a density  $\rho_t(x)$  with the system's entropy at a time  $t$ . However, for deterministic dynamics of the form

$$\dot{x} = f(x) \tag{2.57}$$

there are certain problems with that interpretation.

Firstly, if we attempt to model systems in a thermodynamic environment using thermostated equations, the dynamical variables  $x = (x^{\text{sys}}, \alpha)$  contain auxiliary degrees of freedom  $\alpha$  in addition to the observable ones  $x^{\text{sys}}$ . Hence, the system entropy  $S_t^{\text{sys}}$  should rather refer to the entropy of the marginalized ensemble  $\rho_t|_{x^{\text{sys}}}$ . However, we can only find the dynamical evolution of the joint distribution  $\rho_t$ . For the EMD equations described above, the steady state density is a Gibbs distribution by construction. However, we have no idea of its transient values. For NEMD using modified EMD equations, one even loses such a physical interpretation for the steady-state distribution.

Secondly, even for Hamiltonian dynamics, where no auxiliary variables are present, the Shannon entropy of the ensemble  $\rho$  shows the constant-entropy paradox (2.13).<sup>12</sup> Hence, there is no reason to expect that for more general dynamics such an interpretation would be useful. We will see below that the situation actually gets much worse for non-equilibrium situations.

Although we cannot start with a definition of the system's entropy, one can still establish a connection to thermodynamics. More precisely, in the remainder of this section we discuss the notion of dissipation, which we interpret as the entropy variation in the medium. In particular, we consider its connection to the expansion and contraction of phase space volumes.

#### Hamiltonian dynamics

A first hint comes from Hamiltonian dynamics. In the Boltzmann picture of statistical mechanics, thermodynamic entropy is related to the (logarithm of) accessible phase

---

<sup>12</sup> We have mentioned that the constant-entropy paradox gets resolved if one considers coarse-grained degrees of freedom. A major part of this thesis is concerned with how this generalizes to non-Hamiltonian microscopic dynamics, *cf.* Chapters 3 and 4.



space volume  $\Pi$ :

$$S^B := \log \Pi$$

The Gibbs-Liouville equation (2.10) ensures that phase space volume is conserved for Hamiltonian dynamics. We have already mentioned that Hamilton's equation of motion are consistent with isolated systems. Let us make this statement more precise: The thermodynamic definition of an isolated system is that the entropy exchange with the outside world (*i.e.* the dissipation) vanishes identically. In statistical physics, this statement should hold for isolated systems independent of a particular point in time or any initial configuration. In the discussion of Eq. (2.10) we have introduced the phase space expansion rate

$$\Lambda(x) := \operatorname{div} f,$$

which identically vanishes for Hamiltonian dynamics. Thus, it has (at least some of) the desired properties of a quantity reflecting the entropy exchange with the environment.

### Thermostated dynamics

Let us now look at the thermostated dynamics. To be specific, consider the Nosé–Hoover equations (2.49). There we have

$$\Lambda(x) = \nabla_{\tilde{p}} \cdot \dot{\tilde{p}} = -\frac{\tilde{\zeta}}{m}.$$

In general, this quantity takes both positive and negative values. The same holds for the entropy (or heat) exchange of a closed system with its environment. However, on average there should be no net exchange of heat with the environment when the system has reached equilibrium.

For thermostats that fall into the Nosé–Hoover (or the more general SDC) scheme outlined above, this is exactly the case. Remember that we constructed the dynamics such that  $\tilde{\zeta}$  has a Gaussian distribution around zero in the stationary state (2.52). Thus, the steady-state mean of  $\tilde{\zeta}$  (and hence the average phase space contraction rate) vanishes:

$$\langle \Lambda \rangle_{\infty} = \int_{\Gamma} \Lambda \rho_{\infty} dx = 0. \quad (2.58)$$

### Non-equilibrium

What about non-equilibrium situations? Firstly, we mention that a generic deterministic dynamics (2.57) does not necessarily admit an invariant density  $\rho_{\infty}$ . Rather it might admit an invariant *measure*  $\mu_{\infty}$ . A detailed discussion of measure theory will be the subject of chapter 3. Here, we just mention that the ensemble average  $\langle \Lambda \rangle_{\infty}$  is still well-defined.

Sometimes, instead of ensemble averages one is interested in time averages. For their

## 2. Notions of entropy and entropy production

definition, note that a differential equation  $\partial_t x = f(x)$  formally defines a *flow*  $\Psi(\cdot, t) = \Psi^{(t)}(\cdot)$ . It propagates an initial condition  $x_0$  to the solution function  $x_t = x(t)$  of the differential equation (2.57):

$$\begin{aligned}\Psi: \Gamma \times \mathbb{T} &\rightarrow \Gamma, \\ (x_0, t) &\mapsto x_t = \Psi^{(t)}(x_0).\end{aligned}$$

Additionally, the flow obeys a semi-group structure, *i.e.*  $\Psi(t_2, \Psi(t_1, x)) = \Psi(t_1 + t_2, x)$  for all  $t_1, t_2$  in its time domain  $\mathbb{T}$ . In the following we assume that the flow is invertible, *i.e.*  $\mathbb{T} = \mathbb{R}$ . The time average  $\overline{\varphi}_t^{(\tau)}$  of an observable  $\varphi(x)$  is defined as

$$\overline{\varphi}_t^{(\tau)}(x) = \frac{1}{\tau} \int_t^{t+\tau} \varphi(\Psi^{(s)}(x)) \, ds. \quad (2.59)$$

We are interested in the time-averaged phase-space expansion rate:

$$\overline{\Lambda}_t^{(\tau)}(x) := \frac{1}{\tau} \int_t^{t+\tau} \Lambda(\Psi^{(s)}(x)) \, ds. \quad (2.60)$$

Henceforth, we only consider situations where the limit

$$\overline{\Lambda}_\infty := \lim_{\tau \rightarrow \infty} \overline{\Lambda}_t^{(\tau)} \quad (2.61)$$

exists. For autonomous systems that limit is independent of  $t$ . Further, we are interested in globally attracting systems, where this limit is also independent of  $x$ . By definition it is the sum of all Lyapunov exponents (for a detailed discussion of the latter *cf.* [GH83]). For systems defined on compact phase spaces, this quantity is always non-positive, *i.e.*  $\overline{\Lambda}_\infty \leq 0$ . For a *physical* dynamics the time average agrees with the ensemble average using a physical<sup>13</sup> invariant measure  $\mu_\infty$ .

$$\overline{\Lambda}_\infty = \langle \Lambda \rangle_\infty \leq 0.$$

We have already seen that the *equality*  $\langle \Lambda \rangle_\infty = 0$  holds in the case of the equilibrium MD equations introduced above. The generic case is

$$\overline{\Lambda}_\infty = \langle \Lambda \rangle_\infty < 0. \quad (2.62)$$

In particular, this holds for the equations of motions used for non-equilibrium molecular dynamics (NEMD). With a reference to the standard literature [ES02; JR10], we state that in that case  $\langle \Lambda \rangle_\infty$  obtains the negative value of the thermodynamic dissipation rate  $\Sigma$

<sup>13</sup> A physical measure for an arbitrary dynamics is formally defined in Refs. [You02; BB03]. Physicists usually call a dynamics “ergodic” when it is “mixing” in the sense that time and ensemble averages of physical observables agree. This intuitive notion agrees with the definition of a physical measure in mathematics. There, the term “ergodic” is already reserved for a slightly more general concept, *cf.* Section 2.6.2.

divided by temperature  $T$ . The latter is defined as

$$\Sigma = \sum_{\nu} J_{\nu} A_{\nu} \quad (2.63)$$

where  $J_{\nu}$  denotes the  $\nu$ th thermodynamic current which is driven by its corresponding *conjugate thermodynamic force* or *affinity*  $A_{\nu}$ , cf. Sec. 2.4.2.

The interpretation of  $\Lambda(x)$  as the dissipation rate or entropy change in the environment is based on physical models for (NE)MD. Regardless of any physical-thermodynamic interpretation, we characterize arbitrary deterministic dynamics by their (average) phase space expansion. In the following we will refer to equations of motion  $\dot{x} = f(x)$  that obey equations (2.11), (2.58) and (2.62) as *uniformly conservative*, *conservative* and *dissipative systems*, respectively.

### 2.5.3. Stroboscopic maps and time-discrete dynamics

Finally, we consider the case of discrete-time dynamical systems. A discrete dynamical system can be obtained from a continuous flow  $\Psi(x, t)$  as a *stroboscopic map*  $\Phi(x) := \Psi(x, \Delta t)$  with a *stroboscopic time interval*  $\Delta t$ . Under the assumption that  $\Psi(\cdot, t)$  is defined for all  $t \in \mathbb{R}$ , the map  $\Phi(x)$  is invertible and maps  $\Gamma$  onto  $\Gamma$ . For  $\tau \in \mathbb{Z}$ , we denote by  $\Phi^{(\tau)}(x) := \Psi(x, \tau \Delta t)$ . Note that for positive and negative integers  $\tau$ ,  $\Phi^{(\tau)}$  is the  $|\tau|$ -fold iterate of  $\Phi$  and  $\Phi^{-1}$ , respectively.

For continuous-time flows, we have classified systems by their phase space contraction rate  $\bar{\Lambda}_{\infty}$ . We would like to do the same for discrete dynamics. To that end consider the *Jacobian determinant*  $\hat{j}^{(\tau)}(x)$  of a  $\tau$ -times iterated map  $\Phi$ ,

$$\hat{j}^{(\tau)}(x) := |\det D\Phi^{(\tau)}(x)|, \quad (2.64)$$

where  $Df(x)$  denotes the *Jacobian matrix* of a differentiable function  $f: \Gamma \rightarrow \Gamma$  at a point  $x$ . If  $\Phi(x)$  is a stroboscopic map obtained from an arbitrary dynamics (2.57) with flow  $\Psi(x, t)$  it holds that [Rue99]

$$\bar{\Lambda}_{\infty} = \int_{\Gamma} \rho_{\infty} \operatorname{div} f \, dx = \int_{\Gamma} \rho_{\infty} \log \hat{j}^{(1)} \, dx. \quad (2.65)$$

If  $\Phi$  is not obtained as a stroboscopic map, we lack the notion of the divergence of a vector field  $\operatorname{div} f$  which defines the dynamics. However, the right hand side of Eq. (2.65) is still well-defined. The logarithm of the Jacobian determinant

$$\Lambda_t^{(1)}(x) := \hat{j}^{(1)}(\Phi^{(t-1)}s) \quad (2.66)$$

describes the phase space expansion (in unit time) that a small volume around a point  $x$  encounters at time  $t$ . Hence, it is natural to classify maps  $\Phi$  with reference to  $\log \hat{j}^{(1)}$ : We say a map  $\Phi$  is *uniformly conservative*, if  $\Lambda_t^{(1)}$  vanishes identically. Similarly, we call it *conservative* or *dissipative* if the steady-state average  $\langle \Lambda^{(1)} \rangle_{\infty}$  vanishes or is negative,

## 2. Notions of entropy and entropy production

respectively. Note that the log-Jacobian is also used in the definition of the Lyapunov exponents of discrete systems.

Finally, let us consider the variation of the Shannon entropy of the ensemble for a time-continuous dynamics in the time-interval  $\Delta t$ . It is straightforward to realize that (cf. Eq. (4.19) or [Rue99])

$$\Delta_t^{(1)} H := (H[\rho_t] - H[\rho_{t-1}]) = \Delta t \cdot \int_{\Gamma} \rho_{t-1} \log \hat{J}^{(1)} dx. \quad (2.67)$$

Thus, in the steady state we have

$$\Delta_{\infty}^{(1)} H := \lim_{t \rightarrow \infty} \Delta_t^{(1)} H[\rho_t] = \Delta t \int_{\Gamma} \rho_{\infty} \log \hat{J}^{(1)} dx = \Delta t \langle \Lambda \rangle_{\infty}. \quad (2.68)$$

Hence,  $\Delta_{\infty}^{(1)} H$  describes the asymptotic change of the ensemble entropy  $H[\rho_t]$  in unit time. Without loss of generality, from now on we fix the stroboscopic interval as the unit time, *i.e.*  $\Delta t = 1$ .

For conservative systems (like Hamiltonian or thermostated equilibrium dynamics) Eq. (2.68) evaluates to zero, *cf.* Eq. (2.58). For Hamiltonian dynamics, the entropy  $H[\rho_t]$  retains its original value for *all times*. In contrast, for the EMD dynamics we have discussed above, the entropy of the ensemble  $H[\rho_t]$  necessarily converges to the Gibbs entropy  $S^G$ . By definition, the latter is a (constrained) *maximum entropy distribution*.

For dissipative non-equilibrium systems we have a completely different situation: From Eq. (2.68) we know that in the  $t \rightarrow \infty$  limit the change per unit time of  $H[\rho_t]$  is *negative*. Hence,  $\lim_{t \rightarrow \infty} H[\rho_t] \rightarrow -\infty$ . Rather than having a maximum entropy distribution, the entropy of the ensemble is unbounded from below. The reason for this is that the limit distribution is not a *density* any more, but something which has a fractal distribution on a (generically) fractal support [Rue99]. This is another reason, why outside of equilibrium one must not identify the entropy of the phase space ensemble with the system's entropy. It is further at variance with the requirement that its value must be constant in any (non-equilibrium) steady state.

### 2.5.4. Reversibility and deterministic fluctuation relations

An important aspect of the Nosé–Hoover and many other thermostated (NE)MD models is their reversibility [JR10]. For deterministic invertible dynamical systems, reversibility means that one can reverse the arrow of time by applying a time-reversal operator  $\mathcal{S}: \Gamma \rightarrow \Gamma$  to its microstates.

For MD, the phase space  $\Gamma$  contains elements  $x = (\vec{q}, \vec{p}, \alpha)$  which summarize coordinates  $\vec{q}$ , momenta  $\vec{p}$  (which might be suppressed in underdamped dynamics) as well as the additional auxiliary variables

$$\alpha = (\tilde{\zeta}, \bar{\zeta}, \dots), \quad (2.69)$$

which represent the degrees of freedom of the thermostat.

Let  $\lambda$  denote the usual  $d$ -dimensional volume (Lebesgue measure) on  $\mathbb{R}^d$ . In the following, we are interested in measure-preserving time reversal involutions  $\mathcal{I}$ . Formally, we call a flow  $\Phi^{(t)} = \Psi(\cdot, t)$  (or an iterated map  $\Phi^{(t)} = \Phi \circ \dots \circ \Phi$ ) *reversible* if and only if there is a mapping  $\mathcal{I}$  such that

$$\mathcal{I} \circ \mathcal{I} = \text{id}, \quad (\text{involution}), \quad (2.70a)$$

$$\lambda(\mathcal{I}^{-1}A) = \lambda(A), \quad (\text{measure-preserving}), \quad (2.70b)$$

$$\Phi^{(-t)}x = (\mathcal{I} \circ \Phi^{(t)} \circ \mathcal{I})x, \quad \forall t \in \mathbb{R} \text{ (or } \mathbb{Z}) \quad (\text{time reversal}). \quad (2.70c)$$

In the present context of thermostated MD, the time-reversal operator that fulfils the above properties is given as  $\mathcal{I}(\vec{q}, \vec{p}, \alpha) = (\vec{q}, -\vec{p}, -\alpha)$ . In Section 4.2 we discuss a more abstract reversible dynamics in discrete time.

Reversibility with a measure-preserving involution is the central requirement for the existence of *fluctuation relations* (FR) for deterministic dynamical systems. Similar to the stochastic FR discussed in Section 2.4.3, deterministic FR are statements about the probability to observe a certain value of entropy production in an interval  $[t, t + \tau]$ . Instead of a trajectory-dependent random variable for a stochastic processes, one considers the probability of observing a value  $a$  for a phase space observable  $\bar{\Omega}_t^{(\tau)}$ . Again, the FR compares the probability of finding  $\bar{\Omega}_t^{(\tau)} = a$  with the probability of finding  $-a$ :

$$\frac{\mathbb{P}[\bar{\Omega}_t^{(\tau)} = a]}{\mathbb{P}[\bar{\Omega}_t^{(\tau)} = -a]} = \exp a. \quad (2.71)$$

The most famous examples of deterministic FR are the transient FR by Evans and Searles [ES02] and the steady-state FR by Gallavotti and Cohen [GC95]. A recent result ensures that deterministic FR follow generally from an abstract fluctuation theorem [Woj09]. Like the stochastic FR (2.44), the deterministic formulation (2.71) is a detailed probabilistic version of the second law of thermodynamics.

### The dissipation function

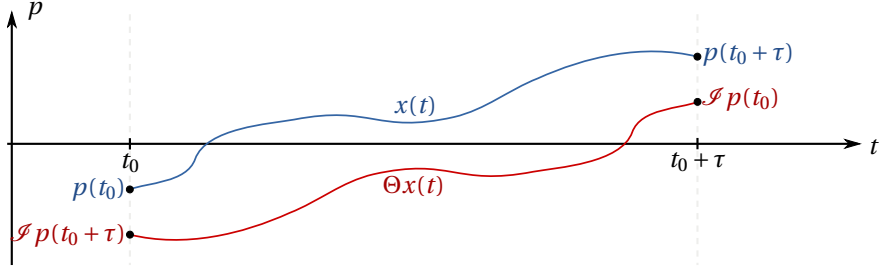
To provide a more concrete example, we discuss the Evans–Searles dissipation function  $\Omega_t^{(\tau)}$  for reversible dynamics (2.57), cf. Refs. [ES02; SRE07]. In order to construct this function, we define a logarithmic ratio of two probability densities  $\rho, \rho'$  on phase space  $\Gamma$ :

$$R[x; \rho, \rho'] := \log \frac{\rho(x)}{\rho'(x)}. \quad (2.72)$$

The average dissipation rate is given by the *time-averaged dissipation function*

$$\bar{\Omega}_t^{(\tau)}[x; \rho] := \frac{1}{\tau} (R[x; \rho, \Theta \rho] - \bar{\Lambda}_t^{(\tau)}(x)), \quad (2.73)$$

## 2. Notions of entropy and entropy production



**Figure 2.3.:** The anti-orbit  $\Theta x(t)$  (red) to the orbit  $x(t)$  (blue) in the time interval  $[t_0, t_0 + \tau]$ . We only display the anti-symmetric (momentum) coordinate  $p$ .

where  $\Theta \rho := \rho \circ \mathcal{S} \circ \Phi^{(\tau)}$ . The first term  $R[x; \rho, \Theta \rho]$  compares a probability density  $\rho$  at a point  $x$  with its value at the conjugate point  $\Theta x := \mathcal{S} \Phi^{(\tau)} x$ . The orbit of the conjugate point  $\Theta x$  is the so-called *conjugate* or *anti-orbit* to the orbit of a point  $x$ . Note that for  $0 \leq t \leq \tau$  we have  $\Phi^{(t)} x' = \mathcal{S} \Phi^{(\tau-t)} x$ . Thus, the anti-orbit consists of the time-reversed points of the corresponding orbit, which are visited in reverse direction, cf. Fig. 2.3. The second term is just the negative value of the time-averaged phase-space expansion in an interval  $[t, t + \tau]$ , see Eq. (2.60).

Depending on the choices of  $\rho$ , one can obtain several of the above mentioned fluctuation relations. The original Evans-Searles FR was obtained for  $\bar{\Omega}_\tau := \bar{\Omega}_0^{(\tau)} [x; \rho_0]$ , where  $\rho_0(x) = \rho_0(\mathcal{S}x)$  is a time-reversal symmetric initial condition [ES02]:

$$\frac{\mathbb{P}[\bar{\Omega}_\tau = a]}{\mathbb{P}[\bar{\Omega}_\tau = -a]} = \exp(a\tau). \quad (2.74)$$

In addition to that, Ref. [SRE07] describes other fluctuation relations that can be obtained by choosing other probability distributions for  $R[x; \rho, \rho']$ .

## 2.6. Measurable dynamical systems

In the previous section, we investigated how phase-space contraction for deterministic dynamics is interpreted as dissipation. The main motivation for this connection to (thermodynamic) entropy came from the equations used in (NE)MD. Independent from physical or thermodynamic interpretations, a purely information-theoretical connection between deterministic dynamical systems and entropy production has been worked out in the last fifty years. Early investigations started with the works of Kolmogorov, who was interested in a way to characterize the *complexity* of dynamical systems [Kol58]. Together with his student Y. Sinai, the concept was further developed and has been summarized by Sinai himself in Ref. [Sin09]. Some similarities of that theory with thermodynamics have been discussed by D. Ruelle [Rue04]. In the following, we only give a brief glimpse into the basic ideas.

### 2.6.1. Mathematical prerequisites

We briefly review the basic concepts of topology and measure theory. A detailed exposition is (freely) available in T. Tao's monograph [Tao11].

Let  $X$  be a set. The set  $\mathcal{P}(X)$  containing all subsets of  $X$  is called the *power set* of  $X$ . In the following, two kinds of subsets of  $\mathcal{P}(X)$  will be important: A  $\sigma$ -*algebra* on a set  $X$  is a family of subsets  $\mathcal{A} \subset \mathcal{P}(X)$  which contains the *empty set*  $\emptyset$  and the entire set  $X$  and is closed under the formation of countable unions and intersections. For any family of sets  $\mathcal{F} \subset \mathcal{P}(X)$ , the family  $\sigma(\mathcal{F})$  is the smallest  $\sigma$ -algebra that contains  $\mathcal{F}$ . Elements of a  $\sigma$ -algebra are called *measurable sets*. A *measurable space*  $(X, \mathcal{A})$  consists of a set and a  $\sigma$ -algebra on that set. We say that a map  $\nu: \mathcal{A} \rightarrow [0, \infty)$  is additive, if its value on the mutual union of a finite family  $\mathcal{F} = (A_i)_i$  of disjoint subsets  $A_i$  is equal to the sum  $\sum_i \nu(A_i)$ . Further, we call the map  $\sigma$ -additive if the above statement also holds for countable families. A *measure* is a map  $\nu: \mathcal{A} \rightarrow [0, \infty)$  which is both additive and  $\sigma$ -additive. A *probability measure*  $\mu$  is a measure that obeys  $\mu(X) = 1$ .

For any map  $f: X_1 \rightarrow X_2$  the pre-image map  $f^{-1}: \mathcal{P}(X_2) \rightarrow \mathcal{P}(X_1)$  is defined as

$$f^{-1}(A_2) = \{x \in X_1 \mid f(x) \in A_2\}.$$

A *measurable function*  $f$  between two measurable spaces  $(X_1, \mathcal{A}_1)$  and  $(X_2, \mathcal{A}_2)$  is a function such that the pre-image of any measurable set is measurable, i.e.  $f^{-1}(A_2) \in \mathcal{A}_1, \forall A_2 \in \mathcal{A}_2$ . A random variable is a measurable function between a probability space  $(X_1, \mathcal{A}_1, \mu)$  and a measurable space  $(X_2, \mathcal{A}_2)$ . For a measurable function  $f: (X_1, \mathcal{A}_1) \rightarrow (X_2, \mathcal{A}_2)$  we define the pushforward of a measure  $\nu$  on  $X_1$  as  $(f)_*(\nu) := \nu \circ f^{-1}$ , where  $\circ$  denotes composition.

A family of subsets  $\mathcal{T} \subset \mathcal{P}(X)$  containing the *empty set*  $\emptyset$  and the entire set  $X$ , which is also closed under the formation of arbitrary unions and finite intersections is called a *topology*. The tuple  $(X, \mathcal{T})$  is called a *topological space* and elements  $B \in \mathcal{T}$  are called *open sets*. A mapping  $f$  between two topological spaces  $(X_1, \mathcal{T}_1)$  and  $(X_2, \mathcal{T}_2)$  is called *continuous*, if the pre-image of an open set is open, i.e.  $f^{-1}(B_2) \in \mathcal{T}_1, \forall B_2 \in \mathcal{T}_2$ . Often, we will assume that  $(X, d)$  is a metric space with some metric  $d$ . In that case, the open sets  $B \in \mathcal{T}$  are determined by the metric. Similarly to the notion of a generated  $\sigma$ -algebra, the topology  $\mathcal{T}(\mathcal{F})$  generated by a family of sets  $\mathcal{F} \subset \mathcal{P}(X)$  is the smallest topology that contains  $\mathcal{F}$ . The  $\sigma$ -algebra  $\sigma(\mathcal{T})$  generated by the open sets is called the *Borel  $\sigma$ -algebra*.

If not indicated differently, we assume that the phase space  $\Gamma$  is a metric space equipped with a topology  $\mathcal{T}_\Gamma$  and the corresponding Borel  $\sigma$ -algebra  $\mathcal{B} = \sigma(\mathcal{T}_\Gamma)$ . A measure on such a space is called a *Borel measure*.

Let  $\mu$  and  $\nu$  be measures for a measurable space  $(\Gamma, \mathcal{B})$ . We say that a measure  $\mu$  is *absolutely continuous* with respect to the measure  $\nu$ , if  $\nu(B) = 0 \Rightarrow \mu(B) = 0$  for all  $B \in \mathcal{B}$

## 2. Notions of entropy and entropy production

and write  $\mu \ll \nu$ . In that case,  $\mu$  has a *density*  $f$  with respect to  $\nu$ , *i.e.*

$$\int_B g \, d\mu = \int_B f g \, d\nu$$

holds for all measurable functions  $g$  and all sets  $B \in \mathcal{B}$ . The function  $f =: \frac{d\mu}{d\nu}$  is called the *Radon–Nikodym derivative* of  $\mu$  with respect to  $\nu$ .

If  $\mu$  is a probability measure defined on the Borel sets and  $\nu = \lambda$  is the Lebesgue measure, the Radon–Nikodym derivative gives the *probability density*  $\rho := \frac{d\mu}{d\lambda}$  of the probability measure  $\mu$ .

### 2.6.2. Measurable and topological dynamical systems

In this subsection we consider  $C^1$ -diffeomorphisms, which are differentiable maps with a differentiable inverse.<sup>14</sup>

Throughout the rest of this section, let  $(\Gamma, \mathcal{B})$  be a measurable space and  $\Phi$  a measurable  $C^1$ -diffeomorphism. The triple  $(\Gamma, \mathcal{B}, \Phi)$  is called a *measurable dynamical system on phase space*  $(\Gamma, \mathcal{B})$ . Similarly, if  $(\Gamma, \mathcal{T})$  is a topological space and  $\Phi$  is a continuous map, we call the triple  $(\Gamma, \mathcal{T}, \Phi)$  a *topological dynamical system*.

Usually, we think of  $\Gamma$  as a separable metric space with  $\mathcal{B}$  denoting the Borel sets. In that case, any measurable function is also continuous and a measurable dynamical system is also a topological one. If a probability measure  $\mu$  is given, we call  $(\Gamma, \mathcal{B}, \Phi, \mu)$  a dynamical system with measure  $\mu$ .

Now we can formally define an *invariant measure*  $\mu_\infty$  as a (probability) measure that is a fixed point of the push-forward, *i.e.*,

$$\Phi_* \mu_\infty \equiv \mu_\infty \circ \Phi^{-1} = \mu_\infty,$$

In that case we call  $(\Gamma, \mathcal{B}, \Phi, \mu_\infty)$  a *measure-preserving dynamical system*.

A special kind of invariant measure is an *ergodic* measure. We call a set  $B$  to be  $\Phi$ -invariant, if  $B = \Phi^{-1}(B)$ . In mathematics, one says that a measure  $\mu_\infty$  is ergodic with respect to  $\Phi$  (or  $\Phi$  is ergodic with respect to  $\mu_\infty$ ) if for any invariant  $B \in \mathcal{B}$

$$\mu_\infty(B) = 0 \quad \text{or} \quad \mu_\infty(B) = 1. \tag{2.75}$$

For ergodic measures, ergodic theorems like the one of Birkhoff then ensure that properly defined asymptotic time-averages agree with averages taken with respect to  $\mu_\infty$ . However, the mathematical definition of ergodicity does *not* imply that the orbit of *all* points densely covers the *entire* phase space, as is sometimes assumed in physics.

The push-forward operator  $\Phi_*$  defines an evolution rule for measures. We will mostly be interested in the evolution of probability measures  $\mu$  which are absolutely continuous

<sup>14</sup> We can relax that to *almost-everywhere* differentiable maps with almost-everywhere differentiable inverse. All that we need is that the log-Jacobian is well-defined almost everywhere. For instance, a Lipschitz continuous functions with a Lipschitz-continuous inverse provides an example.



with respect to a *reference measure*  $\nu$ . Most of the time we choose  $\nu = \lambda$ , where  $\lambda$  denotes the translation invariant (Lebesgue) measure on  $(\Gamma, \mathcal{B})$ . In that case,  $\mu$  has a probability density  $\rho: \Gamma \rightarrow [0, \infty)$ .

Because  $\Phi$  is a measurable diffeomorphism, the  $t$ -times push-forwarded measure

$$\mu_t := \Phi_*^t \mu$$

is *equivalent* to  $\mu$ , i.e.  $\mu$  and  $\mu_t$  are mutually absolutely continuous. One also says that  $\Phi$  is *non-singular with respect to  $\mu$* . In Refs. [ES02; SRE07] the authors call a density  $\rho$  of  $\mu$  to be *ergodically consistent* with  $\Phi$  if (i)  $\Phi$  is non-singular with non-singular inverse and (ii) time-reversal can be represented by a measure-preserving involution (2.70). A non-singular  $\Phi$  yields a sequence of measures  $\{\mu_t\}$  with densities  $\{\rho_t\}$  that evolve according to

$$\rho_t = \frac{\rho_0 \circ \Phi^{-t}}{\hat{j}^{(t)} \circ \Phi^{-t}}, \quad (2.76)$$

where  $\hat{j}^{(t)}$  is the Jacobian determinant (2.64).

This provides the following perspective on observables and their averages: In the present setting, an observable  $\varphi: \Gamma \rightarrow \mathbb{R}$  is a measurable function that assigns real values to points in phase space. Usually, we interpret an observable as the outcome of a measurement procedure. The (Lebesgue) integral over the full phase space is the *ensemble average* of an observable  $\varphi$ :

$$\langle \varphi \rangle_t := \int_{\Gamma} \varphi d\mu_t = \int_{\Gamma} \varphi \rho_t dx. \quad (2.77)$$

The properties of the Jacobian

$$\hat{j}^{(t)} = \frac{1}{\hat{j}^{(-t)} \circ \Phi^t}, \quad (2.78a)$$

$$= \prod_{k=1}^t (\hat{j} \circ \Phi^{t-k}) \equiv \prod_{k=1}^t (\hat{j} \circ \Phi^k), \quad (2.78b)$$

together with the usual transformation rules of the integral, yield for the ensemble average at time  $t$ :

$$\langle \varphi \rangle_t = \int_{\Gamma} (\varphi \circ \Phi^t) \rho_0 dx = \int_{\Gamma} \varphi \rho_t dx. \quad (2.79)$$

Hence, there are two ways to interpret the average  $\langle \varphi \rangle_t$ : In the first, one can think of a time-dependent observable  $\varphi_t := \varphi \circ \Phi^t$  which is averaged with the initial density  $\rho_0$ . In the second one, the observable does not depend on time whereas  $\rho_t$  evolves according to Eq. (2.76). A similar duality exists in quantum mechanics: The first interpretation is called ‘‘Heisenberg picture’’, whereas the latter is named after Schrödinger.

### 2.6.3. Topological and measure-theoretic entropy

In this subsection, we define entropy for both topological and measurable dynamical systems. Although these expressions are called “entropies”, they rather resemble asymptotic time-averages. Hence, they are better understood as rates of change of entropy than as entropy itself. This is a consequence of the fact that these entropies were designed to quantify the (asymptotic) *complexity* of the dynamics, rather than that of an ensemble at a specific point in time.

The *topological entropy* is the appropriate notion of the complexity of a topological dynamical system  $(\Gamma, \mathcal{T}, \Phi)$ . It provides an upper bound to the *metric entropy* defined for the measurable dynamical system  $(\Gamma, \sigma(\mathcal{T}), \Phi)$ . Here, we only give a brief overview. For a more rigorous discussion and explicit proofs we refer the reader to chapters 4–6 of Ref. [Jos06].

Both the definition of the topological and the metric entropy start with the notion of a (minimal) cover of phase space. The elements of that cover can be understood as an approximate description of the position of a point  $x \in \Gamma$ . For instance, they may correspond to a certain value of an observable representing a measurement on the system.<sup>15</sup> If we measure the system initially and after each of  $\tau$  iterations, we obtain a time series  $\underline{\omega}^{(\tau)}$ . The topological entropy specifies the exponential growth rate of the number of distinguishable time series  $\underline{\omega}^{(\tau)}$  with  $\tau$ .

Hence, in a sense it specifies the (rate of) additional *information* we obtain about the topological structure of the flow when we observe longer and longer time series. The measure-theoretic entropy additionally takes a (non-uniform) probability for time series into account.

#### Covers and partitions

Henceforth it will be useful to consider *indexed families of sets*  $\mathcal{F} = (\mathcal{C}_\omega)_{\omega \in \Omega}$ . An element  $\mathcal{C}_\omega \subset \Gamma$  is indexed by an entry  $\omega$ , which is an element of the index set  $\Omega$ . If we do not care about the indexing, we consider  $\mathcal{F} \subset \mathcal{P}(\Gamma)$  as a subset of the power set.

A finite (indexed) family  $\mathcal{F} = (\mathcal{C}_\omega)_{\omega \in \Omega} \subset \mathcal{P}(\Gamma)$  is called a *cover* of  $\Gamma$  if  $\bigcup_{\omega} \mathcal{C}_\omega = \Gamma$ . For a topological space, a cover  $\mathcal{F} \subset \mathcal{T}$  consisting of open sets is called an *open* or *topological cover*. A cover  $\mathcal{F}$  is called *minimal* if it does not contain any true subset  $\mathcal{F}' \subsetneq \mathcal{F}$  which already is a cover. In the following, we will always assume that  $\mathcal{F}$  is minimal. A cover consisting of disjoint sets is called a *partition*. Note that for a connected phase space  $\Gamma$ , a partition can never be an open cover.<sup>16</sup>

For two families of subsets  $\mathcal{F}$  and  $\mathcal{F}'$  we define their *mutual refinement* or *join* as the

<sup>15</sup> We will make that notion more precise in Chapter 3.

<sup>16</sup> Later we will define so-called topological partitions, which are neither partitions nor covers, but turn out to be a useful generalizations. Topological partitions have been introduced by R. Adler [Adl98].

set of all mutual intersections of their elements:

$$\mathcal{F} \vee \mathcal{F}' := \{\mathcal{C} \cap \mathcal{C}' \mid \mathcal{C} \in \mathcal{F}, \mathcal{C}' \in \mathcal{F}'\} \quad (2.80)$$

Note that the join of two partitions again is a partition. Further, the join of two open covers is an open cover, because a topology is closed under finite intersections.

The pre-image of an indexed family of subsets  $\mathcal{F} = (\mathcal{C}_\omega)_{\omega \in \Omega}$  under the map  $\Phi$  is defined as

$$\Phi^{-1}(\mathcal{F}) := \{\Phi^{-1}\mathcal{C}_\omega \mid \omega \in \Omega\} \quad (2.81)$$

Hence, if  $\Phi$  is continuous, the pre-image of a topological cover again is a topological cover. Similarly, if  $\Phi$  is measurable, the pre-image of a family of measurable subsets again is a family of measurable subsets.

We define the *topological entropy of a minimal cover*  $\mathcal{F}$  as the logarithm of the number of its elements:

$$H_{\text{top}}(\mathcal{F}) := \log |\mathcal{F}|$$

The dynamics creates a sequence of refined covers  $(\bigvee_{n=1}^t \Phi^{-n}(\mathcal{F}))_{t \in \mathbb{N}}$ . For a topological dynamical system  $(\Gamma, \mathcal{T}, \Phi)$  and a cover  $\mathcal{F}$  we define the *topological entropy* of the asymptotic dynamical refinement as

$$H_{\text{top}}(\mathcal{F}, \Phi) := \lim_{t \rightarrow \infty} \frac{1}{t} H_{\text{top}} \left( \bigvee_{n=1}^t \Phi^{-n}(\mathcal{F}) \right).$$

As we have mentioned above, this quantity is rather an asymptotic entropy *rate* than an entropy. The *topological entropy of a topological dynamical system*  $(\Gamma, \mathcal{T}, \Phi)$  is the maximum of that quantity over all minimal covers:

$$h_{\text{top}}(\Phi) := \sup_{\mathcal{F} \text{ is a minimal cover}} H_{\text{top}}(\mathcal{F}, \Phi). \quad (2.82)$$

The definition of the measure-theoretic *Kolmogorov–Sinai* entropy (KS-entropy) proceeds in an analogue way, if we replace “cover” by “partition” and “continuous” by “measurable”. To that end let  $(\Gamma, \mathcal{A}, \mu_\infty, \Phi)$  be a measure-preserving dynamical system and  $\mathcal{Q} = (\mathcal{C}_\omega)_{\omega \in \Omega}$  be a finite partition. The measure-theoretic entropy of the partition is

$$H_{\mu_\infty}(\mathcal{Q}) = - \sum_{\mathcal{C} \in \mathcal{Q}} \mu_\infty(\mathcal{C}) \log(\mu_\infty(\mathcal{C})).$$

With that, the *metric entropy of*  $(\Gamma, \mathcal{A}, \mu_\infty, \Phi)$  *with respect to*  $\mathcal{Q}$  reads

$$H_{\mu_\infty}(\mathcal{Q}, \Phi) := \lim_{t \rightarrow \infty} \frac{1}{t} H_{\mu_\infty} \left( \bigvee_{n=1}^t \Phi^{-n} \mathcal{Q} \right).$$

## 2. Notions of entropy and entropy production

Finally, the *Kolmogorov–Sinai entropy* of  $(\Gamma, \Phi, \mu_\infty, \mathcal{Q})$  is obtained as the supremum over all partitions:

$$h_{\mu_\infty}(\Phi) := \sup_{\mathcal{Q} \text{ is a partition}} H_{\mu_\infty}(\mathcal{Q}, \Phi). \quad (2.83)$$

Generally each invariant measure  $\mu_\infty$  yields another value of the Kolmogorov–Sinai entropy. If these values have a maximum, the associated measure is called the *measure of maximum entropy*. The following variational principle states that the topological entropy is an upper bound for the Kolmogorov–Sinai entropy:

### 2.2 Theorem (Variational principle [Din71; Goo71a; Goo71b])

$$h_{\text{top}}(\Phi) = \sup_{\mu_\infty \text{ is } \Phi\text{-invariant}} h_{\mu_\infty}(\Phi) \quad (2.84)$$

Finally we state a result that relates the Kolmogorov–Sinai entropy to phase-space contraction:

**2.3 Theorem (Pesin’s formula [HP08])** Let  $\bar{\Lambda}_\infty^+(x)$  be the sum of all positive Lyapunov exponents associated with a point  $x$  for a map  $\Phi$ , counted according to their multiplicity. Let  $\mu_\infty$  be a  $\Phi$ -invariant probability measure. Pesin’s formula states that

$$h_{\mu_\infty} = \int_\Gamma \bar{\Lambda}_\infty^+ d\mu_\infty. \quad (2.85)$$

Note that Pesin’s formula Eq. (2.85) only talks about the *positive* part of the Lyapunov spectrum. The reason for that is that the entropies defined above distinguish the *direction of time*: The refinements of  $\mathcal{Q}$  are constructed using the *pre-images* rather than the images of partition elements. If  $\Phi$  is invertible with measurable (or continuous) inverse, we can define similar quantities for the reversed dynamical systems  $(\Gamma, \Phi^{-1})$ . Then, Pesin’s formula yields that the sum of the KS-entropy for dynamical system and its reverse agree with the sum of *all* Lyapunov exponents, and hence with the average phase-space contraction in the steady state, *cf.* Ref. [Gas04].

## 2.7. Summary

In the present chapter, we have reviewed different notions of entropy together with the contexts they appear in. We started with the definition of entropy in classical thermodynamics for macroscopic heat engines. After that, we presented the notion of entropy as information, as it arises in a mathematical treatment of the elements of communication. Even much earlier, Gibbs introduced a similar probabilistic notion of (equilibrium) entropy in the context of his statistical ensembles. This idea together with the assumption of

an underlying microscopic Hamiltonian dynamics are the key assumptions of classical statistical physics. Consistency between different macroscopic thermodynamic constraints is achieved in the thermodynamic limit, *i.e.* the case where the particle number becomes very large.

Modern statistical physics aim to generalize the classical concept. More precisely, it drops certain classical assumptions and aims to define appropriate generalizations of entropy and dissipation. For instance, the thermodynamic limit does not apply when one deals with small systems, where noise plays a major role. We reviewed modern model paradigms, using either stochastic or deterministic, but non-Hamiltonian dynamics. Finally, we briefly gave an overview of the abstract mathematical treatment of the latter, *i.e.* the ergodic theory of measurable dynamical systems.

Throughout the chapter we stressed the importance of a distinction between a “system” and its surrounding “medium” by means of operational accessibility. Or, as Oliver Penrose puts it in his work on a deductive treatment of the foundations of statistical mechanics [Pen70]:

“This limitation on our powers of observation is an essential part of statistical mechanics; without it the theory would be no more than a branch of ordinary mechanics.”

In his work, Penrose carefully reviews the assumptions made by (classical) statistical mechanics. The key assumptions are those of *causality of the microscopic dynamics* and the *Markovian postulate* for the mesoscopic dynamics. The microscopic causality manifests itself in the assumptions of a *deterministic* evolution. Penrose focuses on Hamiltonian mechanics as the underlying deterministic evolution, though he is well aware that this is only another (good) approximation. The Markovian postulate characterizes the nature of the statistics of the time series obtained by consecutive measurements on a system: They have limited memory in the sense that the probability of measuring a future state must only depend on the current state and not on the history of past measurements.

While Penrose used these postulates in his deductive treatment of classical statistical mechanics, the present thesis looks at them in the light of more modern paradigms. The success of (non-equilibrium) molecular dynamics simulations with non-Hamiltonian equations of motion encourages a deeper look at more *general* microscopic dynamics. Further, the abstract mathematical concept of entropy in ergodic theory is most interesting for non-Hamiltonian, dissipative dynamics. Thus, Chapters 3 and 4 investigate the behaviour of observable time series produced by reversible deterministic, but not necessarily Hamiltonian dynamics.

Within that framework, Chapter 3 investigates the implications of the Markovian postulate on observable states and microscopic ensembles. Chapter 4 outlines an abstract mathematical framework for entropy and entropy production based on information theory. Although Chapter 4 is largely independent of the Markovian postulate, we explicitly give an analytically tractable example where the Markovian postulate holds exactly. Fur-

## *2. Notions of entropy and entropy production*

ther, we establish a connection to the modern theory of stochastic thermodynamics, which is similarly based on the Markovian postulate.

Stochastic thermodynamics is generally understood as a paradigm for small systems in (non-equilibrium) thermodynamic environments. For such systems, noise is not negligible and may even play a functional role. Consequently, the second half of this thesis, Chapters 5 and 6, is concerned with the quantification of fluctuations in Markovian stochastic thermodynamics.

In the final Chapter 7, we put our results in the context of the bird's-eye perspective given in the present introductory chapter. At that point, the author hopes that the reader will have a good idea of the notions of entropy and dissipation in modern statistical physics; both in general terms and in the particular framework of stochastic thermodynamics. To come back to the initial quote: The present chapter should be understood as the common basis for the discussion we attempt in the final chapter.

### 3. Markovian symbolic dynamics

“ *The crucial [postulate in idealized models of real physical systems] is expressing the assumption, that the successive observational states of a physical system form a Markov chain. This is a strong assumption, [...], but even so, it has been adopted here because it provides the simplest precise formulation of a hypothesis that appears to underlie all applications of probability theory in physics.* ”

---

O. Penrose, *Foundations of Statistical Mechanics*, 1970

#### What is this about?

In the discussion of the previous chapter we have already mentioned the two crucial assumptions of classical statistical mechanics as identified by O. Penrose. The first one was the assumption that the “phase-space density at any time is completely determined by what happened to the system before that time and is unaffected by what will happen to the system in the future” [Pen70]. The second one is the Markovian postulate, which we chose as the initial quote for this chapter.

In his considerations, Penrose thought of Hamiltonian mechanics as the underlying microscopic deterministic evolution. From that perspective, he presented a deductive treatment of the foundations of classical statistical mechanics. As he mentions himself (*cf.* the initial quote), the Markovian postulate is crucial for the success of such a treatment.

In the present chapter, rather than adopting it, we establish requirements under which the Markovian postulate holds rigorously. Stochasticity in the observed time series arises from the sensitive dependence on initial conditions for such deterministic-chaotic systems. In our framework, a measurement outcome is given by the value of an observable  $M$  on an underlying microscopic phase space  $\Gamma$ . Usually, the microscopic evolution rule  $\Psi: x_0 \mapsto x_t$  is described by non-linear deterministic equations  $\dot{x} = f(x)$ . The (NE)MD equations from section 2.5 provided physical examples for such equations. For simplicity, in this chapter we will restrict ourselves to iterated invertible (stroboscopic) maps  $\Phi$ .

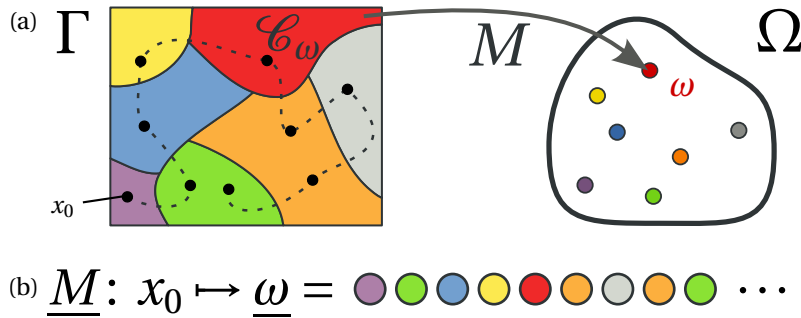
We investigate the implications of the Markovian postulate for measurement observables  $M$  and microscopic ensembles (measures)  $\mu$ . The rigorous treatment demands a larger amount of formality than previous and subsequent chapters. However, we hope

### 3. Markovian symbolic dynamics

that the chapter is self-contained enough, such that it can be read without any further reference. This chapter is supposed to appear as a paper in the *Journal of Statistical Physics* [Alt p].

It is structured as follows: We start by formally introducing our mathematical framework in Section 3.1. In particular, we describe how a measurable dynamical system on phase space together with an observable gives rise to a measurable symbolic dynamical system on a shift-invariant subset of all infinite time series, a so-called subshift. In Section 3.2 we review the notion of partitions created by observables. So-called Markov partitions generate a topology on the space of possible time series, which provides the necessary topological backbone of a Markov chain. In Section 3.3 we equip these spaces with measures that make the symbolic dynamics a Markov chain. In the final Section 3.4 we discuss our findings in the context of earlier results in ergodic theory. Further, we propose operational interpretations of the abstract results in the context of an experimenter taking (idealized) measurements.

#### 3.1. Symbolic stochastic dynamics



**Figure 3.1.:** a) All “microstates”  $x \in \mathcal{C}_\omega$  are mapped to the “measurement result”  $\omega$  by an observable  $M$ . Consequently, a cell  $\mathcal{C}_\omega$  is the pre-image  $M^{-1}(\omega)$  of  $\omega$  for the observable. b) A microscopic (discrete) orbit  $x(t)$  starting at at  $x(0) = x_0$  produces a (discrete) time series  $\underline{\omega} = \underline{M}(x_0)$  of subsequent measurement results.

We start by introducing the basic mathematical notions needed for the formal treatment of the following situation, cf. Fig. 3.1: Consider an observable  $M: \Gamma \rightarrow \Omega$ , which describes the possible outcomes of a measurement on a physical system. In reality, any measurement apparatus only has a finite resolution. Hence, the number  $|\Omega|$  of mesoscopic measurement results  $\omega \in \Omega$  is smaller than the number of microscopic states  $x \in \Gamma$ . In particular, we are interested in the situation where  $N := |\Omega|$  is finite, but may be large. This assumption of finiteness will simplify the mathematical treatment considerably.<sup>1</sup> Now consider an experimenter who records subsequent values of the observable  $M$  in a time series  $\underline{\omega}^{(\tau)} = (\omega_0, \omega_1, \dots, \omega_\tau)$ . Upon repeating the experiment, she is careful to prepare the system each time in the same initial state. Nonetheless, she observes a different time

<sup>1</sup> Moreover, if one takes the finiteness of any numerical representation of data on a piece of paper or in the storage of a computer seriously, this assumption is also physically justified.



series of successive measurement results every time. The reason for this phenomenon lies in the finite resolution of her experimental apparatus: She can only *prepare* a mesoscopic initial condition, *i.e.* one that is experimentally accessible to her. However, she has no chance to ensure that the system is at a certain microscopic configuration.

The mathematical subjects we need in order to formalize this situation are *measure theory* and the derived concept of a *stochastic process*. In accordance with the causality hypothesis, the evolution of the system in phase space is described as an abstract *dynamical system* (*cf.* Section 2.6). The mesoscopic or *coarse-grained* description on the level of time series is the subject of *symbolic dynamics*. Putting everything together we arrive at the central Definition 3.8 of a *stochastic process of observed time series* which are generated by a dynamical system and an observable.

### 3.1.1. Stochastic processes

In Section 2.6 we have already introduced the necessary definitions and notation for measurable dynamical systems. Further, we have given an informal discussion of stochastic processes in Section 2.4. The rigorous treatment in the present chapter requires a much higher degree of formality. In the following, we revisit the theory of Markov chains on finite state spaces. For a detailed review of continuous-time stochastic processes on arbitrary state spaces we refer the reader to Ref. [CB12].

**3.1 Definition (Stochastic process)** Let  $(X, \mathcal{B}, \mu)$  be a probability space. Further, let  $(f_t)_{t \in \mathbb{T}}$  be a family of random variables  $f_t: (X, \mathcal{B}) \rightarrow (\Omega, \mathcal{A})$  indexed by  $t \in \mathbb{T}$ . We call  $(f_t)_{t \in \mathbb{T}}$  a *stochastic process*,  $(X, \mathcal{B}, \mu)$  its *underlying probability space* and  $(\Omega, \mathcal{A})$  its *state space*. Further, for  $x \in X$  the family  $\underline{f}(x) := (f_t(x))_{t \in \mathbb{T}}$  is called the *trajectory* of  $x$ .

In the above definition, the index set  $\mathbb{T}$  was arbitrary. In many cases it is understood as a *time domain* and thus assumed to be a totally ordered set with a (semi-)group structure like  $\mathbb{R}$ ,  $\mathbb{Z}$  or  $\mathbb{N}$ . In the following,  $\mathbb{T}$  will always denote  $\mathbb{Z}$  or  $\mathbb{N}$ , *i.e.* we are interested in *discrete-time stochastic processes*. Similarly, the state space  $\Omega$  can be either discrete or continuous. In the discrete case, we call the process a *stochastic jump process*. Henceforth, we often consider infinite tuples  $\underline{\omega} := (\omega_t)_{t \in \mathbb{T}} \in \Omega^{\mathbb{T}}$  and their finite projections  $\underline{\omega}^{(\tau)} \in \Omega^{\tau+1}$ , which we will refer to as infinite and finite *time series*, respectively.

In the present context of stochastic process, time series appear as the values of the trajectory map  $\underline{f}(\cdot)$  for a given point  $x \in X$ . Yet, the space of infinite time-series  $\Omega^{\mathbb{T}}$  is interesting on its own. Formally, it is a *product space*. Let us recall some useful definitions about the structure (in particular the topology and the  $\sigma$ -algebra) of product spaces:

**3.2 Definition (Product spaces, projection operator, cylinders)** Let  $I$  and  $\Omega$  be arbitrary sets. The Cartesian product of  $\Omega$  over  $I$ ,  $\prod_{i \in I} \Omega$ , is the set containing all generalized tuples  $(\omega_i)_{i \in I}$  such that  $\omega_i \in \Omega$ . The short form  $\Omega^I := \prod_{i \in I} \Omega$  yields a more compact notation. Moreover, we write  $\Omega^I \equiv \Omega^N$  with  $N \in \mathbb{N} \setminus 0$ , if either  $|I| = N$  and we do not care about the structure of the index set  $I$ , or if  $I = \{1, 2, \dots, N\}$ .

### 3. Markovian symbolic dynamics

For  $I' \subset I$ , the *projection map*  $\pi_{I' \leftarrow I}: \Omega^I \rightarrow \Omega^{I'}$  singles out the entries at indices  $i \in I' \subset I$ . If  $I'$  is finite, we say  $\pi_{I' \leftarrow I}$  is *finite*. If  $I' = \{i\}$ , we write  $\pi_i := \pi_{\{i\} \leftarrow I}$  for the projector on the  $i$ th component.

If  $(\Omega, \mathcal{T})$  is a topological space, the pre-image  $\pi_i^{-1}(A)$  of an open-set  $A \in \mathcal{T}$  is called a *cylinder* with base in  $\Omega$ .

The *product topology* is the topology generated by all cylinders, *i.e.* it is the smallest topology such that any finite projection is continuous. Similarly, the *product  $\sigma$ -algebra* is defined to be the smallest  $\sigma$ -algebra such that every finite projection is measurable.

In the present work we consider *only* the case of finite state spaces with  $|\Omega| = N < \infty$ . Then, it is natural to choose both the  $\sigma$ -algebra and topology on the factor set  $\Omega$  to be the power set, *i.e.*  $\mathcal{T} = \mathcal{A} = \mathcal{P}(\Omega)$ . Thus, the product  $\sigma$ -algebra, by definition, is the  $\sigma$ -algebra generated by the cylinders. However, note that the product topology and the product  $\sigma$ -algebra are distinct if  $\mathbb{T}$  is an infinite set.

Now that we have the notion of a measurable structure on the space of time series, we can define the *probability of a trajectory*:

**3.3 Definition (Probabilities of trajectories)** Let  $\underline{f} := (f_t)_{t \in \mathbb{T}}$  be a discrete time jump process on a finite state space  $(\Omega, \mathcal{P}(\Omega))$  with underlying probability space  $(X, \mathcal{B}, \mu)$ , *cf.* Def. 3.1. Further, let  $A \in \mathcal{A}^{\mathbb{T}}$  be an element of the product  $\sigma$ -algebra on  $\Omega^{\mathbb{T}}$ . Then, the probability  $\mathbb{P}(A)$  is defined as the push-forward measure with respect to the trajectory, *i.e.*

$$\mathbb{P}(A) := (\underline{f}^* \mu)(A) \equiv \mu(\underline{f}^{-1}(A)).$$

**3.4 Remark** The Daniell–Kolmogorov theorem, (*cf.* Appendix A.1), ascertains that a consistent choice of the values of the measure  $\mathbb{P}$  fully determines the stochastic process. Then, a prescription of the probabilities  $\mathbb{P}(A)$  of events  $A \in \mathcal{A}^{\mathbb{T}}$  is sufficient, *i.e.* there is no need to explicitly refer to the underlying probability space. In contrast, in the present work we are explicit: The underlying probability space is the phase space of the deterministic microscopic dynamics  $\Phi$  of a physical system together with an ensemble measure. The trajectory  $\underline{f}$  corresponds to the recording of (infinite) time series. We will formally define such an *stochastic process of observed time series* in Def. 3.8.

Usually, one is interested in the probability of so-called *cylinder sets*, which are finite intersections of cylinders. Such sets define stochastic *events* by specifying which *elementary event*  $\omega_t \in \Omega$  may or may not occur at a given position in time  $t \in \mathbb{T}$ . In particular, we are interested in the probabilities of a special sort of cylinder sets:

**3.5 Definition** Let  $\underline{\omega}^{(\tau)} = (\omega_0, \omega_1, \dots, \omega_\tau) \in \Omega^{\tau+1}$  be a finite time series of length  $\tau+1$  and  $t \in \mathbb{T}$ . We define the  *$t$ -shifted forward cylinder* for  $\underline{\omega}^{(\tau)}$  as

$$\begin{aligned} Z_t[\underline{\omega}^{(\tau)}] &:= \bigcap_{k=0}^{\tau} \pi_{t+k}^{-1}\{\omega_k\} \\ &\equiv \{\underline{v} \in \Omega^{\mathbb{T}} \mid v_{t+k} = \omega_k, 0 \leq k \leq \tau\}. \end{aligned}$$

The measure of such a cylinder,  $\mathbb{P}_t[\underline{\omega}^{(\tau)}] = \mathbb{P}(Z_t[\underline{\omega}^{(\tau)}])$ , is called the *probability* for the finite time series  $\underline{\omega}^{(\tau)}$  to *occur* at time  $t$ .

Henceforth, we use the following convention with respect to temporal indices: A *superscript* ( $\tau$ ) denotes the *run length* of finite sequences. A *subscript*  $t$  characterizes a *distinct point* in time. If both indices appear, we usually refer to a finite sequence that starts at time  $t$  and extends to time  $t + \tau$ .

In the rest of this work, we will focus our attention on the arguably simplest (non-trivial) kind of stochastic process, that is a *discrete time Markovian jump process* or *Markov chain* on a finite state space. We say that  $\mathbf{p} = (p_\omega)_{\omega \in \Omega}$  is a *stochastic vector* if it obeys  $0 \leq p_\omega \leq 1$  for all  $\omega \in \Omega$  and it sums to unity, *i.e.*  $\sum_{\omega \in \Omega} p_\omega = 1$ . Further, a stochastic matrix  $\mathbb{W} = (w_{\omega'}^\omega)_{\omega, \omega' \in \Omega}$  obeys  $0 \leq w_{\omega'}^\omega \leq 1$  for all  $\omega, \omega' \in \Omega$  and each row is normalized, *i.e.*  $\sum_{\omega'} w_{\omega'}^\omega = 1$ .

**3.6 Definition (Homogeneous Markov chain)** Let  $\underline{\omega}^{(\tau-1)} \in \Omega^\tau$  and  $\omega_\tau \in \Omega$  be arbitrary. Define  $\underline{\omega}^{(\tau)} := (\underline{\omega}^{(\tau-1)}, \omega_\tau) \in \Omega^{\tau+1}$ . A jump process is called a *homogeneous Markov chain* if there exists a stochastic matrix  $\mathbb{W}$  with entries  $w_{\omega'}^\omega$ , such that

$$\mathbb{P}_t[\underline{\omega}^{(\tau)}] = \mathbb{P}_t[\underline{\omega}^{(\tau-1)}] \cdot w_{\omega_\tau}^{\omega_{\tau-1}}, \forall t \in \mathbb{T} \quad (3.1)$$

The above relation is called the *Markov property*. The numbers  $w_{\omega'}^\omega$  are called the *transition probabilities* of the Markov chain. The stochastic matrix  $\mathbb{W}$  is called the *transition matrix* of the process.

The Markov property is a very strong property of a stochastic process. In particular, it allows for a recursive construction of time-series probabilities from a stochastic vector  $\mathbf{p}_t$  with elements  $p_t(\omega)$ :

$$\mathbb{P}_t[\underline{\omega}^{(\tau)}] = p_t(\omega_0) \prod_{k=1}^{\tau} w_{\omega_k}^{\omega_{k-1}}.$$

Summing over all possible finite time series of a given length  $\tau$  which end in a state  $\omega_{t+\tau}$  yields the connection between the elements of  $\mathbf{p}^{(t)}$  and  $\mathbf{p}^{(t+\tau)}$ :

$$p_{t+\tau}(\omega_{t+\tau}) = \sum_{\underline{\omega}^{(T)}} \left[ p_t(\omega_t) \prod_{k=t+1}^{t+\tau} w_{\omega_k}^{\omega_{k-1}} \right]$$

where  $T := \{t, t+1, \dots, t+\tau-1\} \subset \mathbb{T}$  and the sum symbol  $\sum_{\underline{\omega}^{(T)}}$  is an abbreviation for the multi-sum

$$\sum_{\underline{\omega}^{(T)}} [\cdot] := \prod_{t_i \in T} \left[ \sum_{\omega_{t_i} \in \Omega} [\cdot] \right] := \sum_{\omega_{t_1} \in \Omega} \sum_{\omega_{t_2} \in \Omega} \dots \sum_{\omega_{t_k} \in \Omega} [\cdot].$$

### 3. Markovian symbolic dynamics

The special case for time series  $\underline{\omega}^{(1)} = (\omega, \omega')$  of length  $\tau = 1$  is called the *master equation*, cf. Sec. 2.4.2:

$$p_{t+1}(\omega') = \sum_{\omega} p_t(\omega) w_{\omega'}^{\omega}. \quad (3.2)$$

Markov processes are *memoryless processes*, in the sense that the probability of the next state only depends on the current state. Hence, at any time the process has completely “forgotten” its past.

Let us summarize the present subsection: We have introduced a formal way to treat sequences of random variables  $\underline{f} = (f_t)_{t \in \mathbb{T}}$  as a stochastic process, without specifying the nature of  $f_t$  or the underlying probability space  $X$ . On the other hand, the setting outlined at the beginning of this section, already suggests a certain interpretation: The underlying probability space is the phase space of some physical dynamics, i.e.  $X = \Gamma$ , and  $f_t$  should correspond to subsequent measurements. Such dynamical systems will be the subject of the next subsection.

#### 3.1.2. The stochastic process of observed time series

For the remainder of this chapter, we focus on *dynamical systems in discrete time*, which are characterized by a map  $\Phi$  from phase space  $\Gamma$  onto itself. We further assume that the system is *autonomous*, i.e. the map  $\Phi$  is independent of time  $t$ . Physically, we understand it as a *stroboscopic* map obtained from some physical microscopic dynamics  $\Psi$  evolving in continuous time, cf. Section 2.5.3.

We restate the definition of a measurable dynamical system from Section 2.6:

**3.7 Definition (Measurable dynamical system)** Let  $(\Gamma, \mathcal{B})$  be a measurable space and let  $\Phi: (\Gamma, \mathcal{B}) \rightarrow (\Gamma, \mathcal{B})$  be a measurable map. We call  $(\Gamma, \mathcal{B}, \Phi)$  a *measurable dynamical system*. If a probability measure  $\mu$  is given, we call  $(\Gamma, \mathcal{B}, \Phi, \mu)$  a dynamical system with measure  $\mu$ .

In statistical physics,  $\mu$  is also called a microscopic *statistical ensemble*. Then, it is understood to represent the (distribution of) configurations of a large number of copies of the system.

An observable  $M: (\Gamma, \mathcal{B}) \rightarrow (\Omega, \mathcal{A})$  is a measurable map from phase space  $\Gamma$  to the space of observations  $\Omega$ . As mentioned above, we assume that  $\Omega$  is of finite cardinality  $N$  and choose  $\mathcal{A} = \mathcal{P}(\Omega)$  as the  $\sigma$ -algebra on  $\Omega$ . With these ingredients we formalize the process of taking measurements at equidistant time steps as a stochastic process:

**3.8 Definition (Stochastic process of observed time series)** Let  $(\Gamma, \mathcal{B}, \Phi, \mu)$  be a dynamical system with measure  $\mu$  and  $M: (\Gamma, \mathcal{B}) \rightarrow (\Omega, \mathcal{A})$  a measurable function. The sequence

$$\begin{aligned} \underline{M} &:= (M_t)_{t \in \mathbb{T}} \\ &:= (M \circ \Phi^t)_{t \in \mathbb{T}}. \end{aligned} \quad (3.3)$$

is called the *stochastic process of observed time series*.

Note that  $\underline{M}$  is only well-defined for  $\mathbb{T} = \mathbb{Z}$  if  $\Phi$  is invertible. Henceforth, we will assume this to be the case.

The image of phase space under the trajectory of the process,  $\mathbb{S}_M^{\mathbb{T}} := \underline{M}(\Gamma) \subset \Omega^{\mathbb{T}}$  contains all the infinite time series that an observable  $M$  can produce. Further, we will prove that this set is invariant under shifting the whole sequence one step to the left, possibly dropping the zeroth element if  $\mathbb{T} = \mathbb{N}$ . In general, sets  $\mathbb{S}^{\mathbb{T}} \subset \Omega^{\mathbb{T}}$  that obey such a dynamic shift-invariance are called *subshifts*. The formal treatment of the topology and the dynamics of subshifts is the goal of *symbolic dynamics*.

