Contents

CHAPTER 1

Motivation

Stochastic thermodynamics has become an established branch of nonequilibrium statistical physics. On the theoretical side, it has been discovered that the behavior of mesoscopic systems is governed by surprisingly general relations. Rapid advances in experimental techniques are leading to tests of these relations by manipulating mesoscopic physical systems.

Perhaps as a consequence of this fast development, some aspects of stochastic thermodynamics might seem obscure to noninitiates. Key results in stochastic thermodynamics, such as fluctuation relations, are so general that one might wonder what the underlying physical assumptions really are. In a broader perspective, it might be difficult to understand how the simplicity of stochastic thermodynamics relates to the daunting complexity of traditional nonequilibrium statistical physics. Preliminary answers to these questions are presented in this chapter. We conclude the chapter with an overview of the book structure.

1.1 What is stochastic thermodynamics?

In its simplest form, **stochastic thermodynamics** is a thermodynamic theory for mesoscopic, nonequilibrium physical systems interacting with equilibrium heat reservoirs.

It is useful to dissect this definition:

- *Thermodynamic theory.* As the name suggests, stochastic thermodynamics draws a correspondence between mesoscopic stochastic dynamics and macroscopic thermodynamics. This correspondence is sketched in fig. 1.1.
- *Mesoscopic, nonequilibrium physical systems.* Stochastic thermodynamics deals with **mesoscopic systems**. For our aims, a mesoscopic system is a physical system characterized by energy differences among its states on the order of the thermal energy $k_{\mathrm{B}}T$, where k_{B} is the Boltzmann constant and T is the temperature. Prominent examples of mesoscopic systems are colloidal particles, macromolecules, nanodevices, or systems of chemical reactions at very low densities. Mesoscopic systems can be driven out of equilibrium by an external manipulation, for example, by varying in time the temperature or by controlling them with optical tweezers. More generally, all physical systems that can be described by stochastic evolution equations, where the noise

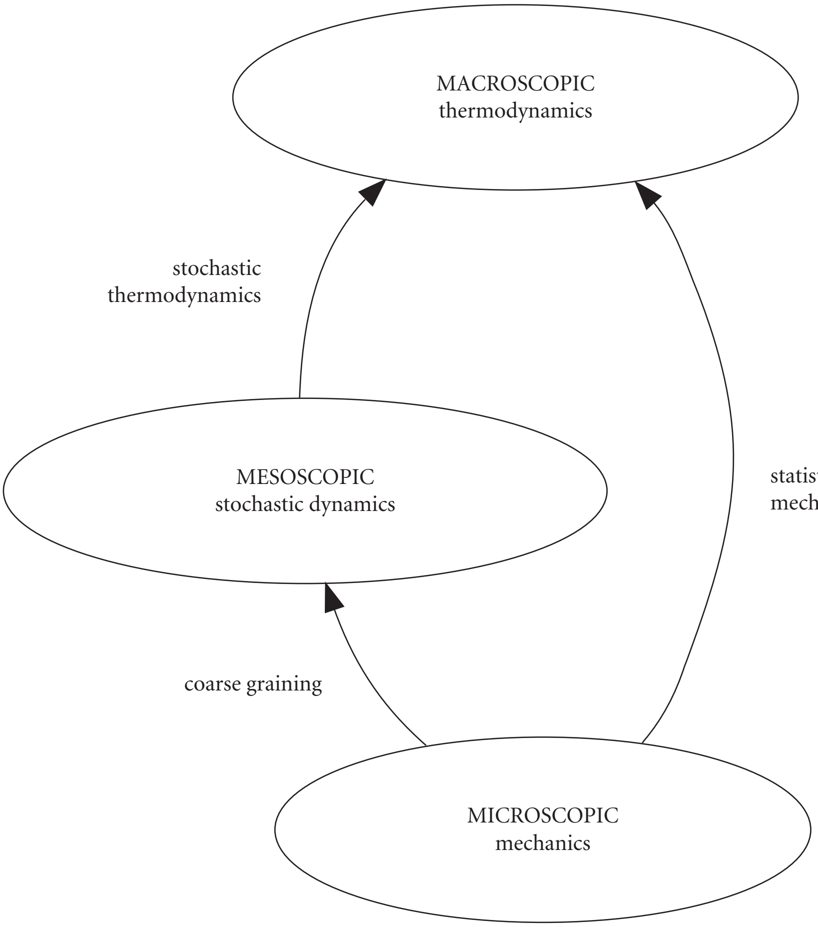

Figure 1.1. Relation between statistical mechanics, coarse-graining techniques, and stochastic thermodynamics. (Inspired by Sekimoto [153].)

models interactions with a heat reservoir, fall into the scope of stochastic thermodynamics.

- *Interacting.* The stochasticity of mesoscopic systems results from interactions with one or multiple reservoirs. Often one does not know the details of these interactions, and the functional form of the noise is dictated by general physical assumptions. Importantly, we always neglect the interaction energy between the system and the bath.
- *Equilibrium heat reservoirs.* We assume that reservoirs relax very quickly to equilibrium compared to the timescales of mesoscopic systems. Therefore, on these timescales, reservoirs are effectively always at equilibrium. This **timescale separation** is key to the simplicity of stochastic thermodynamics. In many cases, this assumption can be justified by a coarse-graining procedure.

Provocatively, one could say that Einstein's paper on Brownian motion was the first paper in history complying with our definition of stochastic thermodynamics. Indeed, Einstein considered the stochastic dynamics of a mesoscopic colloid and used it to draw far-reaching conclusions for the thermodynamics of general nonequilibrium systems.

1.2 Why does it work and why is it useful?

Stochastic thermodynamics associates thermodynamic quantities with mesoscopic physical systems, whose evolution is described by stochastic dynamics, and predicts their properties. This task is considerably simpler than the general problem in nonequilibrium statistical physics, i.e., that of deriving macroscopic dynamics from a "fundamental" microscopic description.

In particular, several fundamental problems that arise in nonequilibrium statistical physics do not even appear in stochastic thermodynamics. One example is understanding how the irreversible nature of macroscopic thermodynamics systems emerges from microscopic dynamics. This conceptual issue is known as *Loschmidt's paradox* and has puzzled physicists since the dawn of thermodynamics. After all, macroscopic systems are made of a large number of elementary particles, and these particles evolve according to microscopic equations that are time-reversible. It is nowadays established that the solution to the Loschmidt paradox originates from the large number of degrees of freedom of thermodynamic systems. However, a rigorous derivation of irreversible macroscopic dynamics starting from reversible microscopic dynamics has proved to be rather difficult and has been carried out without simplifying assumptions only for a limited number of idealized systems. This difficulty does not arise in stochastic thermodynamics, since the stochastic equations that constitute its starting point are already irreversible.

At this point, one might wonder whether stochastic thermodynamics might be *too simple* to be really interesting. In particular, one question is, In which approximation do real mesoscopic physical systems satisfy the hypotheses of stochastic thermodynamics? Although theoretical arguments can partially answer this question, an ultimate response can only come from experiments. Stochastic thermodynamics has been successfully employed to measure equilibrium free energies from nonequilibrium measurements, and the range of mesoscopic physical systems that are experimentally controllable keeps growing. It is our hope that these experiments will clarify how safe it is to apply stochastic thermodynamics to generic mesoscopic systems and which aspects have to be treated with special care.

Due to its simplifying assumptions, stochastic thermodynamics circumvents many technical subtleties of kinetic theory. In this respect, it might appear that it provides less interesting challenges for mathematical physics. However, stochastic thermodynamics has proved to be an interesting playground for advanced mathematical tools to analyze stochastic processes, including control theory, large deviations, and probability measures in the space of trajectories. As shown in this book, these tools are precious to shed light on the nature of nonequilibrium mesoscopic processes.

1.3 Plan of the work

The main goal of this book is to provide a pedagogical introduction to stochastic thermodynamics for graduate students in physics. Our book is structured so that it can be used as a textbook for a graduate course or for independent study. To this aim, we mark those sections of the book that contain more advanced material with the notation (*). These sections can be skipped to ease a first reading or to use the book for a course covering basic concepts only.

Relevant bibliography is collected in the "Further reading" sections at the end of each chapter. These sections refer to works where results discussed in the chapter were originally presented, along with other useful references to deepen the study of specific topics.

One of the best ways to learn a subject is by problem solving. Following this philosophy, we have included an exercise section to complement most chapters. Some of these exercises are meant to be solved with paper and pen, whereas others require computer simulations. These latter exercises assume that the reader is familiar with basic computer programming (in any language). For reasons of space, we do not introduce numerical algorithms other than particularly relevant ones, such as the Gillespie algorithm. As with sections, some exercises are marked with a (*), to warn the reader that their solution is particularly challenging or that it requires concepts from a starred section.

Chapter 2 provides a brief overview of the theories upon which stochastic thermodynamics is built: thermodynamics, statistical mechanics, the theory of stochastic processes, and information theory. This overview is far from exhaustive due to reasons of space. We focus on aspects of these theories of more relevance for stochastic thermodynamics. To avoid overburdening the book with the complexities of stochastic calculus, we mainly focus on physical systems with discrete states that can be described by master equations. Throughout the book, sections requiring knowledge of stochastic processes with continuous state space are always starred.

Chapter 3 introduces the basic concepts of stochastic thermodynamics. We discuss how work, heat, and entropy can be consistently introduced at the level of single stochastic trajectories of systems described by master equations. We show that these quantities satisfy relations that are analogous to the first and second laws of traditional thermodynamics. Particular emphasis is given to the physical interpretation of these quantities.

Chapter 4 is devoted to fluctuation relations, which are perhaps the most celebrated results in stochastic thermodynamics. The connection between entropy production and statistical irreversibility is the core concept of this chapter. We exploit this connection to introduce fluctuation relations in a unified framework.

Chapter 5 discusses manipulation of information at the mesoscopic level. In the first part of the chapter, we introduce counterintuitive physical aspects of information processing, such as Maxwell demons and the Landauer cost of erasing information. In the rest of the chapter, we show how stochastic thermodynamics clarifies these concepts, both in general and in the context of concrete examples.

Chapter 6 is devoted to large deviation theory. Once confined to pure mathematics and statistics, large deviation theory has risen as a fundamental tool in statistical physics and beyond. Many current developments in stochastic thermodynamics heavily rely on large deviation theory. In this chapter, we introduce the theory and discuss its main applications in stochastic thermodynamics.

Chapter 7 presents key experimental results in stochastic thermodynamics. A focus of this chapter is how fluctuation theorems allow for estimating equilibrium free energy from nonequilibrium measurements. We also discuss other groundbreaking experiments that have tested manipulation of information at the mesoscopic level.

Chapter 8 presents a collection of developments of stochastic thermodynamics in several directions. Sections of this chapter are rather independent of each other and provide an introduction to more recent research topics. We have not marked with a (*) sections and exercises of this chapter; however, most of the material should be considered as advanced.

Chapter 9 presents perspectives on open issues and future directions.

CHAPTER 2

Basics

The prerequisites for stochastic thermodynamics are laid down in this chapter. In particular, we briefly review the thermodynamics of macroscopic systems, the statistical description of their behavior, both static and dynamic, and some basic concepts of information theory. We take advantage of this preliminary material to introduce most of the notation used in the rest of the book.

2.1 Thermodynamics

Stochastic thermodynamics describes thermodynamic processes taking place in small systems in contact with reservoirs. Before going any further, we need to specify what we mean by "thermodynamic processes" and "reservoirs," and to introduce other main concepts in traditional thermodynamics.

In its basic form, thermodynamics deals with systems that are made up of a very large number of particles and are in thermodynamic equilibrium, in the sense that their macroscopically observable properties, like density, pressure, etc., do not change with time. It is a common observation that many materials, kept isolated from the environment, reach sooner or later a state characterized by constant values of macroscopic properties. Importantly, systems that reach such an equilibrium state by different manipulations behave from then on in the same way, from the point of view of thermodynamics. *Thermodynamic equilibrium wipes out previous history.* For example, the thermodynamic behavior of two glasses of water with the same density at the same pressure does not depend on whether one of the two has been prepared by melting an ice cube in a glass of water at a higher temperature, while the other has simply kept the state it had when flowing out of the tap. There are exceptions to this rule: some systems, such as structural glasses, keep memory of past manipulations. Their thermodynamic behavior is more complex (and controversial). We do not deal with them.

The state of a thermodynamic system is characterized by a judiciously chosen, small set of macroscopically observed properties: its composition, its mass, its volume, the pressure acting on it, etc. In thermodynamics, knowledge of the values of these quantities in a given equilibrium state is sufficient to predict their values after the system undergoes a thermodynamic transformation. In this sense, traditional thermodynamics is *deterministic*.

For example, we consider a cylinder closed by a piston, holding a **simple fluid**, i.e., a fluid composed of a single chemical species. For such a system, the number n of moles of the fluid (or the number of particles $N = n\,N_{\mathrm{A}}$, where $N_{\mathrm{A}} \approx 6.02 \cdot 10^{23}$ is Avogadro's number), the volume V of the cylinder, and the pressure P applied by the piston are a complete set of thermodynamic observables.

