Irreversibility and information

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von

Léo Granger

geboren am 16.02.1986 in Paris

Eingereicht am 26.11.2013

Eingereicht am 26.11.2013

1. Gutachter: Prof. Dr. Holger Kantz
2. Gutachter: Prof. Dr. Jens-Uwe Sommer

Verteidigt am 28.05.2014

Acknowledgments

My deepest thank goes to Holger Kantz, he was the best PhD supervisor I could imagine. Giving me a complete freedom for my research, being open for new ideas, having his door open as well, spending time for discussion, he allowed me to find and study the problems I am really interested in. Further thanks go to Markus Nieman and Jochen Bröcker who always had time for interesting discussions. I thank Julia Gundermann for proofreading a big part of this thesis. Furthermore, I thank the rest of the TSA group: Stefan Siegert, Anja von Wulffen, Marc Höll, Colm Mulhern, Stephan Bialonski, for passionated discussions at the lunch table, and for stimulating games at the kicker table. And Pablo Sartori for his never ending energy.

I would like to thank the Max Planck Institute for the Physics of Complex System for financial support, for the cheap coffee, and for providing me with a desk, a Computer and with the pens and notebooks necessary for my research.

Einen herzlichen Dank an Fanny, Till, Anton, Sara, Anett, Bubbel, Corinna, Klemens, Franzi, für die nichtwissenschaftliche Seite von meiner Zeit in Dresden.

Et Florent (o< coin).

Finally, last but not least, merci aux deux femmes de ma vie, liebe Susanne et p'tite Camille pour votre soutient infaillible, surtout à la fin de ma thèse.

Contents

1 Introduction

1.1 Irreversibility, information, and entropy

When he first mentioned his “intelligent being” able to act on the molecular scale, James Clerk Maxwell was the first to point out the role of information in a microscopic (atomic) understanding of irreversibility [Max71]. In a simple Gedankenexperiment, he showed that if one was able to use information about the microscopic state of a system, like the positions and momenta of the particles in a gas, then one could trigger processes that would otherwise be impossible, like the creation of a heat flow from a cold body to a hotter one without the expenditure of work. This “intelligent being” was baptised “Maxwell’s Demon” by William Thomson [Tho79] and is now the paradigm for the relation between information processing and irreversibility [LR02]. This relation is symmetric: On the one hand, information about the microscopic state of a macroscopic system can be used in order to “undo” some irreversible process, like creating a heat flow from a cold body to a hotter one or converting heat from one single heat source into mechanical work. On the other hand, information processing must be implemented on physical devices themselves obeying the laws of irreversibility, and should be accompanied by a certain “amount” of irreversibility at least enough to compensate the aforementioned “undoing” made possible.

Interestingly enough, the quantification of irreversibility and information both fall under the name *entropy*. Source of fascination and confusion, entropy is the nightmare of many physics and engineering undergraduate students and the passion of many researchers and hobby scientists. This word has several meanings and is used to denote different concepts. Among other things, entropy denotes the central quantity both of thermodynamics, the physical theory of irreversibility, and of information theory. The confusion (and perhaps the fascination) is increased even further by the fact that information and irreversibility are deeply linked and information theoretic entropy helps to understand thermodynamic entropy.

This thesis is concerned with the irreversibility of certain information processing operations. The two main results of this thesis concern the recording and erasure of information on a physical memory and the measurement, or acquisition of information by a physical device. These operations are shown to be irreversible, and their degree of irreversibility is related to the amount of information processed. The degree of irreversibility of a process is quantified by the amount of thermodynamic entropy produced during that process. The amount of information processed is quantified using quantities derived from the information theoretic entropy. In order to present the results of this thesis, these two kinds of entropies are carefully introduced and a particular attention is given to how information theoretic entropy relates to thermodynamic entropy.

1.2 Overview

The first part of this document, consisting of chapters 2 and 3 introduces the theoretical background while the second part consisting of chapters 4 and 5 presents the results.

Chapter 2 presents the basics of classical thermodynamics and equilibrium statistical mechanics. Classical thermodynamics, introduced in section 2.1, is a phenomenological theory for the irreversibility of macroscopic transformations. The state of a system is described using a few variables called *state variables* like the energy, the volume or the magnetization. A physical process is a change in the state variables. The entropy is the starting-point of our presentation. It is a function of the state variables. Entropy has the fundamental property that it specifies the processes that are physically possible: A process is physically possible if and only if it leads to an increase of the total entropy of all the systems involved. The entropy produced by a given process is a quantitative measure of the irreversibility of the process. The concepts of heat and temperature are derived from the entropy principle. There is a natural link between energy loss (dissipation) and entropy production. This link is particularly simple in *isothermal processes* where the two quantities are proportional.

Material systems are composed of many particles in interaction. The “real”, or “microscopic” state of a system is actually the list of the positions and velocities of all the particles constituting the system. The time evolution of the microscopic state of the system is specified by the Hamilton function, summarizing the characteristics of the particles and their interaction. The thermodynamic description misses a lot of details about the microscopic state of a system and a lot of microscopic configurations are compatible with one thermodynamic state. Equilibrium statistical mechanics, introduced in section 2.2, makes a bridge between the microscopic and the thermodynamic descriptions. A thermodynamic state is seen as an *ensemble* of microscopic states. The aim of statistical mechanics is then to find the distribution of microscopic states correctly describing the thermodynamics of the system. The distribution compatible with thermodynamics is the Boltzmann-Gibbs or *canonical* distribution. This distribution allows to numerically compute the thermodynamic functions starting from a Hamiltonian (microscopic) description.

Describing matter at the molecular level, it is tempting to assume that we can act at this very level. Section 2.3 shows that by making use of information about the microscopic state of a system, Maxwell’s Demon is able to perform processes that are not permitted by classical thermodynamics.

Chapter 3 presents the relatively recent theory of stochastic thermodynamics starting with a short introduction to information theory in section 3.1. Information theory provides theoretical tools able to quantify information. Missing information about a physical quantity is modelled by considering the quantity random. The information missing about that quantity is quantified by its information theoretic or *Shannon* entropy. The concepts of information theory can straightforwardly be applied to statistical mechanics: When one only specifies the thermodynamic state of a system, one assumes that its microscopic state is random. Statistical mechanics tells us that the probability distribution of the microscopic state is the canonical distribution. The information that one misses about the microscopic state is then quantified by its Shannon entropy. It turns out that this Shannon entropy is the thermodynamic entropy given by equilibrium statistical mechanics. Moreover, the canonical distribution is justified from an information theoretic point of view: It is the distribution best representing our knowledge of the microscopic state of a system when one only knows its thermodynamic state. In fact, among all distributions compatible with the thermodynamic state of the system, the canonical distribution maximizes the Shannon entropy.

Stochastic thermodynamics, presented in section 3.2, extends equilibrium statistical mechanics to non equilibrium isothermal processes. It allows for non canonical, or non equilibrium distributions and simply assumes that such a non equilibrium distribution relaxes towards equilibrium according to a linear equation. The various thermodynamic functions are successfully generalized for non equilibrium states. In particular, the thermodynamic entropy of a non equilibrium distribution is simply given by its Shannon entropy. The relaxation towards equilibrium is the source of irreversibility and the entropy production is actually a loss of information about the microscopic state of the system. Stochastic thermodynamics provides an ideal framework for the formalization of the problem of Maxwell's Demon: The acquisition of information about the microscopic state of the system naturally leads to a reduction of its entropy. This reduction is equal to the amount of information provided by the measurement.

Apart from introducing the theoretical framework used to derive the results of the second part, the goal of the first part is also to emphasize the quantitative relation between the irreversibility of a process and the loss or gain of information about the microscopic state of the system. In the second part, this relation is investigated the other way round. It is studied whether and to what extend the loss and gain of information coming from an external source is irreversible. The answer is 'yes', and it is shown that the entropy produced during these processes is always greater than the information lost or gained. The results of chapters 4 and 5 were published respectively in [GK13] and [GK11] in a slightly more restrictive setup.

Chapter 4 deals with information loss. This chapter presents a simple yet general model for a physical memory device and the processes of recording and erasing a piece of information on it. It is shown that the erasing of the information is an irreversible process and that the entropy produced during that process is related to the amount of information erased. The piece of information comes from an external source. It is just a random event such as the result of the flipping of a coin. The memory is modelled as a system obeying the laws of stochastic thermodynamics. The main requirement concerning the erasure of the information is that one should not make use of the information recorded while erasing it. In fact, making use of the information while erasing it would imply that the information is also present elsewhere. The information present in the memory is identified as the correlation of the microscopic state of the memory with the event recorded. Not making use of information contained in the memory amounts to not using information about the microscopic state of the memory and this turns out to be the reason for the irreversibility of the erasure. The relation between irreversibility and information erasure is resolved in time: The rate at which entropy is produced is at least equal to the rate at which information is erased all along the erasure process.

Chapter 5 is concerned with the acquisition of information coming from an external source. A simple model is developed for a physical measurement device able to obtain information about the microscopic state of a system. The measurement process is analyzed and shown to be irreversible. The amount of entropy produced is related to the information processed. Remarkably, there is equality between the two quantities only if they are both zero. The model presented is simple in that we assume no back action of the measurement device on the system on which the measurement is performed. The measurement process is cyclic. A measurement event happens to be separable into two different processes: the loss of the information obtained during the previous measurement cycle and the acquisition of the information of the current cycle. Each of these processes is entropy producing.

2 Classical thermodynamics and statistical mechanics

2.1 Classical thermodynamics

The aim of this section is to give a short introduction to classical thermodynamics. The presentation is inspired by [LY99]. It might differ from traditional introductions to thermodynamics in that thermal engines and maximum efficiencies are not mentioned. Instead, the central concepts of entropy and entropy production are introduced from the very beginning. The presentation ends with the definition of isothermal processes which will be the processes considered in the rest of the document.

This introduction is not meant to be complete. Interested readers might find a complete presentation of classical thermodynamics in [Cal85].

2.1.1 Adiabatic thermodynamics

The irreversibility assumption is formulated for *adiabatic processes*. An adiabatic process is a process during which the system under consideration only exchanges energy with its environment in form of mechanical work. Since purely mechanical transformations are reversible, if an adiabatic process is irreversible, the irreversibility must lie inside the system. The irreversibility assumption is then the following: *There is a function called* entropy*, which only depends on the state of the system under consideration (i.e. on its energy, volume etc. . .) and which is such that an adiabatic process is possible if and only if the entropy of the final state is not less than the entropy of the initial state.*

More formally, let us consider a system over which we have a certain control through a *control parameter* denoted by λ[1]. The control parameter might be the volume of the system, a magnetic or electric field applied to the system. By changing λ, we might change the *energy* E of the system. The energy exchanges between the adiabatic system and its environment can only occur in form of *work*. Hence, in order to drive our system from one value of λ to another, an amount of work has to be performed which is equal to the change in the energy of the system. For instance, in

[1] Throughout this document, we will assume that λ is a scalar quantity, however generalization to several control parameters is straightforward.

order to compress a gas confined in a closed box (i.e. to decrease its volume), it is necessary to perform some work. On the other hand, it is possible to extract some work during the expansion of a gas (i.e. an increase of volume).

The energy of the system and the control parameter completely specify the thermodynamic state of the system. They are called *state variables.* The irreversibility principle is formalized in the assumption that there exists a *state function* (a function of the state variables) called *entropy* and denoted by $S(E, \lambda)$ which can only increase in adiabatic transformations. In other words, it is possible to bring the system from a state (E, λ) to a state (E', λ') if and only if

$$S(E, \lambda) \leq S(E', \lambda'). \tag{2.1}$$

If there is equality, then the transformation is said to be *reversible.* In fact, in that case, it is possible to perform the *reverse process*, i.e. to bring the system from state (E', λ') to state (E, λ). Otherwise, when strict inequality holds, it is not possible to perform the reverse process and the process is said to be *irreversible.* During this process, the entropy of the system increases by an amount $\Delta S = S(E', \lambda') - S(E, \lambda) \geq 0$ and we say that the amount ΔS was *irreversibly produced* or simply *produced* during the process. The amount of entropy produced during a process quantifies the "degree of irreversibility" of the process. The more entropy is produced during a process, the "more" irreversible it is.

The entropy has the following properties:

- It is an increasing function of the energy:

$$\frac{\partial S}{\partial E}(E, \lambda) \geq 0. \tag{2.2}$$

- It is a concave function of the energy:

$$\frac{\partial^2 S}{\partial E^2}(E, \lambda) \leq 0. \tag{2.3}$$

- It is *additive*, i.e. the entropy of a compound system is the sum of the entropies of its constitutive parts.

Saying that the entropy is an increasing function of the energy is equivalent to say that it is always possible to increase the energy of an adiabatic system (i.e. to "heat it up") without changing the control parameter. However, in order to decrease the energy of the system (i.e. to "cool it down"), it is necessary to change the control parameter. As a consequence, the minimum amount of work needed in order to perform a given process is minimum if the process is performed reversibly. In fact, consider a process where λ is driven from λ_0 to λ_1. Let E_0 and E_1 be the initial and final energy of the system. The amount of work performed on the system during this process is $W = E_1 - E_0$. Let $\tilde{E}_1$ be the final energy of the system if the process had been performed reversibly. It is given by the condition:

$$S(E_0, \lambda_0) = S(\tilde{E}_1, \lambda_1). \tag{2.4}$$

The work that would be performed during the reversible process is $W_{\text{rev}} = \tilde{E}_1 - E_0$. We have:

$$S(E_0, \lambda_0) = S(\tilde{E}_1, \lambda_1) \leq S(E_1, \lambda_1). \tag{2.5}$$

Since the entropy is an increasing function of the energy, we have $\tilde{E}_1 \leq E_1$ and hence:

$$W - W_{\text{rev}} = E_1 - \tilde{E}_1 \geq 0, \tag{2.6}$$

with equality if and only if the process was performed reversibly. The extra amount of work performed compared to a reversible process is called the *dissipated work*:

$$W_{\text{d}} = W - W_{\text{rev}} \geq 0. \tag{2.7}$$

This amount of work is lost, it cannot be retrieved performing the reverse process. Similarly to the entropy produced, the work dissipated during a process is a measure of the irreversibility of the process. Both quantities are zero for reversible process and positive for irreversible process. However, the work dissipated seems to be a more practical measure since it is expressed in terms of mechanical energy which can in principle be measured. The relation between the energy dissipated and the entropy produced during a process is quite complicated in general.