### 3.1.3. Symbolic dynamics

Symbolic dynamics is the study of infinite symbol sequences produced by a finite *alphabet*  $\Omega$ . Hence, it is closely related to the study of *messages* in information theory, *cf.* Section 2.2. In the present case, the alphabet  $\Omega$  is the co-domain of the observable  $M$ . However, for now we will forget about  $M$  and the underlying dynamics. Rather, we first characterize certain subsets  $\mathbb{S}^{\mathbb{T}} \subset \Omega^{\mathbb{T}}$  which are invariant under the action of a *shift map*:

**3.9 Definition (Shift, alphabet, shift map)** Let  $\Omega$  be a finite set. The sets  $\Omega^{\mathbb{Z}}$  and  $\Omega^{\mathbb{N}}$  are called the *full shift* and *forward shift* over  $\Omega$ , respectively. The set  $\Omega$  is referred to as an *alphabet* and its elements are called *symbols*.

A finite string  $\underline{\omega}^{(\tau)} = (\omega_t)_{t \in \{0, 1, \dots, \tau\}} \in \Omega^{\tau+1}$  consisting of  $\tau + 1$  symbols is called a *block*.<sup>2</sup>

The *shift map*  $\hat{s}$  acts on (bi-)infinite symbol sequences  $\underline{\omega} = (\omega_t)_{t \in \mathbb{T}} \in \Omega^{\mathbb{T}}$  in shifting the whole sequence by one step to the left:

$$\begin{aligned} \hat{s}: \Omega^{\mathbb{T}} &\rightarrow \Omega^{\mathbb{T}}, \\ (\omega_t)_{t \in \mathbb{T}} &\mapsto (\omega_{t+1})_{t \in \mathbb{T}}. \end{aligned} \tag{3.4}$$

A *subshift* is a shift-invariant subset of the full shift:

**3.10 Definition (Subshift)** Let  $\mathbb{S}^{\mathbb{T}} \subset \Omega^{\mathbb{T}}$  be a subset of the full or forward shift, respectively. If  $\mathbb{S}^{\mathbb{T}}$  is *shift-invariant*, i.e. if  $\hat{s}\mathbb{S}^{\mathbb{T}} = \mathbb{S}^{\mathbb{T}}$ , we call  $\mathbb{S}^{\mathbb{T}}$  a *subshift*. The elements  $\underline{\omega} \in \mathbb{S}^{\mathbb{T}}$  are called *allowed* or *admissible* sequences.

By definition, the set of all cylinders is invariant under the action of the shift map  $\hat{s}$ . Hence, the  $\sigma$ -algebra generated by the cylinders is shift-invariant by construction. Thus,  $\hat{s}$  is measurable and specifies a measurable dynamical system on  $\mathbb{S}^{\mathbb{T}}$ :

**3.11 Definition (Symbolic dynamics)** Let  $\Omega^{\mathbb{T}}$  be a shift space,  $\hat{s}$  the shift map and  $\mathbb{S}^{\mathbb{T}}$  a subshift. Further, let  $\mathcal{A}_{\mathbb{S}^{\mathbb{T}}}$  be the restriction of the product  $\sigma$ -algebra on  $\Omega^{\mathbb{T}}$  to  $\mathbb{S}^{\mathbb{T}}$ . The dynamical system  $(\mathbb{S}^{\mathbb{T}}, \mathcal{A}_{\mathbb{S}^{\mathbb{T}}}, \hat{s})$  is called a *symbolic dynamical system* or just short a *symbolic dynamics*.

<sup>2</sup> Above we used the expressions “finite” and “infinite time series” for what in the framework of symbolic dynamics is called a block and a symbol sequence, respectively. We make no difference in their meaning, and distinguish them just by the mathematical framework they appear in.

### 3. Markovian symbolic dynamics

For the rest of this subsection, we are concerned with special kinds of subshifts. Consider a finite set  $F$  of blocks, possibly of different lengths. We will refer to  $\underline{\omega}^{(\tau)}$  as a *forbidden block* of length  $\tau + 1$ . For finite  $F$  it is obvious that the set

$$\mathbb{S} = \{ \underline{\omega} \in \Omega^{\mathbb{T}} \mid \underline{\omega} \text{ does not contain any block } \underline{\omega}^{(\tau)} \in F \text{ as a sub string} \}$$

is a subshift. More precisely, we call  $\mathbb{S}$  a  $m$ -step subshift of finite type (SFT), where  $m + 1$  denotes the maximal length of the blocks in  $F$ . For a much deeper treatment of this and related subjects, we refer to Refs. [Wei73; Wil73]. In addition, Ref. [BL89] contains a presentation which may be more accessible to physicists.

In the remainder of this work, we will only consider 1-step SFTs. For finite alphabets  $\Omega$ , this assumption can be made without any loss of generality: For an  $m$ -step SFT one simply considers the finite set  $\Omega^m$  as the alphabet of a larger shift space  $(\Omega^m)^{\mathbb{T}}$ . Henceforth we use ‘‘SFT’’ as a synonym for ‘‘1-step SFT’’.

A SFT on an alphabet  $\Omega$  with cardinality  $N$  can be characterized by a simple matrix:

**3.12 Definition (Subshift of finite type)** Let  $\mathbb{S}^{\mathbb{T}}$  be a subshift and let  $\mathbb{A}$  be a  $N \times N$ -matrix with entries  $a_{\omega'\omega''}^{\omega} \in \{0, 1\}$ . The set

$$\mathbb{S}^{\mathbb{T}} := \{ \underline{\omega} = (\omega_t)_{t \in \mathbb{T}} \mid a_{\omega_t \omega_{t+1}}^{\omega_t} = 1 \} \quad (3.5)$$

is called a (1-step) subshift of finite type (SFT). Then, the matrix  $\mathbb{A}$  is called the adjacency matrix of  $\mathbb{S}^{\mathbb{T}}$ .

The reader might have noticed the similarity of Equation (3.5) with the Markov property (3.1): Both state that the possibility to see a block (respectively finite time series)  $\underline{\omega}^{(\tau+1)} := (\underline{\omega}^{(\tau)}, \omega_{\tau+1})$  only depends on the last symbol  $\omega_{\tau}$  of  $\underline{\omega}^{(\tau)}$ . Hence, both equations characterize a memoryless process where the future only depends on the present and not on the past. The Markov property can be understood as specifying the *probability* rather than (or more precisely: in addition to) the *possibility* of the appearance of sequences. It is clear that the topological structure of a subshift  $\mathbb{S}^{\mathbb{T}} = \underline{f}(X)$  created by a Markovian stochastic process  $\underline{f}$  is a SFT. This is the reason why SFTs are also known in the literature as *topological Markov shifts* [AGW77], *topological Markov chains* [Kri80] or *intrinsic Markov chains* [Par64]. If a Markov chain has a transition matrix  $\mathbb{W}$ , the adjacency matrix of the corresponding SFT obeys  $\mathbb{A} = \text{sgn}(\mathbb{W})$ . In that case we say that the transition matrix  $\mathbb{W}$  is *compatible* with the adjacency matrix  $\mathbb{A}$ .

Now let us come back and draw a connection to the stochastic process of observed time series  $\underline{\omega} \in \mathbb{S}_M^{\mathbb{T}}$ :

**3.13 Proposition** Let  $\underline{M} := (M_t)_{t \in \mathbb{T}} = (M \circ \Phi^t)_{t \in \mathbb{T}}$  be the trajectory of a stochastic process of observed time series. Further let  $\mathbb{S}_M^{\mathbb{T}} := \underline{M}(\Gamma)$ . Then, the following statements are true:

- The shift map  $\hat{s}$  on  $\mathbb{S}_M^\mathbb{T}$  plays the role of  $\Phi$  on  $\Gamma$ , i.e.

$$\underline{M} \circ \Phi = \hat{s} \circ \underline{M}. \quad (3.6)$$

- The set of all possible infinite time series,  $\mathbb{S}_M^\mathbb{T}$ , is a subshift.

PROOF For the first statement observe that for  $x \in \Gamma$  we have

$$(\underline{M} \circ \Phi)(x) = (M\Phi^t \Phi x)_{t \in \mathbb{T}} = (M\Phi^{t+1} x)_{t \in \mathbb{T}} = \hat{s}(M\Phi^t x)_{t \in \mathbb{T}} = (\hat{s} \circ \underline{M})(x)$$

and hence  $\hat{s}\mathbb{S}_M^\mathbb{T} = \hat{s}\underline{M}\Gamma = \underline{M}\Phi\Gamma$ . Then, because  $\Phi\Gamma = \Gamma$  we have  $\hat{s}\mathbb{S}_M^\mathbb{T} = \mathbb{S}_M^\mathbb{T}$  and thus the second statement follows.  $\square$

In the next section, we investigate how the choice of the observable  $M$  influences the properties of the stochastic process of measured time series, cf. definition 3.8. Henceforth,  $\mathbb{S}_M^\mathbb{T} := \underline{M}(\Gamma)$  will contain the possible sequences generated by such a process.

### 3.2. Observables and partitions on phase space

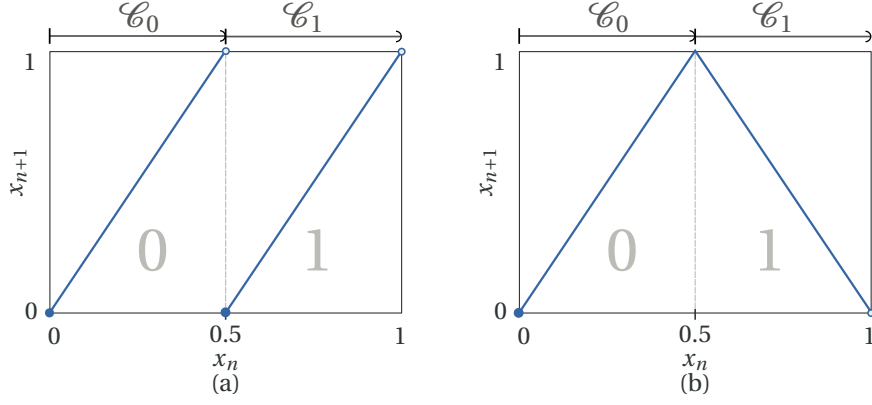
Equipped with the notions of an SFT and that of the stochastic process of observed time series, we can ask the following question: For a given dynamics  $\Phi$ , what are necessary conditions on the observable  $M$  such that the subshift  $\mathbb{S}_M^\mathbb{T}$  is of finite type?

Unfortunately, it turns out that this is not the case in general. Even for the most simple dynamics, there are certain sequences  $\underline{\omega}$  missing from  $\mathbb{S}_M^\mathbb{T}$ , which would be needed to make it an SFT. In a sense, the subshift  $\mathbb{S}_M^\mathbb{T}$  is “too small”: Every  $x \in \Gamma$  has *exactly one* symbolic representation, due to the fact that  $M$  is a function. In order to make  $\mathbb{S}_M^\mathbb{T}$  an SFT — and therefore a subshift with an easy topological structure — we need to admit *multiple* symbolic sequences  $\underline{\omega}$  for points  $x \in \Gamma$ .

In Ref. [Adl98], R. Adler discusses exactly this problem, which is also well-known from the decimal encoding of the real numbers. Consider the unit interval  $\Gamma = [0, 1)$ . Points  $x \in \Gamma$  have decimal representations  $0.\underline{\omega}$  where  $\underline{\omega} \in \{0, 1, 2, \dots, 9\}^\mathbb{N}$  is an semi-infinite string of digits. Obviously, the set  $\{0, 1, 2, \dots, 9\}^\mathbb{N}$  is an SFT, but the encoding of the unit interval by infinite symbols is not unique: The real number  $\frac{1}{2}$  has the two equivalent decimal expansions  $\frac{1}{2} = 0.4\bar{9} \equiv 0.5\bar{0}$ .

Let us formulate this observation in the language of generated time series. To that end, consider the *binary* encoding of the unit interval  $\Gamma = [0, 1)$ . As a dynamics  $\Phi$ , we consider “multiplication by two modulo one”, also known as the *Bernoulli map*, which is depicted in Figure 3.2a). The observable  $M$  is defined to yield 0 if  $0 \leq x < \frac{1}{2}$  and 1 otherwise. It partitions  $\Gamma$  into the intervals  $\mathcal{C}_0 = [0, \frac{1}{2})$  and  $\mathcal{C}_1 = [\frac{1}{2}, 1)$ . By definition,  $\underline{M}: \Gamma \rightarrow \mathbb{S}_M^\mathbb{N}$  is surjective. One can easily verify that the function  $\underline{M}^{-1}: (\omega_t)_{t \in \mathbb{N}} \mapsto \sum_{t \in \mathbb{N}} \omega_t 2^{-(t+1)}$  is its inverse and thus  $\underline{M}$  is a one-to-one map. However, observe that both  $\underline{M}^{-1}(1, 0, 0, \dots) = \underline{M}^{-1}(0, 1, 1, \dots) = 0.5$ . Because  $\underline{M}$  is bijective, one of these infinite sequences (in the

### 3. Markovian symbolic dynamics



**Figure 3.2.:** (a) The graph of the Bernoulli map  $\Phi: x \mapsto (2x \bmod 1)$  on the unit interval  $[0, 1]$ . The partition  $\mathcal{Q} = (\mathcal{C}_0, \mathcal{C}_1) = ([0, 0.5], [0.5, 1])$  for the observable  $M(x) = \chi_{[0.5, 1)}(x)$  creates a subshift  $\mathbb{S}_M^{\mathbb{N}}$ . However, this is not a subshift of finite type as the sequence  $\omega = (0, 1, 1, \dots)$  is not part of  $\mathbb{S}_M^{\mathbb{N}} = \underline{M}(\Gamma)$ . (b) The tent map  $\Phi: x \mapsto 2 \min\{x, 1-x\}$  continuously maps the unit interval  $\Gamma = [0, 1]$  to itself. It is topologically conjugate to the map shown in (a).

present case, the second one) is not an element of  $\mathbb{S}_M^{\mathbb{N}}$  and thus the latter cannot be an SFT. Unlike one may expect, the origin of that problem is *not* the discontinuity at  $x = \frac{1}{2}$ . The same is true for the point  $x = \frac{1}{2}$  in the continuous tent map shown in Fig. 3.2b, which is topologically conjugate to the Bernoulli map.

As one might already guess from these examples, the issue arises for points  $x \in \Gamma$  that lie (or are mapped to) the boundary of partition elements. The solution as presented in Ref. [Adl98] is to define the factor map  $K: (\omega_t)_{t \in \mathbb{N}} \mapsto \sum_{t \in \mathbb{N}} \omega_t 2^{-(t+1)}$  for *all* sequences in the SFT  $\mathbb{S}^{\mathbb{N}} = \{0, 1\}^{\mathbb{N}}$ . In this *extension* of  $M^{-1}$  on  $\mathbb{S}_M^{\mathbb{N}}$  to  $K$  on  $\mathbb{S}^{\mathbb{N}} \supset \mathbb{S}_M^{\mathbb{N}}$  we give up bijectivity for a nicer topology. This is the general idea behind *topological partitions*, which we review in Section 3.2.2.

In the following, it is instructive to look at the structure on phase space generated by  $M$  as opposed to  $M$  itself. In Section 2.6.3 we have already defined the concept of a *partition*  $(\mathcal{C}_\omega)_{\omega \in \Omega}$  of  $\Gamma$  as a subset of  $\mathcal{P}(\Gamma)$  whose elements are disjoint and whose union is  $\Gamma$ .

Now consider the set  $\mathcal{C}_\omega := M^{-1}(\{\omega\})$  which is the pre-image of a singleton subset  $\{\omega\}$ . The *partition induced by  $M$*  is defined as  $\mathcal{Q} := (\mathcal{C}_\omega)_{\omega \in \Omega}$ . We will refer to its elements  $\mathcal{C}_\omega$  as the *cells* induced by  $M$ . Note, that from any (finite) partition indexed by elements  $\omega \in \Omega$ , one can construct a map  $M: \Gamma \rightarrow \Omega$  defined by  $M(x) := \omega$  if  $x \in \mathcal{C}_\omega$ .

Henceforth, we will slightly abuse the notation and drop the set delimiters when we refer to singletons and write, for instance,  $\mathcal{C}_\omega = M^{-1}(\omega)$ . Further, we will use the expressions “observation”, “state” and “symbol” synonymously to refer to  $\omega \in \Omega$ , depending on whether we want to emphasize its role as the value of an observable, the state of a stochastic process or the symbol of an alphabet, respectively.



### 3.2.1. Generating partitions

In this subsection we are concerned with so-called *generating partitions*. Rather than embracing the topological definition of a generating partition (cf. Ref. [Adl98]), we use Sinai's measure-theoretic definition of such partitions [Sin09]: Informally speaking, the elements  $\mathcal{C}_\omega$  of a generating partition  $\mathcal{Q}$  and their (pre-)images under iterations of the dynamics  $\Phi$  generate the measurable structure on phase space.

We need the following elementary result of set theory:

**3.14 Lemma** *Let  $f: X_1 \rightarrow X_2$  be a map and  $\mathcal{F}_1 \subset \mathcal{P}(X_1)$  and  $\mathcal{F}_2 \subset \mathcal{P}(X_2)$ . Then it holds that*

$$\begin{aligned} f^{-1}\left(\bigcup_{B \in \mathcal{F}_2} B\right) &= \bigcup_{B \in \mathcal{F}_2} f^{-1}(B), & f^{-1}\left(\bigcap_{B \in \mathcal{F}_2} B\right) &= \bigcap_{B \in \mathcal{F}_2} f^{-1}(B), \\ f\left(\bigcup_{B \in \mathcal{F}_1} B\right) &= \bigcup_{B \in \mathcal{F}_1} f(B), & f\left(\bigcap_{B \in \mathcal{F}_1} B\right) &\subseteq \bigcap_{B \in \mathcal{F}_1} f(B). \end{aligned}$$

Further,  $f^{-1}(\sigma(\mathcal{F}_2)) = \sigma(f^{-1}(\mathcal{F}_2))$  and  $f(\sigma(\mathcal{F}_1)) \subseteq \sigma(f(\mathcal{F}_1))$ .

**PROOF** The first part of the claim is (trivially) verified using the definition of the image and pre-image of a point. The second claim then follows from the definition of the generated  $\sigma$ -algebra.  $\square$

Now consider the set  $\mathcal{Z}^a = \{\pi_{t-\mathbb{T}}^{-1}(\omega) \mid \omega \in \Omega, t \in \mathbb{T}\} \subset \Omega^{\mathbb{T}}$  consisting of the cylinders whose base (i.e. the image of the projection) is a singleton  $\{\omega\}$ . Because the base of an element  $Z^a \in \mathcal{Z}^a$  cannot be subdivided into non-empty sets, we call them the *atomic cylinders*. All (open) subsets of  $\Omega$  are countable (in fact finite) unions of singletons  $\{\omega\}$ , because  $\Omega$  is finite. Taking the pre-image of a set commutes with set-theoretic operations (cf. Lemma 3.14). This implies that the atomic cylinders are a large enough set to generate the product  $\sigma$ -algebra  $\mathcal{A}^{\mathbb{T}}$  on  $\Omega^{\mathbb{T}}$ , i.e.  $\sigma(\mathcal{Z}^a) = \mathcal{A}^{\mathbb{T}}$ . For the rest of this subsection, we take  $\mathbb{T} = \mathbb{Z}$ , and sequences  $\omega \in \mathbb{S}^{\mathbb{T}}$  are bi-infinite.

We denote the set  $\zeta^a := \bigcup_{t \in \mathbb{Z}} \Phi^{-t} \mathcal{Q}$ , which contains all pre-images and images of a partition  $\mathcal{Q} = M^{-1}\Omega$  under the (iterated) action of  $\Phi$ , as the *atomic cells*. Using Eq. (3.6), we find that  $\underline{M}$  applied to any atomic cell  $\mathcal{C}^a \in \zeta^a$  uniquely identifies an atomic cylinder:

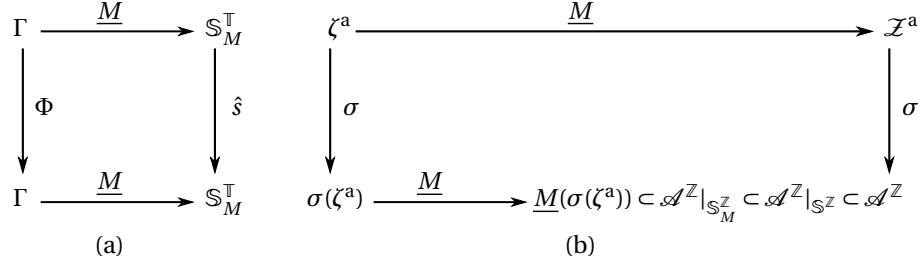
$$\underline{M}\Phi^t \mathcal{C}_\omega = s^t \underline{M}\mathcal{C}_\omega = s^t \pi_0^{-1}\{\omega\} = \pi_{-t}^{-1}\{\omega\} \in \mathcal{Z}^a.$$

Thus, taking the image  $\underline{M}$  of  $\zeta^a$  is a bijection onto  $\mathcal{Z}^a$ , i.e.  $\zeta^a = \underline{M}^{-1}\mathcal{Z}^a$  and  $\underline{M}\zeta^a = \mathcal{Z}^a$ . For a *generating partition*, the atomic cells generate the measurable structure on phase space:

**3.15 Definition (Generating partition)** Let  $(\Gamma, \mathcal{B}, \Phi)$  be a measurable dynamical system and  $\mathcal{Q}$  a partition of  $\Gamma$ . Let  $\zeta^a := \bigcup_{t \in \mathbb{Z}} \Phi^{-t} \mathcal{Q}$  denote the set of atomic cells.

A partition  $\mathcal{Q}$  is called *generating* for  $(\Gamma, \mathcal{B}, \Phi)$ , if  $\sigma(\zeta^a) = \mathcal{B}$ , i.e. if the  $\sigma$ -algebra generated by the atomic cells  $\zeta^a$  agrees with the  $\sigma$ -algebra for the dynamical system.

### 3. Markovian symbolic dynamics



**Figure 3.3.:** (a) A commuting diagram illustrating  $\hat{\sigma} \circ \underline{M} = \underline{M} \circ \Phi$ . If  $\Phi$  is invertible and  $\mathbb{T} = \mathbb{Z}$ , one can reverse the vertical arrows. If  $M$  is obtained from a generating partition w.r.t. the Borel sets, also the horizontal arrows can be reversed. (b) The relation between the atomic cylinders and the atomic cells and the corresponding generated  $\sigma$ -algebras. Every measurable subset of  $\mathbb{S}_M^{\mathbb{Z}}$  is also measurable in the restriction  $\mathcal{A}^{\mathbb{Z}}|_{\mathbb{S}^{\mathbb{Z}}}$  of the product  $\sigma$ -algebra  $\mathcal{A}^{\mathbb{Z}}$ .

The next lemma states that the images of measurable sets on phase space are measurable, cf. Fig. 3.3b:

**3.16 Lemma** Let  $\mathcal{Q}$  be a generating partition for the measurable dynamical system  $(\Gamma, \mathcal{B}, \Phi)$ . Let  $\underline{M} = (M \circ \Phi^t)_{t \in \mathbb{Z}}$  denote the stochastic process of observed time series. Let  $\mathcal{A}_M^{\mathbb{Z}} := \mathcal{A}^{\mathbb{Z}}|_{\mathbb{S}_M^{\mathbb{Z}}}$  be the symbol product  $\sigma$ -algebra restricted to  $\mathbb{S}_M^{\mathbb{Z}}$ . Then,  $\mathcal{B}$  is the pre-image of  $\mathcal{A}_M^{\mathbb{Z}}$  for  $\underline{M}$ , i.e.

$$\underline{M}^{-1}(\mathcal{A}_M^{\mathbb{Z}}) = \mathcal{B}$$

and the image of any measurable set  $B \in \mathcal{B}$  is measurable in  $\mathcal{A}_M^{\mathbb{Z}}$ , i.e.

$$\underline{M}(B) \subseteq \mathcal{A}_M^{\mathbb{Z}}.$$

PROOF For a generating partition, by definition we have  $\sigma(\zeta^a) \equiv \mathcal{B}$ . We already established that the symbolic product  $\sigma$ -algebra  $\mathcal{A}^{\mathbb{Z}}$  is generated by the atomic cylinders. Hence  $\mathcal{A}_M^{\mathbb{Z}} := \sigma(\mathcal{Z}^a)|_{\mathbb{S}_M^{\mathbb{Z}}}$ . The proof of the first statement is straightforward:

$$\begin{aligned} \underline{M}^{-1}(\sigma(\mathcal{Z}^a)|_{\mathbb{S}_M^{\mathbb{Z}}}) &= (\underline{M}^{-1}\sigma(\mathcal{Z}^a)|_{\underline{M}^{-1}\mathbb{S}_M^{\mathbb{Z}}}) \\ &= \sigma(\underline{M}^{-1}\mathcal{Z}^a)|_{\Gamma} = \sigma(\underline{M}^{-1}\mathcal{Z}^a) \\ &= \sigma(\zeta^a) \equiv \mathcal{B}, \end{aligned}$$

where to get to the second line we used that by definition  $\underline{M}^{-1}\mathbb{S}_M^{\mathbb{Z}} \equiv \underline{M}^{-1}\underline{M}\Gamma = \Gamma$  and Lemma 3.14. Also by Lemma 3.14, observe that  $\underline{M}(\sigma(\zeta^a)) \subseteq \sigma(\underline{M}\zeta^a) = \sigma(\mathcal{Z}^a)$ . Further, for  $B \in \sigma(\zeta^a)$  we have (by definition of  $\underline{M}$ ) that  $\underline{M}B \subset \mathbb{S}_M^{\mathbb{Z}}$  and thus  $\underline{M}\sigma(\zeta^a) \subset \mathcal{P}(\mathbb{S}_M^{\mathbb{Z}})$ . Hence, we have established that  $\underline{M}(\sigma(\zeta^a)) \subseteq \sigma(\mathcal{Z}^a) \cap \mathcal{P}(\mathbb{S}_M^{\mathbb{Z}}) \equiv \mathcal{A}_M^{\mathbb{Z}}$  and thus proved the second statement.  $\square$

Note that the notion of a generating partition is always defined with respect to the  $\sigma$ -algebra  $\mathcal{B}$  on phase space. If  $\mathcal{B}$  is the family of Borel sets, we have the following corollary:

**3.17 Corollary** *If  $\mathcal{Q}$  is a generating partition for  $(\Gamma, \mathcal{B}, \Phi)$ , where  $\mathcal{B}$  denotes the Borel sets,  $\underline{M}: \Gamma \rightarrow \mathbb{S}_M^{\mathbb{Z}}$  is an invertible map with measurable inverse.*

PROOF Let  $\mathcal{F}_\Gamma$  and  $\mathcal{F}_M$  contain the singleton subsets of  $\Gamma$  and  $\mathbb{S}_M^{\mathbb{Z}}$ , respectively. Then, to establish invertibility of  $\underline{M}$  we need to show bijectivity of  $\underline{M}: \mathcal{F}_\Gamma \rightarrow \mathcal{F}_M$ . Surjectivity is trivial, because  $\mathbb{S}_M^{\mathbb{Z}}$  is defined to be the image of  $\Gamma$  under  $\underline{M}$ . For injectivity, we show that the pre-image of  $\{\omega\}$  under  $\underline{M}$  is a singleton for any  $\omega \in \mathbb{S}_M^{\mathbb{Z}}$ . We will prove the result by contradiction. To that end let  $x \in \Gamma$ . By surjectivity, there is an element  $\omega \in \mathbb{S}_M^{\mathbb{Z}}$ , such that  $\underline{M}(x) = \omega$ . Now suppose that  $\underline{M}^{-1}\{\omega\}$  is not a singleton. Then, the first statement of Lemma 3.16 ensures that the singleton  $\{x\}$  cannot be an element of  $\mathcal{B}$ . However, this is a contradiction to the assumption that  $\mathcal{B}$  is the Borel  $\sigma$ -algebra. Hence,  $\underline{M}$  is injective and surjective and thus a bijection with inverse  $\underline{M}^{-1}: \mathbb{S}_M^{\mathbb{Z}} \rightarrow \Gamma$ . The fact that this inverse is measurable is a just the second statement of lemma 3.16.  $\square$

Under the assumptions of the above lemma, *i.e.* if  $\underline{M}$  is invertible, one says that it provides a *faithful encoding* of the dynamical system on phase space as a symbolic dynamics. Then, any measure on  $\mathbb{S}_M^{\mathbb{Z}}$  can be pushed forward by  $\underline{M}^{-1}$  to a Borel measure on phase space. In particular, we can *construct* certain measures in a much easier way referring to  $\mathbb{S}$  and  $\hat{s}$  instead of using the dynamics  $\Phi$ .

### 3.2.2. Topological and Markov partitions

The ultimate goal of this chapter is to find measures on  $\Gamma$ , such that the stochastic process of observed time series obeys Markovian statistics. A necessary requirement for this is that the space of allowed time series forms an SFT. However, we have seen that (generating) partitions induced by observables generally do not yield subshifts  $\mathbb{S}_M^{\mathbb{Z}}$  which are of finite type.

From the simple chaotic maps described at the beginning of this section (*cf.* Fig 3.2), we have identified the problem that arise from elements that lie on the boundary of the partition induced by  $M$ . In this subsection, we review how this problem is fixed by slightly enlarging  $\mathbb{S}_M^{\mathbb{Z}}$ .

From the measure-theoretic point of view, boundary points are negligible for sets  $\mathcal{C}_\omega$  with a Riemann-integrable characteristic function  $\chi_\omega(x)$ . Thus, R. Adler suggested the use of *topological partitions* rather than usual partitions [Adl98]. Topological partitions consist of open sets, whose closures cover the phase space:

**3.18 Definition (Topological partition)** Let  $(\Gamma, d)$  be a metric space with topology  $\mathcal{T}$  and let  $\mathcal{Q}_{\text{top}} = (\mathcal{C}_\omega)_{\omega \in \Omega} \subset \mathcal{T}$  be a finite collection of open sets. We call  $\mathcal{Q}$  a *topological partition* if

$$\begin{aligned} (i) \quad & \mathcal{C}_\omega \cap \mathcal{C}_{\omega'} = \emptyset, \omega \neq \omega', \\ (ii) \quad & \Gamma = \bigcup_{\omega \in \Omega} \overline{\mathcal{C}_\omega}, \end{aligned}$$

where the overbar denotes closure of sets.

### 3. Markovian symbolic dynamics

Thus, from a partition  $\mathcal{Q} = (\mathcal{C}_\omega)_{\omega \in \Omega} = M^{-1}(\Omega)$  of a metric phase space  $(\Gamma, d)$  we obtain a topological partition  $\mathcal{Q}_{\text{top}} = \overset{\circ}{\mathcal{Q}} := (\overset{\circ}{\mathcal{C}}_\omega)_{\omega \in \Omega}$ . Like a usual partition, a topological partition on a dynamical system creates a subshift  $\mathbb{S}^{\mathbb{Z}}$ . In order to construct this subshift, it is useful to define the following sets:

**3.19 Definition (Bi-infinitely refined cell)** Let  $\mathcal{Q}_{\text{top}} = (\overset{\circ}{\mathcal{C}}_\omega)_{\omega \in \Omega}$  be a topological partition. For  $\underline{\omega} \in \Omega^{\mathbb{Z}}$  we call the set

$$\mathcal{C}[\underline{\omega}] := \bigcap_{T=0}^{\infty} \overline{\bigcap_{t=-T}^{t=T} \Phi^{-t} \overset{\circ}{\mathcal{C}}_{\omega_t}}, \quad (3.7)$$

the *bi-infinitely refined cell* for  $\underline{\omega}$ .

Because  $\bigcap_{T=0}^{\infty} \overline{\bigcap_{t=-T}^{t=T} \Phi^{-t} \overset{\circ}{\mathcal{C}}_{\omega_t}} \subset \bigcap_{t=-\infty}^{t=\infty} \overline{\Phi^{-t} \overset{\circ}{\mathcal{C}}_{\omega_t}}$  the set  $\mathcal{C}[\underline{\omega}]$  contains only points such that  $\Phi^t x \in \overline{\overset{\circ}{\mathcal{C}}_{\omega_t}}$  [Adl98].<sup>3</sup> Assume for now that the topological partition  $\mathcal{Q}_{\text{top}} = (\overset{\circ}{\mathcal{C}}_\omega)_{\omega \in \Omega}$  is obtained from taking the interior of elements of a proper partition  $\mathcal{Q} = (\mathcal{C}_\omega)_{\omega \in \Omega}$  of phase space. Then, the orbit of any point  $x \in \mathcal{C}[\underline{\omega}]$  visits the *closures*  $\overline{\overset{\circ}{\mathcal{C}}_{\omega_t}} = \overline{\mathcal{C}_{\omega_t}}$  of the partition elements recorded by the time series  $\underline{\omega}$ , *i.e.* it is at least infinitesimally close to an orbit whose iterates  $x_t = \Phi^t x$  are in  $\mathcal{C}_{\omega_t}$ , where  $\omega_t$  is specified by the time series  $\underline{\omega}$ .

With that, we define the set of symbol sequences:

**3.20 Definition (Symbol sequences obtained from a topological partition)** Let  $\mathcal{Q}_{\text{top}}$  be a topological partition and  $\mathcal{C}[\underline{\omega}]$  the set defined in equation (3.7). Further, let  $(\Gamma, \Phi)$  be a dynamical system. Then, we define the *set of  $(\mathcal{Q}_{\text{top}}, \Phi)$ -sequences*  $\mathbb{S}^{\mathbb{Z}} \subset \Omega^{\mathbb{Z}}$  as

$$\mathbb{S}^{\mathbb{Z}} := \{\underline{\omega} = (\omega_t)_{t \in \mathbb{Z}} \mid \mathcal{C}[\underline{\omega}] \neq \emptyset\} \quad (3.8)$$

Thus this set consists of time series  $\underline{\omega}$ , such that there are orbits of points  $x \in \Gamma$  which are at any point in (or at least infinitely close to) the partition elements labelled by the indexed elements of  $\underline{\omega}$ . Adler proved that  $\mathbb{S}^{\mathbb{Z}}$  defined in equation (3.8) is indeed a subshift [Adl98]. Further, because  $\mathcal{C}_\omega \subseteq \overset{\circ}{\mathcal{C}}_\omega$  it contains the subshift  $\mathbb{S}_M^{\mathbb{Z}} \subset \mathbb{S}^{\mathbb{Z}}$ . Hence, the corresponding extended  $\sigma$ -algebra  $\mathcal{A}^{\mathbb{Z}} := \Omega^{\mathbb{Z}}|_{\mathbb{S}^{\mathbb{Z}}}$  contains  $\mathcal{A}_M^{\mathbb{Z}}$ . In the following, when talking about the stochastic symbolic dynamics of observed time series, we will always mean the measurable dynamical system  $(\mathbb{S}^{\mathbb{Z}}, \mathcal{A}^{\mathbb{Z}}, \hat{s})$ . Note, that the first statement of Lemma 3.16 also holds true for the slightly extended measurable symbolic dynamics: If  $\mathcal{Q}$  is generating, the image of any measurable set  $B \in \mathcal{B}$  is measurable in  $\mathcal{A}^{\mathbb{Z}} \supset \mathcal{A}_M^{\mathbb{Z}}$ . We will make use of that fact in Section 3.3.3.

**3.21 Remark** In Ref. [Adl98] the notion of a generating partition is purely topological. However, if  $\mathcal{Q}$  is generating with respect to the Borel  $\sigma$ -algebra on a metric space in the sense of Def. 3.15, then the topological partition  $\mathcal{Q}_{\text{top}} = \overset{\circ}{\mathcal{Q}}$  is generating in the topological sense. For completeness, we mention that in analogy to the example presented in Figure 3.2, a

<sup>3</sup>In general,  $\mathcal{C}[\underline{\omega}] \neq \bigcap_{t=-\infty}^{t=\infty} \overline{\Phi^{-t} \overset{\circ}{\mathcal{C}}_{\omega_t}}$ . However, the missing points are (pre-)images of boundary points and do not matter measure theoretically. Again, we refer to Ref. [Adl98] for the details.

factor map  $K: \mathbb{S}^{\mathbb{Z}} \rightarrow \Gamma$  can be constructed. It maps the time series  $\underline{\omega} \in \mathbb{S}^{\mathbb{Z}}$  to the unique element  $x \in \mathcal{C}[\underline{\omega}]$  and commutes with the dynamics, *i.e.*  $\Phi \circ K = K \circ \hat{s}$ , *cf.* Ref. [Adl98].

Now we can finally define a Markov partition:

**3.22 Definition (Markov partition)** Let  $\mathcal{Q}$  be a partition,  $\hat{\mathcal{Q}}$  the topological partition containing the interiors of sets in  $\mathcal{Q}$ , and  $\mathbb{S}^{\mathbb{Z}} \subset \Omega^{\mathbb{Z}}$  be the subshift generated by  $\hat{\mathcal{Q}}$  via Definition 3.20.

Then, if  $\mathbb{S}^{\mathbb{Z}}$  (and hence also its projection to positive times,  $\mathbb{S}^{\mathbb{N}}$ ) is a subshift of finite type, we call  $\mathcal{Q}$  (or  $\hat{\mathcal{Q}}$ ) a (topological) *Markov partition*.

Although our definition is equivalent to Adler's [Adl98], it is different from previous definitions [Sin68; Bow70]. In those references, a Markov partition is also always a topological generator, *i.e.* a partition that generates the Borel  $\sigma$ -algebra. For the remainder of the present chapter, we will work under the assumptions of Definition 3.22 with a Markov partition  $\mathcal{Q}$ . However, we do not demand that  $\mathcal{Q}$  is generating — neither in the topological, nor in the measure-theoretic sense.

Let us summarize the results of this section. For a dynamical system  $(\Gamma, \mathcal{B}, \Phi)$  we have introduced the notion of a generating partition. If we have such a generating partition, the images of measurable sets  $B \in \mathcal{B}$  under  $\underline{M}$  are measurable sets in the subshift  $\mathbb{S}_M^{\mathbb{Z}}$ . From any partition, we can obtain a topological partition consisting of the interiors of its elements. Via Def. 3.20 a topological partition defines a subshift  $\mathbb{S}^{\mathbb{Z}} \supset \mathbb{S}_M^{\mathbb{Z}}$ , which contains some additional sequences corresponding to orbits that move along the boundaries of partition elements. In the next section, we will assume that we have a Markov partition. Thus the dynamical system  $(\mathbb{S}^{\mathbb{Z}}, \mathcal{A}^{\mathbb{Z}}, \hat{s})$  is a measurable dynamical system on a SFT. Then, the associated (topological) partition on phase space allows for a *stochastic* symbolic dynamics with Markovian statistics for its time series.

### 3.3. Markov measures

In this section, we define probability measures  $\mu$  on SFTs  $\mathbb{S}^{\mathbb{T}}$ , such that the probabilities of  $t$ -shifted forward cylinders,  $\mathbb{P}_t[\underline{\omega}^{(t)}] := \mu(Z_t[\underline{\omega}^{(t)}])$  obey the Markov property (3.1) for  $t \in \mathbb{T}$ . We will see that there is a non-trivial difference between the case of  $\mathbb{T} = \mathbb{N}$  and  $\mathbb{T} = \mathbb{Z}$  when it comes to transient, *i.e.* not shift-invariant measures.

#### 3.3.1. Markov measures for semi-infinite sequences

In the present subsection, we are only interested in the forward dynamics, *i.e.* we set  $\mathbb{T} = \mathbb{N}$  and consider semi-infinite time series. Hence, we always consider the  $\sigma$ -algebra  $\mathcal{A}^{\mathbb{N}}$  on  $\Omega^{\mathbb{N}}$ .

**3.23 Definition (Markov measure)** Let  $\mathbb{S}^{\mathbb{N}} \subset \Omega^{\mathbb{N}}$  be an SFT with adjacency matrix  $\mathbb{A}$ ,  $\tau \in \mathbb{N}$  and  $Z_0[\underline{\omega}^{(\tau)}]$  be a forward cylinder. Further, let  $\mathbb{W} = (w_{\omega'}^{\omega})_{\omega, \omega' \in \Omega}$  be a stochastic matrix

### 3. Markovian symbolic dynamics

compatible with  $\mathbb{A}$  and  $\mathbf{p}_0 = (p_0(\omega))_{\omega \in \Omega}$  a stochastic vector. A measure  $\bar{\mu} = \bar{\mu}(\mathbb{W}, \mathbf{p}_0)$  on  $(\Omega^{\mathbb{N}}, \mathcal{A}^{\mathbb{N}})$  obeying

$$\mathbb{P}_0[\underline{\omega}^{(\tau)}] \equiv \bar{\mu}(Z_0[\underline{\omega}^{(\tau)}]) = p_0(\omega_0) \prod_{k=1}^{\tau} w_{\omega_k}^{\omega_{k-1}}$$

is called a (one-sided) *Markov measure*.

In Appendix A.1 we use the Daniell–Kolmogorov extension theorem to prove that this measure exists and is unique for any given choice of  $\mathbb{W}$  and  $\mathbf{p}_0$ . Note that the forward arrow on  $\bar{\mu}$  is not a vector arrow. Rather, it indicates that the Markov measure is a “forward measure”, *i.e.* it is defined for the *semi-infinite* sequences extending in forward time. Further, in Appendix A.1 we proof the following property of the Markov measure:

**3.24 Proposition** *Let  $\bar{\mu}_0 \equiv \bar{\mu}(\mathbb{W}, \mathbf{p}_0)$  be a Markov measure. Then, for  $t \in \mathbb{N}$*

$$\bar{\mu}_t := s_*^t \bar{\mu}_0 = \bar{\mu}(\mathbb{W}, \mathbf{p}_t)$$

where  $\mathbf{p}_t = \mathbf{p}_0 \mathbb{W}^t$ .

Thus, the Markov measure  $\bar{\mu}_\infty = \bar{\mu}(\mathbb{W}, \mathbf{p}_\infty)$  is stationary if and only if the stochastic vector  $\mathbf{p}_\infty$  is a left eigenvector of  $\mathbb{W}$ . If the SFT  $\mathbb{S}^{\mathbb{N}}$  is *irreducible* (*i.e.* if its adjacency matrix  $\mathbb{A}$  is that of a strongly connected graph),  $\mathbf{p}_\infty$  and thus  $\bar{\mu}_\infty$  are uniquely determined by a compatible transition matrix  $\mathbb{W}$ . If the SFT is additionally aperiodic,  $\lim_{t \rightarrow \infty} (\bar{\mu}_t) = \bar{\mu}_\infty(\mathbb{W}, \mathbf{p}_\infty)$  exists and is the same for *any* (transient) Markov measure. Under these constraints,  $\bar{\mu}_\infty$  was first introduced by W. Parry [Par64]. He further showed that there is a unique stationary Markov measure  $\bar{\mu}_\infty^P$ , which maximizes the Kolmogorov–Sinai entropy (2.83) of the measurable dynamical system  $(\mathbb{S}^{\mathbb{N}}, \mathcal{A}^{\mathbb{N}}, \delta)$ . This measure has become known as the *Parry measure*.

#### 3.3.2. Markov measures for bi-infinite sequences

A stationary Markov measure can be extended to a stationary Markov measure on  $\mathbb{S}^{\mathbb{Z}}$  by slightly changing definition 3.23:

**3.25 Definition (Stationary Markov measure on the bi-infinite sequences)** Let  $\mathbb{S}^{\mathbb{Z}} \subset \Omega^{\mathbb{Z}}$  be an SFT with an adjacency matrix  $\mathbb{A}$ . For  $t \in \mathbb{Z}$  let  $Z_t[\underline{\omega}^{(\tau)}]$  be a  $t$ -shifted forward cylinder. Further, let  $\mathbb{W} = (w_{\omega'}^{\omega})_{\omega, \omega' \in \Omega}$  be a stochastic matrix compatible with  $\mathbb{A}$  and  $\mathbf{p}_\infty = (p_\infty(\omega))_{\omega \in \Omega}$  a stochastic left eigenvector of  $\mathbb{W}$ . A measure  $\mu_\infty = \mu_\infty(\mathbb{W}, \mathbf{p}_\infty)$  obeying

$$\mathbb{P}_t[\underline{\omega}^{(\tau)}] \equiv \mu_\infty(Z_t[\underline{\omega}^{(\tau)}]) = p_\infty(\omega_0) \prod_{k=1}^{\tau} w_{\omega_k}^{\omega_{k-1}}$$

is called a *stationary Markov measure for the bi-infinite shift*.

Again, this measure exists, is unique and shift-invariant (*cf.* appendix A.1). If  $\mathbb{S}^{\mathbb{Z}}$  is irreducible,  $\bar{\mu}_\infty$  is uniquely determined by  $\mathbb{W}$ .

Apart from stationary Markov measures, there are other *non-stationary* extensions to the bi-infinite sequences. Maybe surprisingly, it turns out that any non-stationary extension to the full shift violates the Markov property (3.1) for  $t$  smaller than some finite  $t_0 \in \mathbb{Z}$ . This can be easily understood from the master equation (3.2), which is a consequence of the Markov property. It states that the probability vectors  $\mathbf{p} := \mathbf{p}_t$  and  $\mathbf{p}' := \mathbf{p}_{t+1}$  must obey  $\mathbf{p}' = \mathbf{p}\mathbb{W}$  independently of  $t$ . Clearly, this condition is fulfilled for the stationary Markov measure, *i.e.* when  $\mathbf{p} = \mathbf{p}'$  is a left unity-eigenvector of  $\mathbb{W}$ .

It is instructive to attempt a construction of a non-stationary Markov measure for the full shift in the same spirit as the one-sided Markov measure. Then, we can use the master equation to see where and why such a measure fails to fulfil the Markov property for times lower than some finite time  $t_0$ : Generically, a stochastic matrix  $\mathbb{W}$  is invertible. Hence,  $\mathbb{W}^t$  and thus  $\mathbf{p}_t := \mathbf{p}\mathbb{W}^t$  is well-defined for all  $t \in \mathbb{Z}$  for any (initial) stochastic vector  $\mathbf{p}_0$ . At first glance, this might seem like a reasonable ansatz. However, in general the inverse of a stochastic matrix is itself not a stochastic matrix,<sup>4</sup> with the consequence that  $\mathbf{p}_t$  for  $t < 0$  is not a stochastic vector any more. Generally,  $\mathbb{W}^{-1}$  is not even positive semi-definite. Thus,  $\mathbf{p}_t$  obtains negative entries for some  $t < t_0 \leq 0$ . The same problem persists also for singular stochastic matrices, given that solutions to the master equation can be found at all.

This leads to an important insight: The *only* stationary measure on the full subshift that obeys the master equation (3.2) for all times is a stationary Markov measure. We will provide a physical interpretation of this statement in Section 3.4.2.

Let us finally discuss another kind of extension of the Markov measure to the full shift, which we will refer to as the *two-sided Markov measure*. Such measures can be used to formalize the notion of a stochastic backward process. A backward process specifies the dynamics one would observe if time were to run backward after the initialization at the initial time  $t_0 = 0$ . The notion of a backward process is crucial for a general formulation of stochastic fluctuation relations [Sei12]. In order to define a backward process, we give the following definition:

**3.26 Definition** Let  $\underline{\omega}^{(\tau)} = (\omega_0, \omega_1, \dots, \omega_\tau) \in \Omega^{\tau+1}$  be a finite time series of length  $\tau + 1$  and  $t \in \mathbb{T}$ . We define the *t-shifted backward cylinder* for  $\underline{\omega}^{(\tau)}$  as

$$\tilde{Z}_t[\underline{\omega}^{(\tau)}] := \bigcap_{k=0}^{\tau} \pi_{t-k}^{-1}\{\omega_k\} \quad (3.9)$$

$$\equiv \{\underline{v} \in \Omega^{\mathbb{T}} \mid v_{t-k} = \omega_k, 0 \leq k \leq \tau\}. \quad (3.10)$$

The measure of such a cylinder,  $\tilde{\mathbb{P}}_t[\underline{\omega}^{(\tau)}] = \mathbb{P}(\tilde{Z}_t[\underline{\omega}^{(\tau)}])$ , is called the *probability* for the finite time series  $\underline{\omega}^{(\tau)}$  to *occur in reverse direction* at time  $t$ .

Note that, by definition, we have  $Z_t[\underline{\omega}^{(\tau)}] = \tilde{Z}_{t+\tau}[\mathbb{R}\underline{\omega}^{(\tau)}]$  where  $\mathbb{R}: \Omega^{\tau+1} \mapsto \Omega^{\tau+1}$  is the reversal operator mapping  $(\omega_k)_{0 \leq k \leq \tau} \mapsto (\omega_{\tau-k})_{0 \leq k \leq \tau}$ .

<sup>4</sup> This is only the case if  $\mathbb{W}$  is both orthogonal and double stochastic. Then, the transposed matrix  $\mathbb{W}^T$  is its inverse.

### 3. Markovian symbolic dynamics

Using the backward cylinders, we define:

**3.27 Definition (Two-sided non-stationary Markov measure)** Let  $\mathbb{S}^{\mathbb{Z}} \subset \Omega^{\mathbb{Z}}$  be an SFT with adjacency matrix  $\mathbb{A}$ . Let  $\tau \in \mathbb{N}$  and let  $Z_0[\underline{\omega}^{(\tau)}]$  and  $\tilde{Z}_0[\underline{\omega}^{(\tau)}]$  be forward and backward cylinders, respectively. Let  $\mathbb{W} = (w_{\omega'}^{\omega})_{\omega, \omega' \in \Omega}$  and  $\tilde{\mathbb{W}} = (\tilde{w}_{\omega'}^{\omega})_{\omega, \omega' \in \Omega}$  be stochastic matrices compatible with  $\mathbb{A}$  and  $\mathbb{A}^T$ , respectively. Let  $\mathbf{p}_0 = (p_0(\omega))_{\omega \in \Omega}$  be a stochastic vector. A measure  $\tilde{\mu} = \tilde{\mu}(\mathbb{W}, \tilde{\mathbb{W}}, \mathbf{p}_0)$  obeying

$$\begin{aligned} \mathbb{P}_0[\underline{\omega}^{(\tau)}] &\equiv \tilde{\mu}(Z_0[\underline{\omega}^{(\tau)}]) = p_0(\omega_0) \prod_{k=1}^{\tau} w_{\omega_k}^{\omega_{k-1}} \\ \tilde{\mathbb{P}}_0[\underline{R}\omega^{(\tau)}] &\equiv \tilde{\mu}(\tilde{Z}_0[\underline{R}\omega^{(\tau)}]) = p_0(\omega_0) \prod_{k=1}^{\tau} \tilde{w}_{\omega_{k-1}}^{\omega_k} \end{aligned}$$

and whenever  $Z[\underline{\omega}^{(\tau)}, \underline{R}\mathbf{v}^{(\tau)}] := Z_0[\underline{\omega}^{(\tau)}] \cap \tilde{Z}_0[\underline{R}\mathbf{v}^{(\tau)}] \neq \emptyset$ :

$$\tilde{\mu}(Z[\underline{\omega}^{(\tau)}, \underline{R}\mathbf{v}^{(\tau)}]) = p_0(\omega_0) \prod_{k=1}^{\tau} [w_{\omega_k}^{\omega_{k-1}}] \prod_{k=1}^{\tau} [\tilde{w}_{\mathbf{v}_{k-1}}^{\omega_k}]$$

is called a *two-sided non-stationary Markov measure for the bi-infinite shift*.

In Appendix A.1 we proof that this measure exists and is unique. The following proposition allows for a more intuitive definition of the two-sided Markov measure:

**3.28 Proposition** Let  $\mathbb{S}^{\mathbb{Z}}$  be a subshift and  $\mathbb{S}^+ = \pi_{\mathbb{N} \leftarrow \mathbb{Z}} \mathbb{S}^{\mathbb{Z}} \subset \Omega^{\mathbb{N}}$  its restriction to the forward sequences. For a measurable set  $A \in \mathcal{A}^{\mathbb{Z}}$  define the projections  $A^+ := \pi_{\mathbb{N} \leftarrow \mathbb{Z}} A$  and  $A^- := \pi_{-\mathbb{N} \leftarrow \mathbb{Z}} A$  onto the forward and backward trajectories, respectively. Then,

$$\tilde{\mu}(A) = \frac{\tilde{\mu}^+(A^+) \tilde{\mu}^-(RA^-)}{\tilde{\mu}^+(A^0)}.$$

where  $\tilde{\mu} = \tilde{\mu}(\mathbb{W}, \tilde{\mathbb{W}}, \mathbf{p}_0)$  is a two-sided Markov measure and the measures  $\tilde{\mu}^+ = \tilde{\mu}(\mathbf{p}_0, \mathbb{W})$  and  $\tilde{\mu}^- = \tilde{\mu}(\mathbf{p}_0, \tilde{\mathbb{W}})$  are Markov measures on  $\mathbb{S}^+$  and  $\mathbb{S}^- = \mathbb{R}\mathbb{S}^+$ , respectively.

The stationary Markov measure and the two-sided Markov measure are just two examples of an extension to the full shift. Many other non-stationary extensions of one-sided Markov measures are possible. However, for the purpose of this work we only consider the stationary and the two-sided Markov measure on the full shift. In the next subsection, we will finally connect Markov measures on  $\mathbb{S}^{\mathbb{Z}}$  to a dynamical system on phase space.

#### 3.3.3. (Natural) Markov measures on phase space

For the remainder of this subsection, let  $\mathcal{Q} = M^{-1}\Omega$  be a generating Markov partition for  $(\Gamma, \mathcal{B}, \Phi)$ . Then, Lemma 3.16 ensures that the image of any measurable set  $B \in \mathcal{B}$  is measurable in  $\mathcal{A}^{\mathbb{Z}}$ . Thus, any (Markov) measure  $\mu$  on  $(\Omega^{\mathbb{Z}}, \mathcal{A}^{\mathbb{Z}})$  yields a measure  $\mu_{\Gamma} := \mu \circ \underline{M}$  on  $(\Gamma, \mathcal{B})$ . More precisely, we consider the push-forward measures  $\underline{M}_*^{-1} \tilde{\mu}_{\infty}$  and  $\underline{M}_*^{-1} \tilde{\mu}$  of the stationary and the two-sided Markov measure, respectively:



**3.29 Definition (Markov measure on phase space)** Let  $(\Gamma, \mathcal{B}, \Phi)$  be a measurable dynamical system. Further, let  $(\mathbb{S}^Z, \mathcal{A}^Z, \hat{s})$  be the symbolic dynamical system obtained from a generating Markov partition  $\mathcal{Q} = M^{-1}(\Omega)$ . Let  $\bar{\mu}(\mathbb{W}, \mathbf{p}_\infty)$  and  $\bar{\mu}(\mathbb{W}, \tilde{\mathbb{W}}, \mathbf{p}_0)$  be stationary and two-sided Markov measures on  $\mathbb{S}$ . Then, we call

$$\begin{aligned}\bar{\mu}^\Gamma(\mathbb{W}, \mathbf{p}_\infty) &:= \underline{M}_*^{-1} \bar{\mu}(\mathbb{W}, \mathbf{p}_\infty) \text{ and} \\ \bar{\mu}^\Gamma(\mathbb{W}, \tilde{\mathbb{W}}, \mathbf{p}_0) &:= \underline{M}_*^{-1} \bar{\mu}(\mathbb{W}, \tilde{\mathbb{W}}, \mathbf{p}_0)\end{aligned}$$

*stationary and two-sided Markov measures on phase space, respectively.*

We are mostly interested in Borel measures on phase space. In that case, the topological and measure-theoretic notion of a generating partition agree.<sup>5</sup>

As of now, we have not required anything else of  $\mathbb{W}$  other than that it is compatible with the adjacency matrix  $\mathbb{A}$  of  $\mathbb{S}^Z$ . In the case where  $\mathbb{S}^Z$  is obtained from a phase-space dynamics,  $\mathbb{A}$  only contains information about the topological structure of the latter. However, we would like to identify a natural transition matrix in order to capture the natural behaviour one would expect for real experiments. To that end, we first need to establish what we mean by a “natural” behaviour for a physical dynamical system. In fact, the concept of a *natural measure* already exists and does express exactly what we are looking for [You02; BB03]:

**3.30 Definition (Natural measure)** A probability measure  $\mu_\Phi$  is called *natural* (with respect to  $\Phi$  and the reference measure  $\nu$ ) if there exists an open subset  $U \in \mathcal{F}$  such that for any absolutely continuous measure  $\mu \ll \nu$  with support  $U_\mu \subset U$  we have:

$$\frac{1}{\tau} \sum_{t=0}^{\tau} (\Phi^t)_* \mu \xrightarrow{\tau \rightarrow \infty} \mu_\Phi.$$

In other words,  $\mu_\Phi$  is a stable fixed point of the push-forward operator  $\Phi_*$ . It can be constructed from the action of  $\Phi_*$  on absolutely continuous measures. The open subset  $U$  is called the *basin of attraction* of  $\mu_\Phi$ : All measures with support in  $U$  are “attracted” towards  $\mu_\Phi$  as time passes. In the following, we take the Lebesgue measure as a reference, *i.e.*  $\nu = \lambda$ . Hence, the above definition means that all measures which admit a probability density in  $U$  “average” to  $\mu_\Phi$ .

Let us now assume that  $U = \Gamma$ , *i.e.* the measure has a basin of attraction that is the whole phase space. Further, we assume that the measure is ergodic. Under these assumptions, we define the *natural transition matrix*:

**3.31 Definition (Natural transition matrix)** Let  $(\Gamma, \mathcal{B}, \Phi)$  be a measurable dynamical system with a Markov partition  $\mathcal{Q} = (\mathcal{C}_\omega)_{\omega \in \Omega}$ . Let  $\mu_\Phi$  be an ergodic natural measure with basin of

<sup>5</sup>Note that in Def. 3.29 we do not require that the partition generated the Borel sets.

### 3. Markovian symbolic dynamics

attraction  $U = \Gamma$ . Then, the matrix  $\mathbb{Q} = \mathbb{Q}(\mu_\Phi, \mathcal{Q})$  with entries

$$q_{\omega'}^\omega := \frac{\mu_\Phi(\Phi^{-1}\mathcal{C}_{\omega'} \cap \mathcal{C}_\omega)}{\mu_\Phi(\mathcal{C}_\omega)}$$

is called the *natural transition matrix*.

Ergodicity of  $\mu_\Phi$  together with the fact that  $\Phi$  is onto, ensures that  $\mathbb{Q}$  is also ergodic and hence irreducible and aperiodic. Thus, there exists a unique *natural stationary measure*  $\bar{\mu}_\Gamma := \underline{M}_*^{-1}\bar{\mu}(\mathbb{Q})$ .

In Section 4.2, we encounter a special natural Markov measure on phase space defined for invertible dynamics  $\Phi$ . Recalling Proposition 3.13, denote by  $\widetilde{M} := (M \circ \Phi^{-t})_{t \in \mathbb{T}}$  the trajectory map of the stochastic symbolic dynamics generated by  $\Phi^{-1}$ . Further, the (topological) partition associated to  $M$  defines the subshift  $\widetilde{\mathcal{S}} \supset \widetilde{\mathcal{S}}_M := \widetilde{M}(\Gamma)$ . Note that the latter contains all the sequences of the subshift  $\mathcal{S} \supset \mathcal{S}_M = M(\Gamma)$  in reverse order, *i.e.*  $\widetilde{\mathcal{S}} = \mathbb{R}(\mathcal{S})$ . In particular, if  $\mathcal{S}$  is an SFT with adjacency matrix  $\mathbb{A}$ ,  $\mathbb{R}\mathcal{S}$  is an SFT characterized by  $\mathbb{A}^\top$ .

Hence, if  $\mathcal{Q}$  is a (generating) Markov partition for  $\Phi$  it is also one for the inverse  $\Phi^{-1}$ . Thus, if the natural transition matrix  $\mathbb{Q}$  for  $\Phi$  exists, so does the natural transition matrix  $\widetilde{\mathbb{Q}}$  defined using the natural measure  $\mu_{\Phi^{-1}}$ . This enables the definition of the *natural two-sided Markov measure* for invertible  $\Phi$ :

**3.32 Definition** Let  $\mathbb{Q}$  and  $\widetilde{\mathbb{Q}}$  be the natural transition matrices with respect to  $\Phi$  and  $\Phi^{-1}$ , respectively. Then, the measure  $\bar{\mu}(\mathbb{Q}, \widetilde{\mathbb{Q}}, \mathbf{p}_0)$  and its phase space analogue,  $\bar{\mu}^\Gamma(\mathbb{Q}, \widetilde{\mathbb{Q}}, \mathbf{p}_0)$  are called the *natural two-sided measures with the initial condition  $\mathbf{p}_0$*  on the symbolic and phase space dynamics, respectively.

The natural two-sided Markov measure has already been used implicitly in the literature on so-called *multibaker maps* [Gas05; VTB97; VTB98; BTV98; TV00; Vol02; Col+11]. In Chapter 4 we discuss the role of the two-sided natural measure for *network multibaker maps*, which are the most general variant of multibaker maps. Additionally, we discuss the (stochastic) thermodynamic properties of and the notion of entropy for such maps in a general information-theoretic framework.

## 3.4. Discussion

After the technical considerations above, we put our findings in the context of previous mathematical and physical results. Firstly, we revisit the concept of natural measures  $\mu_\Phi$ , which we used to define the natural transition matrix. In particular, we discuss their relation to the natural stationary Markov measure  $\bar{\mu}_\infty(\mathbb{Q}, \mathbf{p}_\infty)$ .

After that, we relate our abstract results to the *operational* framework outlined at the beginning of Section 3.1. There, we see how our abstract mathematical results obtain natural physical interpretations.

### 3.4.1. Connection to ergodic theory

An ergodic natural measure with full support on the whole phase space is the attractor of measures that are absolutely continuous with respect to the Lebesgue measure. Then, any probability measure with a density will converge to the natural measure. Further, the natural transition matrix is well-defined, *cf.* Def. 3.31. Note that generically the natural measure is *not* absolutely continuous.

For the remainder of this subsection, we stick with the above assumptions. Then, the Markov chain defined by the natural transition matrix is both ergodic and aperiodic with unique stationary distribution  $\mathbf{p}_\infty$ . Hence, any extensions of a one-sided natural Markov measure eventually converges to the unique stationary Markov measure  $\bar{\mu}_\infty(\mathbb{W}) := \bar{\mu}_\infty(\mathbb{W}, \mathbf{p}_\infty)$ .

Let us now assume that the partition  $\mathcal{Q}$ , which we use to define the natural transition matrix, generates the Borel  $\sigma$ -algebra. Further, let  $\mu$  be an extension of a (non-stationary) natural one-sided Markov measure  $\mu(\mathbb{Q}, \mathbf{p})$ . In that case,  $\mu^\Gamma := \mu \circ \underline{M}$  is a Borel measure on phase space. Two questions naturally arise:

- What is the relation of the natural measure  $\mu_\Phi$  and

$$\mu_\infty^\Gamma := \lim_{t \rightarrow \infty} (\Phi^t)_* \mu^\Gamma,$$

given the latter exists?

- Under which conditions is a non-stationary measure  $\mu^\Gamma$  absolutely continuous with respect to the Lebesgue measure?

Let us first assume a positive answer to the second question. In that case,  $\mu^\Gamma$  necessarily converges to the natural measure under push-forward by  $\Phi$  and we have  $\mu_\infty^\Gamma = \mu_\Phi$ . Moreover,  $\mu$  converges to the unique natural stationary measure  $\bar{\mu}_\infty(\mathbb{Q})$  defined on the whole shift, *i.e.* for  $\mathbb{T} = \mathbb{Z}$ . Because  $\underline{M} \circ \Phi = \hat{s} \circ \underline{M}$ , we find that the natural stationary measure on phase space actually *agrees* with the natural measure, *i.e.*

$$\bar{\mu}_\infty^\Gamma(\mathbb{Q}) := \bar{\mu}_\infty(\mathbb{Q}) \circ \underline{M} = \mu_\infty^\Gamma = \mu_\Phi. \quad (3.11)$$

However, this fact crucially depends on the fact that  $\mu^\Gamma$  was absolutely continuous in the first place. Remember that we demanded that  $\mu^\Gamma$  is a pull-back to phase space of an (arbitrary) extension of a natural one-sided measure  $\bar{\mu}(\mathbb{Q})$ . The latter is defined on the symbolic  $\sigma$ -algebra  $\mathcal{A}^\mathbb{N}$  for the forward sequences. This  $\sigma$ -algebra uniquely defines the restricted  $\sigma$ -algebra  $\mathcal{B}^\mathbb{N} = \underline{M}^{-1}(\mathcal{A}^\mathbb{N})$ .