Thermodynamics deals with two main kinds of transformations:

Adiabatic transformations. In adiabatic transformations, energy is exchanged between the system and the environment only in the form of work. This means that the system is **thermally isolated**, i.e., enclosed by walls that do not allow the exchange of heat. Values of observables like pressure, volume, etc. can be mechanically altered. In our example, this means that the walls of the cylinder do not allow any uncontrolled interaction with the surroundings, and only the position of the piston can be changed. We then let the piston rest in the new position, and we wait until the system reaches a new equilibrium state. The motion of the piston can be very slow (quasi-static transformation) or abrupt, meaning that the intermediate states are not necessarily equilibrium states. In either case, the transformation brings the system from one equilibrium state to a new equilibrium state, whose properties depend on the details of the transformation.

Heat exchange. In the case of heat exchange, the system interacts in an uncontrolled way with another thermodynamic system, for example, a similar container with nonisolating walls. We say that the two systems are put "in contact." This interaction in general involves energy exchange between the two systems. In this case, we must specify which macroscopic variables are kept constant, because it is not possible in general to keep, e.g., both the volume and the pressure constant.

Any transformation can be decomposed into a succession of (possibly infinitesimal) adiabatic transformations and heat exchanges.

We now consider two systems $\mathcal{S}_1$ and $\mathcal{S}_2$ that are put in contact with a much larger system $\mathcal{S}_0$, one after the other, in such a way that they both reach equilibrium. We assume that $\mathcal{S}_0$ is so large that its thermodynamic state is not significantly affected by being put in contact with either $\mathcal{S}_1$ or $\mathcal{S}_2$, whereas the thermodynamic states of $\mathcal{S}_1$ or $\mathcal{S}_2$ may in principle change. Experience shows that if, after this procedure, $\mathcal{S}_1$ and $\mathcal{S}_2$ are put in contact with each other, their thermodynamic states remain unaltered. This observation is summarized by the **zeroth law of thermodynamics**:

> Two systems, each in thermodynamic equilibrium with a third one, are in equilibrium with each other.

The zeroth law of thermodynamics allows us to define a quantity Θ that assumes the same value in systems in thermodynamic equilibrium with one another. This quantity can be made observable if we put these systems in touch with a reference system $\mathfrak{T}$, small enough not to perturb their equilibrium state, and then measure a macroscopic quantity (e.g., the volume) of $\mathfrak{T}$. The system $\mathfrak{T}$ is called **thermometer**, and Θ is an "empirical temperature." We call the large system $\mathcal{S}_0$ a **heat reservoir** characterized by a given value of Θ.

In general, there are multiple ways to perform a thermodynamic transformation from one equilibrium state to another. For example, we can increase the pressure of a fluid with a fixed value of V by putting it in contact with a heat reservoir at an empirical temperature Θ, or by letting an electrical current go through a resistor immersed in it, until a thermometer reads the same value of Θ. According to our postulates, the system reaches the same equilibrium state in all these cases. In particular, the energy contained in the system (its **internal energy**) is the same. In an adiabatic transformation, conservation of energy imposes that the change ΔE in the internal energy between the initial and the final state must be equal to the work W^{ad} performed on the system:

$$\Delta E = E_{\mathrm{f}} - E_0 = W^{\mathrm{ad}}. \tag{2.1}$$

We use in this book the convention that work is considered positive if it is performed *on* the system and negative if it is performed *by* the system. In macroscopic thermodynamic systems, the internal energy E is **extensive**, i.e., it is proportional to the size of the system (as measured by V or N, as long as the density is fixed).

In a nonadiabatic transformation between two equilibrium states, the work W performed on the system is in general different from W^{ad}. Therefore, a certain amount of energy exchanged by interactions between the system and its surroundings is not taken into account in W. We identify this energy Q with the heat exchanged in the transformation. We thus have

$$Q = W - \Delta E. \tag{2.2}$$

With this convention, Q is positive if it is released by the system and negative if it is provided to the system. Equation (2.2) embodies the **first law of thermodynamics**, which may be expressed as follows:

> The change of the internal energy of a system is equal to the difference between the work done on it and the heat released by it.

Our sign convention for work and heat is the opposite of that commonly used in traditional thermodynamics, but it turns out to be the most natural in stochastic thermodynamics. The reason is that, historically, thermodynamics originated from the study of thermal engines, and the emphasis was on the conversion of heat into work. Conversely, in stochastic thermodynamics, one is usually interested in describing dissipative systems, which convert work into heat.

The internal energy E is one of the macroscopic observables. For a simple fluid, knowledge of n, P, and V allows one to evaluate E. It is possible to invert this relation to express, e.g., P or Θ as a function of n, V, and E.

Given an equilibrium state $\mathcal{E}_0$ of a thermodynamic system, there are states $\mathcal{E}$ that cannot be reached from $\mathcal{E}_0$ via an adiabatic transformation, while the inverse transformation $\mathcal{E} \longrightarrow \mathcal{E}_0$ is possible. In this sense, thermodynamic systems possess an intrinsic *irreversibility*. A main goal of thermodynamics is to characterize the set of states that can be reached from a given state by an arbitrary combination of adiabatic transformations and heat exchanges.

The irreversibility of thermodynamic transformations is characterized by the concept of **entropy**. The entropy S is a function of the thermodynamic state of an equilibrium system. The irreversibility of thermodynamic transformations is captured by the **second law of thermodynamics**:

> The total entropy of a thermally isolated system cannot decrease.

The entropy S has the following properties:

Additivity. If a system is made up of several subsystems $\mathcal{S}_1, \dots, \mathcal{S}_k, \dots$, each at equilibrium, the entropy of the total system is equal to the sum of the entropies of the subsystems:

$$S\left(\bigcup \mathcal{S}_k\right) = \sum_k S(\mathcal{S}_k). \tag{2.3}$$

As a consequence of additivity, the entropy of a homogeneous system is extensive, i.e., it is proportional to the system size. This implies that, upon rescaling the other extensive observables, such as the number of particles, the volume, the internal energy, etc., by a factor $\lambda > 0$, one has

$$S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N). \tag{2.4}$$

Monotonicity. S increases as E increases when all other variables are kept constant. This property holds for the vast majority of practical cases. There exist some intriguing systems where S can decrease with E, but we do not deal with them in this book.

Concavity. Entropy is a concave function. This means that, given two equilibrium states $\mathcal{E}_0 = (N, V_0, E_0)$ and $\mathcal{E}_1 = (N, V_1, E_1)$ of the same system and any real number α between 0 and 1, the intermediate state $\mathcal{E}_\alpha$ characterized by $(N, (1-\alpha)V_0 + \alpha V_1, (1-\alpha)E_0 + \alpha E_1)$ satisfies

$$S(\mathcal{E}_\alpha) \geq (1-\alpha)S(\mathcal{E}_0) + \alpha S(\mathcal{E}_1), \qquad 0 \leq \alpha \leq 1. \tag{2.5}$$

Properties of concave and convex functions are summarized in appendix A.1.

The second law of thermodynamics also implies several important properties of thermodynamic systems:

Temperature. Systems in thermal equilibrium with each other share the same value of the quantity

$$\frac{1}{T} = \frac{\partial S}{\partial E}, \tag{2.6}$$

if the other extensive quantities (N, V, etc.) are kept constant. To prove this result, we consider two systems $\mathcal{S}_1$ and $\mathcal{S}_2$ in contact with each other and adiabatically insulated from their surroundings. Their total energy $E = E_1 + E_2$ is fixed. If they are in mutual equilibrium, their total entropy cannot grow upon exchanging energy in the form of heat. Thus, at equilibrium, we have

$$\frac{\partial (S_1(E_1) + S_2(E - E_1))}{\partial E_1} = \frac{1}{T_1} - \frac{1}{T_2} = 0. \tag{2.7}$$

Therefore, T acts as the empirical temperature Θ that we defined before. The monotonicity of S implies that T cannot be negative for the cases we consider.

Heat exchange. If two systems in contact with each other do not exchange work, the system with a larger value of T provides heat to the system with a smaller value of T. We call E_1 and E_2 the initial internal energies of the two systems and T_1 and T_2 their initial temperatures. After being in contact for some time, they reach equilibrium at energy values E_1' and E_2'. The total entropy is given by

$$\begin{aligned} S = S_1(E_1') + S_2(E_2') &\leq S_1(E_1) + \frac{E_1' - E_1}{T_1} + S_2(E_2) + \frac{E_2' - E_2}{T_2} \\ &= S_1(E_1) + S_2(E_2) + \left(E_1' - E_1\right)\left(\frac{1}{T_1} - \frac{1}{T_2}\right), \end{aligned} \tag{2.8}$$

where we use the concavity of S and the fact that $E_1 + E_2 = E_1' + E_2'$. The second law imposes that

$$S_1(E_1') + S_2(E_2') \geq S_1(E_1) + S_2(E_2) \tag{2.9}$$

and therefore

$$\left(E_1' - E_1\right)\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \geq 0. \tag{2.10}$$

Thus, if $E_1' - E_1 > 0$, we have $1/T_1 > 1/T_2$, which corresponds to $T_1 < T_2$: the body with higher T releases energy to that with lower T. This means that T behaves as a temperature scale: hotter bodies are characterized by larger values of T, and heat flows naturally from them to colder bodies. This fact is summarized by the **Clausius statement** of the second law of thermodynamics:

> Heat can never pass from a colder to a warmer body without some other change, connected therewith, occurring at the same time.

In fact, T defined in this way coincides with the absolute temperature scale and is called the **temperature** from now on.

As an illustration of these ideas, we consider a cylinder containing n moles of an ideal gas. An **ideal gas** is a fluid in which the pressure P, the volume V, the number of particles N, and the temperature T satisfy the **equation of state**

$$P = \frac{N k_{\mathrm{B}} T}{V}. \tag{2.11}$$

We change the volume of the gas by an infinitesimal quantity $\mathrm{d}V$ by performing on it an infinitesimal amount of work

$$\mathrm{d}W = -P\,\mathrm{d}V. \tag{2.12}$$

If the gas is thermally isolated, we have $\mathrm{d}E = -P\,\mathrm{d}V$. It turns out that the internal energy of an ideal gas depends only on its temperature T. Thus we obtain

$$\left.\frac{\partial E}{\partial V}\right)_T = \left.\frac{\partial E}{\partial V}\right)_S + \left.\frac{\partial E}{\partial S}\right)_V \left.\frac{\partial S}{\partial V}\right)_T = 0, \tag{2.13}$$

where $\partial X/\partial Y)_Z$ is the partial derivative of X with respect to Y, taken at constant Z. By combining eqs. (2.12) and (2.6), we obtain

$$\left.\frac{\partial E}{\partial V}\right)_S = -P; \qquad \left.\frac{\partial E}{\partial S}\right)_V = T. \tag{2.14}$$

Therefore,

$$\left.\frac{\partial S}{\partial V}\right)_T = \frac{P}{T}. \tag{2.15}$$

If the volume V of the system changes from V_0 to V_f at constant T, its entropy S changes by

$$\Delta S = \int_{V_0}^{V_\mathrm{f}} \mathrm{d}V\, \frac{P}{T} = N k_\mathrm{B} \int_{V_0}^{V_\mathrm{f}} \frac{\mathrm{d}V}{V} = N\, k_\mathrm{B} \ln \frac{V_\mathrm{f}}{V_0}. \tag{2.16}$$

Therefore, if $V_\mathrm{f} > V_0$, the transformation $V_0 \longrightarrow V_\mathrm{f}$, but not its reverse, can take place in a thermally isolated system.

Another facet of these postulates is the **Kelvin-Planck statement** of the second law:

> It is impossible to devise a cyclically operating heat engine, the only effect of which is to absorb energy in the form of heat from a single thermal reservoir and to deliver an equivalent amount of work.

Indeed, such a device would produce negative entropy during a cycle. To show that, we imagine enclosing the device and the reservoir with a wall, so that the entire system is thermally isolated. During the cycle, the device transfers a positive amount W of work to the environment and reduces the internal energy of the reservoir by the same amount. Since the entropy of the device does not change and the entire system is isolated, the total entropy change is $S^\mathrm{res} = -W/T < 0$, in contradiction with the second law.

2.2 Thermodynamic efficiency

An important application of thermodynamics is the study of engines and their efficiency. We generally call an **engine** a physical machine that operates cyclically and converts one form of energy into another. Historically, the most important example is **heat engines**, i.e., machines that cyclically convert heat into work. The Kelvin statement of the second law of thermodynamics implies that heat engines must necessarily operate using at least two heat reservoirs at different temperatures T_hot and T_cold, with $T_\mathrm{hot} > T_\mathrm{cold}$.