2.1.2 Thermal contact

The additivity of the entropy implies that two systems 1 and 2 with energies E_1 and E_2 and entropies $S_1(E_1)$ and $S_2(E_2)$ can be combined to form a compound system with entropy[2]

$$S_{\text{tot}}(E_1, E_2) = S_1(E_1) + S_2(E_2). \tag{2.8}$$

The two systems are said to be in *thermal contact* if they are allowed to exchange energy. The energy exchanged between two systems in thermal contact is called *heat transfer*.

Heat transfer can only happen in such a way that the total entropy (2.8) increases. Imagine that an infinitesimally small amount $\mathrm{d}E$ is transferred from system 1 to system 2. Additivity implies that the total variation of the entropy is the sum of the

[2] We omit for a moment the possible dependence on the control parameters since we are not going to consider their variations.

variations of the entropies of the two subsystems:

$$\mathrm{d}S_{\mathrm{tot}} = \mathrm{d}S_1 + \mathrm{d}S_2 = \frac{\partial S_1}{\partial E_1}\mathrm{d}E_1 + \frac{\partial S_2}{\partial E_2}\mathrm{d}E_2. \tag{2.9}$$

During this process, system 1 looses amount dE of energy and system 2 receives the same amount. Hence:

$$\mathrm{d}E_1 = -\mathrm{d}E \tag{2.10}$$

and

$$\mathrm{d}E_2 = \mathrm{d}E. \tag{2.11}$$

Equation (2.9) thus becomes:

$$\mathrm{d}S_{\mathrm{tot}} = \left(\frac{\partial S_2}{\partial E_2} - \frac{\partial S_1}{\partial E_1}\right)\mathrm{d}E. \tag{2.12}$$

This quantity is positive if and only if dE has the same sign as $\frac{\partial S_2}{\partial E_2} - \frac{\partial S_1}{\partial E_1}$. Hence, the only heat exchanges possible are the ones in the direction of increasing $\frac{\partial S}{\partial E}$.

The quantity $\frac{\partial S}{\partial E}$ is called the *inverse temperature* and is usually denoted by the Greek letter β. The inverse temperature of a system in state (E, λ) is defined by:

$$\beta(E, \lambda) = \frac{\partial S}{\partial E}(E, \lambda). \tag{2.13}$$

As the entropy, it is a state function. Since the entropy is an increasing function of the energy, the inverse temperature is positive:

$$\beta(E, \lambda) \geq 0. \tag{2.14}$$

Furthermore, it is a decreasing function of the energy since the entropy is concave.

To be possible, an infinitesimal amount of heat dE transferred from system 1 to system 2 has to satisfy:

$$\left(\beta_2(E_2) - \beta_1(E_1)\right)\mathrm{d}E \geq 0. \tag{2.15}$$

Hence dE can be positive if and only if the inverse temperature of system 1 is lower than the inverse temperature of system 2. Since β is a decreasing function of the energy, any heat transfer tends to reduce the difference $\beta_2 - \beta_1$.

If the two systems have the same inverse temperature: $\beta_1(E_1) = \beta_2(E_2)$, then no heat exchange is allowed between them. In fact, due to the concavity of the entropy, such a heat exchange would necessarily lead to a reduction of the total entropy. The two systems are then said to be in *thermal equilibrium*. The joint state of two systems in thermal equilibrium is completely specified by the total energy $E_{\mathrm{tot}} = E_1 + E_2$ (and the possible set of control parameters). The entropy of this state is then given by the maximum value of (2.8) over the possible repartitions of the total energy between the two systems:

$$S_{\mathrm{tot}}^{\mathrm{eq}}(E_{\mathrm{tot}}) = \max_{E_1+E_2=E_{\mathrm{tot}}} S_1(E_1) + S_2(E_2). \tag{2.16}$$

This maximum is reached when the inverse temperatures of the two systems are equal and we can then define the inverse temperature of the joint system by applying equation (2.13) to the joint system:

$$\beta_{\mathrm{tot}}(E_{\mathrm{tot}}) = \frac{\partial S_{\mathrm{tot}}^{\mathrm{eq}}}{\partial E_{\mathrm{tot}}}(E_{\mathrm{tot}}), \tag{2.17}$$

and verify that it is indeed equal to the inverse temperature of the two subsystems: $\beta_{\mathrm{tot}}(E_{\mathrm{tot}}) = \beta_1(E_1) = \beta_2(E_2)$.

The inverse temperature is an “index of equilibrium” in the sense that two systems are in equilibrium if and only if they have the same inverse temperature. Moreover, the inverse temperature gives the direction in which energy (heat) transfers are possible.

We assume that when two systems are put into thermal contact, they will exchange heat until they are in equilibrium, i.e. until their inverse temperatures are equal. This is the “thermalization assumption”.

The absolute temperature is defined as the inverse of the inverse temperature:

$$T(E,\lambda) = \frac{1}{\beta(E,\lambda)}. \tag{2.18}$$

Historically, T was introduced before β which explains the name “inverse temperature” for β. In the following, we will use T or β indifferently. Moreover, we express the temperature in the same units as the energy implying that we consider the entropy to be dimensionless.

Systems having a high temperature are said to be “hot” and systems having a low temperature are said to be “cold”. With this definitions, inequality (2.15) implies that heat spontaneously flows from hot bodies to colder ones.

2.1.3 Isothermal thermodynamics

When a system is so big, that its inverse temperature is not influenced by heat exchanges with another system, it is called a *heat bath*, *heat reservoir* or *thermostat*. When a heat bath at inverse temperature β_{r} receives an amount ΔE_{r} of energy, its entropy varies according to

$$\Delta S_{\mathrm{r}} = \beta_{\mathrm{r}} \Delta E_{\mathrm{r}}, \tag{2.19}$$

where the subscript “r” stands for “reservoir”.

Equilibrium with a heat bath

Consider a system with an entropy function $S(E,\lambda)$ in contact with a heat bath at inverse temperature β_r. An amount Q of heat transferred from the heat bath to the system leads to the production of an amount of entropy equal to

$$\Delta S_\mathrm{tot} = \Delta S + \Delta S_\mathrm{r} = \Delta S - \beta_\mathrm{r} Q, \tag{2.20}$$

where $\Delta S = S(E+Q,\lambda) - S(E,\lambda)$ is the variation of the entropy of the system and $\Delta S_\mathrm{r} = \beta_\mathrm{r}\Delta E_\mathrm{r} = -\beta_\mathrm{r} Q$ is the variation of the entropy of the heat bath. The entropy production (2.20) can be written as the variation of the following function:

$$\psi(\beta_\mathrm{r},\lambda,E) = S(E,\lambda) - \beta_\mathrm{r} E. \tag{2.21}$$

In fact, one can check that $\Delta S_\mathrm{tot} = \psi(\beta_\mathrm{r},\lambda,E+Q) - \psi(\beta_\mathrm{r},\lambda,E)$. As a consequence, heat exchanges between the system and the heat bath are only possible in the direction increasing ψ. The function ψ is called *Massieu function*.

Due to the concavity of the entropy, the Massieu function $\psi(\beta_\mathrm{r},\lambda,E)$ has a unique maximum as a function of E for every β_r and λ. From the previous paragraph, we know that the maximum is attained when the inverse temperature of the system is equal to the inverse temperature of the heat bath, in other words, for $E = E_\mathrm{eq}(\beta_\mathrm{r},\lambda)$ satisfying:

$$\beta(E_\mathrm{eq},\lambda) = \frac{\partial S}{\partial E}(E_\mathrm{eq},\lambda) = \beta_\mathrm{r}. \tag{2.22}$$

When this condition is satisfied, no heat exchange is possible between the system and the heat bath. The system is then said to be *in equilibrium* with the heat bath. The thermalization assumption implies that when a system is put into contact with a heat bath at inverse temperature β_r, it will exchange heat with the heat bath until it reaches equilibrium, i.e. until the inverse temperature of the system will be equal to β_r.

When the system is in equilibrium with the heat bath, then its thermodynamic state is completely specified by the inverse temperature of the heat bath, β_r (which is also the inverse temperature of the system) and the control parameter λ. The energy of the system is then a function of β_r and λ and it is given by equation (2.22). The equilibrium value of the Massieu function defines a new state function:

$$\psi_\mathrm{eq}(\beta_\mathrm{r},\lambda) = \max_E \psi(\beta_\mathrm{r},\lambda,E) = \psi(\beta_\mathrm{r},\lambda,E_\mathrm{eq}). \tag{2.23}$$

In the following, we will drop the subscripts “r” and “eq” and $\psi(\beta,\lambda)$ will denote the Massieu function of a system in equilibrium with a heat bath at inverse temperature β.

Work dissipation and entropy production in isothermal processes

Consider the following process: Our system is in equilibrium with a heat bath at inverse temperature β and we change the control parameter from λ_0 to λ_1. Such a process is called *isothermal*. We would like to investigate the amount of work needed to perform such a process and to compare it to the amount of entropy produced by the process.

The system and the heat bath can be viewed as an adiabatic compound system. Hence, the work performed is equal to the total change in energy in the two systems:

$$W = \Delta E + \Delta E_{\mathrm{r}} = \Delta E - Q, \tag{2.24}$$

where $\Delta E = E(\beta, \lambda_1) - E(\beta, \lambda_0)$ is the variation of the energy of the system and $\Delta E_{\mathrm{r}} = -Q$ is the amount of heat transferred from the heat bath to the system during the process. Since the energy of the system is a state function, its change just depends on the initial and final values of its state variables, i.e. of β and λ.

The total entropy produced during the process is the sum of the changes in entropy of the two systems:

$$\Delta S_{\mathrm{tot}} = \Delta S + \Delta S_{\mathrm{r}}. \tag{2.25}$$

As the energy, the entropy of the system is a state variable, hence its variation depends only on the initial and final values of the state variables: $\Delta S = S(\beta, \lambda_1) - S(\beta, \lambda_0)$. From the definition of the heat bath, its entropy variation is equal to:

$$\Delta S_{\mathrm{r}} = \beta \Delta E_{\mathrm{r}} = -\beta Q. \tag{2.26}$$

Equation (2.25) can be rearranged into:

$$\beta Q = \Delta S - \Delta S_{\mathrm{tot}} \leq \Delta S. \tag{2.27}$$

Hence, the variation of the entropy of the system is an upper bound to the quantity βQ, i.e. there is a maximum amount of heat that the system can receive during the process, and this maximum depends only on the initial and final values of β and λ, i.e. it is the variation of a state function. Equivalently, one can say that there is a minimum amount of heat that has to be transferred from the system to the environment. The maximum for Q is attained when $\Delta S_{\mathrm{tot}} = 0$, i.e. for reversible processes. It is equal to:

$$Q_{\mathrm{rev}} = \frac{\Delta S}{\beta} = T \Delta S. \tag{2.28}$$

Equation (2.19) tells us that heat baths always satisfy this condition. Hence heat baths act reversibly.

From equation (2.24) and the fact that Q is bound from above, inequality (2.27), we get that W is bound from below:

$$W = \Delta E - Q = \Delta E - Q_{\text{rev}} + T\Delta S_{\text{tot}} \geq \Delta E - Q_{\text{rev}}. \tag{2.29}$$

Hence, there is a minimum amount of work that has to be provided to the system in order to perform this process. This minimum is attained for reversible processes and is equal to:

$$W_{\text{rev}} = \Delta E - Q_{\text{rev}} = \Delta E - T\Delta S. \tag{2.30}$$

This is the variation of the following state function:

$$F(T, \lambda) = E(\beta, \lambda) - TS(\beta, \lambda), \tag{2.31}$$

where $T = 1/\beta$ is the *absolute temperature* of the heat bath. The function F is called *free energy*, or *Helmholtz free energy*. Its variations along an isothermal process give the reversible work of the process. The free energy is linked to the equilibrium Massieu function (2.23) by $\psi(\beta, \lambda) = -\beta F(T, \lambda)$, or equivalently $F(T, \lambda) = -T\psi(\beta, \lambda)$. The work dissipated during the process is then:

$$W_{\text{d}} = W - W_{\text{rev}} = W - \Delta F. \tag{2.32}$$

Combining equations (2.29) and (2.30), we obtain:

$$W_{\text{d}} = T\Delta S_{\text{tot}}. \tag{2.33}$$

In words: The work dissipated during an isothermal process is proportional to the entropy produced during the process.

It is not surprising that the work performed is minimal for a reversible process. In fact, isothermal processes are particular cases of adiabatic processes and hence the results of paragraph 2.1.1 still hold. However isothermal processes are somewhat simpler than adiabatic processes. Instead of being given by the isoentropic condition (2.4), the reversible work is given by the state function F given by equation (2.31). Moreover, the work dissipated during an isothermal process is proportional to the entropy produced.

Legendre transformation

It is possible to describe the state of the system using the inverse temperature β instead of the energy E. Then, the (equilibrium) Massieu function $\psi(\beta, \lambda)$ contains the same information as the entropy $S(E, \lambda)$. The function ψ is the *Legendre transform* of the function S. It is defined by:

$$\psi(\beta, \lambda) = S(E(\beta, \lambda), \lambda) - \beta E(\beta, \lambda), \tag{2.34}$$

where $E(\beta, \lambda)$ is obtained by inverting $\beta(E, \lambda) = \frac{\partial S}{\partial E}(E, \lambda)$ with respect to E. The partial derivatives of the function ψ are:

$$\frac{\partial \psi}{\partial \beta}(\beta, \lambda) = -E(\beta, \lambda) \tag{2.35}$$

and

$$\frac{\partial \psi}{\partial \lambda}(\beta, \lambda) = \frac{\partial S}{\partial \lambda}(E(\beta, \lambda), \lambda). \tag{2.36}$$

The simplest way of seeing this, although not very rigorous, is to compute the differential of $\psi = S - \beta E$:

$$\mathrm{d}\psi = \mathrm{d}S - \beta \mathrm{d}E - E\mathrm{d}\beta. \tag{2.37}$$

The differential of S reads:

$$\mathrm{d}S = \beta \mathrm{d}E + A\mathrm{d}\lambda, \tag{2.38}$$

where we have set $A = \frac{\partial S}{\partial \lambda}$. Inserting this into equation (2.37) above yields:

$$\mathrm{d}\psi = -E\mathrm{d}\beta + A\mathrm{d}\lambda. \tag{2.39}$$

And thus $-E$ and A are indeed the partial derivatives of ψ with respect to β and λ.