The most refined sets in  $\mathcal{B}^\mathbb{N}$  are the pre-images of semi-infinite sequences  $\underline{\omega} \in \mathbb{Q}^\mathbb{N}$  under  $\underline{M}$ . That is, they are sets containing points with the same *symbolic future*. Such sets lie on the same (local) stable manifold. Specifying a measure on  $\mathcal{B}^\mathbb{N}$  can thus be understood at defining a *marginalized* measure, where the local stable manifold is integrated out. What remains is the topological structure that governs the *future* of points, which are also

### 3. Markovian symbolic dynamics

known as the (local) unstable manifolds. One also says that a measure on  $\mathcal{B}^{\mathbb{N}}$  defines a *conditional measure on the unstable manifolds* [You02].

The stable and unstable manifolds always intersect transversally. In a sense, they define a curvilinear transversal coordinate system [Adl98]. Extending a one-sided Markov measure to the bi-infinite sequences thus means specifying a structure along the stable directions. Or to put it differently: Specifying the *past* of a system defines a density along the stable directions, specifying the *future* does the same along the unstable directions.

Let us now choose the measure  $\mu$  as an extension of  $\bar{\mu}(\mathbb{Q})$  such that along the stable direction we have a density with respect to the (restricted) Borel measure. Then, the question whether  $\mu$  converges to the natural measure can be reformulated: Is the measure  $\bar{\mu}^{\Gamma}(\mathbb{Q})$  (restricted to  $\mathcal{B}^{\mathbb{N}}$ ) absolutely continuous with respect to  $\lambda$  (restricted to  $\mathcal{B}^{\mathbb{N}}$ )? Or equivalently: Does the measure  $\mu^{\Gamma}$ , constructed as an extension of a *natural* one-sided Markov measure, have absolutely continuous conditional measures on the unstable manifolds?

A positive answer to that question for *invariant* measures defines (a generalization) of a so-called SRB measure [You02]. The latter were discovered by Sinai, Ruelle and Bowen in 1970s in the context of Anosov systems and Axiom-A flows [Sin72; BR75]. The natural measure has been proposed to be the proper generalization of the SRB measure in a general context [You02; BB03].

These facts suggest that it is useful to think of natural measures as stationary Markov measures defined for a generating Markov partition. In fact, the SRB measure can be defined using generating Markov partitions for hyperbolic systems [Sin72; GC95]. Hence, it seems natural to ask whether under the present assumptions, the invariant natural measure on phase space  $\mu_{\infty}$  agrees with the natural measure  $\mu_{\Phi}$ .

What we know is that both measures (by definition) agree on sets of the form  $\mathcal{C}_t[(\omega, \omega')] \equiv \mathcal{C}_t[\omega] \cap \mathcal{C}_{t+1}[(\omega')]$ . However, as the family of sets

$$\{\mathcal{C}_t[(\omega, \omega')] \mid t \in \mathbb{Z}, \omega, \omega' \in \Omega\}$$

is not closed under intersections, this does not imply that the measures are the same, cf. Lemma A.1.

Further, it is possible to construct ergodic invariant measures with full support on  $\Gamma$  which are not (pull-backs) of Markov measures. The easiest example is provided by considering the pull-back of a stationary  $m$ -step Markov measure [BC75]. The latter is a Markov measure constructed for an alphabet  $\Omega^m$ , where the symbols are allowed blocks of  $m$  symbols that may appear in some 1-step SFT.

Although the author did not find a rigorous proof during the preparation of this thesis, he believes that under the present assumptions  $\mu_{\infty}$  agrees with  $\mu_{\Phi}$ . If this is in fact true, we would have an alternative approach to the natural measure. That said, we continue with a more physics-motivated discussion of our results.

### 3.4.2. Operational interpretation

In this subsection, we discuss our findings in the light of the experimental situation outlined at the beginning of Section 3.1. In particular we consider the situation where an experimenter records time series which appear to obey Markovian statistics. As we have seen in Section 2.4, such experimental situations are often modelled as Markovian stochastic processes. In the remainder of this section we discuss the implicit assumptions on the microscopic dynamics, the microscopic ensemble and the nature of the observable in the light of our mathematical treatment above.

Before we continue, let us review the assumptions of our idealized framework of the measurement process: The map  $\Phi$  is understood as a stroboscopic map obtained from the real dynamics which proceeds in continuous time. To ensure that an autonomous (*i.e.* time-independent) dynamics  $\Phi$  is a good approximation of the real situation, we require a huge temporal precision on the (stroboscopic) measurement intervals. Further, the measurement needs to be either perfectly reproducible (because then the disturbance of the system by the measurement apparatus is the same at each iterated measurement) or completely interaction free.

Obviously, neither of these conditions will ever be rigorously fulfilled. However, we are still able to have these requirements fulfilled in a *gedankenexperiment*, and look for non-trivial interpretations.

#### Observables, partitions and the measurement process

We start by discussing the nature of observables and the corresponding partitions of phase space. For the moment suppose that we know the microscopic dynamics  $\Phi$  and we also have a theoretical model (*i.e.* a well-defined observable  $M$ ) for a real measurement apparatus. Generically, such an observable will not induce a Borel-generating Markov partition on phase space. Just by considering the resolutions which are available in typical experiments, we would expect the partition to be far too coarse to generate the Borel-sets. However, no one can ever say if the statistics of experimentally observed time series *really* are Markovian. The only thing we might be able to say is that they *appear* memoryless within the finite time span of a given experiment.

**Generating partitions and measurability** Let us continue with the notion of “measurability”, both in the abstract mathematical and the operational sense. By the former we mean the concept of a  $\sigma$ -algebra on phase space, whereas by the latter we mean the experimental observation of time series, *i.e.* the recording of subsequent (elementary) measurement results  $\omega$ . Generating partitions (in the sense of the measure-theory) provide the connection between these two notions. The advantage of the measure-theoretic Definition 3.15 over the topological one is that it holds for arbitrary  $\sigma$ -algebras  $\mathcal{B}$  on phase space. In contrast to the topological definition, we do not require  $\mathcal{B}$  to be as refined as the Borel sets. All that is needed is that  $\mathcal{B}$  agrees with the  $\sigma$ -algebra  $\sigma(\zeta^a)$  generated by

### 3. Markovian symbolic dynamics

the atomic cells  $\zeta^a$ .

To appreciate the operational interpretation of  $\sigma(\zeta^a)$ , note that an atomic cell  $\mathcal{C}_t[(\omega)]$  contains exactly the points in phase space, which yield the measurement result  $\omega$  at time  $t$ . The generated  $\sigma$ -algebra  $\sigma(\zeta^a)$  thus contains countable unions and intersections of these “elementary events”. Measurable sets  $A \in \sigma(\zeta^a)$  thus describe events which can be formulated as (countable) Boolean statements about these elementary events. For instance,  $\sigma(\zeta^a)$  contains events like “We observe the same measurement result  $\omega$  for each time in the discrete interval  $(t_0, t_0 + 1, \dots, t_0 + \tau)$ ” or “Neither result  $\omega$  nor  $\omega'$  do occur at time  $t = t_0$ ”. Also statements about infinite time series like “The measurement result  $\omega$  never appears after some time  $t = t_0$ ” are possible.

If a partition is generating, then by definition  $\sigma(\zeta^a) = \mathcal{B}$ . This means that the “usual” descriptions of operationally accessible measurement events (like the examples given above) agree with the mathematically measurable sets on phase space. Or to put it differently: If a partition is generating, then all events that can be described by Boolean predicates are measurable in the mathematical sense. In fact, it is hard to describe an event that is not an element of the  $\sigma$ -algebra generated by the elementary measurement results.

If we understand measurability in that operational sense, it seems sensible to *choose*  $\mathcal{B} := \sigma(\zeta^a)$  as the measurable structure on phase space. Then, the mathematical and operational notions of measurability coincides and *every* partition induced by *any* observable is generating.

However, there is a caveat to this interpretation: Usually we have an initial time  $t = 0$  where we prepare the system and then record only for  $t \geq 0$ . Hence, to be precise we must consider the partition generated by the atomic cells in forward time only. It is worthwhile to mention the following consequence of a theorem by Kolmogorov–Sinai for invertible  $\Phi$ : The mere existence of a one-sided generator for a dynamical system  $(\Gamma, \mathcal{B}, \Phi)$  implies that its dynamical notion of entropy (*cf.* Sec. 2.6.3) vanish identically.

Finally, consider the case where one is fortunate enough to have a high-resolution measurement apparatus whose phase space partition generates the Borel  $\sigma$ -algebra. In that case, single points in phase space are represented by elements in the  $\sigma$ -algebra. Unfortunately, these events are statements about infinite time series and hence not operationally accessible: Their recording might surpass the time scale of a typical graduate student ...

**Markov partitions and local equilibrium** Next, let us discuss the Markovian postulate. We have already seen that this implies a topological constraint on the structure that the observable induces on phase space: The corresponding partition must be a Markov partition. We stress again, that we do not require that this partition generates the Borel sets, *i.e.* its elements do not need to be “small”.

Hence, on first sight, this requirement seems not too restrictive.<sup>6</sup> However, let us come

---

<sup>6</sup> For instance, it would allow the trivial partition obtained by a constant observable  $M(x) = 1, \forall x$ .

back to the example of the Bernoulli map, Fig. 3.2a). If we just slightly misplace the partition such that  $\mathcal{C}_0 = [\varepsilon, \frac{1}{2})$  the sequence  $\underline{\omega} = (0, 0, \dots)$  (which formerly belonged to the fixed point at  $x = 0$ ) would not be allowed. We cannot expect that we are fortunate enough to obtain such an alignment in reality.

However, we observe that Markovianity holds (at least for all practical purposes) for many single-molecule experiments [Sei11]. As we have discussed in Section 2.4, the loss of memory is a necessary consequence of the physical assumption of *local equilibrium*. More precisely, it relies on the notion of an (infinite) *separation of time scales*: We assume that within the stroboscopic measurement interval, the ensemble constrained to an observable (“local”) state (approximately) relaxes to a stationary<sup>7</sup> (“equilibrium”) distribution. If there exists a stationary distribution which attracts most initial conditions, it necessarily loses the information of the latter.

Obviously, a system with a smooth evolution in continuous time never relaxes to such a stationary distribution in *finite* times. However, it might already “mix” points in phase space to a sufficient degree, such that for all practical purposes a stationary distribution is a valid approximation. A form of this “smearing out” of points over the elements of the partition also manifests in the *n-fold intersection property* used by Adler for his (equivalent) definition of Markov partitions [Adl98].

Finally, it is interesting to ask *why* Markovian statistics are so commonly found in experiments. In the final Chapter in Section 7.3.1 we revisit the Markovian postulate as a kind of “anthropic principle” imposed by the *scientific method* [Pop02].

### The arrow of time and stationarity on two levels

In the present subsection, we are concerned with reversibility and the *arrow of time* as it appears on the two levels of description. We start with the invertible microscopic evolution on phase space, where we demanded *determinism*. On that level, we have a symmetry with respect to the direction of time. In principle, by applying a time-reversal operator (*cf.* Sec. 2.5.4), one can always find the original ensemble at time  $t_0$  from an evolved microscopic ensemble at time  $t_0 + \tau$ . In that sense, an *arbitrary* (normalized) initial ensemble at time  $t_0$  plays the same role as the push-forwarded ensemble at time  $t_0 + \tau$ .

In a popular article on the *thermodynamic* arrow of time, J. Lebowitz points out that the irreversibility on the observational level originates from particular choices of initial conditions [Leb93]. Here we extend this discussion to the Markovian postulate. First of all, note that a Markovian dynamics breaks time-symmetry in a peculiar way: Take an *arbitrary* stochastic vector  $\mathbf{p}_{t_0}$  as an initial condition. Because the transition matrix is stochastic, we know that any subsequent distribution  $\mathbf{p}_{t_0+\tau} = \mathbf{p}_{t_0} \mathbb{W}^\tau$  is equally stochastic. Further, as stochastic matrices are generically invertible, we can obtain  $\mathbf{p}_{t_0}$  from  $\mathbf{p}_{t_0+\tau}$ . However, if we use the inverse of  $\mathbb{W}$  to obtain a vector  $\mathbf{p}_t$  for any  $t < t_0$ , we cannot be sure

<sup>7</sup> There is also the corresponding notion of a “local steady state”, *cf.* [HS01], which might be more appropriate here.

### 3. Markovian symbolic dynamics

if this vector is still stochastic. Moreover, one can always find a time  $t_1 \leq t_0$  such that  $\mathbf{p}_t$  is not stochastic for  $t < t_1$  — unless  $\mathbf{p}_{t_0} = \mathbf{p}_\infty$  is already a left unity eigenvector to  $\mathbb{W}$ . Hence, unlike in the microscopic case,  $t_0$  distinguishes a particular point in time where we know that  $\mathbf{p}_t$  is stochastic for  $t \in [t, \infty)$ . Without loss of generality, in the remainder of this chapter we choose  $t_0 = 0$  and refer to it as the *present*.

The necessity of a “present”, *i.e.* a distinguished initial point in time, for generic Markovian dynamics manifests in the fact that there is no non-stationary measure that yields Markovian statistics for all  $t \in \mathbb{Z}$ . Operationally, the initial point  $t = 0$  is distinguished by the *preparation procedure*. Hence, we interpret the mathematical statement in an operational way: The manipulations on a system during the experimental preparation of a certain observable state are a non-Markovian process.

We also want to stress the distinction between microscopic and observational steady states. First of all, it is clear that the former implies the latter. However, we can only *observe* the statistics of time series. *Physically*, the microscopic states should not play a role, though they might be very useful as tools in the mathematical description [Rue99; Rue04]. Moreover, we have no chance to prepare them even with the best experimental equipment, because in general they concentrate finite probability on infinitesimal regions in phase space. Equally, as they are obtained only in the  $t \rightarrow \infty$  limit, we can never wait long enough for such a state to evolve naturally.

Finally, let us stress that it is not needed to have microscopic stationarity for observational stationarity: *Any* non-stationary extension of a one-sided stationary Markov measure to the full shift yields observational stationarity. Moreover, if the natural measure is an SRB measure (*i.e.* it has absolutely continuous density along the unstable directions), we can initialize *non-stationary* absolutely continuous measures yielding *stationary* observational statistics. In Section 4.3.3 we discuss this result in the light of the foundations of the stochastic thermodynamics of Markov chains, *cf.* the introductory remarks in Section 2.4.2.

## 3.5. Summary

In the present chapter, we have outlined a gedankenexperiment where an experimenter performs perfectly reproducible successive measurements on a physical system. In many experimental situations, such observations seem to obey Markovian Statistics. The main goal of this chapter was the rigorous discussion of the consistency of two common assumptions: Firstly, the assumption of microscopic causality or determinism implies that the underlying microscopic dynamics is prescribed by a deterministic (measurable) dynamical system [Pen70]. Secondly, the Markovian postulate for the observed process is central for the foundations of (classical) statistical mechanics [Pen70] as well as for modern stochastic thermodynamics [Sei12], *cf.* Sec. 2.4.

In summary, the constraints imposed by the Markovian postulate are two-fold: The partition of phase space induced by the observable must be a Markov partition in the



sense of Definition 3.22. However, it does not need be a particularly fine partition. More precisely, we do not require that the partition generates the Borel  $\sigma$ -algebra.

For observables which induce a Markov partition, we identified measures on phase space which yield Markovian statistics for the observed time series after a preparation at  $t = 0$ . We found that there is a large class of measures which fulfil this requirement. Measures within that class differ in terms of their past, *i.e.* the statistics of events happening at times  $t < 0$ . Moreover, we have discussed how “measurability” in the mathematical and the operational sense agree, if we accept  $\sigma$ -algebras on phase space that are coarser than the Borel sets. In that case, one can still find Borel measures on phase space that yield Markovian statistics. It is only important that their restriction on the  $\sigma$ -algebra generated by the atomic cells agrees with the values of the Markov measure. Hence, in addition to the ambiguity with respect to the past, we have an ambiguity with respect to the “fine structure” within measurable events. Consequently, this leads to a large class of ensembles that obey the Markovian postulate.

In spite of that, we must admit that observables corresponding to real measurements on physical systems do not induce Markov partitions in the mathematical sense. However, Markovian statistics seem to be readily observed — at least *for all practical purposes*. We discussed this fact in the context of the assumption of local equilibrium. Further we have hinted at a “Markovian anthropic principle”, which will be discussed in more detail in Section 7.3.1.

In the next chapter, we consider the same idealized framework of taking measurements as in the present chapter. However, we will not assume that the coarse-grained observations behave in a Markovian way. Rather, the subject of the next chapter is an information-theoretical analysis of our idealized framework. More precisely, we are going to compare the uncertainty we have about the system after iterated measurements with the information contained in its actual microscopic configuration. Moreover, we present an analytically tractable model dynamics with an observable that induces a generating Markov partition. Thus, the next chapter will also provide concrete examples for the results obtained in the present chapter.



## 4. An information-theoretical approach to stochastic thermodynamics

“ [...] it was necessary to think of probability theory as extended logic, because then probability distributions are justified in terms of their demonstrable information content, rather than their [...] frequency connections.”

---

E. T. Jaynes, *A Backward Look on The Future*, 1993

### What is this about?

In the previous Chapter 3, we started our discussion of the *microscopic* foundations of ST. More precisely, we were interested in the following question: When does a microscopic-deterministic dynamics yield Markovian statistics for the time series of coarse-grained observables?

In the present chapter, rather than investigating the microscopic foundation of the Markovian postulate, we discuss the connection of microscopic and mesoscopic notions of entropy. In Chapter 2 we have already reviewed the identification of entropy and entropy production in statistical mechanics. There, we advocated the distinction of system and medium by their *observability* in experiments, *cf.* Section 2.3. We motivated the notions of entropy used in stochastic thermodynamics in Section 2.4.2. In addition, Section 2.5 reviewed the identification of *dissipation* in molecular dynamics simulations with the phase space contraction prescribed by a microscopic model dynamics  $\Phi$ .

Here, we present an approach where these identifications emerge naturally. In our reasoning we apply some of Jaynes' ideas, as summarized in the initial quote [Jay93]. Instead of using phenomenological thermodynamic notions of entropy, we base our reasoning on information theory as a theory of statistical inference.

Often, the ensembles of statistical mechanics are interpreted in the so-called *frequentist* picture, where the phase space density is interpreted as characterizing the statistics of trials obtained by measurements. However, microscopic configurations (*i.e.* points in a system's phase space) are never operationally accessible. Hence, we follow Jaynes instead and interpret microscopic ensembles as our best estimate of the real microscopic situation — given any prior knowledge about the dynamics.

#### 4. An information-theoretical approach to stochastic thermodynamics

The present chapter is structured as follows: We begin by using the deterministic microscopic framework presented in the previous chapter in order to formalize the measurement process. In this endeavour, we rely on Jaynes' approach to statistical physics as a theory of inference. Consequently, we introduce two different phase-space ensembles, which reflect the information available on the microscopic and the mesoscopic levels of description, respectively: On the mesoscopic level, we obtain a *coarse-grained ensemble*  $\rho_t^{\text{cg}}$  based on the statistics of mesoscopic observations, that gives rise to a coarse-grained, *i.e.* inferred, entropy  $S_t^{\text{cg}}$ . In contrast, if we have knowledge about the microscopic dynamics  $\Phi$ , we can also calculate the “real”, *fine-grained* evolution  $\rho_t^{\text{fg}}$  of an initial ensemble and thus the *fine-grained* entropy  $S_t^{\text{fg}}$ . In addition, the comparison of these ensembles by the means of a Kullback–Leibler divergence yields a time-dependent relative entropy  $S_t^{\text{rel}}$ .

The entropies obtained in this way are averages over the entire phase space. Considering the phase space cylinders introduced in the previous chapter, we obtain more *detailed* notions of entropy and entropy variation. Like the phase space cylinders, these detailed entropies are associated with observable time-series  $\underline{\omega}^{(\tau)}$ . Eventually, they will take the role of the entropic  $\tau$ -chains, which have been introduced in the context of stochastic thermodynamics (ST) in Section 2.4.2.

In order to have a clear information-theoretic interpretation of our results, we motivate the definition of four *fundamental*  $\tau$ -chains. We see that the detailed versions of  $S_t^{\text{fg}}$ ,  $S_t^{\text{cg}}$  and  $S_t^{\text{rel}}$  can be obtained as linear combinations of these fundamental  $\tau$ -chains. In addition, we discuss how they yield *thermodynamic*  $\tau$ -chains for the variation of the entropy in a system and its medium. Under the assumption of a physical, *i.e.* reversible microscopic dynamics we finally arrive at a central result of this work: A conjecture regarding the generic microscopic foundations of ST, which involves the concept of the natural two-sided Markov measure introduced in the previous chapter.

Another central aspect of the present chapter is the introduction of network multibaker maps (NMBM). We use them as abstract model dynamics, which are both analytically tractable and versatile in mimicking more complicated situations. In particular, NMBM can be tuned to exhibit the hallmarks of the physical microscopic dynamics used in molecular dynamics simulations: They can be made time-reversible with a measure-preserving involution. Further, one can tune them to be uniformly conservative, conservative or dissipative, where the latter case is the generic one. Consequently, we use NMBM to exemplify the ideas presented in the current and the previous chapter.

### 4.1. A general information-theoretic framework

#### 4.1.1. The measurement process revisited

Let us return to the experimental situation described in the previous chapter: A scientist performs measurements on a system by recording the output  $\omega_t \in \Omega$  of a measurement apparatus at equidistant times  $t$  in a time-series  $\underline{\omega}^{(\tau)}$ . Upon repeating the experiment, she

samples the probabilities  $\mathbb{P}_t[\underline{\omega}^{(\tau)}]$  for the measured time-series between time  $t$  and  $t + \tau$ .

At the initial time  $t = t_0 = 0$  of preparation, the system finds itself in a certain microscopic state  $x \in \Gamma$ . However, the scientist never has access to this microscopic information. What she knows is the initial observable state  $\omega_0$ .

Hence, her initial measurements specify the initial *visible* or *observable* ensemble. For instance, she may ensure a specific initial condition  $\omega_{\text{init}}$  by dropping all trials where  $\omega_0 \neq \omega_{\text{init}}$ . Alternatively, the histogram of the initial measurements provides (a frequentist approach to) an initial distribution  $\mathbf{p}_0 = (p_0(\omega))_{\omega \in \Omega}$ .

In addition to the distribution  $\mathbf{p}_0$  for observable states, we require *microscopic ensembles* to specify a distribution on the phase space  $\Gamma$  of the underlying microscopic dynamics. In contrast to the observable ensemble, the microscopic ensemble cannot be sampled in a frequentist way. From Jaynes' point of view, the microscopic ensemble formally expresses our expectation about the microscopic distribution based on whatever information is available [Jay57]. Or differently put: It has to be *inferred* in a way consistent with our knowledge of physics and mathematics.

A distribution that maximizes the (differential) Shannon entropy with respect to a set of constraints (which formalize prior knowledge) is the *least biased* or *maximally non-committal* prior [Jay03]. In thermodynamics, the Gibbs distribution (2.15) provides an example: It is the least biased prior with respect to the available *macroscopic* thermodynamic information.

In that light, let us review the (mesoscopic) information available at the initial time  $t = 0$ . On the one hand, we have the coarse-grained information specified by the initial observable ensemble  $\mathbf{p}_0$ . This yields the constraint

$$\int \chi_\omega \rho_0 dx \stackrel{!}{=} p_0(\omega), \quad (4.1)$$

for the initial microscopic density  $\rho_0$ , where  $\chi_\omega$  is the indicator function for  $\mathcal{C}_\omega$ . Recall that  $\mathcal{C}_\omega$  is the set of phase space points yielding the measurement result  $\omega$ .

On the other hand, there might be additional information (or assumptions) regarding the observable *dynamics* or the thermodynamic interpretation of individual measurement results. In statistical mechanics, a common assumption is that of local equilibrium (LE), *cf.* Secs. 2.4.1 and 3.4.2. In its most common form, LE assumes that the (marginalized) distribution  $\rho_0(x|\omega) =: \rho_\omega: \mathcal{C}_\omega \rightarrow \mathbb{R}$  of microstates  $x \in \mathcal{C}_\omega$  is a Gibbsian, *i.e.* a maximum-entropy (MaxEnt) distribution — compatible with the physical interpretation of the measurement result  $\omega$ .

After a measurement at time  $t = 1$ , we obtain a new observable distribution  $\mathbf{p}_1$ . Using the principles described above, we get an updated, inferred distribution  $\rho_0^{\text{cg}}$ . However, this inferred distribution is usually not the same as distribution  $\rho_1^{\text{fg}}$ , which is obtained by the microscopic dynamics acting on  $\rho_0$ . Figure 4.1 illustrates this situation.

#### 4. An information-theoretical approach to stochastic thermodynamics

Consistency requires that the sampled observable probabilities  $\mathbf{p}_t$  obey

$$\int_{\mathcal{C}_\omega} \rho_t^{\text{cg}} dx \stackrel{!}{=} p_t(\omega) \stackrel{!}{=} \int_{\mathcal{C}_\omega} \rho_t^{\text{fg}} dx. \quad (4.2)$$

The first equality sign expresses consistency between the coarse-grained ensemble and measurements. The second equality expresses the fact that the microscopic theory is *physically valid*: If its predictions do not agree with our measurements, the theory should better be discarded. Consequently, Eq. (4.2) ensures that the microscopic model  $\Phi$  has been obtained via the *scientific method* [Pop02].

In the present thesis, we do not report on (numerical) experiments on particular systems. Consequently, we cannot sample the statistics of time-series in order to obtain the observable distribution  $\mathbf{p}_t$  at different moments in time. Instead, we consider some microscopic (not explicitly specified) model dynamics  $\Phi$ , which we assume to be a good physical theory. Hence, equation 4.2 defines the observable ensemble  $\mathbf{p}_t$  at time  $t$ .

In the previous Chapter we have looked for conditions on  $\Phi$  and  $\rho_0$  such that  $\mathbf{p}_t$  evolves according to Markovian statistics. Here, we initially drop this assumption and do not impose any further requirements on the microscopic dynamics  $\Phi$  or the measurement observable  $M$ . Instead, we discuss how entropies quantify our knowledge of the physics of a system, without invoking thermodynamic arguments. Only later we return to dynamics which yield Markovian observable statistics — and conjecture how the framework outlined so far may provide a microscopic foundation of Markovian ST.

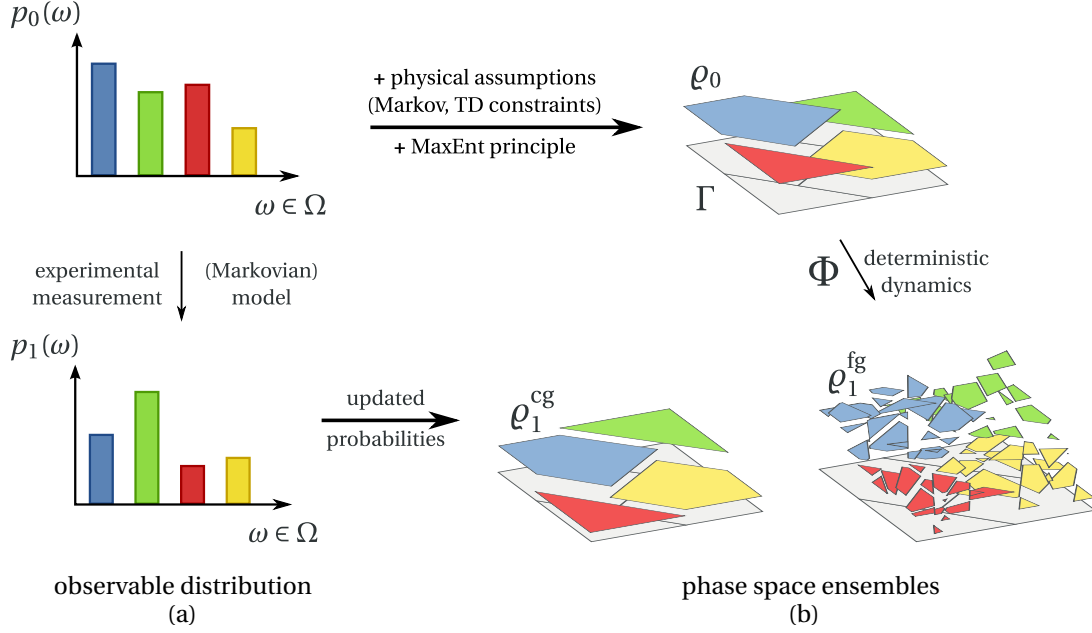
##### 4.1.2. Fine- and coarse-grained entropy

We start with the formal definitions of the fine- and coarse-grained ensembles motivated above. In Chapter 2 we defined the (differential) entropy  $H[\rho]$  with respect to a density  $\rho$  rather than using a (probability) measure  $\mu$ . Hence, we have implicitly chosen a *reference measure*  $\nu$  and the density  $\rho = \frac{d\mu}{d\nu}$  amounts to the Radon–Nikodym derivative of  $\mu$  with respect to  $\nu$ . In addition, we discussed in Section 3.3.3 how a reference measure  $\nu$  is needed for the definition of a natural measure.

Throughout the present chapter, we choose the Borel sets  $\mathcal{B}$  as a  $\sigma$ -algebra and use the translation invariant (Lebesgue) measure  $\lambda = \nu$  as a reference. Further, we assume that the phase-space dynamics  $\Phi$  is invertible and non-singular with respect to  $\lambda$ . This ensures that the iterated Jacobian determinant  $\hat{J}^{(\tau)}(x)$  of  $\Phi$  is well-defined for all  $\tau \in \mathbb{Z}$  and almost all  $x \in \Gamma$ .

The *fine-grained measure at time  $t$*  is nothing else than the image (pushforward) of an initial probability measure  $\mu_0$ :

$$\mu_t^{\text{fg}} := (\Phi^t)_* \mu_0 \equiv \mu_0 \circ \Phi^{-t}. \quad (4.3)$$



**Figure 4.1.:** Inferring microscopic distributions from mesoscopic and macroscopic constraints. (a) Top: The mesoscopic information comes in the form of an observable distribution  $p_t(\omega)$ . As observable are experimentally accessible, we can prepare them as a mesoscopic initial condition  $\mathbf{p}_0$ . Bottom: At a subsequent time step,  $\mathbf{p}_1$  is obtained from a new measurement or the prediction of a (Markovian) mesoscopic model. (b) Top: From a mesoscopic initial condition, the microscopic initial density  $\rho_0$  is inferred. The microscopic density is a maximum-entropy distribution that respects both i) the observable distribution and ii) any additional (dynamical, macroscopic or thermodynamic) constraints. Bottom left: At a subsequent time step, we infer the coarse-grained density  $\rho_t^{cg}$  again by a maximum-entropy principle. Bottom right: In contrast, the initial density is propagated by the microscopic deterministic dynamics  $\Phi$  to yield the fine-grained density  $\rho_1^{fg}$ , which shows a more complicated structure.

Its density is determined by Equation (2.76), i.e.

$$\rho_t^{fg}(x) := \frac{d\mu_t^{fg}}{dv} = \frac{\rho_0^{fg}(\Phi^{-t}(x))}{\hat{j}^{(t)}(\Phi^{-t}x)}. \quad (4.4)$$

In the following, we need to condition the fine-grained density on *phase space cylinders*  $\mathcal{C}_t[\underline{\omega}^{(\tau)}] \subset \Gamma$ . A phase space cylinder contains the initial conditions  $x$ , such that the (finite) time-series  $\underline{\omega}^{(\tau)}$  occurs at time  $t$  in the trajectory  $\underline{M}(x)$ . Recall that  $\mathcal{C}_t[\underline{\omega}^{(\tau)}]$  is the pre-image of the symbolic cylinder  $Z_t[\underline{\omega}^{(\tau)}]$  from Definition 3.5:

$$\mathcal{C}_t[\underline{\omega}^{(\tau)}] := \underline{M}^{-1}(Z_t[\underline{\omega}^{(\tau)}]). \quad (4.5)$$

For a probability measure  $\mu$  on  $(\Gamma, \mathcal{B})$ , we define the *conditioned measure*

$$\mu_t|_{\underline{\omega}^{(\tau)}}(B) := \frac{\mu(\Phi^{-t}B \cap \mathcal{C}_t[\underline{\omega}^{(\tau)}])}{\mu(\mathcal{C}_t[\underline{\omega}^{(\tau)}])} \equiv \frac{\mu_t(B \cap \mathcal{C}_0[\underline{\omega}^{(\tau)}])}{\mathbb{P}_t[\underline{\omega}^{(\tau)}]} \quad (4.6)$$

#### 4. An information-theoretical approach to stochastic thermodynamics

with density  $\varrho_t|_{\omega^{(t)}}(x) := \frac{d\mu_t|_{\omega^{(t)}}}{dv}$ .

For real systems, we know neither  $\Phi$  nor  $\varrho_0$ , so the fine-grained distribution  $\mu_t^{\text{fg}}$  cannot be inferred. All that we know are individual values of the measurement observable  $M$ . From performing measurements on a large number of system, we can estimate the distribution  $\mathbf{p}_t = (p_t(\omega))_{\omega \in \Omega}$ .

In Section 3.2 we discussed the partition  $\mathcal{Q} = (\mathcal{C}_\omega)_{\omega \in \Omega} = M^{-1}(\Omega)$  induced by the observable  $M$ . Denote by  $\mu_\omega^{\text{ic}} := \mu_0|_{\mathcal{C}_\omega}$  the (unknown) *initial measure conditioned on cell*  $\mathcal{C}_\omega$  and by  $\varrho_\omega^{\text{ic}}$  its density. The usual assumption for  $\varrho_\omega^{\text{ic}}$  ensures that it is a Gibbs distribution  $\varrho^{\text{G}}$ , *i.e.* a maximum-entropy distribution, *cf.* Eq. (2.15). From the point of information theory, constrained maximum entropy distributions are the least biased estimates which are compatible with the given prior information expressed by the constraints [Jay57].

In the thermodynamic context, this prior information is physical and usually amounts to a certain knowledge about the macroscopic state of the system, *e.g.* its volume or its temperature. For more abstract considerations without the reference to physics, such constraints arise from (known) model symmetries. In the mathematical framework introduced in the previous chapter, a measurement result  $\omega$  provides us with incomplete information about the system's position in phase space: We know that its microstate  $x$  must belong to the partition element  $\mathcal{C}_\omega$ .

In addition to any other prior knowledge we have, this induces an additional constraint on the microscopic measure: The measure capturing this information must be supported on  $\mathcal{C}_\omega$  only. We call the maximum-entropy measure  $\mu_\omega^{\text{pr}}$  supported on cell  $\mathcal{C}_\omega$  the *prior measure* for state  $\omega$ . The entropy associated to its density  $\varrho_\omega^{\text{pr}}$  is called the *assumed internal entropy* of state  $\omega$ . Note that the assumed entropy of a state is independent of time. It is therefore independent of the “real” internal entropy  $H[\varrho_t^{\text{fg}}|_{\mathcal{C}_\omega}]$ , which we obtain by constraining  $\varrho_t^{\text{fg}}$  to  $\mathcal{C}_\omega$ .

Combining the information of the measurement at time  $t$  with the prior leads to the *coarse-grained measure*  $\mu_t^{\text{cg}}$ , *cf.* Fig. 4.1. It is fully defined by its density

$$\varrho^{\text{cg}}(x) := \sum_{\omega \in \Omega} \chi_\omega(x) p_t(\omega) \varrho_\omega^{\text{pr}}(x). \quad (4.7)$$

The probability of an observable state is  $p_t(\omega) = \mu^{\text{fg}}(\mathcal{C}_\omega)$ . Because the prior density  $\varrho_\omega^{\text{pr}}$  is normalized on  $\mathcal{C}_\omega$ ,  $\varrho^{\text{cg}}$  automatically fulfils Eq. (4.2).

The fine- and coarse-grained measures give rise to the *fine- and coarse-grained entropies*

$$S_t^{\text{fg}} := H[\varrho_t^{\text{fg}}] \text{ and } S_t^{\text{cg}} := H[\varrho_t^{\text{cg}}], \quad (4.8)$$

where  $H[\varrho]$  denoting the differential Shannon entropy (2.3).

Given the family of maximum-entropy priors  $(\varrho_\omega^{\text{pr}})_{\omega \in \Omega}$  for each of the cells, the microscopic density fully specifies the coarse-grained density. The additional information is expressed by the Kullback–Leibler divergence of  $\varrho^{\text{cg}}$  from  $\varrho^{\text{fg}}$ . In the following, we refer to



it as the *relative entropy*

$$S_t^{\text{rel}} := D_{\text{KL}} \left[ \rho_t^{\text{fg}} \parallel \rho_t^{\text{cg}} \right] \geq 0. \quad (4.9)$$

### 4.1.3. The fundamental and derived entropic $\tau$ -chains

In Section 2.4.2, we have introduced observables  $\varphi[\underline{\omega}^{(\tau)}]$  that depend on (parts of) the random trajectory  $\underline{\omega}^{(\tau)}$  generated by a noise history in a stochastic process. Moreover, we focused on the interpretation of the quantities  $s_t^{\text{sys}}[\underline{\omega}^{(\tau)}]$  and  $s_t^{\text{med}}[\underline{\omega}^{(\tau)}]$  as the entropy variations in the system and its medium, respectively. The expressions yielding their definition were motivated by the fact that their (time-series) averages amount to the variation of the entropies  $\Delta_{t+\tau} S^{\text{sys}}$  and  $\Delta_{t+\tau} S^{\text{med}}$ , respectively.

In the present section, we aim to achieve the same for the (variations of the) entropies  $S_t^{\text{fg}}$ ,  $S_t^{\text{cg}}$  and  $S_t^{\text{rel}}$ . More precisely, we define time-series dependent observables  $s_t^{\text{fg}}$ ,  $s_t^{\text{cg}}$  and  $s_t^{\text{rel}}$  yielding  $S_t^{\text{fg}}$ ,  $S_t^{\text{cg}}$  and  $S_t^{\text{rel}}$  as their averages.

In order to achieve this goal, it is convenient to introduce some notation, which is properly formalized in Appendix A.2. A  $\tau$ -chain  $\varphi^{(\tau)}$  is a function

$$\begin{aligned} \varphi^{(\tau)} : \Omega^{\tau+1} \times \mathbb{T} &\rightarrow \mathbb{R}, \\ (\underline{\omega}^{(\tau)}, t) &\mapsto \varphi_t[\underline{\omega}^{(\tau)}]. \end{aligned}$$

Note that for a more concise notation, we drop the temporal index on  $\varphi$  when talking about its value  $\varphi_t[\underline{\omega}^{(\tau)}]$ , as the run length  $\tau$  is specified explicitly by the symbol  $\underline{\omega}^{(\tau)}$ .

The run-length index ( $\tau$ ) is important for the definition of the *canonical sequence*  $(\varphi_t^{(\tau)})_{t \in \mathbb{T}}$  of  $\tau$ -chains which are obtained from a (time-dependent) state observable  $\varphi_t(\omega) : \Omega \times \mathbb{T} \rightarrow \mathbb{R}$ . The elements  $\varphi_t^{(\tau)}$  of a canonical sequence evaluate  $\varphi_t$  at the state of the system at time  $t + \tau$ :

$$\varphi_t^{(\tau)}[\underline{\omega}^{(\tau)}] := \varphi_{t+\tau}(\omega_\tau). \quad (4.10)$$

In the following, we will encounter certain  $\tau$ -chains which are obtained as canonical sequences of state observables (*i.e.* zero-chains). Consequently, in our wording we synonymously refer to  $\varphi_t(\omega)$  and its canonical  $\tau$ -chain.

Before we discuss the  $\tau$ -chains associated to the fine-grained, coarse-grained and relative entropy, we introduce four *fundamental* entropic  $\tau$ -chains. We call them *fundamental*, because they have a clear information-theoretic interpretation in our framework for taking measurements on an underlying microscopic-deterministic system. From the fundamental  $\tau$ -chains, we construct the *derived*  $\tau$ -chains  $s_t^{\text{fg}}$ ,  $s_t^{\text{cg}}$  and  $s_t^{\text{rel}}$ .

The first fundamental  $\tau$ -chain is the *visible self-information*:

$$s_t^{\text{vis}}(\omega) := -\log p_t(\omega). \quad (4.11)$$

#### 4. An information-theoretical approach to stochastic thermodynamics

It quantifies the *uncertainty* or *surprisal* of finding the state  $\omega$  at time  $t$ , if we know the distribution  $\mathbf{p}_t$ . We call it visible, because  $\mathbf{p}_t$  is experimentally accessible through measurements.<sup>1</sup>

Further, we have time-independent *assumed internal entropy* of a state  $\omega$  associated to the prior distribution  $\varrho_\omega^{\text{pr}}$ :

$$s^{\text{pr}}(\omega) := H[\varrho_\omega^{\text{pr}}] = - \int_{\mathcal{C}_\omega} \varrho_\omega^{\text{pr}} \log \varrho_\omega^{\text{pr}} dx. \quad (4.12)$$

Another fundamental quantity is the so-called *cross entropy* between the prior  $\varrho_\omega^{\text{pr}}$  and the real fine-grained entropy on cell  $\mathcal{C}_\omega$ :

$$s_t^\times(\omega) := - \int_{\mathcal{C}_\omega} \varrho_t^{\text{fg}}|_{\mathcal{C}_\omega} \log(\varrho_\omega^{\text{pr}}) dx. \quad (4.13)$$

It quantifies the mismatch of the (constrained) real microscopic cell and the assumed prior on that cell.

The three quantities introduced so far are state variables. As just mentioned, we identify them with their canonical (sequence of)  $\tau$ -chains  $(s_t^{\text{vis}})^{(\tau)}$ ,  $(s^{\text{pr}})^{(\tau)}$  and  $(s_t^\times)^{(\tau)}$ .

The fourth fundamental observable is already defined as a  $\tau$ -chain. It quantifies the expansion of a phase space cylinder  $\mathcal{C}_0[\underline{\omega}^{(\tau)}]$ , *i.e.* the (average) phase space expansion experienced by the points that give rise to a time series  $\underline{\omega}^{(\tau)}$  in the interval  $[t, t + \tau]$ :

$$\lambda_t^{(\tau)}[\underline{\omega}^{(\tau)}] := \int_{\mathcal{C}_0[\underline{\omega}^{(\tau)}]} \varrho_t^{\text{fg}}|_{\underline{\omega}^{(\tau)}} \log(\hat{J}^{(\tau)}) dx. \quad (4.14)$$

#### Averages

The probability of observing a time-series  $\underline{\omega}^{(\tau)}$  in the interval  $[t, t + \tau]$  is defined as the fine-grained measure  $\mu_t^{\text{fg}}$  of the associated phase space cylinder, *cf.* Definition 3.5:

$$\mathbb{P}_t[\underline{\omega}^{(\tau)}] := \mu_t^{\text{fg}}(\mathcal{C}_0[\underline{\omega}^{(\tau)}]).$$

The average of a  $\tau$ -chain is defined as (*cf.* (A.1))

$$\langle\langle \varphi^{(\tau)} \rangle\rangle_t^{(\tau)} := \sum_{\underline{\omega}^{(\tau)} \in \Omega^{\tau+1}} \mathbb{P}_t[\underline{\omega}^{(\tau)}] \varphi_t[\underline{\omega}^{(\tau)}].$$

Lemma A.12 ensures that the members of a canonical sequence obey

$$\langle\langle \varphi^{(\tau)} \rangle\rangle_t^{(\tau)} = \langle \varphi \rangle_{t+\tau}. \quad (4.15)$$

<sup>1</sup>At the moment we refrain from calling it the system's entropy, although this would be the correct interpretation.

Hence, the assumed entropy and the cross entropy have time-series averages that agree with the state averages at time  $t + \tau$

$$\langle\langle s^{\text{pr}} \rangle\rangle_t^{(\tau)} = \langle s^{\text{pr}} \rangle_{t+\tau} \equiv \sum_{\omega} [p_{t+\tau}(\omega) s^{\text{pr}}(\omega)] \quad (4.16a)$$

$$\langle\langle s^{\times} \rangle\rangle_t^{(\tau)} = \langle s^{\times} \rangle_{t+\tau} \equiv \sum_{\omega} [p_{t+\tau}(\omega) s_{t+\tau}^{\times}(\omega)] \quad (4.16b)$$

Similarly, the time-series average of the visible self-information is the Shannon entropy of the visible ensemble at time  $t + \tau$ :

$$\langle\langle s^{\text{vis}} \rangle\rangle_t^{(\tau)} = \sum_{\omega} [p_{t+\tau}(\omega) \log p_{t+\tau}(\omega)] \equiv H[\mathbf{p}_{t+\tau}]. \quad (4.16c)$$

Finally, the time-series average of  $\lambda_t^{(\tau)}$  yields the average total phase space expansion between  $t$  and  $t + \tau$ :

$$\langle\langle \lambda \rangle\rangle_t^{(\tau)} = \int_{\Gamma} \rho_t^{\text{fg}} \log(\hat{J}^{(\tau)}) dx = \Lambda_t^{(\tau)} \quad (4.16d)$$

From the four fundamental sequences of  $\tau$ -chains, we define the  $\tau$ -chains for the coarse-grained, fine-grained and relative entropy

$$(s_t^{\text{cg}})^{(\tau)} := (s_t^{\text{vis}})^{(\tau)} + (s_t^{\text{pr}})^{(\tau)}, \quad (4.17a)$$

$$(s_t^{\text{fg}})^{(\tau)} := \lambda_t^{(\tau)} + S_t^{\text{fg}} \quad \text{and} \quad (4.17b)$$

$$(s_t^{\text{rel}})^{(\tau)} := (s_t^{\text{vis}})^{(\tau)} - (s_t^{\text{fg}})^{(\tau)} + (s_t^{\times})^{(\tau)}, \quad (4.17c)$$

respectively. In order to justify their names, we look at their respective time-series averages, which obey

$$\langle\langle s^{\text{cg}} \rangle\rangle_t^{(\tau)} = S_{t+\tau}^{\text{cg}}, \quad (4.18a)$$

$$\langle\langle s^{\text{fg}} \rangle\rangle_t^{(\tau)} = S_{t+\tau}^{\text{fg}}, \quad (4.18b)$$

$$\langle\langle s^{\text{rel}} \rangle\rangle_t^{(\tau)} = S_{t+\tau}^{\text{rel}}. \quad (4.18c)$$

Equations (4.18) are obtained as the result of the following calculations. We start with Eq. (4.17a):

$$\begin{aligned} S_t^{\text{cg}} &= - \int_{\Gamma} \left( \sum_{\omega} \llbracket x \in \mathcal{C}_{\omega} \rrbracket p_{\omega}^{(t)} \rho_{\omega}^{\text{pr}} \right) \log \left( \sum_{\omega} \llbracket x \in \mathcal{C}_{\omega} \rrbracket p_{\omega}^{(t)} \rho_{\omega}^{\text{pr}} \right) dx \\ &= \sum_{\omega} \int_{\mathcal{C}_{\omega}} p_{\omega}^{(t)} \rho_{\omega}^{\text{pr}}(x) (-\log p_{\omega}^{(t)} - \log \rho_{\omega}^{\text{pr}}(x)) dx \\ &= \sum_{\omega} [p_{\omega}^{(t)} (-\log p_{\omega}^{(t)})] + \sum_{\omega} \left[ p_{\omega}^{(t)} \left( - \int_{\mathcal{C}_{\omega}} \rho_{\omega}^{\text{pr}}(x) \log \rho_{\omega}^{\text{pr}}(x) dx \right) \right] \\ &= \langle s_t^{\text{vis}}(\omega) + s^{\text{pr}}(\omega) \rangle_t \end{aligned}$$

#### 4. An information-theoretical approach to stochastic thermodynamics

Finally, equation (4.15) yields the claim. As mentioned in Section 2.5.3, an easy calculation shows

$$\begin{aligned}
S_{t+\tau}^{\text{fg}} &= - \int_{\Gamma} \varrho_{t+\tau}^{\text{fg}}(x) \log \varrho_{t+\tau}^{\text{fg}}(x) dx = - \int \frac{\varrho_t^{\text{fg}}(\Phi^{-\tau}(x))}{\hat{j}^{(\tau)}(\Phi^{-\tau}(x))} \log \frac{\varrho_t^{\text{fg}}(\Phi^{-\tau}(x))}{\hat{j}^{(\tau)}(\Phi^{-\tau}(x))} dx \\
&= - \int_{\Gamma} \varrho_t^{\text{fg}} \log \frac{\varrho_t^{\text{fg}}(x)}{\hat{j}^{(\tau)}(x)} dx = - \int_{\Gamma} \varrho_t^{\text{fg}} \log \varrho_t^{\text{fg}}(x) dx + \int_{\Gamma} \varrho_t^{\text{fg}} \log \hat{j}^{(\tau)}(x) dx \\
&= S_t^{\text{fg}} + \Lambda_t^{(\tau)} = \left\langle S_t^{\text{fg}} + \lambda \right\rangle_t^{(\tau)}. \tag{4.19}
\end{aligned}$$

In the first line we used Equation (2.76). To obtain the second line we used the transformation theorem for integrals. Finally, we used Eq. (4.16d) in the last line and thus prove Eq. (4.18b). For Eq. (4.18c) we use the averages of the fundamental sequences of observables (4.16):

$$\begin{aligned}
\left\langle s^{\text{rel}} \right\rangle_t^{(\tau)} &= \left\langle s^{\text{vis}} \right\rangle_t^{(\tau)} - \left\langle s^{\text{fg}} \right\rangle_t^{(\tau)} + \left\langle s^{\times} \right\rangle_t^{(\tau)} \\
&= - \sum_{\omega} [p_{t+\tau}(\omega) \log p_{t+\tau}(\omega)] - S_{t+\tau}^{\text{fg}} + S_{t+\tau}^{\times}.
\end{aligned}$$

Now, we split the phase space integrals  $\int_{\Gamma}$  into the integrals over the partition elements  $\sum_{\omega} \int_{\mathcal{E}_{\omega}}$  to obtain

$$\begin{aligned}
\left\langle s^{\text{rel}} \right\rangle_t^{(\tau)} &= \sum_{\omega} \left[ \int_{\mathcal{E}_{\omega}} \varrho_{t+\tau}^{\text{fg}} \left( -\log p_{t+\tau}(\omega) + \log \varrho_{t+\tau}^{\text{fg}} - \log \varrho_{\omega}^{\text{pr}} \right) dv \right] \\
&= \sum_{\omega} \left[ \int_{\mathcal{E}_{\omega}} \varrho_{t+\tau}^{\text{fg}} \log \left( \frac{\varrho_{t+\tau}^{\text{fg}}}{p_{t+\tau}(\omega) \varrho_{\omega}^{\text{pr}}} \right) dv \right] \\
&= - \int_{\Gamma} \varrho_{t+\tau}^{\text{fg}} \log \left( \frac{\sum_{\omega} \chi_{\omega} p_{t+\tau}(\omega) \varrho_{\omega}^{\text{pr}}}{\varrho_{t+\tau}^{\text{fg}}} \right) dv \\
&= D_{\text{KL}} \left[ \varrho_{t+\tau}^{\text{fg}} \parallel \varrho_{t+\tau}^{\text{cg}} \right] \\
&= S_{t+\tau}^{\text{rel}}
\end{aligned}$$

which proofs Eq. (4.18c).

Let us summarize this subsection. We established  $\tau$ -chains for a set of fundamental information-theoretic quantities which appear naturally within our framework of the measurement process. From those (time-series dependent) fundamental expressions we constructed the three derived  $\tau$ -chains  $s^{\text{cg}}$ ,  $s^{\text{fg}}$  and  $s^{\text{rel}}$ . We showed that an average over time-series running from time  $t$  to  $t + \tau$  amounts to the value of the coarse-grained, fine-grained and relative entropy at the final time, respectively. In the next section, we calculate their temporal variations.

#### 4.1.4. Temporal variations

In Section 2.4.2 we have introduced time-series dependent expressions  $\delta s^{\text{sys}}$  and  $\delta s^{\text{med}}$  in the context of stochastic thermodynamics. Their respective averages yield the temporal

variations  $\Delta S^{\text{sys}}$  and  $\Delta S^{\text{med}}$ . Now we do the same for the changes of the fine-grained, coarse-grained and relative entropy.

Let  $(\varphi^{(\tau)})_{\tau \in \mathbb{N}}$  be a sequence of  $\tau$ -chains, which does *not* need to be canonical, cf. (4.10). For a given finite time-series  $\underline{\omega}^{(\tau)} = (\omega_0, \omega_1, \dots, \omega_\tau)$ , the elements of the corresponding *sequence of the variations*  $(\delta^{(\tau)}\varphi)_{\tau \in \mathbb{N}^+}$  are defined as

$$\delta^{(\tau)}\varphi_t[\underline{\omega}^{(\tau)}] := \varphi_t^{(\tau)}[(\omega_0, \omega_1, \dots, \omega_{\tau-1}, \omega_\tau)] - \varphi_t^{(\tau-1)}[(\omega_0, \omega_1, \dots, \omega_{\tau-1})]. \quad (4.20)$$

The fundamental variation  $\tau$ -chains associated to the fundamental chains  $s^{\text{vis}}$ ,  $s^{\text{pr}}$ ,  $s^\times$  and  $\lambda$  read:

$$\delta^{(\tau)}s_t^{\text{vis}}[\underline{\omega}^{(\tau)}] = \log \frac{p_{t+\tau-1}(\omega_{\tau-1})}{p_{t+\tau}(\omega_\tau)}, \quad (4.21a)$$

$$\delta^{(\tau)}s_t^{\text{pr}}[\underline{\omega}^{(\tau)}] = s^{\text{pr}}(\omega_\tau) - s^{\text{pr}}(\omega_{\tau-1}), \quad (4.21b)$$

$$\delta^{(\tau)}s_t^\times[\underline{\omega}^{(\tau)}] = s_{t+\tau}^\times(\omega_\tau) - s_{t+\tau-1}^\times(\omega_{\tau-1}), \quad (4.21c)$$

$$\delta^{(\tau)}\lambda_t[\underline{\omega}^{(\tau)}] = \int_{\mathcal{C}[\underline{\omega}^{(\tau)}]} \varrho_t^{\text{fg}} \Big|_{\underline{\omega}^{(\tau)}} \log(\hat{J}^{(1)} \circ \Phi^{(t-1)}) \, dx, \quad (4.21d)$$

Thus, the variations of the derived quantities (4.17) amount to:

$$\delta^{(\tau)}s_t^{\text{cg}} \equiv \delta^{(\tau)}s_t^{\text{vis}} + \delta^{(\tau)}s_t^{\text{pr}} = \log \frac{p_{t+\tau-1}(\omega_{\tau-1})}{p_{t+\tau}(\omega_\tau)} + s^{\text{pr}}(\omega_\tau) - s^{\text{pr}}(\omega_{\tau-1}), \quad (4.22a)$$

$$\delta^{(\tau)}s_t^{\text{fg}} \equiv \delta^{(\tau)}\lambda_t = \int_{\mathcal{C}[\underline{\omega}^{(\tau)}]} \varrho_t^{\text{fg}} \Big|_{\underline{\omega}^{(\tau)}} \log(\hat{J}^{(1)} \circ \Phi^{(t-1)}) \, dx \quad \text{and} \quad (4.22b)$$

$$\delta^{(\tau)}s_t^{\text{rel}} \equiv \delta^{(\tau)}s_t^{\text{vis}} - \delta^{(\tau)}s_t^{\text{fg}} + \delta^{(\tau)}s_t^\times, \quad (4.22c)$$

Lemma A.15 in Appendix A.2 ensures that the time-series average  $\langle\langle \delta s \rangle\rangle_t^{(\tau)}$  of  $\delta^{(\tau)}s$  equals the temporal variations  $\Delta_{t+\tau} S = S_{t+\tau} - S_{t+\tau-1}$  of the time-dependent average  $S_t := \langle\langle s \rangle\rangle_t^{(\tau)}$ . Hence, the time-series averages of the  $\tau$ -chains (4.22) yield the variations of the corresponding entropies:

$$S_{t+\tau}^{\text{cg}} - S_{t+\tau-1}^{\text{cg}} = \Delta S_{t+\tau}^{\text{cg}} = \langle\langle \delta s^{\text{cg}} \rangle\rangle_t^{(\tau)}, \quad (4.23a)$$

$$S_{t+\tau}^{\text{fg}} - S_{t+\tau-1}^{\text{fg}} = \Delta S_{t+\tau}^{\text{fg}} = \langle\langle \delta s^{\text{fg}} \rangle\rangle_t^{(\tau)}, \quad (4.23b)$$

$$S_{t+\tau}^{\text{rel}} - S_{t+\tau-1}^{\text{rel}} = \Delta S_{t+\tau}^{\text{rel}} = \langle\langle \delta s^{\text{rel}} \rangle\rangle_t^{(\tau)}. \quad (4.23c)$$

Finally, let us discuss these averages in more detail. The variation of the coarse-grained entropy

$$\begin{aligned} \Delta S_{t+\tau}^{\text{cg}} &= \Delta S_{t+\tau}^{\text{vis}} + \Delta S_{t+\tau}^{\text{pr}} \\ &= \sum_{\omega, \omega'} \left[ \mathbb{P}_{t+\tau-1}[(\omega, \omega')] \left( \log \frac{p_{t+\tau-1}(\omega)}{p_{t+\tau}(\omega')} + s^{\text{pr}}(\omega') - s^{\text{pr}}(\omega) \right) \right] \end{aligned} \quad (4.24)$$

consists of the change of the visible and the assumed internal entropy. The former is ob-

#### 4. An information-theoretical approach to stochastic thermodynamics

tained as the average over the well known logarithmic ratio of the ensemble probabilities before and after the transition. The latter is just the difference of the assumed internal entropies of the respective cells.

We have already encountered the variation of the fine-grained entropy at several points in this thesis:

$$\Delta S_{t+\tau}^{\text{fg}} = \langle\langle \delta \lambda \rangle\rangle_t^{(\tau)} = \Lambda_{t+\tau-1}^{(1)} = \int_{\Gamma} \rho_{t+\tau-1}^{\text{fg}} \log(\hat{J}^{(1)}) dx. \quad (4.25)$$

It has the same value as the averaged phase-space expansion rate and can be calculated from the logarithm of the Jacobian determinant.

Finally, consider the variation of the relative entropy:

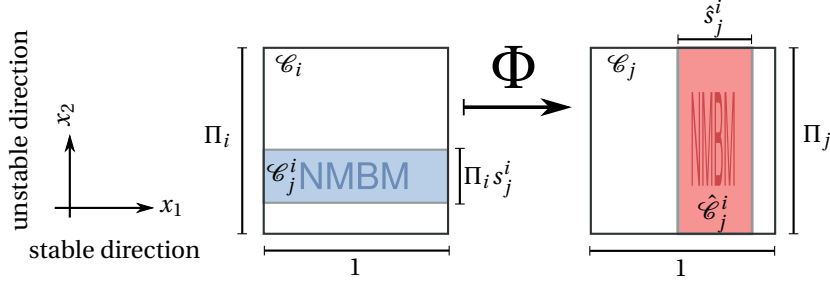
$$\begin{aligned} \Delta S_t^{\text{rel}} &= \int_{\Gamma} \rho_t^{\text{fg}} \log\left(\frac{\rho_t^{\text{fg}}}{\rho_t^{\text{cg}}}\right) - \rho_{t-1}^{\text{fg}} \log\left(\frac{\rho_{t-1}^{\text{fg}}}{\rho_{t-1}^{\text{cg}}}\right) dx \\ &= \int_{\Gamma} \rho_{t-1}^{\text{fg}} \log\left(\frac{\rho_t^{\text{fg}} \circ \Phi}{\rho_t^{\text{cg}} \circ \Phi}\right) - \rho_{t-1}^{\text{fg}} \log\left(\frac{\rho_{t-1}^{\text{fg}}}{\rho_{t-1}^{\text{cg}}}\right) dx \\ &= \int_{\Gamma} \rho_{t-1}^{\text{fg}} \log\left(\frac{\rho_t^{\text{fg}} \circ \Phi}{\rho_{t-1}^{\text{fg}}} \frac{\rho_{t-1}^{\text{cg}}}{\rho_t^{\text{cg}} \circ \Phi}\right) dx \\ &= \int_{\Gamma} \rho_{t-1}^{\text{fg}} \log\left(\frac{1}{\hat{J}^{(1)}(x)} \frac{\rho_{t-1}^{\text{cg}}(x)}{\rho_t^{\text{cg}}(\Phi(x))}\right) dx. \end{aligned} \quad (4.26)$$

Note that though  $S_t^{\text{rel}}$  is a Kullback–Leibler divergence and hence always positive, this is not generally true for the variation  $\Delta S_t^{\text{rel}}$ . This is easily seen from the following example: Consider that at some point in time  $t > 0$  we have  $\rho_t^{\text{fg}} = \rho_t^{\text{cg}}$ . In that case, in general  $\rho_{t-1}^{\text{fg}} \neq \rho_{t-1}^{\text{cg}}$  and hence  $0 = S_t^{\text{rel}} < S_{t-1}^{\text{rel}}$ . Hence, the positivity of  $\Delta S_t^{\text{rel}}$  crucially depends on the microscopic initial condition  $\rho_0$ . We will discuss the issue of positivity in detail in Section 4.3. Before we do so, however, let us illustrate the results of the present section using a concrete example.

## 4.2. Network multibaker maps

In this section, we exemplify the information-theoretical framework presented in Section 4.1 on a concrete model. As our chaotic microscopic dynamics, we use a variant of so-called multibaker maps, which were originally introduced by Hopf [Hop48]. Gaspard stressed their role as a generic example of a strongly mixing hyperbolic map [Gas05]. In analogy to models of transport theory, Vollmer and co-workers introduced reversible multibaker maps [VTB98; Vol02]. Being reversible, they share many of the features of the (NE)MD models discussed in Section 2.5.

A multibaker map consists of a (countable) number of rectangular *cells*. The dynamics maps rectangular subsets of each cell to rectangular subsets in *adjacent* cells. Historically, multibaker maps are arranged on a regular one-dimensional lattice with either open or



**Figure 4.2.:** Dynamics of the network multibaker map. For each of the  $|V_i|$  cells adjacent to the cell indexed by  $i$ , a horizontal strip (blue)  $\mathcal{C}_j^i \subset \mathcal{C}_i$  of relative height  $s_j^i$  is affine-linearly mapped to a vertical strip (red)  $\hat{\mathcal{C}}_j^i \subset \mathcal{C}_j$  of width  $\hat{s}_j^i$ . The horizontal coordinate  $x_1$  specifies the position along the unstable direction, i.e. the direction where  $\Phi$  is expanding. Similarly, phase space is contracted along the stable, vertical direction  $x_2$ . The deformation of a strip is made visible by the distortion of the letters “NMBM”.

periodic boundary conditions. Several variants of this linearly arranged multibaker maps exist in the literature [Gas05; Vol02; Col+11].

Here, we generalize this setting to more complex phase-space topologies. More precisely, we extend the neighbouring relations between the cells of a multibaker maps to arbitrary networks of states. In spite of being more general, these *network multibaker maps* (NMBM) still admit an explicit analytical treatment of the evolution of their phase-space densities. For multibaker maps on linear chains, such calculations have been the subject of earlier work [VTB97; Vol02].

Generalizing the approach followed in Ref. [RTV00], we focus on the statistical behaviour of certain sets of microscopic orbits rather than on global averages. More precisely, we consider the evolution of the microscopic trajectories that belong to a phase space cylinder  $\mathcal{C}_t[\underline{\omega}^{(t)}]$ . We will see how the expressions used in Markovian ST emerge naturally as a result.