We consider a heat engine that is alternately put in contact with two heat reservoirs. During each cycle, the system receives an amount of energy E_hot from the hot reservoir and releases an amount of energy E_cold to the cold one. At the end of the cycle, the engine returns to the same state it had at the beginning of the cycle. Therefore, the work $W = E_\mathrm{cold} - E_\mathrm{hot}$ performed on the engine in a cycle is equal to the net total heat released to the reservoirs (remember our sign convention!). The total change in entropy in a cycle, S^tot, is given by the change ΔS^sys of the entropy of the system plus the entropy change S^res of the reservoirs:

$$S^\mathrm{tot} = \Delta S^\mathrm{sys} + S^\mathrm{res}. \tag{2.17}$$

Here and in the following we denote with ΔX the change of a state function X. The entropy change of the reservoir S^res is not a state function, since the internal energy and thus the energy of a reservoir can change without altering its temperature. Therefore, neither is S^tot a state function. In a cycle, ΔS^sys vanishes, whereas the entropy of the reservoirs changes by

$$S^\mathrm{res} = -\frac{E_\mathrm{hot}}{T_\mathrm{hot}} + \frac{E_\mathrm{cold}}{T_\mathrm{cold}} \geq 0. \tag{2.18}$$

The extracted work attains its maximum when the entropy increase S_{cold} of the reservoir at lower temperature is the opposite of the entropy decrease S_{hot} of the reservoir at higher temperature:

$$S_{\text{cold}} = \frac{E_{\text{cold}}}{T_{\text{cold}}} \geq -S_{\text{hot}} = \frac{E_{\text{hot}}}{T_{\text{hot}}}. \tag{2.19}$$

This condition implies

$$-W \leq E_{\text{hot}} \left(1 - \frac{T_{\text{cold}}}{T_{\text{hot}}}\right). \tag{2.20}$$

Traditionally, the **thermal efficiency** η^{th} of a heat engine is defined as the ratio between the extracted work and the energy absorbed from the hot reservoir:

$$\eta^{\text{th}} = -\frac{W}{E_{\text{hot}}}. \tag{2.21}$$

Equation (2.20) implies that the maximal thermal efficiency is determined by the temperatures of the hot and cold reservoirs, independent of the amount of energy exchanged during a cycle:

$$\eta^{\text{th}} \leq \eta_{\text{C}}^{\text{th}} = 1 - \frac{T_{\text{cold}}}{T_{\text{hot}}}, \tag{2.22}$$

where $\eta_{\text{C}}^{\text{th}}$ is the **thermal Carnot efficiency**.

It is interesting to consider more general engines whose energy currencies are not limited to work and heat. This is especially true in stochastic thermodynamics, where information can also be exchanged for work. To deal with these more general engines, we define the **efficiency** η_S in terms of *entropy* rather than energy. To this aim, we consider an engine operating between two arbitrary reservoirs. During a cycle, the engine extracts an amount S^{in} of entropy from a reservoir and releases an amount S^{out} of entropy to another reservoir. In this framework, we define the efficiency as

$$\eta_S = -\frac{S^{\text{in}}}{S^{\text{out}}}. \tag{2.23}$$

With this definition of efficiency, eq. (2.22) becomes

$$\eta_S \leq \eta_{\text{C}}, \tag{2.24}$$

where in this case the **Carnot efficiency** is simply

$$\eta_{\text{C}} = 1. \tag{2.25}$$

This definition of efficiency allows us to characterize engines operating between two arbitrary reservoirs.

2.3 Free energy and nonequilibrium free energy

The thermodynamic behavior of a system is identified once we know the expression of the internal energy E as a function of entropy S, particle number N, volume V, and other thermodynamically relevant observables. If the system is put in contact with a heat reservoir at temperature T, its internal energy E and entropy S are determined by the interaction with the reservoir. In this case, it is convenient to express the thermodynamic properties in terms of T rather than S. Since $T = \partial E/\partial S$ (where the other thermodynamic observables are kept constant), this change of variable can be achieved via a Legendre transformation. Properties of the Legendre transformation are summarized in appendix A.2. The **free energy** is defined as the opposite of the Legendre transform of the internal energy:

$$F(T, N, V) = E(S, N, V, \ldots) - T\,S, \tag{2.26}$$

where S is expressed as a function of T and of the other observables as the solution of the equation

$$\left.\frac{\partial E}{\partial S}\right)_{N,V,\ldots} = T, \tag{2.27}$$

and the dots stand for possible other thermodynamic observables. In the following, for simplicity, we limit ourselves to the basic observables S, V, and N. Given $F(T, N, V)$, the entropy S is given by

$$S(T, N, V) = -\left.\frac{\partial F}{\partial T}\right)_{N,V}, \tag{2.28}$$

whereas the pressure is given by

$$P = -\left.\frac{\partial F}{\partial V}\right)_{T,N}. \tag{2.29}$$

Therefore, the free energy fully describes the thermodynamic state, in the sense that it permits us to reconstruct all thermodynamic observables. Functions with this property are called **thermodynamic potentials**. Applying the symmetry of partial derivatives of F to relations like eqs. (2.28) and (2.29), we obtain other useful relations between thermodynamic quantities, such as

$$\left.\frac{\partial^2 F}{\partial V\,\partial T}\right)_N = -\left.\frac{\partial S}{\partial V}\right)_T = \left.\frac{\partial^2 F}{\partial T\,\partial V}\right)_N = -\left.\frac{\partial P}{\partial T}\right)_V. \tag{2.30}$$

Such equalities are known as **Maxwell relations**. In particular, one can use these relations to show that the equation of state of an ideal gas implies that its entropy depends on T and V in the form $S(T, V) = Nk_{\mathrm{B}} \ln V + S_0(T)$, where $S_0(T)$ does not depend on V.

We now consider a system initially at equilibrium with values S, N, and V of the thermodynamic observables that is brought to a new equilibrium by putting it in touch with a heat reservoir at temperature T. In the new equilibrium state, the free energy has the value $F(T, N, V)$ and the value of the entropy satisfies eq. (2.28). During equilibration, energy is exchanged as heat with the reservoir and as work with the external environment. We wish to characterize this exchange.

By the second law, the total entropy change of the system plus the reservoir cannot be negative:

$$S^{\mathrm{tot}} = \Delta S^{\mathrm{sys}} + S^{\mathrm{res}} \geq 0. \tag{2.31}$$

We call W the work performed on the system and Q the heat released to the reservoir. The first law imposes that

$$\Delta E = W - Q = W - T\,S^{\mathrm{res}}, \tag{2.32}$$

where we use the fact that the heat reservoir is always at equilibrium at temperature T. As a consequence of eqs. (2.31) and (2.32), we obtain

$$W \geq \Delta E - T\,\Delta S^{\mathrm{sys}}. \tag{2.33}$$

Therefore, the equilibrium state can be identified as the one in which the expression $E - T\,S$ attains its minimum. Evaluating the equilibrium value S of the system entropy at this minimum, we retrieve the condition given by eq. (2.27). The value of $E - T\,S$ at its minimum is equal to the free energy F at temperature T. Mathematically, this result descends from the fact that Legendre transforms satisfy a variational principle (cf. eq. (A.29)). These results justify defining the **nonequilibrium free energy** by

$$F^{\mathrm{neq}}(T, S, N, V) = E(S, N, V) - T\,S \tag{2.34}$$

for arbitrary values S of the entropy. The equilibrium free energy F is then obtained by the variational principle

$$F(T, N, V) = \min_{S} F^{\mathrm{neq}}(T, S, N, V). \tag{2.35}$$

The nonequilibrium free energy is not a thermodynamic potential. One reason is that it simultaneously depends on some quantities characterizing the system (such as S) and others that characterize the reservoir (such as T). We defined F^{neq} for a system initially at equilibrium, for which S is thermodynamically defined and which is brought out of equilibrium by allowing its contact with a heat reservoir. We discuss in section 5.2 a generalization of nonequilibrium free energy to mesoscopic systems prepared in an arbitrary nonequilibrium state.

There exist thermodynamic potentials other than the free energy. They differ in the thermodynamic quantities that are kept fixed by external reservoirs. For example, the appropriate thermodynamic potential for a system in contact with a heat reservoir and kept at a fixed pressure P is the **Gibbs free energy**

$$G(T, P, N) = E + P\,V - TS, \tag{2.36}$$

where E, V, and S are expressed as functions of T, P, N. Beyond heat reservoirs, we can also consider particle reservoirs, which are able to exchange with the system an unlimited amount of particles of a given chemical species without changing their properties and while remaining at thermodynamic equilibrium. They are characterized by the values T of their absolute temperature and μ of their **chemical potential**. The chemical potential μ of a system with free energy F that can exchange a single chemical species is given by

$$\mu = \left.\frac{\partial F}{\partial N}\right)_{T,V}, \tag{2.37}$$

where N is the number of molecules of the considered species. For a system containing a single chemical species at fixed temperature T and pressure P, the chemical potential is equal to the Gibbs free energy per particle:

$$\mu(T,P) = \frac{G(T,P,N)}{N}. \tag{2.38}$$

For systems exchanging multiple chemical species, a distinct chemical potential can be assigned to each one of them. Each chemical potential is defined by a formula similar to eq. (2.37) in which we keep constant the number of molecules of all exchangeable species but one, and take the derivative with respect to that one. By the same reasoning we followed for the free energy, it can be shown that equilibrium in the presence of particle reservoirs corresponds to the minimum of

$$\Phi(E,T,V,\mu_1,N_1,\mu_2,N_2,\ldots) = E - TS - \sum_i \mu_i N_i, \tag{2.39}$$

where the sum runs over all exchanged chemical species.

2.4 Statistical mechanics

Statistical mechanics links the microscopic description of a macroscopic system at equilibrium to its thermodynamic behavior. We consider a macroscopic system and assume for simplicity that its microscopic states ξ (also called **microstates**) are discrete, $\xi \in \{1,2,\ldots\}$. Macroscopically, the thermodynamic equilibrium state is identified by the values of macroscopic observables, like the internal energy E, the volume V, the number of particles N, etc. At the microscopic level, the system incessantly and rapidly changes its microstate according to its dynamics. Therefore, we cannot do better than assign it a **statistical state**, i.e., a probability distribution over the microstates. We denote by p_ξ the probability of a discrete microstate ξ. If the variables ξ are continuous, the probability density is denoted by $p(\xi)$. In either case, we denote by $\langle f(\xi)\rangle$ the expectation of the function $f(\xi)$ over the probability distribution of ξ. We also use the notations p_ξ^{eq} and $\langle f(\xi)\rangle^{\mathrm{eq}}$ whenever we want to stress that a probability distribution corresponds to thermodynamic equilibrium. We briefly review the properties of probability distributions in appendix A.3.

The **fundamental postulate of statistical mechanics** stipulates that an isolated system at thermodynamic equilibrium can be found with equal probability in any of the microstates ξ compatible with given values of the thermodynamic observables, and that the thermodynamic entropy S is related to the number $\mathcal{W}$ of the microstates that satisfy this condition by the relation

$$S = k_{\mathrm{B}} \ln \mathcal{W}. \tag{2.40}$$

Here k_{B} is the Boltzmann constant:

$$k_{\mathrm{B}} \approx 1.384 \cdot 10^{-23}\ \mathrm{J/K}. \tag{2.41}$$

This probability distribution over the microstates is known as the **microcanonical distribution** (or **microcanonical ensemble**). The term *ensemble* is used to stress that we are effectively replacing a single system and its detailed dynamical behavior with a large collection of statistically identical systems, such that their distribution over the microstates is constant in time—in agreement with thermodynamic equilibrium at the macroscopic level. This change of description is appropriate if the system is large enough. This condition is formalized by the **thermodynamic limit**, in which one imagines the size of the system (measured by the number of particles N) to go to infinity, keeping constant the ratios V/N, $E/N, \dots,$ of extensive variables.

Starting from the microcanonical distribution, one can show that the equilibrium state of a system in equilibrium with a reservoir at temperature T is described by the **Maxwell-Boltzmann** (or **canonical**) **distribution** (or **canonical ensemble**)

$$p_\xi^{\mathrm{eq}} = \frac{\mathrm{e}^{-\epsilon_\xi/k_{\mathrm{B}}T}}{Z}, \tag{2.42}$$

where ϵ_ξ is the energy of microstate ξ and the denominator is the **partition function**

$$Z = \sum_\xi \mathrm{e}^{-\epsilon_\xi/k_{\mathrm{B}}T}. \tag{2.43}$$

The partition function is related to the free energy F by

$$F = -k_{\mathrm{B}}T \ln Z. \tag{2.44}$$

This relation allows us to write the equilibrium distribution p^{eq} in the form

$$p_\xi^{\mathrm{eq}} = \mathrm{e}^{(F-\epsilon_\xi)/k_{\mathrm{B}}T}. \tag{2.45}$$

One of the simplest thermodynamic systems is the **ideal gas**. An ideal gas is made of N point-like particles of mass m that interact weakly, so that their potential energy is negligible compared to their kinetic energy. The state of a particle i is identified by its momentum $\vec{p}_{r,i}$ and its position $\vec{r}_i$. In evaluating the partition function, we have to take into account that the particles are not distinguishable. Thus the partition function must be multiplied by a factor $1/N!$, because configurations that differ only by the exchange of particles should not be considered different. In order to make the expression of Z dimensionless, we introduce an elementary phase-space volume h for each degree of freedom, where h is Planck's constant. This value is chosen to make a connection with the behavior of quantum systems. Using Stirling's approximation for the factorial, we obtain

$$\begin{aligned} F = -k_{\mathrm{B}}T \ln Z = -k_{\mathrm{B}}T \ln \frac{1}{N!} \int \prod_{i=1}^{N} \left[\frac{\mathrm{d}\vec{r}_i \, \mathrm{d}\vec{p}_{r,i}}{h^3} \exp\left(-\frac{p_{r,i}^2}{2mk_{\mathrm{B}}T} \right) \right] \\ = Nk_{\mathrm{B}}T \ln \left[\frac{N}{\mathrm{e}\, V} \left(\frac{h^2}{2\pi m k_{\mathrm{B}}T} \right)^{3/2} \right], \end{aligned} \tag{2.46}$$

where $\mathrm{e} = 2.7818\dots$ is the basis of natural logarithms. The chemical potential μ is obtained by taking the derivative of F with respect to N, as shown in eq. (2.37):

$$\mu = \left.\frac{\partial F}{\partial N}\right)_{T,V} = k_{\mathrm{B}}T \ln \frac{N}{V} + \mu^{(0)}, \tag{2.47}$$

where $\mu^{(0)}$ denotes terms that are independent of the concentration.