Before going further, we have to say a few words about the function A. It is a state function. For an adiabatic system, it is given by:

$$A(E, \lambda) = \frac{\partial S}{\partial \lambda}(E, \lambda). \tag{2.40}$$

For an isothermal system, it is given by:

$$A(\beta, \lambda) = \frac{\partial \psi}{\partial \lambda}(\beta, \lambda). \tag{2.41}$$

Here, and throughout this document, we abusively use the same letter A, whether its arguments are (E, λ) or (β, λ) in order to keep the notations light. But it should be understood that $A(E, \lambda)$ and $A(\beta, \lambda)$ *are different functions of their arguments*. The rigorous form to write it is equation (2.36). Similarly, we will use the notation $S(\beta, \lambda)$ instead of the more rigorous but heavier $S(E(\beta, \lambda), \lambda)$ to denote the entropy of an isothermal system. The quantity A is linked to the reversible work of an infinitesimal transformation. Define the *pressure*[3] $P = A/\beta$. The reversible work of an infinitesimal transformation, where the control parameter is changed by $\mathrm{d}\lambda$ is:

$$\delta W_{\text{rev}} = -P\mathrm{d}\lambda. \tag{2.42}$$

This is true for an adiabatic and for an isothermal transformation. Setting $\mathrm{d}S = 0$

[3] In fact, if λ is the volume, then P is indeed the thermodynamic pressure.

in equation (2.38), one obtains:

$$\delta W_{\mathrm{rev}}^{\mathrm{adia}} = \mathrm{d}E = -\frac{A}{\beta}\mathrm{d}\lambda = -P\mathrm{d}\lambda. \tag{2.43}$$

For an infinitesimal isothermal transformation, the reversible work is given by:

$$\delta W_{\mathrm{rev}}^{\mathrm{iso}} = \frac{\partial F}{\partial \lambda}\mathrm{d}\lambda = -\frac{1}{\beta}\frac{\partial \psi}{\partial \lambda}\mathrm{d}\lambda = -P\mathrm{d}\lambda. \tag{2.44}$$

The minus signs in the equations above are due to historical reasons: when thermodynamics was developed, people were interested in the amount of work they could extract during a given transformation.

2.2 Equilibrium statistical mechanics

2.2.1 From the microscopic to the macroscopic description

An adiabatic system is actually composed of many particles that interact with each other and are possibly subjected to some external field force. The dynamical state of the system is described by the set of positions $q = \{q_i\}$ and momenta $p = \{p_i\}$ of all the particles. The energy of the system is given by the Hamilton function or *Hamiltonian* $E = \mathcal{H}_\lambda(q,p)$. The Hamiltonian contains all the details concerning the dynamics of the particles including their interaction and possible external fields applied on the system. Driving the system means to change its Hamiltonian. It therefore depends on the control parameter λ. The evolution of the state of the system is given by the Hamiltonian equations:

$$\begin{aligned} \frac{\mathrm{d}q_i}{\mathrm{d}t} &= \frac{\partial \mathcal{H}_\lambda}{\partial p_i} \\ \frac{\mathrm{d}p_i}{\mathrm{d}t} &= -\frac{\partial \mathcal{H}_\lambda}{\partial q_i}. \end{aligned} \tag{2.45}$$

The Hamiltonian equations of motions are time reversible and for a constant value of λ, the energy is conserved (it is an *integral of the motion*). When λ is varied, the energy of the system changes and hence work is performed on the system.

In macroscopic systems such as considered in thermodynamics, the number of degrees of freedoms is enormously big, around 10^{24}. Hence, compared to the thermodynamic description, where the state of the system is given by a couple of state variables (the energy and the work coordinates), the Hamiltonian description is a highly detailed one. Passing from the Hamiltonian to the thermodynamic description, we thus loose an enormous amount of information about the state of the system considered. The set $x = (q,p)$ of coordinates and momenta of the constitutive particles of the system is the *microscopic configuration* of the system. Later, we will also use the terms *microscopic state* or *micro-state* to denote it.

It is the purpose of statistical mechanics, introduced in this section, to make the bridge between the microscopic and the thermodynamic descriptions. The following presentation is essentially the same as the one given by Gibbs in his original work [Gib02] with a few simplifications.

2.2.2 Gibb's canonical ensemble

In order to link the thermodynamic and the Hamiltonian description, Gibbs considered an ensemble of independent systems, all having the same Hamiltonian, but being in possibly different microscopic configurations [Gib02]. Such an ensemble is characterised by the density ρ over the phase space. The fraction of ensemble members having their microscopic configuration in a region of infinitesimal size $\mathrm{d}x$ around some x is given by $\rho(x)\mathrm{d}x$. The purpose of statistical mechanics is to find the density, or *distribution*, correctly describing a system in a given thermodynamic state.

The energy of the ensemble is given by:

$$E_\lambda[\rho] = \int \rho(x)\mathcal{H}_\lambda(x)\mathrm{d}x. \tag{2.46}$$

The distribution ρ evolves according to the Liouville equation[4]:

$$\frac{\partial \rho}{\partial t} - \sum_{i=1}^{r} \left(\frac{\partial \rho}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} - \frac{\partial \rho}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} \right) = 0. \tag{2.47}$$

Gibbs identified the macroscopic state of a thermodynamic system with the distribution ρ over its microscopic states. An equilibrium state is a distribution that is invariant under the Liouville equation (2.47). Such an invariant distribution is a function of the constants of motion only. Assuming that the energy is the only constant of motion, an invariant density has the form $\rho(x) \propto f(\mathcal{H}(x))$.

Two systems with Hamiltonian $\mathcal{H}_1$ and $\mathcal{H}_2$ in equilibrium at the same temperature can be combined to form an equilibrium system with Hamiltonian $\mathcal{H}_{\text{tot}} = \mathcal{H}_1 + \mathcal{H}_2$. Here, we make no statement about what the temperature is in terms of the microscopic configuration of the system; we assume the two systems to have the same temperature in order to make sure that the compound system is an equilibrium one. Since the two systems are not interacting, the joint distribution of their microscopic states is the product $\rho_{\text{tot}}(x_1, x_2) = \rho_1(x_1)\rho_2(x_2)$. Stating that this distribution is an equilibrium one is saying that it has the form $\rho_{\text{tot}}(x_1, x_2) \propto f(\mathcal{H}_{\text{tot}}(x_1, x_2))$. Hence

[4] Actually, Gibbs was the first to introduce this equation, but it is based on a theorem of Liouville

the function f has to satisfy the property

$$f(\mathcal{H}_1 + \mathcal{H}_2) = f(\mathcal{H}_1)f(\mathcal{H}_2). \tag{2.48}$$

The functions satisfying this relation are exponential functions $f(y) = \exp(ay)$. Thus, an equilibrium ensemble will be given by the so called *canonical distribution*:

$$\rho_{a,\lambda}(x) = \frac{\exp(a\mathcal{H}_\lambda(x))}{Z(a, \lambda)}, \tag{2.49}$$

where

$$Z(a, \lambda) = \int \exp(a\mathcal{H}_\lambda(x))\mathrm{d}x \tag{2.50}$$

is the normalization constant and is called the *partition function*.

The parameter a is an index of equilibrium in the sense that two equilibrium ensembles having the same a are in equilibrium with each other. They can be combined to form an equilibrium compound system which will have the same value for a. Hence, a is a function of β.

The parameters a and λ are state variables because they completely specify the equilibrium ensemble. The partition function (2.50) is a state function. However, it is multiplicative and not additive. In fact, if two systems are canonically distributed with parameter a and partition functions $Z_1(a)$ and $Z_2(a)$, then the system obtained by combining them is also canonically distributed with parameter a and partition function $Z_{\mathrm{tot}}(a) = Z_1(a)Z_2(a)$. In order to obtain an additive state function, one should consider the logarithm of the partition function:

$$R(a, \lambda) = \log Z(a, \lambda). \tag{2.51}$$

In order to understand the physical meaning of R, we have to analyze its variations with respect to small variations of its parameters a and λ:

$$\mathrm{d}R = \frac{\partial R}{\partial a}(a, \lambda)\mathrm{d}a + \frac{\partial R}{\partial \lambda}(a, \lambda)\mathrm{d}\lambda. \tag{2.52}$$

Using the definition of R (2.51) and of the partition function (2.50), we obtain for the derivative of R with respect to a:

$$\frac{\partial R}{\partial a} = \frac{1}{Z}\frac{\partial Z}{\partial a} = \int \mathcal{H}_\lambda(x)\frac{\exp\left(a\mathcal{H}_\lambda(x)\right)}{Z}\mathrm{d}x = \int \rho_{a,\lambda}(x)\mathcal{H}_\lambda(x)\mathrm{d}x = E(a, \lambda), \tag{2.53}$$

which is just the energy of the ensemble.

The derivative of R with respect to the control parameter reads:

$$\frac{\partial R}{\partial \lambda} = a \int \rho_{a,\lambda}(x) \frac{\partial \mathcal{H}_\lambda}{\partial \lambda}(x) \mathrm{d}x. \tag{2.54}$$

To interpret this quantity, consider the following process. The control parameter λ is driven from some initial value λ_0 to some final value λ_1. Furthermore, we assume that during the whole process, *the ensemble remains in equilibrium* with the parameter a being held constant. Hence, all along the process, the ensemble is canonically distributed with the parameter a. We now wish to calculate the value of the work performed during such a process. The work performed on one ensemble member in micro-state x when λ is varied by an infinitesimally small amount $\mathrm{d}\lambda$ is given by:

$$\delta W_\lambda(x) = \frac{\partial \mathcal{H}_\lambda}{\partial \lambda}(x) \mathrm{d}\lambda. \tag{2.55}$$

Its ensemble average is:

$$\langle \delta W_\lambda \rangle = \left(\int \rho_{a,\lambda}(x) \frac{\partial \mathcal{H}_\lambda}{\partial \lambda}(x) \mathrm{d}x \right) \mathrm{d}\lambda = \frac{1}{a} \frac{\partial R}{\partial \lambda}(a, \lambda) \mathrm{d}\lambda. \tag{2.56}$$

As a consequence, the average work performed during the process is given by:

$$\langle W \rangle = \int_{\lambda_0}^{\lambda_1} \langle \delta W_\lambda \rangle = \frac{1}{a} \int_{\lambda_0}^{\lambda_1} \frac{\partial R}{\partial \lambda}(a, \lambda) \mathrm{d}\lambda = \frac{\Delta R}{a}, \tag{2.57}$$

where $\Delta R = R(a, \lambda_1) - R(a, \lambda_0)$. The process just described is *reversible.* In fact, if we drive the control parameter from the final value λ_1 back to the initial value λ_0 and we assume that the ensemble is canonically distributed with the parameter a being held constant all along the process, then we would be able to extract the very same amount $\langle W \rangle$ of work from the system. The quantity R/a gives the reversible work to perform in a transformation where a is held constant. Hence, we have:

$$\frac{1}{a} \frac{\partial R}{\partial \lambda} = -P = -\frac{A}{\beta}, \tag{2.58}$$

yielding an expression for the pressure in terms of the dynamical (Hamiltonian) properties of the system.

The differential of R (2.52) thus becomes:

$$\mathrm{d}R = E \, \mathrm{d}a - \frac{a}{\beta} A \, \mathrm{d}\lambda. \tag{2.59}$$

We arrive at the conclusion that *the dependence of R in $-a$ and λ is the same as the dependence of ψ in β and λ.* In other words, if we assume $\beta = -a$ and

$\psi(\beta, \lambda) = R(-a, \lambda)$, we recover the results of classical (reversible) thermodynamics derived in the previous section. *The canonical ensemble provides a microscopic model for the reversible isothermal thermodynamics.*

2.2.3 Canonical statistical mechanics

A system in equilibrium with a heat bath at inverse temperature β is described by a *canonical ensemble*. It's microscopic configurations are distributed according to the canonical distribution:

$$\rho_{\beta,\lambda}(x) = \frac{\exp\left(-\beta \mathcal{H}_\lambda(x)\right)}{Z(\beta, \lambda)}, \tag{2.60}$$

where $Z(\beta, \lambda)$ is the *partition function* and is given by normalization:

$$Z(\beta, \lambda) = \int \exp\left(-\beta \mathcal{H}_\lambda(x)\right) \mathrm{d}x. \tag{2.61}$$

The *equilibrium Massieu function* is the logarithm of the partition function:

$$\psi(\beta, \lambda) = \log Z(\beta, \lambda). \tag{2.62}$$

The free energy is related to the Massieu function and hence to the partition function as:

$$F(T, \lambda) = -T\psi(\beta, \lambda) = -T \log Z(\beta, \lambda). \tag{2.63}$$

The ensemble average, or equilibrium value of the energy is given by the thermodynamic relation (2.35):

$$E(\beta, \lambda) = -\frac{\partial \psi}{\partial \beta}(\beta, \lambda) = -\frac{1}{Z(\beta, \lambda)} \frac{\partial Z}{\partial \beta}(\beta, \lambda). \tag{2.64}$$

Using the thermodynamic definition of the Massieu function, equation (2.34) we can find an explicit expression for the entropy of the system:

$$S(\beta, \lambda) = \psi(\beta, \lambda) + \beta E(\beta, \lambda). \tag{2.65}$$

Using the definition of ψ, the normalization of $\rho_{\beta,\lambda}$ and of the definition of the average energy, we obtain:

$$S(\beta, \lambda) = \log Z + \int \rho_{\beta,\lambda}(x) \beta \mathcal{H}_\lambda(x) \mathrm{d}x = -\int \rho_{\beta,\lambda}(x) \log \frac{e^{-\beta \mathcal{H}_\lambda(x)}}{Z} \mathrm{d}x. \tag{2.66}$$

Hence, the final expression of the canonical entropy also called *Gibbs entropy* is:

$$S(\beta, \lambda) = -\int \rho_{\beta,\lambda}(x) \log \rho_{\beta,\lambda}(x) \mathrm{d}x. \tag{2.67}$$

This is, in a slightly modified way, the derivation Gibbs presented in his celebrated *Elementary principle in Statistical Mechanics* [Gib02].