#### 4.2.1. Formulation of the model

The definition of the model starts with an arbitrary network specified by a directed graph  $G = (V, E)$  on  $N$  vertices  $i \in V := \{1, 2, \dots, N\}$ , and edges  $e \in E \subset V \times V$ . To keep notation at bay, we assume that the graph is simple, i.e. that there is at most one edge connecting vertex  $i$  to vertex  $j$ . However, note that the prescription of a NMBM trivially extends to graphs with more than one edge between two given vertices.

Now denote by  $V_i := \{j \in V : (i, j) \in E\}$  the set of vertices that are connected to a state  $i$  and by  $|V_i|$  the *degree* of vertex  $i$ . The microscopic phase space consists of rectangular cells  $\mathcal{C}_i := [0, 1] \times \Pi_i \cdot [0, 1]$  with area (i.e. Lebesgue measure)  $\Pi_i$  for each vertex  $i$ . The overall *phase space*  $\Gamma$  of the system is the disjoint union  $\Gamma := \bigsqcup_{i=1}^N \mathcal{C}_i$ . Hence, a point  $x \in \Gamma$  is a triple  $x = (x_1, x_2, i)$  where  $(x_1, x_2) \in \mathcal{C}_i$ .

A NMBM,  $\Phi: \Gamma \rightarrow \Gamma$ , deterministically maps phase-space points  $x \in \mathcal{C}_i$  to adjacent cells  $\mathcal{C}_j$ . It is specified geometrically (cf. Fig. 4.2) by dividing each cell  $\mathcal{C}_i$  into  $|V_i|$  horizontal

#### 4. An information-theoretical approach to stochastic thermodynamics

strips of finite relative height  $s_j^i > 0$  and an offset  $b_j^i = \sum_{k < j} s_k^i$ :

$$\mathcal{C}_j^i := \begin{cases} [0, 1] \times \Pi_i \cdot [b_j^i, b_j^i + s_j^i] & i \in V, j \in V_i, \\ \emptyset & \text{else.} \end{cases} \quad (4.27)$$

The dynamics  $\Phi$  maps each horizontal strip,  $\mathcal{C}_j^i \subset \mathcal{C}_i$ , to a *vertical* strip,  $\hat{\mathcal{C}}_j^i := \Phi \mathcal{C}_j^i \subset \mathcal{C}_j$  defined as

$$\hat{\mathcal{C}}_j^i := [\hat{b}_j^i, \hat{b}_j^i + \hat{s}_j^i] \times \Pi_j \cdot [0, 1]. \quad (4.28)$$

The number  $\hat{s}_j^i \in [0, 1]$  denotes the relative width of the vertical strips in cell  $\mathcal{C}_j$  and thus fulfil  $\sum_i \hat{s}_j^i = 1$ . The *offsets* read  $\hat{b}_j^i := \sum_{k < i} \hat{s}_j^k$ . To obtain an analytically tractable model, we choose  $\Phi$  to act in an affine-linear way on the cells. This means that points in any strip are mapped such that the horizontal direction is contracted uniformly by a factor  $\hat{s}_j^i < 1$  whereas the vertical direction is expanded by a factor  $(s_j^i)^{-1} > 1$ . Formally,

$$\begin{aligned} \Phi: \Gamma &\rightarrow \Gamma \\ (x_1, x_2, i) &\mapsto \left( \hat{b}_j^i + \hat{s}_j^i x_1, \frac{\Pi_j}{s_j^i} \left( \frac{x_2}{\Pi_i} - b_j^i \right), j \right) \end{aligned} \quad (4.29)$$

for  $b_j^i < \frac{x_2}{\Pi_i} \leq b_{j+1}^i$ . Hence, a NMBM is fully defined by the numbers  $\Pi_i$ ,  $s_j^i$  and  $\hat{s}_j^i$ . Note that the matrix  $A$  with entries  $a_{ij} = \text{sgn}(s_j^i) = \text{sgn}(\hat{s}_j^i)$  is the adjacency matrix of the graph  $G$ .

#### 4.2.2. Reversibility and further constraints

By imposing further constraints on the numbers  $s_j^i$ ,  $\hat{s}_j^i$  and  $\Pi_i$ , one can implement additional features into the dynamics. More precisely, one can abstractly realize time-reversible dynamics similar to the models used in (NE)MD, *cf.* Section 2.5. For the remainder of this work, let us make the following assumption for the graph  $G$ : (i)  $G$  is connected, *i.e.* that its adjacency matrix is irreducible. (ii)  $G$  is dynamically reversible, *i.e.* the presence of a directed edge  $e := (i, j) \in E$  implies that  $-e := (j, i) \in E$ , too. (iii) Each vertex is connected to itself, *i.e.*  $(i, i) \in E, \forall i \in V$ .

Then, the dynamics can be made reversible with a measure-preserving involution  $\mathcal{I}$ . It has been shown [TV00; Vol02] that a necessary requirement for the existence of such an involution is the symmetry

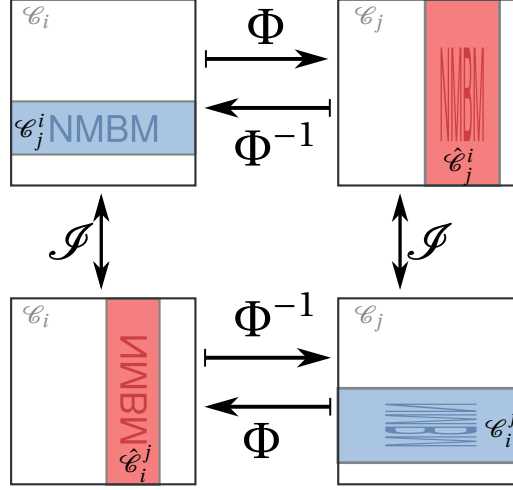
$$\hat{s}_j^i = s_i^j, \forall i, j. \quad (4.30)$$

Then, the map

$$\begin{aligned} \mathcal{I}: \Gamma &\rightarrow \Gamma \\ (x_1, x_2, i) &\mapsto (1 - \Pi_i^{-1} x_2, \Pi_i(1 - x_1), i), \end{aligned} \quad (4.31)$$

fulfils the conditions (2.70), *i.e.* it is a measure-preserving ( $\det(D\mathcal{I}) = 1$ ) involution that





**Figure 4.3.:** A reversible NMBM obeying Eq. (4.30) features a volume preserving involution  $\mathcal{S}$ , cf. Eq. (4.31), which interchanges horizontal and vertical stripes in one cell, i.e.  $\mathcal{S}\hat{\mathcal{C}}_j^i = \mathcal{C}_i^j$ . The dynamics  $\Phi$  is only volume preserving, if Eq. (4.33) holds as well. Again, the letters “NMBM” reflect how the dynamics and the involution acts on individual points  $x \in \Gamma$ .

facilitates time-reversal. Moreover, it acts locally on the cells, i.e. points in a cell  $\mathcal{C}_i$  stay there under the action of  $\mathcal{S}$ . The consequences of this symmetry are depicted in Figure 4.3.

For a generic choice of the parameters, the dynamics is dissipative. This fact is not influenced by the assumption of reversibility as prescribed by Equation (4.30). However, by imposing further constraints we are able to mimic conservative and uniformly conservative dynamics, which mimic isolated systems and other non-driven systems, respectively.

From the geometric specification of the multibaker map or from (4.29) it is clear that

$$\hat{j}^{(1)}(x) = \frac{|\hat{\mathcal{C}}_j^i|}{|\mathcal{C}_j^i|} = \frac{\Pi_j s_j^i}{\Pi_i s_j^i} \text{ for } x \in \mathcal{C}_j^i. \quad (4.32)$$

This means, a reversible multibaker map is uniformly conservative if and only if

$$\Pi_i s_j^i \equiv \Pi_j s_i^j. \quad (4.33)$$

In order to check whether a NMBM is non-uniformly conservative, which is a weaker property, we need to know the steady-state distribution. We will see in Section 4.2.5 that reversible NMBM are conservative if and only if the relative volumes  $s_j^i$  fulfil the Kolmogorov cycle criterion (2.35), i.e.

$$\prod_{t=1}^{\tau} s_{i_t}^{i_{t-1}} = \prod_{t=1}^{\tau} s_{i_{t-1}}^{i_t}. \quad (4.34)$$

#### 4. An information-theoretical approach to stochastic thermodynamics

for any sequence of  $\tau + 1$  states obeying  $i_0 = i_\tau$ .

In summary, NMBM are very versatile and can be used to reproduce the key features of general reversible dynamics. This is not limited to the case of reversible dynamics with a measure-preserving involution  $\mathcal{I}$ . For instance, Gaspard and co-workers also introduced non-reversible uniformly conservative maps of the unit square to itself [GW93; GD95]. We just mention that the most general form of NMBM also contains these maps as a special case.

##### 4.2.3. NMBM observables, priors and initial conditions

In the previous section we demonstrated the versatility of NMBMs to serve as a toy model for more realistic dynamics. Now, we choose an observable  $M: \Gamma \rightarrow \Omega$  that associates observations (measurement results)  $\omega \in \Omega$  to points  $x \in \Gamma$  in phase space. In Chapter 3 we have discussed in detail, how an observable  $M$  induces a phase space partition  $\mathcal{Q} = M^{-1}(\Omega)$ . The elements  $\mathcal{C}_\omega$  of that partition contain the pre-images of elements in  $\Omega$ .

For NMBM it is natural to choose  $\Omega = V$  and define  $M(x) := i$  for  $x \in \mathcal{C}_i$ . Then, the induced partition  $\mathcal{Q} = (\mathcal{C}_\omega)_{\omega \in \Omega}$  agrees with the partition  $\mathcal{Q} = (\mathcal{C}_i)_{i \in V}$  we used to define the NMBM. For *reversible* network multibaker maps, this partition also has an important symmetry: The involution  $\mathcal{I}$  factorizes on the partition elements (*cf.* Fig. 4.3), *i.e.*

$$\mathcal{I}\mathcal{C}_\omega = \mathcal{C}_\omega, \forall \omega \in \Omega. \quad (4.35)$$

In that case we say that the partition  $\mathcal{Q} = (\mathcal{C}_\omega)_{\omega \in \Omega}$  is *absolutely  $\mathcal{I}$ -invariant*.<sup>2</sup>

In the discussion of the previous chapter (Sec. 3.4.2) we have discussed properties of partition induced by *real* measurements on real (*i.e.* physical) systems. From that point of view, the assumption of an absolutely  $\mathcal{I}$ -invariant partition seems rather restrictive. However, we have also seen that the thermostated equations of motion used in (NE)MD are reversible with a measure-preserving involution (*cf.* Sec. 2.5.1 and Ref. [JR10]).

In that case, the measure-preserving time-reversal involution is given by flipping momenta and (possibly) auxiliary variables. Note that then *any* (MD-)observable which depends *only* on configurational degrees of freedom obeys the symmetry (4.35) on the induced cells. That is, any partition constructed in such a way is absolutely  $\mathcal{I}$ -invariant. This factorization of the involution over the partition is the reason why such a dynamics *always* satisfies a fluctuation theorem for the phase space expansion of phase space cylinders [Woj09]. For reversible multibaker maps on a one-dimensional lattice this was first realized in Ref. [RTV00].

Now let us discuss the priors we use in the specification of the initial (and coarse-grained) measure. For our abstract dynamics we have no special physical model in mind. Hence, there are no further constraints reflecting any additional knowledge. The

<sup>2</sup> This is a stronger condition than  $\mathcal{I}$ -invariance  $\mathcal{I}\mathcal{Q} = \mathcal{Q}$ , *i.e.*  $\forall \mathcal{C} \in \mathcal{Q}: \exists \mathcal{C}' = \mathcal{I}\mathcal{C} \in \mathcal{Q}$ .

compatible maximum-entropy distributions  $\varrho_\omega^{\text{pr}}$  are uniform on each cell:

$$\varrho_\omega^{\text{pr}} = \Pi_\omega^{-1}. \quad (4.36)$$

Thus, the assumed entropy for each cell can be obtained in a ‘‘Boltzmannian’’ way as the logarithm of the associated phase-space volume:

$$s^{\text{pr}}(\omega) = \log \Pi_\omega. \quad (4.37)$$

In order to calculate the other fundamental  $\tau$ -chains (4.11)–(4.13), we need to specify the initial condition  $\varrho_0$ . To be consistent with the arguments brought forward in Section 4.1, we take the coarse-grained density as our initial ensemble, *i.e.*

$$\varrho_0^{\text{fg}} = \varrho_0 = \varrho_0^{\text{cg}} = \sum_{\omega \in \Omega} [\chi_\omega(x) p_t(\omega) \Pi_\omega^{-1}]. \quad (4.38)$$

#### 4.2.4. Evolution of the densities

Let us recall Definition 3.3 of the trajectory  $\underline{M}^{(\tau)}: \Gamma \rightarrow \Omega^{\tau+1}$ . It maps an initial condition  $x$  to its observable time-series of length  $\tau$ . The latter is obtained from measurements on the iterations of  $x$  produced by the successive action of the map  $\Phi$ :

$$\underline{M}^{(\tau)} = (M \circ \Phi^t)_{t \in \{0, 1, \dots, \tau\}}$$

Then, for a point  $x \in \Gamma$  with time-series  $\underline{\omega}^{(\tau)} = \underline{M}^{(\tau)} x$ , equations (4.4) and (4.38) together with (4.32) yield the fine-grained density:

$$\begin{aligned} \varrho_\tau^{\text{fg}}(\Phi^\tau x) &= \varrho_0(x) \left( \prod_{t=1}^{\tau} \hat{j}^{(1)}(\Phi^t(x)) \right)^{-1} \\ &= \frac{p_0(\omega_0)}{\Pi_{\omega_\tau}} \prod_{t=1}^{\tau} \left( \frac{s_{\omega_t}^{\omega_{t-1}}}{s_{\omega_{t-1}}^{\omega_t}} \right). \end{aligned} \quad (4.39)$$

The phase space cylinder  $\mathcal{C}_0[\underline{\omega}^{(\tau)}] := (\underline{M}^{(\tau)})^{-1} \{\underline{\omega}^{(\tau)}\}$  contains all the points  $x \in \Gamma$  that give rise to the time-series  $\underline{\omega}^{(\tau)}$  starting at  $t = 0$ . Its  $\tau$ -fold iterated image is the image of the (negatively shifted) forward cylinder  $\mathcal{C}_{-\tau}[\underline{\omega}^{(\tau)}]$

$$\mathcal{C}_{-\tau}[\underline{\omega}^{(\tau)}] := \Phi^{(\tau)}(\mathcal{C}_0[\underline{\omega}^{(\tau)}]) = \underline{M}^{-1}(Z_{-\tau}[\underline{\omega}^{(\tau)}]). \quad (4.40)$$

Note that it can be defined recursively as

$$\mathcal{C}_{-\tau}[\underline{\omega}^{(\tau)}] \equiv \Phi(\mathcal{C}_{-\tau-1}[\underline{\omega}^{(\tau-1)}] \cap \mathcal{C}_{\omega_k}^{\omega_{k-1}}), \quad (4.41)$$

where  $\mathcal{C}_0[(\omega_0)] = \mathcal{C}_{\omega_0}$ .

Consider the volume (Lebesgue-measure) of the phase space cylinder  $\mathcal{C}_{-\tau}[\underline{\omega}^{(\tau)}]$ . One can easily verify that for any  $t \leq \tau \in \mathbb{N}$ ,  $\mathcal{C}_{-t}[\underline{\omega}^{(t)}]$  is a *vertical strip*. Intersecting it with the

#### 4. An information-theoretical approach to stochastic thermodynamics

horizontal strip  $\mathcal{C}_{\omega_k}^{\omega_{k-1}}$  reduces its volume by a factor  $s_{\omega_k}^{\omega_{k-1}} < 1$ . Through contraction and expansion in horizontal and vertical directions, respectively,  $\Phi$  changes the volume by another factor  $\hat{s}_{\omega_k}^{\omega_{k-1}} / s_{\omega_k}^{\omega_{k-1}}$ . Hence,  $\lambda(\mathcal{C}_{-t}[\underline{\omega}^{(t)}]) = s_{\omega_t}^{\omega_{t-1}} \lambda(\mathcal{C}_{-t}[\underline{\omega}^{(t-1)}])$  and after iteration to  $t = \tau$ ,

$$\lambda(\mathcal{C}_{-\tau}[\underline{\omega}^{(\tau)}]) = \Pi_{\omega_\tau} \prod_{t=1}^{\tau} \hat{s}_{\omega_t}^{\omega_{t-1}} = \Pi_{\omega_\tau} \prod_{t=1}^{\tau} s_{\omega_{t-1}}^{\omega_t}, \quad (4.42)$$

where for the last equality we used Eq. (4.30). The probability of observing a time-series  $\underline{\omega}^{(\tau)}$  at starting at time  $t = 0$  is

$$\begin{aligned} \mathbb{P}_0[\underline{\omega}^{(\tau)}] &= \int_{\mathcal{C}_0[\underline{\omega}^{(\tau)}]} \rho_t^{\text{fg}} dx = \int_{\mathcal{C}_{-\tau}[\underline{\omega}^{(\tau)}]} \rho_t^{\text{fg}}(\Phi^t x) dx \\ &= (\lambda(\mathcal{C}_{-\tau}[\underline{\omega}^{(\tau)}])) \cdot (\rho_t^{\text{fg}}(\Phi^t x)) \\ &= \left( \Pi_{\omega_\tau} \prod_{t=1}^{\tau} s_{\omega_{t-1}}^{\omega_t} \right) \cdot \left( \frac{p_0(\omega_0)}{\Pi_{\omega_\tau}} \prod_{t=1}^{\tau} \left( \frac{s_{\omega_t}^{\omega_{t-1}}}{s_{\omega_{t-1}}^{\omega_t}} \right) \right) \\ &= p_0(\omega_0) \prod_{t=1}^{\tau} s_{\omega_t}^{\omega_{t-1}}. \end{aligned} \quad (4.43)$$

Note that in the second line we used that the fine-grained density  $\rho^{\text{fg}}(\Phi^t x)$  is the same for all  $x \in \mathcal{C}_{-\tau}[\underline{\omega}^{(\tau)}]$ .

Equation (4.43) reveals an interesting fact: The probabilities evolve according to a Markov chain for a transition matrix  $\mathbb{W}$  with elements  $w_{\omega'}^\omega = s_{\omega'}^\omega$ . This means that our assumed initial distribution must be the extension of a one-sided Markov measure with transition matrix  $\mathbb{W}$ . We will come back to that observation in Section 4.3.3.

#### 4.2.5. The entropic $\tau$ -chains and their variations

Now we have all the ingredients to write down the fundamental and derived entropic  $\tau$ -chains. For brevity, we only state the variations. They are obtained by applying Eqs. (4.21) to expressions (4.11)–(4.13), which are evaluated using the expressions (4.32), (4.37) and (4.39):

$$\delta^{(\tau)} s_t^{\text{vis}} = \log \frac{p_{t+\tau-1}(\omega_{\tau-1})}{p_{t+\tau}(\omega_\tau)}, \quad (4.44a)$$

$$\delta^{(\tau)} s_t^{\text{pr}} = \delta^{(\tau)} s_t^{\times} = \log \frac{\Pi_{\omega_\tau}}{\Pi_{\omega_{\tau-1}}}, \quad (4.44b)$$

$$\delta^{(\tau)} \lambda_t = \log \frac{\Pi_{\omega_\tau} s_{\omega_{\tau-1}}^{\omega_\tau}}{\Pi_{\omega_{\tau-1}} s_{\omega_\tau}^{\omega_{\tau-1}}}. \quad (4.44c)$$

From the expressions for the fundamental variations we obtain the variations of the derived quantities as the linear combinations (4.22):

$$\delta^{(\tau)} s_t^{\text{cg}} = \log \frac{p_{t+\tau-1}(\omega_{\tau-1})}{p_{t+\tau}(\omega_\tau)} + \log \frac{\Pi_{\omega_\tau}}{\Pi_{\omega_{\tau-1}}}, \quad (4.45a)$$

$$\delta^{(\tau)} s_t^{\text{fg}} = \log \frac{s_{\omega_{\tau-1}}^{\omega_\tau}}{s_{\omega_\tau}^{\omega_{\tau-1}}} + \log \frac{\Pi_{\omega_\tau}}{\Pi_{\omega_{\tau-1}}}, \quad (4.45b)$$

$$\delta^{(\tau)} s_t^{\text{rel}} = \log \frac{p_{t+\tau-1}(\omega_\tau) s_{\omega_\tau}^{\omega_{\tau-1}}}{p_{t+\tau}(\omega_\tau) s_{\omega_{\tau-1}}^{\omega_\tau}}. \quad (4.45c)$$

Equation (4.43) tells us that the relative strip volumes  $s_{\omega'}^\omega$  are the transition probabilities of the Markovian process that generates the observed time series. In order to connect the expressions obtained here with the ones used in Markovian ST (*cf.* Section 2.4.2), we define:

$$\delta^{(\tau)} s_t^{\text{sys}} = \log \frac{p_{t+\tau-1}(\omega_{\tau-1})}{p_{t+\tau}(\omega_\tau)} = \delta^{(\tau)} s_t^{\text{vis}} \equiv \delta^{(\tau)} s_t^{\text{cg}} - \delta^{(\tau)} s_t^{\text{pr}}, \quad (4.46a)$$

$$\delta^{(\tau)} s_t^{\text{med}} = \log \frac{s_{\omega_\tau}^{\omega_{\tau-1}}}{s_{\omega_{\tau-1}}^{\omega_\tau}} = -\delta^{(\tau)} \lambda_t + \delta^{(\tau)} s_t^{\text{pr}}, \quad (4.46b)$$

$$\delta^{(\tau)} s_t^{\text{tot}} = \log \frac{\phi_{\omega_\tau}^{\omega_{\tau-1}}(t+\tau-1)}{\phi_{\omega_{\tau-1}}^{\omega_\tau}(t+\tau)} = \delta^{(\tau)} s_t^{\text{vis}} - \delta^{(\tau)} \lambda_t + \delta^{(\tau)} s_t^\times. \quad (4.46c)$$

This is a remarkable result: We have reproduced the  $\tau$ -chains used in ST within an general information-theoretical framework for deterministic dynamics. Admittedly, the identifications made in Eqs. (4.46) were made in order to be consistent with Markovian ST. Consequently, one wonders how much of this result is truly general and how much is due to the special structure of NMBM as a model dynamics.

Regardless of the generality of this result, it yields the proof of Eq. (4.34): For NMBM, the average phase space contraction in the steady state is equivalent to the average entropy change in the medium identified in ST. In the latter framework, the Kolmogorov criterion (2.35) ensures reversibility and hence detailed balance. This in turn ensures that the average phase-space contraction in the steady state vanishes, *i.e.* that the NMBM is conservative.

In the following section, we discuss how to generalize this results to arbitrary microscopic dynamics, which we consider physical. Amongst other things, we motivate Eqs. (4.46) *without* referring to ST. Rather, we base our argument on the identification of dissipation with phase-space contraction, as is common in thermostated MD [ES02; SRE07; JR10].

### 4.3. Discussion

Let us now discuss how the information-theoretical framework outlined in the present chapter may serve as a microscopic foundation for ST. First, we give a motivation of Equations (4.46) without an explicit reference to ST. After that, we focus on the expression for the total entropy production as a relative entropy. In order to connect the present results to the those of Chapter 3, we discuss how the natural two-sided measure appears naturally for NMBM. Finally, we comment on the influence of the reference measure on our results.

#### 4.3.1. Consistent identification of system and medium entropy

In the introduction to the present chapter we emphasized that the information-theoretic formalism introduced here is independent of ST. However, we have motivated the definitions (4.46) to ensure consistency with Markovian ST. In contrast, now we motivate them from the point using the general perspective on the distinction of system and medium presented in Section 2.3.2. In addition, we make use of the arguments in Section 2.5, where we identified thermodynamic dissipation with phase space contraction.

To be consistent with the notion of a system's entropy as the entropy of a coarse-grained, experimentally accessible and thus *visible* ensemble, we set:

$$\begin{aligned} (s^{\text{sys}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] &:= (s^{\text{vis}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] \\ &\equiv -\log p_{t+\tau}(\omega_\tau). \end{aligned} \quad (4.47)$$

The entropy of the medium is more complicated. Firstly it should take into account the integrated dissipation into the medium,  $s^{\text{flow}}$ . This term accounts for the macroscopic irreversibility characterized by the calorimetrically accessible heat flow from the system to the medium. In correspondence to the (NE)MD models discussed in Section 2.5 we thus set:

$$(s^{\text{flow}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] := -\lambda_t^{(\tau)}[\underline{\omega}^{(\tau)}]. \quad (4.48)$$

Moreover, the medium entropy should contain a hidden contribution  $s^{\text{hidd}}$ , which arises due to our (subjective) assumptions of the distribution on the cells. As we mentioned in Section 2.2, the cross-entropy (4.13) quantifies this mismatch, which we interpret as one contribution to the hidden entropy. Additionally, we have a second contribution that amounts to the entropy  $H[\varrho_t^{\text{fg}}] = S_t^{\text{fg}}$  of the microscopic ensemble at time  $t$ . The difference of these two terms defines the *hidden entropy*:

$$(s^{\text{hidd}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] := (s^\times)_t^{(\tau)}[\underline{\omega}^{(\tau)}] - S_t^{\text{fg}}. \quad (4.49)$$

Hence, the medium entropy consists of the contributions

$$\begin{aligned} (s^{\text{med}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] &:= (s^{\text{flow}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] + (s^{\text{hidd}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] \\ &\equiv - \int_{\mathcal{C}_0[\underline{\omega}^{(\tau)}]} \rho_t^{\text{fg}} \Big|_{\underline{\omega}^{(\tau)}} \log(\hat{J}^{(\tau)}) \, dx + s^\times(\omega_\tau) - S_t^{\text{fg}}. \end{aligned} \quad (4.50)$$

In the context of Markovian ST, Seifert pointed out that such a hidden contribution follows from the assumption of local equilibrium [Sei11]. In contrast, our argument explicitly establishes the role of the deterministic dynamics and the real microscopic distribution.

The total entropy is the sum of the contributions assigned to the medium and the system:

$$\begin{aligned} (s^{\text{tot}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] &:= (s^{\text{sys}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] + (s^{\text{med}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] \\ &\stackrel{(4.50)}{\equiv} (s^{\text{vis}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] - \lambda_t^{(\tau)}[\underline{\omega}^{(\tau)}] + (s^\times)_t^{(\tau)}[\underline{\omega}^{(\tau)}] - S_t^{\text{fg}} \\ &\stackrel{(4.17b)}{\equiv} (s^{\text{vis}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] + (s^\times)_t^{(\tau)}[\underline{\omega}^{(\tau)}] - (s^{\text{fg}})_t^{(\tau)}[\underline{\omega}^{(\tau)}] \\ &\stackrel{(4.17c)}{\equiv} (s^{\text{rel}})_t^{(\tau)}[\underline{\omega}^{(\tau)}]. \end{aligned} \quad (4.51)$$

This is an important result: It identifies the total entropy as a relative entropy between the real (fine-grained) density and our assumed coarse-grained density. Note that the latter is “subjective” in the sense that it depends on the maximum-entropy priors  $\rho_\omega$ .

The total entropy is a Kullback–Leibler divergence and thus always positive. Moreover, we see that the value of the fine-grained entropy  $S_t^{\text{fg}}$  at time  $t$ , which we identified as a part of the hidden entropy production, cancels in the corresponding variation  $\delta^{(\tau)} s_t^{\text{tot}}$ . In addition, the NMBM dynamics presented in Section 4.2 provides a consistency check: For this analytically tractable dynamics, the identifications (4.46), which we have just motivated generally, agree with the expressions used in ST.

### 4.3.2. Positivity of the variation of the total entropy

In the previous Subsection 4.3.1, we motivated the identification of the total entropy with the relative entropy. Let us discuss this fact in the light of the second law of thermodynamics. In that perspective, we should have that  $\Delta^{(\tau)} S_t^{\text{rel}} \geq 0$  for all  $t, \tau > 0$ . The present author conjectures that the following (non-rigorous) argument can be transformed into a proper proof.

Without loss of generality we choose  $t = 0$  as the initial time where the system is prepared. That point in time is special, because in the framework presented above it represents the point in time where the coarse-grained density  $\rho_0^{\text{cg}}$  agrees with the fine-grained density  $\rho_0^{\text{fg}}$ . The initial density was chosen as a distribution with maximum entropy. By definition, a maximum entropy (MaxEnt) distribution is the least biased distribution compatible with a set of constraints formalizing prior knowledge about the system [Jay03]. Hence, we understand a MaxEnt distribution as the one with the least

#### 4. An information-theoretical approach to stochastic thermodynamics

amount of “intrinsic structure” among all compatible distributions. In the course of the dynamics (unless they are uniformly conservative), phase space is contracted and expanded. Thus, additional structure is introduced into each cell. Therefore, in the course of time the difference of the real distribution within a cell  $\rho_t^{\text{fg}}|_{\mathcal{C}_\omega}$  and the maximum entropy prior becomes more and more pronounced.

This difference of one probability distribution with respect to another is quantified by the Kullback–Leibler divergence. Hence, the relative entropy and thus the total entropy should always increase. Consequently, a second law stating that  $\Delta^{(r)} S_t^{\text{rel}} \geq 0$  should also hold in this case.

Note that this holds for the dynamically reversible NMBM we have discussed above. They reproduce the expressions known from ST and we have already discussed in Section 2.4.2. As such the total entropy production is a KL-divergence and thus always positive, cf. Equation (2.42c).

#### 4.3.3. Foundations of Markovian stochastic thermodynamics

After having motivated the identification of system and medium entropies in the last previous in general terms, we return to Markovian dynamics. For NMBM, we have seen that the expressions (4.47), (4.50) and (4.51) reproduce the expressions known from ST, cf. Eqs. (4.46). Although NMBM constitute an abstract model without a physical justification, they intuitively represent the dynamics of more general (hyperbolic) dynamics. In particular, we have demonstrated how NMBM can be made reversible and further tuned to show features of conservative and dissipative systems.

In the light of Chapter 3, let us review the properties of NMBMs that yield Equations (4.44): Firstly, a necessary requirement for the Markovian evolution of time-series is that the observable defining the coarse-grained states  $\omega \in \Omega$  induces a Markov partition  $\mathcal{Q}$ .<sup>3</sup> Secondly, the fact that  $\mathcal{Q}$  is absolutely  $\mathcal{S}$ -invariant with a measure-preserving time-reversal involution  $\mathcal{S}$  ensures the validity of Eq. (4.35). For NMBM it yields the relation  $s_j^i = \hat{s}_i^j$ , which relates the relative weights of images and pre-images of the strips. Thirdly, even if the topological requirements for a Markovian evolution of the observed time-series is fulfilled, we still need the right initial conditions.

Let us formulate this conditions in a more general context. We have already discussed the issue of Markov partitions for real, *i.e.* physical systems in Section 3.4.2. For now, assume that this assumption is fulfilled at least for all practical purposes.

Regarding the existence of a measure-preserving time-reversal involution  $\mathcal{S}$  and the existence of an absolutely  $\mathcal{S}$ -invariant partition, recall the discussion of thermostated equations of motions in Section 2.5.1. We have seen that *any* observable that depends only on the coordinates yields a partition that obeys  $\mathcal{S}\mathcal{C}_\omega = \mathcal{C}_\omega$ . Hence, absolute  $\mathcal{S}$ -invariance might be in fact a generic symmetry of partitions induced by physical observables on physical microscopic dynamics. Additionally, recent work points out that care has to be

<sup>3</sup>Note that the natural partition for NMBM is also *generating*, but this is not needed for the argument.



taken when applying the framework of ST to systems where momenta of particles are treated explicitly [KN13].

The most subtle, but arguably most important question regards the choice of initial conditions. In Chapter 3 we have discussed Markov measures on phase space. These are measures, such that (forward) time-series yield Markovian statistics. The natural Markov measure used the notion of a natural measure on phase space. The next paragraph comments on the significance of the natural two-sided measure introduced in Definition 3.32.

**Significance of the two-sided natural measure for reversible dynamics** The natural two-sided measure is constructed using the natural transition matrices  $\mathbb{Q}$  and  $\tilde{\mathbb{Q}}$  obtained from the natural measure  $\mu_\Phi$  for  $\Phi$  and the natural measure  $\mu_{\Phi^{-1}}$  for the inverse dynamics  $\Phi^{-1}$ . Suppose the natural measure for both  $\Phi$  and  $\Phi^{-1}$  is an SRB measure, *i.e.* it has absolutely continuous densities along the unstable manifolds. The unstable manifolds of  $\Phi^{-1}$  are the stable manifolds of  $\Phi$  and vice versa.

For a dynamics featuring a measure-preserving time-reversal involution  $\mathcal{S}$  which acts locally on the partition elements, this yields a geometric interpretation of the natural measure. First note that every measure preserving time-reversal involution  $\mathcal{S}$  ensures that  $\tilde{\mu} := \mu \circ \mathcal{S}$  is absolutely continuous, if  $\mu$  is absolutely continuous. Hence, by the definition of the natural measure and the time-reversal symmetry, we have

$$\begin{aligned} \mu_\Phi \circ \mathcal{S} &= \lim_{\tau \rightarrow \infty} \left( \frac{1}{\tau} \sum_{i=1}^{\tau} \mu \circ \Phi^{-1} \right) \circ \mathcal{S} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{i=1}^{\tau} \mu \circ \mathcal{S} \circ \Phi = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{i=1}^{\tau} \tilde{\mu} \circ \Phi \\ &= \mu_{\Phi^{-1}}. \end{aligned}$$

This means that applying  $\mathcal{S}$  allows us to switch between  $\mu_{\Phi^{-1}}$  and  $\mu_\Phi$ . If  $\mathcal{Q}$  is absolutely  $\mathcal{S}$ -invariant, this relation factorizes on the partition elements, *i.e.* it holds also if we constrain the natural measures to any cell  $\mathcal{C}_\omega$ . Within each cell, a natural Markov measure conditioned on the forward cylinders (*i.e.* the unstable manifolds) has to be proportional to the natural measure on this cell. The natural two-sided measure is a special way to specify the (transversal) density along the stable manifolds of  $\Phi$ , *i.e.* the unstable manifolds of  $\Phi^{-1}$ : On the latter, it obtains the density of the natural measure  $\mu_{\Phi^{-1}} = \mu_\Phi \circ \mathcal{S}$ . Hence, for an absolutely  $\mathcal{S}$ -invariant partition the natural two-sided measure constitutes an  $\mathcal{S}$ -invariant initial condition, *cf.* the discussion of the dissipation function in Section 2.5.4.

For a piecewise linear dynamics like NMBM, the natural measure  $\mu_\Phi$  conditioned on the unstable manifolds is piecewise constant [BB03]. With the above symmetry, the natural measure corresponds to the uniform initialization on each cell. Thus, for NMBM the coarse-grained measure obtained from the maximum entropy distributions on each cell is a two-sided natural Markov measure.

Let us return to Definition 3.31 of the natural transition matrix. Given an absolutely

#### 4. An information-theoretical approach to stochastic thermodynamics

$\mathcal{I}$ -invariant partition obeying Eq. (4.35), we see that

$$\begin{aligned} q_{\omega'}^{\omega} &\equiv \frac{\mu_{\Phi}(\Phi^{-1}\mathcal{C}_{\omega'} \cap \mathcal{C}_{\omega})}{\mu_{\Phi}(\mathcal{C}_{\omega})} = \frac{\mu_{\Phi}(\mathcal{C}_{\omega'} \cap \Phi\mathcal{C}_{\omega})}{\mu_{\Phi}(\mathcal{C}_{\omega})} \\ &= \frac{\mu_{\Phi^{-1}}(\mathcal{I}\mathcal{C}_{\omega'} \cap \mathcal{I}\Phi\mathcal{C}_{\omega})}{\mu_{\Phi^{-1}}(\mathcal{I}\mathcal{C}_{\omega})} = \frac{\mu_{\Phi^{-1}}(\mathcal{C}_{\omega'} \cap \Phi^{-1}\mathcal{C}_{\omega})}{\mu_{\Phi^{-1}}(\mathcal{C}_{\omega})} \equiv \tilde{q}_{\omega'}^{\omega}. \end{aligned}$$

The numbers  $\tilde{q}_{\omega'}^{\omega}$  are the entries of the natural transition matrix for the backward process  $\tilde{\mathbb{Q}}$ . Because the adjacency matrix of the backward process is the transpose of the adjacency matrix of the forward process, we have

$$\mathbb{A} = \text{sgn } q_{\omega'}^{\omega} = \text{sgn}(\tilde{q}_{\omega'}^{\omega}) = \mathbb{A}^T. \quad (4.52)$$

Hence, an absolutely invariant partition yields a *dynamically* reversible stochastic process. For a NMBM, the natural measure  $\mu_{\Phi^{-1}}$  for the inverse dynamics is constant along the  $x_2$  direction. Hence, the relative *volumes* of vertical strips  $\hat{s}_{\omega'}^{\omega}$  correspond to the entries of the natural matrix of the backward process  $\tilde{q}_{\omega'}^{\omega}$ . Then, the symmetry (4.30) expresses the fact that  $s_{\omega'}^{\omega} = q_{\omega'}^{\omega} = \tilde{q}_{\omega'}^{\omega} = \hat{s}_{\omega'}^{\omega}$ .

**A general conjecture** In Chapter 2 we have pointed out that the medium entropy relates to the irreversibility of the system's transitions. For a jump process, it generically compares a forward jump  $\omega \rightarrow \omega'$  to the corresponding backward jump  $\omega' \rightarrow \omega$  in an appropriately chosen backward process.

Hence, for the entropy change in the medium it seems natural to consider the logarithmic ratio:

$$s^{\text{med}} \sim \log \frac{q_{\omega'}^{\omega}}{\tilde{q}_{\omega'}^{\omega}}.$$

Indeed, for reversible NMBM this just yields the desired expression, *i.e.*  $s^{\text{med}} \sim \log \frac{s_{\omega'}^{\omega}}{s_{\omega}^{\omega'}}$ . We formally conjecture that this should also hold generally:

**4.1 Conjecture (Foundations of dynamically reversible, Markovian ST)** Let  $(\Gamma, \mathcal{B}, \Phi)$  be a dynamical system with a measure-preserving time-reversal involution  $\mathcal{I}$ . Let  $\mathcal{Q}$  be an absolutely  $\mathcal{I}$ -invariant Markov partition. Under the assumption that we take the two-sided natural Markov measures as initial conditions and as priors, we postulate that:

$$\delta^{(\tau)} s_t^{\text{med}}[\underline{\omega}^{(\tau)}] := -\delta^{(\tau)} \lambda_t[\underline{\omega}^{(\tau)}] + \delta^{(\tau)} s_t^{\text{pr}}[\underline{\omega}^{(\tau)}] \quad (4.53)$$

$$= \log \frac{q_{\omega'}^{\omega}}{q_{\omega}^{\omega'}}, \quad (4.54)$$

where  $q_{\omega}^{\omega'}$  are the elements of the natural transition matrix for  $\Phi$  and  $\mathcal{Q}$ .

If proven true, this conjecture provides a proper deterministic foundation of stochastic thermodynamics. It further enables the systematic study of the relaxation of certain

assumptions. The obvious generalization is to consider  $\mathcal{S}$ -invariant partitions which are not absolutely  $\mathcal{S}$ -invariant. This gives rise to a stochastic process which is not dynamically reversible.

In Secs. 4.2.3 and 2.5.4 we have mentioned Wojtkowski's abstract fluctuation theorem for the phase space contraction [Woj09]. For NMBM, the latter is connected to the entropy production identified in ST and thus reproduces the stochastic fluctuation relations. Actually, Wojtkowski's theorem is formulated for a more general situation where the backward microscopic dynamics  $\Psi$  does not need to be the inverse process  $\Phi^{-1}$ . All that is needed that  $\Phi$  and  $\mathcal{S} \circ \Psi \circ \mathcal{S}$  are related by a measure-preserving involution  $\mathcal{S}$ . For instance, this situation is found when the dynamics is driven out of equilibrium by magnetic rather than electrical fields. From the perspective of ST, the natural measure induced by  $\Psi$  then yields the transition rates that define the appropriate backward process. We expect this to provide a dynamical picture of the stochastic master fluctuation relations, which has been formulated by Seifert in Ref. [Sei12]. A more detailed discussion on that subject can be found in Section 7.2.4.

However, there is a caveat regarding the maximum entropy priors: For NMBM, the uniform distribution on each cell is both a maximum-entropy prior as well as the natural two-sided measure on phase space. This fact has its origin in the linearity of NMBM. For any abstract dynamics, which is non-linear, the natural two-sided measure has a more complicated structure, *i.e.* it is *not* uniform on the cells. In the beginning of the present chapter, we argued that the prior measure needs to maximize entropy with respect all the *additional knowledge* we have about the dynamics.

Throughout this work we emphasized that *physical* microscopic dynamics — like Hamiltonian dynamics or (NE)MD equation — exhibit certain features: They allow for a measure-preserving involution. Given a measurement observable, that is invariant under time-reversal, this involution factorizes on the cells of the induced partition. Stochastic thermodynamics additionally assumes a Markovian evolution of the observable states. In that regard, the natural two-sided Markov measure appears as the natural candidate for a maximum-entropy measure that reflects these additional constraints.

#### 4.3.4. Influence of the reference measure

Finally, we comment on the influence of the reference measure in our considerations. In a recent work, Polettini investigates how the choice of a reference measure formalizes another aspect of subjectivity. This subjectivity is reflected in the fact that probability densities  $\rho$  and thus entropies  $H[\rho]$  depend on the choice of reference [Pol12; PV13]. In addition, the value of the differential entropy change upon coordinate transformations  $T: \Gamma \rightarrow \Gamma$ .

Polettini argues that the choice of a reference measure needs to be understood as expressing some prior assumption about the system. The Lebesgue measure thereby reflects the *microcanonical* prior, where all states are assumed to have equal a-priori likelihood to

#### 4. An information-theoretical approach to stochastic thermodynamics

appear.<sup>4</sup> In his interpretation, the choice of prior constitutes a gauge transformation. We will come back to that interpretation in the following Chapter 5. In ST, the expressions associated to the entropy (changes) in the system and the medium change under the action of the gauge transformation, *i.e.* the choice of reference [Pol12]. However, the expression associated with the change of the *total entropy* is gauge-invariant.

In the present framework, this statement is equally true if we identify the total entropy with the relative entropy, *cf.* Eq. (4.9), *i.e.* a Kullback–Leibler divergence. As such it has the property to be *invariant* under the choice of the reference measure or coordinate transformations. Hence, also Polettini’s considerations point to the validity of the identifications made here.

#### 4.4. Summary

In this chapter, we have proposed an information-theoretic framework to constitute the microscopic foundations of stochastic thermodynamics. We started by formalizing the process of taking repeated measurements in the spirit of Jaynes’ view of statistical physics as a theory of statistical inference. In this framework, we used a deterministic microscopic dynamics to motivate fundamental information-theoretic  $\tau$ -chains  $s^{\text{vis}}$ ,  $s^{\text{pr}}$ ,  $s^{\times}$  and  $\lambda$ . These  $\tau$ -chains are observables that depend on finite time-series  $\underline{\omega}^{(\tau)}$  of length  $\tau$ .

From the fundamental  $\tau$ -chains we were able to construct the derived chains  $s^{\text{cg}}$ ,  $s^{\text{fg}}$  and  $s^{\text{rel}}$  as linear combinations. The derived chains average to the entropy of the inferred coarse-grained ensemble, the real microscopic ensemble and their relative entropy, respectively. We further motivated the expressions  $s^{\text{sys}}$ ,  $s^{\text{med}}$  and  $s^{\text{tot}}$ , which correspond to the entropy of the system, the medium and the sum of both, respectively. In that step, we used the connection of dissipation and phase space contraction in physical microscopic models.

In order to exemplify our considerations, we introduced network multibaker maps as a versatile and generic, yet analytically tractable model systems. We showed how for these maps our abstract thoughts are consistent with stochastic thermodynamics and how the entropic concept of the latter emerge naturally. Further, we used network multibaker maps to motivate the importance of the natural two-sided Markov measure, which was introduced in Chapter 3.

This in turn lead to Conjecture 4.1 which formalizes the microscopic foundations of ST in a general information-theoretical framework. At this point, we have concluded the first part of the thesis regarding the *microscopic* foundations. In the following two Chapters 5 and 6 we are concerned with the mathematical structure of Markovian ST on finite state spaces.

Hence, the discussion there is independent of the hypothesis of a deterministic phase space dynamics we have assumed in the previous and present Chapters 3 and 4. However,

---

<sup>4</sup> Ref. [PV13] beautifully explains this fact using the different perspective that parents and their children have regarding the “disorder” they perceive in the children’s playing room.

we will continue using notions of entropy and entropy production as consistency criteria. Interestingly, (an extension of) the gauge invariance discussed in Section 4.3.4 will play a central role.



## 5. The structure of Markov jump processes

“ *Ich will jetzt beweisen, daß die Auflösung der Gleichungen [auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird], sich allgemein angeben lassen.* ”

---

G. Kirchhoff, *Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird*, 1847

### What is this about?

In the present chapter we are concerned with the algebraic and topological structure of Markovian dynamics on finite networks. In Section 2.4.2 we have seen how the network of states for a model of ST can be represented as an (undirected) graph. Graphs are further commonly used to represent electrical circuits.

The introductory quote of Kirchhoff is from his work on the distribution of electrical currents in circuits build out of resistors and batteries. An attempted translation reads:

I am going to prove that the solution of the equations [encountered in the study of the distribution of galvanic currents] can be stated in a general way.

The results proved by Kirchhoff are commonly known as Kirchhoff's current and voltage law, respectively.

Nowadays these results have been generalized by mathematicians as the *matrix-tree theorem* of graph theory [Tut98]. As such, they are algebraic-topological results that hold for arbitrary graphs — independent of the physical or mathematical problem they represent. Here, we discuss their relevance for Markov processes on finite state spaces. Consequently, they are going to be useful for the understanding of the structure of models used in ST.

We start the presentation in Kirchhoff's original setting, *i.e.* the study of electrical networks. The electrical currents running in an electrical circuit have their analogue in the probability currents of a stationary Markov process. Building up on this idea, the first result of the present chapter is a complete analogy between electrical circuits built from resistors and batteries with dynamically reversible Markov processes. Moreover, this analogy carries over to ST.

## 5. The structure of Markov jump processes

After that we investigate the algebraic structure of the *physical observables* corresponding to measurable currents in ST. A result is the generalization of the so-called Schnakenberg decomposition [Sch76] to arbitrary physical observables [Alt+12].

The classical Schnakenberg decomposition is a result formulated for the steady-state ensemble averages of physical observables. Here, we generalize this result further to statements about *fluctuations* of physical observables away from their expectation value. The main tool in our analysis will be the theory of large deviations for Markov processes. A central object of that theory is the *rate function*  $I_\varphi(x)$  for physical observables  $\varphi$ , which is a statement about their entire spectrum of fluctuations.

The main result of this treatment can be summarized as follows: Firstly, we provide a fully analytical approach to the fluctuation spectrum of arbitrary physical observables, which does *not* involve the solution of a complicated eigenvalue problem. Secondly, we show that the fluctuation spectrum of *any* physical observable is determined by the fluctuation spectrum of the probability currents on a special set of edges.

Parts of the results presented in the present chapter have already been published by the author in Ref. [Alt+12]. The main theorems have been obtained in the context of the Artur Wachtel's M.Sc. thesis [Wac13], which was jointly supervised by Jürgen Vollmer and the author of the present thesis. The proofs of the main results can be found there and in an unpublished manuscript by these authors [WVA14].

### 5.1. Kirchhoff's laws and an electrical analogy

In the previous Chapters 3 and 4 we have discussed the deterministic microscopic origins of (Markovian) stochastic processes. Here, we complement this perspective with a discussion of the structure of a stochastic network. More precisely, we focus on (non-equilibrium) steady states of continuous time Markovian processes. Let us briefly review the set-up for this chapter:

Similar to Section 2.4.2, let  $\omega \in \Omega = \{1, 2, \dots, N\}$  denote the elements of a finite state space. The probability distribution  $\mathbf{p}_t$  evolves according to the continuous-time master equation (2.28)

$$\partial_t \mathbf{p}_t = \mathbf{p}_t \mathbb{W}.$$

For continuous-time dynamics, the matrix  $\mathbb{W}$  is a *rate* matrix, rather than a stochastic matrix. Its off-diagonal elements  $w_{\omega'}^\omega$ ,  $\omega \neq \omega'$  are the time-independent transition rates. The diagonal element  $w_\omega^\omega = -\sum_{\omega'} w_{\omega'}^\omega = -\langle \tau_\omega \rangle^{-1}$  amounts to the negative of the escape rate from state  $\omega$ . The inverse of the escape rate is the average *staying time*: For a Markov process, the staying time  $\tau_\omega$  for a state  $\omega$  obeys an exponential distribution  $\mathbb{P}[\tau_\omega = t] \propto \exp\left(-\frac{t}{\langle \tau_\omega \rangle}\right)$ . Similar to the previous chapters, we allow only one type of transition between any two states. A generalization to multiple types of transitions is possible and has been discussed in Refs. [AG07; EB10; FDP11; Esp12]. As we have discussed above, the



reversibility of microscopic physical laws motivates dynamical reversibility, *i.e.*  $w_{\omega'}^{\omega} > 0 \Leftrightarrow w_{\omega}^{\omega'} > 0$ .

We visualize the network of states as a graph  $G = (V, E)$  with vertices  $v \in V$  and edges  $e \in E$ . In the present context of Markovian processes, we identify  $V \sim \Omega$  and draw an edge  $e = (\omega, \omega')$  wherever  $w_{\omega'}^{\omega} > 0$ . At time  $t$  the system is in state  $\omega$  with a probability  $p_t(\omega)$ . If the network is connected (*cf.* Section 2.4.2), there exists a unique (invariant) steady-state distribution  $\mathbf{p}_{\infty}$  [Fel68].

### 5.1.1. Steady states and Kirchhoff's current law

Henceforth, we focus on the steady state of a Markov process on a finite set of states  $\omega \in \Omega$ . With the invariant distribution  $\mathbf{p}_{\infty}$ , the Master equation (2.28) in matrix form reads

$$0 = \mathbf{p}_{\infty} \mathbb{W}. \quad (5.1)$$

Note that for the time-continuous case the invariant density is a left eigenvector of the rate matrix  $\mathbb{W}$  for the eigenvalue *zero*.

It is instructive to discuss this equation also by looking at the individual components of the vector equation (5.1). With the definition of the steady state probability fluxes

$$\phi_{\omega'}^{\omega} := p_{\infty}(\omega) w_{\omega'}^{\omega}, \quad \omega \neq \omega',$$

and the steady state currents

$$J_{\omega'}^{\omega} := \phi_{\omega'}^{\omega} - \phi_{\omega}^{\omega'},$$

the master equation (5.1) in component form reads

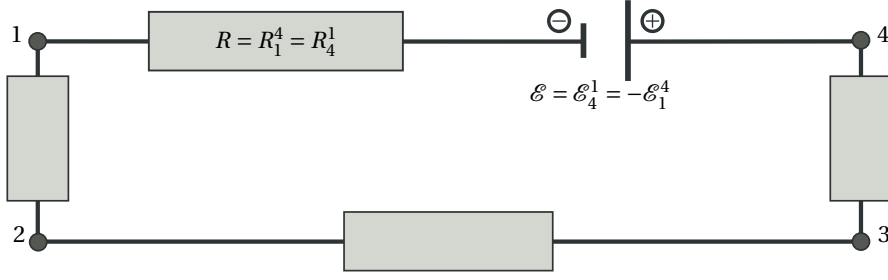
$$\sum_{\omega'} J_{\omega'}^{\omega} = \sum_{\omega'} \left[ \phi_{\omega'}^{\omega} - \phi_{\omega}^{\omega'} \right] = 0, \quad \forall \omega. \quad (5.2)$$

Because of probability conservation, the net current  $\sum_{\omega'} J_{\omega'}^{\omega}$  arriving at each vertex  $\omega'$  in the graph must vanish. Equivalently, the *influx*  $\sum_{\omega'} \phi_{\omega'}^{\omega}$  into a state  $\omega$  equals its *outflux*  $\sum_{\omega'} \phi_{\omega}^{\omega'}$ .

By definition, an *equilibrium system* fulfils detailed balance. Then, we have that  $J_{\omega'}^{\omega} = 0$  for all  $\omega, \omega' \in \Omega$ , *i.e.* each individual term in the above sum in Eq. (5.2) vanishes. In contrast, for a non-equilibrium steady state (NESS), there are at least some edges  $(\omega, \omega')$  that support non-vanishing currents  $J_{\omega'}^{\omega}$ .

For currents flowing in electrical networks, Eq. (5.2) is known as *Kirchhoff's current law* [Kir47]. It has its origin in the conservation of electrical charge. In the next subsection we will elaborate further on the analogy between electrical and stochastic networks.

## 5. The structure of Markov jump processes



**Figure 5.1.:** An electrical network. Each wire connecting two vertices  $\omega$  and  $\omega'$  has a positive resistance  $R_{\omega'}^{\omega} = R_{\omega}^{\omega'}$ . In addition, a battery-like element with electromotance  $\mathcal{E} = \mathcal{E}_4^1 = -\mathcal{E}_1^4$  between vertices 1 and 4 results in a current  $J = J_4^1$  with the same sign as  $\mathcal{E}$ .

### 5.1.2. Kirchhoff's second law and an electrical analogy

For our electrical analogy, let us review Kirchhoff's second law. This law is also known as the *mesh rule* or *voltage law* and relates the voltage drops  $V$  due to electrical resistance to the electromotive force  $\mathcal{E}$  of a battery.

Consider an electrical circuit consisting of battery-like elements and resistors like the one shown in Fig. 5.1. We label the vertices by  $\omega \in \Omega$ . Then, a wire connecting two vertices is identified by an edge  $(\omega, \omega')$ . A battery-like element on an edge  $(\omega, \omega')$  is characterized by its *electromotive force* or *electromotance*  $\mathcal{E}_{\omega'}^{\omega}$ . It can be either positive or negative, where the sign characterizes the direction of the flow of the mobile charges. Consequently, the matrix containing the electromotances for all wires is anti-symmetric, i.e.  $\mathcal{E}_{\omega'}^{\omega} = -\mathcal{E}_{\omega}^{\omega'}$ . In contrast, the *Ohmian resistances*  $R_{\omega'}^{\omega} = R_{\omega}^{\omega'}$  are positive and symmetric. If no current is flowing between two vertices  $\omega$  and  $\omega'$ , a battery with electromotance  $\mathcal{E}_{\omega'}^{\omega}$  creates a voltage difference  $\Delta U_{\omega'}^{\omega} = -\mathcal{E}_{\omega'}^{\omega}$ . The voltage  $U_{\omega}$  is measured with respect to an arbitrary reference potential  $U_0 \equiv 0$ . If a current  $J_{\omega'}^{\omega}$  is flowing over a resistor  $R_{\omega'}^{\omega}$ , the *absolute value* of the difference to the reference potential  $\Delta U_{\omega'}^{\omega}$  changes by the amount  $-V_{\omega'}^{\omega} = R_{\omega'}^{\omega} J_{\omega'}^{\omega}$ :

$$\Delta U_{\omega'}^{\omega} = -\mathcal{E}_{\omega'}^{\omega} - V_{\omega'}^{\omega} = -\mathcal{E}_{\omega'}^{\omega} + R_{\omega'}^{\omega} J_{\omega'}^{\omega}. \quad (5.3)$$

Note that the motance has the same sign as the current it produces and hence we have  $\text{sgn}(\mathcal{E}_{\omega'}^{\omega}) = \text{sgn}(R_{\omega'}^{\omega} J_{\omega'}^{\omega}) \equiv -\text{sgn}(V_{\omega'}^{\omega})$ . The dissipated power  $P_{\omega'}^{\omega}$  along edge  $(\omega, \omega')$  equals the product of the current times the negative of the voltage drop  $V_{\omega'}^{\omega}$ :

$$P_{\omega'}^{\omega} = -V_{\omega'}^{\omega} J_{\omega'}^{\omega} = \mathcal{E}_{\omega'}^{\omega} J_{\omega'}^{\omega} \geq 0. \quad (5.4)$$

The above mentioned *voltage law* then states that the difference of reference potential  $U_{\omega}$  along any cycle  $\alpha^{(\tau)} = (\omega_0, \dots, \omega_{\tau-1}, \omega_{\tau})$  in the network is conserved:

$$0 = \sum_{i=1}^{\tau} \Delta U_{\omega_i}^{\omega_{i-1}}. \quad (5.5)$$

An equivalent statement is that the sum off all voltage drops  $V_{\omega_i}^{\omega_{i-1}}$  equals the sum of the

**Table 5.1.:** Electrical and thermodynamic analogies, cf. Ref. [Alt+12]. FED denotes free energy differences as in Hill's theory, [Hil77]. Note that the analogy is based in the steady state values of the fluxes and currents.

symbol	definition	thermodynamic	electric
$U_\omega$	$-\log p_\infty(\omega)$	state variable	potential
$\Delta U_{\omega'}^\omega$	$\log[p_\infty(\omega) / p_\infty(\omega')]$	difference of a state variable	potential difference
$J_{\omega'}^\omega$	$\phi_{\omega'}^\omega - \phi_\omega^{\omega'}$	current	
$A_{\omega'}^\omega$	$\log[\phi_{\omega'}^\omega / \phi_\omega^{\omega'}]$	affinity, <i>gross FED</i>	—
$\mathcal{E}_{\omega'}^\omega$	$\log[w_{\omega'}^\omega / w_\omega^{\omega'}]$	motance, <i>basic FED</i>	electromotance
$V_{\omega'}^\omega$	$-\mathcal{E}_{\omega'}^\omega$	negative motance	voltage drop
$R_{\omega'}^\omega$	$V_{\omega'}^\omega / J_{\omega'}^\omega$	—	resistance
$P_{\omega'}^\omega$	$-V_{\omega'}^\omega J_{\omega'}^\omega$	—	dissipated power
$P$	$\frac{1}{2} \sum_{\omega, \omega'} P_{\omega'}^\omega$	med/tot entropy production	tot. dissipated power

electromotances  $\mathcal{E}_{\omega_i}^{\omega_{i-1}}$  along any cycle  $\alpha^{(\tau)}$ :

$$\sum_{i=1}^{\tau} \mathcal{E}_{\omega_i}^{\omega_{i-1}} = \sum_{i=1}^{\tau} V_{\omega_i}^{\omega_{i-1}}. \quad (5.6)$$

Let us now come back to the case of Markov processes. In Section 2.4.2 we have introduced the *affinity*  $A_{\omega'}^\omega$  and the *motance*  $\mathcal{E}_{\omega'}^\omega$ , of an edge  $(\omega, \omega')$ . We saw how they appear in Schnakenberg's network theory [Sch76] and discussed their relation to the (change of) entropy in ST. In the context of a steady-state Markov process, it is possible to define quantities on edges and vertices that fulfil expressions that are the analogues of Kirchoff's laws. We define the following quantities for any vertex  $\omega$  and any edge  $(\omega, \omega')$  where the steady-state current  $J_{\omega'}^\omega$  does not vanish:

$$U_\omega := -\log p_\infty(\omega), \quad (5.7a)$$

$$\Delta U_{\omega'}^\omega := \log \frac{p_\infty(\omega)}{p_\infty(\omega')}, \quad (5.7b)$$

$$V_{\omega'}^\omega := -\mathcal{E}_{\omega'}^\omega = -\log \frac{w_{\omega'}^\omega}{w_\omega^{\omega'}}, \quad (5.7c)$$

$$R_{\omega'}^\omega := \frac{V_{\omega'}^\omega}{J_{\omega'}^\omega}. \quad (5.7d)$$

The definitions above fulfil the properties of their electrical counter-parts: The “resistance” matrix  $R$  is positive and symmetric, whereas the motance matrix  $\mathcal{E}$  as well as the current matrix  $J$  are anti-symmetric. By definition they fulfil the relations (5.3), (5.5) and (5.6).<sup>1</sup> With a factor  $\frac{1}{2}$  due to the double-counting of edges, the dissipated power (5.4) of an

<sup>1</sup>This even holds for the transient case, if we substitute  $p_\infty$  by  $p_t$ .

## 5. The structure of Markov jump processes

electrical network in the steady state amounts to

$$\begin{aligned} P &:= \frac{1}{2} \sum_{\omega, \omega'} P_{\omega'}^{\omega} = -\frac{1}{2} \sum_{\omega, \omega'} J_{\omega'}^{\omega} V_{\omega'}^{\omega} \\ &= \sum_{\omega, \omega'} \phi_{\omega'}^{\omega} \log \frac{w_{\omega'}^{\omega}}{w_{\omega}^{\omega'}} \end{aligned} \quad (5.8)$$

This expression agrees with the medium entropy production for a stationary Markov process. Note that because we are considering a steady state, the entropy change in the system vanishes and Eq. (5.8) similarly agrees with the *total* entropy production.

Different electrical analogies have been presented in the literature and are suitable for different purposes (see e.g. [ZS07; Fel68]). In Table 5.1 we summarize our analogy, which has first appeared in a slightly less general form in Ref. [Alt+12]. In addition, we give the *thermodynamic* interpretations as discussed in Section 2.4.2 and in Hill's work [Hil77].

## 5.2. Cycles and trees as the fundamental building blocks of networks

In the introduction to this chapter, we mentioned that Kirchhoff's results are a special case of more general *matrix-tree theorems* [Tut98]. In the present section, we provide an abstract algebraic framework for treating networks represented as graphs.

Let us recall the notation introduced in Section 2.4.2. A *directed graph*  $G^d = (V, E^d)$  consists of vertices  $v \in V$  and edges  $e = (v, v') \in E^d$ . For an *undirected graph*  $G^u = (V, E^u)$  the edges  $e = \{v, v'\} \in E^u$  are not ordered. Henceforth, we will only consider bi-directional graphs, where the presence of an edge  $(v, v') \in E^d$  implies that  $(v', v) \in E^d$ , too.

In contrast to the presentation in Section 2.4.2, we use the symbol  $v_i \in V$  rather than  $\omega$  to refer to a vertex. In doing so, we stress the generality of the present algebraic discussion: It is completely independent from the interpretation of the graph as a stochastic network. Similarly, we choose an (arbitrary) enumeration for the edges  $e_m \in E$ .

### 5.2.1. Anti-symmetric observables on the edges

As a motivation for what follows, let us introduce the notion of observables associated to edges. The quantities defined in Table 5.1 can be understood as real-valued functions  $\varphi$  on the set of directed edges:

$$\begin{aligned} \varphi: E^d &\rightarrow \mathbb{R}, \\ e = (v, v') &\mapsto \varphi(e) = \varphi_{v'}^v. \end{aligned} \quad (5.9)$$

More abstractly, they can be interpreted as (time-independent) *one-chains*, cf. Chapter 4 and Appendix A.2.

In particular, we are interested in *anti-symmetric observables*. For a graph which corresponds to a *dynamically reversible* Markov process representing a mesoscopic description of a physical system, such observables have a physical interpretation. To appreciate this

## 5.2. Cycles and trees as the fundamental building blocks of networks

fact, recall our motivation for a dynamically reversible graph in Section 2.4.2: The states  $\omega$  represent the values of some physical observable which depends only on the symmetric coordinates. Physical currents are the instantaneous changes of measurable quantities. Hence, they have to be anti-symmetric with respect to time-reversal. This implies that the value associated to an observable transition  $\omega \rightarrow \omega'$  must take the negative value if time were to run backward. Thus, in the following we will refer to anti-symmetric observables defined on the edges as *physical observables*.

With the definition of the *negative edge*  $-e := (v', v)$  for any directed edge  $e = (v, v')$ , anti-symmetric observables obey

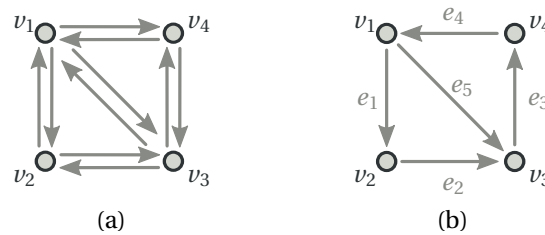
$$\varphi(e) = \varphi_{v'}^v = -\varphi_v^{v'} = \varphi(-e). \quad (5.10)$$

The definition of the negative (*i.e.* inverse) element of a physical observable  $\varphi$  allows for an abstract treatment in the framework of linear algebra.