Values of thermodynamic observables can be obtained by taking appropriate derivatives of the partition function. For example, the equation of state is obtained by taking the derivative of F, as expressed by eq. (2.46), with respect to V:

$$P = -\left.\frac{\partial F}{\partial V}\right)_{T,N} = k_{\mathrm{B}}T \left.\frac{\partial \ln Z}{\partial V}\right)_{T,N} = \frac{N\,k_{\mathrm{B}}T}{V}. \tag{2.48}$$

Similarly, defining $Z(\beta) = \sum_{\xi} \mathrm{e}^{-\beta\epsilon_\xi}$, we have

$$\left.\frac{\partial \ln Z}{\partial \beta}\right|_{\beta=1/k_{\mathrm{B}}T} = -\frac{1}{Z}\sum_{\xi} \epsilon_\xi\, \mathrm{e}^{-\epsilon_\xi/k_{\mathrm{B}}T} = -\left\langle \epsilon \right\rangle^{\mathrm{eq}}. \tag{2.49}$$

We identify the average $\langle \epsilon \rangle^{\mathrm{eq}}$ with the thermodynamic value E of the internal energy. On the one hand, taking a further derivative, we obtain

$$\left.\frac{\partial^2 \ln Z(\beta)}{\partial \beta^2}\right|_{\beta=1/k_{\mathrm{B}}T} = \left\langle \epsilon^2 \right\rangle^{\mathrm{eq}} - \left(\langle \epsilon \rangle^{\mathrm{eq}}\right)^2 \geq 0. \tag{2.50}$$

On the other hand, an explicit evaluation of the derivatives yields

$$\left.\frac{\partial^2 \ln Z(\beta)}{\partial \beta^2}\right|_{\beta=1/k_{\mathrm{B}}T} = k_{\mathrm{B}}T^2 \left.\frac{\partial \left\langle \epsilon \right\rangle^{\mathrm{eq}}}{\partial T}\right)_{N,V,\ldots}. \tag{2.51}$$

Comparing eqs. (2.50) and (2.51), we obtain

$$k_{\mathrm{B}}T^2 \left.\frac{\partial \left\langle \epsilon \right\rangle^{\mathrm{eq}}}{\partial T}\right)_{N,V,\ldots} = \left\langle \epsilon^2 \right\rangle^{\mathrm{eq}} - \left(\langle \epsilon \rangle^{\mathrm{eq}}\right)^2. \tag{2.52}$$

This relation expresses a thermodynamic derivative (on the left-hand side) in terms of a microscopic fluctuation (on the right-hand side). It is an elementary example of relations that are collectively known as **fluctuation-dissipation relations**. Importantly, eq. (2.52) tells us that the variance of the distribution of ϵ grows proportionally to its average, and therefore to the system size. Therefore, the relative uncertainty on the energy

$$\frac{\sqrt{\left\langle \epsilon^2 \right\rangle^{\mathrm{eq}} - \left(\langle \epsilon \rangle^{\mathrm{eq}}\right)^2}}{\langle \epsilon \rangle^{\mathrm{eq}}} = \sqrt{\frac{k_{\mathrm{B}}T^2}{\left(\langle \epsilon \rangle^{\mathrm{eq}}\right)^2} \left.\frac{\partial \left\langle \epsilon \right\rangle^{\mathrm{eq}}}{\partial T}\right)_{N,V,\ldots}} \tag{2.53}$$

scales like the inverse square root of the system size. As a consequence, also taking into account the smallness of k_{B}, energy fluctuations are negligible for macroscopic systems. Thus, although statistical mechanics describes systems with probability distributions at the microscopic level, its predictions are deterministic at the macroscopic level. This is also true for macroscopic systems out of equilibrium.

If a system is in contact with reservoirs exchanging extensive quantities other than energy, averages and fluctuations of these other quantities can be evaluated following a similar strategy. We consider, for example, a system exchanging energy and particles with a reservoir characterized by a temperature T and a chemical potential μ. The equilibrium distribution of such a system is given by the **grand canonical ensemble**

$$p_\xi^{\mathrm{eq}} = \frac{1}{Z^{\mathrm{gc}}}\, \mathrm{e}^{-(\epsilon_\xi - \mu N_\xi)/k_{\mathrm{B}}T}, \tag{2.54}$$

where we define the grand canonical partition function as a function of α and β

$$Z^{\mathrm{gc}} = \sum_\xi \mathrm{e}^{\alpha N_\xi - \beta \epsilon_\xi}, \tag{2.55}$$

with $\beta = 1/k_{\mathrm{B}}T$ and $\alpha = \mu/k_{\mathrm{B}}T$. The derivatives of the logarithm of the partition function return averages and variances of thermodynamic observables in this case too:

$$\left.\frac{\partial \ln Z^{\mathrm{gc}}}{\partial \alpha}\right|_{\alpha=\mu/k_{\mathrm{B}}T,\beta=1/k_{\mathrm{B}}T} = \langle N\rangle^{\mathrm{eq}}; \tag{2.56a}$$

$$\left.\frac{\partial \ln Z^{\mathrm{gc}}}{\partial \beta}\right|_{\alpha=\mu/k_{\mathrm{B}}T,\beta=1/k_{\mathrm{B}}T} = -\,\langle \epsilon\rangle^{\mathrm{eq}}; \tag{2.56b}$$

and

$$\left.\frac{\partial^2 \ln Z^{\mathrm{gc}}}{\partial \alpha^2}\right|_{\alpha=\mu/k_{\mathrm{B}}T,\beta=1/k_{\mathrm{B}}T} = \left\langle N^2\right\rangle^{\mathrm{eq}} - \left(\langle N\rangle^{\mathrm{eq}}\right)^2; \tag{2.57a}$$

$$-\left.\frac{\partial^2 \ln Z^{\mathrm{gc}}}{\partial \alpha\, \partial \beta}\right|_{\alpha=\mu/k_{\mathrm{B}}T,\beta=1/k_{\mathrm{B}}T} = \langle N\,\epsilon\rangle^{\mathrm{eq}} - \langle N\rangle^{\mathrm{eq}}\,\langle \epsilon\rangle^{\mathrm{eq}}; \tag{2.57b}$$

$$\left.\frac{\partial^2 \ln Z^{\mathrm{gc}}}{\partial \beta^2}\right|_{\alpha=\mu/k_{\mathrm{B}}T,\beta=1/k_{\mathrm{B}}T} = \left\langle \epsilon^2\right\rangle^{\mathrm{eq}} - \left(\langle \epsilon\rangle^{\mathrm{eq}}\right)^2. \tag{2.57c}$$

These relations allow us to estimate relative fluctuations of thermodynamic observables. The symmetry of thermodynamic derivatives yielding the Maxwell relation (2.30) corresponds to the symmetry of the covariance of fluctuations, as in the example of eq. (2.57).

In the canonical ensemble, the internal energy E is a fluctuating quantity. Its distribution can be evaluated by means of the so-called **Boltzmann-Einstein principle**. To introduce it, we first associate each value of the internal energy with the entropy of the corresponding microcanonical ensemble:

$$S(E) = k_{\mathrm{B}} \ln \mathcal{W}(E) = k_{\mathrm{B}} \ln \sum_\xi \delta(\epsilon_\xi - E), \tag{2.58}$$

where $\delta(x)$ is the Dirac delta function. We substitute this result in the expression for the probability of E in the canonical ensemble:

$$\begin{aligned} p^{\mathrm{eq}}(E) &= \sum_{\xi} p_{\xi}^{\mathrm{eq}}\, \delta(\epsilon_{\xi} - E) = \mathrm{e}^{(F-E)/k_{\mathrm{B}}T} \sum_{\xi} \delta(\epsilon_{\xi} - E) \\ &= \exp\left[-\frac{E - TS(E) - F}{k_{\mathrm{B}}T}\right]. \end{aligned} \tag{2.59}$$

The argument of the exponential on the right-hand side of eq. (2.59) is the difference between the nonequilibrium free energy F^{neq} with the given value of E and the equilibrium free energy F. This reasoning can be extended to multiple observables. For example, if we look for the joint probability distribution of E and an arbitrary macroscopic observable A, we obtain

$$p^{\mathrm{eq}}(E, A) = \exp\left[-\frac{E - TS(E, A) - F}{k_{\mathrm{B}}T}\right], \tag{2.60}$$

where

$$S(E, A) = k_{\mathrm{B}} \ln \sum_{\xi} \delta(\epsilon_{\xi} - E)\, \delta(a_{\xi} - A) \tag{2.61}$$

is the entropy of a constrained microcanonical ensemble in which the values of E and of A are both fixed. In this way, the evaluation of entropy can be used to estimate probabilities.

2.5 Stochastic dynamics

In stochastic thermodynamics, we study the dynamics of mesoscopic physical systems subject to random interactions with a heat reservoir. Because of this source of randomness, at a given time t a system can be found in a given discrete state x with probability $p_x(t)$. If the variable x is continuous, we denote by $p(x;t)$ the probability density of finding the system in x at time t.

The distribution $p_x(t)$ is just one way of describing the dynamics. Another way is to study trajectories of the system:

$$\boldsymbol{x} = (x(t)). \tag{2.62}$$

This notation means that $\boldsymbol{x}$ in boldface (the trajectory) identifies the whole function $x(t)$ over a given time interval. In the following, we often use this notation to distinguish the whole function $\boldsymbol{x}$ from its instantaneous value $x(t)$. Trajectories of a stochastic system are characterized by some degree of randomness. We refer to the dynamics of a stochastic system, described either in terms of time-dependent probability distributions or in terms of trajectories, as a **stochastic process**.

Many stochastic processes of physical interest possess a useful simplifying property called the **Markov property**. A Markov process is a stochastic process that has a finite memory, i.e., one in which knowledge of the recent past fully determines the statistics of the system in the present. Given an increasing sequence of time instants $(t_0, t_1, \ldots, t_n, t)$, we denote by $p_{x;t|x_n;t_n,x_{n-1};t_{n-1},\ldots,x_1;t_1,x_0;t_0}$ the conditional probability that the system is in a discrete state x at time t, given that it was in state x_n at time

t_n, in state x_{n-1} at time $t_{n-1}, \ldots$, and in state x_0 at time t_0. For any increasing sequence of time instants, a Markov process satisfies the condition

$$p_{x;t|x_n;t_n,x_{n-1};t_{n-1},\ldots,x_1;t_1,x_0;t_0} = p_{x;t|x_n;t_n}. \tag{2.63}$$

This means that, in a Markov process, given the value of x at a given time t_n, the evolution of the system at a later time is independent of events that occurred at earlier times. The same definition of a Markov process holds if x is continuous. The conditional probability density $p_{x;t|x_n;t_n}$ of a Markov process satisfies a simple relation. We pick an intermediate time t' such that $t_0 < t' < t$. By the law of total probabilities (A.35) and by the Markov property (2.63), we obtain

$$p_{x;t|x_0;t_0} = \sum_{x'} p_{x;t|x';t'}\, p_{x';t'|x_0;t_0}. \tag{2.64}$$

Equation (2.64) is called the **Chapman-Kolmogorov** equation. Therefore, the knowledge of $p_{x;t|x';t'}$ and of the distribution $p_x(t_0)$ at an initial time t_0 allows one to evaluate the distribution $p_x(t)$ at arbitrary times $t > t_0$. The conditional probability $p_{x;t|x';t'}$ is also called the **propagator** or **Green function** of the process. Conservation of probability requires

$$\sum_{x} p_{x;t|x';t'} = 1, \qquad \forall x' \text{ and } \forall t' \le t. \tag{2.65}$$

Applying the Chapman-Kolmogorov equation (2.64) to a time interval of infinitesimal duration dt, we obtain

$$p_x(t+\mathrm{d}t) = \sum_{x'} p_{x;t+\mathrm{d}t|x';t}\, p_{x'}(t), \tag{2.66}$$

where $p_{x;t+\mathrm{d}t|x';t}$ can be written in the form

$$p_{x;t+\mathrm{d}t|x';t} = \delta^{\mathrm{K}}_{xx'} + \mathrm{d}t\, L_{xx'}(t), \tag{2.67}$$

where $\delta^{\mathrm{K}}_{xx'}$ is the Kronecker delta and $L_{xx'}(t)$ is a matrix that satisfies

$$\begin{aligned} L_{xx'}(t) &\ge 0, && \text{if } x \neq x'; \\ L_{xx}(t) &= -\sum_{x'\,(\neq x)} L_{x'x}, && \forall x. \end{aligned} \tag{2.68}$$

The matrix $L_{xx'}$ is called the **generator** of the process. Thus $p_x(t)$ satisfies an evolution equation of the form

$$\frac{\mathrm{d}p_x(t)}{\mathrm{d}t} = \sum_{x'} L_{xx'}(t)\, p_{x'}(t), \tag{2.69}$$

which has the solution

$$p_x(t) = \sum_{x'} G_{xx'}(t,t')\, p_{x'}(t'), \tag{2.70}$$

where the Green function is obtained by formally integrating eq. (2.69) with the initial condition $G_{xx'}(t',t') = \delta^{\mathrm{K}}_{xx'}$:

$$G_{xx'}(t,t') = p_{x;t|x';t'} = \left(\mathbb{T}\, \mathrm{e}^{\int_{t'}^{t} \mathrm{d}t''\, L(t'')} \right)_{xx'}. \tag{2.71}$$

Here (expression)$_{xx'}$ denotes the matrix element of expression, and $\mathbb{T}$ denotes the time-ordered product:

$$\begin{aligned}\left(\mathbb{T}\, \mathrm{e}^{\int_{t'}^{t} \mathrm{d}t''\, L(t'')} \right)_{xx'} &= \delta^{\mathrm{K}}_{xx'} + \int_{t'}^{t} \mathrm{d}t_0\, L_{xx'}(t_0) \\ &\quad + \int_{t_0}^{t} \mathrm{d}t_1 \int_{t'}^{t} \mathrm{d}t_0 \sum_{x_0} L_{xx_0}(t_1) L_{x_0x'}(t_0) + \cdots .\end{aligned} \tag{2.72}$$

In the case where the generator does not depend on time, the Green function depends only on the time difference $(t - t')$.