With the help of statistical mechanics, it is possible to link the temperature to the motion of the particles constituting the system. Assume that the system has r degrees of freedom (typically, $r = 3N$ where N is the number of particles), the Hamiltonian is then:

$$\mathcal{H}_\lambda(q,p) = K(p) + V_\lambda(q), \tag{2.68}$$

where $K(p) = \sum_{i=1}^{r} \frac{p_i^2}{2m}$ is the kinetic energy of the system (m is the mass of the particles) and $V_\lambda(q)$ is the potential energy of the system. The partition function factorizes into:

$$Z(\beta,\lambda) = \int e^{-\beta\mathcal{H}_\lambda(q,p)}\mathrm{d}q\mathrm{d}p = \int e^{-\beta K(p)}\mathrm{d}p \int e^{-\beta V_\lambda(q)}\mathrm{d}q = Z_{\mathrm{kin}}(\beta)Z_{\mathrm{pot}}(\beta,\lambda), \tag{2.69}$$

where Z_{kin} is the part due to the kinetic energy and Z_{pot} due to the potential energy. The kinetic part factorizes further:

$$\begin{aligned} Z_{\mathrm{kin}}(\beta) = \int e^{-\beta K(p)}\mathrm{d}p = \int e^{-\beta\sum_{i=1}^{r}\frac{p_i^2}{2m}}\mathrm{d}p_1\ldots\mathrm{d}p_r = \\ \int e^{-\beta\frac{p_1^2}{2m}}\mathrm{d}p_1\ldots\int e^{-\beta\frac{p_r^2}{2m}}\mathrm{d}p_r = \left(\int e^{-\beta\frac{p_1^2}{2m}}\mathrm{d}p_1\right)^r = Z_{\mathrm{kin}}^1(\beta)^r \end{aligned} \tag{2.70}$$

The canonical distribution factorizes as well:

$$\rho_{\beta,\lambda}(q,p) = \rho_{\beta,\lambda}^{\mathrm{pot}}(q)\rho_\beta^1(p_1)\ldots\rho_\beta^1(p_r), \tag{2.71}$$

where

$$\rho_{\beta,\lambda}^{\mathrm{pot}}(q) = \frac{e^{-\beta V_\lambda(q)}}{Z_{\mathrm{pot}}(\beta,\lambda)}, \tag{2.72}$$

and

$$\rho_\beta^1(p_1) = \frac{e^{-\beta\frac{p_1^2}{2m}}}{Z_{\mathrm{kin}}^1(\beta)}. \tag{2.73}$$

Note that $\rho_\beta^1(p_1)$ is a Gaussian distribution with 0 mean and a variance equal to

m/β. The average kinetic energy of the system is given by:

$$\langle K\rangle = \int \rho_{\beta,\lambda}(q,p)K(p)\mathrm{d}q\mathrm{d}p = \int \rho^{\mathrm{pot}}_{\beta,\lambda}(q)\rho^{1}_{\beta}(p_1)\dots\rho^{1}_{\beta}(p_r)\left(\sum_{i=1}^{r}\frac{p_i^2}{2m}\right)\mathrm{d}q\,\mathrm{d}p_1\dots\mathrm{d}p_r$$
$$= \frac{1}{2m}\sum_{i=1}^{r}\left(\int \rho^{1}_{\beta}(p_i)p_i^2\,\mathrm{d}p_i\right) = \frac{r}{2m}\underbrace{\int \rho^{1}_{\beta}(p_1)p_1^2\,\mathrm{d}p_1}_{m/\beta}. \tag{2.74}$$

The last integral is the variance of the distribution $\rho^1_\beta(p_1)$ which is equal to m/β as we just noticed. Hence, the average kinetic energy is equal to:

$$\langle K\rangle = \frac{r}{2\beta} = \frac{r}{2}T. \tag{2.75}$$

The temperature is thus twice the average kinetic energy per degree of freedom.

2.3 Maxwell's Demon and Szilard's engine

The microscopic description of matter suggests that it is possible to manipulate each atom separately. Doing so, however, would allow to reduce the entropy of an adiabatic system. The first one to realize this was James Clerk Maxwell. In his *Mechanical theory of heat*, he describes a simple Gedankenexperiment. Imagine a container containing a gas at some temperature T. The container is split into two halves by a separating wall. On this wall, there is a small trap door. The trap door is operated by a mechanism which opens it if:

- A particle faster than average passes from the left to the right half of the container,
- a particle slower than average passes from the right to the left half of the container.

Otherwise, the trap door is kept closed. As time passes, there is a net energy flow from the left to the right half of the container. The average velocity, and thus kinetic energy, of the particles in the left half decreases while in the right half, it increases. Since the temperature is proportional to the average kinetic energy, see eq. (2.75), the temperature of the left half of the container decreases while the temperature of the right half increases. The energy flows from the colder to the hotter half. The net result of this process is a decrease in the entropy of the gas inside the container.

One can imagine other ways to decrease the entropy of an adiabatic system in a similar way. For instance, it is possible to compress a gas without performing work

Figure 2.1: Maxwell's Demon. Courtesy of Anne Gärtner.

by waiting for the precise moment when there is no particle in front of the piston and by moving the piston at such moments only.

Such a process should not be possible without being accompanied by the production of an amount of entropy at least equal to the entropy reduction in the container.

In his original book, Maxwell spoke of “an intelligent being” acting on the door [Max71]. William Thomson (Lord Kelvin) called this being a “demon” [Tho74, Tho79]. In the contemporary literature, it is usually called “Maxwell’s Demon”, even though some computer with a receptor would do the job perfectly [LR02].

In a paper published in 1929, the Hungarian physicist Leó Szilard gave a quantitative relation between the amount of entropy reduction and the information needed to achieve this reduction in a simple toy model [Szi29]. Let a closed volume containing a gas of one particle be in contact with a heat bath a temperature T. At some point, a partition in inserted in the middle of the container. Then, the operator looks whether the particle is on the right side or on the left side of the partition. Note that each of these events occurs with probability 1/2. If the particle is found on the right side, then the operator performs an isothermal expansion to the left, else the isothermal expansion is performed to the right. Since the volume of the gas is doubled during this expansion, it yields an amount of work

$$W = T \log 2. \tag{2.76}$$

At the end of the expansion, the system is in the same state as at the very beginning of the process. Hence, this process is an isothermal cycle whose net effect is to furnish some positive amount of work. Since this is a cycle, the reversible work associated to it is 0 so that the dissipated work (and hence the entropy production) is negative.

These Gedankenexperiments show that one has to be very careful when mixing the microscopic and the thermodynamic scale. When one only specifies the thermodynamic state of a system, the information about its microscopic state is very limited. These Gedankenexperiments suggest that the use of more information about the microscopic state allows one to change the thermodynamic state in a way that could lead to a reduction of the entropy. This relation between information and entropy will be made quantitative in the next chapter.

3 Stochastic thermodynamics

Stochastic thermodynamics is the simplest theory describing non-equilibrium isothermal processes. It allows to explicitly compute the entropy produced during some arbitrary processes and it extends equilibrium statistical mechanics to non-equilibrium situations. Moreover, it provides an ideal framework to formalize Maxwell's Demon and the interplay between information and irreversibility.

Stochastic thermodynamics applies to "small" systems, where the energies involved are of the order of the temperature. Examples of applications include colloidal particles, bio-polymers such as DNA, RNA or proteins, molecular motors or single electron transistors, electrical circuits at low intensities. These system are in contact with an environment (water or air) and they are too small to modify its temperature. Hence the environment acts as a heat bath in a rather good approximation.

The presentation given here emphasizes the role of information from the beginning. It is shown that when the system is not in equilibrium, i.e. not described by a canonical distribution, then we have *more information* about its microscopic state than just the value of its thermodynamic state variables. The non equilibrium dynamics of the macroscopic state of the system is then the dynamics of our information about its microscopic state. The expression for the entropy production given by stochastic thermodynamics is shown to be equal to the loss of information about the microscopic state of the system. The problem of Maxwell's Demon is then formalized easily. Since the amount of information we have about the microscopic state of the system is a characteristic of its macroscopic (equilibrium or not) state, it is natural that by changing this amount of information, we can change the macroscopic state.

A recent review of stochastic thermodynamics with applications to biological systems can be found here [Sei12]. Short papers presenting the formalism include [GS97, Cro98, Sei08, VdBE10, EVdB10, Esp12]. The problem of Maxwell's Demon is formalized in [SU10, HP11b, HP11a, SU12b, ES12]. The interplay between information processing and entropy production is developed more generally in [SU09, SU12a, SU13].

3.1 A short introduction to information theory

This presentation naturally starts with an introduction to information theory. This section aims at introducing the mathematical tools used to quantify information. The central concept of information theory is called (*information theoretic*) entropy and was first introduced by Claude Shannon [Sha48]. At the end of the section, we show how information theory can be applied to equilibrium statistical mechanics. A good and short introduction to information theory can be found in [Khi57] and a detailed presentation in [CT06].

3.1.1 Shannon entropy

Let X be a random variable with N possible outcomes $\{x_1, \cdots, x_N\}$. Let p_X be the distribution of X:

$$p_X(x_i) = P(X = x_i) \geq 0. \tag{3.1}$$

The probability distribution p_X satisfies normalization:

$$\sum_{i=1}^{N} p_X(x_i) = 1. \tag{3.2}$$

We would like to find a number $H(X)$ quantifying our uncertainty about the outcome of X. Any reasonable measure of the uncertainty should satisfy a small number of basic properties:

1. $H(X)$ should depend on X only though its distribution p_X: $H(X) = H(p_1, \cdots, p_N)$ where we have set $p_i = p_X(x_i)$. In fact, the uncertainty about the outcome of a random variable does not depend on whether it is a ball drawn from an urn or the result of the rolling of a die. It only depends on the probabilities of the different outcomes.

2. Our uncertainty about the outcome of X is maximum if all the possible outcomes are equally probable. Hence, the function $H(p_1, \cdots, p_N)$ should be maximum for $p_i = 1/N$ for all i.

3. If X and Y are two random variables, our uncertainty about the outcome of the couple (X, Y) is the sum of our uncertainty about the outcome of X and of our average uncertainty about the outcome of Y knowing the outcome of X:

$$H(X, Y) = H(X) + \sum_{i=1}^{N} p_X(x_i) H(Y|X = x_i). \tag{3.3}$$

   Here, $H(Y|X = x_i)$ is the uncertainty about the outcome of Y that remains when we know that $X = x_i$. In other words, it is the uncertainty associated with the distribution of Y conditionned on the fact that $X = x_i$. In particular, if X and Y are independent, then

$$H(X, Y) = H(X) + H(Y), \tag{3.4}$$

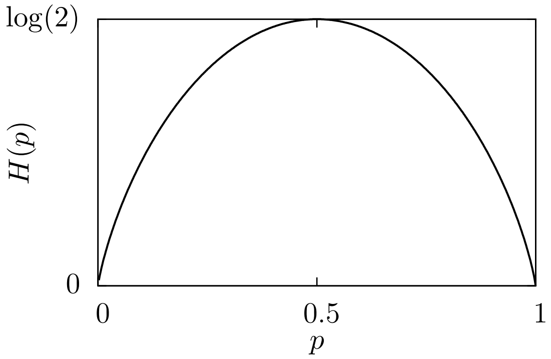

Figure 3.1: The Shannon entropy $H(p) = -p \log p - (1-p)\log(1-p)$ of a binary random variable where one event occurs with probability p and the other with probability $1-p$.

meaning that our uncertainty about the ouctome of the couple (X, Y) is the sum of our uncertainties about the individual outcomes of X and Y.

4. Adding one event of zero probability does not change our uncertainty:

$$H(p_1, \cdots, p_N, 0) = H(p_1, \cdots, p_N). \tag{3.5}$$

It can be shown (see for instance [Khi57]) that any function satisfying these four properties, also called *Khinchin's axioms*, has the form:

$$H(X) = -\sum_{i=1}^{N} p_X(x_i) \log p_X(x_i), \tag{3.6}$$

with the convention $p \log p = 0$ if $p = 0$. This quantity was first introduced by Claude Shannon in 1948 [Sha48]. Shannon called this quantity "entropy" in analogy to the expression for the entropy appearing in statistical mechanics, equation (2.67). In the following, we will call $H(X)$ the "Shannon entropy of X". The Shannon entropy of a random variable with two possible outcomes, one of them occurring with probability p and the other with probability $1-p$ reduces to:

$$H(p) = -p \log p - (1-p)\log(1-p). \tag{3.7}$$

This quantity is plotted as a function of p on figure 3.1. The Shannon entropy of X is zero if and only if the outcome of X is certain, i.e. if $p_X(x_k) = 1$ for a certain k and $p_X(x_i) = 0$ for all $i \neq k$. On the other hand it is maximum if X is equally distributed, i.e. if $p_X(x_i) = 1/N$ for every i. In this case, $H(X) = \log N$.

3.1.2 Conditional entropy and mutual information

Let Y be a second random variable with M possible outcomes $y_1, \cdots, y_M$. Let $p_{X,Y}$ be the joint distribution of X and Y:

$$p_{X,Y}(x_i, y_j) = P(X = x_i \text{ and } Y = y_j). \tag{3.8}$$

The probability $p_X(x_i)$ that $X = x_i$ independently of the outcome of Y is given by the *marginal distribution* of X:

$$p_X(x_i) = \sum_{j=1}^{M} p_{X,Y}(x_i, y_j). \tag{3.9}$$

The marginal distribution p_Y of Y is given by a similar relation.

The two random variables X and Y are said to *independent* if their joint distribution is the product of the marginal distributions:

$$p_{X,Y}(x_i, y_j) = p_X(x_i)p_Y(y_j). \tag{3.10}$$

The result of the rolling of two dice is a simple example of two independent random variables. The result of the first die does not influence the result of the second die. For instance, the probability that the result of the second die is a six is the same, whether the result of the first die is a six or not. In other words, the knowing the result of one of the dice does not give any information about the result of the other die.

If equation (3.10) does not hold, the random variables X and Y are said to be *depend on each other.* In this case, the outcome of one of the random variable "influences" the outcome of the other. Let us illustrate this point with a simple example. Consider an urn containing k black balls and l white balls. We randomly draw two balls from this urn. Let X = "color of the second ball" and Y = "color of the first ball". If Y = "black", then the probability that X = "black" is:

$$P(X = \text{“black”}|Y = \text{“black”}) = \frac{k-1}{k+l-1}, \tag{3.11}$$

whereas if Y = "white", the probability that X = "black" is given by:

$$P(X = \text{“black”}|Y = \text{“white”}) = \frac{k}{k+l-1}. \tag{3.12}$$

On the other hand, the overall probability that X = "black", i.e. the marginal probability of this event is:

$$P(X = \text{“black”}) = \frac{k}{k+l}. \tag{3.13}$$

One can easily verify that:

$$\begin{aligned} P(X = \text{“black”}) = P(X = \text{“black”}|Y = \text{“black”})P(Y = \text{“black”})+ \\ P(X = \text{“black”}|Y = \text{“white”})P(Y = \text{“white”}). \end{aligned} \tag{3.14}$$

In this example, we see that the probability of the outcome of X is not the same whether we know the outcome of Y or not. And if we know the outcome of Y, the probability of X = “black” is not the same depending on the outcome of Y. The outcome of Y does not influence the outcome of X in the sense that it acts like a “force” on it but rather in the sense that it restricts the possibilities for the outcome of X. In other words, knowing the outcome of Y *provides some information* about the outcome of X.