The *averages* of such quantities have the interpretation of physical currents per unit time. Examples include the current running through a wire in an electrical network, the heat flow in a thermodynamic system, reaction rate of chemical compounds or the velocity of a molecular motor.

However, we can also understand them as the increments of a counting process [HS07]. A jump along an edge  $e = (v, v')$  increases a physical observable by an value  $\varphi(e)$ . Upon a jump along the reverse edge  $-e$ , the variable is decreased by the same amount. The corresponding examples then are the transported charge, the motance interpreted as a basic free energy difference, the change in the number of certain chemical molecules and the distance of a mechanical step, respectively.

### 5.2.2. Algebraic graph theory



**Figure 5.2.:** The relation between bi-directional and oriented graphs. a) In a bi-directional graph  $(V, E^d)$  of a dynamically reversible Markovian jump process, we have that  $e \in E^d$  implies that also  $-e \in E^d$ . b) An oriented graph  $(V, E^o)$  is obtained by arbitrarily choosing one of the directed edges between any two vertices  $v$  and  $v'$ .

Let us formally define the space of all anti-symmetric observables

$$\mathcal{O} := \left\{ \varphi: E^d \rightarrow \mathbb{R} \mid \varphi \text{ is anti-symmetric} \right\}. \quad (5.11)$$

Anti-symmetric observables are completely defined by their values on a set of *oriented*

## 5. The structure of Markov jump processes

edges  $E^0$  obtained by arbitrarily picking an ordered pair  $(v, v')$  for each unordered pair  $\{v, v'\} \in E^u$ . For the dynamically reversible (*i.e.* bi-directional) graphs we are interested in, we have  $E^d = E^0 \cup -E^0$ . The space of anti-symmetric observables  $\mathcal{O}$  is isomorphic to the space of functions defined on the oriented edges  $\mathbb{R}^{E^0}$ . Consequently, we identify them with each other in what follows.

The notion of the negative  $-e$  of an edge  $e$  allows us to treat  $\mathcal{O}$  as a linear (vector) space. If we identify an edge  $e \in E^0$  with its indicator function

$$e: E^0 \rightarrow \mathbb{R},$$

$$e' \mapsto \delta_{e,e'} = \begin{cases} 1 & e = e', \\ 0 & e \neq e', \end{cases}$$

we can use the set of all oriented edges  $E^0$  as a basis for the space of anti-symmetric variables  $\mathcal{O}$ . An element  $\varphi \in \mathcal{O}$  can hence be written as

$$\varphi = \sum_{e_0 \in E^0} \varphi(e_0) e_0 = \sum_{m=1}^M \varphi_m e_m, \quad (5.12)$$

where  $M := |E^0|$  is the number of oriented edges and  $(e_m)_{1 \leq m \leq M}$  assigns some arbitrary enumeration to the oriented edges. This allows for the definition of a bilinear *scalar product* for any two anti-symmetric observables  $\varphi, \psi \in \mathcal{O}$ :

$$\langle \varphi, \psi \rangle := \sum_{m=1}^M [\varphi_m \psi_m]. \quad (5.13)$$

Note that the basis given by  $E^0$  is orthonormal with respect to that scalar product.

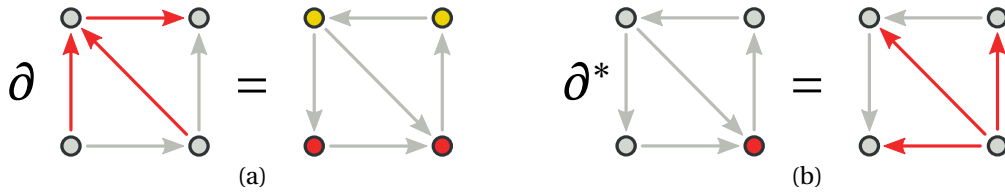
In analogy to our algebraic treatment of the edges, we do the same for the vertices. Again, we identify a vertex  $v \in V$  with its indicator function  $v: V \rightarrow \mathbb{R}; v' \mapsto \delta_{v,v'}$ . Consequently, the space of all real functions on the vertices,  $\mathcal{U} := \mathbb{R}^V$  is identified with the linear space spanned by the  $v \in V$ . Similarly, the scalar product for  $\mathcal{U}$  treats the vertices as an orthonormal basis.

Any linear operator between linear spaces is fully defined by its action on the respective basis elements.<sup>2</sup> Hence, the so-called *boundary operator*

$$\begin{aligned} \partial: E^0 &\rightarrow \mathcal{U} \\ (v, v') &\mapsto v - v' \end{aligned} \quad (5.14)$$

linearly extends to an operator  $\partial: \mathcal{O} \rightarrow \mathcal{U}$ . It is dual (with respect to the natural scalar products on  $\mathcal{O}$  and  $\mathcal{U}$ ) to the *co-boundary operator*  $\partial^*: \mathcal{U} \rightarrow \mathcal{O}$ . For illustrative purposes, in Figure 5.3 we display the action of the boundary and the co-boundary operator on edges and vertices, respectively.

<sup>2</sup> The matrix representation of operators in quantum mechanics is a well-known example.



**Figure 5.3.:** The action of the boundary and the co-boundary exemplified on the graph of Figure 5.2. a) The boundary operator acts on the directed edges marked in red, whose linear combination is an element of the space  $\mathcal{O}$ . b) The co-boundary acts on vertex space  $V$ . An edge or vertex with weight 0 is marked grey, a weight of 1 is indicated red, while a weight of  $-1$  is yellow. Note that a negative weight on an edge is equivalent to a positive weight with reverse orientation.

The definition of the boundary operator allows us to express Kirchhoff's current law and thus the stationarity condition for master equations (5.2) as

$$\partial J = 0.$$

It can be understood as the discrete analogue of the divergence of a vector field. The *cycle space*  $\mathcal{Z}$  contains the *divergence-free* currents:

$$\mathcal{Z} := \ker \partial \subset \mathcal{O}, \quad (5.15)$$

where “ $\ker \partial$ ” denotes the kernel (null-space) of the linear boundary operator  $\partial$ .

The image of the dual co-boundary operator,

$$\mathcal{Y} := \text{im } \partial^* \subset \mathcal{O} \quad (5.16)$$

is called the *co-cycle space* and can be interpreted as the set of *gradient fields*. We will discuss the relation to vector calculus and field theory in more detail in Section 5.4.2.

Duality of  $\partial$  and  $\partial^*$  ensures that the cycles and co-cycles form an orthogonal decomposition of the space of anti-symmetric observables  $\mathcal{O}$ , i.e.  $\mathcal{Y} = \mathcal{Z}^\perp$ . Thus we can decompose the space of anti-symmetric observables as

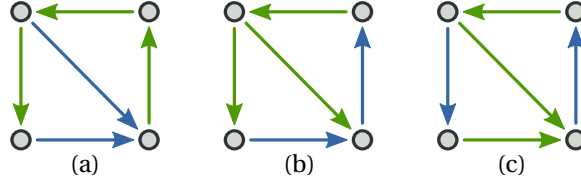
$$\mathcal{Z} \oplus \mathcal{Y} = \mathcal{Z} \oplus \mathcal{Z}^\perp = \mathcal{O}.$$

Consequently, if we pick arbitrary elements  $z \in \mathcal{Z}$  and  $y \in \mathcal{Y}$  of the cycle and co-cycle space, respectively, we have  $\langle z, y \rangle = 0$ .

### 5.2.3. Trees and chords

In the previous subsection, we have defined cycles as elements of the kernel of the boundary operator. This agrees with the picture of a cycle as “something that has no boundary”. However, the usual intuition of a cycle is that of a closed path  $\alpha^{(\tau)} = (\omega_0, \omega_1, \dots, \omega_\tau)$  with  $\omega_0 = \omega_\tau$ . In Section 2.4.2 we have mentioned *fundamental cycles* and that they play an important role in Schnakenberg's network theory for master equations [Sch76]. In this

## 5. The structure of Markov jump processes



**Figure 5.4.:** Three different spanning trees for the graph given in Figure 5.2. The edges  $\mathcal{T}$  of the different trees are marked green, while the chords  $\mathcal{H}$  are depicted blue. Every other spanning tree of the graph, up to symmetries, looks like one of the depicted ones. In the following we will mainly consider the spanning tree 5.4a.

subsection, we will make clear what we mean by a fundamental cycle and how this notion fits into the general algebraic framework.

Fundamental cycles naturally emerge from the topology of the graph. To make that statement precise, we consider *subgraphs*  $G' \subset G$  of a graph  $G$ . We say that a graph  $G' = (V', E')$  is a subgraph of  $G = (V, E)$  if  $V' \subset V$  and  $E' \subset E$ .

A *circuit* is a connected graph  $(V, E^0)$  where the number of edges and vertices is the same, *i.e.*  $|E^0| = |V|$ . Consequently, every vertex has exactly two neighbours. Intuitively, a circuit can be understood as the graph of a discretised circle whose edges are not necessarily all pointing in the same direction.

A *tree* is a connected graph  $(V, \mathcal{T})$  that contains no circuits as subgraphs. It necessarily satisfies  $|\mathcal{T}| = |V| - 1$ . Every connected graph  $G = (V, E^0)$  has a tree  $(V, \mathcal{T})$  as a subgraph which contains all vertices  $V$  of  $G$ . Such a tree is called a *spanning tree* for the graph  $G$ .

For a given graph which itself is not a tree, the choice of the spanning tree is not unique. Figure 5.4 shows different spanning trees of the oriented graph shown in Figure 5.2b. The set of edges of  $G$  that do not belong to a given tree, are called its *chords*  $\eta \in \mathcal{H} = E \setminus \mathcal{T}$ . Adding a chord  $\eta$  to a spanning tree creates the subgraph  $(V, \mathcal{T} \cup \{\eta\})$  which contains exactly one circuit. By aligning the edges on that circuit to point in the same direction as  $\eta$ , we obtain the *fundamental cycle*  $\zeta_\eta$ . The fundamental cycle is a subgraph of the bidirectional graph  $G = (V, E^d)$ .

For the example of the spanning trees depicted in Figure 5.4, the fundamental cycles are shown in Figure 5.5. A fundamental cycle has no boundaries and is thus an element of the cycle space. Note that all fundamental cycles are linearly independent (as they do not share any chords) and their number is  $|\mathcal{H}| = |E^0| - |V| + 1$ . This number is called the *first Betti* or *cyclomatic* number and is a topological constant. It coincides with the dimension of the cycle space  $\mathcal{Z}$ , *cf.* Ref. [Tut98]. Hence, the set of fundamental cycles provide a basis for  $\mathcal{Z}$ .

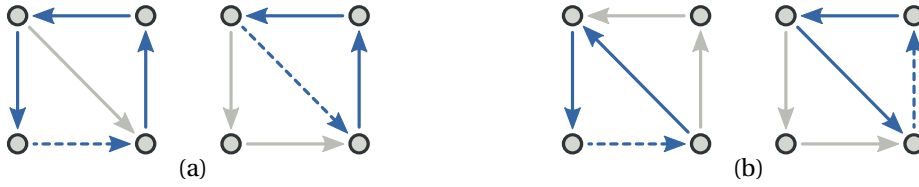
Consequently, every abstract cycle  $z \in \mathcal{Z}$  can be written as a linear combination of fundamental cycles

$$z = \sum_{\eta \in \mathcal{H}} \zeta_\eta z(\eta), \quad (5.17)$$

where  $z(\eta)$  is the value of  $z$  on chord  $\eta$ . The relation (5.17) yields Schnakenberg's *cycle*



## 5.2. Cycles and trees as the fundamental building blocks of networks



**Figure 5.5:** Fundamental cycles of the spanning trees in figure 5.4: every chord  $\eta \in \mathcal{H}$  (dashed here, blue in figure 5.4) generates a fundamental cycle, also marked in blue here. The grey edges and vertices are not part of the fundamental cycles. Note that both spanning trees in figure 5.4b and c generate the exact same fundamental cycles, but with different chords. In contrast, the spanning tree in figure 5.4a shares only one fundamental cycle with the spanning trees b and c.

decomposition for the steady state probability current  $J$  [Sch76]:

$$J = \sum_{\eta \in \mathcal{H}} \zeta_{\eta} J(\eta). \quad (5.18)$$

It implies that the steady state current  $J$  is completely defined by its values  $J(\eta)$  on the chords. Equation (5.17) is the appropriate generalization.

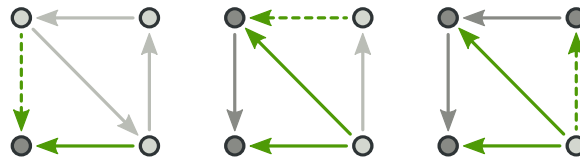
The orthogonal decomposition of  $\mathcal{O}$  into abstract cycles and co-cycles implies that

$$\langle z, y \rangle = 0, \quad \forall z \in \mathcal{Z}, y \in \mathcal{Y}. \quad (5.19)$$

Using the intuition provided by the fundamental cycles, this equation is a generalization of Kirchhoff's second law. In order to appreciate this fact, note that any cycle  $z$  is a linear combination of fundamental co-cycles and thus

$$0 = \langle z, y \rangle = \sum_{\eta \in \mathcal{H}} z(\eta) \langle \zeta(\eta), y \rangle, \quad \forall z \in \mathcal{Z}, y \in \mathcal{Y}. \quad (5.20)$$

Then, any term appearing in the sum is a discrete integral around the edges of the cycle  $\zeta(\eta)$ , i.e.  $\langle \zeta(\eta), y \rangle = \sum_{e \in \zeta(\eta)} y(\eta)$ . Thus, expression (5.19) is just a weighted sum of observables “integrated” around (independent) cycles in the graph.



**Figure 5.6:** Fundamental co-cycles of the spanning tree in figure 5.4a: every edge  $\tau \in \mathcal{T}$  (dashed here, green in figure 5.4) generates a fundamental co-cycle, also marked in green here. The grey edges and vertices are not part of the fundamental co-cycles.

A similar, but slightly less intuitive approach allows us to construct a basis of the co-cycle space  $\mathcal{Y}$ . The removal of an edge  $\tau \in \mathcal{T}$  from a spanning tree yields a disconnected subgraph consisting of exactly two connected components. Note that these components can be as small as a single vertex without any edges. Within the set of edges  $\mathcal{H} \cup \{\tau\}$

## 5. The structure of Markov jump processes

there is a subset of edges containing the edge  $\tau$  together with all other edges that connect these two components. This subset can be understood as the *boundary* between the two components. Reorienting the boundary edges to align with  $\tau$  yields the *fundamental co-cycle* corresponding to edge  $\tau$ . Figure 5.6 illustrates this construction.

### 5.3. Quantification of fluctuations of physical observables

The algebraic structure of physical observables on finite networks is reflected by the stochastic properties of Markov processes. In particular, it allows us to quantify the fluctuations of physical observables in an analytic way. In what follows, a fluctuation means the behaviour of a system away from its average.

Mathematically, the fluctuations of a random variable  $X$  are contained in its probability distribution  $\rho^X(x)$ . For distributions which are strongly peaked around a certain value, its *mean* (or *expectation*)

$$\mathbb{E}(X) := \langle x \rangle = \int x \rho^X(x) dx \quad (5.21)$$

and its *variance*

$$\text{var}(X) := \langle (x - \mathbb{E}(X))^2 \rangle = \int (x - \mathbb{E}(X))^2 \rho^X(x) dx \quad (5.22)$$

are a good characterization of the distribution. The mean defines a typical value and the variance (or its square root, the so-called *standard deviation*) characterize the fluctuations around that value. However, for distributions that have non-negligible probability in the *tails* of the distribution, one requires more information to capture their atypical behaviour.

#### 5.3.1. Cumulants of random variables

The mean and the variance of a distribution are the first two *cumulants* of a distribution. The general definition for an  $\mathbb{R}^d$ -valued random variable  $\mathbf{X} = (X_i)_{i \in \{1, 2, \dots, d\}}$  proceeds via the cumulant-generating function (CGF)

$$g_{\mathbf{X}}: \mathbb{R}^d \rightarrow \mathbb{R},$$

$$\mathbf{q} \mapsto \log \langle \exp(\mathbf{q} \cdot \mathbf{X}) \rangle. \quad (5.23)$$

The *joint cumulants* are obtained as the partial derivatives of the CGF with respect to the components of  $\mathbf{q}$  evaluated at  $\mathbf{q} = \mathbf{0}$ :

$$\kappa(X_{i_1}, X_{i_2}, \dots, X_{i_\ell}) := \frac{\partial}{\partial q_{i_1}} \frac{\partial}{\partial q_{i_2}} \dots \frac{\partial}{\partial q_{i_\ell}} g_{\mathbf{X}}(\mathbf{0}), \quad (5.24)$$

We call the (countable) set of all cumulants the *fluctuation spectrum* of a distribution. If the components  $X_i$  of  $\mathbf{X}$  are independent random variables, the mixed joint cumulants

vanish. Moreover, the cumulants are multi-linear in their arguments. We emphasize that the CGF and thus the fluctuation spectrum contains *all* the information of the original distribution.

The cumulant-generating function  $g_{\mathbf{X}}(\mathbf{q})$  is (non-strictly) convex and satisfies  $g_{\mathbf{X}}(\mathbf{0}) = 0$  for all  $\mathbf{X}$ . If  $\mathbf{X}$  is scalar (*i.e.*  $d = 1$ ), one defines the  $\ell$ th cumulant

$$\kappa(\underbrace{X, \dots, X}_{\ell \text{ times}}) := \kappa_{\ell}(X). \quad (5.25)$$

### 5.3.2. Asymptotic properties in stochastic processes

In Chapter 3 we have formally defined a stochastic process as a sequence of random variables  $(X_t)_{t \in \mathbb{T}}$ . Here, we consider the case of sequences of random vectors  $(\mathbf{X}_t)_{t \in \mathbb{T}}$  and  $\mathbb{T} = \mathbb{R}^+$  or  $\mathbb{T} = \mathbb{N}$ . We are interested in the distribution  $\rho_{\mathbf{X}}^{(\tau)}$  of the time-average

$$\bar{\mathbf{X}}^{(\tau)} := \frac{1}{\tau} \int_{t=0}^{\tau} \mathbf{X}_t dt,$$

where for  $\mathbb{T} = \mathbb{N}$  the integral turns into a sum. More precisely, we are interested in its asymptotic properties for  $\tau \rightarrow \infty$ .

If the individual random vectors  $\mathbf{X}_t = \mathbf{X}$  are independent and identically distributed (i.i.d.) with mean  $\mathbb{E}(X) = \mu$ , the *strong law of large numbers* ensures that

$$\lim_{\tau \rightarrow \infty} \bar{\mathbf{X}}^{(\tau)} \xrightarrow{\text{a.s.}} \mu.$$

This means that fluctuations around the mean value asymptotically vanish. In that case, the *central limit theorem* (CLT) makes this statement more precise, as it specifies an asymptotic distribution. If  $\sigma^2 = \text{var}(X)$  denotes the variance of  $\mathbf{X}$ , one formulation of the CLT states that

$$\lim_{\tau \rightarrow \infty} \sqrt{\tau} \left( \bar{\mathbf{X}}^{(\tau)} - \mu \right) \xrightarrow{\text{in distribution}} \mathcal{N}(0, \sigma^2), \quad (5.26)$$

where  $\mathcal{N}(\mu, \sigma^2)$  is a *normally distributed* random variable with mean  $\mu$  and variance  $\sigma^2$ . A consequence of the CLT is the square-root scaling behaviour of the variance  $\text{var}(\bar{\mathbf{X}}^{(\tau)})$  with  $\tau$ .

Large-deviations theory (LDT) generalizes scaling statements about the distribution of sequences of random variables. A comprehensive introduction to the topic can be found in Ref. [Tou11]. The standard textbook on the subject and its relation to (classical) statistical physics is the one by Ellis, *cf.* Ref. [Ell05]. The relation is discussed in even more detail in the review article [Tou09].

Consider a sequence of random vectors  $(\bar{\mathbf{X}}^{(\tau)})_{t \in \mathbb{T}}$ . We say that this sequence obeys a *Large deviation principle* if the limit

$$I_{\mathbf{X}}(\mathbf{x}) := - \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \log \left[ \rho_{\mathbf{X}}^{(\tau)}(\mathbf{x}) \right]. \quad (5.27)$$

## 5. The structure of Markov jump processes

exists (at least in an open neighbourhood of  $\mathbb{R}^d$  around the origin). The function  $I_X(\mathbf{x})$  is called *rate function*. If a large deviations principle holds, the probability density  $\rho_X^{(\tau)}$  obeys the scaling

$$\rho_X^{(\tau)}(\mathbf{x}) \propto \exp[-\tau I_X(\mathbf{x}) + R(\tau, \mathbf{x})], \quad (5.28)$$

where  $R(\tau, \mathbf{x}) \sim o(\tau)$  scales sub-linearly with  $\tau$ . Hence, for large  $\tau$  we can approximately write  $\rho_X^{(\tau)}(\mathbf{x}) \propto \exp[-\tau I_X(\mathbf{x})]$ .

For the case of i.i.d. random variables, we have seen that asymptotically the variance, *i.e.* its second cumulant vanishes. In that case, the CLT (5.26) implies the same for all higher cumulants. In fact, this is generally true for random variables that obey a large deviation principle.

Hence, the fluctuation spectrum for such random variables is not useful in the asymptotic limit  $\tau \rightarrow \infty$ . However, the scaling relation (5.27) suggests the definition of the *scaled cumulant-generating function* (SCGF):

$$\lambda_X(\mathbf{q}) := \lim_{\tau \rightarrow \infty} \frac{1}{\tau} g_{\bar{\mathbf{X}}^{(\tau)}}(\tau \mathbf{q}) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \log \left\langle \exp \left[ \tau \mathbf{q} \cdot \bar{\mathbf{X}}^{(\tau)} \right] \right\rangle. \quad (5.29)$$

where again  $\mathbf{q} \in \mathbb{R}^d$ . Like the normal cumulant-generating function, the SCGF is convex and  $\lambda_X(\mathbf{0}) = 0$  for every family of random vectors. The cumulant-generating function  $g_X(\mathbf{q})$  of a random variable  $X$  is always a smooth function. However, the scaled limit  $\lambda_X(\mathbf{q})$  is not necessarily differentiable any more.

If  $\lambda_X(\mathbf{q})$  is differentiable on  $\mathbb{R}^d$ , the Gärtner–Ellis Theorem then assures that the sequence  $\left( \bar{\mathbf{X}}^{(\tau)} \right)_{\tau \in \mathbb{T}}$  satisfies a Large Deviation Principle [Ell05]. Moreover, its rate function

$$I_X(\mathbf{x}) = \mathbf{x} \cdot \mathbf{q}(\mathbf{x}) - \lambda_X(\mathbf{q}(\mathbf{x})) \quad (5.30)$$

is obtained as the Legendre transform of  $\lambda$ . Note that the functional dependence  $\mathbf{q}(\mathbf{x})$  is given by inverting the relation  $\mathbf{x} = \nabla \lambda_X(\mathbf{q})$ .

In that case, we define the *scaled fluctuation spectrum* as the set of *scaled cumulants*:

$$c(X_{i_1}, X_{i_2}, \dots, X_{i_\ell}) := \frac{\partial}{\partial q_{i_1}} \frac{\partial}{\partial q_{i_2}} \dots \frac{\partial}{\partial q_{i_\ell}} \lambda_X(\mathbf{0}). \quad (5.31)$$

The scaled cumulants directly inherit multi-linearity from the cumulants. From the definition in equation (5.29) we immediately infer the scaling of the cumulants to be

$$c(X_{i_1}, X_{i_2}, \dots, X_{i_\ell}) = \lim_{\tau \rightarrow \infty} \tau^{\ell-1} \kappa(X_{i_1}^{(\tau)}, X_{i_2}^{(\tau)}, \dots, X_{i_\ell}^{(\tau)}). \quad (5.32)$$

So far we dealt with general sequences of random vectors  $\bar{\mathbf{X}}^{(\tau)}$ . In the next section we will explicitly consider the sequence of time-averages of increasing length. If these time-averages are obtained for physical observables in a Markov process, we have analytical access to the SCGF. Henceforth, we are always interested in the characterization of

the asymptotic fluctuations. Hence, we omit the word “scaled” and refer simply to the *fluctuation spectrum* of an observable.

### 5.3.3. Large deviation theory of Markovian jump processes

In this section we consider the time-averages of the physical observables for Markov processes in continuous time. In Section 5.2.1 we introduced the notion of a physical observable  $\varphi \in \mathcal{O}$  as an anti-symmetric variable defined on the edges of a graph:

$$\begin{aligned} \varphi: E &\rightarrow \mathbb{R} \\ (\omega, \omega') &\mapsto \varphi_{\omega'}^{\omega} \end{aligned}$$

For an ergodic Markov chain on a finite state space we define the time-average as

$$\bar{\varphi}^{(\tau)} := \frac{1}{\tau} \sum_{n=1}^{N(\tau)} \varphi_{\omega_n}^{\omega_{n-1}}. \quad (5.33)$$

The sum is over the states  $(\omega_n)_{n \in \{0,1,\dots,n(\tau)\}}$  that occur in the trajectory  $\underline{\omega}^{(\tau)}$  for the interval  $[0, \tau]$ . For dynamics in continuous time the number of jumps  $n(\tau)$  is itself a random variable. The distribution of these time averages always satisfies a Large deviation principle [Tou11]. Moreover, the scaled cumulant-generating function  $\lambda_{\varphi}(q)$  is differentiable and can be calculated as an eigenvalue of an appropriately defined matrix.

To be concrete, we consider the multi-variate case where  $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_d) \in \mathcal{O}^d$  is a  $d$ -tuple of anti-symmetric observables. Its component-wise time average (5.33) is a family of random vectors  $\bar{\boldsymbol{\varphi}}^{(\tau)}$  taking values in  $\mathbb{R}^d$ . The multivariate SCGF  $\lambda_{\boldsymbol{\varphi}}(\mathbf{q})$  is given by the unique dominant eigenvalue of the *tilted matrix*  $\mathbb{W}_{\boldsymbol{\varphi}}(\mathbf{q})$  with components [Tou11]

$$(\mathbb{W}_{\boldsymbol{\varphi}}(\mathbf{q}))_j^i := w_j^i \exp(\mathbf{q} \cdot \boldsymbol{\varphi}_j^i) = w_j^i \exp\left(\sum_{\ell=1}^d q_{\ell} \varphi_{\ell}(e_j^i)\right). \quad (5.34)$$

The expression  $\varphi_{\ell}(e_j^i)$  stands for the value of the  $\ell$ th random variable in  $\boldsymbol{\varphi}$  on the edge  $e_j^i$ , along which transitions occur at rate  $w_j^i$ .

The scaled cumulants are obtained as the partial derivatives of the dominant eigenvalue. Generally, the analytic calculation of the dominant eigenvalue of a big matrix is difficult. At this point one usually applies numerical algorithms to find (approximations) to  $\lambda_{\boldsymbol{\varphi}}(\mathbf{q})$  [Tou11]. Thus, in general it is difficult to find the large-deviation function  $I_{\boldsymbol{\varphi}}(\mathbf{x})$ .

The key result of the present chapter is an easy method for the determination of the fluctuation spectrum of *any* physical observable  $\boldsymbol{\varphi}$  — without explicitly referring to  $I_{\boldsymbol{\varphi}}(\mathbf{x})$ . The crucial point is to realize that all of the relevant information is already contained in the coefficients  $a_i(\mathbf{q})$  of the characteristic polynomial

$$\chi_{\mathbb{W}_{\boldsymbol{\varphi}}(\mathbf{q})}(\lambda) := \sum_{i=0}^N a_i(\mathbf{q}) \lambda^i := \det(\lambda \mathbb{1} - \mathbb{W}_{\boldsymbol{\varphi}}(\mathbf{q})). \quad (5.35)$$

## 5. The structure of Markov jump processes

The roots of the characteristic polynomial are the eigenvalues  $\lambda$  of the tilted matrix  $\mathbb{W}_\varphi(\mathbf{q})$ . The characteristic equation

$$\chi_{\mathbb{W}_\varphi(\mathbf{q})}(\lambda) \stackrel{!}{=} 0$$

contains all the information about the largest eigenvalue  $\lambda_\varphi(\mathbf{q})$  of the tilted matrix. Since it is the cumulant-generating function, it fulfils  $\lambda_\varphi(\mathbf{0}) = 0$ . This relation uniquely determines the *solution branch* of Eq. (5.3.3) which gives the  $\mathbf{q}$ -dependence of  $\lambda_\varphi(\mathbf{q})$ . Thus, we can use the Implicit Function Theorem to extract the scaled cumulants iteratively from the coefficients  $a_i(\mathbf{q})$  of  $\chi_{\mathbb{W}_\varphi(\mathbf{q})}$ . Their  $\mathbf{q}$ -dependence is directly analytically accessible from the determinant formula for matrices.

The most important case is that of a single current-like observable, where both  $\bar{\varphi}^{(\tau)}$  and  $q$  are scalar. Its first two scaled cumulants read

$$c_1(\varphi) = -\frac{a'_0}{a_1}, \quad (5.36a)$$

$$c_2(\varphi) = \frac{-2 a_2 c_1^2(\varphi) - 2 c_1(\varphi) a'_1 - a''_0}{a_1} = 2 \frac{a'_1 a'_0}{a_1^2} - 2 \frac{a_2 (a'_0)^2}{a_1^3} - \frac{a''_0}{a_1}, \quad (5.36b)$$

where all of the  $a_i$  and their derivatives must be evaluated at  $q = 0$ . Higher cumulants can be calculated iteratively. For a more detailed account on the analytical calculation of the fluctuation spectrum we refer to the M.Sc. thesis [Wac13] and the pre-print [WVA14].

Let us stress the following fact: The only ingredients needed for the calculation of the asymptotic fluctuation spectrum are the transition matrix  $\mathbb{W}$  and the observable  $\varphi$ . At no point we explicitly need the (steady-state) distribution  $\mathbf{p}_\infty$  of the Markov chain. In the previous chapter, we have seen that the latter is completely determined by the currents on a set of fundamental chords. Moreover, we have seen how a physical observable can be decomposed into fundamental cycles. The next subsection deals with an important consequence of this fact for the fluctuation spectra of time-averages  $\bar{\varphi}^{(\tau)}$  taken for physical observables  $\varphi \in \mathcal{O}$ .

### 5.3.4. Cycles and fluctuations of physical observables

In this subsection, we clarify how the topological structure of the Markovian jump process influences the fluctuation spectrum of scaled cumulants. The results in this section have mainly been developed by A. Wachtel. We state them without proof and refer the reader to Ref. [WVA14].

In Section 5.2.2 of the previous chapter we have introduced algebraic concept of a cycle  $z \in \mathcal{Z}$ . Remember that we identified the space of anti-symmetric observables  $\mathcal{O}$  with the vector space  $\mathbb{R}^{E^0}$ , where  $E^0$  was a set of oriented edges of a bi-directional graph. In that abstract sense, a cycle was identified with an element of the subspace  $\mathcal{Z} \subset \mathcal{O}$  defined as the kernel of the boundary operator  $\partial$ . The orthogonal complement of  $\mathcal{Z}$  in  $\mathcal{O}$  is the co-cycle space  $\mathcal{Y}$ . This orthogonal decomposition of  $\mathcal{O}$  guarantees that any anti-symmetric

### 5.3. Quantification of fluctuations of physical observables

observable can be written as a unique sum  $\varphi = z + y$  of a cycle  $z \in \mathcal{Z}$  and a co-cycle  $y \in \mathcal{Y}$ .

Because the (scaled) cumulants are multi-linear, we can calculate the fluctuation spectrum of  $\varphi$  from the fluctuation spectrum of  $z$  and  $y$ . By ergodicity of the Markov chain, we have  $\lim_{\tau \rightarrow \infty} \varphi^{(\tau)} \xrightarrow{\text{a.s.}} \langle \varphi \rangle_\infty$ . As the first scaled cumulant agrees with the first cumulant of  $\varphi$ , it can be written as the scalar product of  $\varphi$  with the steady state current  $J$ :

$$\kappa_1(\varphi) = c_1(\varphi) = \mathbb{E}(\varphi) = \langle \varphi \rangle_\infty = \langle \varphi, J \rangle.$$

We already know that for a co-cycle  $y \in \mathcal{Y}$ , Kirchhoff's second law holds, *i.e.*

$$\kappa_1(y) = \langle y, J \rangle = 0 \tag{5.37}$$

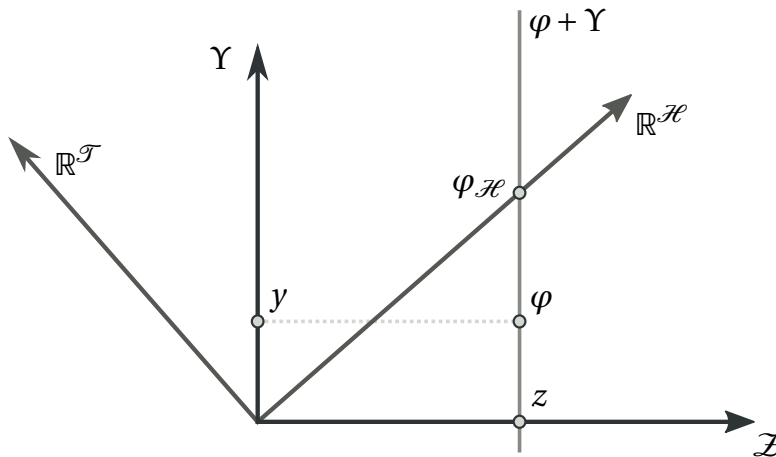
The following proposition thus generalizes Kirchhoff's second law to the entire fluctuation spectrum:

**5.1 Proposition** *Let  $y \in \mathcal{Y}$  be a co-cycle. Then its scaled cumulant-generating function  $\lambda_y(q) \equiv 0$  vanishes. Thus, all of its scaled cumulants vanish.*

Andrieux and Gaspard [AG07] proved a special case of this result. Similar considerations can be found in Ref. [FDP11]. The general proof has the following corollary [Wac13; WVA14]:

**5.2 Corollary** *The scaled cumulant-generating function  $\lambda_\varphi(q)$  of a current-like observable  $\varphi \in \mathcal{O}$  satisfies  $\lambda_\varphi(q) = \lambda_z(q)$  where  $z \in \mathcal{Z}$  is the unique cycle part of  $\varphi$ .*

In other words, the SCGF and thus the rate function of all the observables in the subspace  $\varphi + \mathcal{Y} \subset \mathcal{O}$  agree with that of both  $\varphi$  and its cycle part  $z$ . Regarding the asymptotic fluctuations, this can be understood as a form of gauge invariance. Consequently, for calculating  $\lambda_\varphi$  we can use any representative within the class  $\varphi + \mathcal{Y}$ .



**Figure 5.7.:** Geometrical interpretation of the chord representation: Projecting  $\varphi$  in parallel to  $\mathcal{Y}$  onto the chord space  $\mathbb{R}^{\mathcal{H}}$  yields  $\varphi_{\mathcal{H}}$ . For more worked examples and additional theory, cf. the pre-print [Pol14].

## 5. The structure of Markov jump processes

Although we have seen that both  $\varphi$  and its cyclic part  $z$  have the same fluctuation spectra, neither choice is necessarily the most convenient one for the purpose of analytical calculations. A good candidate in that respect is the *chord representation* of an observable  $\varphi$ ,

$$\varphi_{\mathcal{H}} := \sum_{\eta \in \mathcal{H}} \langle \varphi, \zeta_\eta \rangle \eta, \quad (5.38)$$

where  $\zeta_\eta \in \mathcal{Z}$  is the fundamental cycle corresponding to the chord  $\eta \in \mathcal{H}$ . The numbers  $\langle \varphi, \zeta_\eta \rangle$  appearing in the chord representation are obtained as the components of oblique projections of  $\varphi$  onto the space  $\mathbb{R}^{\mathcal{H}}$  spanned by the chords. Oblique projections, are not necessarily, but rather project parallel to a linear subspace as depicted in Figure 5.7. In our case, the oblique projection is parallel to  $\Upsilon$ , cf. also reference [Pol14].

The chord representation  $\varphi_{\mathcal{H}}$  of  $\varphi$  is the unique element in the intersection  $\mathbb{R}^{\mathcal{H}} \cap \varphi + \Upsilon$ . By Corollary 5.2, the scaled cumulant-generating functions of  $\varphi$  and  $\varphi_{\mathcal{H}}$  agree. Moreover, the chord representation  $\varphi_{\mathcal{H}}$  is supported on at most  $b = |\mathcal{H}| = |E| - |V| + 1$  edges, where  $\varphi$  generically takes non-trivial values on all edges.

This in turn may reduce the effort to calculate the scaled cumulants: Denote the component of the oblique projection onto  $\eta$  by  $\varphi_\eta := \langle \varphi, \zeta_\eta \rangle$ . We summarize them in the  $b$ -dimensional vector  $\mathbf{\Phi} = (\varphi_{\eta_1}, \varphi_{\eta_2}, \dots, \varphi_{\eta_b})^\top$ , where  $b = \dim \mathcal{Z}$  is also the dimension of the cycle space. Writing also  $\mathbf{H} = (\eta_1, \eta_2, \dots, \eta_b)^\top$ , the chord representation can be thought of as the matrix product  $\varphi_{\mathcal{H}} = \mathbf{\Phi}^\top \mathbf{H}$ . Note that this matrix product formally looks like a scalar product  $\langle \cdot, \cdot \rangle$ , but its value is an element of  $\mathbb{R}^{\mathcal{H}} \cap \varphi + \Upsilon$ .

The scaled cumulant-generating functions satisfy  $\lambda_\varphi(q) = \lambda_{\varphi_{\mathcal{H}}}(q) = \lambda_{\mathbf{H}(\mathbf{\Phi}q)}$ . Consequently, in order to determine the scaled cumulants  $c_1(\varphi), c_2(\varphi), \dots, c_\ell(\varphi)$  of  $\varphi$ , it is sufficient to calculate the vector  $\mathbf{\Phi}$  and the joint scaled cumulants  $c(\eta_{i_1}, \dots, \eta_{i_\ell})$  of the chords up to the order  $\ell$ . Due to multi-linearity, the scaled cumulants of  $\varphi$  read

$$c_\ell(\varphi) = \left( \prod_{i=1}^{\ell} \sum_{\eta_i \in \mathcal{H}} \varphi_{\eta_i} \right) c(\eta_1, \eta_2, \dots, \eta_\ell) \quad (5.39)$$

such that

$$c_1(\varphi) = \sum_{\eta \in \mathcal{H}} \varphi_\eta c_1(\eta) = \sum_{\eta \in \mathcal{H}} \varphi_\eta J(\eta), \quad (5.40a)$$

$$c_2(\varphi) = \sum_{\eta_1 \in \mathcal{H}} \sum_{\eta_2 \in \mathcal{H}} \varphi_{\eta_1} \varphi_{\eta_2} c(\eta_1, \eta_2). \quad (5.40b)$$

For the first cumulant — the expectation value — the representation in equation (5.40a) is called the *Schnakenberg decomposition* [Sch76]. Equation (5.39) generalizes this decomposition to the entire fluctuation spectrum.



## 5.4. Discussion

In the discussion of the present chapter, we focus on three aspects. Firstly, we comment on other cycle decompositions that have been used in the literature, and how they relate to the Schnakenberg decomposition (5.18). Secondly, we give an intuition about the role of the boundary and co-boundary operator. Finally, we review the significance of our main results for ST.

### 5.4.1. Different cycle decompositions and the flux-cycle transform

The decomposition of the steady-state current  $J_{\omega'}^{\omega}$  into its fundamental cycles using Eq. (5.17) is just one way of a cycle decomposition. Different variants have been proposed in the literature [Sch76; Kal93; Kal06; JQQ04].

**Decompositions using fundamental cycles** The Schnakenberg decomposition of the current (5.18) is a special case of the general decomposition using fundamental cycles. It has been used by several authors in the context of Markov processes [AG07; FDP11]. Algorithms have been proposed for the efficient construction of fundamental cycles in (large) random graphs [Pat69]. A recent preprint [Pol14] reviews the relevance of (oblique) projections in the context of ST. It also discusses the duality relation between the boundary and co-boundary operators  $\partial$  and  $\partial^*$  in more detail. An explicit method for the construction of the steady state currents is reviewed in Ref. [ZS07].

In summary, we can say that the ideas pioneered by Kirchhoff [Kir47] are still relevant nowadays in both abstract mathematics and theoretical physics. In mathematics, the result is usually known under the name of the matrix-tree or matrix-forest theorem [Tut98]. The basis for the Schnakenberg decomposition which uses the notion of a set of fundamental cycles was laid by Hill [Hil77]. In addition, Hill has also considered another set of cycles which we briefly review in the next paragraph.

**Hill's stochastic cycles** Kalpazidou [Kal06] and Jiang *et. al.* [JQQ04] have reviewed the significance of cycles for Markov processes. In particular, they present several methods for a decomposition of the network of states using cycles. These authors distinguish between stochastic and deterministic decomposition algorithms.

The extended set of cycles used by Hill includes all distinct closed paths formed by a random walk on the network of states [Hil77]. This allows for the treatment of networks with absorbing states. Hill emphasized that for biological processes, the knowledge of which cycles are completed how many times before an absorbing state is reached is important. As an example consider a model for an enzyme which includes the process of the enzyme's degradation as an absorbing state. The number of completed cycles than quantifies the biological activity during an enzyme's life time. Moreover, the stochastic dynamics of cycle completion can be formalized as an equivalent Markov process on a transformed state space which has the cycles as its vertices, *cf.* Refs. [JQQ04; Kal06].

## 5. The structure of Markov jump processes

**Flux cycles and the cycle transformation** In Ref. [Alt+12] we introduced another cycle decomposition, complementary to both Hill's and Schnakenberg's cycles. Rather than being based on the steady-state currents  $J$ , it is based on the steady state fluxes  $\phi$ . Consequently, we call it the *flux-cycle decomposition*. The number of flux cycles in this decomposition is generally larger than the number of fundamental cycles (*i.e.* the cyclomatic number) but smaller than the number of all distinct closed paths.

Similarly to the latter approach, the cycles in the flux-cycle decomposition for a graph  $G$  can be viewed as the vertices of a transformed graph  $G'$ . The dynamics on the new graph  $G'$  are those of an equilibrium system, *i.e.* a system where detailed balance holds and no currents are flowing. A set of flux cycles can be constructed using the algorithm published in Ref. [Alt+12], where also consider an analogy to networks of public transport.

In Ref. [Alt+12] we also discuss how changes of weights assigned by the algorithm correspond to the change of the dominant paths in a network. We illustrate this fact using the example of a periodic *totally asymmetric exclusion process* (TASEP) with two particles on four sites. The change in the decomposition amounts to the change of the cycles that dominate the steady state upon changing an external driving. For the TASEP example, it results in the change of “gait” observed in the movement of the two particles. In the context of biological systems the notion of dominant cycles has been discussed in Ref. [Hil79].

### 5.4.2. Analogy to field theory

Let us discuss the role of the boundary and co-boundary operators  $\partial: \mathcal{C} \rightarrow \mathcal{U}$  and  $\partial^*: \mathcal{U} \rightarrow \mathcal{C}$  in physical terms. They are maps between the *edge space*  $\mathcal{C} \sim \mathbb{R}^{E^0}$  and the *vertex space*  $\mathcal{U} \sim \mathbb{R}^V$ . In an analogy to continuous field theory, these spaces correspond to the notion of *vector* and *scalar* fields.

In a dynamic field theory like hydrodynamics, one distinguishes between currents of conserved and non-conserved quantities. Let  $\vec{z}$  denote the current of a conserved quantity. Then, the continuity equation for the steady state reads

$$\nabla \cdot \vec{z} = 0. \quad (5.41)$$

If we interpret the boundary operator  $\partial$  as a discrete divergence, the discrete analogue of Eq. (5.41) is the definition of the elements of the cycle space Eq. (5.15). Hence, the cycle space contains the discrete analogue of *currents of conserved quantities*, which amount to *divergence-free fields* in a steady state.

Similarly, the co-boundary operator is a discrete version of the vector gradient. In field theory, the gradient acts on scalar fields  $U$  to produce irrotational *gradient fields*

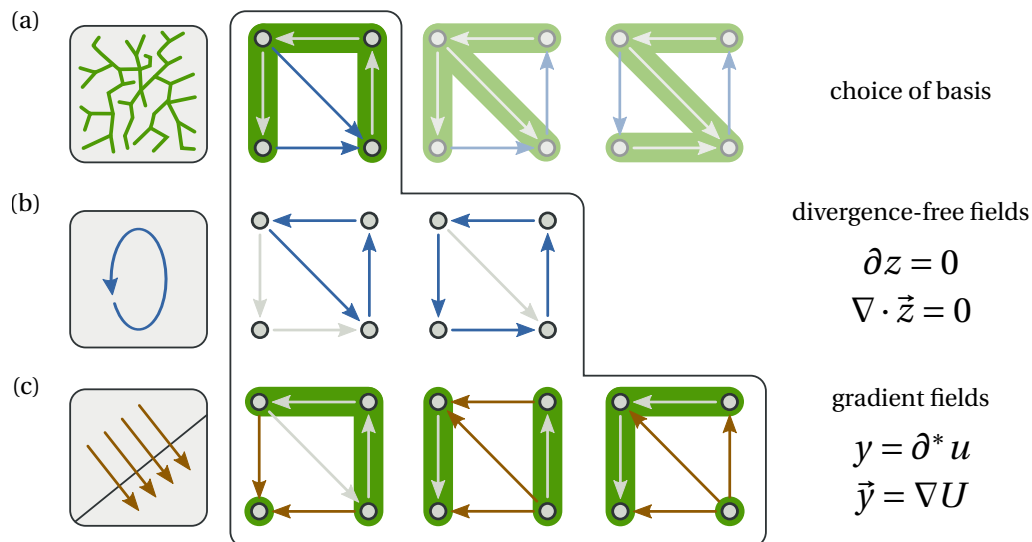
$$\vec{y} := \nabla U. \quad (5.42)$$

Again, the above equation is the continuous version of the definition of the *co-cycle space*

in Eq. (5.16). This is why co-cycles take the role of potential differences. Kirchhoff's second law (5.5) expresses this fact for the difference  $\Delta U_\omega^w$  of the reference voltage  $U_\omega = -\log p_\infty(\omega)$  associated to a vertex.

Finally, the duality of  $\partial$  and  $\partial^*$  ensures that the cycles and co-cycles form an orthogonal decomposition of the space of anti-symmetric observables  $\mathcal{O}$ . Consequently, we can write  $\varphi = z + y$  with  $z \in \mathcal{Z}$  and  $y \in \mathcal{Y}$  for any  $\varphi \in \mathcal{O}$ . This decomposition is the discrete analogue of the Helmholtz decomposition for vector fields in  $\mathbb{R}^3$ , which can be found in any standard textbook on fluid mechanics, e.g. Ref. [Ach90].

Figure 5.8 illustrates these analogies. It also motivates the interpretation of co-cycles as *tidal currents* which was introduced by Poletti [Pol14]. From that point of view, fluxes can be characterized by their cyclic contribution as well as their tendency to “flood” across a boundary, cf. Fig. 5.8c). Note that generically a current  $\varphi \in \mathcal{O}$  has both tidal and cyclic contributions. Yet, only the latter are important for the (asymptotic) characterization of fluctuations in non-equilibrium steady states.



**Figure 5.8:** Trees, cycles and co-cycles for discrete graphs. a) A spanning tree (green) determines the basis of physical, i.e. anti-symmetric, observables. The corresponding chords are depicted in blue. Alternative choices of a spanning tree (semi-transparent) yield different chords and a different set of basis vectors. b) The fundamental cycles constitute a basis of the cycle space. In the analogy of vector calculus, cycles  $z \in \mathcal{Z}$  are divergence-free vector fields expressing the stationary currents of conserved quantities. c) The co-cycles  $y \in \mathcal{Y}$  can be visualized as tidal currents across a boundary (brown). They correspond to gradient fields in vector calculus.

### 5.4.3. Relevance for stochastic thermodynamics

Finally, we discuss the relevance of our results for ST. More precisely, we discuss how they provide useful tools for the study of dynamically reversible Markov processes in general. Consequently, here we do not refer to any particular *model* or a particular (physical) observable. We will come back to these considerations in much more detail

## 5. The structure of Markov jump processes

in the following Chapter 6. There, we apply our general results to models of molecular motors in non-equilibrium environments.

Before we start the discussion let us emphasize once more that our methods are completely *analytical*. They can be implemented in any computer algebra system to yield *exact results*.

### Access to the steady-state distribution

None of the results presented in the present chapter requires access to the steady state distribution of a Markov process. Usually, the steady state distribution  $\mathbf{p}_\infty$  is obtained as the left eigenvalue of  $\mathbb{W}$  corresponding to the zero eigenvalue. As the eigenvalue is known, the determination of  $\mathbf{p}_\infty$  can be obtained by solving a set of linear equations. The fact that the stationary distribution is accessible using combinatorial aspects of graphs was already known to Hill [Hil77]. A nice presentation of the method can be found in Ref. [ZS07].

For the purpose of *symbolic* calculations in computer algebra systems, there is another elegant way to obtain the stationary distribution: Consider the random variable characterizing the fraction of time the system stays in state  $v_k$ . Its SCGF is obtained as the dominant eigenvalue of the matrix  $\mathbb{W}_{v_k}(q)$  with entries

$$(\mathbb{W}_{v_k}(q))_j^i := w_j^i \exp(q\delta_{jk}). \quad (5.43)$$

The first cumulant of the staying time distribution characterizes the average occupancy of state  $v_k$  in the limit of large times. For an ergodic system, it thus amounts to the steady state probability  $p_\infty(k)$ . This result holds in general, *i.e.* not only for dynamically reversible Markov processes [Tou11].

The essence of our approach applies here as well: We do not need to calculate the largest eigenvalue of the tilted matrix. Rather all necessary information is contained in the characteristic polynomial  $\chi_k(\lambda, q)$  of  $\mathbb{W}_{v_k}$ . Equation (5.36a) then reduces to

$$p_\infty(k) = c_1(v_k) = -\frac{a'_0}{a_1} \equiv -\frac{\frac{\partial \chi_k}{\partial q}}{\frac{\partial \chi_k}{\partial \lambda}} \Bigg|_{\lambda=q=0}. \quad (5.44)$$

### Benefits of the chord representation

The chord representation  $\varphi_{\mathcal{H}}$  of an observable  $\varphi$  takes only non-vanishing values on the chords. That means, it is supported by at most  $|\mathcal{H}| = |E^0| - |V| + 1$  edges, which is the number of independent cycles. For any observable supported on more than  $|\mathcal{H}|$  edges, the chord representation might be more suitable for the purpose of calculation. The value of  $\varphi_{\mathcal{H}}$  on a chord  $\eta$  is  $\langle \varphi, \zeta_\eta \rangle$ , which is simply the integral of the observable along the fundamental cycle  $\zeta_\eta$ .

Moreover, another simplification for the purpose of calculation can be obtained by a suitable choice of spanning tree. Assume we are only interested in the fluctuation

spectrum of certain observables, of which we know that they are conserved along some particular cycles of the network. In that case, it is convenient to choose a spanning tree such that these cycles are elements of the basis of fundamental cycles. This choice of basis ensures that in the chord representation of the observables the corresponding chords do not show up. In turn, this reduces the number of terms appearing in the general expression (5.40a) for  $\ell$ th joint scaled cumulant.

## 5.5. Summary

In the present chapter we discussed the algebraic and topological structure of Markovian dynamics on finite networks. As a motivation, we considered Kirchhoff's laws for electrical circuits and gave a complete analogy between the latter and the steady state of Markov processes. On the abstract level, Kirchhoff's laws are formulated using an abstract boundary operator  $\partial$  and its dual  $\partial^*$ . These two latter operators can be understood as the discrete version of the vector divergence and gradient for vector and scalar fields, respectively.

Physical observables  $\varphi \in \mathcal{O}$  are the analogue of the vector fields representing currents in field theory. In field theory, steady state currents of conserved quantities are divergence free. In analogy to the Helmholtz composition from fluid mechanics, one can always write  $\varphi = z + y$ , where  $z \in \mathcal{Z}$  is its divergence-free contribution. The contribution  $y \in \mathcal{Y} = \mathcal{Z}^\perp$  is called the co-cyclic contribution to  $\varphi$ .

The main part of the present chapter discussed the quantification of fluctuations of physical observables. We found that the fluctuation spectrum, and equivalently the rate function  $I_\varphi$  of a physical observable  $\varphi \in \mathcal{O}$  only depends on its cyclic part  $z$ . Hence, with respect to their fluctuations, two observables  $\varphi = z + y$  and  $\varphi' = z + y'$  are equivalent, if their cyclic contribution  $z$  agrees. Consequently, we can choose  $\varphi \in z + \mathcal{Y}$  such that it is most convenient for the purpose of calculation. In that context, we present a generalization of the Schnakenberg decomposition for arbitrary scaled cumulants of arbitrary observables.

A corollary follows from the multi-linearity of cumulants. It states that in order to obtain the fluctuation spectrum of an arbitrary physical observable  $\varphi$ , it is enough to know the fluctuation spectrum of the currents  $J$  on a subset of edges.

The results in this chapter are both analytic and constructive. Rather than having to solve an eigenvalue problem to obtain the fluctuation spectrum, it suffices to differentiate the coefficients of a polynomial. Our method thus neither requires advanced combinatorics nor the solution of any linear equation or an eigenvalue problem.

In the following Chapter 6, we demonstrate the power of our methods on models of ST. More precisely, we study several models for the molecular motor kinesin. Our methods allow us to analytically obtain a phase diagram for the operation modes of the motor in dependence of its chemo-mechanical driving forces. Moreover, our methods allow a much more detailed look at the structure of the phase diagram than what was possible before.



## 6. Modelling molecular motors

“ [Can] a physicist who [has learnt] the statistical foundation of his science [and] begins to think about organisms and about the way they behave [make] any relevant contributions? [...] It will turn out that he can. ”

---

E. Schrödinger, *What is Life?*, 1944

### What is this about?

In small systems, fluctuations are relevant and have recently attracted thorough attention from scientists in various fields. The molecular machinery of life, *i.e.* the biological macromolecules responsible for any metabolic activity in living cells has arguably been at the centre of this research. *In vivo*, these systems are always found in a chemical environment far from equilibrium provided by the cell's cytosol. In recent years, scientists have also performed *in vitro* experiments on these systems, thus opening the door for a quantitative treatment.

The initial quote by Schrödinger is about the role of theoretical physics for the study of living systems. He affirms that statistical physics can yield valuable contributions to the study of the arguably most complex system — life. He himself provides one of the best examples of such a contribution. In his essay “What is life?” Schrödinger used his concepts from statistical mechanics in order to discuss the structure of the hereditary substance [Sch92]. His conclusion was that the latter must have the form of an “aperiodic crystal” — ten years before Watson and Crick unveiled the structure of DNA.

Another contribution from physics to biology comes in the form of simplified models for complex systems. The models for small systems formulated within the framework of ST is just one example. In his pioneering contributions, Hill already focused on models of the molecular machinery of life [Hil77]. At that time, however, fluctuations in those systems were not yet accessible by measurements.

More recent experiments provide ways to compare model predictions with actual data. Hence, a good model must yield more than only a correct account for the average behaviour of a system. Additionally, it must capture the behaviour of the random fluctuating trajectories observed in single realizations of an experiment.

For biological systems, which have evolved for millions of years, this is particularly important. In the light of evolution [Dob73], it is fair to assume that fluctuation may play

## 6. Modelling molecular motors

a crucial role for the *function* of living systems. Consequently, we start the present chapter with a discussion of functional fluctuations.

After that, we focus on the fluctuations arising in the model of one of the most well-studied small biological systems, namely the molecular motor *kinesin*. Using the tools we have established throughout the previous Chapter 5, we perform a quantitative study of the fluctuating properties in various models for kinesin. In particular we give a detailed account on kinesin's phase diagram, which is spanned by the external chemical and mechanical forces driving the motor.

We are further interested in the comparison of more complicated models with both simpler *reduced* models and *minimal* models obtained from first principles. Simplified models can be obtained via a coarse-graining procedure. Minimal models are more fundamental and in a sense maximally non-committal with respect to the physical information we have about a system: The topological properties of a physically motivated network of states, the thermodynamic forces exerted by the medium and observable physical currents.

Using a classical model for kinesin [LL07] as an example, we find that the simpler models are extremely good in capturing the fluctuating behaviour of the original model — for values ranging over more than twenty logarithmic decades. We end with a discussion of the visible and hidden structure in phase-diagrams. In particular, we discuss the benefits of considering a certain signal-to-noise ratio, which has a direct interpretation in the theory of non-linear response.

Some of the content of the present chapter is discussed in the pre-print of a manuscript prepared by A. Wachtel, J. Vollmer and the present author [WVA14]. The non-linear response theory of ST is the topic of another publication in preparation by the same authors [WAV p]. The plots of kinesin's phase diagram were compiled by A. Wachtel using *Mathematica*<sup>TM</sup>. As such, they are obtained from analytical expressions.

### 6.1. Fluctuations in small systems

Systems in thermodynamic environments are always subject to noise. As we have discussed in detail in Chapter 2, noise is the result of the interaction of the system's degrees of freedom with its environment. There we also introduced the notion of the medium which comprises any unobserved degrees of freedom. Generally speaking, fluctuations originate from the unpredictable behaviour of these degrees of freedom.

On macroscopic scales, however, we do not resolve this microscopic behaviour and fluctuations are not visible. The smaller the systems we observe become, the larger the effect of fluctuations. Within the last twenty years or so, novel microscopy techniques allowed probing of systems on ever smaller length and time scales. Though these systems may be *small*, most of them are by no means *simple*.



### 6.1.1. Thermodynamic aspects of molecular motors

Throughout this chapter, we consider the thermodynamic properties of molecular motors. Generally speaking, a molecular motor is a macromolecular machine involved in some intra-cellular activity. Such activities can vary widely from intracellular transport to the locomotion of the whole cell.

To perform these tasks a molecular motor needs an energy source. The most common fuel used by the machinery of life is *adenosine triphosphate* (ATP). It is produced by another molecular motor called ATP synthase which itself is powered by a gradient in proton concentration through a membrane. In green plants, this gradient is provided by *photosynthesis* in the chloroplasts. For eukaryotic organisms, ATP is produced in the mitochondria, as a part of the cellular respiration, *i.e.* the conversion of energy provided by nutrients like sugars or fat. For a detailed account on the molecular biology of the cell *cf.* Ref. [Alb+07].

Many molecular motors have catalytic sites where the exothermic hydrolysis of ATP into *adenosine diphosphate* (ADP) and inorganic phosphate (P) provides the motor's energy. Given their fundamental role for life, thermodynamic aspects of such small engines like their *efficiency* or *power* have recently attracted much attention of both the physics and biology community [How97; Qia05; Sei05; LL07; LLV09; Sei11; SS07; ELB09].

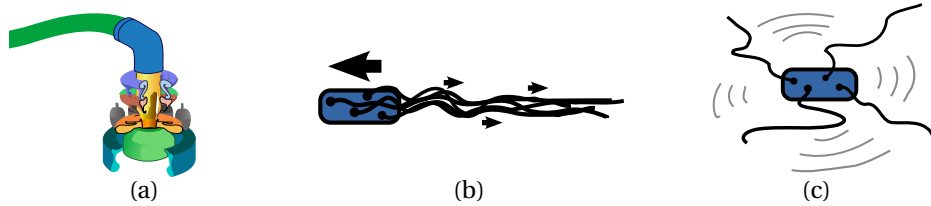
The notions of the efficiency and power of an engine have their origins in the thermodynamics of macroscopic engines. Both small and macroscopic engines operate in a cyclic fashion. Upon the completion of a cycle the system returns to its initial state, while in the mean time some form of energy is converted into another.

However, the difference in size also yields important differences in their functioning. For a macroscopic motor the environment is usually divided into several heat reservoirs. The role of the fuel is simply to provide a temperature gradient in which the motor operates. In contrast, molecular motors usually work in an isothermal environment which acts both as a heat bath and as a source of (chemical) energy. Moreover, the mechanical motion of a molecular machine is strongly influenced by its surroundings.

In addition, inertia plays a major role in sustaining the mechanical motion of a macroscopic motor. In contrast, for small micro-organisms and even smaller molecular motors inertial effects are damped out by a viscous environment. A comprehensive account of the fundamental concepts of locomotion in overdamped situations can be found in Purcell's work on "Life at low Reynolds number" [Pur77].

Another difference of molecular motors and macroscopic engines is the importance of thermal fluctuations. While these fluctuations are negligible for macroscopic engines, they lead to unexpected events in the case of small systems. For instance, rather than hydrolysing an ATP molecule and performing work against some force, a molecular motor may perform this task while *synthesizing* ATP. Surely such an event can never break the first law of thermodynamics, which states that energy is conserved. Consequently, the energy needed to synthesize ATP, which is an exothermic process, must have been taken *from*

## 6. Modelling molecular motors



**Figure 6.1.:** a) Schematic representation of the molecular motor that drives the flagellum of the bacterium *E. Coli*. Its base is immersed in the cell membrane. Flagella are homogeneously distributed over the body of the bacterium. b) If all flagella rotate in the same direction, they entangle to form a large propeller. This enables the bacterium to “run” in a straight line. c) Fluctuations in the rotation direction of some of the motors lead to a disentangling of the propeller. The bacterium tumbles and reorients its direction of motion.

*the environment*. Such *anti-dissipative* events are very uncommon on the macroscopic scale. However, rather than being impossible, they are just extremely improbable. The fluctuation relations we have briefly discussed in Chapter 2 express a symmetry between dissipative ( $Q := k_B T \Delta s^{\text{tot}} > 0$ ) and anti-dissipative ( $Q := k_B T \Delta s^{\text{tot}} < 0$ ) events: The ratio of observing the latter over the former becomes exponentially small upon increasing the magnitude  $|Q|$  of an (anti-)dissipative event. The typical scale of  $Q$  for a molecular motor is of the order of the thermal energy  $k_B T$ , which is much lower than typical macroscopic energy scales. Hence, anti-dissipative events are much more probable for small systems than for macroscopic engines.

### 6.1.2. Functional fluctuations

In the light of evolution, it is not surprising that fluctuations in biological systems may be *functional*. Evolution itself provides us with a first example of such a functionality: Random beneficial mutations (fluctuations) are selected and fixated by external factors, thus increasing the so-called *fitness* (reproduction rate). Theoretical evolution itself has become a major field of research and is built on the tools of statistical physics [Hal10; Gey14; ML10]. However, in the rest of this section we discuss another example of the functionality of fluctuations occurring in a special kind molecular motor.

Bacteria and other micro-organisms that live in a liquid medium, propel themselves by the motion of extracellular filaments called flagella. In many cases the appendages are themselves passive and their motion is the result of the motion of a molecular motor at the base of the flagellum. For the case of the bacterium *E. Coli*, flagella are attached all over the cell’s body, as depicted schematically in Figure 6.1. On average, the motors are synchronized, *i.e.* all the motors and thus the flagella move in the same counter-clockwise direction. In that case, the flagella entangle and form a screw-like propeller and the bacterium performs a movement in a straight line. We say it performs the “run” motion depicted in Figure 6.1b. In order to change its direction of motion, the bacterium goes into a “tumbling” movement, *cf.* 6.1c.

After the tumbling motion, the bacterium performs another run in a new direction. Alternating run and tumble motions result in a random walk of the organism in its spatial

environment. The origin of this “run and tumble” movement has been found to be a random fluctuation of one or few of the molecular motors driving the flagella [Lar+74]: Every now and then one motor may perform a counter-clockwise motion which leads to a disentanglement of the screw propeller. During the process of disentanglement and reformation of the propeller the bacterium tumbles. Moreover, the probability of tumbling to occur is dependent on the chemical environment of the motors. On average, this may lead to a bias in the direction of motion called *chemotaxis* [MK72]. Chemotaxis enables the bacterium to actively move towards sources of food and thus is an essential *function* of the cell. Hence we see that fluctuations may play a pivotal role in the functionality of living matter.