Markov processes with continuous state space are also defined by the condition (2.63). In this case, we denote the propagator by $p(x;t|x';t')$, and the Chapman-Kolmogorov equation takes the form

$$p(x;t|x_0;t_0) = \int \mathrm{d}x'\, p(x;t|x';t')\, p(x';t'|x_0;t_0). \tag{2.73}$$

The evolution equation now reads

$$\frac{\partial}{\partial t} p = \mathcal{L} p, \tag{2.74}$$

where the generator $\mathcal{L}$ is a linear operator that in general includes derivatives with respect to x. In the next three sections, we discuss separately, and in more detail, Markov processes with discrete and continuous state space.

2.6 Master equations

Master equations describe the evolution of Markov processes in continuous time with discrete states. A master equation is defined by the **jump rates** $k_{xx'}$ from discrete state x' to x. The jump rate is proportional to the conditional probability that a jump $x' \longrightarrow x$ takes place in an infinitesimal time interval $(t, t + \mathrm{d}t)$, given that the system is in state x' at time t. Specifically, the jump rates (or simply *rates*) are related to the propagator by

$$p_{x;t+\mathrm{d}t|x';t} = k_{xx'}\, \mathrm{d}t, \qquad x \neq x'. \tag{2.75}$$

The rates can in principle depend on time. Because of the normalization condition, eq. (2.65), the probability of remaining in a given state x in an infinitesimal time interval must be equal to $1 - \mathrm{d}t \sum_{x'} k_{x'x} = 1 - \mathrm{d}t\, k^{\mathrm{out}}_x$, where we define the **escape rate** from state x:

$$k^{\mathrm{out}}_x = \sum_{x'} k_{x'x}. \tag{2.76}$$

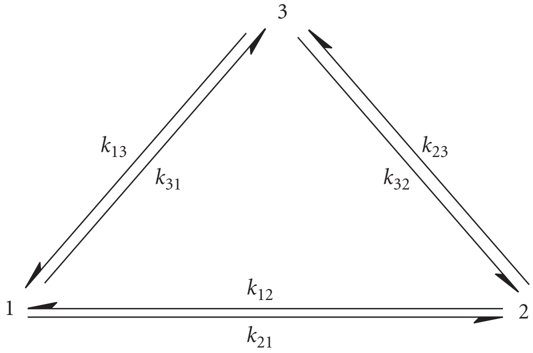

Figure 2.1. Jump network of a system with three states.

Given the rates, we construct the evolution equation for the probability $p_x(t)$ by evaluating the net probability flow reaching the state x. A state x receives an **inflow** of probability from other states at rate $\sum_{x'\,(\neq x)} k_{xx'} p_{x'}(t)$, and returns an **outflow** at rate $\sum_{x'\,(\neq x)} k_{x'x} p_x(t)$. The net probability flow is given by the inflow minus the outflow, leading to the **master equation**

$$\frac{\mathrm{d}}{\mathrm{d}t} p_x(t) = \sum_{x'\,(\neq x)} \left[k_{xx'}\, p_{x'}(t) - k_{x'x}\, p_x(t) \right]. \tag{2.77}$$

We now assume that the rates $k_{xx'}$ are independent of time, and look at the behavior of the probability distribution $p_x(t)$ for $t \to \infty$. A **stationary** distribution p_x^{st} is a probability distribution that satisfies

$$\sum_{x'\,(\neq x)} \left[k_{xx'}\, p_{x'}^{\mathrm{st}} - k_{x'x}\, p_x^{\mathrm{st}} \right] = 0, \qquad \forall x, \tag{2.78}$$

and that is therefore a solution of the master equation (2.77) constant in time.

It is useful to represent a master equation via a **jump network**, where the nodes represent the states x and the arrows $x' \longrightarrow x'$ represent possible jumps, i.e., jumps with nonvanishing rates. We draw two opposite arrows between a pair of states if jumps in both directions are possible. An example of a jump network with three states is shown in fig. 2.1, where each arrow $x' \longrightarrow x$ is associated with a nonzero rate $k_{xx'}$.

We say that a jump network is **connected** if any states x can be reached from any other state x' by means of a sequence of jumps with nonvanishing rates. In the following, we always assume that this property holds, since disconnected master equations represent multiple noninteracting physical systems that can be studied independently. Under such assumptions, and provided that the number of states is finite, the **Perron-Frobenius theorem** asserts that the leading eigenvalue, i.e., the eigenvalue with the largest real part, is purely real and nondegenerate. Moreover, its associated eigenvector can be chosen to have strictly positive entries. In the case of master equations with a finite number of states, the leading eigenvalue must be zero, otherwise the equation would not preserve normalization. The normalized eigenvector can therefore be interpreted as the stationary probability distribution. If the number of states is infinite, the stationary distribution may not exist: this is the case, for example, of a particle diffusing on an infinite line, when the rates of jumps to the left or to the right are independent of its position. If the number of states is finite, the real parts of all eigenvalues except the leading one are negative, and therefore the stationary distribution is approached exponentially fast in time:

$$\lim_{t\to\infty} p_x(t) = p_x^{\mathrm{st}}, \qquad \forall x. \tag{2.79}$$

The Perron-Frobenius theorem is not limited to generators of Markov processes. In particular, it does not require the generator to preserve normalization, i.e., that its columns sum to one. An elementary proof of these properties that does not explicitly rely on the Perron-Frobenius theorem is reported in appendix A.5.

The master equation can be seen as a continuity equation for the probability. This interpretation becomes more transparent by introducing the **probability current**

$$J_{xx'}(t) = k_{xx'} p_{x'}(t) - k_{x'x} p_x(t), \qquad x \neq x', \tag{2.80}$$

which quantifies the net probability flow from state x' to state x at time t. In terms of the probability currents, the master equation can be rewritten in the compact form

$$\frac{\mathrm{d}}{\mathrm{d}t} p_x(t) = \sum_{x'\,(\neq x)} J_{xx'}(t). \tag{2.81}$$

Equation (2.81) states that, at any given time, the rate change of the probability of being in a state x is given by the total net flow to state x from all other states x'. If the probability distribution p is the stationary one, we have

$$\sum_{x'\,(\neq x)} J_{xx'} = 0, \qquad \forall x. \tag{2.82}$$

If for any allowed jump $x \longrightarrow x'$ (i.e., such that $k_{x'x} > 0$) the reverse jump is also allowed ($k_{xx'} > 0$), the system exhibits **microscopic reversibility**. Most of the systems studied in stochastic thermodynamics possess this property. In this case, we often represent the jump network by a graph in which each possible jump between two states is represented by an undirected **edge** between the corresponding nodes. We discuss in section 8.5 how to handle microscopically irreversible systems.

If the stationary distribution satisfies, beyond eq. (2.78), the stronger condition

$$k_{xx'} p_{x'}^{\mathrm{st}} = k_{x'x} p_x^{\mathrm{st}}, \qquad \forall x \neq x', \tag{2.83}$$

then in the stationary state the probability current $J_{xx'}$ vanishes for each pair (x, x') of different states. The condition in eq. (2.83) is known as the **detailed balance condition**. If it is satisfied, the stationary distribution is called the **equilibrium distribution** and we denote it by p^{eq}. A master equation admits an equilibrium distribution if, for any sequence $(x_0, x_1, \ldots, x_n)$ of states all different from one another, we have

$$k_{x_0x_1} k_{x_1x_2} \cdots k_{x_nx_0} = k_{x_0x_n} k_{x_nx_{n-1}} \cdots k_{x_1x_0}. \tag{2.84}$$

This condition requires in particular that microscopic reversibility is satisfied. The condition of eq. (2.84) might seem obscure at first, but it becomes clearer in the context of a specific jump network as in fig. 2.1. Because of the conservation of probability, the sum of currents arriving at each node must vanish in the stationary state; see eq. (2.82). This implies that nonvanishing stationary currents can only flow in loops. This requires in particular that the jump network contains a sequence of distinct nodes, each of them connected to the previous one by an edge, and where an edge connects the last one to the first. This sequence defines a **cycle** in the network. A nonvanishing current in a cycle can only be sustained if the rates "pushing" the currents clockwise and counterclockwise do not balance. The balance conditions for the rates along each cycle of the network are indeed given by eq. (2.84). This also means that, if there are no cycles (e.g., if the network is linear or treelike), then the stationary solution always satisfies detailed balance conditions, provided that microscopic reversibility holds. When the condition (2.84) is not satisfied for some cycles, the probability currents along these cycles do not vanish in the stationary state. In this case, the stationary distribution is also called a **nonequilibrium steady state**.

In thermodynamic systems, the condition of detailed balance is often associated with thermodynamic equilibrium, and the stationary distribution p_x^{eq} appearing in eq. (2.83) is the Boltzmann distribution, eq. (2.42). We further discuss this point in section 3.1. Then, assuming that $k_{x'x} > 0$, eq. (2.83) implies that

$$\frac{k_{xx'}}{k_{x'x}} = \frac{p_x^{\mathrm{eq}}}{p_{x'}^{\mathrm{eq}}} = \mathrm{e}^{-(\epsilon_x - \epsilon_{x'})/k_{\mathrm{B}}T}. \tag{2.85}$$

Although the master equation (2.77) is linear, solving it explicitly can be difficult when the number of states is large. To tackle it numerically, it is convenient to simulate an ensemble of random trajectories rather than integrating the master equation itself. Trajectories of a master equation can be very efficiently simulated by means of the **Gillespie algorithm**, which is briefly described in appendix A.6.

2.7 Trajectories of master equations

During its evolution in a time interval $[t_0, t_{\mathrm{f}}]$, a system described by a master equation visits a sequence $x_0, x_1, \dots, x_{\mathrm{f}}$ of states (fig. 2.2). We call t_k the random time at which the system jumps from state x_{k-1} to state $x_k \neq x_{k-1}$. The system is in state x_k during a time interval $t_k \le t < t_{k+1}$. Thus the trajectory $\boldsymbol{x}$ in the given time interval is made of a sequence of **dwells**, where the system remains in the same state, separated by **jumps**, where the system changes state, as shown in fig. 2.2. The whole trajectory $\boldsymbol{x}$ is then encoded in the sequence

$$\boldsymbol{x} = ((x_0, t_0), (x_1, t_1), \dots, (x_{\mathrm{f}}, t_n), t_{\mathrm{f}}). \tag{2.86}$$

We wish to evaluate the probability density $\mathcal{P}_{\boldsymbol{x}}$ of the trajectory $\boldsymbol{x}$. To this aim, we discretize the time interval into $\mathcal{N}$ intervals of short duration $\Delta t = \mathcal{T}/\mathcal{N}$, where $\mathcal{T} = t_{\mathrm{f}} - t_0$ is the duration of the whole time interval. We then approximate the trajectory by the sequence $\boldsymbol{x}^{\mathrm{dsc}} = (x_0, x_1, \dots, x_{\mathcal{N}})$ of states at the time t_ℓ, where $t_\ell = t_0 + \ell\,\Delta t$. For each small interval $(t_\ell, t_\ell + \Delta t)$, the conditional probability that $x(t_\ell + \Delta t) = x$, given that $x(t_\ell) = x'$, is given by $p_{x;t+\Delta t|x';t}$. The probability of the discrete trajectory is then

$$\mathcal{P}_{\boldsymbol{x}^{\mathrm{dsc}}|x_0} = \prod_{\ell=1}^{\mathcal{N}} p_{x_\ell;t_{\ell-1}+\Delta t|x_{\ell-1};t_{\ell-1}}. \tag{2.87}$$

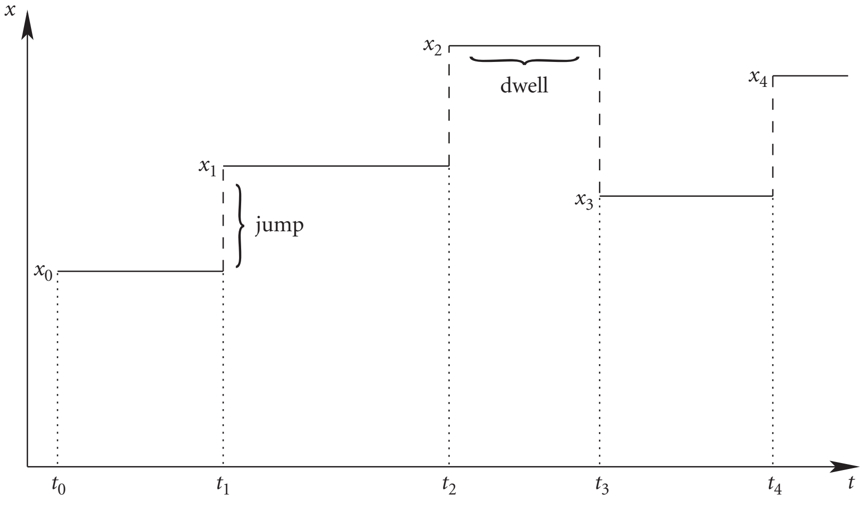

Figure 2.2. Example of a trajectory of a master equation.