The notation $P(X = x_i|Y = y_j)$ introduced above means “the probability that $X = x_i$ knowing that $Y = y_j$”. In the following we will use the different notations:

$$P(X = x_i|Y = y_j) = p_{X|Y}(x_i|y_j) = p_{X|Y=y_j}(x_i) \tag{3.15}$$

to denote the probability that $X = x_i$ conditioned on $Y = y_j$. The conditional distributions $p_{Y|X}$ and $p_{X|Y}$ are related to the joint distribution $p_{X,Y}$ via:

$$p_{X,Y}(x_i, y_j) = p_{X|Y}(x_i|y_j)p_Y(y_j) = p_{Y|X}(y_j|x_i)p_X(x_i). \tag{3.16}$$

When reordered in the following way:

$$p_{X|Y}(x_i|y_j) = \frac{p_{Y|X}(y_j|x_i)p_X(x_i)}{p_Y(y_j)} \tag{3.17}$$

this equality is called *Bayes theorem.* By definition, the conditional distributions are probability distributions and are thus normalized:

$$\sum_{i=1}^{N} p_{X|Y}(x_i|y_j) = 1 \tag{3.18}$$

for any j.

We have seen that if one knows that $Y = y_j$, then the probability that $X = x_i$ changes from $p_X(x_i)$ to $p_{X|Y}(x_i|y_j)$. As a consequence, our uncertainty about the outcome of X changes. Before we know the outcome of Y, the uncertainty we have about the outcome of X is quantified by its Shannon entropy $H(X)$ given by equation (3.6). When we know that $Y = y_j$, the uncertainty we have about the outcome of X

is quantified by the Shannon entropy of the conditioned distribution $p_{X|Y=y_j}$:

$$H(X|Y=y_j) = -\sum_{i=1}^{N} p_{X|Y}(x_i|y_j)\log p_{X|Y}(x_i|y_j). \tag{3.19}$$

The entropy of X conditioned on Y is the average over the outcomes of Y of the quantity (3.19) above:

$$H(X|Y) = \sum_j p_Y(y_j) H(X|Y=y_j). \tag{3.20}$$

This quantity is smaller than the unconditioned entropy of X:

$$H(X|Y) \leq H(X) \tag{3.21}$$

with equality if and only if X and Y are independent. We will show this in the next paragraph. The inequality (3.21) means that on average, knowing the outcome of Y reduces our uncertainty over the outcome of X. We interpret this reduction in uncertainty as *information* and define the *mutual information* between X and Y as the average reduction in uncertainty about the outcome of X upon knowing the outcome of Y:

$$I(X,Y) = H(X) - H(X|Y) \geq 0. \tag{3.22}$$

The mutual information is symmetric: $I(X,Y) = I(Y,X)$, i.e. X possesses as much information about Y as Y possesses about X. We are going to show this in the next paragraph as well.

3.1.3 Relative entropy

A concept that will prove extremely useful in the following is the concept of *relative entropy* or Kullback-Leibler divergence from one distribution to another distribution. Let p and q be two distributions over X. The relative entropy from p to q is defined as [CT06]:

$$D[p\|q] = \sum_{i=1}^{N} p(x_i)\log\frac{p(x_i)}{q(x_i)}. \tag{3.23}$$

This quantity is defined if and only if $q(x_i) \neq 0$ for every i for which $p(x_i) \neq 0$. It is non negative and is zero if and only if the two distributions are identical, i.e. if $p(x_i) = q(x_i)$ for every i. To prove this we use of the fact that $\log y \leq y - 1$ with equality if and only if $y = 1$. Hence we have:

$$\log\frac{p(x_i)}{q(x_i)} = -\log\frac{q(x_i)}{p(x_i)} \geq 1 - \frac{q(x_i)}{p(x_i)}. \tag{3.24}$$

Averaging with respect to p yields:

$$D[p\|q] \geq \sum_{i=0}^{N} p(x_i)\left(1 - \frac{q(x_i)}{p(x_i)}\right) = \sum_{i=1}^{N} p(x_i) - \sum_{i=1}^{N} q(x_i) = 0, \tag{3.25}$$

where we have used the fact that p and q are normalized, i.e. $\sum_i p(x_i) = \sum_i q(x_i) = 1$. There is equality in the inequality (3.25) if and only if there is equality in the inequality (3.24) for every i, which happens if and only if $p(x_i) = q(x_i)$ for every i. Hence we have proved that the relative entropy (3.23) is non negative and is zero if and only if p and q are identical.

The relative entropy $D[p\|q]$ can be interpreted as a measure of the amount of information we miss when we think that X is distributed according to q when it is distributed according to p in reality. It can also be interpreted as the amount of information that we gain when we learn that X is distributed according to p when we thought that it was distributed according to q. For instance, consider two correlated random variables X and Y as in the previous paragraph. We are interested in X and we assume that we are able to see a realization of Y. Before we know the outcome of Y, we can only tell that X is distributed according to its marginal distribution $q = p_X$. After we see the outcome of Y, for instance $Y = y_j$, we can say that X is distributed according to $p_{X|Y=y_j}$. The amount of information we gain about X seeing $Y = y_j$ is then $D[p_{X|Y=y_j}\|p_X]$. This quantity averaged over the outcome of Y is equal to:

$$\begin{aligned}
&\sum_{j=1}^{M} p_Y(y_j) D[p_{X|Y=y_j}\|p_X] = \\
&\sum_{j=1}^{M} p_Y(y_j) \sum_{i=1}^{N} p_{X|Y}(x_i|y_j) \log p_{X|Y}(x_i|y_j) - \sum_{i=1}^{N}\left(\sum_{j=1}^{M} p_{X|Y}(x_i|y_j) p_Y(y_j)\right) \log p_X(x_i) \\
&\qquad = H(X) - H(X|Y) = I(X,Y).
\end{aligned} \tag{3.26}$$

This is precisely the mutual information between X and Y as defined by equation (3.22). From this expression, we see that it is a positive quantity, since it is the average of a relative entropy, which is itself a positive quantity. Moreover, it is zero if and only if $p_{X|Y=y_j}$ is identical to p_X for every j, hence if and only if X and Y are independent. To see that the mutual information is symmetric, we use equation

(3.26) above together with equation (3.16):

$$\begin{aligned} I(X,Y) &= \sum_{j=1}^{M}\sum_{i=1}^{N} p_{X|Y}(x_i|y_j)p_Y(y_i)\log\frac{p_{X|Y}(x_i|y_j)}{p_X(x_i)} \\ &= \sum_{j=1}^{M}\sum_{i=1}^{N} p_{X,Y}(x_i,y_j)\log\frac{p_{X,Y}(x_i,y_j)}{p_X(x_i)p_Y(y_j)} = D[p_{X,Y}\|p_Xp_Y] \end{aligned} \tag{3.27}$$

The last expression is symmetric with respect to the exchange of X and Y, thus $I(X,Y) = I(Y,X)$. This expression gives us a new interpretation of the mutual information between X and Y. Since $I(X,Y)$ is the relative entropy from the joint distribution $p_{X,Y}$ to the product p_Xp_Y of the marginal distributions, it is the amount of information we would miss if we thought that X and Y were independent. In fact, if we thought X and Y were independent we would think that they are jointly distributed according to the product p_Xp_Y instead of $p_{X,Y}$.

3.1.4 Continuous random variables

These concepts can be generalized for continuous random variables. Let X be a real valued random variable with a probability density function, or distribution ρ_X. For two real numbers a and b with $a \leq b$, the probability that X lies between a and b is obtained by integrating the distribution from a to b:

$$P(X \in [a,b]) = \int_a^b \rho_X(x)\mathrm{d}x. \tag{3.28}$$

More generally, if X is a multidimensional random vector, the probability that X lies in a given set A is given by:

$$P(X \in A) = \int_A \rho_X(x)\mathrm{d}x. \tag{3.29}$$

The distribution should be normalized:

$$\int \rho_X(x)\mathrm{d}x = 1. \tag{3.30}$$

Here and every time the integration domain is not specified it should be understood as being the whole domain.

Differential and relative entropy

Equation (4.1) can be generalized in the following way:

$$H[\rho_X] = -\int \rho_X(x)\log\rho_X(x)\mathrm{d}x, \tag{3.31}$$

The quantity $H[\rho_X]$ is called *differential entropy* in the literature. It is the analogue of the Shannon entropy for continuous random variables. However, it misses some important properties of the Shannon entropy. The differential entropy is not invariant under coordinate transformation. Moreover it can take any real value, even negative ones. However, differential entropy still measures the degree of randomness in the sense that a random variable having a greater differential entropy is “more random” than a variable having a lower differential entropy. The differential entropy retains another crucial feature of its discrete counterpart, namely that it decreases upon conditioning on a second random variable. This decrease still gives the mutual information between the two random variables. As we will see, the mutual information is invariant under coordinate transformation.

The relative entropy can be generalized to continuous variables as well. Let ρ_1 and ρ_2 be two probability density functions. The relative entropy from ρ_1 to ρ_2 is defined as:

$$D[\rho_1\|\rho_2] = \int \rho_1(x)\log\frac{\rho_1(x)}{\rho_2(x)}\mathrm{d}x. \tag{3.32}$$

As in the discrete case, the relative entropy is defined if and only if $\rho_2(x) \neq 0$ for every x for which $\rho_1(x) \neq 0$. Unlike the differential entropy (3.31), the relative entropy (3.32) is invariant under coordinate transformation. Moreover, as its discrete counterpart, it is non negative and is zero if and only if ρ_1 and ρ_2 are identical[1].

Mutual information

Let X and Y be two continuous random variables. Let $\rho_{X,Y}$ be their joint probability density function. In other words:

$$P(X \in A \text{ and } Y \in B) = \int_A \left(\int_B \rho_{X,Y}(x,y)\mathrm{d}y\right)\mathrm{d}x. \tag{3.33}$$

The marginal distribution of X is obtained from $\rho_{X,Y}$ by integrating out y:

$$\rho_X(x) = \int \rho_{Y,X}(x,y)\mathrm{d}y. \tag{3.34}$$

[1] Rigorously speaking, $D[\rho_1\|\rho_2] = 0$ if and only if ρ_1 and ρ_2 are equal almost everywhere.

The marginal distribution of Y is obtained similarly, by integrating out x. As in the discrete case, the distribution of X conditioned on $Y = y$ is given by Bayes' theorem:

$$\rho_{X|Y=y}(x) = \rho_{X|Y}(x|y) = \frac{\rho_{X,Y}(x,y)}{\rho_Y(y)} = \frac{\rho_{Y|X}(y|x)\rho_X(x)}{\rho_Y(x)}. \tag{3.35}$$

The differential entropy of X conditioned on $Y = y$ is then:

$$H[\rho_{X|Y=y}] = -\int \rho_{X|Y}(x|y)\log\rho_{X|Y}(x|y)\mathrm{d}x. \tag{3.36}$$

The entropy of X conditioned on Y, which we denote $H[\rho_{X|Y}]$, is the average over Y of the quantity above:

$$H[\rho_{X|Y}] = -\int \rho_Y(y)H[\rho_{X|Y=y}]\mathrm{d}y. \tag{3.37}$$

As already mentioned, this quantity is smaller than the marginal entropy $H[\rho_X]$ of X and the difference between the two is the mutual information between X and Y:

$$I(X,Y) = H[\rho_X] - H[\rho_{X|Y}] \geq 0. \tag{3.38}$$

To see that this expression is non negative, we note that the continuous equivalent of equation (3.26) holds:

$$I(X,Y) = \int \rho_Y(y)D[\rho_{X|Y=y}\|\rho_X]\mathrm{d}y. \tag{3.39}$$

Moreover, using the same reasoning as for discrete random variables, we find:

$$I(X,Y) = D[\rho_{X,Y}\|\rho_X\rho_Y]. \tag{3.40}$$

As a consequence, the mutual information between two continuous random variables is invariant under coordinate transformations. Moreover, it is symmetric, non negative and it is zero if and only if X and Y are independent.

3.1.5 Application to equilibrium statistical mechanics

These information theoretic concepts can be used to reinterpret equilibrium statistical mechanics. Consider a system in equilibrium with a heat bath at inverse temperature β. We assume that we know its Hamiltonian $\mathcal{H}_\lambda$. The thermodynamic state of the system is given by the two state variables β and λ. We assume that we know the thermodynamics of the system, i.e. we know the functional form of the state functions $E(\beta,\lambda)$ and $S(\beta,\lambda)$. What can we say about the microscopic configuration of system? The system is composed of particles that all have a unique position and a momentum. What can we say about theses positions and momenta? This question can be rephrased the following way: What uncertainty do we have about the positions and momenta of all the particles constituting our system (i.e. about its microscopic configuration) when we only know the thermodynamic state variables β and λ?

To us, the microscopic configuration of the system is random. The question is then, what is its distribution? Obviously any distribution ρ describing the state of the system should be compatible with the information we have. In particular, it should predict that the system has the energy $E(\beta, \lambda)$:

$$E_\lambda[\rho] = \int \rho(x)\mathcal{H}_\lambda(x)\mathrm{d}x = E(\beta, \lambda). \tag{3.41}$$

As soon as we assume that the microscopic states of the system are distributed according to ρ, we implicitly assume that our uncertainty about the microscopic state of the system is:

$$H[\rho] = -\int \rho(x) \log \rho(x)\mathrm{d}x, \tag{3.42}$$

the Shannon (differential) entropy of ρ. In a paper published in 1957, an American physicist named E. T. Jaynes proposed that we should choose the distribution maximizing the entropy (3.42) under the constraint (3.41) [Jay57]. His argument is that if we did not do so, we would implicitly make the assumption that we have more information about the microscopic state of the system than just the average energy.

It turns out that the distribution maximizing the entropy for a given value of the average energy *is the canonical distribution* given by equation (2.60):

$$\rho_{\beta,\lambda}(x) = \frac{\exp(-\beta\mathcal{H}_\lambda(x))}{Z(\beta, \lambda)}, \tag{3.43}$$

where $Z(\beta, \lambda) = \int \exp(-\beta\mathcal{H}_\lambda(x))\mathrm{d}x$ is the partition function. It ensures that the distributions $\rho_{\beta,\lambda}$ is normalized. Moreover, as we saw in the previous chapter, the entropy $S(\beta, \lambda)$ of the system is linked to the distribution $\rho_{\beta,\lambda}$ through the relation (2.67):

$$S(\beta, \lambda) = -\int \rho_{\beta,\lambda}(x) \log \rho_{\beta,\lambda}(x)\mathrm{d}x = H[\rho_{\beta,\lambda}]. \tag{3.44}$$

In other words, the Gibbs entropy of the system as a function of its thermodynamic state (β, λ) is just the Shannon entropy of the canonical distribution. Hence, the thermodynamic entropy measures the information that we miss about the microscopic configuration of a system when we only know its thermodynamic state.