In Section 5.3 of the previous chapter we have introduced large-deviations theory as way to quantify fluctuations. Moreover, we have developed an analytical tool-kit to calculate the fluctuation spectrum. In the following, we use our formalism in order to characterize the fluctuating statistics of biologically relevant physical observables. However, before treating multiple models of the molecular motor kinesin in detail, we introduce a general method of model reduction which is sensitive to fluctuations.

## 6.2. Fluctuation-sensitive model reduction

We have discussed the general idea of model reduction by *coarse graining* already in Chapter 2. There, we were mostly concerned in the coarse graining of continuous degrees of freedom  $x$  into a finite number of effective variables  $\omega$ . The present section is concerned with the coarse graining of Markov chains, *i.e.* we already start with a finite state space  $\Omega$ . After the coarse-graining, we obtain a new state space  $\tilde{\Omega}$  with  $|\tilde{\Omega}| < |\Omega|$ .

As we have seen, coarse graining is often based on the separation of time scales between the original and the coarse-grained level of description. Several coarse-graining procedures for Markov jump processes which are based on time-scale separation have been presented in the recent literature [PV08; Pug+10; Esp12]. In general, a coarse-grained view of a Markovian model is not Markovian any more, due to some memory contained in the probabilities of the fundamental states that make up a coarse-grained state. Only in the limit of an infinite separation of time-scales, Markovianity is rigorously recovered on the coarse-grained level.

Esposito’s approach [Esp12] explicitly discusses this limit. In contrast, in the work of Pigolotti and co-workers [PV08; Pug+10] Markovianity is already built into the coarse-grained model. In order to compare the original with the coarse-grained model, they performed stochastic simulations of both models. From these simulations they obtained finite-time approximations to the large deviation function for the motance, *i.e.* the observable identified with the *entropy variation in the medium*. They found that while certain reduction steps hardly change the fluctuations, others do so in a rather drastic way.

In Ref. [AV12], we presented an alternative coarse-graining procedure. Our approach is in the spirit of Pigolotti’s work in the sense that reduced description is formulated as a

## 6. Modelling molecular motors

Markovian model. However, rather than being based on the separation of time scales, it is based on topological considerations and Schnakenberg's decomposition for the entropy. Further thermodynamic consistency arguments were used to close the equations.

In contrast to Pigolotti's work [PV08], we found that our approach yields a fluctuation-sensitive coarse-graining procedure [AV12]. Moreover, this result seemed to be independent of the concrete model or observable under consideration. In the previous Chapter 5 we presented an algebraic-topological approach to the large deviation theory of Markov jump processes. In fact, these methods were developed in order to quantify the small differences between the original and reduced models. In this section, we briefly review the ideas of Ref. [AV12] and discuss them in the light of the new results obtained in Section 5.3.4.

### 6.2.1. Heuristic motivation

In this subsection we present the heuristic argument for a coarse-graining. We have seen before (*cf.* Section 2.4.2) how entropies are identified in stochastic models of physical systems in thermodynamic environments. Moreover, we have seen in Chapter 4 how entropies and the abstract definitions of motance/affinity also arise without the need for a thermodynamic model. Hence, we consider the anti-symmetric motance and (steady state) affinity matrix  $\mathcal{E}$ ,  $A \in \mathcal{O}$  with elements  $\mathcal{E}_{\omega'}^{\omega} = \log \frac{w_{\omega'}^{\omega}}{w_{\omega}^{\omega'}}$  and  $A_{\omega'}^{\omega} = \log \frac{\phi_{\omega'}^{\omega}}{\phi_{\omega}^{\omega'}}$  as the relevant physical observables. Our method is based on the following (heuristic) requirements:

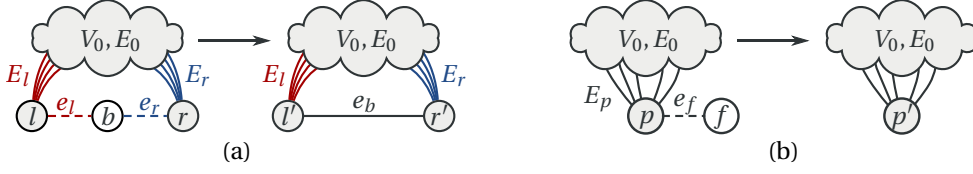
- (i)a *Cycle topology*: The number and mutual connections of cycles are preserved.
- (i)b *Cycle affinities*: The algebraic values of the affinity of any cycle is preserved.
- (ii) *Locality*: Fluxes and probabilities may only change locally
- (iii) *Trajectories*: The system's entropy variation along single trajectories is the same between the models.

Requirement (i) describes consistency within the stochastic models and with macroscopic thermodynamics. It is motivated by the role of cyclic currents as the fundamental building blocks of Markov jump processes, *cf.* also the discussion in Sec. 5.4.3. The Schnakenberg decomposition of the entropy production in the steady state is based on the cycle affinities, are directly related to the thermodynamic forces driving the system.

Requirement (ii) is motivated in the light of the connection to an underlying microscopic phase space: Imagine that the coarse-graining acts only on a part of the (observable, mesoscopic) states and leaves others untouched. Then, the dynamics on the elements of the phase-space partition that correspond to untouched states should not change. This is the notion of *locality* expressed in requirement (ii).

The last requirement (iii) is motivated in the light of stochastic thermodynamics. The system's entropy variation along trajectories is preserved. It constitutes the final relation needed to close our equations.

## 6.2.2. Target topologies and coarse-grained observables



**Figure 6.2.:** The reduction of vertices (white) appearing in certain subgraphs does not change the cycle structure of the graph: a) Bridges ( $b$ ) are vertices which have exactly two neighbours ( $l, r$ ), which are themselves not connected. b) Leaves ( $v$ ) have only one neighbour ( $p$ ). Note that trees are composed out of bridges and leaves only.

The requirement (i) restricts our coarse-graining procedure to target only vertices with a certain topology of its neighbours. More precisely, we apply it only to states appearing in certain subgraphs. Our target topologies are called *bridges* and *leaves*, which we illustrate in Figure 6.2. A bridge is defined as a vertex that is only connected to two neighbouring vertices, which are not themselves adjacent. The latter requirement is important, because the smallest non-trivial cycle consists of three states. A *leaf* is a vertex that has only a single neighbour.

Note that a tree is a subgraph that can be reduced to a single vertex by subsequently removing leaves. Further, for any connected graph one can find a spanning tree that includes any bridge and leaf states together with their adjacent edges. In that regard, dropping leaves and edges corresponds to dropping edges of the spanning tree, while preserving the chords of a graph. In the light of the discussion in the Section 5.3.4, this appears to be a good choice in order to preserve the (asymptotic) fluctuations.

## Observables and bridges

By changing the mesoscopic state space in our model reduction, we also need to define the new coarse-grained observables  $\tilde{\varphi} \in \tilde{\mathcal{O}}$  for the observables  $\varphi \in \mathcal{O}$  of the original model. In the following, we denote the “left” and “right” neighbours of a bridge state  $b$  by  $l$  and  $r$ , respectively. The remaining vertices are summarized in the vertex subset  $V_0$ , cf. Fig. 6.2.

$$\tilde{\varphi}_{r'}^l = \varphi_r^l + \varphi_l^r + (d_l - d_r), \quad (6.1a)$$

$$\tilde{\varphi}_i^{n'} = \varphi_i^n + d_n \quad \text{for } i \in V_0, n \in \{l, r\}, \quad (6.1b)$$

$$\tilde{\varphi}_j^i = \varphi_j^i \quad \text{for } i, j \in V_0. \quad (6.1c)$$

The constants  $d_l$  and  $d_r$  act as gauges that do not change the macroscopic observations. They only change the part of  $O$  that lies in the co-cycle space. In the phase-space picture, they depend on the microscopic dynamics and the chosen re-partitioning of phase space. As we do not know these details, we choose  $d_l \equiv d_r \equiv 0$  for simplicity.

## 6. Modelling molecular motors

### Observables and leaves

We also need a rule for coarse-graining the observables if we drop leaves. As on average the system jumps into the leaf state as often as it leaves it again. Hence, we set:

$$\tilde{\varphi}_i^{p'} = \varphi_i^p + d_p \quad \text{for } i \in V_0, \quad (6.2a)$$

$$\tilde{\varphi}_j^i = \varphi_j^i \quad \text{for } i, j \in V_0. \quad (6.2b)$$

Again,  $d_p$  is a co-cyclic gauge and for simplicity we choose  $d_p \equiv 0$ .

### 6.2.3. The coarse-graining algorithm

In Ref. [AV12] we presented the coarse-graining prescription for bridge states. In the CG procedure we absorb the bridge into its neighbours leading to new states  $l'$  and  $r'$ . Thus, the transition rates for the sets of edges  $E_l$  and  $E_r$  connecting  $l$  and  $r$  to the rest of the network have also to be adjusted.

This has to be done in accordance with requirement (i)b, *i.e.*  $A_\alpha = A'_\alpha$  where  $A_\alpha$  is the cyclic affinity for any cycle  $\alpha$ . Demanding the conservation of fluxes along any edge  $e \in E \setminus \{e_l, e_r\}$  that is not adjacent to the bridge state  $b$ , Eq. (2.37) yields

$$\frac{\phi_{r'}^{l'}}{\phi_{l'}^{r'}} \stackrel{!}{=} \frac{\phi_b^l \phi_r^b}{\phi_b^r \phi_l^b}. \quad (6.3a)$$

Any trajectory passing through the two edges  $(l, b)$  and  $(b, r)$  in the original model will be a trajectory through  $(l', r')$  in the coarse-grained model.

In ST, the change of entropy associated to the system's state is the difference of the logarithms, *cf.* Eq. (2.48a). In the following,  $\mathbf{p} = \mathbf{p}_\infty$  always denotes the stationary distribution. Hence, for brevity of notation we write  $p_\omega$  instead of  $p_\infty(\omega)$ , as we have done in previous chapters. Condition (iii) demands that the change in system's entropy along a trajectory entering the bridge on one side and re-emerging on the other side is the same between the models. In terms of the invariant probabilities, this amounts to

$$\frac{p_l}{p_r} \stackrel{!}{=} \frac{p_{l'}}{p_{r'}}. \quad (6.3b)$$

A priori, other closures of the form  $p_{l'}/p_{r'} = c$  are also possible but lack the advantage of an interpretation in ST.

Without loss of generality we assume that there is a positive net current  $J_{\text{bridge}} = J_b^l = J_r^b > 0$  flowing through the bridge from the left to the right neighbour state. Together with the steady-state balance condition (3.2) and the locality assumption, the Eqs. (6.3)

uniquely determine all rate constants of the coarse-grained model [AV12]:

$$w_{n'}^i = w_n^i \quad \text{for } i \in V_0, n \in \{l, r\}, \quad (6.4a)$$

$$w_i^{n'} = w_i^n / g \quad \text{for } i \in V_0, n \in \{l, r\}, \quad (6.4b)$$

$$w_{r'}^l = (J_{\text{bridge}} + m) / (g p_l), \quad (6.4c)$$

$$w_{l'}^r = m / (g p_r), \quad (6.4d)$$

where

$$g = (p_l + p_r + p_b) / (p_l + p_r), \quad (6.5a)$$

$$m = \phi_b^r \phi_l^b / (J_{\text{bridge}} + \phi_b^r + \phi_l^b). \quad (6.5b)$$

The determination of the new rates can also be expressed as a reduction procedure for the motance:

$$\tilde{\mathcal{E}}_{r'}^l = \mathcal{E}_r^l + \mathcal{E}_l^r, \quad (6.6a)$$

$$\tilde{\mathcal{E}}_i^{n'} = \mathcal{E}_i^n - \log g \quad \text{for } i \in V_0, n \in \{l, r\}, \quad (6.6b)$$

$$\tilde{\mathcal{E}}_j^i = \mathcal{E}_j^i \quad \text{for } i, j \in V_0. \quad (6.6c)$$

Eq. (6.6a) expresses that the dissipation along a trajectory which passes through the bridge is conserved. Eq. (6.6b) states that along the edges  $\mathcal{E}_n$ ,  $n \in \{l, r\}$ , there is an additional contribution  $-\log(g)$  to  $\tilde{\mathcal{E}}_i^n$ , which is the same for both neighbours due to the closure (6.3b). Eq. (6.6c) expresses locality and is independent of the closure.

For the case of leaves, locality of probability demands that  $p_{p'} \stackrel{!}{=} p_p + p_f$ . In order to preserve the fluxes, the new rates need to obey

$$w_{p'}^i = w_p^i \quad \text{and} \quad (6.7a)$$

$$w_i^{p'} = \frac{p_p w_i^p}{p_p + p_f} \quad \text{for } i \in V_0. \quad (6.7b)$$

### 6.3. Applications to kinesin

In order to be concrete, we illustrate our methods using the well-established model of the molecular motor kinesin [LL07; LL09]. Kinesin is one of the most well-studied molecular motors. It is responsible for the transport of cargo in eukaryotic cells. As such, it plays a major role also in the reproductive cycle of cells, *cf.* Refs. [Alb+07; Yil+04; CC05] and the references therein for further reading.

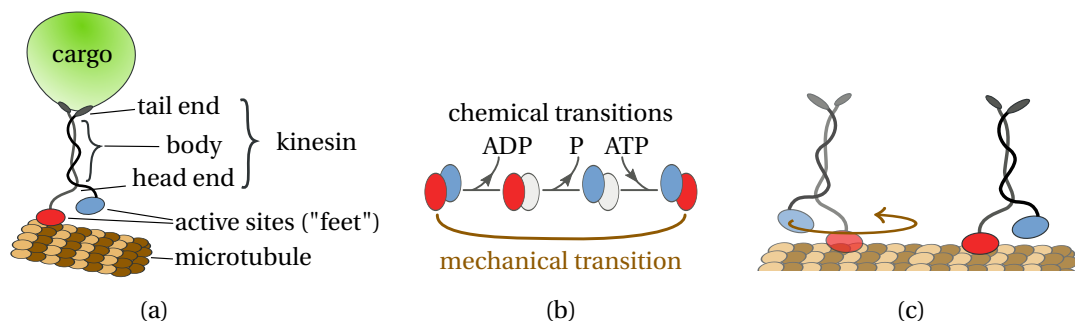
Figure 6.3a) shows a schematic representation of kinesin, which is a macromolecule composed of several subunits. The “feet” or *motor domains* are its active sites constituting

## 6. Modelling molecular motors

the *head end* of the molecule. The “body” of the molecule is a stalk connecting the head end to the tail end. At the latter one finds binding site for kinesin’s payload, which is a vesicle containing other chemical compounds.

The active sites bind to intracellular filaments called *microtubules*, which are one on the central structural elements of the cytoskeleton. Microtubules are themselves hollow cylinders which consist of polymerized dimers called *tubulin*. They further show *polarity* due to the helical arrangement of the tubulin components. Hence, the two ends of a microtubule are distinguishable.

The active sites act as a catalyst for the hydrolysis of adenosine triphosphate (ATP) into adenosine diphosphate (ADP) and inorganic phosphate (P). This exothermic reaction provides the energy for a mechanical transition along the microtubule. The polarity of the microtubule thereby ensures a preferred direction of motion. The mechanics of kinesin’s mechanical step have only recently been understood [Yil+04]. They arise from a change of the active sites’ binding affinity to the microtubule that goes along with the hydrolysis reaction. Additionally, during this reaction the molecule undergoes a conformational change. The complex interplay of this mechanisms leads to the motion depicted in Figure 6.3b).



**Figure 6.3.:** Kinesin is a molecular motor featuring two chemically active sites (“feet”). Different colours at kinesin’s active sites represent the different chemical compositions. Here, we display the typical behaviour of the motor where forward movement is linked to the hydrolysis of ATP (red) into ADP (blue). (a) Schematic representation of kinesin attached to a microtubule. (b) Dominant mechano-chemical pathway of kinesin under physiological conditions: The net attachment, hydrolysis and release of one ATP molecule drives a conformational change, leading to a mechanical step. (c) Kinesin walks in a “hand-over-hand” motion, where the mechanical transition exchanges the leading with the trailing active site.

### 6.3.1. A model for kinesin

Liepelt and Lipowsky [LL07] have constructed a chemo-mechanical Markov model for kinesin, which is shown in Figure 6.4a). Their minimal model included six states, each of which corresponding to a different chemical composition of the active sites. Upon a mechanical step of distance  $l \approx 8$  nm, the position of the leading and trailing “foot” is interchanged.

In order to explore kinesin’s response to different non-equilibrium conditions, Carter



and Cross designed an experiment where a mechanical force  $F$  acts on the motor [CC05]. A positive force  $F > 0$  thereby indicates that it acts *against* kinesin's natural direction of motion.

Under physiological chemical conditions and in the absence of an external force, on average one molecule of ATP is hydrolysed per mechanical step. However, the chemical and the mechanical cycle are not *tightly coupled* [LL09]. It is possible to have a futile hydrolysis of ATP, with no mechanical step taking place. In the model, this is reflected by the presence of three cycles, two of which are independent, *cf.* Fig. 6.4. The *forward cycle*  $\mathcal{F}$  corresponds to the standard motion under physiological conditions. The *backward cycle*  $\mathcal{B}$  describes the situation where fuel consumption leads to a backward rather than a forward step. Upon completion of the *dissipative slip cycle*  $\mathcal{D}$ , two molecules of ATP are hydrolysed while no mechanical transition takes place.

The construction of the six-state model shown in Figure 6.4a) is found in Ref. [LL07]. For the determination of the transition rates and their dependence on chemical concentration and mechanical load, experimental data was used. Appendix B reviews the arguments brought forward in Ref. [LL07] to construct a simpler four-state model.

In the following Sections 6.3.2 and 6.3.3, we illustrate our methods using different models for kinesin. At first, we demonstrate the power of the coarse-graining procedure introduced in the previous Section 6.2 as published in Ref. [AV12]. For both the original and several reduced models we plot the rate function  $I_\varphi(x)$  obtained for the time-averages of several observables  $\varphi$ . In addition, we demonstrate the convergence of the scaling form of sampled probability distributions to  $I_\varphi$ . Remarkably, not only the asymptotic statistics (characterized by the rate function) but also the sampled distributions for finite times agree to a very high degree between the models.

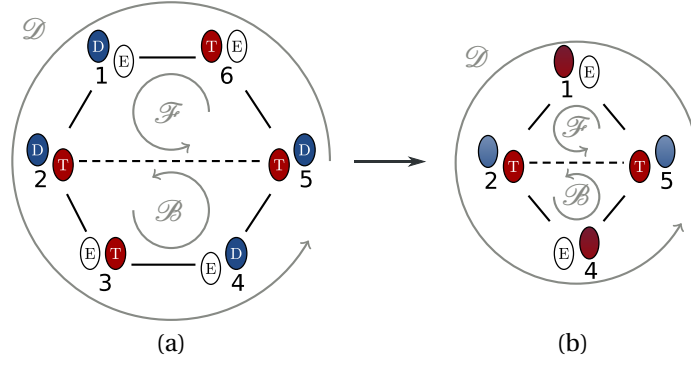
Secondly, we investigate the full *phase diagram* of kinesin, which is spanned by the values of the driving parameters  $\mu$  and  $F$ . Motivated by the success of our coarse-graining procedure, we conclude that a minimal model with four-states is enough. In order to get rid of the ambiguity when it comes to choosing *which* bridge states are removed, we introduce a simple four-state model, *cf.* Appendix B.

### 6.3.2. Fluctuations in the coarse-grained model

In Ref. [AV12] we applied the coarse-graining procedure outlined in Section 6.2 to the bridge states of the six-state model. Note that both the forward and the backward cycle contain two bridges that lie next to each other. Upon reduction of one of the neighbouring bridge states, the resulting state is not a bridge any more, as its neighbours are mutually connected. Hence, there is an ambiguity when it comes to the choice, which states to remove. As an example, in Fig. 6.4b we show the removal of the bridge states 6 and 3, yielding a coarse-grained model on four states.

Figure 6.5 shows the effect of the coarse-graining procedure for a set of parameters that correspond to physiological conditions. More precisely, we considered the time averages

## 6. Modelling molecular motors



**Figure 6.4.:** (a) 6-state model of kinesin [LL07] with its cycles. The dashed line represents the mechanical transition which allows the motor to move along the microtubule. All other transition are chemical. The cycle  $\mathcal{F}$  is called the forward cycle. It corresponds to the process that is usually found under physiological conditions shown in Fig. 6.3b: The hydrolysis of one ATP molecule leads to a forward step. In the lower cycle  $\mathcal{B}$ , an ATP molecule is hydrolysed and the motor performs a backward step. The outer cycle  $\mathcal{D}$  describes the futile hydrolysis of two ATP molecules. (b) Coarse-grained description with bridges 3 and 6 reduced. Note that all cycles are preserved.

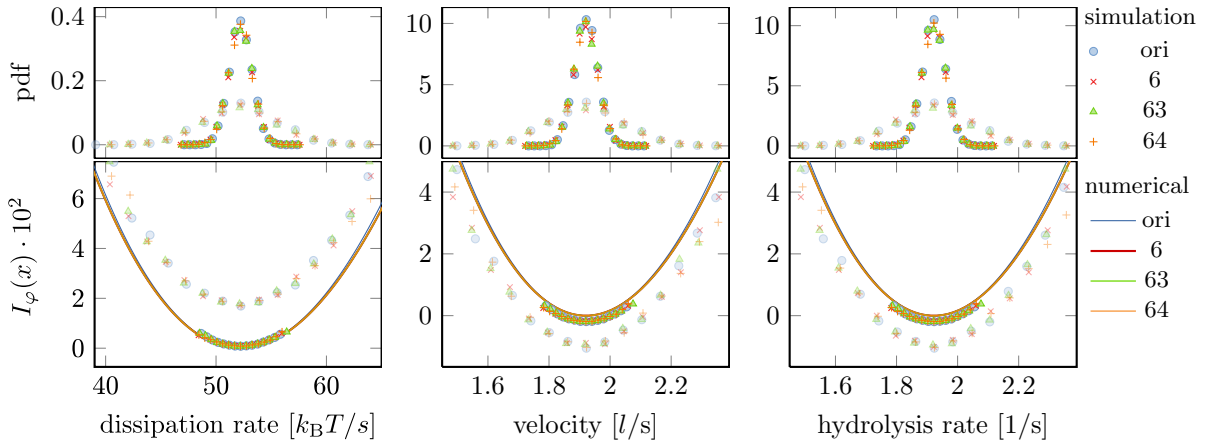
of three observables: The hydrolysis rate, the velocity and the dissipation rate. The former two were obtained by counting the transitions along the respective edges in the graph, whereas the latter is the time-average (5.33) associated with the motance matrix. We will discuss the *chord representations* of these observables in detail below.

The data presented in Figure 6.5 illustrates several results. First of all, the data obtained for different implementations of the coarse-graining lie on top of the data obtained in the original model. This holds true for both the asymptotic statistics characterized by a rate function  $I_\varphi(x)$  as well as for the finite-time probability distributions  $\rho_\varphi^{(r)}(x)$ . While the former were calculated using Equation (5.30), the latter are sampled from the simulation of a large number of stochastic trajectories.

The results show the efficiency of our fluctuation-preserving coarse-graining: Firstly, it preserves the whole fluctuation spectrum (encoded in the rate function) to a very good degree. This is true not only for the dissipation, but also for other observables. The good agreement is a consequence of the generalized Schnakenberg decomposition (5.39). Secondly, the results do not change between the original model on six states, and the different five- and four-state models obtained from the removal of bridge states. This underlines the importance of the cycle topology, which we emphasized in our heuristic motivation of the coarse-graining procedure.

In conclusion, a minimal model on four states that respects the cycle topology and the affinities of the six-state model is sufficient to capture the fluctuations of a more complicated model. Because we do not allow multiple transitions between states, a non-tightly coupled model featuring multiple cycles cannot be formulated on only three states. This motivates the construction of minimal four-state model directly from experimental data which is shown in Appendix B.

This minimal model has several advantages over a four-state model obtained from the



**Figure 6.5:** Simulation and numerical results for dissipation rate, moving velocity and hydrolysis rate of the kinesin model taken from Ref. [AV12]. Data is shown for the original 6-state model (“ori”), a 5-state model with state 6 reduced (“6”) and two 4-state models with state 6,3 or 6,4 reduced (“63” and “64”). The rate constants for the original model are taken from Ref. [LL07] describing the data in Ref. [CC05] for chemical concentrations  $c_{\text{ADP}} = c_{\text{P}} = c_{\text{ATP}} = 1 \mu\text{M}$  and stepping size  $l \approx 8 \text{ nm}$ . The top row shows the sampled pdf for  $\tau \approx 1200 \text{ s}$  (opaque symbols) and  $\tau \approx 120 \text{ s}$  (transparent symbols). The bins with the width of half an empirical standard deviations are centred around the empirical mean. For the simulation we sampled  $N = 5000$  trajectories. The bottom row shows convergence of rescaled data (cf. Eq. (5.28)) to the rate function  $I_\varphi(x)$  (solid lines) obtained via Legendre transform (implemented using Ref. [Luc97]) of the SCGF  $\lambda_\varphi(q)$ .

coarse-graining procedure. Firstly, there is no ambiguity with respect to the choice of the removed bridge states. Secondly, when removing a bridge state, one distributes it among its two neighbours. In that process, the neighbours lose their former interpretation as being the equivalence classes belonging to a certain observable result. This is also reflected by the new transition rates: They lose their direct interpretation as the kinetic rates of a specific chemical reaction.

### 6.3.3. Analytical treatment of kinesin’s phase diagram

In the previous section, we discussed the fluctuation spectrum of the 6-state kinesin model and its reduction to 5-state and 4-state models for a special set of (physiological) parameters, cf. Figure 6.5. In Ref. [LL09], Liepelt and Lipowsky analysed the “phase diagram” of their model obtained by varying the forces that drive the system. The effective parameters of the model are the thermodynamic affinities characterized by the (non-dimensionalized) chemical potential difference  $\mu$  and a mechanical load  $f$  [LL07]. With the equilibrium constant  $K_{\text{eq}}$  for the hydrolysis of ATP, one defines

$$\mu := \log \frac{K^{\text{eq}}[\text{ATP}]}{[\text{ADP}][\text{P}]}, \quad (6.8)$$

## 6. Modelling molecular motors

where  $[X]$  stands for the chemical concentration of compound  $X$  in the solution. The non-dimensionalized force is defined as

$$f := \frac{lF}{k_B T} \quad (6.9)$$

where  $l$  is the distance of kinesin's mechanical step and  $F$  is the applied external force. The non-dimensionalized driving parameters constitute the affinities of the cycles of the model. One also refers to this thermodynamic consistency requirement as the ‘‘Schnakenberg conditions’’ [AG07; FDP11].

The Schnakenberg decomposition (5.40a) ensures that the affinities of all other cycles are obtained as linear combinations of the affinities of the fundamental cycles [Sch76]. Note that these affinities in turn are linear combinations of the (dimensionless) chemical and mechanical drivings acting on the system. They correspond to net (free) energy difference in the environment (divided by  $k_B T$ ) upon the completion of a thermodynamic cycle, cf. also Ref. [Sei11].

After a mechanical step, the work performed by the external force on the system amounts to  $f k_B T = F \cdot l$ . Similarly, after a hydrolysis reaction the chemical potential in the environment has changed by a factor  $k_B T \mu = k_B T \log \frac{K_{\text{eq}}[\text{ATP}]}{[\text{ADP}][\text{P}]}$ . To the present author's knowledge, Hill first realized and appreciated this fact. He calls this phenomenon the ‘‘transduction of (free) energy’’ [Hil77]. Consequently, we prefer the term Hill–Schnakenberg conditions or simply *thermodynamic balance conditions*. Formally, one may express this fact as

$$\mathcal{E}_\alpha := \sum_{e \in \alpha} \mathcal{E}_e \stackrel{!}{=} \sum_{\nu} n_\alpha^\nu (\Delta A)^\nu, \quad (6.10)$$

where  $\mathcal{E}_\alpha$  is the motance of a cycle,  $\Delta A^\nu$  refers to a difference in a (thermodynamic) potential and  $n_\alpha^\nu \in \mathbb{Z}$  is an integer.

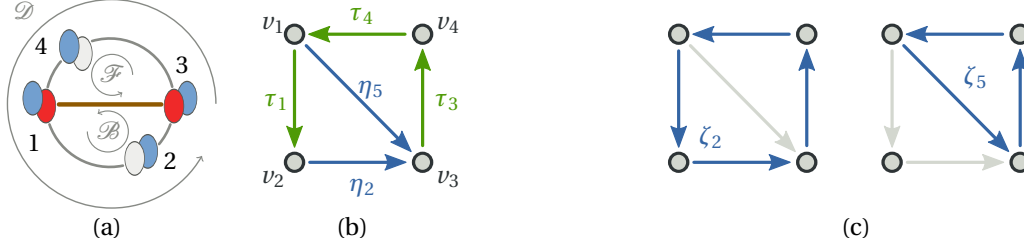
### Chord representation of the observables

Throughout the present Section 6.3.3 we plot results obtained using the six-state model. In the next Section 6.3.4 we compare these results to minimal and coarse-grained models formulated on only four states.

For the sake of simplicity, however, already in the present section we exemplify the application of our method using a four-state model. A detailed treatment can be found in Ref. [WVA14]. Here, we focus on the chord representation (5.38) of biologically relevant physical observables. Moreover, using the four-state model we ensure continuity with respect to the previous chapter, where we exemplified cycles and co-cycles using a graph on four vertices.

Figure 6.6 shows an abstract representation of a minimal four-state model alongside its abstract representation as a directed graph. We chose the enumeration of the vertices in Fig. 6.6a) to be consistent with Fig. 6.6b) and hence with all the examples used in Chapter 5.

This allows us to identify a basis of fundamental cycles basis of fundamental cycles  $(\zeta_2, \zeta_5)$  which corresponds to the dissipative slip and forward cycle, respectively.



**Figure 6.6.:** (a) Our minimal 4-state model for kinesin. Details of the parametrization can be found in Appendix B. Note that the enumeration is different than in Figure 6.4. However, it is consistent with the graph shown in (b). Note that this graph was used as an example throughout Chapter 5. The chosen spanning tree  $(V, \{\tau_1, \tau_3, \tau_4\})$  is marked in green. The fundamental chords  $\eta_2, \eta_5$  are shown in blue. (c) The fundamental cycles  $\zeta_2$  and  $\zeta_5$  correspond to the dissipative slip cycle  $\mathcal{D}$  and the forward cycle  $\mathcal{F}$ , respectively.

The chord representation (5.38)  $\varphi_{\mathcal{H}}$  of a physical observable  $\phi$  simplifies the calculation the fluctuations. We demonstrate our method using the same observables considered already in Figure 6.5: The displacement rate  $d := l\eta_5$  and the hydrolysis rate  $h := \tau_1 + \tau_3$  count the centre of mass movement ( $l = 8 \text{ nm}$ ) and the numbers of ATP molecules hydrolysed, respectively. The dissipation rate is calculated as the time-average of the motance  $\mathcal{E}_{\omega'}^{\omega} \mapsto \log \frac{w_{\omega'}^{\omega}}{w_{\omega}^{\omega'}}$ .

The chord representations  $\varphi_{\mathcal{H}} = \langle \zeta_2, \varphi \rangle \eta_2 + \langle \zeta_5, \varphi \rangle \eta_5$  for those observables read

$$d_{\mathcal{H}} = l \langle \zeta_2, \eta_5 \rangle + l \langle \zeta_5, \eta_5 \rangle \eta_5 = l \eta_5, \quad (6.11a)$$

$$h_{\mathcal{H}} = \langle \zeta_2, \tau_1 + \tau_3 \rangle \eta_2 + \langle \zeta_5, \tau_1 + \tau_3 \rangle \eta_5 = 2\eta_2 + \eta_5, \quad (6.11b)$$

$$\mathcal{E}_{\mathcal{H}} = \langle \zeta_2, \mathcal{E} \rangle \eta_2 + \langle \zeta_5, \mathcal{E} \rangle \eta_5 = 2\mu\eta_2 + (\mu - f)\eta_5. \quad (6.11c)$$

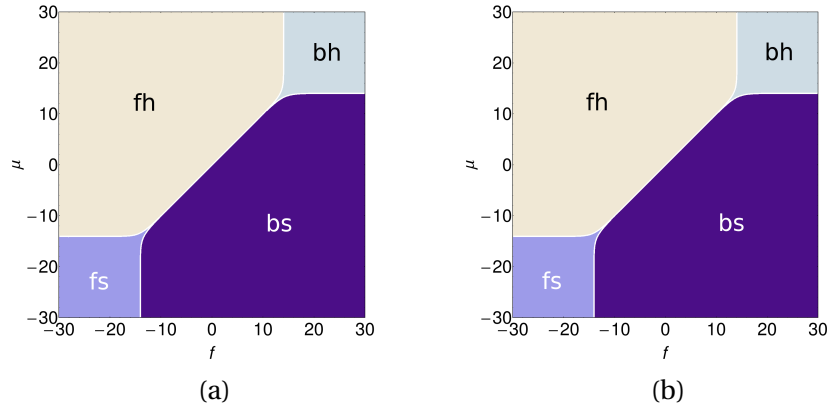
The first two lines are directly evident from Figure 6.6. The third line reflects the thermodynamic balance conditions (6.10). These conditions enter explicitly in the construction of the minimal four-state model as well as the construction of the original six-state model, cf. Appendix B and Ref. [LL07], respectively.

### A phase diagram and beyond

Based on Equations 5.39 derived in Section 5.3.4, we can calculate the asymptotic fluctuation spectrum of the (time-averages) of these currents analytically. All that we need are the chord representations (6.11) of the observables we are interested in and the scaled cumulants for the currents on the fundamental chords. For brevity, we will only deal with the dominant part of the spectrum, *i.e.* with the first and second scaled cumulants.

Using the first scaled cumulants, *i.e.* the steady-state averages, in Figure 6.7a) we reproduce a result by Liepelt and Lipowsky, regarding the operational modes of kinesin [LL09]. The modes are defined by the signs of the first cumulants  $c_1(d)$  and  $c_1(h)$  of the displacement and hydrolysis rate. They distinguish the direction of the motor's velocity along the

## 6. Modelling molecular motors



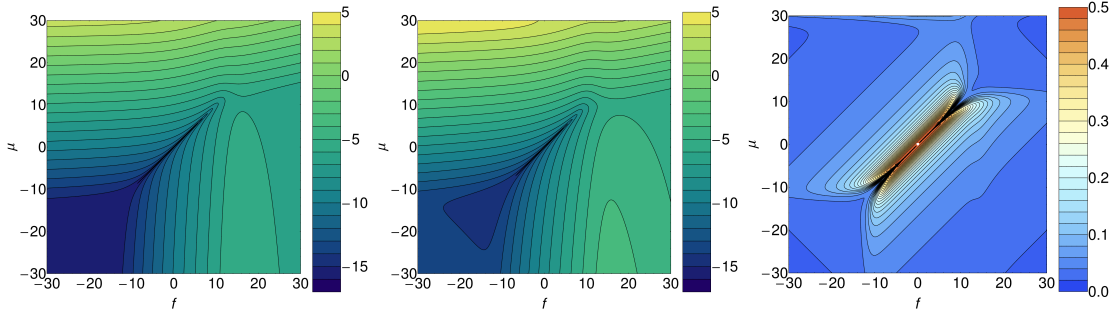
**Figure 6.7.:** Phase diagram of the molecular motor kinesin in the  $(f, \mu)$ -plane. The left and right plots are analytical results obtained for the well-established six-state model [LL07] and a minimal four-state model, cf. Appendix B, respectively. The different regions in the diagram reflect the different tuples  $(\text{sgn}(c_1(d)), \text{sgn}(c_1(h)))$ , i.e. the signs of the average motor velocity and the average hydrolysis rate. They correspond to the four operational modes of kinesin: forward stepping and hydrolysis ( $\text{fh} \equiv (+, +)$ ), backward stepping and hydrolysis ( $\text{bh} \equiv (-, +)$ ), forward stepping and synthesis ( $\text{fs} \equiv (+, -)$ ), and backward stepping and synthesis ( $\text{bs} \equiv (-, -)$ ), cf. also Fig. 6.9. In this phase diagram, no difference is visible between the two models.

microtubule and whether ATP is produced or consumed, respectively. The top left part of the diagram corresponds to the usual operation mode of kinesin under physiological conditions: The motor moves forward on the microtubule while hydrolysing ATP. Even if we apply a force ( $f > 0$ ) pulling the motor back this motion is sustained. However, for sufficiently large forces the motor runs backward while still hydrolysing ATP (top right). In the other large region on the lower right, kinesin synthesizes ATP from ADP and P. In the part of this region where the force is positive, we can say that kinesin acts as a “chemical factory”: Mechanical work is converted into chemical energy stored in ATP. Finally, the lower left region corresponds to ATP synthesis during forward motion. We see that for this behaviour there must be both a force that pushes the motor while also the chemical concentrations in the environment highly favour ATP synthesis.

As a comparison, Figure 6.7b shows the phase diagram obtained with our minimal four-state model. To the eye, it is indistinguishable from Figure 6.7a) obtained for the six-state model. Thus in order to compare the models in more detail, we need additional information than just the sign of the first cumulants.

Our method provides this additional information. Before we discuss the actual *values* of the first and second cumulants for the displacement and the hydrolysis, we look at the (non-dimensionalized) dissipation provided by the motance matrix  $\mathcal{E}$ . The dissipation is in a sense the most natural scalar observable to look at. It is completely determined by the logarithmic ratio of the rates alone. Hence, it is well-defined for *any* dynamically reversible Markovian model and does neither require a physical interpretation of the transition rates nor of the system as such.

Figure 6.8 shows our results. The plot on the left shows the decadic *logarithm* of the first



**Figure 6.8.:** Dissipation in the six-state model corresponding to the physical observable  $\mathcal{E}$ . The axis the plots correspond to the chemical and mechanical driving parameters  $\mu$  and  $f$ , respectively. Left: The decadic logarithm  $\log_{10} c_1(\mathcal{E}) = \log_{10} \langle s^{\text{tot}} \rangle_{\infty} = \log_{10} \langle s^{\text{med}} \rangle_{\infty}$ . Centre: The decadic logarithm  $\log_{10} c_2(\mathcal{E})$  characterizing the strength of fluctuations in the dissipation. Right: The signal-to-noise ratio  $c_1(\mathcal{E})/c_2(\mathcal{E})$ .

scaled cumulant  $\log_{10} c_1(\mathcal{E})$ , *i.e.* the expectation value of the total entropy production in the steady state. The black contour lines thus correspond to values separated by one order of magnitude. Hence, the values we plot range over more than twenty decades. The centre plot for the logarithm of the second scaled cumulant  $\log_{10} c_2(\mathcal{E})$ , which characterizes the (scaled) variance in the dissipation rate, ranges over similar values. Qualitatively, it looks rather similar to the first cumulant. In order to find more structure, the rightmost plot shows a signal-to-noise ratio (SNR) calculated as the ratio  $c_1(d)/c_2(d)$ . Note that we use a *linear* scale here.

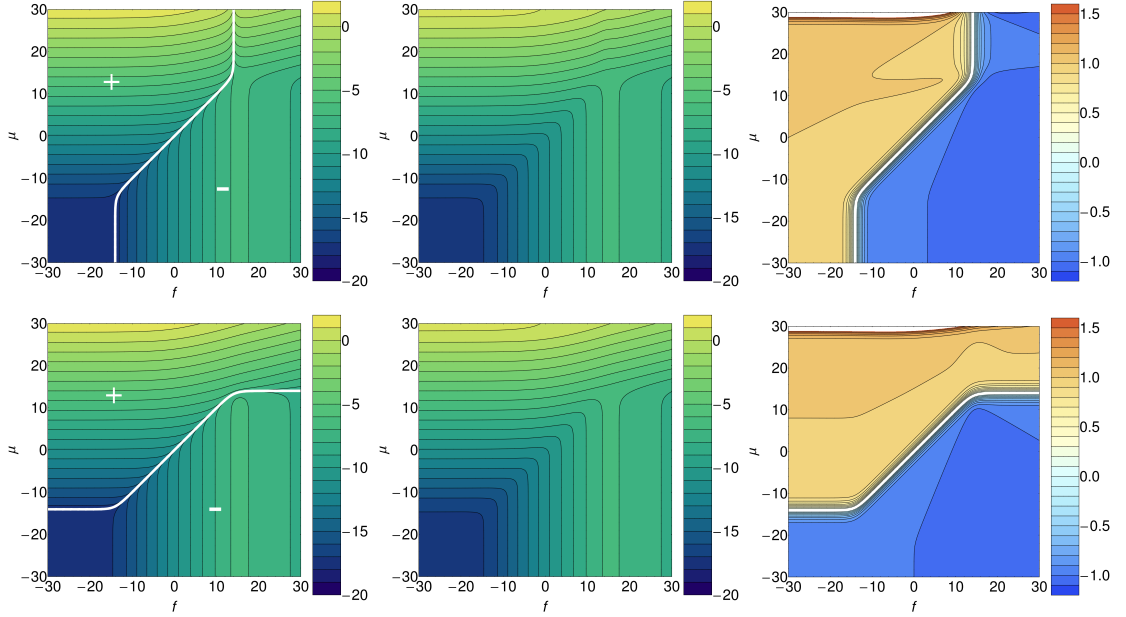
Before we interpret these results, let us first take a look at the corresponding plots for the observables that define the operation modes of the motor. Figure 6.9 shows similar plots for the logarithms of first and second cumulants  $|c_1|, c_2$  as well as the SNR  $c_1/c_2$  for the displacement  $d$  (top row) and hydrolysis rate  $h$  (bottom row), respectively.

On first sight, the plots for the first and second cumulants look qualitatively similar to the ones shown in Fig. 6.8: The values range over the large range spanning more than twenty logarithmic decades. Moreover, the overall features illustrated by the contour lines (*e.g.* the existence of a local maximum in the absolute value of the cumulants around  $f \approx 15$ ) are common to all plots.

However, there is a caveat to this statement. Unlike the average entropy production, the average displacement rate (*i.e.* the velocity) and the average hydrolysis rate change their signs. Consequently, the logarithm of the absolute value of the first cumulant  $|c_1|$  shows a continuous line of singularities. The position of these singularities is marked by a white line in the first column of Fig. 6.9. Note that the white lines are by definition the lines separating the different modes of kinesin, shown in the phase-diagram in Fig. 6.7a).

Let us discuss the structure found in the cumulants in more detail. The non-negative dissipation rate vanishes at the centre of the diagram, where we have thermodynamic equilibrium  $f = \mu = 0$  and the motor does neither move nor catalyse any chemical reactions on average. Further, the dissipation is very small along the first diagonal  $f = \mu$  in a neighbourhood around the origin.

## 6. Modelling molecular motors



**Figure 6.9.:** First two cumulants of the mechanical displacement  $d$  (top row) and hydrolysis  $h$  (bottom row) in the six-state model. The first column shows the absolute value of the first scaled cumulants  $c_1(d)$  and  $c_1(h)$ , corresponding to the average motor velocity and hydrolysis rate, respectively. The sign is indicated as an inset, with the two regions separated by the thick solid line corresponding to parameters such that  $c_1 = 0$ , cf. also the phase diagram Fig. 6.7a). The presentation of the data is similar to the presentation in Fig. 6.8, i.e. the middle and right columns show the second cumulant  $c_2$  and the SNR  $c_1/c_2$ , respectively.

It is a consequence of the quasi-tight coupling mechanical and chemical transitions for kinesin: It is very natural for the motor to behave in the way shown in Fig. 6.3b, as expressed by the succession of states along the forward cycle. To appreciate this fact, note that the affinity of the forward cycle  $\zeta_5$  is given by  $\mathcal{E}(\zeta_5) = \mu - f$ . Hence, the diagonal in the phase diagram corresponds to a vanishing affinity along that cycle. Similarly, in the centre region the values of the velocity and hydrolysis rate vanish somewhere very close to this diagonal.

Moreover, for not too high driving forces we have an (approximate) reflection symmetry along the diagonal, which holds for the first and second cumulant of all the observables we considered here. In conclusion, the (affinity of the) forward cycle dominates the dynamics of kinesin — at least in regions where the absolute values of the non-dimensionalized drivings are smaller than about ten.

Another prominent feature is the region of low dissipation in the lower left corner of the phase diagram. It is the region where the system on average runs along the reversed backwards cycle, i.e. on average it moves forwards while synthesizing ATP. It does so, however in an extremely slow fashion, as indicated by the small values of the velocity and hydrolysis rate (cf. the left column of Fig. 6.9). Hence, the reason for the small dissipation rate in that part of the phase diagram is the slowness of the kinetics. Note that this feature also shows in the plots for the second cumulants: In the kinetically hindered region in the



lower left, the system is essentially “frozen”. Upon increasing the chemical driving, the system starts moving again.

More structure is visible in the third column where we plot the signal-to-noise-ratio (SNR)  $c_1/c_2$ . Nevertheless, we stop the discussion of the structure of kinesin’s phase-diagram at this point. Additional remarks regarding the significance of the SNR are made in the discussion in Section 6.4.1. Before we come to that point, however, let us see whether these features are preserved in simplified models.

#### 6.3.4. Simplified models

In the present section, we discuss the structure of kinesin’s phase diagram in simplified models. The main focus thereby lies on the minimal model constructed in Appendix B, where we carefully made sure to follow the physical arguments for the construction of the six-state model in Ref. [LL07]. After that, we quickly discuss simplified models obtained by the fluctuation-sensitive coarse-graining approach presented in Section 6.2.

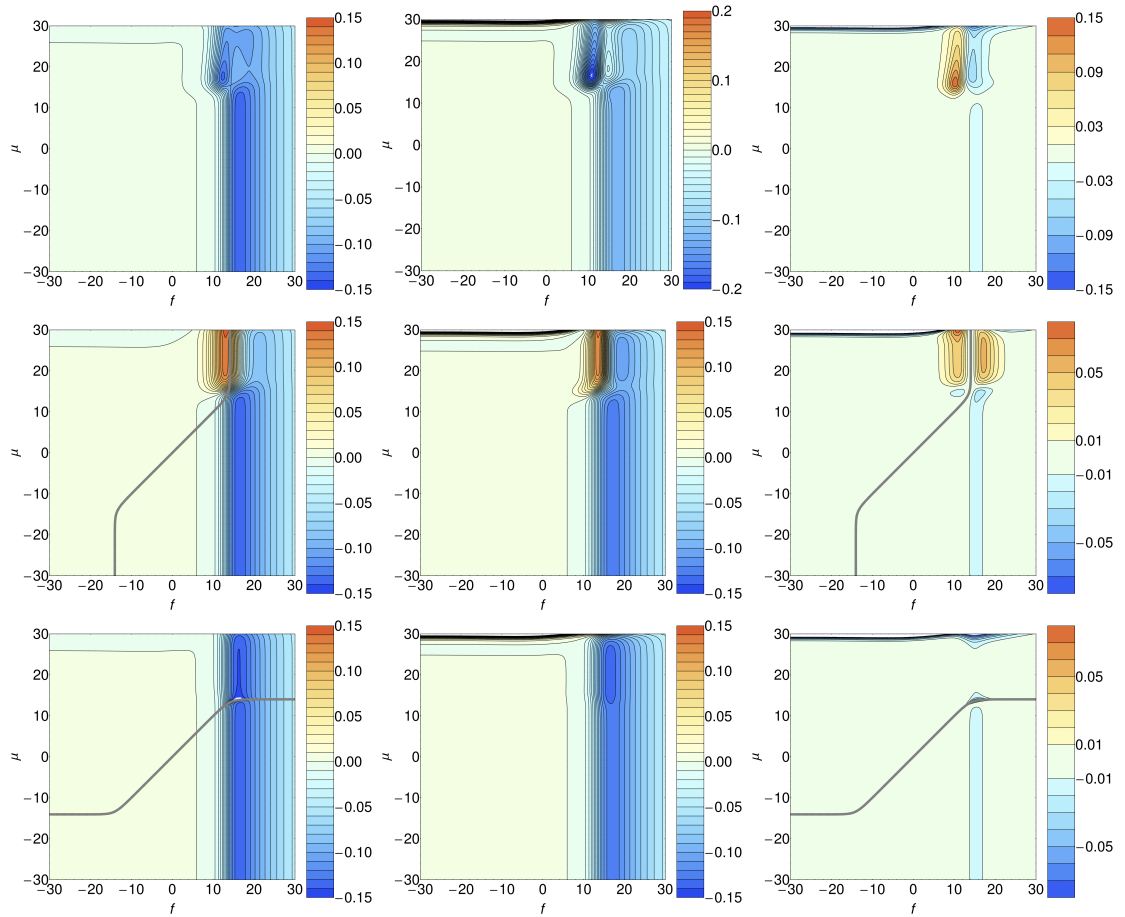
##### Comparison to the minimal model

We have already seen in Figure 6.7 that the phase diagram of a minimal four-state model agrees very well with the phase diagram of the more involved six-state model. In order to see whether we also have an agreement in the more detailed structure, we consider relative errors between the models. For any quantity  $X$  that takes the values  $X_6$  and  $X_4$  in the six and four-state model, respectively, we define the relative error  $\delta X := \frac{X_4 - X_6}{X_6} = \frac{X_4}{X_6} - 1$ . For our case,  $X$  represents the (decadic logarithms of the) first and second cumulants as well as the SNR. For these observables, we have discussed  $X_6$  already in Figures 6.8 and 6.9.

Figure 6.10 shows the relative errors  $\delta X$  for those quantities. In the entire region depicted in the phase diagram, the relative error is bounded below approximately 15%. This alone is a remarkable result, given the fact that the values range over more than twenty logarithmic decades.

Even more spectacular is the fact that in a large part of the phase-diagram the mismatch is even less than one percent. Especially for the SNRs this low error is achieved almost everywhere. An exception to this very low error can be observed in the vicinity of the *stalling force*, which is defined by the force necessary to stop the motor from moving. For physiological chemical conditions, *i.e.* a (non-dimensionalized) chemical potential difference  $\mu$  between 20 and 30, we observe mismatches of around 10 to 20%. This is well within the expected uncertainty in the experimental data both models are built on, *cf.* Ref. [CC05]. In conclusion, we see that a four-state model is as effective as a more involved six-state model — if the physics are the same.

## 6. Modelling molecular motors



**Figure 6.10.:** Relative errors  $\delta X := \frac{X_4 - X_6}{X_6}$  for the quantities  $X_4$  calculated in the minimal four-state model with respect to their values  $X_6$  in the six-state model. From top to bottom we show the errors for the dissipation  $\mathcal{E}$ , the displacement  $d$  and the hydrolysis  $h$ , respectively. The columns (left to right) correspond to the first cumulant  $c_1$ , the second cumulant  $c_2$  and the SNR  $c_1/c_2$ . The relative error for all quantities is bounded below 20% everywhere. In large parts of the phase diagram, it is less or of the order of a couple of percent.

### Comparison to the coarse-grained model

We omit showing similar plots for the five- and four-state models obtained using our approach to model reduction presented in Section 6.2. Qualitatively the results are the same, as we would expect from the good agreement for a typical set of physiological parameters shown in Fig. 6.5. Further, we have no physical arguments that would prefer the reduction of one particular bridge state over any other.

Moreover, note that *by construction* our coarse-graining procedure preserves the currents on all chords and hence in the entire network. Then, the Schnakenberg decomposition (5.40a) ensures that the first cumulant of *any* physical observables agrees between the models. Or differently put: The relative error  $\delta c_1$  identically vanishes for any physical observable.

## 6.4. Discussion

We close this chapter with a discussion of our results. At first, we give some additional remarks on the role of the SNR obtained as the ratio  $c_1/c_2$  of the first two (scaled) cumulants of a physical observable. In particular, we show how it reveals additional structure hidden in the phase diagram of non-equilibrium systems modelled by the means of ST. After that, we formulate “take-home messages” regarding the formulation of Markovian models for small (biological) systems.

### 6.4.1. The significance of the SNR

In the following we consider the SNR, *i.e.* the ratio  $\frac{c_1(g)}{c_2(g)}$  of the first and second cumulant of a physical observable  $g \in \mathcal{O}$ . First, we motivate why the SNR reveals more of the hidden structure contained in phase diagrams of non-equilibrium situations. After that, we briefly discuss its role in the theory of non-linear response. Finally, we discuss the special role of the motance  $\mathcal{E}$  for arbitrary dynamically reversible Markov processes and that of the displacement for molecular motors performing a linear motion.

#### Revealing hidden structure

We have seen that plotting the SNR shows features which are not directly visible in the plots of the cumulants. We observed that the order of magnitude of the first and second cumulants are roughly the same throughout the phase diagram. This fact can be easily understood from the scaling of the SCGF with respect to the transition matrix.

To that end, note that the thermodynamic balance conditions (6.10) ensure that  $\exp(\mathcal{E}_\omega^\omega)$ , *i.e.* the ratio of forward and backward transition rates scales with the driving forces. Disregarding correct mathematical notation, one can say  $\mathbb{W} \propto \exp|\mathcal{E}|$  where  $|\mathcal{E}|$  reflects (the magnitude) of external driving. It is an easy mathematical exercise to show that the SCGF and hence the entire fluctuation spectrum is homogeneous of first order in  $\mathbb{W}$ . Hence, a rescaling of  $\mathbb{W}$  by a common factor  $\exp|\mathcal{E}|$  amounts to a rescaling of all the cumulants by

## 6. Modelling molecular motors

the same factor. Plotting the SNR (and thus hiding the general exponential trend), we are more sensitive to the detailed structure of the phase diagram (like hidden symmetries, *cf.* the next paragraph).

### The role of the motance

As an example for the additional structure found in the SNR, consider Figure 6.8, where we show our results for the dissipation. The observable corresponding to the dissipation is the motance  $\mathcal{E}$ . The latter plays a special role: It solely depends on the transition rates of the stochastic process and thus is a *neutral* observable: One only requires that the Markov process under consideration is dynamically reversible, but not necessary that it constitutes a *model* for any physical system. It is well-defined by the transition rates for any dynamically reversible Markov process used in ST, *cf.* also Ref. [Pol12].

In Section 6.3.3 we have seen that for moderate values of the driving there is an (approximate) reflection symmetry at the first diagonal of the phase diagram. We argued that this symmetry originates from the quasi-tight coupling of kinesin's chemical and mechanical transitions. However, this symmetry is broken for sufficiently high driving parameters. We explained this fact with a change of kinesin's dominant cycles, *cf.* also Refs. [Alt+12] and [Hil79].

The SNR for the dissipation exhibits this symmetry also for much higher values of the driving. Moreover, we find an additional second mirror symmetry in the phase diagram. We have the following explanation for this fact: The symmetry along the first diagonal still expresses the tight-coupling of ATP hydrolysis with forward stepping. The existence of the additional symmetry arises from the existence of a second fundamental cycle. More precisely, it indicates the existence of two loosely coupled fundamental cycles. In the case of completely uncoupled cycles, the symmetry  $\mathcal{E}_\alpha \mapsto -\mathcal{E}_\alpha$ ,  $J_\alpha \mapsto -J_\alpha$  is obeyed for each individual cycle  $\alpha$ . As kinesin's cycles *are* coupled, the symmetry only holds approximately.

### Displacement, drift and diffusion

Many molecular motors perform a one-dimensional motion along a linear track. Kinesin is just one example, but there are many others like dynein or RNA-polymerase [CMB08]. In Markovian models, one can always consistently identify the step length (which may be zero) associated to any transition.

As for kinesin, one can then write the displacement  $d \in \mathcal{O}$  as an anti-symmetric observable. Usually one is interested in the drift velocity  $V$  and the diffusion constant  $D$  of the motor. The drift velocity  $V = c_1 d$  is just the first cumulant of the displacement. The diffusion constant is, up to a factor of two, the second cumulant, *cf.* also Ref. [Der83]. This fact is directly visible from the scaling of the *mean square displacement*  $\kappa_2$  along a

trajectory  $\underline{\omega}^{(\tau)}$  of run-length  $\tau$ , which amounts to

$$l^2 c_2(\eta_5) = c_2(d) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \kappa_2 \left( \tau \bar{d}^{(\tau)} \right) = \lim_{\tau \rightarrow \infty} \frac{l^2}{\tau} \kappa_2 \left( \sum_{e \in \underline{\omega}^{(\tau)}} \eta_5(e) \right) = 2D,$$

where  $e \in \underline{\omega}^{(\tau)}$  sums over the directed edges passed by  $\underline{\omega}^{(\tau)}$ . There is plenty of literature on the calculation of velocity  $V$  and diffusion constant  $D$  in Markovian models [CMB08; BH12]. Our approach generalizes the treatments presented in these references. In our method, one does not need to bother with the combinatorial complexity due to the topology of the graph: It is hidden in the coefficients  $a_n$  of the characteristic polynomial of the tilted matrix  $\mathbb{W}_d(q)$ , cf. Equation (5.34).

Moreover, for the kinesin model, the displacement  $d$  constitutes an example where the choice of an appropriate spanning tree can simplify calculations, cf. the discussion in Sec. 5.4.3. Recall that the choice of spanning tree defines the fundamental cycles. In the present case, the fundamental cycles  $\zeta_2$  and  $\zeta_5$  correspond to the dissipative slip and the forward cycle, respectively. The former involves only chemical transitions and hence  $d$  vanishes on  $\zeta_2$ . Consequently,  $\eta_2$  does not appear in its chord representation  $d_{\mathcal{J}^c} = d \equiv l\eta_5$ . Note that this is not the case for the spanning trees depicted in Figure 5.4b/c. In that case, the fundamental cycles are the forward and backward cycle, which both involve a mechanical transition.

Another interesting quantity is given by the inverse of the SNR for the displacement  $d$ : It yields the typical length scale above which drift dominates diffusion. In the more general context, the inverse SNR is also known in the literature as the *Fano factor*. It was introduced originally for particle detection in high-energy physics [Fan47].

Recently, the notion of a Fano factor has also been used in the context of stochastic transport and chemical systems [RDD05; Fan+10; QK14]. As the inverse of the SNR, it diverges when the signal (expressed by the first cumulant) in the denominator passes through zero. In contrast, the SNR seems to show no singularities indicating that the denominator, *i.e.* the second cumulant is always positive.

Interpreted as a transport coefficient like the diffusion constant, the positivity of the second (self-)cumulant  $c_2(\varphi)$  of a physical observable  $\varphi$  is clear: A negative diffusion constant (indicating a negative mean *square* displacement) is not possible. However, there is a caveat to this statement regarding the origin of the phase diagram, where all driving forces vanish: At equilibrium, the symmetry imposed by detailed balance ensures that *all* scaled cumulants vanish identically, due to the fact that the SCGF  $\lambda_\phi(q) \equiv 0$  is constantly zero. However, this non-generic feature is a mathematical peculiarity that has its origin in the formulated of the theory. Outside of (and arbitrarily close to) equilibrium, transition-rate independent physical observables  $\phi$  have a positive second moment and the SNR is well-defined, cf. also Figures 6.8 and 6.9.

### Linear and non-linear response in ST

The lack of divergence of the SNR even at thermodynamic equilibrium (where both cumulants vanish) is evident from the interpretation of the second cumulant as a *transport coefficient* or generalized susceptibility. Let us briefly comment on this interpretation. A more detailed discussion can be found in the (yet unpublished) manuscript [WAV p].

For any observable, the second cumulant has a direct interpretation in terms of response theory. In particular, the SNR allows us to characterize fluctuation-dissipation relations in situations far from equilibrium. In Section 2.4.1 we have discussed the Einstein relation, which connects the strength of a transport coefficient (in that case, the mobility or inverse drag  $\zeta$ ) to the strength of fluctuations (in that case given by the diffusion constant). The general relation is formulated in Equation 2.26, which states that the noise correlations amount to the elements of the mobility matrix multiplied by temperature and a numerical factor of 2. It is a consequence of the theory of linear response.

For abstract ST, the mobility is given as the derivative of the first cumulant  $c_1$  with respect to a driving force. Recall that all physical currents can be obtained as linear combinations of the currents associated to the family  $(\eta_\alpha)_{\alpha \in \{1,2,\dots,|\mathcal{A}|\}}$  of fundamental chords. The associated abstract driving force is the motance  $\mathcal{E}_\alpha$  of the fundamental cycle  $\zeta_\alpha$ . Consequently, we define the mobility matrix:

$$\mathbb{M}_{\alpha,\beta} := \frac{\partial c_1(\eta_\alpha)}{\partial \mathcal{E}_\beta} \quad (6.12)$$

Note, that  $\mu_{\alpha,\beta}$  amounts to an entry of a generalized Onsager matrix  $\mathbb{L}$ , cf. Eq. (2.22).

The fluctuation-dissipation theorem [CS98] ensures that close to equilibrium we have linear response, *i.e.*

$$\mathbb{M}_{\alpha,\beta} = 2c_2(\eta_\alpha, \eta_\beta). \quad (6.13)$$

Moreover, we can apply the fluctuation relations discussed in Sec. 2.4.3 to determine the region in phase space where we expect this relation to hold: In the limit of vanishing driving, the distribution of the dissipation is a centred Gaussian one. The fluctuation relations [LS99; Sei05] for the steady-state dissipation then ascertains that the scaled variance of the limiting distribution must approach two times the mean. Hence, a value close to  $\frac{1}{2}$  in the SNR of the dissipation  $\mathcal{E}$  amounts to the region of linear response. Results in the same spirit have been obtained for deterministic dynamics [Gal98; Rue99]. For the current set-up, linear response has been discussed in a paper by Lebowitz and Spohn on the fluctuation relation for stochastic dynamics [LS99].

#### 6.4.2. Consequences for models

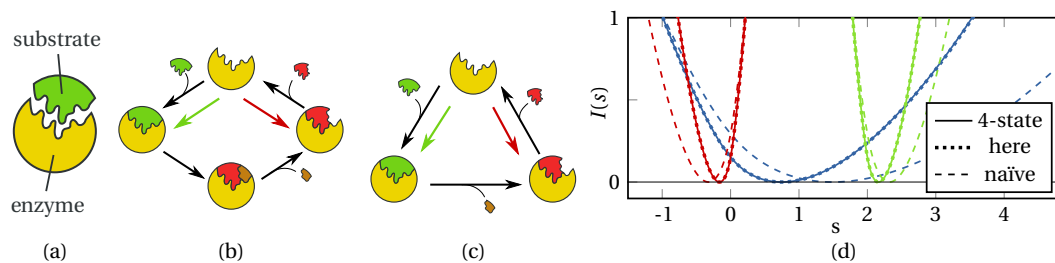
Equations (6.10) relate the affinities of cycles to (linear combinations) of the thermodynamic forces. They relate the microscopic rates to the macroscopic conditions found in

the medium. Hence, they must hold in any model describing the same physical situation, independent of the mesoscopic resolution.

However, thermodynamic balance is actually a *thermostatic* statement about (local) equilibrium distributions. As such, they are insufficient to specify any kinetic, *i.e.* dynamic properties of the model. The observable dynamic quantities are the average currents associated with physical observables. The Schnakenberg decomposition (5.40a) ensures that these currents are completely determined by the average probability currents flowing on the fundamental chords.

Hence, for thermodynamic consistency on the level of the (average) thermodynamic currents one must always take the probability currents into account. If one fails to do so while inferring a simpler from a more complex model, one gets inconsistent results. The next paragraph discusses an example.

### Naïve coarse-graining of bridge states



**Figure 6.11.:** a) Schematic representation of an enzyme that binds a substrate. b) A simple one-cycle catalysis model. Upon binding the substrate, it is catalytically split into two subunits. The release of the small subunit happens almost instantly, whereas the large subunit stays attached for a longer time. c) A reduced three-state model. Catalytic splitting and release of the small subunit yield a combined transition. d) Rate functions for the dissipation (blue, center), the association of the large subunit (red, left) and the association of the substrate rate (green, shifted to the right by  $s_0 = 2$ ). As a parametrization we choose all transition rates as unity, with exception of the fast transition  $w_D^{\ominus} = 100$  for the release of the small subunit. The method presented in Sec. 6.2 preserves fluctuation to a high degree. The naïve choice Eq. (6.15), however, results in severe changes even in the first cumulant.