From the master equation, we approximate the conditional probability by

$$p_{x;t+\Delta t|x';t} \approx \delta^{\mathrm{K}}_{xx'} + \Delta t\, L_{xx'}(t), \tag{2.88}$$

where $L_{xx'}(t)$ is the generator. This expression yields conditional probabilities that are properly normalized. Therefore, the probability of the trajectory $\boldsymbol{x}$ is approximately equal to

$$\mathcal{P}_{\boldsymbol{x}|\boldsymbol{x_0}} \approx \mathcal{P}_{\boldsymbol{x}^{\mathrm{dsc}}|x_0} \approx \prod_{\ell=1}^{\mathcal{N}} \left(\delta^{\mathrm{K}}_{x_\ell x_{\ell-1}} + \Delta t\, L_{x_\ell x_{\ell-1}}(t_{\ell-1})\right). \tag{2.89}$$

On the one hand, we can explicitly evaluate the products over each dwell:

$$\prod_{\ell \in \,\mathrm{dwell}} p_{x_\ell;t_{\ell-1}+\mathrm{d}t|x_\ell;t_{\ell-1}} = \prod_\ell \left(1 - \mathrm{d}t_\ell\, k^{\mathrm{out}}_{x_\ell}\right) \approx \mathrm{e}^{-\sum_\ell k^{\mathrm{out}}_{x_\ell}(t)\,\mathrm{d}t_\ell} \approx \mathrm{e}^{-\int \mathrm{d}t\, k^{\mathrm{out}}_{x(t)}(t)}. \tag{2.90}$$

On the other hand, the probability that the system undergoes a jump from x' to x in the short time interval $[t_\ell, t_\ell + \mathrm{d}t]$ is given by $k_{xx'}\,\mathrm{d}t\, p_{x'}(t_\ell)$ to first order in $\mathrm{d}t$. Therefore, the factors contributing to the probability of a trajectory $\boldsymbol{x}$ due to the sequence of dwells and jumps are

$$\begin{aligned} \mathcal{P}_{\boldsymbol{x}} = {} & \mathrm{e}^{-\int_{t_n}^{t_\mathrm{f}} \mathrm{d}t'\, k^{\mathrm{out}}_{x_n}(t')} k_{x_n x_{n-1}}(t_n)\, \mathrm{e}^{-\int_{t_n}^{t_{n-1}} \mathrm{d}t'\, k^{\mathrm{out}}_{x_{n-1}}(t')} \dots \\ & \times \mathrm{e}^{-\int_{t_2}^{t_1} \mathrm{d}t'\, k^{\mathrm{out}}_{x_1}(t')} k_{x_1 x_0}(t_1)\, \mathrm{e}^{-\int_{t_1}^{t_0} \mathrm{d}t'\, k^{\mathrm{out}}_{x_0}(t')} p_{x_0}(t_0). \end{aligned} \tag{2.91}$$

This expression includes the probability of the initial state $p_{x_0}(t_0)$. The choice of writing the factors on the right-hand side of eq. (2.91) in temporal order from right to left might seem awkward for a reader used to writing in the Latin alphabet from left to right. However, it becomes quite natural when thinking of the probability density as a product of matrix elements, one for each infinitesimal time step, acting on the initial probability $p_{x_0}(t_0)$.

We define the integral over trajectories in the following way:

$$\int \mathcal{D}\boldsymbol{x} \cdots = \sum_{n=0}^{\infty} \sum_{x_0, x_1, \ldots, x_n} \int_{t_0}^{t_2} \mathrm{d}t_1 \int_{t_2}^{t_3} \mathrm{d}t_2 \cdots \int_{t_{n-1}}^{t_\mathrm{f}} \mathrm{d}t_n \cdots . \tag{2.92}$$

With this definition, the probability density $\mathcal{P}_{\boldsymbol{x}}$ satisfies the normalization condition

$$\int \mathcal{D}\boldsymbol{x}\, \mathcal{P}_{\boldsymbol{x}} = 1. \tag{2.93}$$

Indeed, comparing eqs. (2.91) and (2.92) with eqs. (2.71) and (2.72), we obtain

$$\begin{aligned} \int \mathcal{D}\boldsymbol{x}\, \mathcal{P}_{\boldsymbol{x}} &= \sum_{x_\mathrm{f} x_0} \left(\mathbb{T}\, \mathrm{e}^{\int_{t_0}^{t_\mathrm{f}} \mathrm{d}t\, L(t)} \right)_{x_\mathrm{f} x_0} p_{x_0}(t_0) \\ &= \sum_{x_\mathrm{f} x_0} p_{x_\mathrm{f};t_\mathrm{f}|x_0;t_0}\, p_{x_0}(t_0) = 1. \end{aligned} \tag{2.94}$$

2.8 Fokker-Planck equation (*)

The stochastic dynamics of systems with continuous state space and continuous trajectories is described by the Fokker-Planck equation and by stochastic differential equations. These tools were introduced in physics to investigate **Brownian motion**. Einstein described Brownian motion as the phenomenon by which "bodies of microscopically visible size suspended in a liquid perform [random] movements of such magnitude that they can be easily observed in a microscope." We follow his line of reasoning to derive the Fokker-Planck equation.

We consider a particle that moves along a one-dimensional continuous coordinate x. We call $p(x;t)$ the probability density of the position x of a particle at time t. We first assume that the particle is immersed in a uniform fluid and is not subject to external applied forces. Due to random interactions with the fluid particles, during a short time interval of duration Δt, the particle experiences a random displacement Δx. If Δt is very small, the displacement Δx is largely in the direction of the initial velocity of the particle due to inertia. However, interactions with the particles of the fluid soon wipe out this dependence. Therefore, if Δt is large enough but still small, we expect this displacement to be independent in nonoverlapping time intervals. We call $\psi(\Delta x)$ the probability distribution of Δx over these time intervals. In the absence of externally applied forces, $\psi(\Delta x)$ must be an even function of Δx due to symmetry. We express $p(x;t+\Delta t)$ in terms of $p(x;t)$ and the displacement distribution

$$p(x;t+\Delta t) = \int \mathrm{d}\Delta x\, \psi(\Delta x)\, p(x-\Delta x;t). \tag{2.95}$$

Since Δt is small, Δx is also small. We therefore expand $p(x-\Delta x;t)$ in a Taylor series to second order:

$$p(x;t+\Delta t) \approx \int \mathrm{d}\Delta x\, \psi(\Delta x) \left[p(x;t) - \Delta x \frac{\partial}{\partial x} p(x;t) + \frac{1}{2} \Delta x^2 \frac{\partial^2}{\partial x^2} p(x;t) \right]. \qquad (2.96)$$

We then perform the integration over Δx. Since $\psi(\Delta x)$ is an even function, the term proportional to Δx vanishes upon integration. We thus obtain

$$p(x;t+\Delta t) - p(x;t) \approx \frac{1}{2} \left\langle \Delta x^2 \right\rangle \frac{\partial^2}{\partial x^2} p(x;t). \qquad (2.97)$$

We now divide eq. (2.97) by Δt and take the limit $\Delta t \to 0$. In taking this limit, we make the crucial assumption that

$$\lim_{\Delta t \to 0} \frac{\left\langle \Delta x^2 \right\rangle}{2\Delta t} = D, \qquad (2.98)$$

where D is a finite quantity called the **diffusion constant**. If this assumption holds, we obtain the **diffusion equation**:

$$\frac{\partial}{\partial t} p(x;t) = D \frac{\partial^2}{\partial x^2} p(x;t). \qquad (2.99)$$

Starting from a localized initial condition $p(x;t_0) = \delta(x - x_0)$, the solution of the diffusion equation is a Gaussian distribution,

$$p(x;t) = \frac{1}{\sqrt{4\pi D\mathcal{T}}} \exp\left(-\frac{(x-x_0)^2}{4D\mathcal{T}} \right), \qquad (2.100)$$

where $\mathcal{T} = t - t_0$.

In a more general case, for instance when the particle is subject to forces, the distribution of the displacements Δx might depend on x, t, and might not necessarily be even. We express it as

$$p(\Delta x; \Delta t, x, t) = p(x + \Delta x, t + \Delta t | x, t). \qquad (2.101)$$

We assume that the following limits exist and define the **drift** and **diffusion coefficients**, respectively, by

$$v(x,t) = \lim_{\Delta t \to 0} \frac{\langle \Delta x \rangle_{x,t}}{\Delta t}; \qquad (2.102a)$$

$$D(x,t) = \lim_{\Delta t \to 0} \frac{\left\langle \Delta x^2 \right\rangle_{x,t}}{2\,\Delta t}. \qquad (2.102b)$$

The averages $\langle \ldots \rangle_x$ are taken over the distribution of displacements. Then the dynamics is described by the **Fokker-Planck equation** (also called the **Kolmogorov forward equation**):

$$\frac{\partial}{\partial t} p(x;t) = -\frac{\partial}{\partial x} \left[v(x,t)\, p(x;t) \right] + \frac{\partial^2}{\partial x^2} \left[D(x,t)\, p(x;t) \right]. \qquad (2.103)$$

Equation (2.103) is derived in appendix A.7. The first term in eq. (2.103) is the drift and is associated with the local mean velocity of particles. The second term is the diffusion, which represents the "random," unbiased component of the motion.

The Fokker-Planck equation can also be written as a continuity equation,

$$\frac{\partial}{\partial t} p(x;t) = -\frac{\partial}{\partial x} J(x;t), \tag{2.104}$$

where the probability current is defined by

$$J(x,t) = v(x,t)\, p(x;t) - \frac{\partial}{\partial x}\left[D(x,t)\, p(x;t)\right]. \tag{2.105}$$

The Fokker-Planck equation is a second-order partial differential equation. To solve it, we need to specify the initial condition, i.e., the distribution $p(x;t_0)$, and the **boundary conditions**. Often the Fokker-Planck equation is studied in the infinite interval $x \in (-\infty, \infty)$, for which the boundary conditions are simply that the probability vanishes as $|x| \to \infty$.

The solution $p(x;t|x_0;t_0)$ of the Fokker-Planck equation with initial condition $\delta(x - x_0)$ at $t = t_0$ is the **propagator**, which expresses the conditional probability density of finding the particle close to x at time $t > t_0$, given that it was at x_0 at time t_0. In some cases, one is interested in the dependence of this probability density on the earlier state x_0 and the earlier time t_0. This dependence is governed by the **Kolmogorov backward equation**:

$$-\frac{\partial}{\partial t_0} p(x;t|x_0;t_0) = v(x_0,t_0)\,\frac{\partial}{\partial x_0} p + D(x_0,t_0)\,\frac{\partial^2}{\partial x_0^2} p. \tag{2.106}$$

We derive this equation in appendix A.7.

We now consider cases where the coefficients $v(x,t)$ and $D(x,t)$ of the Fokker-Planck equation do not depend on time. The stationarity condition reads

$$\frac{\partial}{\partial x} J^{\mathrm{st}}(x) = 0, \tag{2.107}$$

where the stationary current is

$$J^{\mathrm{st}}(x) = v(x)\, p^{\mathrm{st}}(x) - \frac{\partial}{\partial x}\left[D(x)\, p^{\mathrm{st}}(x)\right]. \tag{2.108}$$

When considering an infinite interval, the stationary solution does not necessarily exist. A prominent example is the diffusion equation (2.99), whose solution (2.100) does not tend to a limiting form as $\mathcal{T} \to \infty$.

In some cases, the Fokker-Planck equation also admits a stationary solution with a vanishing stationary current:

$$J^{\mathrm{st}}(x) = 0, \qquad \forall x. \tag{2.109}$$

Equation (2.109) is the detailed balance condition for Fokker-Planck equations. In general, detailed balance is a rather restrictive condition. However, there are physically relevant scenarios where this property holds, such as many one-dimensional systems and most systems at thermodynamic equilibrium, as discussed at the end of section 2.6. In the one-dimensional case, when detailed balance holds, the stationary solution is given by

$$p^{\mathrm{st}}(x) \propto \frac{1}{D(x)} \exp\left(\int_{x_0}^{x} \mathrm{d}x' \, \frac{v(x')}{D(x')}\right), \tag{2.110}$$

where x_0 is arbitrary and the proportionality constant is determined by the normalization condition $\int_{-\infty}^{\infty} \mathrm{d}x \, p^{\mathrm{st}}(x) = 1$.