The easiest way to prove that the canonical distribution maximizes the Shannon entropy for a given value of the energy is to compute the relative entropy $D[\rho\|\rho_{\beta,\lambda}]$ of an arbitrary distribution ρ to a canonical distribution $\rho_{\beta,\lambda}$. A short calculation yields:

$$D[\rho \| \rho_{\beta,\lambda}] = \beta \left(E_\lambda[\rho] - E(\beta, \lambda) \right) - \left(H[\rho] - S(\beta, \lambda) \right). \tag{3.45}$$

Since the relative entropy is positive, it is clear from the expression above that if $E_\lambda[\rho] = E(\beta, \lambda)$, then $H[\rho] \leq S(\beta, \lambda)$. Moreover, there is equality if and only if $D[\rho \| \rho_{\beta,\lambda}] = 0$, i.e. if ρ is identical to $\rho_{\beta,\lambda}$. Furthermore, from equation (3.45) above, we see that if $H[\rho] = S(\beta, \lambda)$, then $E_\lambda[\rho] \geq E(\beta, \lambda)$ with equality if and only if ρ and $\rho_{\beta,\lambda}$ are identical. In other words, the canonical distribution minimizes the energy for a given Shannon entropy.

3.2 Stochastic thermodynamics

3.2.1 Motivation

Let us now return to our original problem, namely the production of entropy and the dissipation of work in an isothermal process. As usual, consider a system in equilibrium with a heat bath at inverse temperature β. An isothermal process is performed on the system by changing the control parameter according to some time dependence $\lambda(t)$. Let λ_0 be the initial value of the control parameter and λ_1 its final value. In order to carry out such a process, we have to perform an amount of work W at least equal to the change in free energy:

$$W \geq F(\lambda_1) - F(\lambda_0) = \Delta F. \tag{3.46}$$

From now on, we drop the β dependence in all the state functions since we are not going to consider any variations of β. While the minimum amount of work we have to perform only depends on the initial and final values λ_0 and λ_1 of the control parameter, the amount of extra work $W_\mathrm{d} = W - \Delta F$ we have to put in actually depends on how we perform the process, i.e. on the time dependence of λ. Furthermore, the total amount of entropy produced by this process is $\Delta S_\mathrm{tot} = \beta W_\mathrm{d}$. What can we say about the microscopic state of the system during the process? Can we relate the amount of entropy production to the microscopic state of the system?

At the beginning of the process, the microscopic state of the system is distributed according to the canonical distribution[2] ρ_{λ_0} given by equation (3.43). At the end of the process, the system is also in equilibrium with the heat bath, hence its microscopic state is distributed according to the canonical distribution ρ_{λ_1}. But what happens in between? What distribution should we assign to the microscopic state of the system at some intermediate time t? Should we assign to it the canonical distribution $\rho_{\lambda(t)}$? If the microscopic state of the system was distributed according to the canonical

[2] We also drop the β dependence of the canonical distribution. From now on, ρ_λ will always denote the canonical distribution with inverse temperature β and control parameter λ given by equation (3.43) or (2.60).

distribution $\rho_{\lambda(t)}$ all along the process, then the process would be reversible and the work performed would be equal to ΔF. Hence, in general, we should assume that the microscopic state of the system is distributed according to some $\rho(x,t)$ at time t, which is different from the equilibrium distribution $\rho_{\lambda(t)}(x)$.

This leads us to the two following important remarks: (i) Irreversibility and entropy production do not occur if the system under consideration is not driven out of equilibrium. Hence, in order to get some understanding of entropy production in terms of microscopic states, the Gibbs' equilibrium statistical mechanics has to be somehow extended to non equilibrium situations. (ii) The fact that we have to describe the system using a non equilibrium distribution $\rho(x,t)$ means that *we have additional information* about the microscopic state of the system as compared to only knowing the current value $\lambda(t)$ of the control parameter (and β of course). The distribution $\rho(x,t)$ is the distribution of microscopic states of the system *knowing how the system was prepared.*

Nevertheless, the question remains: which distribution $\rho(x,t)$ should we choose? Or, what equation does $\rho(x,t)$ obey? This is equivalent to ask: How does our information about the microscopic state of the system evolve?

The thermalization hypothesis has the following consequence: if by some way we know that the microscopic states of the system are distributed according to some non equilibrium distribution ρ_0 (because we have prepared the system to be so) and we let the system in contact with the heat bath keeping the control parameter at a constant value λ, then we expect the system to eventually reach equilibrium. Hence we expect that the microscopic state of the system is eventually distributed according to the canonical distribution ρ_λ. Mathematically, this means that if λ is held constant, then the distribution $\rho(x,t)$ should converge towards $\rho_\lambda(x)$ for any initial condition $\rho_0(x)$. Physically, this means that all the information we initially have about the microscopic state of the system by knowing it is distributed according to ρ_0 eventually disappears. When the system has reached equilibrium, the only information that remains are the values of λ and β.

In the following we will call *micro-state* the microscopic configuration of a thermodynamic system. Typically, a micro-state is the set of positions and momenta of all the particles constituting the system. However, a micro-state might also be some coarse grained quantity such as a reaction coordinate. In any case, a micro-state is something that we neither have access to nor control over. Like before, we will denote a micro-state by the letter x. Each micro-state x has a certain energy $\mathcal{H}_\lambda(x)$. We assume that we know the function $\mathcal{H}_\lambda$ and that we can control it through the control parameter λ. On the other hand, we will call *macro-state* a probability distribution over the micro-states. We will usually denote macro-states with the Greek letter ρ.

Each macro-state ρ has an average energy:

$$E_\lambda[\rho] = \int \rho(x)\mathcal{H}_\lambda(x)\mathrm{d}x. \tag{3.47}$$

Moreover, each macro-state ρ describes a certain state of knowledge we have about the micro-state of the system. The uncertainty we have about the micro-state of the system when it is in macro-state ρ is given by the Shannon entropy $H[\rho]$ of ρ. The canonical distribution ρ_λ is the macro-state of a system in equilibrium with the heat bath. For a given macro-state ρ, the relative entropy

$$D[\rho\|\rho_\lambda] = \int \rho(x)\log\frac{\rho(x)}{\rho_\lambda(x)}\mathrm{d}x \geq 0 \tag{3.48}$$

can be seen as a measure of the "distance to equilibrium" or of the amount of "non-equilibriumness" of ρ. No matter how we prepare our system, we assume that if we leave it in contact with a heat bath, it will relax towards equilibrium and its micro-states will eventually be distributed according to the canonical distribution ρ_λ.

3.2.2 Master equations

Equation of evolution

The simplest equation compatible with the thermalization hypothesis is the following linear equation:

$$\frac{\partial\rho}{\partial t}(x,t) = \int R_{\lambda(t)}(x,x')\rho(x',t)\mathrm{d}x', \tag{3.49}$$

Where the coefficients $R_\lambda(x,x')$ are such that, for λ constant, the distribution $\rho(x,t)$ converges towards the canonical distribution ρ_λ. In a general process, the control parameter depends on time and hence the coefficients R_λ as well, as explicitly written in equation (3.49). In the following, though, we will often omit the explicit time dependence of λ to keep the notations light but it should always be understood that it can vary in time. The dynamics described by equation (3.49) can be summarized into two components: driving and relaxation. By manipulating the control parameter we can tune the instantaneous equilibrium distribution $\rho_{\lambda(t)}$. This is the driving. As a response, the system always tries to converge towards the current equilibrium distribution. This is the relaxation.

The equation (3.49) is phenomenological. It is just the most simple equation describing the relaxation of an arbitrary distribution ρ_0 towards the canonical distribution ρ_λ. The coefficients R_λ should be obtained either though experimental data, or they should be derived from a fully microscopic model including the system and the heat bath by making some approximations. From now on, we assume that equation

(3.49) correctly describes the dynamics of a system coupled to a heat bath and that we know the function R_λ.

In any case, the function R_λ has to satisfy a certain number of properties. The distribution $\rho(x,t)$ should be normalized at every time t:

$$\int \rho(x,t)\mathrm{d}x = 1. \tag{3.50}$$

Hence, the time derivative of the quantity above should be zero. Using equation (3.49), we obtain:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\int \rho(x,t)\mathrm{d}x = \int \frac{\partial \rho}{\partial t}(x,t)\mathrm{d}x &= \int \left(\int R_\lambda(x,x')\rho(x',t)\mathrm{d}x'\right)\mathrm{d}x \\ &= \int \left(\int R_\lambda(x,x')\mathrm{d}x\right)\rho(x',t)\mathrm{d}x' = 0. \end{aligned} \tag{3.51}$$

This quantity should be zero for any distribution $\rho(x',t)$. Hence, the function R_λ has to satisfy

$$\int R_\lambda(x,x')\mathrm{d}x = 0 \tag{3.52}$$

for every λ and every x'.

Obviously, the equilibrium distribution ρ_λ should be invariant under equation (3.49) if λ is held constant. This implies:

$$\int R_\lambda(x,x')\rho_\lambda(x')\mathrm{d}x' = 0 \tag{3.53}$$

for every λ and every x. The canonical distribution ρ_λ should be the only invariant distribution, i.e. it should be the only distribution satisfying (3.53). Furthermore, if λ is held fixed, then any distribution should converge towards the canonical distribution ρ_λ. A sufficient condition for this is that R_λ satisfies the so called *detailed balance* condition:

$$R_\lambda(x,x')\rho_\lambda(x') = R_\lambda(x',x)\rho_\lambda(x). \tag{3.54}$$

This condition expresses the fact that the heat bath is in equilibrium. In the following, we will assume it without more justification. However, many of the results can be generalized to a situation where detailed balance does not hold.

Dynamics at the microscopic level

Under these assumptions, how does the dynamics look like at the microscopic level? Assume that the system is in micro-state x_0 at time t, where is it at some later time? Assuming that the system is in micro-state x_0 amounts to say that its micro-state is distributed according to the so called Dirac distribution $\rho(x,t) = \delta(x - x_0)$ centered on x_0. The Dirac distribution has the following properties: $\delta(x) = 0$ for $x \neq 0$ and it is normalized:

$$\int \delta(x)\mathrm{d}x = 1. \tag{3.55}$$

The Dirac distribution satisfies the following property: For any function f,

$$\int \delta(x)f(x)\mathrm{d}x = \int \delta(x)f(0)\mathrm{d}x = f(0)\int \delta(x)\mathrm{d}x = f(0), \tag{3.56}$$

because $\delta(x)f(x) = \delta(x)f(0)$ since $\delta(x) = 0$ for $x \neq 0$. In fact, the Dirac distribution is actually defined through relation (3.56).

The distribution $\rho(x, t + \mathrm{d}t)$ is, to first order in $\mathrm{d}t$:

$$\rho(x, t + \mathrm{d}t) = \rho(x, t) + \frac{\partial \rho}{\partial t}(x, t)\mathrm{d}t = \rho(x, t) + \left(\int R_\lambda(x, x')\rho(x', t)\mathrm{d}x'\right)\mathrm{d}t, \tag{3.57}$$

where we have inserted equation (3.49) in the second equality to compute $\partial_t \rho$. Inserting $\rho(x, t) = \delta(x - x_0)$ in the equation (3.57) above and using the property (3.56) yields:

$$\rho(x, t + \mathrm{d}t) = \delta(x - x_0) + R_\lambda(x, x_0)\mathrm{d}t. \tag{3.58}$$

Hence for $x \neq x_0$, $R_\lambda(x, x_0)\mathrm{d}x\,\mathrm{d}t$ is the probability to find the system in a neighborhood of size $\mathrm{d}x$ around x at time $t + \mathrm{d}t$ given that it was in x_0 at time t. In other words, for a portion A of phase space not including x_0, the quantity:

$$P(A, t + \mathrm{d}t|x_0, t) = \left(\int_A R_\lambda(x, x_0)\mathrm{d}x\right)\mathrm{d}t \tag{3.59}$$

is the probability that the micro-state of the system is in A at time $t + \mathrm{d}t$ given that it was x_0 at time t. Hence, $R_\lambda(x, x_0)$ is the probability per unit time (and unit phase space volume) that the system "jumps" from micro-state x_0 to micro-state x. It is sometimes referred as "rate" in the literature.

Driving and relaxation

As already mentioned, the dynamics has two components: Driving and relaxation. It is useful to make these two components visible when computing the time derivative of observables. The observables we will be interested in in the following generally have the form:

$$A_\lambda[\rho] = \int f(\rho(x, t), \lambda(t))\mathrm{d}x, \tag{3.60}$$

where f is a real function of its two real arguments. Examples of functionals of this form include the energy $E_\lambda[\rho]$, the Shannon entropy $H[\rho]$ (which does not explicitly depend on the control parameter λ) or the distance to equilibrium $D[\rho\|\rho_\lambda]$.

Differentiating (3.60) with respect to time yields:

$$\frac{\mathrm{d}}{\mathrm{d}t}A_\lambda[\rho] = \dot{\lambda}\frac{\partial}{\partial\lambda}A_\lambda[\rho] + \int \frac{\partial\rho}{\partial t}(x,t)\frac{\partial f}{\partial\rho}(\rho(x,t),\lambda(t))\mathrm{d}x, \tag{3.61}$$

where

$$\frac{\partial}{\partial\lambda}A_\lambda[\rho] = \int \frac{\partial f}{\partial\lambda}(\rho(x,t),\lambda(t))\mathrm{d}x \tag{3.62}$$

is the partial derivative of A with respect to λ. The two terms appearing on the right hand side of equation (3.61) correspond respectively to the two components of the dynamics.

The first term only involves the time derivative of λ; it is the contribution due to the driving. The second term only involves the (partial) time derivative of ρ; it is the contribution due to the relaxation. We introduce the following notation to denote this second term:

$$\begin{aligned} \left.\frac{\partial}{\partial t}A_\lambda[\rho]\right|_{\lambda(t)} &= \int \frac{\partial\rho}{\partial t}(x,t)\frac{\partial f}{\partial\rho}(\rho(x,t),\lambda(t))\mathrm{d}x \\ &= \iint R_{\lambda(t)}(x,x')\rho(x',t)\frac{\partial f}{\partial\rho}(\rho(x,t),\lambda(t))\mathrm{d}x\mathrm{d}x'. \end{aligned} \tag{3.63}$$

In the second equality we have used the master equations (3.49) to express $\frac{\partial\rho}{\partial t}$. Equation (3.63) gives the evolution of $A_\lambda[\rho]$ under the "frozen dynamics", i.e. it is the time derivative of $A_\lambda[\rho]$ if the system is in macro-state $\rho(x,t)$ and the control parameter is "frozen" at the value $\lambda = \lambda(t)$. Using this notation, the time derivative (3.61) of $A_\lambda[\rho]$ takes the more compact form:

$$\frac{\mathrm{d}}{\mathrm{d}t}A_\lambda[\rho] = \dot{\lambda}\frac{\partial}{\partial\lambda}A_\lambda[\rho] + \left.\frac{\partial}{\partial t}A_\lambda[\rho]\right|_{\lambda(t)}, \tag{3.64}$$

where its two contributions, driving and relaxation, are well separated.