The catalytic cycle of ATP hydrolysis on kinesin's active sites is an example of a general enzymatic activity. Figure 6.11 shows the catalytic splicing of a *substrate* molecule into two subunits at the active site of an enzyme. We assume that one subunits are always released in the same order, for instance due to steric effects. Often the substrate splits into a small and a large subunit, similar to the case for the hydrolysis of ATP in P and ADP. The small subunit is released immediately after splicing, whereas the larger part stays bound to the enzyme for a longer time.

In Ref. [LL08], the authors consider the catalytic cycle at kinesin's active sites. Denote by  $D$ ,  $T$  and  $E$  the states where ADP, ATP and nothing, respectively are bound to the active site. In addition, we use the symbol  $\Theta$  to denote the state where ATP is already split but

## 6. Modelling molecular motors

both ADP and P are still attached to kinesin's active site.

Kinesin provides an example for a very fast of the small subunit P. Consequently, it is natural to remove the (bridge) state  $\Theta$  in a coarse-graining procedure. For the rates of the new transition connecting  $T$  and  $D$  the choice

$$\frac{w_{D'}^{T'}}{w_{T'}^{D'}} = \frac{w_{\Theta}^T w_D^{\Theta}}{w_T^{\Theta} w_{\Theta}^D} \quad (6.14)$$

preserves the affinity of the single cycle, if all rates are left constant.

For three linearly connected states Hill proposed to choose

$$w_{D'}^{T'} = \frac{w_{\Theta}^T w_D^{\Theta}}{\langle \tau_{\Theta} \rangle}, \quad (6.15a)$$

$$w_{T'}^{D'} = \frac{w_T^{\Theta} w_{\Theta}^D}{\langle \tau_{\Theta} \rangle}, \quad (6.15b)$$

if the staying time  $\langle \tau_{\Theta} \rangle$  is small [Hil77]. However, a linear chain always fulfils detailed balance, and hence does not carry any currents.

Out of equilibrium one has to be careful with using Eq. (6.15) as a prescription for the new rates. Generically, adapting these rates for the new transition while leaving all others constant massively changes the currents running through the network. Consequently, we expect the entire fluctuation spectrum to change rather drastically, which is in fact the case.

In Figure 6.11d the solid lines show the rate function obtained for different physical observables in a simple parametrization of the enzymatic reaction network depicted in Figure 6.11b. Dashed and dotted lines correspond to reduced three-state models, cf. Figure 6.11c. More precisely, the dashed lines show the naïve choice prescribed by Eq. (6.15). Similarly, the dotted lines amount to a reduced model obtained with our fluctuation-sensitive coarse-graining procedure.

The change in the currents is generic for the choice given by (6.14), if the other transition rates are *unmodified*. Recall that in the heuristic derivation in of our coarse-graining algorithm Sec. 6.2 we demanded locality. Locality amounts to the requirement that the steady-state probabilities are not changed in the part of the network which is unaffected by the coarse-graining. However, locality cannot be achieved, unless some other transition rates are changed.

Though our fluctuation-sensitive uses further constraints than only locality, the moral from the data presented Fig. 6.11 is the following: The thermodynamic balance conditions 6.10 for the (fundamental) cycles of the network ensure thermodynamic consistency, which is a *static* requirement. A model's dynamic properties are characterized by the steady-state currents of physical (and hence in principle measurable) currents. The generalized Schnakenberg decomposition (5.40a) for their expectation values relates them to the steady-state probability currents  $J$ . The preservation of the currents of *all* physical observables is given if probability current  $J$  agrees on a fundamental set of chords.



Consequently, the static and dynamic requirements specify  $2|\mathcal{H}|$  independent constraints, where  $|\mathcal{H}|$  is the cyclomatic number. Any two models for the same physical situations, should (at least approximately) obey these constraints. A direct consequence of these constraints is the preservation of the dissipation in the steady state. We will discuss the significance of that statement in more detail in the final Section 7.3.

### Some modest advice for model construction

A main result of the present chapter is that a four-state model is sufficient for the modelling of kinesin. More precisely, we need a four-state model that shares the same cycle topology than the original model. We have also seen that the cycle topology determines how different currents are coupled to each other. A non-tightly coupled model thus always needs at least two (independent) cycles.

The minimal four-state model constructed in Appendix B qualitatively captures all the features of the more complicated model. In the parametrization of the model we followed the same physical arguments than Liepelt and Lipowsky adapted for constructing their model in Ref. [LL07]. The quantitative agreement between the models is remarkable: Relative errors of a few percent are typically much smaller than the experimental uncertainty in the measurement of the (chemical) transition rates.

Hence, we propose the following paradigm for constructing good models: At first, determine how experimentally accessible currents couple to the external driving forces. To that end, try to measure or estimate the coupling matrix  $L_{\alpha\beta}$  between different measurable currents and external forces.

After that, design a network of (observable) states where changes between states can be attributed to observable currents. This step is the hardest: it requires a certain degree of physical, chemical or biological knowledge or intuition. The coupling matrix  $L_{\alpha\beta}$  will be useful in that regard.

After the construction, reduce the network to the minimal topology by removing all bridges and leaves. The topology of the network then tells you about the structure in the *abstract* matrix  $\mu_{\alpha,\beta}$  introduced in Equation (6.12). The physical forces and observable currents which are assigned to transitions between states connect the measured matrix  $L_{\alpha\beta}$  to  $\mu_{\alpha,\beta}$ . This can be used as another consistency check.

Finally, try to measure, predict or by means of thermodynamic balance (6.10) infer the transition rates between states. Symmetry considerations may help in that approach. When applying the thermodynamic balance, make sure that you have a *physical reason* why to use the constraint to infer a particular transition and not another.<sup>1</sup>

A (Markovian) model constructed in this way minimizes the amount of additional assumptions. It is thus maximally non-committal with respect to our missing knowledge.

<sup>1</sup> We stress this here because of the following reason: In both Ref. [LL07] and in App. B one *chooses to infer* the rate for the transition modelling ATP release from the trailing active site by thermodynamic balance. If any other rate on the backward cycle is inferred in that way rather than determined by symmetry, the physics described by the model are different. The application of our method then shows that in that case other cycles become dominant for high values of the driving, and the phase diagram changes drastically.

## 6. Modelling molecular motors

Moreover, one can use this paradigm in order to systematically *test* assumptions about small (biological) systems in fluctuating environments. Though not more than a sketch, we hope that the ideas outlined here may guide scientists to design more and more accurate models. In a sense we are back with Schrödinger's quote: We hope that our abstract physical-mathematical considerations might contribute to improving our understanding of life — if even just a tiny bit.

### 6.5. Summary

Fluctuations, *i.e.* a behaviour away from the expected average, are non-negligible in small systems in thermodynamic environments due to the influence of thermal noise. For biological systems, which have evolved over millions of years, such fluctuations are often functional. The molecular machinery of life, which includes so-called molecular motors, ensures the function of living cells. ST can be understood as a paradigm for designing models of such systems as well as for their thermodynamic interpretation.

Static properties of a model are prescribed by the thermodynamic properties in the medium. The requirement of thermodynamic balance manifests in the Hill–Schnakenberg conditions, *cf.* Eq. (6.10). The dynamic properties of physical currents on an ensemble level are completely determined by the probability currents running in the network. The latter in turn are determined by the currents on a set of fundamental chords, reflected in the Schnakenberg decomposition (5.40a). Together these requirements yield a set of consistency criteria regarding the asymptotic ensemble properties of different models.

The (approximate) preservation of fluctuations between two models requires additional conditions. A heuristic motivation of the latter yields the fluctuation-sensitive coarse-graining procedure presented in Section 6.2. Additional constraints can also be obtained as the result of experimental data. Minimal models are constructed using thermodynamic consistency and all other available information, while ensuring the simplest required cycle topology of a network.

A main emphasis of this chapter was on the application of our ideas to molecular motors. Using the tools from the previous Chapter 5, we compared an established model of the motor protein kinesin with simplified models. Our results are remarkable: Our reduced models captured the fluctuations of the original model to a very high degree. Mismatches in relevant quantities are bounded below 15% for all the values of the external driving we considered. This is even more spectacular given the fact that the *values* of observable currents change over about twenty logarithmic decades.

In the discussion, we were concerned with the physical relevance of the SNR of an observable, *i.e.* ratio of its first and second cumulant. We mentioned its role in thermodynamic response theory and showed how the SNR of the motance can be used to infer the linear response regime. We further discussed how our results simplified previous approaches to the calculation of drift and diffusion in models of molecular motors which perform a one-dimensional motion. Finally, we formulated a set of ideas to be applied in

the future design of models.

At this point, we conclude the main part of the present work. We give a summary and a final discussion in the following final Chapter 7. In addition, we also give an outlook on further research and take a final bird's eye perspective on the topics discussed in the present thesis.



## 7. Conclusion and outlook

“ Computers may be thought of as engines for transforming free energy into waste heat and mathematical work. ”

---

C. H. Bennett, *The thermodynamics of computation*, 1981

The present chapter intends to bridge the gap to our introductory remarks in Chapter 1. In Section 7.1 we start by telling the reader what we have told him in the previous Chapters 2–6. This summary can be understood as a brief synopsis of the discussions and summaries that were provided at the of each of these chapters.

After that, Section 7.2 provides a review as to how our results are placed amongst the current literature — at least from the author’s knowledge of the latter. As an outlook, we discuss promising directions for follow-up work.

Finally, we conclude this thesis with an author’s perspective in Section 7.3. Instead of focussing on the rigorous results obtained so far, this personal perspective provides the “bigger picture” as perceived by the author, who views complex systems as information processing devices — as already hinted at in the epigraph of the present chapter.

### 7.1. Summary

Chapters 1 and 2 provided the necessary background for this thesis. They reviewed the state of the art of modern statistical physics in general and ST in particular. Even though many fascinating aspects of this active field of research could not be covered, the cited references should provide a good starting point for further reading.

The main part of the thesis split into two major parts. Chapters 3 and 4 discussed deterministic microscopic foundations of ST. In Chapters 5 and 6 we focused on the structure of Markovian ST on finite state spaces.

Before discussing anything new, let us summarize what we have discussed so far.

#### 7.1.1. Microscopic foundations of stochastic thermodynamics

##### The foundations of Markovian stochastic thermodynamics

A crucial assumption for statistical physics in general and stochastic thermodynamics in particular is the *Markovian postulate* [Pen70]. It states that observable time-series

## 7. Conclusion and outlook

corresponding to measurements on thermodynamic systems obey Markovian statistics. It is equivalent to the claim that observable states do not contain any memory of their past.

From the point of a deterministic evolution for observationally inaccessible microscopic states, this is a highly non-trivial statement. In thermodynamics, the Markovian postulate is closely linked to the hypothesis of local equilibrium (LE). The latter assumes that on the time-scale of observations, the distribution of the microscopic degrees of freedom has relaxed to a constrained equilibrium ensemble — at least for all practical purposes.

In Chapter 3 we investigated the microscopic implications of the Markovian postulate. We considered deterministic dynamics  $\Phi$  in discrete time which map phase space  $\Gamma$  bijectively onto itself. An observable  $M: \Gamma \rightarrow \Omega$  maps microscopic states  $x \in \Gamma$  to a finite number of observable states  $\omega \in \Omega$ . The observable thereby partitions phase space into equivalence classes  $\mathcal{C}_\omega$  which are indexed by the values of the measurement result.

An observable time-series  $\underline{\omega}^{(\tau)} = (\omega_0, \omega_1, \dots, \omega_\tau)$  of length  $\tau$  summarizes subsequent observations. Different microscopic initial conditions generically yield different time-series. The main result of the chapter states the requirements on the observable  $M$  and a probability distribution  $\rho$  for the microscopic states, such that the time-series  $\underline{\omega}^{(\tau)}$  obey Markovian statistics.

We found that for a given dynamics  $\Phi$  and an appropriate observable  $M$ , there are many distributions  $\rho$  that fulfil this requirement. Generically, these distributions are non-stationary and yield Markovian statistics only after some distinct point in time  $t_0$ , which corresponds to the time of the preparation of the system. This statement provides an example for an operational interpretation of our abstract results. In addition, we discussed our results from the perspective of modern ergodic theory.

### Stochastic thermodynamics as a theory of statistical inference

In 1957, Jaynes proposed the view of statistical mechanics as a theory of statistical inference, formulated in the language of information theory [Jay57]. Indeed, his ideas provide a self-consistent foundational framework, which formalizes Gibbs' approach to statistical ensembles [Gib48]. However, he stresses that the probability densities  $\rho$  which characterize phase space ensembles must not be interpreted in a frequentist way. Rather, they should be understood as maximally non-committal statements about our expectations of the probability of microscopic states — given our previous knowledge about the microscopic physics that govern their dynamics.

In Chapter 4 we propose a tentative information-theoretic framework for ST. The formal background is provided by the mathematical treatment of observed time-series in Chapter 3. We use Jaynes' notion of maximum-entropy priors to infer the distribution of microstates  $x \in \mathcal{C}_\omega$ .

From these priors we construct the coarse-grained ensemble  $\rho_t^{\text{cg}}$  to express our expectation about the distributions of microstates  $x \in \Gamma$ . Subsequent measurements provide new information about the state of the system. Consequently, we obtain an

update rule that specifies the temporal evolution of  $\rho_t^{\text{cg}}$ . The coarse-grained ensemble is inferred without a detailed knowledge of the microscopic dynamics  $\Phi$ . However, a given dynamics  $\Phi$  specifies a deterministic evolution rule for the initial ensemble, which is expressed by the fine-grained ensemble  $\rho_t^{\text{fg}}$ .

In our information-theoretic perspective, the uncertainty expressed in the coarse- and fine-grained ensemble is quantified by their differential entropy. Their relative entropy, *i.e.* the Kullback–Leibler divergence of  $\rho^{\text{fg}}$  from  $\rho^{\text{cg}}$  quantifies the mismatch between the microscopic and an inferred description based on coarser models or measurements. We found that these entropies can be obtained as time-series averages of entropic  $\tau$ -chains, *i.e.* random variables that depend on finite time-series  $\underline{\omega}^{(\tau)}$ . Motivated by (non-equilibrium) molecular dynamics simulation, we related phase-space contraction and dissipation with the entropy change in the hidden, unobservable degrees of freedom. A corollary is the identification of the relative entropy with the total entropy.

As an example for the application of the mathematical framework we introduced network multibaker maps (NMBM) as a versatile model dynamics. Reversible NMBM share the mathematical properties of the equations of motion used in molecular dynamics. We explicitly showed how ST emerges in the context of NMBM. After a discussion of the results of Chapter 3 in the light of NMBM, we postulate that our results apply more generally. In particular, we conjecture that under the right assumptions, ST naturally emerges from physical microscopic dynamics, *i.e.* dynamics that are time-reversible with a measure-preserving time-reversal involution.

### 7.1.2. Structure and models of stochastic thermodynamics

#### Kirchhoff's laws, cycles and fluctuations

In Chapter 5 we dealt with the structure of Markovian jump processes on finite state spaces. The topology of the network of states is determined by the transition probabilities and can be visualized as a graph. In the steady state, probability currents on that graph resemble electrical currents in an electrical circuit. For the latter, Kirchhoff's first and second law state that (i) the current balances at each vertex and (ii) that the integrated difference of a reference voltage vanishes along a loop. We find that Kirchhoff's laws equally apply for Markov processes and provide a complete electrical analogy.

The reason for the applicability of Kirchhoff's results is the algebraic structure of the network. In an abstract sense, electrical currents and voltage drops are anti-symmetric observables defined on the edges of a graph. In the context of ST, ensemble averages of anti-symmetric observables correspond to physical currents.

For such physical observables algebraic graph theory provides powerful tools for abstract structural investigations. For instance, the abstract space of all physical observables  $\mathcal{O}$  decomposes into two orthogonal components  $\mathcal{Z}$  and  $\mathcal{Y}$ . The former contains so-called *cycles*, whereas the latter contains the *co-cycles*. Cycles are the analogue of divergence-free currents in field theory, hence they correspond to the currents of conserved quantities.

## 7. Conclusion and outlook

Co-cycles are obtained as the discrete gradients of potentials  $\mathcal{U}$  defined on the vertices of the graph. An abstract formulation of Kirchhoff's first law holds for all abstract cycles  $z \in \mathcal{Z}$ , whereas Kirchhoff's second law generalizes to observables  $y \in Y$ .

We are particularly interested in the fluctuations of physical currents, *i.e.* the stochastic deviations from their averages. In mathematical terms, we consider the probability distribution associated to the finite-time averages  $\overline{\varphi}^{(\tau)}$  of physical observables  $\varphi \in \mathcal{O}$ . A time-average is obtained from integrating  $\varphi$  along the transitions defined by a stochastic time-series  $\underline{\omega}^{(\tau)}$ . For ergodic Markov processes, time-averages converge to a distinct value  $\langle \varphi \rangle_\infty$  in the asymptotic limit  $\tau \rightarrow \infty$ . This value agrees with the steady-state ensemble average of the corresponding physical current.

Consequently, the asymptotic distribution is a  $\delta$ -peak. For any finite time  $\tau$ , the probability distribution of  $\overline{\varphi}^{(\tau)}$  has a finite width. For Markov processes, its convergence to the  $\delta$ -distribution is governed by a large deviation principle, which amounts to a scaling form characterized by a *rate function*  $I_\varphi(x)$ .

The rate function is completely determined by a set of *scaled cumulants*  $c_n(\varphi)$ , which we call the *fluctuation spectrum* of  $\varphi$ . Two of the main results of Chapter 5 regard the latter: Firstly, we showed that the fluctuation spectrum of  $\varphi = z_\varphi + y_\varphi$  only depends on its component  $z_\varphi \in \mathcal{Z}$  in cycle space. Secondly, we demonstrated that the spectrum of any physical observable  $\varphi$  is completely determined by the cumulants of the probability currents.

The importance of cycles in the network of states of Markovian processes yields several cycle decompositions. Probably the most well-known is the Schnakenberg decomposition for the average steady state probability currents. We generalized this decomposition to the entire fluctuation spectrum of arbitrary physical observables. Moreover, we introduced the chord representation  $\varphi_{\mathcal{X}}$  of  $\varphi \in \mathcal{O}$  and explain its benefits for the purpose of calculations.

### Models of molecular motors

In Chapter 5 we developed the necessary tools for the quantification of *fluctuations* in models of small systems under non-equilibrium conditions. Chapter 6 applied these methods to models of actual (bio-)physical systems. Currently, the study of molecular motors is attracting a lot of interdisciplinary attention. In biology, fluctuations are not only relevant but often even important for the *function* of living systems. In quantifying fluctuations, statistical physics provides biologists and modellers with new tools for the study of life.

More precisely, our methods help us to separate good from less good models. The minimal requirement on any model used in ST is that the calculated average currents agree with ensemble measurements. Consequently, two different models of the same physical system must in this respect also agree *with each other*. The results of the previous chapter emphasized the importance of the topology of cycles in the network of states. The



latter is characterized by the connections of a set of (fundamental) cycles. The motance or affinity of a cycle amounts to the logarithm of forward to backward transition rates integrated along its edges. Physically, the motance of any cycle is a linear combination of the *thermodynamic forces* that drive the system out of equilibrium. This thermodynamic balance requirement is also known as the Hill–Schnakenberg conditions.

The results of Chapter 6 can be summarized as follows: The ensemble behaviour of a model is fully characterized by the topology of its fundamental cycles together with their affinities and associated steady-state currents. We say that two models are equivalent on the ensemble level, if they share these properties.

Based on this *necessary* requirement, we presented a coarse-graining algorithm for Markov processes. In addition to the currents  $J$  and affinities, this approach additionally preserves the fluxes  $\phi$ . With these additional constraints, we find that in addition to the *exact* preservation of the averages of any physical observable, also its fluctuations are preserved to a very good degree.

In order to be concrete, we exemplified our abstract results using a well-known model for the molecular motor *kinesin*. We explore its phase-diagram, *i.e.* its dependence on chemical and mechanical forces that drive the system out of equilibrium. In that context we illustrate the ability of our coarse-graining procedure to preserve fluctuations. We further construct a minimal model for kinesin based on the same data and the same physical assumptions used in the construction of the original model [LL07]. Remarkably, the relative error of the simplified versions with respect to the original one is only a few percent throughout the phase diagram — in most parts it is even much lower.

In comparison with the typical uncertainties in an experimental measurements of kinetic rates, this mismatch is negligible. In conclusion, we propose a guiding principle for the construction of physical models from available data.

## 7.2. Outlook

### 7.2.1. Possible generalizations

In the first part of the present thesis we discussed the microscopic deterministic foundations of ST. For simplicity, we assumed a stroboscopic picture, *i.e.* a dynamics evolving in discrete time. Much of the ergodic theory for measurable dynamical systems was first formulated using discrete-time maps and later generalized to flows. For instance, the SRB measure was first formulated for Anosov maps [Sin68] and then extended for continuous Axiom-A flows [Sin72; BR75]. We expect that this is also the case for our set-up. However, given the limited temporal resolution of any real experiment, the discrete-time case might be more natural.

We further demanded discreteness (in fact finiteness) of the space of observations, *i.e.* the space of possible measurement outcomes. In that case, an observable time-series is generated by a jump process. In Section 2.4 we have briefly considered ST for continuous

## 7. Conclusion and outlook

state spaces using Langevin and Fokker–Planck equations. Consequently, an extension of the algebraic framework introduced in Chapters 5 and 6 to continuous state spaces is desirable, but seems rather involved.

In contrast, we expect the formulation of the results in Chapters 5 and 6 for the case of discrete *time* to be straightforward. The algebraic-topological treatment in Chapter 5 is independent of the temporal evolution. Also the time-discrete formulation of the large deviation principle resembles the continuous case. The main difference is that the SCGF is obtained as the largest eigenvalue of the tilted stochastic matrix, rather than as the logarithm of the corresponding tilted transition matrix [Tou11].

### 7.2.2. Network multibaker maps as a versatile tool for ergodic theory

Let us briefly comment on the role of network multibaker maps (NMBM) as a model dynamics. In our opinion, NMBM constitute a sufficient representation of the dynamics between elements of Markov partitions for generic uniformly hyperbolic maps. Consequently, they can provide a good pictorial representation of any system that satisfies the “Chaotic Hypothesis” of Gallavotti–Cohen [GC95].

This is interesting for several reasons. Network multibaker maps are both analytically tractable and formulated on a two-dimensional phase space, which can easily be sketched. Moreover, a two-dimensional phase space is sufficient to exhibit transversal stable and unstable manifolds. Transversality is an important aspect of the topology of a Markov partition [Adl98].

Further, NMBM are extremely versatile: Not only can we design them to be (uniformly) conservative or dissipative, but we can also make them time-reversible. Throughout this thesis we have argued that time-reversal is a hallmark of a *physical* microscopic dynamics. In fact, the authors of Ref. [MV03] formulated the need for a uniformly hyperbolic model dynamics with time-reversal. NMBM provide this example.

Finally, for any NMBM which is based on a *simple* graph, the cells representing the vertices form a generating Markov partition. That means that every symbol sequences generated by a Markov jump process has a corresponding phase-space points. After choosing the parameters  $s_j^i$  which define a NMBM, one has full information about its symbolic (equivalent) dynamics. Hence, one immediately knows whether or not to expect certain features of chaotic dynamics (like homo- and heteroclinic orbits) and where to find them in phase space. Further, NMBM constitute a constructive example for Theorem 3.2. of Ref. [BB03] about the existence of a deterministic representation of stochastic cellular automata.

### 7.2.3. Towards a dynamical picture of local equilibrium

We argued how NMBM serve as a representation for systems that fall under the chaotic hypothesis. The chaotic hypothesis implies the existence of a Markov partition  $\mathcal{Q} = (\mathcal{C}_\omega)_{\omega \in \Omega}$  for physical systems. However, the partition  $\mathcal{V} = (\mathcal{V}_k)$  which is induced by a real

physical observable is usually much coarser. Let us assume that each  $\mathcal{V}_k = \bigcup_{\omega \in \Omega_k} \mathcal{C}_\omega$  is comprised of a subset of “microscopic”<sup>1</sup> cells  $(\mathcal{C}_\omega)_{\omega \in \Omega_k}$  indexed by  $\omega \in \Omega_k \subset \Omega$ . In order to distinguish them from the microscopic cells  $\mathcal{C}_\omega$ , we call  $\mathcal{V}_k$  an observable cell. Further, we consider an observable time-scale  $\tau_{\text{obs}} \gg \tau_{\text{mic}}$ , where we use the stroboscopic time interval  $\Delta\tau \equiv \tau_{\text{mic}}$  as the microscopic time-scale.

Transitions  $k \rightarrow k'$  between observable cells  $\mathcal{V}_k$  on observable time-scales are generally non-Markovian. However, there might be situations where “internal” transitions  $\omega_k \rightarrow \omega'_k$  with  $\omega_k, \omega'_k \in \Omega_k$  happen much faster than transitions  $\omega_k \rightarrow \omega'_{k'}$ ,  $k \neq k'$  between different cells. If these time-scales are properly separated, one can approximate observable transitions  $k \rightarrow k'$  on a coarse-grained time scale  $\tau_{\text{obs}} \gg \tau_{\text{mic}}$  as a Markov process.

In Ref. [Esp12], Esposito investigates this situation for Markov jump process in continuous time. He derives a renormalized form for the total entropy production  $\delta S^{\text{tot}} = \delta S_{\text{hom}}^{\text{tot}} + \delta S_{\text{inhom}}^{\text{tot}}$ , which consists of homogeneous and an inhomogeneous term. The homogeneous term  $\delta S_{\text{hom}}^{\text{tot}}$  is formally identical with the usual expression (2.42c) used in ST. The probabilities  $p_k$  appearing in  $\delta S_{\text{hom}}^{\text{tot}}$  are the probabilities of observable states  $k$  rather than the finer microscopic states  $\omega$ . Similarly, the microscopic transition rates  $w_{\omega'}^\omega$  are replaced by the observable transition rates  $V_{k'}^k$ . In the limit of infinite time-scale separation these probabilities become time-independent and homogeneous and thus define the observable Markov process.

However, the inhomogeneous term  $\delta S_{\text{inhom}}^{\text{tot}}$  does not vanish in this limit. The following additional requirements are needed:

1. Internal transitions  $\omega_k \rightarrow \omega'_k$  obey detailed balance.
2. Transitions between different cells  $\omega_k \rightarrow \omega'_{k'}$  happen in a certain regular way.

The “regular way” can be formulated as a time-reversal symmetry for the transitions between microscopic states conditioned on an observable transition.

Network multibaker maps allow us to translate these notions into the terms of an underlying phase space dynamics. A sufficient separation of time-scales allows us to formulate an approximate autonomous fast dynamics  $\Phi_k$  for the microstates  $x \in \mathcal{V}_k$ . Requirement 1) then amounts to  $\Phi_k$  being a conservative map, *i.e.* a map which features an equilibrium distribution as its steady state.

The phase space formulation of requirement 2) is not that obvious. However, the existence of such an additional requirement is already interesting on its own. It emphasizes the requirement of an additional symmetry regarding the microscopic realization of transitions between observable states: To the author’s knowledge, this has not been mentioned anywhere in the literature yet.

In conclusion, we sketched how Esposito’s coarse-graining procedure together with the NMBM perspective provides a possible dynamical picture of LE. Moreover, the inhomogeneous term allows *quantitative* statements about how well the LE hypothesis is satisfied

<sup>1</sup>Note that in this context “microscopic” does *not* refer to a point in the phase space of a deterministic dynamical system, but to an element  $\mathcal{C}_\omega$  of a Markov partition.

## 7. Conclusion and outlook

in a given physical system. Working out the details of this dynamical perspective on LE will be an objective for future work by the present author.

### 7.2.4. Deterministic foundations of stochastic fluctuations relations

In addition to a more detailed picture of LE, our work provides a new perspective on the fluctuation relations, both in the deterministic as well as in the stochastic setting. In order to appreciate this fact, note that dynamically reversible Markov processes and time-reversal symmetric NMBM are equivalent. For the sake of brevity of the following argument, consider a NMBM with cells  $\mathcal{C}_i$  of equal size  $\Pi_i = 1$ . Then, a dynamically reversible rate matrix  $\mathbb{W}$  containing transition probabilities  $w_{\omega}^{\omega'}$ , equivalently defines a reversible network multibaker with relative strip widths  $w_j^i = s_j^i = \hat{s}_i^j$ .

For dynamically reversible Markov processes, a number of fluctuation relations are known. The most general one holds for the total entropy. In the NMBM setting, it translates into a fluctuation theorem for the relative entropy. The change in relative entropy, Eq. (4.26), can be written as an integral over two contributions: One that accounts for the negative of the phase space expansion rate and one that describes the ratio of two inferred densities. Hence, it has the form of a phase space average of the generalized dissipation Evans–Searles dissipation function [ES02]. Thus, we have established a tentative connection between the latter and the general transient fluctuation relation for stochastic dynamics.

In addition, reversible NMBM fulfil the chaotic hypothesis. Moreover, they can be made dissipative and as such have a natural (SRB) measure as a *microscopic* steady state. Trivially, the stochastic processes is also stationary and obeys the fluctuation relation for the steady state where  $\Delta S^{\text{tot}} = \Delta S^{\text{med}}$ . In that setting, it is a constructive example for a system where the Gallavotti–Cohen fluctuation relation holds [GC95].

In general, the existence of a time-reversal involution which factors on the elements of an absolutely  $\mathcal{I}$ -invariant partition implies an abstract fluctuation theorem [Woj09]. In fact, this abstract fluctuation theorem holds in the more general case where we have two bijective phase space dynamics  $\Psi$  and  $\Phi$ , which are conjugate to each other. Conjugacy in this general setting means that there is a time-reversal involution  $\mathcal{I}$  such that  $\Psi = \mathcal{I} \circ \Phi \circ \mathcal{I}$ . Here, we have always demanded that the conjugate dynamics  $\Psi = \Phi^{-1}$  is also the inverse one.

The more general setting opens up a way to study the microscopic foundations of the “Master fluctuation relation” proposed by Seifert [Sei12]. In the Master fluctuation relation, one is not required to consider reversed trajectories generated by the *same* stochastic process. Rather, the reversed process is any process obtained from the original one by applying an involution-symmetry to the model. Typically, the action of the time-reversal symmetry on control parameters includes the inversion of electric or magnetic fields. Further, if  $\lambda(t)$  specifies a protocol (*i.e.* a deterministic change of the systems parameters in a non-autonomous model),  $\lambda(t)$  may be reversed in a generalized time-

reversed, *conjugate stochastic process*. Consider now any observable  $\varphi$  with a defined parity, *i.e.* where the action of the time-reversal  $\mathcal{I}(\varphi) = \pm\varphi$  is an involution. The master fluctuation relation then holds for all such observables  $\varphi$  such that  $\mathcal{I}(\varphi)$  has the same physical interpretation in the conjugate process. Examples of such quantities are (besides the total entropy production) are the work or the heat [Sei12].

Wojtkowski's conjugate map  $\Psi$  then acts as the deterministic dynamics that gives rise to the conjugate stochastic process. Hence, we conjecture that the abstract fluctuation theorem provides a microscopic basis for the Master fluctuation relation. The appropriate two-sided Markov measure in that case needs to be formulated with respect to the natural measure  $\mu_\Psi = \mu_\Phi \circ \mathcal{I}$  of  $\Psi$  rather than with respect to  $\mu_{\Phi^{-1}}$ .

### 7.2.5. Information, complexity and neuroscience

Finally, let us draw a connection to neuroscience. The Lyapunov spectrum quantifies the average phase space expansion in ergodic systems. In Section 2.6.3 we mentioned how Pesin's formula connects Kolmogorov–Sinai entropy to the Lyapunov spectrum and thus to phase space contraction. We finally showed how this connects to a stochastic description on the level of a Markov partition, *cf.* also Ref. [GW93].

These notions have just recently been investigated from the perspective of theoretical neuroscience [MW10; LLSB13]. Mathematically, the cortex can be thought of as a high-dimensional systems capable of performing complex computations [VS96]. A statistical mechanics perspective in the spirit of Jaynes' ideas can be found in Ref. [Tka+13].

We expect huge developments in this very modern field in the years to follow. The dynamical systems framework (*e.g.* the usage of SRB measures and KS entropy) applied to neuronal computation promises novel qualitative and quantitative results. An overview about the models used by theoretical neuroscientists can be found in a recent review [Wo+14]. Network multibaker maps may also provide a novel perspective on these information-processing non-equilibrium systems.

## 7.3. A summarizing (personal) perspective

Finally, let us come back to Bennett's initial quote on computers as machines that turn available free energy into waste heat and mathematical work. We slightly generalize its statement in the following way: In our opinion, an interesting thermodynamic system (like a computer or a molecular motor) turns free energy into some useful work and some waste heat. Surely, this opinion is as subjective as the meaning of the notion of *useful* work is ambiguous. We just vaguely understand it as the work necessary to generate some form of *pattern*. A pattern is something that is created (and may be sustained) in spite of entropic decay of structure dictated by the second law, *cf.* also Ref. [NP77] by Nicolis and Prigogine. It can be as complex as a living being or as mundane as the output of a trivial mathematical calculation.

## 7. Conclusion and outlook

A reverse statement of the initial quote then reads: A thermodynamic system under sustained non-equilibrium conditions is an engine that converts (free) energy in waste heat and the generation of patterns. If we understand a pattern as encoding some form of information, we might even say: An (interesting) non-equilibrium thermodynamic system is an information processing device, which computes patterns and produces waste heat *via* the consumption of (free) energy.

### 7.3.1. On the ubiquity of Markovian statistics

In Section 3.4.2 we have already commented on the ubiquity of Markovian statistics for measured time-series. We hinted at a certain anthropic principle expressing that scientist observe Markovianity because they look for it. After all, Markovian stochastic processes are the most well-understood ones and are readily used for the purpose of modelling — as we have seen throughout this work.

In the present section, we pick up on that discussion. In his book on the foundations of (classical) statistical mechanics [Pen70], Penrose stresses that the Markovian postulate ensures reproducibility of experimental results. In order to appreciate this fact, suppose that an initial preparation procedure always leaves the system in some observational state  $\omega \in \Omega$ . Note that these states are defined with respect to the measurement apparatus that a scientist A uses to record time series  $\underline{\omega}^{(\tau)}$ .

Now imagine another scientist B, who works on the same system with the same measurement apparatus. Scientist B wants to reproduce some results reported by scientist A. Hence, she needs to be able to prepare initial conditions that yield experiments showing the same statistical properties. However, this is only possible if the initial condition  $\omega$  reported by A does not carry knowledge of their preparation protocol.

This is the main argument behind the Markovian postulate: If the dynamics on the level of the observable states shows Markovian behaviour, the system admits statistically regular and hence reproducible experimental trials. In a sense, Markovian observable states are maximally non-committal with respect to the preparation procedure.

From that perspective the “Markovian anthropic principle” is nothing else than the Scientific Method: Scientific statements need to be reproducible and thus experimentally falsifiable [Pop02]. Consequently, they must be statistically regular, in the sense that the prepared initial state fully specifies the probabilities of future observed time-series.

Local equilibrium, which comes with the time-scale separation between microscopic and observable dynamics, is one way to ensure this statistical regularity. However, in a real experiment the temporal resolution might just be “too good” for the system to relax between two subsequent observations: Observable time series with entries  $\omega \in \Omega$  are not (1-step) Markovian. In that case, one commonly summarizes  $k$  subsequent measurement results  $(\omega_t, \omega_{t+1}, \dots, \omega_{t+k-1})$  into a new observable state  $\underline{\omega}^{(k-1)} \in \Omega^k$ .

Often, time-series recorded in terms of these new observables are (at least to a very good approximation) described by Markovian statistics. Mathematically, this is equivalent to a

refinement  $\bigvee_{t=0}^{k-1} [\Phi^{-t}(\mathcal{Q})]$  of the partition  $\mathcal{Q}$  induced by the measurement observable  $M$ .

A discussion of observable states which points in the same direction was attempted by Shalizi and Moore in the framework of “Computational Mechanics” [SM03]. More precisely, the concept of refining partitions by summarizing subsequent observations into new observable states is the idea behind the construction of *causal states* [SC01; SM03]. These states allow for an *optimal* predictions of future events, *cf.* Ref. [SC01]. Hence, the idea of a causal state is somewhat analogue to our information-theoretical discussion of ST in Chapter 4. Interestingly, causal states have been argued to be the observable states of minimal “thermodynamic depth”, *i.e.* the process of their preparation is the least complex [CS99].

In analogy to deterministic Turing machines, the network of causal states is called an  $\varepsilon$ -machine. These machines can be constructed from data or a probabilistic specification of a dynamics. The stochastic transitions between causal states are Markovian. Some argue that an “ $\varepsilon$ -machine” with a finite number of causal states is simple another name for a Markov chain.

In our opinion,  $\varepsilon$ -machines are a way to interpret Markov chains as *computing devices*. This interpretation then allows for a connection to measures of (computational) complexity, like the Chaitin–Kolmogorov complexity or the notion of Kolmogorov–Sinai entropy [SC01]. These measures of complexity in turn can be used for the characterization of patterns generated by stochastic or deterministic algorithm as outlined in a recent review article by Crutchfield [Cru12].

### 7.3.2. On information processing systems

Motivated by the perspective of Markovian ST as an  $\varepsilon$ -machine, we discuss thermodynamic systems as information processing devices. In particular, we focus on information processing in living systems: Molecular motors and other biological devices that are involved in the cell’s regulatory feedback processes.

In the introductory Chapter 1 we have mentioned Landauer’s principle [Lan61]. It states that any irreversible (*i.e.* non-invertible) logical operation dissipates energy. Bennett summarizes the argument as follows [Ben03]: Reversible operations on the memory of a computer yield a decrease of entropy in its information-bearing degrees of freedom (IBDF). Consider for instance the erasure of a piece of memory, facilitated by resetting every bit to a neutral binary state, say “1”. If we assume the binary data to be essentially random with “0” and “1” appearing at the same frequencies, the average information per bit is  $\log 2 = 1$  bit. After erasure, we know that any bit is in state “1” and thus the uncertainty (or information) per bit is exactly zero. Consequently, the entropy of the IBDF has decreased by an amount of  $\log 2$  per bit. The second law of thermodynamics then ensures that the entropy of the medium (*i.e.* the non-information bearing degrees of freedom, NIBDF) has to be increased by at least the same amount. In essence, this is Landauer’s principle for the thermodynamic cost of irreversible computation. In an

## 7. Conclusion and outlook

isothermal environment, Landauer's limit on the entropy  $S_L = \log 2$  yields a lower bound  $Q^L := k_B T \log 2$  for the dissipated heat per bit.

Although it is possible to design *reversible* computers in a gedankenexperiment [FT82], this is neither practically possible nor useful: A reversible computer cannot delete anything stored in its memory. As such, it will be useless as soon as the latter has reached its storage capacity — unless this memory is erased and Landauer's principle holds.

For real (irreversible) computers engineered by humans, Landauer's limit for the minimum heat is not of technical relevance: It is completely negligible against the dissipated power due to resistivities in the electrical circuits and the energy lost as heat in other conversion processes. However, a recent work inspired by Landauer's original setting demonstrated that his limit can be reached experimentally [Bér+12] — if the experimental set-up is prepared carefully enough.

Even before that experiment, scientists have achieved an experimental realization of Maxwell's demon, *i.e.* a device that turns information about a system into useful work [Toy+10]. Abstractly, a Maxwell demon is a kind of feedback control. Similarly, Landauer's principle can be understood as a special case of a generalized second law for systems with feedback [SU10; HP11; SU12; Sag12].

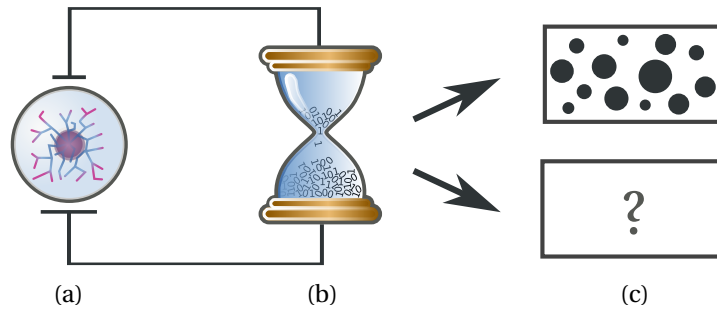
In the experiments mentioned above, the set-ups of the measurement device and its feedback control are very elaborate. Yet, there are machines that work very closely to Landauer's limit. In contrast to digital computers and experimental set-ups, they have not been designed by humans, but by evolution: In every living cell, *polymerases* are enzymes that copy, transcribe and replicate genetic information. They can be understood as molecular motors that run along single-stranded pieces of RNA or DNA, while copying information from the template strand onto the new strand.

A letter in the genetic alphabet corresponds to one of four nucleic acids. Hence, the information is  $\log 4 = 2 \log 2 = 2$  bit per copying event. In Ref. [Ben82], Bennett calculated that the thermodynamic cost of genetic copying maybe as small as  $20k_B T$  per nucleotide, *i.e.*  $10k_B T$  per bit. This is not so far away from Landauer's limit which amounts to  $1k_B T$  per bit.

In biological processes, not only thermodynamic efficiency but also the error rate is important [Ben82]. Relatively high error rates are admissible when genetic information is used as the template for protein synthesis. However, for DNA replication much lower error rates are needed in order to prevent too many unwanted mutations. In real cells, the error rate is reduced by additional molecular motors which run along the freshly synthesized strand and check for errors. This proof-reading scheme has recently been treated in a model of stochastic thermodynamics [SP13].

The copying of genetic information is just one example where biological systems perform computations. Generally, any regulatory mechanism can be understood as information processing through some feedback loop. One example which has been recently studied along these lines is sensory adaptation in the chemotactic response of E. Coli [Lan+12; Sar+14; BHS14], *cf.* also Section 6.1.2 of the present thesis. From an abstract perspective,





**Figure 7.1.:** A (non-equilibrium) thermodynamic system as an information processing device. A source of (free) energy (a) enables a thermodynamic system (b) to produce patterns in information-bearing, accessible degrees of freedom (c, top) at the cost of information dumped to the unobservable medium (c, bottom). The latter process is called dissipation. Here, we interpret it as information written to non-information bearing degrees of freedom.

an adapting system learns how to predict and react accordingly to changes of its environment. Prediction and learning have also been discussed in the framework of stochastic thermodynamics [Sti+12; HBS14].

### 7.3.3. On dissipation

In conclusion, recent work emphasizes the intimate connection between entropy and information — and thus between entropy production and information processing. In thermodynamics, entropy is connected to heat and entropy production to dissipation. After all, the intuitive notion of heat is information-theoretical in the first place: Heat is energy stored in inaccessible and thus non-information bearing degrees of freedom. Dissipation is information lost to the latter.

Useful computations need to be irreversible, because information has to be deleted at some point. Landauer’s principle and its extensions state that this comes at the cost of dissipation, which is the transfer of information into the realm of inaccessibility. Maxwell’s demon is a hypothetical apparatus which by some means has access to that realm. Experimental realizations of Maxwell’s demon stress that the definition of “inaccessibility” is operational rather than objective — as is the distinction between system and medium, which we have adapted throughout this thesis.

We argue that this apparent subjectivity is not a problem for physics or science in general. In contrast, it reflects the role of *observations* and *models* in the method of scientific discovery. A model is a mathematical formalization of empiric rules about the evolution of the world around us. Hence, it enables us to make predictions about the future state of systems, given a certain knowledge about the system’s present state.

The knowledge about this state is described by an ensemble. Its information content is described by the corresponding differential entropy. Thereby one gets different numerical values for the entropy, depending on the ensemble and for instance, the dimensionality of the state space of the model. However, the exact value of the entropy is not important.

## 7. Conclusion and outlook

What counts, is how this information is changing in the course of time — either due to insights from measurements or due to the evolution of the system. ST is a framework to process this information.

Dissipation is a key quantity in this information processing. It formalizes the information written to unobservable degrees of freedom — in contrast to the information contained in the pattern formed by the statistics of the information-bearing observable degrees of freedom. An (irreversible) computation performed on the latter requires some information being dumped to the former, *cf.* Figure 7.1.

In order to compare different *models* of thermodynamic systems, one usually refers to its observable *steady state*. Then, the (subjective) choice of an initial condition does not play any role. However, in this situation we do not observe any *changes* in the statistics of observable patterns described by these models. Consequently, from looking at the observable steady state it is impossible to formulate *dynamical* statements about the information *processed* by the system. We can, however, make dynamical statements regarding the change of entropy in the *unobservable* degrees of freedom, formalized by the dissipation rate.

Ultimately, this is the argument that leads us to say that different models of the same system are consistent if they share the value of the dissipation rate. In the first part of the present thesis we have witnessed how the correct identification of dissipation between to levels of description leads to the emergence of ST. The second part emphasized how dissipation, which is obtained as a bi-linear form involving thermodynamic currents and forces, acts as a consistency criterion between stochastic models on different resolutions.

This is the final conclusion for this work: Viewing dissipation in a complex physical system as information processing is more than an analogy. It formalizes the predictive capabilities of the mathematical models we use to describe nature. Ultimately, we see it as a consequence of the mathematical logic of scientific discovery.

# Acknowledgements

Many people were involved in the perception of this thesis. This is the place to thank them.

First and foremost, I want to thank Jürgen Vollmer for his supervision during the last years. His physical intuition and perfect amount of mathematical rigorosity have substantially influenced both the style of this thesis and my way of my physical thinking. By sharing his connections to scientists working in various fields, he facilitated my entry in the world of scientific research. I am also grateful for the amount of freedom and responsibility I have enjoyed working in his group.

Many thanks go to the members of this group and generally to all the people on the Yellow Floor. I found my scientific home in the Department of Complex Fluids at the Max Planck Institute for Dynamics and Self-Organization. As the head of the department, Stephan Herminghaus always supported my theoretical work and enabled me to share it with people all over the world. In that context let me also thank the other funding agencies, namely the German Academic Exchange Service (DAAD), the Göttingen Graduate School for Neurosciences, Biophysics, and Molecular Biosciences (GGNB) and the Studienstiftung des Deutschen Volkes.

Throughout my time as a Ph.D. candidate I could rely on my thesis committee and especially my second supervisor Marc Timme for both good collaborations and career advices. Further thanks go to all the members of my thesis examination committee for listening to my story. Let me explicitly thank Stefan Kehrein for agreeing to act as a second referee for this thesis.

I am further indebted to the people who helped me to organize myself. Without Barbara Kutz, Monika Teuteberg, Antje Erdmann and Frauke Bergmann I would have dissipated much more energy in the bureaucratic medium surrounding institutional research in Göttingen.

Throughout the last years I was lucky to enjoy discussions with many people in and outside science. Lukas Geyrhofer and David Hofmann were often the first people to hear about my most recent thoughts over one of many dinners we have prepared together. Talking to fellow scientists has constantly changed and challenged my perspective. I particularly acknowledge discussions with Christian Bick, Guillaume Lajoie, Hugo Touchette, Jonathan Dawes, Matteo Poletini, Nigel Goldenfeld, Pablo Sartori, Susanne Still and Udo Seifert.

Two people are still missing in this admittedly incomplete list. The first one is Lamberto Rondoni, who I especially want to thank for many vivid discussions and his warm

## 7. Conclusion and outlook

hospitality in Torino. Ultimately, he has challenged me to pursue the programme that has eventually become the first part of the present thesis.

Secondly, I am indebted to Artur Wachtel in so many ways. His knowledge of physics and mathematics as well as his incredible motivation made him one of my most valued collaborators. Moreover, his technical skills with *Mathematica*<sup>TM</sup>,  $\LaTeX$  and computers in general contributed majorly to the final shape of the present work.

The latter was also influenced by the many people that provided me with feedback on the manuscript, namely Artur, David, Guillaume, Jakob, Jürgen, Laura and Lukas. Thank you all very much for your constructive criticism.

I would also like to express my gratitude towards my many friends in Göttingen and elsewhere. Thank you so much for making these last years a wonderful time, which I will never forget. Thank you also for your support on every front; you were always there for me when I needed you.

Finally, thanks go to my family, who has always supported and encouraged me on my way in and outside of science — even if that meant that I would spend much less time with them than they deserved. I love you and I will always be grateful for everything you have done for me.

Bernhard Altaner  
Göttingen, June 2014

# A. Mathematical appendix

## A.1. Construction of the Markov measure

The author of this thesis is not aware of an explicit construction of non-stationary Markov measure in the literature. Usually, the one-sided or stationary Markov measures are introduced like we did in the main text, *cf.* Defs. 3.23 and 3.25. However, the proof of existence and uniqueness is omitted with a reference to the “usual extension theorems”. In this appendix we explicitly prove existence and uniqueness of the measures defined in definitions 3.23, 3.25 and 3.27.

We first show the existence and uniqueness of certain measures (whose definition seems rather complicated) using the Daniell–Kolmogorov extension theorem. Then, we prove that these measures are the same as the ones defined in the main text, by using the following lemma:

**A.1 Lemma** *Let  $\mathcal{F} \subset \mathcal{P}(X)$  be a  $\pi$ -system, i.e. a non-empty family of subsets of  $X$  which is closed under finite intersections. Then, a measure on  $\sigma(\mathcal{F})$  is uniquely defined by its values on  $\mathcal{F}$ .*

### A.1.1. The basic theorems

We start with some preliminary definitions: A *semi-algebra* on a space  $X$  is a family of sets  $\mathcal{A} \subset \mathcal{P}(X)$  that contains the empty set  $\emptyset$  and is closed under intersections. Further it is *semi-closed* under the formation of complements, i.e.  $\forall A, B \in \mathcal{A} : A \setminus B = \bigcup_{k=1}^K C_k$ , where  $(C_k)_{k \in \{1, 2, \dots, K\}} \subset \mathcal{A}$  is a finite family of mutually disjoint subsets. A probability *pre-measure* on a semi-algebra  $\mathcal{A}$ , is a  $\sigma$ -additive map  $\hat{\mu} : \mathcal{A} \rightarrow [0, 1]$  that obeys  $\hat{\mu}(\emptyset) = 0$ .

The first theorem we need is a generalized version of Carathéodory’s extension theorem (*cf.* Ref. [CB12], Proposition A 25). It states that a pre-measure  $\hat{\mu}$  on a semi-algebra can be uniquely extended to a measure on the generated  $\sigma$ -algebra:

**A.2 Theorem (Carathéodory)** *Let  $\mathcal{A} \subset \mathcal{P}(X)$  be a semi-algebra and  $\hat{\mu}$  a probability pre-measure on  $\mathcal{A}$ . Then, there exists a unique probability measure  $\mu : \mathcal{A} \rightarrow [0, 1]$  on  $\sigma(\mathcal{A})$ , obeying  $\mu|_{\mathcal{A}} = \hat{\mu}$ , i.e.  $\mu$  restricted to  $\mathcal{A}$  agrees with  $\hat{\mu}$ .*

The other tool we need is the Daniell–Kolmogorov extension theorem for inner regular measures. Inner regularity is not a very restrictive property and states that measurable sets can be approximated from within by compact sets:

## A. Mathematical appendix

**A.3 Definition (Inner regular measure)** Let  $((X, \mathcal{A}), \mathcal{T})$  be a measurable space  $(X, \mathcal{A})$  with topology  $\mathcal{T}$ , such that any open set is measurable, *i.e.*  $\mathcal{T} \subset \mathcal{A}$ . A measure  $\nu$  on  $(X, \mathcal{A})$  is called *inner regular*, if for any set  $A \in \mathcal{A}$

$$\nu(A) = \sup \{ \nu(K) \mid K \subset A \text{ is compact} \}.$$

Now we can formulate the version of the Daniell-Kolmogorov theorem, which can be found in T. Tao's introductory textbook [Tao11]:

**A.4 Theorem (Daniell–Kolmogorov)** Let  $F$  be an arbitrary set and  $((X_t, \mathcal{A}_t), \mathcal{T}_t)_{t \in F}$  be an indexed family of measurable spaces  $(X_t, \mathcal{A}_t)$  with respective topologies  $\mathcal{T}_t$ . For all nested finite subsets  $T' \subset T \subset F$ , let  $\mu_T$  be an inner regular probability measure on the product  $\sigma$ -algebra  $\mathcal{A}_T := \bigotimes_{t \in T} \mathcal{A}_t$  which obeys

$$(\pi_{T' \leftarrow T})_* \mu_T = \mu_{T'},$$

where  $\pi_{T' \leftarrow T}$  is the projection map. Then, there exists a unique probability measure  $\mu_F$  on  $\mathcal{A}_F = \bigotimes_{t \in F} \mathcal{A}_t$  such that  $(\pi_{T \leftarrow F})_* \mu_F = \mu_T$  for all  $T \subset F$ .

### A.1.2. Equivalent definitions of the Markov measures

We start with a minimal structure on finite products of  $\Omega$ :

**A.5 Lemma** Let  $\Omega$  and  $T \subset \mathbb{T}$  be finite sets. Let  $\Omega^T = \prod_{t \in T} \Omega$  be the product space consisting of generalized tuples  $\underline{\omega}_T$  with components  $\omega_t = \pi_t \underline{\omega}_T \in \Omega$ . Let  $\mathcal{P}_T = \mathcal{P}(\Omega^T)$  and  $\mathcal{R}_T \subset \mathcal{P}_T$  be the set containing the empty set  $\emptyset$ , the entire set  $\Omega^T$  and all singleton subsets  $\{\underline{\omega}_T\} \subset \Omega^T$ . Then,  $\mathcal{R}_T$  is a semi-algebra and  $\sigma(\mathcal{R}_T) = \mathcal{P}_T$ .

PROOF Because  $\Omega$  and  $T$  are finite, so are  $\Omega^T$  and  $\mathcal{P}_T$ . Hence, each element in  $\mathcal{P}_T$  can be constructed as the union of a finite number of singleton sets  $\{\underline{\omega}_T\}$ . In consequence,  $\mathcal{R}_T$  is semi-closed under the formation of complements, closed under intersections and contains the empty set and the entire set by definition. Therefore  $\mathcal{R}_T$  is a semi-algebra which generates the whole power set as its  $\sigma$ -algebra.  $\square$

On these (finite) sets we define a families of pre-measures:

**A.6 Definition (Markov pre-measure)** Let  $T \subset \mathbb{N}$  be a finite set and  $\mathcal{R}_T$  be as in lemma A.5. Let  $t_{\max}$  be the largest integer in  $T$  and  $T_{\max} := \{0, 1, \dots, t_{\max}\}$ . Let  $A \in \mathcal{R}_T$ . Further, let  $\mathbb{W}$  be a  $N \times N$  stochastic matrix with entries  $w_{\omega'}^\omega$  and  $\mathbf{p} \equiv \mathbf{p}_0$  a stochastic vector. We define functions  $\bar{\mu}_T: \mathcal{R}_T \rightarrow [0, 1]$  as follows:

$$\bar{\mu}_T(A) := \begin{cases} 0, & A = \emptyset, \\ 1, & A = \Omega^T, \\ \sum_{\underline{\omega} \in (T_{\max} \setminus T)} \left[ p_0(\omega_0) \prod_{t=1}^{t_{\max}} w_{\omega_t}^{\omega_{t-1}} \right], & A = \{\underline{\omega}_T\}. \end{cases}$$

If  $T = T_{\max}$ , the set of summation variables is empty and in the sum there appears only one term, which has all  $\omega_t$  specified by  $A = \{\underline{\omega}_T\}$ .

We need to show that these set functions are indeed (probability) pre-measures. Then, theorem A.2 ensures that they can be uniquely extended to measures on  $\mathcal{A}_T := \mathcal{P}(\Omega^T)$ .

**A.7 Lemma** *The set function  $\mu_T = \mu_T(\mathbb{W}, \mathbf{p})$  introduced in definition A.6 is a pre-measure on  $\mathcal{R}_T$ .*

PROOF By definition  $\mu_T(\emptyset) = 0$ . Hence, we only need to show  $\sigma$ -additivity, which is the same as additivity because  $\mathcal{R}_T$  is finite. Then, the Carathéodory extension theorem A.2 yields the required result.

Additivity is also easy to see: Let  $\mathcal{F} = (A^i)_{i \in I}$  be a finite, disjoint family of sets  $A^i \in \mathcal{R}_T$  whose union,  $A = \bigcup_i A^i$ , is an element of  $\mathcal{R}_T$ . If  $A = \emptyset$ , the family must consist only of the empty set and additivity is trivial. If  $A = \{\underline{\omega}_T\}$  is a singleton,  $\mathcal{F}$  contains exactly one element which is  $\{\underline{\omega}_T\}$  and again, additivity is trivial. The only remaining case is that  $A = \Omega^T$  and hence  $\bar{\mu}_T(A) = 1$ . In that case,  $\mathcal{F}$  consists of *all* singleton sets and hence  $\mathcal{F}$  can be indexed by  $\underline{\omega}_T \in \Omega^T$ . Therefore,

$$\begin{aligned} \sum_{A^i \in \mathcal{F}} \mu_T(A^i) &= \sum_{\underline{\omega}(T)} \mu_T(\{\underline{\omega}_T\}) \\ &= \sum_{\underline{\omega}(T)} \sum_{\underline{\omega}(T_{\max} \setminus T)} \left[ p(\omega_0) \prod_{t=1}^{t_{\max}} w_{\omega_t}^{\omega_{t-1}} \right] \\ &= \sum_{\omega_0 \in \Omega} \sum_{\omega_1 \in \Omega} \cdots \sum_{\omega_{t_{\max}} \in \Omega} \left[ p(\omega_0) \prod_{t=1}^{t_{\max}} w_{\omega_t}^{\omega_{t-1}} \right] = 1. \end{aligned}$$

In the last line, we just used the fact that  $\mathbf{p}$  is a stochastic vector and  $\mathbb{W}$  is a stochastic matrix. Thus, we have shown additivity which ensures that  $\mu_T$  is a pre-measure on  $\mathcal{R}_T$ .  $\square$

At this stage, we constructed a measure space  $(\Omega^T, \mathcal{P}_T, \bar{\mu}_T)$  for any finite  $T \subset \mathbb{N}$ . To apply the Daniell–Kolmogorov theorem we have to show inner regularity of the measures and the compatibility criterion:

**A.8 Lemma** *Let  $T \subset \mathbb{Z}$  be a finite subset. Then, the measure space  $(\Omega^T, \mathcal{P}_T, \bar{\mu}_T)$  is inner regular with respect to the discrete topology  $\mathcal{T}_T = \mathcal{P}_T$ .*

PROOF Any measure defined on the power set of any finite set is inner regular, because all sets are compact in the discrete topology on a finite set.  $\square$

**A.9 Corollary** *There is a unique measure  $\bar{\mu} := \bar{\mu}(\mathbb{W}, \mathbf{p})$  on  $(\Omega^{\mathbb{N}}, \mathcal{A}^{\mathbb{N}})$ , such that its restriction to  $\mathcal{P}_T$  is  $\bar{\mu}_T(\mathbb{W}, \mathbf{p})$  for all finite  $T \subset \mathbb{N}$ .*

PROOF Lemmata A.5–A.8 ensure that all of the assumptions for the Daniell–Kolmogorov theorem A.4 with the exception of the compatibility criterion are already satisfied. The

## A. Mathematical appendix

only thing left to show is that

$$(\pi_{T' \leftarrow T})_* \bar{\mu}_T = \bar{\mu}_{T'}, \quad \forall T' \subset T$$

for any finite  $T \subset \mathbb{N}$ . In our consideration we can restrict ourselves to  $A \in \mathcal{R}_{T'}$ . The Carathéodory extension ensures that this agreement on the semi-algebras  $\mathcal{R}_{T'}$  carries on to the family of measures defined on the generated  $\sigma$ -algebras. First, let  $A = \emptyset$  and observe that  $(\pi_{T' \leftarrow T})_* \bar{\mu}_T(\emptyset) = 0 = \bar{\mu}_{T'}(\emptyset)$ . Now let  $A = \Omega^{T'}$  and hence  $\pi_{T' \leftarrow T}^{-1} A = \Omega^T$  and therefore  $(\pi_{T' \leftarrow T})_* \bar{\mu}_T(\Omega^{T'}) = 1 = \bar{\mu}_{T'}(\Omega^{T'})$ .

The interesting case is  $A = \{\underline{\omega}_{T'}\}$  where  $\underline{\omega}_{T'} \in \Omega^{T'}$ . For convenience, we introduce the following notation: Let  $T \subset \mathbb{N}$  be finite and  $T_1, T_2 \subset T$  be disjoint subsets. Let  $\underline{\omega}_{T_1} \in \Omega^{T_1}$  and  $\underline{\omega}_{T_2} \in \Omega^{T_2}$ . The expression  $\underline{\omega}_{T_1} \oplus \underline{\omega}_{T_2} \in \Omega^T$  denotes the unique element satisfying  $\pi_{T_1 \leftarrow T}(\underline{\omega}_{T_1} \oplus \underline{\omega}_{T_2}) = \underline{\omega}_{T_1}$  and  $\pi_{T_2 \leftarrow T}(\underline{\omega}_{T_1} \oplus \underline{\omega}_{T_2}) = \underline{\omega}_{T_2}$ . Let  $T'' = T \setminus T'$  and find that

$$\begin{aligned} (\pi_{T' \leftarrow T})_* \bar{\mu}_T(\{\underline{\omega}_{T'}\}) &\equiv \bar{\mu}_T(\pi_{T' \leftarrow T}^{-1}(\{\underline{\omega}_{T'}\})) \\ &= \bar{\mu}_T\left(\bigcup_{\underline{\omega}_{T''} \in \Omega^{T''}} \{\underline{\omega}_{T''} \oplus \underline{\omega}_{T'}\}\right) \\ &= \sum_{\underline{\omega}_{T''}} [\bar{\mu}_T(\{\underline{\omega}_{T''} \oplus \underline{\omega}_{T'}\})] \\ &= \sum_{\underline{\omega}_{T''}} \sum_{\underline{\omega}_{(T_{\max} \setminus T)}} \left[ p(\omega_0) \prod_{t=1}^{t_{\max}} w_{\omega_t}^{\omega_{t-1}} \right] \\ &= \sum_{\underline{\omega}_{(T_{\max} \setminus T')}} \left[ p(\omega_0) \prod_{t=1}^{t_{\max}} w_{\omega_t}^{\omega_{t-1}} \right] \underbrace{\sum_{\underline{\omega}_{(T_{\max} \setminus T'_{\max})}} \left[ \prod_{t=t'_{\max}+1}^{t_{\max}} w_{\omega_t}^{\omega_{t-1}} \right]}_{=1} \\ &= \bar{\mu}_{T'}(\{\underline{\omega}_{T'}\}). \end{aligned}$$

Since we have shown compatibility, the Daniell–Kolmogorov extension theorem finally yields the desired result. The rearrangement of the sums over index sets can be seen from

$$\begin{aligned} T'' \cup (T_{\max} \setminus T) &= (T \setminus T') \cup (T_{\max} \setminus T) \\ &= T_{\max} \setminus T' \\ &= (T_{\max} \setminus T'_{\max} \cup T'_{\max}) \setminus T' = (T'_{\max} \setminus T') \cup ((T_{\max} \setminus T'_{\max}) \setminus T') \\ &= (T'_{\max} \setminus T') \cup (T_{\max} \setminus T'_{\max}) \quad \square \end{aligned}$$

By definition, this measure agrees with the set function defined in Def. 3.23 on all 0-shifted  $\tau$ -cylinders. It is easy to see that these sets together with the empty set are closed under intersection: Just note that for any given  $\tau$ , the family  $\mathcal{Z}_0^{(\tau)}$  is a partition of  $\Omega^{\mathbb{N}}$ . Further, the set  $\mathcal{Z}_0^{(\tau')}$  is a refinement of  $\mathcal{Z}_0^{(\tau)}$  whenever  $\tau' > \tau$ . Consequently, if one takes two arbitrary elements of  $\{\mathcal{Z}_0^{(\tau)}\}_{\tau \in \mathbb{N}}$ , they are either disjoint or one is contained within the other. Hence,  $\{\mathcal{Z}_0^{(\tau)}\}_{\tau \in \mathbb{N}}$  is a  $\pi$ -system and uniqueness is guaranteed by Lemma A.1.