2.9 Langevin equation (*)

Trajectories of a Brownian particle with drift coefficient $v(x, t)$ and diffusion coefficient $D(x, t)$ are solutions of the **Langevin equation**

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(x, t) + \sigma(x, t) \, \xi(t), \tag{2.111}$$

where $\sigma(x, t)$ is a function related to $D(x, t)$ and $\xi(t)$ is a random quantity. We impose that this random quantity is unbiased, $\langle \xi(t) \rangle = 0$, $\forall t$. We also assume that displacements in nonoverlapping time intervals are uncorrelated. This amounts to requiring $\langle \xi(t) \, \xi(t') \rangle = 0$ for t and t' sufficiently far apart. Assuming that this holds whenever $t \neq t'$, we obtain $\langle \xi(t) \xi(t') \rangle \propto \delta(t - t')$. We set the proportionality constant equal to 1 by suitably defining $\sigma(x, t)$. A variable $\xi(t)$ satisfying these properties is called **white noise**. In the mathematical literature, Langevin equations are an example of **stochastic differential equations**.

Equation (2.111) is ill defined as it stands, since the variance of the random quantity $\xi(t)$ is a delta function in time. The mathematical tools to deal with stochastic differential equations were developed well after the work of Langevin and are the subject of the theory of **stochastic calculus**. Stochastic calculus is nowadays a rather developed field of mathematics. In this section, we concentrate on the most important and useful results without going too deep into the mathematical details.

We formally evaluate the time integral of $\xi(t)$:

$$W(t) = \int_{t_0}^{t} \mathrm{d}t' \, \xi(t') = \int_{t_0}^{t} \mathrm{d}W. \tag{2.112}$$

The random process $W(t)$ is called the **Wiener process**. It has vanishing average

$$\langle W(t) \rangle = \left\langle \int_{t_0}^{t} \mathrm{d}t' \, \xi(t') \right\rangle = \int_{t_0}^{t} \mathrm{d}t' \, \langle \xi(t') \rangle = 0. \tag{2.113}$$

Since $\xi(t)$ is unbiased on average, so is its time integral $W(t)$. The second moment of $W(t)$ is given by

$$\begin{aligned} \langle W^2(t) \rangle &= \left\langle \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t} \mathrm{d}t'' \, \xi(t') \xi(t'') \right\rangle = \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t} \mathrm{d}t'' \, \langle \xi(t') \xi(t'') \rangle \\ &= \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t} \mathrm{d}t'' \, \delta(t' - t'') = \mathcal{T}, \end{aligned} \tag{2.114}$$

where $\mathcal{T} = t - t_0$. In general, it can be shown that the distribution of $W(t)$ is Gaussian:

$$p(W(t) = w) = \frac{1}{\sqrt{2\pi \mathcal{T}}} \, \mathrm{e}^{-w^2/2\mathcal{T}}. \tag{2.115}$$

Mathematicians prefer to formally define $W(t)$ as a process whose increments over a time interval of duration $\mathcal{T}$ are Gaussian random variables with zero mean and variance $\mathcal{T}$. They allow physicists to use $\xi(t)$ as an ill-defined "derivative" of $W(t)$.

We now consider a Langevin equation without drift:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma \, \xi(t), \tag{2.116}$$

where we assume that the coefficient σ is constant. Integrating from $t = t_0$ to a generic time t, we obtain

$$x(t) = x_0 + \sigma \, W(t). \tag{2.117}$$

Thus the distribution of $x(t)$ is a Gaussian centered in x_0 and with variance $\sigma^2 \left\langle W^2(t) \right\rangle = \sigma^2 \mathcal{T}$. By comparing with the diffusion equation, eq. (2.99) and its solution, we conclude that this Langevin equation describes the trajectories of a diffusion equation with a diffusion coefficient

$$D = \sigma^2/2. \tag{2.118}$$

By going through the derivation of the Fokker-Planck equation in section 2.8, one sees that the solution of the Langevin equation (2.111) with constant σ describes a process satisfying the Fokker-Planck equation (2.103) with the same drift $v(x, t)$ and with diffusion coefficient $D = \sigma^2/2$.

Processes where σ depends on x are often called Langevin equations with **multiplicative noise** in the physics literature. This term is somewhat misleading as it appears to be limited to the case in which $\sigma \propto x$. In the presence of multiplicative noise, the interpretation of eq. (2.111) is mathematically subtle. Depending on its interpretation, it may or may not represent the behavior of the solutions of eq. (2.103). This is due to the fact that, for short time intervals $\mathrm{d}t$, the typical increments of W are much larger than $\mathrm{d}t$, and thus the second term of the Langevin equation dominates over the first one.

Formally, solving a Langevin equation requires the evaluation of **stochastic integrals** of the form

$$\int_{t_0}^{t_\mathrm{f}} \mathrm{d}t \, \xi(t) f(x(t), t) = \int_{t_0}^{t_\mathrm{f}} \mathrm{d}W(t) f(x(t), t), \tag{2.119}$$

where $f(x, t)$ is a given function. The ambiguity in the interpretation of the Langevin equation stems from the ill-defined nature of stochastic integrals. There are two major conventions to resolve this ambiguity and therefore to assign a precise interpretation to the corresponding Langevin equation:

Ito convention. The Ito convention is defined by the prescription

$$I_\mathrm{I} = \lim_{\mathrm{d}t \to 0} \sum_{k=0}^{\mathcal{N}} [W(t_k + \mathrm{d}t) - W(t_k)] f(x(t_k), t_k), \tag{2.120}$$

where $t_i = t_0 + i\,\mathrm{d}t$ and the sum runs over the $\mathcal{N} = (t_\mathrm{f} - t_0)/\mathrm{d}t$ intervals of duration $\mathrm{d}t$ in which we divide the interval $[t_0, t_\mathrm{f}]$. In the Ito convention, the function $f(x(t), t)$ is evaluated at the beginning of each infinitesimal time interval $[t_k, t_k + \mathrm{d}t]$. We denote the Ito convention by a dot product symbol:

$$I_\mathrm{I} = \int_{t_0}^{t_\mathrm{f}} \mathrm{d}W \cdot f(x(t'), t'). \tag{2.121}$$

Stratonovich convention. The Stratonovich convention is defined by

$$\begin{aligned} I_\mathrm{S} &= \lim_{\mathrm{d}t\to 0} \sum_{k=0}^{\mathcal{N}} [W(t_k + \mathrm{d}t) - W(t_k)] f\left(\frac{x(t_k + \mathrm{d}t) + x(t_k)}{2}, t_k + \frac{\mathrm{d}t}{2}\right) \\ &= \lim_{\mathrm{d}t\to 0} \sum_{k=0}^{N} [W(t_k + \mathrm{d}t) - W(t_k)] \frac{1}{2}\left[f(x(t_k + \mathrm{d}t), t_k + \mathrm{d}t) + f(x(t_k), t_k)\right]. \end{aligned} \tag{2.122}$$

In the Stratonovich convention, $f(x, t)$ is evaluated at the midpoint of each infinitesimal interval $[t_k, t_k + \mathrm{d}t]$. We can equivalently evaluate the function $f(x, t)$ as the average of its values at the boundaries of the interval: the difference between these two prescriptions vanishes as $\mathrm{d}t \to 0$. We denote the Stratonovich convention by a circle product symbol:

$$I_\mathrm{S} = \int_{t_0}^{t_\mathrm{f}} \mathrm{d}W \circ f(x(t'), t'). \tag{2.123}$$

In conventional (nonstochastic) integrals, the choices (2.120) and (2.122) yield the same result as $\mathrm{d}t \to 0$. This is not necessarily the case for stochastic integrals. Therefore, a Langevin equation is not fully defined unless one declares whether stochastic integrals are interpreted according to the discretization of eq. (2.120) or (2.122).

The Langevin equation (2.111) is equivalent to the Fokker-Planck equation (2.103) under the Ito convention. As shown in appendix A.8, the Fokker-Planck equation corresponding to the Stratonovich convention looks slightly different:

$$\frac{\partial}{\partial t} p(x;t) = -\frac{\partial}{\partial x}\big(v(x,t)\,p(x;t)\big) + \frac{1}{2}\frac{\partial}{\partial x}\left[\sigma(x,t)\,\frac{\partial}{\partial x}\big(\sigma(x,t)\,p(x;t)\big)\right]. \tag{2.124}$$

In this interpretation, the current is defined by

$$J(x,t) = w(x,t)\,p(x;t) - D(x,t)\frac{\partial}{\partial x} p(x;t), \tag{2.125}$$

where the diffusion coefficient is given by $D(x,t) = \sigma^2(x,t)/2$ as in the Ito convention, and

$$w(x,t) = v(x,t) - \frac{1}{2}\sigma(x,t)\frac{\partial}{\partial x}\sigma(x,t) = v(x,t) - \frac{1}{2}\frac{\partial}{\partial x}D(x,t). \tag{2.126}$$

When σ (or D) does not depend on x, eq. (2.124) reduces to eq. (2.103), confirming the equivalence of the two interpretations in this special case.

There are rules for transforming a Langevin equation interpreted in the Ito convention to an equivalent one interpreted in the Stratonovich convention and back. As shown in appendix A.8, one obtains the rule that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(x,t) + \frac{1}{2}\,\sigma(x,t)\frac{\partial}{\partial x}\sigma(x,t) + \sigma(x,t)\,\xi \quad \text{(Ito)}$$

is equivalent to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(x,t) + \sigma(x,t)\,\xi \quad \text{(Stratonovich).} \tag{2.127}$$

The extra term necessary to change convention is often called the **noise-induced drift**.

Which convention is the most appropriate to describe a given natural phenomenon? The short answer is "it depends." One can show that, considering a noise source with a finite correlation time and then taking the limit of vanishing correlation time, the resulting Langevin equation is of the Stratonovich type. Therefore, Stratonovich is the correct interpretation for many systems affected by noise characterized by a negligible but finite correlation time. On the other hand, the evaluation of Stratonovich integrals requires knowledge of the "future," i.e., of the value of functions after the initial instant of each time step; see eq. (2.122). This aspect is unrealistic when the noise is generated by the physical process itself. In general, when modeling a system with multiplicative noise, it is safer to derive rigorously the Langevin equation and therefore to make sure that the interpretation is correct.

Traditionally, most physicists tend to prefer to work with the Stratonovich convention, whereas mathematicians tend to prefer the Ito convention. It is useful to be aware of the pros and cons of this choice. The Stratonovich convention has the advantage of respecting the chain rule, i.e., $\mathrm{d}[f(x)]/\mathrm{d}t = f'(x)\,\mathrm{d}x/\mathrm{d}t$, even if x follows a Langevin equation. In the case of Ito, assuming that $x(t)$ is the solution of eq. (2.111), the derivative $\mathrm{d}f/\mathrm{d}t$ of $f(x(t))$ is instead given by

$$\frac{\mathrm{d}f}{\mathrm{d}t} = v(x,t)f'(x) + \frac{1}{2}\sigma^2(x,t)f''(x) + \sigma(x,t)f'(x)\,\xi. \tag{2.128}$$

This relation is known as the **Ito formula**. Therefore, the Ito approach introduces a complication that is not present in the Stratonovich one. It has however the advantage that expectations of stochastic integrals, such as

$$\left\langle \int \mathrm{d}W \cdot f(x) \right\rangle,$$

always vanish in the Ito convention. This is in general not true for the Stratonovich convention, where such integrals must be evaluated on a case-by-case basis.

2.10 Information

It is often useful to quantify the "uncertainty" associated with a certain probability distribution over the states of a system $\mathcal{S}$. A measure of uncertainty must satisfy three reasonable properties: it must vanish when the system is known to be in one state, it must be maximal when the distribution is uniform over all possible states, and, if the system is made up of two independent systems, it must be the sum of the two uncertainties. Claude Shannon showed that the only function of the probabilities that satisfies these three requirements (up to a multiplicative constant) is the **Shannon entropy**

$$H(\mathcal{S}) = -\sum_x p_x \ln p_x, \tag{2.129}$$

where the sum runs over all the states of system $\mathcal{S}$. The larger $H(\mathcal{S})$, the larger the uncertainty about $\mathcal{S}$, up to a maximum $H(\mathcal{S}) = \ln \mathcal{N}$ when the probability is uniformly distributed over $\mathcal{N}$ states. With a slight abuse of notation, we interchangeably use $H(\mathcal{S})$ or $H(p)$ to denote the Shannon entropy of a system $\mathcal{S}$ having distribution $p = (p_x)$ over its states. In the latter case, we use the notation $p = (p_x)$ to denote by the symbol p the entire vector p_x. In the simple case of a binary variable ($x \in \{0, 1\}$) with equal probabilities $p_0 = p_1 = 1/2$, one has

$$H(\mathcal{S}) = \ln 2. \tag{2.130}$$

This quantity is called a **bit** of information.