In equation (3.48), we introduced $D[\rho\|\rho_\lambda]$ as a measure of the distance of ρ to the equilibrium ρ_λ. While the driving might increase or decrease this quantity, the relaxation can only decrease it:

$$\left.\frac{\partial}{\partial t}D[\rho\|\rho_\lambda]\right|_{\lambda(t)} \leq 0, \tag{3.65}$$

as we show now. Define

$$\Sigma(t) = -\left.\frac{\partial}{\partial t}D[\rho\|\rho_\lambda]\right|_{\lambda(t)}. \tag{3.66}$$

Σ measures the rate of decrease of the distance to equilibrium. Using the normalization condition (3.51), one can show that:

$$\Sigma(t) = -\int \frac{\partial \rho}{\partial t}(x,t) \log \frac{\rho(x,t)}{\rho_{\lambda(t)}(x)} \mathrm{d}x. \tag{3.67}$$

From this expression, using the master equations (3.49) and the detailed balance condition (3.54) we arrive at the following expression:

$$\Sigma = -\iint R_\lambda(x,x')\rho(x') \log \frac{R_\lambda(x',x)\rho(x)}{R_\lambda(x,x')\rho(x')} \mathrm{d}x\mathrm{d}x', \tag{3.68}$$

where we have omitted the time dependence in ρ and in λ to keep the notations light. Using the identity

$$-\log X \geq 1 - X \tag{3.69}$$

for

$$X = \frac{R_\lambda(x',x)\rho(x)}{R_\lambda(x,x')\rho(x')}, \tag{3.70}$$

we obtain:

$$\Sigma \geq \iint R_\lambda(x,x')\rho(x')\mathrm{d}x\mathrm{d}x' - \iint R_\lambda(x',x)\rho(x)\mathrm{d}x\mathrm{d}x' = 0. \tag{3.71}$$

This proves (3.65). There is equality in (3.71) if and only if there is equality in (3.69), i.e. if $X = 1$. This happens if and only if ρ satisfies the detailed balance condition (3.54), i.e. if ρ is the current equilibrium. Note that we have just proved that for λ constant, $\rho(x,t)$ converges towards $\rho_\lambda(x)$ since $D[\rho\|\rho_\lambda]$ decreases towards zero.

As we will show in the next paragraph, Σ is the rate of entropy production, i.e. the amount of entropy produced per unit time. In the literature it is usually defined through equation (3.68) first introduced in [Sch76]. Here, we showed that it measures the rate at which the macro-state of the system tends to relax towards the current equilibrium state.

3.2.3 Thermodynamics

Now that we have an equation for ρ and a model for the microscopic dynamics, we can return to our problem of isothermal work dissipation introduced in paragraph 3.2.1 page 40. The question is: How much work do we have to perform in order to carry out a protocol $\lambda(t)$? Can we show that inequality (3.46) is satisfied? Can we quantify the dissipated work $W_\mathrm{d} = W - \Delta F$ and relate it to the microscopic dynamics?

The work performed per unit time on the system is given by the variation of the energy that is due to the driving:

$$\dot{W} = \dot{\lambda}\frac{\partial}{\partial\lambda}E_\lambda[\rho]. \tag{3.72}$$

The total work performed during the process is then:

$$W = \int \dot{W}\mathrm{d}t, \tag{3.73}$$

where the integration is carried out over the whole duration of the process. Our strategy is to try to compare the work performed per unit time (3.72) to the variation of some "non equilibrium free energy". Inspired by the equilibrium situation, we generalize Gibbs' expression for the equilibrium thermodynamic entropy, equation (2.67), for non equilibrium states:

$$S[\rho] = H[\rho]. \tag{3.74}$$

Using this definition of the entropy, we generalize the equilibrium expression (2.31) of the free energy to non equilibrium states as:

$$F_\lambda[\rho] = E_\lambda[\rho] - TS[\rho], \tag{3.75}$$

where T is the temperature of the heat bath. Using equation (3.45) we first remark that:

$$F_\lambda[\rho] - F(\lambda) = TD[\rho\|\rho_\lambda], \tag{3.76}$$

where $F(\lambda) = F_\lambda[\rho_\lambda]$ is the equilibrium free energy given by (2.63) on page 24. Equation (3.76) implies that the equilibrium distribution minimizes the free energy (3.75).

The quantity $F_\lambda[\rho]$ is a functional of the form (3.60), hence its time variation can be split into the form (3.64):

$$\frac{\mathrm{d}}{\mathrm{d}t}F_\lambda[\rho] = \dot{\lambda}\frac{\partial}{\partial\lambda}F_\lambda[\rho] + \left.\frac{\partial}{\partial t}F_\lambda[\rho]\right|_{\lambda(t)}. \tag{3.77}$$

Since the entropy $S[\rho]$ does not explicitly depend on λ, we get using equation (3.75):

$$\dot{\lambda}\frac{\partial}{\partial\lambda}F_\lambda[\rho] = \dot{\lambda}\frac{\partial}{\partial\lambda}E_\lambda[\rho] = \dot{W}. \tag{3.78}$$

On the other hand, using equation (3.76) we obtain:

$$\left.\frac{\partial}{\partial t}F_\lambda[\rho]\right|_{\lambda(t)} = T\left.\frac{\partial}{\partial t}D[\rho\|\rho_\lambda]\right|_{\lambda(t)} = -T\Sigma, \tag{3.79}$$

since the equilibrium free energy $F(\lambda)$ does not depend on ρ. Equations (3.77), (3.78) and (3.79) can be combined into:

$$\dot{W} = \frac{\mathrm{d}}{\mathrm{d}t} F_\lambda[\rho] + T\Sigma \geq \frac{\mathrm{d}}{\mathrm{d}t} F_\lambda[\rho]. \tag{3.80}$$

In other words, the variations of the non equilibrium free energy (3.75) give a lower bound to the amount of work performed per unit time on the system. The equality is reached if and only if $\Sigma = 0$, i.e. if ρ is the current equilibrium state. Hence, the variations of $F_\lambda[\rho]$ provide the highest lower bound to the amount of work performed per unit time and the name "free energy" is justified. Integrating equation (3.80) over the whole duration of the process yields inequality (3.46). Moreover we get an expression for the dissipated work in terms of the microscopic dynamics:

$$W_\mathrm{d} = W - \Delta F = T \int \Sigma(t) \mathrm{d}t. \tag{3.81}$$

This quantity is positive since $\Sigma(t) \geq 0$ and it is zero if and only $\Sigma(t) = 0$ all along the process, i.e. if the system remains in equilibrium over the whole process.

Using equations (3.75), (3.78) and (3.80), one can show:

$$\dot{W} - \frac{\mathrm{d}}{\mathrm{d}t} F_\lambda[\rho] = T \frac{\mathrm{d}}{\mathrm{d}t} S[\rho] - \left. \frac{\partial}{\partial t} E_\lambda[\rho] \right|_{\lambda(t)} = T\Sigma. \tag{3.82}$$

The quantity:

$$\left. \frac{\partial}{\partial t} E_\lambda[\rho] \right|_{\lambda(t)} = \dot{Q} \tag{3.83}$$

is the rate of change of the energy due to the relaxation. Hence, it is the amount of heat received by the system per unit time. Equation (3.82) can be rewritten in the form:

$$\dot{S} = \frac{\dot{Q}}{T} + \Sigma \geq \frac{\dot{Q}}{T}, \tag{3.84}$$

where we have used the simplified notation $\dot{S} = \frac{\mathrm{d}}{\mathrm{d}t} S[\rho]$. Equation (3.84) above is a generalization of inequality (2.27) page 17 to non-equilibrium situations. The variations of the entropy $S[\rho]$ yields the lowest upper bound to the amount of heat received by the system. Hence, $S[\rho]$ is a "good" generalization of the thermodynamic entropy to non equilibrium states. The quantity Σ introduced in equation (3.66) is equal to the rate of entropy production:

$$\dot{S}_\mathrm{tot} = \dot{S} - \frac{\dot{Q}}{T} = \Sigma \geq 0. \tag{3.85}$$

The total amount of entropy produced is obtained by integrating this relation over the whole duration of the process:

$$\Delta S_{\mathrm{tot}} = \Delta S - \frac{Q}{T} = \int \Sigma(t)\mathrm{d}t. \tag{3.86}$$

Recalling the definition of Σ:

$$\dot{S}_{\mathrm{tot}} = -\left.\frac{\partial}{\partial t} D[\rho \| \rho_\lambda]\right|_{\lambda(t)} \geq 0. \tag{3.87}$$

The rate of thermodynamic entropy production is equal to rate at which the system relaxes towards the current equilibrium state. If λ is held constant, then the total entropy produced during the relaxation from some non-equilibrium initial distribution ρ to the equilibrium ρ_λ is then:

$$\Delta S_{\mathrm{tot}} = D[\rho \| \rho_\lambda]. \tag{3.88}$$

This quantity is the total amount of information about the micro-state of the system that we loose during the relaxation.

We see that the formalism introduced in the previous paragraph reproduces accurately the results of classical isothermal thermodynamics and gives a new interpretation of the entropy production in terms of relaxation towards equilibrium. This formalism now allows us to deal with non-equilibrium states and a natural question is now: What is the minimum amount of work to perform in a transition between two non-equilibrium states? Imagine a process bringing the system from some non-equilibrium macro-state ρ_0 to some other ρ_1 whereby the control parameter was varied from λ_0 to λ_1. Integrating equation (3.80) yields:

$$W = \Delta F + T \int \Sigma(t)\mathrm{d}t \geq \Delta F, \tag{3.89}$$

where $\Delta F = F_{\lambda_1}[\rho_1] - F_{\lambda_0}[\rho_0]$ is the difference in *non-equilibrium free energy* along the process. As in the transition between equilibrium states, the equality is reached if and only if $\Sigma(t) = 0$ during the whole process which only happens if the system is in equilibrium all along the process. Here is the strategy in order to reach the lower bound in inequality (3.89) [HITD10, THD10, EVdB11]:

1. Instantaneously change λ from λ_0 to $\tilde{\lambda}_0$ such that ρ_0 is equilibrium: $\rho_0 = \rho_{\tilde{\lambda}_0}$.

2. Infinitely slowly drive λ from $\tilde{\lambda}_0$ to $\tilde{\lambda}_1$ which is such that ρ_1 is equilibrium $\rho_1 = \rho_{\tilde{\lambda}_1}$.

3. Instantaneously change λ from $\tilde{\lambda}_1$ to its final value λ_1.

The process described above is a combination of infinitely slow and infinitely fast processes. In fact, these processes (and combinations of these processes) are the only "reversible" processes in the sense that they saturate inequality (3.89). By manipulating (3.89), one can show that:

$$\Delta S = \frac{Q}{T} + \int \Sigma(t)\mathrm{d}t, \tag{3.90}$$

is still valid as well in the transition between non equilibrium states. As a conclusion, we can say that the non equilibrium entropy $S[\rho]$ and free energy $F_\lambda[\rho]$ thus defined successfully generalize their equilibrium counterparts to non-equilibrium states.

3.3 Measurement and feed-back

The current macro-state $\rho(x,t)$ of the system represents the information we have about the micro-state of the system. The master equation (3.49) actually describes how our information about the micro-state of the system evolves in time. When we say that the macro-state of the system is $\rho(x,t)$, we mean that we know how it was prepared. We were able to integrate the master equation (3.49) until time t. A legitimate question is now: What happens if we "have a look" at the micro-state of the system; if somehow we are able to get some information about the micro-state of the system, not by preparing it, but by "measuring" it?

Consider the following situation. Our system is in equilibrium, the control parameter being at λ. At some point we perform a measurement on the system. The measurement outcome y is a random variable that depends on the micro-state x occupied by the system at the moment of the measurement. If the measurement is error free, y is a deterministic function of x. However, due to measurement errors this is not generally the case. We note $p(y|x)$ the probability density function of y when the system is in micro-state x at the time of the measurement. The function $p(y|x)$ is a characteristic of the measurement device and it is linked to the measurement errors. The more sharply peaked it is around some value $\bar{y}(x)$, the smaller are the measurement errors.

Knowing the measurement outcome y increases our information about the micro-state currently occupied by the system. Before knowing the measurement outcome, our only information about the system is that it is in equilibrium with the heat bath, hence we assume its micro-states to be distributed according to the equilibrium distribution ρ_λ. The information we obtain by knowing the measurement outcome is incorporated by replacing the a priori distribution ρ_λ by the conditional distribution ρ_y given by Bayes' rule (3.35):

$$\rho_y(x) = \rho(x|y) = \frac{p(y|x)\rho_\lambda(x)}{p_\mathrm{m}(y)}, \tag{3.91}$$

where $p_{\mathrm{m}}(y) = \int p(y|x)\rho_\lambda(x)\mathrm{d}x$ is the marginal distribution of y, i.e. it gives the a priori probability that a given value of the measurement outcome occurs.

In general, the distribution ρ_y is different from ρ_λ. Hence, simply by performing a measurement on a system initially in equilibrium, we have put it in a non equilibrium macro-state. This is not so surprising since we have *defined* equilibrium as the state where the information we have about the system is minimal. Since ρ_y is non equilibrium, its free energy is higher than the free energy of the initial equilibrium state ρ_λ, see equation (3.76):

$$\Delta F(y) = F_\lambda[\rho_y] - F(\lambda) = TD[\rho_y \| \rho_\lambda] \geq 0. \tag{3.92}$$

Hence, by performing a measurement on the system *we were able to increase its free energy without performing work* which violates equation (3.89). By driving the system from ρ_y back to its initial equilibrium state ρ_λ, a maximum amount of work equal to $\Delta F(y)$ can be extracted. This way, it is possible to extract work in a cyclic process involving only one heat bath, which would be impossible without measurement.