### A.1.3. Measures on the bi-infinite sequences

So far we have only shown the argument for the one-sided case of Def. A.6. Due to the fact that the construction always proceeds using finite  $T \subset \mathbb{T}$ , the proofs for  $\mathbb{T} = \mathbb{Z}$  proceed along the same lines. Throughout this subsection let  $T \subset \mathbb{Z}$  be finite,  $\mathcal{R}_T$  be the semi-algebra on  $\Omega^T$  that contains the singletons, the full set and the empty set. Further, let  $t_{\max} = \max(\max(T), 0)$  and  $t_{\min} = \min(\min(T), 0)$  and  $T_{\text{def}} := \{t_{\min}, t_{\min} + 1, \dots, t_{\max}\}$ . We define the set functions corresponding to Definition 3.25 in the main text:

**A.10 Definition (Stationary Markov pre-measure for the bi-infinite sequences)** Let  $\mathbf{p}_\infty$  be a stochastic left eigenvector for the unity eigenvalue of a stochastic matrix  $\mathbb{W}$ .

For  $A \in \mathcal{R}_T$  we define functions  $\bar{\mu}_T: \mathcal{R}_T \rightarrow [0, 1]$  as follows:

$$\bar{\mu}_T(A) := \begin{cases} 0, & A = \emptyset, \\ 1, & A = \Omega^T, \\ \sum_{\underline{\omega}_{T_{\text{def}} \setminus T}} [p_\infty(\omega_0) \left( \prod_{t=t_{\min}+1}^{t_{\max}} w_{\omega_t}^{\omega_{t-1}} \right)], & A = \{\underline{\omega}_T\}. \end{cases}$$

Similarly, the set functions corresponding to definition 3.27 read:

**A.11 Definition (Two-sided Markov pre-measure)** Let  $\mathbf{p}$  be a stochastic vector and let  $\mathbb{W}$  and  $\tilde{\mathbb{W}}$  be  $N \times N$  stochastic matrices compatible with some adjacency matrix  $\mathbb{A}$  respectively its transpose  $\mathbb{A}^T$ .

For  $A \in \mathcal{R}_T$  we define functions  $\bar{\mu}_T: \mathcal{R}_T \rightarrow [0, 1]$  as follows:

$$\bar{\mu}_T(A) := \begin{cases} 0, & A = \emptyset, \\ 1, & A = \Omega^T, \\ \sum_{\underline{\omega}_{T_{\text{def}} \setminus T}} [p(\omega_0) \left( \prod_{t=1}^{t_{\max}} w_{\omega_t}^{\omega_{t-1}} \right) \left( \prod_{t=1}^{-t_{\min}} \tilde{w}_{\omega_{-t}}^{\omega_{-t+1}} \right)], & A = \{\underline{\omega}_T\}. \end{cases}$$

As said above, the proof of existence is directly analogous to the one-sided Markov measure. It remains to be shown that the measure defined here is the equivalent to the one in the main text. We start with the two-sided measure. It is defined on the set of 0-shifted forward and backward cylinders, as well as on intersections of elements of the two. Thus, it is defined on a  $\pi$ -system and yields a unique measure. Further, it is easy to check that Def. A.11 is consistent with Def. 3.27 and thus the measures are equivalent. Finally, the same argument together with the fact that backward cylinders can be rewritten as shifted forward cylinders for the reversed sequences ensures that Def. A.10 is consistent with Def. 3.25. The same argument is used to prove Proposition 3.28.

It remains to be shown that  $s_*^t \bar{\mu}(\mathbb{W}, \mathbf{p}_0) = \bar{\mu}(\mathbb{W}, \mathbf{p}_t)$  for  $t \in \mathbb{Z}$ , which was the statement of Proposition 3.24. But this is clear from definition A.6 and the fact that  $s^{-t} Z_0[\underline{\omega}^{(t)}] = Z_t[\underline{\omega}^{(t)}]$ .

## A.2. Sequences of $\tau$ -chains and their variations

Stochastic thermodynamics uses the notion of time-series dependent observables for the identification of fluctuating entropies, *cf.* Sec. 2.4.2. For the case of stochastic jump processes in discrete time, these observables are examples of  $\tau$ -chains. Formally, a  $\tau$ -chain is a measurable function

$$\begin{aligned} \varphi^{(\tau)}: \Omega^{\tau+1} \times \mathbb{Z} &\rightarrow \mathbb{R}, \\ (\underline{\omega}^{(\tau)}, t) &\mapsto \varphi_t[\underline{\omega}^{(\tau)}]. \end{aligned}$$

The  $(\tau + 1)$ -point average or *time-series average* is defined as

$$\langle\langle \varphi^{(\tau)} \rangle\rangle_{t_0}^{(\tau)} := \sum_{\underline{\omega}^{(\tau)} \in \Omega^{\tau+1}} \mathbb{P}_{t_0}[\underline{\omega}^{(\tau)}] \varphi_{t_0}[\underline{\omega}^{(\tau)}]. \quad (\text{A.1})$$

where  $\mathbb{P}_{t_0}[\underline{\omega}^{(\tau)}]$  is the probability for a time-series  $\underline{\omega}^{(\tau)}$  to occur at time  $t_0$ . For brevity of notation we usually write  $\langle\langle \varphi \rangle\rangle_{t_0}^{(\tau)}$  instead of  $\langle\langle \varphi^{(\tau)} \rangle\rangle_{t_0}^{(\tau)}$ , when no ambiguity can arise. In Section 2.4.2 we saw that jump averages reduce to state averages, if the observable is a one-chain, *i.e.* if it depends only on the initial and final state of a jump  $\omega \rightarrow \omega'$ . The general result is the following:

**A.12 Lemma** *Let  $\varphi_t^{(\tau)} = \varphi_{t+k}(\omega_k)$  be a  $\tau$ -chain, which depends only on the  $k$ th component  $\omega_k$  of a block  $\underline{\omega}^{(\tau)}$ . Then, the time-series average reduces to the state average*

$$\langle\langle \varphi \rangle\rangle_t^{(\tau)} = \langle \varphi \rangle_{t+k}.$$

PROOF Let  $T = \{0, 1, \dots, \tau\} \setminus \{k\}$ .

$$\begin{aligned} \langle\langle \varphi^{(\tau)} \rangle\rangle_t^{(\tau)} &= \sum_{\underline{\omega}^{(\tau)}} \mathbb{P}_t[\underline{\omega}^{(\tau)}] \varphi_t[\underline{\omega}^{(\tau)}] \\ &= \sum_{\omega_k} \varphi_{t+k}(\omega_k) \sum_{\underline{\omega}^{(T)}} \mathbb{P}_t[\underline{\omega}^{(\tau)}] \\ &= \sum_{\omega_k} \varphi_{t+k}(\omega_k) p_{t+k}(\omega_k) = \langle \varphi \rangle_{t+k} \end{aligned}$$

To get to the last line we used the rule of marginal probability, *i.e.* the fact that  $\mu$  is a measure and that  $p_t(\omega_k) = \mathcal{C}_t[(\omega_k)] = \bigcup_{\underline{\omega}^{(T)}} \mathcal{C}_t[\underline{\omega}^{(\tau)}]$ .  $\square$

In the following we consider the relation between state averages  $\langle \varphi_t \rangle$  and time-series averages  $\langle\langle \varphi_t \rangle\rangle$ . More precisely, we are interested in averages of certain *sequences* of  $\tau$ -chains.

Note that any state observable (*i.e.* any 0-chain)  $\varphi_t(\omega)$  induces such a *canonical sequence*, *cf.* Sec. 4.1.3:

**A.13 Definition (Canonical sequence)** Let  $\varphi_t : \Omega \rightarrow \mathbb{R}$  be a 0-chain parametrized by a time index  $t$ . Let  $t_0, \tau \in \mathbb{N}$  and denote by  $\omega_\tau$  the final state of  $\underline{\omega}^{(\tau)}$ . The sequence  $(\varphi_{t_0}^{(\tau)})_{\tau \in \mathbb{N}}$  with elements

$$\varphi_{t_0}^{(\tau)}[\underline{\omega}^{(\tau)}] := \varphi_{t_0+\tau}(\omega_\tau),$$

is called the *canonical sequence of time-series observables* for  $\varphi_t$ .

For the canonical sequence, Lemma A.12 ensures that

$$\langle\langle \varphi^{(\tau)} \rangle\rangle_{t_0}^{(\tau)} = \langle \varphi \rangle_{t_0+\tau}. \quad (\text{A.2})$$

The notion of a sequence of  $\tau$ -chains is needed for the consistent definition of the *temporal variation* associated with a  $\tau$ -chain. Before we discuss the *sequence of variations* associated to a sequence of  $\tau$ -chains, we define the *temporal variation*  $\Delta F$  of a function  $F(\tau) : \mathbb{N} \rightarrow \mathbb{R}$  as

$$\Delta F(\tau) := F(\tau) - F(\tau - 1). \quad (\text{A.3})$$

We define a similar notion for sequences of  $\tau$ -chains:

**A.14 Definition** Let  $(\varphi_t^{(\tau)})_{t \in \mathbb{N}}$  be a sequence of  $\tau$ -chains. The sequence of time-series observables  $(\delta_t^{(\tau)} \varphi)_{t \in \mathbb{N}^+}$  with elements

$$(\delta_t^{(\tau)} \varphi_t)[(\underline{\omega}^{(\tau-1)}, \omega_\tau)] := \varphi_t^{(\tau)}[(\underline{\omega}^{(\tau-1)}, \omega_\tau)] - \varphi_t^{(\tau-1)}[\underline{\omega}^{(\tau-1)}]$$

is called the *temporal variation* of the sequence  $(\varphi_t^{(\tau)})_{t \in \mathbb{N}}$ .

Again, for a more compact notation we may drop the superscript  $(\tau)$  on  $\delta_t^{(\tau)} \varphi_t$  if the meaning is clear from the context.

With these definition we are able to state and proof the following result. It relates the variations of (explicitly time-dependent) averages for sequences of  $\tau$ -chains with the average of the variation of the  $\tau$ -chain:

**A.15 Lemma (Variations and averages)** Let  $(\varphi_t^{(\tau)})_{t \in \mathbb{N}}$  be a sequence of  $\tau$ -chains. Let  $(\delta_t^{(\tau)} \varphi_t)_{t \in \mathbb{N}^+}$  be its temporal variation. Define  $F_{t_0}(\tau) := \langle\langle \varphi_t^{(\tau)} \rangle\rangle_{t_0}^{(\tau)}$ . Then, we have

$$\Delta_{t_0+\tau-1}^{(1)} F := (\Delta F_{t_0})(\tau) = \langle\langle \delta_t^{(\tau)} \varphi_t \rangle\rangle_{t_0}^{(\tau)}.$$

PROOF Let  $\underline{\omega}^{(\tau-1)} := \pi_{\{0,1,\dots,\tau-1\} \leftarrow \{0,1,\dots,\tau\}}(\underline{\omega}^{(\tau)})$  denote the first  $\tau$  components of  $\underline{\omega}^{(\tau)}$ . First

## A. Mathematical appendix

observe that

$$\begin{aligned}
\langle\langle \varphi^{(\tau-1)} \rangle\rangle_{t_0}^{(\tau)} &= \sum_{\underline{\omega}^{(\tau)} \in \Omega^{\tau+1}} \left[ \mathbb{P}_{t_0} [\underline{\omega}^{(\tau)}] \varphi_{t_0}^{(\tau-1)} [\underline{\omega}^{(\tau-1)}] \right] \\
&= \sum_{\underline{\omega}^{(\tau-1)} \in \Omega^\tau} \sum_{\omega_\tau} \left[ \mathbb{P}_{t_0} [\underline{\omega}^{(\tau-1)}] \cdot \mathbb{P}_{t_0+\tau-1} [\omega_{\tau-1} \rightarrow \omega_\tau] \varphi_{t_0}^{(\tau-1)} [\underline{\omega}^{(\tau-1)}] \right] \\
&= \sum_{\underline{\omega}^{(\tau-1)} \in \Omega^\tau} \left[ \mathbb{P}_{t_0} [\underline{\omega}^{(\tau-1)}] \varphi_{t_0}^{(\tau-1)} [\underline{\omega}^{(\tau-1)}] \right] \cdot \sum_{\omega_\tau} [\mathbb{P}_{t_0+\tau-1} [\omega_{\tau-1} \rightarrow \omega_\tau]] \\
&= \sum_{\underline{\omega}^{(\tau-1)} \in \Omega^\tau} \left[ \mathbb{P}_{t_0} [\underline{\omega}^{(\tau-1)}] \varphi_{t_0}^{(\tau-1)} [\underline{\omega}^{(\tau-1)}] \right] = \langle\langle \varphi^{(\tau-1)} \rangle\rangle_{t_0}^{(\tau-1)},
\end{aligned}$$

where  $\mathbb{P}_{t_0+\tau-1} [\omega \rightarrow \omega'] := \frac{\mathbb{P}_{t_0+\tau-1}[(\omega, \omega')]}{p_{t_0+\tau-1}(\omega)}$  is the conditional probability for symbol  $\omega'$  to occur at time  $t_0 + \tau$  given  $\omega$  has occurred at time  $t_0 + \tau - 1$ . With that we obtain the desired result:

$$\begin{aligned}
(\Delta F_{t_0})(\tau) &= F_{t_0}(\tau) - F_{t_0}(\tau-1) = \langle\langle \varphi_t^{(\tau)} \rangle\rangle_{t_0}^{(\tau)} - \langle\langle \varphi_t^{(\tau-1)} \rangle\rangle_{t_0}^{(\tau-1)} \\
&= \langle\langle \varphi_t^{(\tau)} \rangle\rangle_{t_0}^{(\tau)} - \langle\langle \varphi_t^{(\tau-1)} \rangle\rangle_{t_0}^{(\tau)} = \langle\langle \varphi_t^{(\tau)} - \varphi_t^{(\tau-1)} \rangle\rangle_{t_0}^{(\tau)} \\
&= \langle\langle \delta^{(\tau)} \varphi_t \rangle\rangle_{t_0}^{(\tau)} \quad \square
\end{aligned}$$

Note that if  $\left( \varphi_t^{(\tau)} \right)_{\tau \in \mathbb{N}}$  is a canonical sequence for a 0-chain  $\varphi_t(\omega)$ , the average depends only on the sum of  $t_0$  and  $\tau$ , cf. Eq. (A.2).

## B. Construction of a minimal model for kinesin

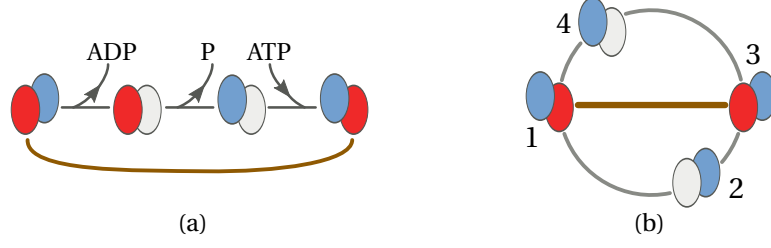
In this section we review the construction and parametrization of the four-state model for the molecular motor kinesin. We follow the arguments of Liepelt and Lipowsky [LL07] to construct the model, and use the experimental data from Ref. [CC05] to obtain numerical values for the parameters. Further we rely on Hill–Schnakenberg balance conditions (6.10) to infer some rates from the others. Even with this consistency requirements, we still lack one first order rate constant in order to specify all numerical rates of the model. This parameter is obtained by a comparison to the six-state model for physiological values of the chemical concentrations at vanishing mechanical driving. Comparing the results obtained with that choice to previous results [LL09], in Section 6.3.3 we find that this choice — made for physiological conditions — is good *globally*, that is everywhere in the phase diagram.

### B.1. State space of the four-state model

Kinesin is a molecular motor “walking” along a one-dimensional “track” by hydrolysing adenosine triphosphate (ATP) into adenosine diphosphate (ADP) and inorganic phosphate (P). For a more detailed exposition of the mechanics of the kinesin step, see Ref. [Yil+04; CC05]. It is important that mechanical stepping and chemical hydrolysis are *not* tightly coupled. Physically, this means that it is possible to have “futile” hydrolysis events in which the motor does not take any step. Mathematically, this means that there must be at least two (fundamental) cycles in any discrete stochastic model. Disregarding multiple transitions between states, the simplest compatible topology, is given by the four-state network used as an example throughout Chapter 5.

In that model, the states are labelled by the chemical composition of kinesin’s active groups: They can either be empty (E) or have ATP or ADP attached, see Figure B.1. In the following, we repeat the arguments given by Liepelt and Lipowsky [LL07; LL09] in order to arrive at a parametrization of our four-state model. In general, transition rates  $w_j^i$  (unit:  $\frac{1}{s}$ ) are parametrized by first-order rate constants  $\kappa_j^i$  which are multiplied concentration and/or force-dependent factors  $C_j^i$  and  $\Phi_j^i$ , respectively.

## B. Construction of a minimal model for kinesin



**Figure B.1.:** a) Chemical pathway involving (from left to right) detachment of ADP (blue) from the leading foot, hydrolysis of ATP (red) at the trailing foot with release of P, and finally attachment of ATP to the leading foot. From the final configuration the system is likely to undergo a mechanical transition (brown) which reverts the order of the feet yielding kinesin's mechanical step. b) In our model, we summarize the first two steps (attachment of ADP, hydrolysis of ATP and release of P) into one transition. Symmetry and the demand for a non-tightly coupled model give rise to the four-state model (right).

### Chemical transitions

For a chemical transition, we have

$$w_j^i := \kappa_j^i \cdot C_j^i \cdot \Phi_j^i(f). \quad (\text{B.1})$$

Here,  $\kappa_j^i$  is the first order rate constant (see table B.1) for the transition ( $i \rightarrow j$ ). Further,

$$C_j^i := \begin{cases} \prod_X [X] & \text{if compound } X \text{ is attached during the transition,} \\ 1 & \text{else,} \end{cases} \quad (\text{B.2})$$

where  $[X]$  denotes the (fixed) concentration of compound  $X$  in the surrounding medium. The factor

$$\Phi_j^i(f) = \frac{2}{1 + \exp[\chi_j^i f]} \quad (\text{B.3})$$

depends on the (non-dimensional) applied force  $f = lF/(k_B T)$  and symmetric mechanical parameters  $\chi_j^i = \chi_i^j$ . Values for the mechanical parameters are the ones used in Ref. [LL07] to account for the data of Ref. [CC05], see table B.1.

### Mechanical transition

The mechanical transition lacks the chemical attachment factor  $C_j^i$  and has a slightly different force-dependent factor  $\tilde{\Phi}(f)$ . The combined rates are thus:

$$w_j^i := \kappa_j^i \cdot \tilde{\Phi}_j^i(f), \quad (\text{B.4})$$

Mechanical transition	$\kappa_3^1 = 3 \cdot 10^5$	$\kappa_1^3 = 0.24$
Chemical transitions (forward cycle)	$\kappa_4^1 = 100$ $\kappa_3^4 = c = 2.52 \cdot 10^6$	$\kappa_1^4 = 2.0$ $\kappa_4^3 = \frac{K_{\text{eq}} \kappa_3^4 \kappa_4^1 \kappa_1^3}{\kappa_1^4 \kappa_3^1} = 49.3$
Chemical transitions (backward cycle)	$\kappa_2^3 = \left(\frac{\kappa_1^3}{\kappa_3^1}\right)^2$ $\kappa_4^1 = 6.4 \cdot 10^{-11}$ $\kappa_1^2 = c = 2.52 \cdot 10^6$	$\kappa_3^2 = \kappa_1^4 = 2.0$ $\kappa_2^1 = \kappa_4^3 = 49.3$
Mechanical load	$\chi_4^3 = \chi_3^4 = \chi_2^1 = \chi_1^2 = 0.15$	$\chi_1^4 = \chi_4^1 = \chi_3^2 = \chi_2^3 = 0.25$

**Table B.1.:** Parameters of the four-state model for kinesin. All first-order reaction rates  $\kappa$  are given in units of  $\text{s}^{-1}$  or, if attachment of chemicals is involved,  $\text{s}^{-1} \mu\text{M}^{-1}$ . The equilibrium constant of the ATP hydrolysis reaction is  $K_{\text{eq}} = 4.9 \cdot 10^{11} \mu\text{M}$ . The parameter  $\theta = 0.65$  enters the mechanical factor of the transition rates.

with the mechanical (load distribution) factor

$$\tilde{\Phi}_j^i(f) := \begin{cases} \exp(-\theta f), & \text{if } (i \rightarrow j) = (1 \rightarrow 3) \\ \exp((1 - \theta)f), & \text{if } (i \rightarrow j) = (3 \rightarrow 1). \end{cases} \quad (\text{B.5})$$

Again, the choice for the parameter  $\theta = 0.65$  corresponds to the experimental data in Ref. [CC05].

## B.2. Choice of parameters

The parametrization of the mechanical transition is taken from the work of Liepelt and Lipowsky [LL07] to reflect experiments [CC05]. However, the choice of first-order rate constants for the chemical transitions requires adaption for our simpler model. The transition  $(\text{ADP}, \text{ATP}) \rightarrow (\text{ADP}, \text{E})$  is present in both our and the original six-state model. For this transition we use the same values as in Ref. [LL07]. Thus, we only need to find a good parametrization of the rate  $w_3^4$  for the transition  $(\text{ADP}, \text{E}) \rightarrow (\text{ATP}, \text{ADP})$  and its reverse. In the force-free case the mechanical parameters drop out. Hence, the first-order constant for one of the transitions  $3 \rightleftharpoons 4$  can be obtained from the Hill–Schnakenberg conditions (6.10) [Sch76]. In equilibrium, *i.e.* when there are no driving forces acting on the system, these conditions ensure that the (chemical) affinity of every cycle vanishes. For the upper (forward) cycle  $(1 \rightarrow 3 \rightarrow 4)$  this statement can be cast into the expression

$$\frac{\kappa_4^3 \kappa_1^4 \kappa_3^1}{\kappa_3^4 \kappa_4^1 \kappa_1^3} \stackrel{!}{=} K_{\text{eq}}, \quad (\text{B.6})$$

where  $K_{\text{eq}} = 4.9 \cdot 10^{11} \mu\text{M}$  is the chemical equilibrium constant for the ATP hydrolysis reaction. Thus, we still have the freedom to choose one of the rate constants of the forward cycle. Hence,  $c \equiv \kappa_4^3$  is the above mentioned fit parameter. It is fixed by demanding the motor's velocity  $V$  in the four-state model to be identical to that of the six-state model in the force-free case ( $f = 0$ ) and at the physiological concentrations  $[\text{ATP}] = [\text{ADP}] = [\text{P}] = 1 \mu\text{M}$ .

## B. Construction of a minimal model for kinesin

By symmetry, the chemical first-order rates of the lower cycle ( $3 \rightarrow 1 \rightarrow 2$ ) are chosen to be the same as that of the upper cycle, with the exception of  $\kappa_2^3$ . The latter determines the likelihood for the system to let go of the ATP molecule rather than (after releasing ADP from the other head) hydrolysing it. Because the balance condition (B.6) is required to hold also for the lower cycle, we use it to determine the missing rate  $\kappa_2^3$ . This is the same reasoning as in Ref. [LL07].

Finally, we have to estimate the missing parameter  $\chi_4^3 = \chi_3^4$  for the combined transition (ADP, ATP)  $\rightarrow$  (ADP, E). Liepelt and Lipowsky used the same parameter ( $\chi = 0.15$ ) for both of its chemical substeps. We take this as an argument to use  $\chi_4^3 = \chi_2^1 = 0.15$  for the combined rate and its reversed counter-part in the lower cycle. All model parameters are summarized in table B.1.



# List of frequently used symbols

$\langle\langle \cdot \rangle\rangle$	Time-series average
$\langle \cdot \rangle$	Ensemble average
$\emptyset$	Empty set
$\mathbb{A}$	Adjacency matrix
$A$	Affinity (generalized thermodynamic force)
$\mathcal{C}_\omega$	Phase space cell, partition element
$D_{\text{KL}}$	Kullback–Leibler divergence, relative entropy
$\mathcal{E}$	Motance, <i>i.e.</i> logarithmic ratio of forward and backward transition rates
$\Gamma$	Phase space of a deterministic dynamical system
$G$	Graph
$H[\cdot]$	Shannon entropy
$I_\varphi(x)$	Rate function for the current-like observable $\varphi$
$\mathcal{I}$	(Measure-preserving) time-reversal involution
$J$	(Probability) current
$\Lambda$	Phase space expansion (rate)
$\lambda_\varphi(q)$	Scaled cumulant generating function for the current-like observable $\varphi$
$M$	Observable for the measurement process
$\mu$	Probability measure
$\mathbb{N}$	Natural numbers (including zero)
$\Omega$	Space of observations, sample space, alphabet

*List of frequently used symbols*

$\omega$	Observable state, letter, symbol
$\underline{\omega}$	Observable time-series, symbolic sequence
$\mathbb{P}$	Probability measure, probability of an event
$\phi$	Probability flux
$\Phi$	Deterministic microscopic dynamics
$\varphi$	(Current-like) observable
$\boldsymbol{p}$	Stochastic vector, discrete probability distribution
$\Psi$	Flow of a continuous dynamical system
$\mathbb{R}$	Real numbers
$\rho$	Probability density
$S$	Entropy obtained as an ensemble average
$s$	Entropic $\tau$ -chain, <i>i.e.</i> a time-series dependent random variable with an entropic interpretation
$\delta s$	(Temporal) variation of an entropic $\tau$ -chain
$\Delta S$	(Temporal) variation of an ensemble entropy
$\text{sgn}$	Sign function
$\mathbb{T}$	Time domain
$w_{\omega'}^{\omega}$	Probability or rate for the transition $\omega \rightarrow \omega'$
$\mathbb{W}$	Transition matrix, rate matrix
$x$	Microscopic state
$\mathbb{Z}$	Integers

# Bibliography

- [Ach90] D. J. Acheson. *Elementary fluid dynamics*. Oxford University Press, 1990 (cit. on p. 139).
- [Adl98] R. L. Adler. “Symbolic dynamics and Markov partitions”. In: *B. Am. Math. Soc.* 35.1 (1998), pp. 1–56 (cit. on pp. 58, 71–73, 75–77, 84, 87, 178).
- [AG07] D. Andrieux and P. Gaspard. “Fluctuation theorem for currents and Schnakenberg network theory”. In: *J. Stat. Phys.* 127.1 (2007), pp. 107–131 (cit. on pp. 11, 120, 135, 137, 156).
- [AGW77] R. L. Adler, L. W. Goodwyn, and B. Weiss. “Equivalence of topological Markov shifts”. In: *Israel. J. Math.* 27.1 (1977), pp. 49–63 (cit. on p. 70).
- [Alb+07] B. Alberts et al. *Molecular biology of the cell*. Garland, New York, 2007 (cit. on pp. 145, 151).
- [Alt p] B. Altaner. “Nonstationary Markov measures”. (in preparation) (cit. on pp. 27, 64).
- [Alt+12] B. Altaner et al. “Network representations of nonequilibrium steady states: Cycle decompositions, symmetries, and dominant paths”. In: *Phys. Rev. E* 85.4 (2012), p. 41133 (cit. on pp. 120, 123, 124, 138, 164).
- [Ast97] R. D. Astumian. “Thermodynamics and kinetics of a Brownian motor”. In: *Science* 276.5314 (1997), pp. 917–922 (cit. on p. 9).
- [AV12] B. Altaner and J. Vollmer. “Fluctuation-Preserving Coarse Graining for Biochemical Systems”. In: *Phys. Rev. Lett.* 108 (22 2012), p. 228101 (cit. on pp. 147, 148, 150, 151, 153, 155).
- [BB00] Y. M. Blanter and M. Büttiker. “Shot noise in mesoscopic conductors”. In: *Phys. Rep.* 336.1 (2000), pp. 1–166 (cit. on p. 26).
- [BB03] M. Blank and L. Bunimovich. “Multicomponent dynamical systems: SRB measures and phase transitions”. In: *Nonlinearity* 16.1 (2003), p. 387 (cit. on pp. 50, 81, 84, 113, 178).
- [BC75] R. Bowen and J.-R. Chazottes. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Vol. 470. Springer, 1975 (cit. on pp. 14, 84).
- [Bek03] J. Bekenstein. “Information in the holographic universe”. In: *Sci. Am.* 289.2 (2003), pp. 58–65 (cit. on p. 23).
- [Ben03] C. Bennett. “Notes on Landauer’s principle, reversible computation, and Maxwell’s Demon”. In: *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 34.3 (2003), pp. 501–510 (cit. on pp. 15, 183).

## Bibliography

- [Ben82] C. Bennett. “The thermodynamics of computation—a review”. In: *Int. J. Theor. Phys.* 21.12 (1982), pp. 905–940 (cit. on pp. 15, 184).
- [BH12] N. Boon and R. Hoyle. “Exact dynamic properties of molecular motors”. In: *J. Chem. Phys.* 137.8 (2012), p. 84102 (cit. on pp. 9, 165).
- [BHS14] A. C. Barato, D. Hartich, and U. Seifert. “Efficiency of cellular information processing”. In: *arXiv* 1405.7241 (2014) (cit. on p. 184).
- [BL89] H. Bai-Lin. *Elementary symbolic dynamics and chaos in dissipative systems*. Vol. 13. World Scientific, 1989 (cit. on p. 70).
- [Bou02] R. Bousso. “The holographic principle”. In: *Rev. Mod. Phys.* 74.3 (2002), p. 825 (cit. on p. 23).
- [Bow70] R. Bowen. “Markov partitions for Axiom A diffeomorphisms”. In: *Am. J. Math.* 92.3 (1970), pp. 725–747 (cit. on p. 77).
- [BR75] R. Bowen and D. Ruelle. “The ergodic theory of Axiom A flows”. In: *Invent. Math.* 29.3 (1975), pp. 181–202 (cit. on pp. 84, 177).
- [BS95] C. Beck and F. Schögl. *Thermodynamics of chaotic systems: an introduction*. Vol. 4. Cambridge University Press, 1995 (cit. on pp. 14, 25).
- [BT05] C. Braga and K. Travis. “A configurational temperature Nosé-Hoover thermostat”. In: *J. Chem. Phys.* 123 (2005), p. 134101 (cit. on p. 47).
- [BTV98] W. Breyermann, T. Tél, and J. Vollmer. “Entropy balance, time reversibility, and mass transport in dynamical systems”. In: *Chaos* 8.2 (1998), pp. 396–408 (cit. on p. 82).
- [Bér+12] A. Bérut et al. “Experimental verification of Landauer’s principle linking information and thermodynamics”. In: *Nature* 483.7388 (2012), pp. 187–189 (cit. on pp. 15, 184).
- [CB12] V. Capasso and D. Bakstein. *An introduction to continuous-time stochastic processes: theory, models, and applications to finance, biology, and medicine*. Springer, 2012 (cit. on pp. 65, 189).
- [CC05] N. Carter and R. Cross. “Mechanics of the kinesin step”. In: *Nature* 435.7040 (2005), pp. 308–312 (cit. on pp. 151, 153, 155, 161, 197–199).
- [CFS82] I. P. Cornfeld, S. V. Fomin, and Y. G. Sinai. *Ergodic theory*. Springer, 1982 (cit. on p. 14).
- [Cil+13] S. Ciliberto et al. “Heat Flux and Entropy Produced by Thermal Fluctuations”. In: *Phys. Rev. Lett.* 110.18 (2013), p. 180601 (cit. on p. 9).
- [CJP10] S. Ciliberto, S. Joubaud, and A. Petrosyan. “Fluctuations in out-of-equilibrium systems: from theory to experiment”. In: *J. Stat. Mech.* 2010.12 (2010), P12003 (cit. on pp. 9, 10, 26).
- [CL00] P. M. Chaikin and T. C. Lubensky. *Principles of condensed matter physics*. Vol. 1. Cambridge University Press, 2000 (cit. on pp. 35, 36).
- [Cla54] R. Clausius. “Ueber eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmetheorie”. In: *Ann. Phys.-Leipzig*. 169.12 (1854), pp. 481–506 (cit. on p. 22).

- [Cla65] R. Clausius. “Ueber verschiedene für die Anwendung bequemer Formen der Hauptgleichungen der mechanischen Wärmetheorie”. In: *Ann. Phys.-Leipzig*. 201.7 (1865), pp. 353–400 (cit. on p. 13).
- [CMB08] Y. R. Chemla, J. R. Moffitt, and C. Bustamante. “Exact solutions for kinetic models of macromolecular dynamics”. In: *J. Phys. Chem. B* 112.19 (2008), pp. 6025–6044 (cit. on pp. 164, 165).
- [Col+11] M. Colangeli et al. “Steady state fluctuation relations and time reversibility for non-smooth chaotic maps”. In: *J. Stat. Mech.* 2011.04 (2011), P04021 (cit. on pp. 82, 103).
- [Cro99] G. Crooks. “Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences”. In: *Phys. Rev. E* 60.3 (1999), p. 2721 (cit. on pp. 26, 43).
- [Cru12] J. P. Crutchfield. “Between order and chaos”. In: *Nat. Phys.* 8.1 (2012), pp. 17–24 (cit. on p. 183).
- [CS98] H. B. Callen and H. Scott. “Thermodynamics and an Introduction to Thermostatistics”. In: *Am. J. Phys.* 66 (1998), pp. 164–167 (cit. on pp. 28, 166).
- [CS99] J. P. Crutchfield and C. R. Shalizi. “Thermodynamic depth of causal states: Objective complexity via minimal representations”. In: *Phys. Rev. E* 59.1 (1999), p. 275 (cit. on p. 183).
- [CW51] H. B. Callen and T. A. Welton. “Irreversibility and generalized noise”. In: *Phys. Rev.* 83.1 (1951), pp. 34–40 (cit. on p. 33).
- [Der83] B. Derrida. “Velocity and diffusion constant of a periodic one-dimensional hopping model”. In: *J. Stat. Phys.* 31.3 (1983), pp. 433–450 (cit. on p. 164).
- [Din71] E. I. Dinaburg. “On the relations among various entropy characteristics of dynamical systems”. In: *Izvestiya: Mathematics* 5.2 (1971), pp. 337–378 (cit. on p. 60).
- [Dob73] T. Dobzhansky. “Nothing in biology makes sense except in the light of evolution”. In: *The American Biology Teacher* 35 (1973), pp. 125–129 (cit. on p. 143).
- [EB10] M. Esposito and C. Van den Broeck. “Three faces of the second law. I. Master equation formulation”. In: *Phys. Rev. E* 82.1 (2010), p. 11143 (cit. on p. 120).
- [Ein05] A. Einstein. “Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen”. In: *Ann. Phys.-Leipzig*. 322.8 (1905), pp. 549–560 (cit. on pp. 31, 32).
- [ELB09] M. Esposito, K. Lindenberg, and C. Van den Broeck. “Universality of efficiency at maximum power”. In: *Phys. Rev. Lett.* 102.13 (2009), p. 130602 (cit. on p. 145).
- [Ell05] R. Ellis. *Entropy, large deviations, and statistical mechanics*. Vol. 1431. 821. Taylor & Francis, 2005 (cit. on pp. 11, 131, 132).
- [EM90] D. J. Evans and G. Morriss. *Statistical mechanics of nonequilibrium liquids*. Cambridge University Press, 1990 (cit. on pp. 13, 48).
- [ES02] D. Evans and D. Searles. “The fluctuation theorem”. In: *Adv. Phys.* 51.7 (2002), pp. 1529–1585 (cit. on pp. 48, 50, 53, 54, 57, 109, 180).

## Bibliography

- [Esp+12] M. Esposito et al. “Stochastically driven single-level quantum dot: A nanoscale finite-time thermodynamic machine and its various operational modes”. In: *Phys. Rev. E* 85.3 (2012), p. 31117 (cit. on p. 9).
- [Esp12] M. Esposito. “Stochastic thermodynamics under coarse graining”. In: *Phys. Rev. E* 85.4 (2012), p. 41125 (cit. on pp. 120, 147, 179).
- [Fan+10] D. Fange et al. “Stochastic reaction-diffusion kinetics in the microscopic limit”. In: *P.N.A.S.* 107.46 (2010), pp. 19820–19825 (cit. on p. 165).
- [Fan47] U. Fano. “Ionization yield of radiations. II. The fluctuations of the number of ions”. In: *Phys. Rev.* 72.1 (1947), p. 26 (cit. on p. 165).
- [FDP11] A. Faggionato and D. Di Pietro. “Gallavotti–Cohen-Type Symmetry Related to Cycle Decompositions for Markov Chains and Biochemical Applications”. In: *J. Stat. Phys.* 143 (2011), pp. 11–32 (cit. on pp. 11, 120, 135, 137, 156).
- [Fel68] W. Feller. *An Introduction to Probability Theory and Its Applications*. 3rd ed. Vol. 1. Wiley New York, 1968 (cit. on pp. 121, 124).
- [FT82] E. Fredkin and T. Toffoli. “Conservative logic”. In: *Int. J. Theor. Phys.* 21.3 (1982), pp. 219–253 (cit. on p. 184).
- [FW11] R. Frigg and C. Werndl. “Entropy — A Guide for the Perplexed”. In: *Probabilities in Physics* (2011) (cit. on p. 15).
- [Gal98] G. Gallavotti. “Chaotic dynamics, fluctuations, nonequilibrium ensembles”. In: *Chaos* 8.2 (1998), pp. 384–392 (cit. on p. 166).
- [Gas04] P. Gaspard. “Time-Reversed Dynamical Entropy and Irreversibility in Markovian Random Processes”. In: *J. Stat. Phys.* 117 (2004), pp. 599–615 (cit. on p. 60).
- [Gas05] P. Gaspard. *Chaos, scattering and statistical mechanics*. Vol. 9. Cambridge University Press, 2005 (cit. on pp. 82, 102, 103).
- [GC13] C. Ganguly and D. Chaudhuri. “Stochastic thermodynamics of active Brownian particles”. In: *Phys. Rev. E* 88.3 (2013), p. 32102 (cit. on p. 9).
- [GC95] G. Gallavotti and E. Cohen. “Dynamical ensembles in stationary states”. In: *J. Stat. Phys.* 80 (5 1995), pp. 931–970 (cit. on pp. 53, 84, 178, 180).
- [GD95] P. Gaspard and J. Dorfman. “Chaotic scattering theory, thermodynamic formalism, and transport coefficients”. In: *Phys. Rev. E* 52.4 (1995), p. 3525 (cit. on p. 106).
- [Gey14] L. Geyrhofer. “Quantifying Evolutionary Dynamics”. PhD thesis. Georg-August-Universität Göttingen, 2014 (cit. on p. 146).
- [GH83] J. Guckenheimer and P. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Springer-Verlag, 1983 (cit. on p. 50).
- [Gib48] J. W. Gibbs. *The collected works of J. Willard Gibbs*. Vol. 1. Yale University Press, 1948 (cit. on pp. 28, 29, 174).

- [GM84] S. de Groot and P. Mazur. *Non-equilibrium thermodynamics*. Dover, 1984 (cit. on pp. 10, 13, 26, 34).
- [GO97] M. Grmela and H. C. Öttinger. “Dynamics and thermodynamics of complex fluids. I. Development of a general formalism”. In: *Phys. Rev. E* 56 (6 1997), pp. 6620–6632 (cit. on p. 9).
- [Goo71] T. N. Goodman. “Relating topological entropy and measure entropy”. In: *B. Lond. Math. Soc.* 3.2 (1971), pp. 176–180 (cit. on p. 60).
- [Goo71] L. W. Goodwyn. “The product theorem for topological entropy”. In: *T. Am. Math. Soc.* 158 (1971), pp. 445–452 (cit. on p. 60).
- [GW93] P. Gaspard and X.-J. Wang. “Noise, chaos, and  $(\epsilon, \tau)$ -entropy per unit time”. In: *Phys. Rep.* 235.6 (1993), pp. 291–343 (cit. on pp. 106, 181).
- [Hal10] O. Hallatschek. “The noisy edge of traveling waves”. In: *P.N.A.S.* 108.5 (2010), p. 1783 (cit. on p. 146).
- [HBS14] D. Hartich, A. C. Barato, and U. Seifert. “Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a Maxwell’s demon interpretation”. In: *J. Stat. Mech.* 2014.2 (2014), P02016 (cit. on pp. 15, 185).
- [Her14] S. Herminghaus. *Wet Granular Matter: A Trule Complex Fluid*. Vol. Volume 6. Series in Soft Condensed Matter. World Scientific, 2014 (cit. on p. 15).
- [Hil77] T. Hill. *Free energy transduction in biology*. Academic Press New York, 1977 (cit. on pp. 31, 40, 41, 123, 124, 137, 140, 143, 156, 168).
- [Hil79] T. L. Hill. “Steady-state phase or cooperative transitions between biochemical cycles”. In: *P.N.A.S.* 76.2 (1979), pp. 714–716 (cit. on pp. 138, 164).
- [Hoo83] W. G. Hoover. “Nonequilibrium molecular dynamics”. In: *Annu. Rev. Phys. Chem.* 34.1 (1983), pp. 103–127 (cit. on pp. 13, 48).
- [Hop48] E. Hopf. *Ergodentheorie*. Chelsea Publishing Company, 1948 (cit. on pp. 14, 15, 102).
- [How97] J. Howard. “Molecular motors: structural adaptations to cellular functions”. In: *Nature* 389.6651 (1997), pp. 561–567 (cit. on p. 145).
- [HP08] B. Hasselblatt and Y. Pesin. “Pesin entropy formula”. In: *Scholarpedia* 3.3 (2008), p. 3733 (cit. on p. 60).
- [HP11] J. Horowitz and J. Parrondo. “Designing optimal discrete-feedback thermodynamic engines”. In: *New. J. Phys.* 13.12 (2011), p. 123019 (cit. on pp. 9, 15, 184).
- [HS01] T. Hatano and S.-i. Sasa. “Steady-state thermodynamics of Langevin systems”. In: *Phys. Rev. Lett.* 86.16 (2001), p. 3463 (cit. on p. 87).
- [HS07] R. Harris and G. Schütz. “Fluctuation theorems for stochastic dynamics”. In: *J. Stat. Mech.* 2007 (2007), P07020 (cit. on p. 125).
- [Jar97] C. Jarzynski. “Nonequilibrium equality for free energy differences”. In: *Phys. Rev. Lett.* 78.14 (1997), p. 2690 (cit. on p. 26).

## Bibliography

- [Jay03] E. T. Jaynes. *Probability theory: the logic of science*. Cambridge university press, 2003 (cit. on pp. 14, 93, 111).
- [Jay57] E. T. Jaynes. “Information Theory and Statistical Mechanics”. In: *Phys. Rev.* 106 (4 1957), pp. 620–630 (cit. on pp. 14, 93, 96, 174).
- [Jay93] E. T. Jaynes. “A backward look to the future”. In: *Physics and Probability*. Cambridge University Press, 1993, pp. 261–75 (cit. on p. 91).
- [Jos06] J. Jost. *Dynamical systems: examples of complex behaviour*. Springer, 2006 (cit. on p. 58).
- [JQQ04] D. Jiang, M. Qian, and M.-P. Qian. *Mathematical theory of nonequilibrium steady states*. Springer, Berlin Heidelberg, 2004 (cit. on p. 137).
- [JR10] O. Jepps and L. Rondoni. “Deterministic thermostats, theories of nonequilibrium systems and parallels with the ergodic condition”. In: *J. Phys. A* 43 (2010), p. 133001 (cit. on pp. 46, 48, 50, 52, 106, 109).
- [Kal06] S. Kalpazidou. *Cycle representations of Markov processes*. Vol. 28. Springer, Berlin, 2006 (cit. on p. 137).
- [Kal93] S. Kalpazidou. “On Levy’s theorem concerning positiveness of transition probabilities of Markov processes: the circuit processes case”. In: *J. Appl. Probab.* 30 (1993), pp. 28–39 (cit. on p. 137).
- [Khi13] A. I. Khinchin. *Mathematical foundations of statistical mechanics*. Courier Dover Publications, 2013 (cit. on pp. 14, 28).
- [Khi57] A. I. Khinchin. *Mathematical foundations of information theory*. Vol. 434. Courier Dover Publications, 1957 (cit. on p. 24).
- [Kir47] G. Kirchhoff. “Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird”. In: *Ann. Phys.-Leipzig*. 148.12 (1847), pp. 497–508 (cit. on pp. 121, 137).
- [KN13] K. Kawaguchi and Y. Nakayama. “Fluctuation theorem for hidden entropy production”. In: *Phys. Rev. E* 88 (2 2013), p. 22147 (cit. on p. 113).
- [Kol58] A. N. Kolmogorov. “A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces”. In: *Dokl. Akad. Nauk SSSR (NS)*. Vol. 119. 1958, pp. 861–864 (cit. on p. 54).
- [Kri80] W. Krieger. *On the subsystems of topological Markov chains*. Cambridge University Press, 1980 (cit. on p. 70).
- [Kur98] J. Kurchan. “Fluctuation theorem for stochastic dynamics”. In: *J. Phys. A* 31 (1998), p. 3719 (cit. on pp. 10, 31, 43).
- [Lan+12] G. Lan et al. “The energy-speed-accuracy trade-off in sensory adaptation”. In: *Nat. Phys.* 8.5 (2012), pp. 422–428 (cit. on p. 184).
- [Lan08] P. Langevin. “Sur la théorie du mouvement brownien”. In: *CR Acad. Sci. Paris* 146.530-533 (1908), pp. 530–533 (cit. on pp. 31, 35).



- [Lan61] R. Landauer. “Irreversibility and heat generation in the computing process”. In: *Ibm. J. Res. Dev.* 5.3 (1961), pp. 183–191 (cit. on pp. 15, 183).
- [Lar+74] S. Larsen et al. “Change in direction of flagellar rotation is the basis of the chemotactic response in *Escherichia coli*.” In: *Nature* 249.452 (1974), p. 74 (cit. on p. 147).
- [Leb93] J. Lebowitz. “Boltzmann’s entropy and time’s arrow”. In: *Phys. Today* 46 (1993), pp. 32–32 (cit. on p. 87).
- [Lip+01] J. Liphardt et al. “Reversible unfolding of single RNA molecules by mechanical force”. In: *Science* 292.5517 (2001), pp. 733–737 (cit. on p. 9).
- [LL07] S. Liepelt and R. Lipowsky. “Kinesin’s network of chemomechanical motor cycles”. In: *Phys. Rev. Lett.* 98.25 (2007), p. 258102 (cit. on pp. 144, 145, 151–155, 157, 158, 161, 169, 177, 197–200).
- [LL08] R. Lipowsky and S. Liepelt. “Chemomechanical Coupling of Molecular Motors: Thermodynamics, Network Representations, and Balance Conditions”. In: *J. Stat. Phys.* 130 (1 2008), pp. 39–67 (cit. on pp. 9, 167).
- [LL09] S. Liepelt and R. Lipowsky. “Operation modes of the molecular motor kinesin”. In: *Phys. Rev. E* 79.1 (2009), p. 11917 (cit. on pp. 151, 153, 155, 157, 197).
- [LLSB13] G. Lajoie, K. K. Lin, and E. Shea-Brown. “Chaos and reliability in balanced spiking networks with temporal drive”. In: *Phys. Rev. E* 87.5 (2013), p. 52901 (cit. on p. 181).
- [LLV09] R. Lipowsky, S. Liepelt, and A. Valleriani. “Energy Conversion by Molecular Motors Coupled to Nucleotide Hydrolysis”. In: *J. Stat. Phys.* 135.5 (2009), pp. 951–975 (cit. on p. 145).
- [LS99] J. Lebowitz and H. Spohn. “A Gallavotti–Cohen type symmetry in the large deviation functional for stochastic dynamics”. In: *J. Stat. Phys.* 95 (1999), p. 333 (cit. on pp. 10, 31, 43, 166).
- [Luc97] Y. Lucet. “Faster than the fast Legendre transform, the linear-time Legendre transform”. In: *Numer. Algor.* 16.2 (1997), pp. 171–185 (cit. on p. 155).
- [Mae04] C. Maes. “On the origin and the use of fluctuation relations for the entropy”. In: *Poincaré Seminar 2003: Bose-Einstein condensation-entropy*. Birkhäuser, Basel, 2004, p. 145 (cit. on pp. 10, 11, 26, 28, 31, 43).
- [MK72] R. M. Macnab and D. Koshland. “The gradient-sensing mechanism in bacterial chemotaxis”. In: *P.N.A.S.* 69.9 (1972), pp. 2509–2512 (cit. on p. 147).
- [ML10] V. Mustonen and M. Lässig. “Fitness flux and ubiquity of adaptive evolution”. In: *P.N.A.S.* 107.9 (2010), pp. 4248–4253 (cit. on p. 146).
- [Mor65] H. Mori. “Transport, Collective Motion, and Brownian Motion”. In: *Prog. Theor. Phys.* 33 (1965), pp. 423–455 (cit. on p. 35).
- [MV03] C. Maes and E. Verbitskiy. “Large deviations and a fluctuation symmetry for chaotic homeomorphisms”. In: *Commun. Math. Phys.* 233.1 (2003), pp. 137–151 (cit. on p. 178).

## Bibliography

- [MW10] M. Monteforte and F. Wolf. “Dynamical entropy production in spiking neuron networks in the balanced state”. In: *Phys. Rev. Lett.* 105.26 (2010), p. 268104 (cit. on p. 181).
- [NP77] G. Nicolis and I. Prigogine. *Self-organization in nonequilibrium systems*. Vol. 191977. Wiley New York, 1977 (cit. on p. 181).
- [Ons31] L. Onsager. “Reciprocal Relations in Irreversible Processes. I.” In: *Phys. Rev.* 37 (4 1931), pp. 405–426 (cit. on p. 33).
- [Par64] W. Parry. “Intrinsic markov chains”. In: *T. Am. Math. Soc.* 112.1 (1964), pp. 55–66 (cit. on pp. 70, 78).
- [Pat69] K. Paton. “An Algorithm for Finding a Fundamental Set of Cycles of a Graph”. In: *Commun. Acm.* 12 (1969), p. 514 (cit. on p. 137).
- [Pen70] O. Penrose. *Foundations of statistical mechanics; a deductive treatment*. Oxford: Pergamon Press, 1970 (cit. on pp. 12, 30, 61, 63, 88, 173, 182).
- [Pol12] M. Poletini. “Nonequilibrium thermodynamics as a gauge theory”. In: *Europhys. Lett.* 97.3 (2012), p. 30003 (cit. on pp. 42, 115, 116, 164).
- [Pol14] M. Poletini. “Cycle/cocycle oblique projections on oriented graphs”. In: *arXiv* 1405.0899 (2014) (cit. on pp. 135–137, 139).
- [Pop02] K. Popper. *The logic of scientific discovery*. Psychology Press, 2002 (cit. on pp. 87, 94, 182).
- [Pug+10] A. Puglisi et al. “Entropy production and coarse graining in Markov processes”. In: *J. Stat. Mech.* 2010 (2010), P05015 (cit. on p. 147).
- [Pur77] E. M. Purcell. “Life at low Reynolds number”. In: *Am. J. Phys.* 45.1 (1977), pp. 3–11 (cit. on p. 145).
- [PV08] S. Pigolotti and A. Vulpiani. “Coarse graining of master equations with fast and slow states”. In: *J. Chem. Phys.* 128 (2008), p. 154114 (cit. on pp. 147, 148).
- [PV13] M. Poletini, and F. Vedovato. “Of dice and men. Subjective priors, gauge invariance, and nonequilibrium thermodynamics”. In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013*. Ed. by M. Pilotelli and G. Beretta. 2013 (cit. on pp. 115, 116).
- [Qia05] H. Qian. “Cycle kinetics, steady state thermodynamics and motors—a paradigm for living matter physics”. In: *J. Phys.: Condens. Matter* 17 (2005), S3783 (cit. on pp. 9, 145).
- [QK14] H. Qian and S. Kou. “Statistics and related topics in single-molecule biophysics”. In: *Annual Review of Statistics and Its Application* 1 (2014), pp. 465–492 (cit. on p. 165).
- [RDD05] P.-E. Roche, B. Derrida, and B. Douçot. “Mesoscopic full counting statistics and exclusion models”. In: *Eur. Phys. J. B* 43.4 (2005), pp. 529–541 (cit. on p. 165).
- [Rom+12] P. Romanczuk et al. “Active brownian particles”. In: *Eur. Phys. J.: Special Topics* 202.1 (2012), pp. 1–162 (cit. on p. 9).

- [RTV00] L. Rondoni, T. Tél, and J. Vollmer. “Fluctuation theorems for entropy production in open systems”. In: *Phys. Rev. E* 61.5 (2000), pp. 4679–4682 (cit. on pp. 103, 106).
- [Rue04] D. Ruelle. *Thermodynamic formalism: the mathematical structure of equilibrium statistical mechanics*. Cambridge University Press, 2004 (cit. on pp. 14, 54, 88).
- [Rue99] D. Ruelle. “Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics”. In: *J. Stat. Phys.* 95.1 (1999), pp. 393–468 (cit. on pp. 15, 30, 51, 52, 88, 166).
- [Rug97] H. Rugh. “Dynamical Approach to Temperature”. In: *Phys. Rev. Lett.* 78.5 (1997), p. 772 (cit. on p. 47).
- [Rén61] A. Rényi. “On measures of entropy and information”. In: *Fourth Berkeley Symposium on Mathematical Statistics and Probability*. 1961, pp. 547–561 (cit. on p. 25).
- [Sag12] T. Sagawa. *Thermodynamics of information processing in small systems*. Springer Verlag, 2012 (cit. on pp. 15, 184).
- [Sar+14] P. Sartori et al. “Thermodynamic costs of information processing in sensory adaption”. In: *arXiv* 1404.1027 (2014) (cit. on p. 184).
- [SC01] C. R. Shalizi and J. P. Crutchfield. “Computational mechanics: Pattern and prediction, structure and simplicity”. In: *J. Stat. Phys.* 104.3 (2001), pp. 817–879 (cit. on p. 183).
- [Sch76] J. Schnakenberg. “Network theory of microscopic and macroscopic behavior of master equation systems”. In: *Rev. Mod. Phys.* 48.4 (1976), pp. 571–585 (cit. on pp. 31, 40, 41, 120, 123, 127, 129, 136, 137, 156, 199).
- [Sch92] E. Schrödinger. *What is life?: With mind and matter and autobiographical sketches*. Cambridge University Press, 1992 (cit. on p. 143).
- [SDC07] A. Samoletov, C. Dettmann, and M. Chaplain. “Thermostats for “slow” configurational modes”. In: *J. Stat. Phys.* 128.6 (2007), pp. 1321–1336 (cit. on pp. 45–47).
- [SDC10] A. Samoletov, C. Dettmann, and M. Chaplain. “Notes on configurational thermostat schemes.” In: *J. Chem. Phys.* 132.24 (2010), p. 246101 (cit. on p. 47).
- [Sei05] U. Seifert. “Entropy production along a stochastic trajectory and an integral fluctuation theorem”. In: *Phys. Rev. Lett.* 95.4 (2005), p. 40602 (cit. on pp. 10, 11, 31, 34, 42, 43, 145, 166).
- [Sei08] U. Seifert. “Stochastic thermodynamics: principles and perspectives”. In: *Eur. Phys. J. B* 64 (3 2008), pp. 423–431 (cit. on pp. 9, 31).
- [Sei11] U. Seifert. “Stochastic thermodynamics of single enzymes and molecular motors”. In: *Eur. Phys. J. E* 34.3 (2011), pp. 1–11 (cit. on pp. 9, 12, 87, 111, 145, 156).
- [Sei12] U. Seifert. “Stochastic thermodynamics, fluctuation theorems, and molecular machines”. In: *Rep. Prog. Phys.* 75 (2012), p. 126001 (cit. on pp. 9, 11, 26, 31, 32, 34–36, 43, 79, 88, 115, 180, 181).
- [Sek98] K. Sekimoto. “Langevin equation and thermodynamics”. In: *Sup. Prog. Theor. Phys.* 130 (1998), pp. 17–27 (cit. on pp. 10, 34).

## Bibliography

- [Sha48] C. E. Shannon. “A Mathematical Theory of Communication”. In: *At.&T. Tech. J.* 27 (1948), 379—423, 623–656 (cit. on pp. 14, 23).
- [Sin09] Y. Sinai. “Kolmogorov-Sinai entropy”. In: *Scholarpedia* 4.3 (2009), p. 2034 (cit. on pp. 54, 73).
- [Sin68] Y. G. Sinai. “Markov partitions and C-diffeomorphisms”. In: *Funct. Anal. Appl.* 2.1 (1968), pp. 61–82 (cit. on pp. 77, 177).
- [Sin72] Y. G. Sinai. “Gibbs measures in ergodic theory”. In: *Russ. Math. Surv.* 27.4 (1972), p. 21 (cit. on pp. 14, 84, 177).
- [SM03] C. R. Shalizi and C. Moore. “What is a macrostate? Subjective observations and objective dynamics”. In: *arXiv cond-mat/0303625* (2003) (cit. on p. 183).
- [Smo15] M. Smoluchowski. “Über Brownsche Molekularbewegung unter Einwirkung äußerer Kräfte und deren Zusammenhang mit der verallgemeinerten Diffusionsgleichung”. In: *Ann. Phys* 48 (1915), p. 1103 (cit. on pp. 31, 32).
- [SP13] P. Sartori and S. Pigolotti. “Kinetic versus Energetic Discrimination in Biological Copying”. In: *Phys. Rev. Lett.* 110.18 (2013), p. 188101 (cit. on p. 184).
- [Spe+07] T. Speck et al. “Distribution of entropy production for a colloidal particle in a nonequilibrium steady state”. In: *Europhys. Lett.* 79.3 (2007), p. 30002 (cit. on p. 9).
- [SRE07] D. J. Searles, L. Rondoni, and D. J. Evans. “The steady state fluctuation relation for the dissipation function”. In: *J. Stat. Phys.* 128.6 (2007), pp. 1337–1363 (cit. on pp. 53, 54, 57, 109).
- [SS07] T. Schmiedl and U. Seifert. “Efficiency at maximum power: An analytically solvable model for stochastic heat engines”. In: *Europhys. Lett.* 81 (2007), p. 20003 (cit. on p. 145).
- [Sti+12] S. Still et al. “Thermodynamics of Prediction”. In: *Phys. Rev. Lett.* 109.12 (2012), p. 120604 (cit. on p. 185).
- [SU10] T. Sagawa and M. Ueda. “Generalized Jarzynski equality under nonequilibrium feedback control”. In: *Phys. Rev. Lett.* 104.9 (2010), p. 90602 (cit. on pp. 15, 184).
- [SU12] T. Sagawa and M. Ueda. “Nonequilibrium thermodynamics of feedback control”. In: *Phys. Rev. E* 85.2 (2012), p. 21104 (cit. on p. 184).
- [Szi29] L. Szilard. “Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen”. In: *Zeitschrift für Physik* 53.11-12 (1929), pp. 840–856 (cit. on p. 15).
- [Tao11] T. Tao. *An introduction to measure theory*. Vol. 126. AMS Bookstore, 2011 (cit. on pp. 55, 190).
- [Tka+13] G. Tkačik et al. “The simplest maximum entropy model for collective behavior in a neural network”. In: *J. Stat. Mech.* 2013.03 (2013), P03011 (cit. on p. 181).
- [TM71] M. Tribus and E. C. McIrvine. “Energy and information”. In: *Sci. Am.* 225.3 (1971), pp. 179–188 (cit. on p. 21).

- [Tou09] H. Touchette. “The large deviation approach to statistical mechanics”. In: *Phys. Rep.* 478.1-3 (2009), pp. 1–69 (cit. on pp. 11, 131).
- [Tou11] H. Touchette. “A basic introduction to large deviations: Theory, applications, simulations”. In: *arXiv* 1106.4146 (2011) (cit. on pp. 131, 133, 140, 178).
- [Toy+10] S. Toyabe et al. “Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality”. In: *Nat. Phys.* 6 (2010), p. 988 (cit. on pp. 9, 15, 184).
- [Tsa88] C. Tsallis. “Possible generalization of Boltzmann-Gibbs statistics”. In: *J. Stat. Phys.* 52.1-2 (1988), pp. 479–487 (cit. on p. 25).
- [Tut98] W. T. Tutte. *Graph theory as I have known it*. Oxford University Press, 1998 (cit. on pp. 119, 124, 128, 137).
- [TV00] T. Tél and J. Vollmer. “Entropy balance, multibaker maps, and the dynamics of the Lorentz gas”. In: *Hard Ball Systems and the Lorentz Gas*. Springer Berlin, 2000, pp. 367–418 (cit. on pp. 82, 104).
- [Var66] S. S. Varadhan. “Asymptotic probabilities and differential equations”. In: *Commun. Pur. Appl. Math.* 19.3 (1966), pp. 261–286 (cit. on p. 11).
- [VK92] N. G. Van Kampen. *Stochastic processes in physics and chemistry*. Vol. 1. Elsevier, 1992 (cit. on pp. 10, 12, 31).
- [Vol02] J. Vollmer. “Chaos, spatial extension, transport, and non-equilibrium thermodynamics”. In: *Phys. Rep.* 372.2 (2002), pp. 131–267 (cit. on pp. 82, 102–104).
- [VS96] C. van Vreeswijk and H. Sompolinsky. “Chaos in neuronal networks with balanced excitatory and inhibitory activity”. In: *Science* 274.5293 (1996), pp. 1724–1726 (cit. on p. 181).
- [VTB97] J. Vollmer, T. Tél, and W. Breymann. “Equivalence of Irreversible Entropy Production in Driven Systems: An Elementary Chaotic Map Approach”. In: *Phys. Rev. Lett.* 79 (15 1997), pp. 2759–2762 (cit. on pp. 82, 103).
- [VTB98] J. Vollmer, T. Tél, and W. Breymann. “Entropy balance in the presence of drift and diffusion currents: an elementary chaotic map approach”. In: *Phys. Rev. E* 58.2 (1998), p. 1672 (cit. on pp. 82, 102).
- [Wac13] A. Wachtel. “Fluctuation Spectra and Coarse Graining in Stochastic Dynamics”. In: *arXiv* 1312.0115 (2013) (cit. on pp. 120, 134, 135).
- [WAV p] A. Wachtel, B. Altaner, and J. Vollmer. “Highly nonlinear and negative differential response in a model for the molecular motor kinesin”. (in preparation) (cit. on pp. 144, 166).
- [Wei73] B. Weiss. “Subshifts of finite type and sofic systems”. In: *Monatshefte für Mathematik* 77.5 (1973), pp. 462–474 (cit. on p. 70).

## Bibliography

- [Wil73] R. Williams. “Classification of subshifts of finite type”. In: *Recent Advances in Topological Dynamics*. Springer, 1973, pp. 281–285 (cit. on p. 70).
- [Woj09] M. P. Wojtkowski. “Abstract fluctuation theorem”. In: *Ergod. Theor. Dyn. Syst.* 29.01 (2009), pp. 273–279 (cit. on pp. 53, 106, 115, 180).
- [Wol+14] F. Wolf et al. “Dynamical models of cortical circuits”. In: *Curr. Opin. Neurobiol.* 25 (2014), pp. 228–236 (cit. on p. 181).
- [WVA14] A. Wachtel, J. Vollmer, and B. Altaner. “Determining the Statistics of Fluctuating Currents: General Markovian Dynamics and its Application to Motor Proteins”. In: *arXiv* 1407.2065 (2014) (cit. on pp. 120, 134, 135, 144, 156).
- [Yil+04] A. Yildiz et al. “Kinesin walks hand-over-hand”. In: *Science* 303 (2004), p. 676 (cit. on pp. 151, 152, 197).
- [You02] L.-S. Young. “What are SRB measures, and which dynamical systems have them?” In: *J. Stat. Phys.* 108.5 (2002), pp. 733–754 (cit. on pp. 50, 81, 84).
- [ZS07] R. Zia and B. Schmittmann. “Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states”. In: *J. Stat. Mech.* 2007 (2007), P07012 (cit. on pp. 124, 137, 140).
- [Zwa61] R. Zwanzig. “Memory Effects in Irreversible Thermodynamics”. In: *Phys. Rev.* 124.4 (1961), pp. 983–992 (cit. on p. 35).

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>.