If the system $\mathcal{S}$ is at thermal equilibrium, its thermodynamic entropy S is proportional to the Shannon entropy of its distribution over the microstates

$$S = k_{\mathrm{B}}\, H(\mathcal{S}) = -k_{\mathrm{B}} \sum_\xi p_\xi^{\mathrm{eq}} \ln p_\xi^{\mathrm{eq}}, \tag{2.131}$$

where k_{B} is the Boltzmann constant. Indeed, evaluating $H(\mathcal{S})$ with p^{eq} given by (2.45) leads to

$$H(\mathcal{S}) = -\sum_\xi p_\xi^{\mathrm{eq}} \frac{F - \epsilon_\xi}{k_{\mathrm{B}}T} = \frac{1}{k_{\mathrm{B}}T}\left(\langle \epsilon \rangle^{\mathrm{eq}} - F\right) = \frac{S}{k_{\mathrm{B}}}, \tag{2.132}$$

since $F = \langle \epsilon \rangle^{\mathrm{eq}} - TS$. Equation (2.131) is known as the **Gibbs relation**.

The dissimilarity between two probability distributions $p = (p_x)$ and $q = (q_x)$ over the states of the same system is measured by the **Kullback-Leibler divergence** (also called the **relative entropy**),

$$D_{\mathrm{KL}}(p\|q) = \sum_x p_x \ln \frac{p_x}{q_x}. \tag{2.133}$$

The Kullback-Leibler divergence has the following properties:

- $D_{\mathrm{KL}}(p\|q) \geq 0$;
- $D_{\mathrm{KL}}(p\|q) = 0$ only if $p_x = q_x$, $\forall x$.

To prove these results, we consider that $-\ln x$ is a convex function of x and therefore satisfies the **Jensen inequality** $\langle -\ln x \rangle \geq -\ln \langle x \rangle$; see appendix A.1. We therefore obtain

$$D_{\mathrm{KL}}(p\|q) = -\sum_x p_x \ln \frac{q_x}{p_x} \geq -\ln \sum_x p_x \frac{q_x}{p_x} = -\ln \sum_x q_x = -\ln 1 = 0. \tag{2.134}$$

In eq. (2.134), equality holds only if $q_x = p_x$ for all x, since $-\ln x$ is a strictly convex function. One has in general $D_{\mathrm{KL}}(p\|q) \neq D_{\mathrm{KL}}(q\|p)$. For instance, if $q_x > 0$, $\forall x$, while $p_x = 0$ for some x, $D_{\mathrm{KL}}(p\|q)$ is finite, but $D_{\mathrm{KL}}(q\|p)$ diverges.

We now consider two systems $\mathcal{S}_1$ and $\mathcal{S}_2$ whose states are respectively denoted by x and y. The information shared by the two systems is measured by the **mutual information**

$$I(\mathcal{S}_1 : \mathcal{S}_2) = \sum_{x,y} p_{x,y} \ln \frac{p_{x,y}}{p_x p_y}, \tag{2.135}$$

where $p_{x,y}$ is the joint probability distribution of the two systems and p_x, p_y are their **marginal distributions**; see appendix A.3. The mutual information can be seen as the Kullback-Leibler divergence between the joint distribution $p_{x,y}$ and the product of the marginal distributions p_x and p_y. Therefore, it is nonnegative and vanishes only if the two variables x and y are independent. The mutual information is symmetric: $I(\mathcal{S}_1 : \mathcal{S}_2) = I(\mathcal{S}_2 : \mathcal{S}_1)$. It can also be expressed in terms of the conditional probability distribution $p_{x|y}$ of the state of $\mathcal{S}_1$, given that of $\mathcal{S}_2$, by

$$I(\mathcal{S}_1 : \mathcal{S}_2) = \sum_{x,y} p_{x,y} \ln p_{x|y} - \sum_{x} p_x \ln p_x. \tag{2.136}$$

The second term on the right-hand side is the Shannon entropy $H(\mathcal{S}_1)$ of $\mathcal{S}_1$. The first term is minus the **conditional entropy** of $\mathcal{S}_1$, given $\mathcal{S}_2$:

$$H(\mathcal{S}_1|\mathcal{S}_2) = -\sum_{x,y} p_{x,y} \ln p_{x|y}. \tag{2.137}$$

When the two systems are statistically independent, i.e., $p_{x,y} = p_x p_y$, we have $H(\mathcal{S}_1|\mathcal{S}_2) = H(\mathcal{S}_1)$, and the mutual information vanishes. We moreover have

$$I(\mathcal{S}_1 : \mathcal{S}_2) = H(\mathcal{S}_1) - H(\mathcal{S}_1|\mathcal{S}_2) = H(\mathcal{S}_2) - H(\mathcal{S}_2|\mathcal{S}_1). \tag{2.138}$$

The Shannon entropy $H(\mathcal{S}_1, \mathcal{S}_2)$ of the joint distribution of the states of $\mathcal{S}_1$ and $\mathcal{S}_2$ is called the **joint entropy** of the two variables. It satisfies the relation

$$H(\mathcal{S}_1, \mathcal{S}_2) = H(\mathcal{S}_1) + H(\mathcal{S}_2|\mathcal{S}_1), \tag{2.139}$$

which is known as the **chain rule**. By exploiting (2.138), we also have

$$H(\mathcal{S}_1, \mathcal{S}_2) = H(\mathcal{S}_1) + H(\mathcal{S}_2) - I(\mathcal{S}_1 : \mathcal{S}_2). \tag{2.140}$$

Equation (2.140) clarifies that the mutual information quantifies the reduction of the uncertainty of the pair $(\mathcal{S}_1, \mathcal{S}_2)$ due to their mutual dependence. A similar relation holds for the Kullback-Leibler divergence:

$$D_{\mathrm{KL}}(p(\mathcal{S}_1, \mathcal{S}_2)\|q(\mathcal{S}_1, \mathcal{S}_2)) = D_{\mathrm{KL}}(p(\mathcal{S}_1)\|q(\mathcal{S}_1)) + D_{\mathrm{KL}}(p(\mathcal{S}_2|\mathcal{S}_1)\|q(\mathcal{S}_2\|\mathcal{S}_1)), \tag{2.141}$$

where

$$D_{\mathrm{KL}}(p(\mathcal{S}_2|\mathcal{S}_1)\|q(\mathcal{S}_2\|\mathcal{S}_1)) = \sum_{x,y} p_{x,y} \ln \frac{p_{x|y}}{q_{x|y}}. \tag{2.142}$$

This result can be verified by a direct calculation. An immediate consequence of eq. (2.141) is that the Kullback-Leibler divergence between two systems cannot increase if we average out some variables.

Given three systems $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$, whose states are respectively denoted by x, y, and z, the **conditional mutual information** between $\mathcal{S}_1$ and $\mathcal{S}_2$, given $\mathcal{S}_3$, is defined by

$$I(\mathcal{S}_1 : \mathcal{S}_2 | \mathcal{S}_3) = \sum_{x,y,z} p_{xyz} \ln \frac{p_{xy|z}}{p_{x|z} p_{y|z}}, \tag{2.143}$$

where $p_{x|z}$ and $p_{y|z}$ are conditional marginal distributions, i.e.,

$$p_{y|z} = \sum_x p_{xy|z}; \qquad p_{x|z} = \sum_y p_{x,y|z}. \tag{2.144}$$

The conditional mutual information $I(\mathcal{S}_1 : \mathcal{S}_2 | \mathcal{S}_3)$ is nonnegative and vanishes only if the joint conditional probability distribution of x and y, given z, factorizes into the product of the conditional marginal distributions. This happens, for instance, if $\mathcal{S}_1$ depends on $\mathcal{S}_2$ only via $\mathcal{S}_3$. The mutual information also satisfies a chain rule, which reads

$$I(\mathcal{S}_1, \mathcal{S}_2 : \mathcal{S}_3) = I(\mathcal{S}_1 : \mathcal{S}_3) + I(\mathcal{S}_2 : \mathcal{S}_3 | \mathcal{S}_1), \tag{2.145}$$

where $I(\mathcal{S}_1, \mathcal{S}_2 : \mathcal{S}_3)$ is the mutual information between $\mathcal{S}_1 \cup \mathcal{S}_2$ and $\mathcal{S}_3$. Derivation of eq. (2.145) is left as exercise 2.7.

2.11 Further reading

The aim of this chapter is limited to giving a bird's-eye perspective on theories that lie at the foundation of stochastic thermodynamics. Several books explain much more extensively the main concepts presented in each of the sections of this chapter, and sometimes choosing among the many classic references might be a matter of personal taste. Callen [28] and Pippard [130] are established textbooks in thermodynamics. Feynman's lectures in physics [54] provide an original introduction to thermodynamics and statistical mechanics, including ideas that, in hindsight, were seminal for stochastic thermodynamics. De Groot and Mazur [37] is a good reference on "nonstochastic" nonequilibrium thermodynamics. Landau et al. [95], Chandler [29], and Peliti [125] are classic references on statistical mechanics.

Equation (2.52) is a basic example of a fluctuation-dissipation relation. These relations are of paramount importance in nonequilibrium statistical physics, as discussed throughout this book. Marini Bettolo Marconi et al. [110] provide a comprehensive review on fluctuation-dissipation relations.

The basic concepts of stochastic processes were introduced into physics by the pioneering work of Einstein [44] on Brownian motion (English translation in [46]). Nowadays the theory of stochastic processes is rather well developed, even if it is still playing a relatively marginal role in many physics curricula. There are many excellent books on stochastic processes that present the theory at different mathematical levels and with slightly different angles. Berg [15] is a basic introduction that focuses on the physical concepts rather than on the mathematics. Gardiner [59], van Kampen [171], and Risken [139] are references of a more mathematical nature. Øksendal

[120] is an excellent, though even more advanced book. Other books focus on specific aspects of the theory of stochastic processes. One important example is first-passage time problems, discussed in Redner [138].

The seminal paper by Shannon [154] already contains the main ideas in information theory. Khinchin [87] provides a more systematic introduction to Shannon's theory. Cover and Thomas [32] and MacKay [105] present information theory in a more modern and extended way.

2.12 Exercises

2.1 Consider a master equation with three states, $x \in \{0, 1, 2\}$, and with constant rates $k_{xx'}$ $(x \neq x')$, none of which vanishes. Write down the explicit form of the master equation. Evaluate the steady-state probability distribution p_x^{st}, $x \in \{0, 1, 2\}$ and the corresponding probability current $J^{\mathrm{st}} = k_{x'x} p_x^{\mathrm{st}} - k_{xx'} p_{x'}^{\mathrm{st}}$ with $x' = x + 1 \mod 3$. Derive the conditions on the rates k such that detailed balance is satisfied, i.e., such that J^{st} vanishes.

2.2 Consider a master equation with four states, $x \in \{0, 1, 2, 3\}$, and with rates $k_{xx'}$ $(\forall x \neq x')$, none of which vanishes. Show that if $k_{xx'} = k_{x'x}$, then the unique stationary distribution is $p_0 = p_1 = p_2 = p_3 = 1/4$. Show with an example that this is not necessarily the case if some of the rates vanish.

2.3 A collection of N white balls and N black balls, with $N \geq 3$, are randomly distributed in two urns, so that each contains N balls. At each step t, one ball is extracted from each urn; the two balls are swapped and put back in the urns. Denote by x_t the number of white balls in the first urn. Show that x_t is a Markov process and express its jump rates $k_{xx'} = p(x_{t+1} = x | x_t = x')$. Discuss whether the process satisfies detailed balance. Evaluate the stationary distribution p_x^{st} and the decay rate of the correlation function $\langle x_t x_0 \rangle - \langle x_0 \rangle^2$, where x_0 is drawn from the stationary distribution.

2.4 Consider an infinite sequence of independent, identically distributed real-valued random variables $y = (y_0, y_1, \ldots)$. Another sequence $x = (x_0, x_{1,2}, \ldots)$ is recursively defined by

$$x_0 = y_0; \qquad x_n = y_n + a\, x_{n-1},$$

where $0 < a < 1$.

a. Show that, if the distribution of y_ℓ is Gaussian, the distribution of x_ℓ approaches for $\ell \to \infty$ a stationary Gaussian distribution and evaluate its parameters.
b. Show that, if y_ℓ is *not* Gaussian distributed, the stationary distribution of x_ℓ, if it exists, is not Gaussian.

2.5 A Brownian particle is confined in a one-dimensional potential. Its position $x(t)$ satisfies the following Langevin equation in the Stratonovich representation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\Gamma k x + \sigma\, \xi(t),$$

where $\sigma > 0$ and $\xi(t)$ is Gaussian random white noise, satisfying $\langle \xi(t) \rangle = 0$ and $\left\langle \xi(t)\, \xi(t') \right\rangle = \delta(t - t')\ \forall t, t'$. Write the formal solution of the equation for $t \geq 0$ as a function of the initial condition $x(0) = x_0$ and of the realization $\xi(t')$ $(0 \leq t' \leq t)$ of the noise. Evaluate $\langle x(t) \rangle$ and $\left\langle (x(t) - \langle x(t) \rangle)^2 \right\rangle$ for $t \to \infty$. Assuming that the potential is harmonic, $U(x) = \frac{1}{2} k x^2$, find the relation that Γ and σ must satisfy for the particle to reach the equilibrium distribution $p^{\mathrm{eq}}(x) \propto \exp(-U(x)/k_{\mathrm{B}}T)$ for $t \to \infty$. Write down the Fokker-Planck equation associated with the Langevin equation and show that the equilibrium distribution is a solution.

2.6 Prove eq. (2.110) for a system satisfying a one-dimensional Fokker-Planck equation (2.103), assuming that the detailed balance condition eq. (2.109) is satisfied.

2.7 Prove the chain rule for the mutual information, eq. (2.145).

Index