The average increase in free energy over the possible measurement outcomes is:

$$\Delta F_{\mathrm{meas}} = \int p_{\mathrm{m}}(y)\Delta F(y)\mathrm{d}y = T\int p_{\mathrm{m}}(y)D[\rho_y\|\rho_\lambda]\mathrm{d}y = TI, \tag{3.93}$$

where we recognize the mutual information I between the measurement outcome and the micro-state of the system, see equation (3.39). Equation (3.93) tells us, that when we perform a measurement on the system, we increase its free energy by an amount that is proportional to the information provided by the measurement. The maximum amount of work that we can extract on average by driving the system back to its initial state ρ_λ is then also TI. Such a process where a measurement is performed on the system and the system is then driven back to its equilibrium state is sometimes called “information to work conversion”.

Note that the system does not need to be initially in equilibrium. In fact, consider a measurement performed on a system initially in an arbitrary macro-state ρ. Right after the measurement, the system is in state ρ_y given by:

$$\rho_y(x) = \frac{p(y|x)\rho(x)}{p_{\mathrm{m}}(y)}, \tag{3.94}$$

where as in (3.91), $p_{\mathrm{m}}(y) = \int p(y|x)\rho(x)\mathrm{d}x$ is the marginal distribution of y. The average over the measurement outcome of the free energy right after the measurement is:

$$\int p_{\mathrm{m}}(y)F_\lambda[\rho_y]\mathrm{d}y = \int p_{\mathrm{m}}(y)E_\lambda[\rho_y]\mathrm{d}y - T\int p_{\mathrm{m}}(y)S[\rho_y]\mathrm{d}y. \tag{3.95}$$

It is easy to show that:

$$\int p_{\mathrm{m}}(y)E_\lambda[\rho_y]\mathrm{d}y = E_\lambda[\rho], \tag{3.96}$$

which means that the measurement does not change the energy of the system on average. On the contrary:

$$\int p_{\mathrm{m}}(y)S[\rho_y]\mathrm{d}y = S[\rho] - I \leq S[\rho] \tag{3.97}$$

where I is the mutual information (3.38) between the micro-state of the system and the measurement outcome. Equation (3.97) says that measuring the state of the system reduces its entropy. Hence, equation (3.93) is still valid if the initial state is non equilibrium:

$$\Delta F_{\mathrm{meas}} = \int p_{\mathrm{m}}(y)F_\lambda[\rho_y]\mathrm{d}y - F_\lambda[\rho] = TI \geq 0. \tag{3.98}$$

In other words, performing a measurement on the system increases its free energy without the need of performing work. The increase in free energy is proportional to the amount of information provided by the measurement. If the system is returned to its initial state ρ, an amount of work up to $\Delta F_{\mathrm{meas}} = TI$ can be extracted.

3.4 Conclusion

Stochastic thermodynamics generalizes Gibbs' equilibrium statistical mechanics to non-equilibrium isothermal processes. A general macroscopic state of a system is represented by a probability density over the microscopic states of the system. If the density is not the current canonical distribution, then the system is out of equilibrium. A non equilibrium macro-state is supposed to relax towards the equilibrium state according to a linear equation. Under these assumptions it is possible to compute the rate at which entropy is irreversibly produced. Moreover, it is possible to generalize the state functions energy, entropy and free energy to non-equilibrium states.

Information theory allows to quantify the information we miss about the microscopic state of a system when we only know its thermodynamic state (specified by the values of the state variables β and λ). It turns out that the canonical distribution of equilibrium statistical mechanics is the probability distribution correctly representing our knowledge of the microscopic state of the system when we only know the values of β and λ.

When we know that a system is in a non-equilibrium state because we know how it was prepared, we have more information about its micro-state than if we only know the values of β and λ. When the system relaxes from this non-equilibrium state towards the equilibrium one, this extra information is lost. The amount of entropy produced during this relaxation, computed using stochastic thermodynamics, is equal to the amount of information lost (quantified using information theory).

If we manage to get some information about the micro-state of the system by performing some measurement, then this information can be converted into work. In fact, measuring the state of the system increases its free energy by an amount proportional to the amount of information obtained through the measurement.

In the following, we will study the thermodynamics of some information processing operation. We will use stochastic thermodynamics to model the device carrying the information.

4 Thermodynamics of the recording and the erasure of information

In the previous chapter, we saw that the irreversibility of thermodynamic processes is due to the fact that we loose information about the microscopic state of the system considered. This chapter investigates the relation between irreversibility and information loss the other way round. The question is whether the erasure of any information is irreversible, and whether there is relationship between the "amount" of information erased and the "amount" of irreversibility of the process (i.e. the amount of entropy produced).

The pioneering work in this respect is due to Rolf Landauer [Lan61]. Working at IBM, Landauer was interested in the minimal energetic costs of digital computation. His major contribution, now known as *Landauer's principle*, was to realize that logically irreversible operations can only be implemented by physically irreversible processes[1]. Quantitatively, Landauer's principle asserts that the erasure of one bit is necessarily accompanied by the generation of $T \log 2$ of heat, where T is the temperature of the environment [Shi95, Pie00, DL09, DBE13]. A bit, or *binary digit*, is physically modelled by a system that has access to two states, 0 and 1, used to encode digital information. The erasure of a bit is the process of bringing the bit to a definite state, e.g. 0, *independently of its initial value.* This operation is also called the *reset to 0 operation.* At the end of the operation, the initial value of the bit it definitely lost. Landauer's principle was demonstrated experimentally only recently [BAP$^+$12, OLT$^+$12].

Based on Landauer's original idea, this chapter investigates the irreversibility of information erasure. However, the information considered here comes from an arbitrary external source and is not necessarily digital. Moreover, the irreversibility of the process is measured in terms of entropy production rather than heat generation. A physical model for a memory device is presented and the processes of recording and erasing information are described. The "amount" of information contained in the memory is well defined all along the erasure process. In fact, the erasure of information is not instantaneous and the amount of information present in the memory

[1] According to Landauer [Lan94], in the fifties it was common to think that the processing of one bit of information leads to the generation of $T \log 2$ of heat. Landauer's contribution was to realize that it is only the *erasure* of information that leads to the generation of heat. Reversible computations can, in principle, be implemented without heat generation.

continuously decreases from a maximum value to zero. Stochastic thermodynamics allows to compute the rate of entropy production all along the erasure process and this rate turns out to be greater than the rate at which the amount of information decreases. This work was published in a more specific setup in [GK13].

4.1 Recording and erasing information on a physical memory

4.1.1 Setup

Here we define what is meant by recording and erasing information. A source randomly emits a symbol α_k out of N possible symbols $\alpha_1, \cdots, \alpha_N$. The probability that α_k is emitted is P_k. The source might be a measurement apparatus and the symbol emitted would be the outcome of a measurement. However, in the following, α_k could be any random variable and the results will not depend on the nature of source. Let H be the Shannon entropy of the probability distribution $\{P_k\}$:

$$H = -\sum_{k=1}^{N} P_k \log P_k. \tag{4.1}$$

It measures our a priori uncertainty about the symbol emitted. As we will see, H can be seen as the average "amount" of information that we want to record.

We want to record the symbol that has been emitted on a physical memory. Eventually, we want to erase the content of the memory. The memory should be a physical system that can be put into at least N different states $\varphi_1, \cdots, \varphi_N$, each corresponding to one of the possible symbols. The state φ_k is said to *encode* the symbol α_k. Moreover, it is convenient to allow for one more state φ_0, called the *standard state* which is the state of the memory when it is empty, i.e. at the beginning of the recording and at the end of the erasure.

The recording process is the process of driving the memory from the standard state to the state encoding the symbol that has appeared. The erasure process is the process of driving the memory back to the standard state *without making use of the information stored*. Hence, the erasure process should be independent of the symbol that has been emitted and that was stored. In fact, if this information is used during the erasure, then it has to be present somewhere outside the memory. If we wish the information to be erased from any medium onto which it is recorded, we will have to perform a process that is independent of the information at some point. This independence of the erasure process in the information stored implies that the erasure process cannot be the time reversed of the recording process (which has to depend on the information to be recorded). For this reason, the erasure process is necessarily accompanied by a certain amount of entropy production which will be quantified in the following.

4.1.2 The memory: encoding, recording and erasing

The memory should be a material system that we can control and which obeys the laws of thermodynamics. We model it as a thermodynamic system in contact with a heat bath at inverse temperature β. Its energy is given by a Hamiltonian function $\mathcal{H}_\lambda(x)$ over its microscopic states $\{x\}$. The Hamiltonian can be controlled through the control parameter λ. The distribution $\rho(x,t)$ over the microscopic states of the memory, or macroscopic state, evolves according to the master equation (3.49).

Through a suitable tuning of the control parameter λ, it is possible to control the distribution $\rho(x,t)$. Hence, in order to encode the different symbols $\{\alpha_k\}$, we use a set of distributions $\{\varphi_k(x)\}$ over the micro-states of the memory.

For the information to be unambiguously stored, the states encoding the different symbols have to be perfectly distinguishable. This means that the corresponding distributions should not overlap: For a given microscopic state x, there is only one k such that $\varphi_k(x) > 0$. If this is not the case, say for a given x, $\varphi_k(x) > 0$ and $\varphi_{k'}(x) > 0$, then the information is not perfectly reliably stored. However, even in this case, we can quantify the "amount" of information recorded. We will address this issue after having discussed the erasure.

We are now able to specify the recording process. At the beginning of the process, the memory is in the standard state $\varphi_0(x)$. At the end of the process, we want it to be in the state $\varphi_k(x)$ encoding the symbol α_k that appeared. This process is implemented by changing the control parameter. From time t_i to time t_rec, it is varied from its initial value λ_0 to some final value λ_rec according to some time dependence $\lambda_k(t)$. The function $\lambda_k(t)$ is the *recording protocol.* Note that the protocol $\lambda_k(t)$ has to depend on k, else it is not possible to bring the memory in different states. However we require that the final value λ_rec *does not depend on the symbol that is recorded* and is thus the same for every α_k. In fact, at the end of the recording process, we want to be able to manipulate the memory *without knowing which symbol is stored.*

At the end of the recording process, *the memory is out of equilibrium.* In fact, at the end of the recording process, we want to allow the memory to be described by one out of N different distributions, but for each value λ of the control parameter, *there is only one equilibrium state* given by the canonical distribution (2.60). In the following, we will quantify the minimum amount of "non-equilibriumness" needed in order to store a certain amount of information. If we want to store the information for a long period of time, we probably want the $\{\varphi_k(x)\}$ to be metastable with a long life-time. Else the memory would tend to relax towards the unique equilibrium state right after the recording process and the information would get lost rather quickly.

The erasure process is a process bringing the memory back to the standard state *without making use of the information stored.* In light of the previous paragraph, we already guess that one way of erasing the information is to let the memory relax towards the unique equilibrium state. In general, we might want to control the erasure process by applying a certain protocol $\lambda(t)$. The crucial feature of the erasure process is that the protocol *should not depend on the information stored*, i.e. it should not depend on which α_k was recorded. For simplicity, we assume that we start the erasure process right after the recording process and that at the end of the erasure process, the control parameter is returned to its initial value λ_0. From time t_{rec} to time t_{f}, the control parameter is driven from $\lambda(t_{\mathrm{rec}}) = \lambda_{\mathrm{rec}}$ to $\lambda(t_{\mathrm{f}}) = \lambda_0$ in such a way that the final macro-state of the memory is the standard state $\varphi_0(x)$.

4.1.3 The information contained in the memory

Let us now quantify the amount of information contained in the memory during the erasure process. The question we would like to answer is the following: Assume that we are able to know in which micro-state the memory is at some intermediate time of the erasure process. What can we infer about the symbol that was recorded?

Let $\rho_k(x,t)$ be the macro-state of the memory at time t of the erasure process if the symbol α_k was recorded. It is the probability distribution over the micro-states of the memory at time t of the erasure process *conditioned on α_k being recorded* and it is obtained by propagating the distribution $\varphi_k(x)$ with the master equation (3.49) with the erasure protocol $\lambda(t)$. If we know that the memory is in micro-state x at time t, then we can infer the probability $P(k|x;t)$ that the symbol α_k was recorded. It is given by the Bayes' rule:

$$P(k|x;t) = \frac{\rho_k(x,t)P_k}{\rho_{\mathrm{m}}(x,t)}, \tag{4.2}$$

where $\rho_{\mathrm{m}}(x,t) = \sum_k P_k \rho_k(x,t)$ is the marginal distribution over the micro-states of the memory at time t. The presence of $\rho_{\mathrm{m}}(x,t)$ in the expression above ensures that the probability distribution $P(k|x;t)$ is normalized: $\sum_k P(k|x;t) = 1$. We will discuss the physical meaning of the marginal distribution $\rho_{\mathrm{m}}(x,t)$ in the section 4.3 below. For now, let us just remark that it is a linear combination of the distributions $\rho_k(x,t)$ and these distributions all obey the same *linear* master equations. Consequently, the marginal distribution $\rho_{\mathrm{m}}(x,t)$ obeys the same master equation as well. At the beginning of the erasure process it is given by

$$\rho_{\mathrm{m}}(x,t_{\mathrm{rec}}) = \varphi_{\mathrm{m}}(x) = \sum_k P_k \varphi_k(x), \tag{4.3}$$

and at some subsequent time, it can be obtained by propagating this initial value with the master equation (3.49) with the erasure protocol $\lambda(t)$.

At the beginning of the erasure process, $t = t_{\mathrm{rec}}$, $P(k|x;t_{\mathrm{rec}}) = 0$ or 1 depending on whether x belongs to the support of φ_k of not. At that time, knowing the micro-state of the memory allows to infer the symbol that was recorded with certainty.

The information contained in the memory is maximum. At the end of the erasure process, $t = t_\mathrm{f}$, we have that $\rho_k(x, t_\mathrm{f}) = \varphi_0(x)$ for each k and thus $\rho_\mathrm{m}(x, t_\mathrm{f}) = \varphi_0(x)$. As a consequence, $P(k|x; t_f) = P_k$ for each k, meaning that the knowledge of the micro-state x of the memory does not change our a priori knowledge of the symbol that appeared. At that moment, one can safely say that the memory does not contain anymore information about the symbol that originally appeared and got recorded.

At some intermediate time t, the information contained in the memory is not maximum anymore, but it is not necessarily zero. The uncertainty about the symbol that was emitted upon knowing that the memory is micro-state x at time t is quantified by the Shannon entropy of the probability distributions $P(k|x; t)$:

$$h_\mathrm{er}(x, t) = -\sum_k P(k|x; t) \log P(k|x; t). \tag{4.4}$$

The amount of information erased until time t is the average uncertainty about the symbol that was stored upon knowing the position of the particle at time t:

$$I_\mathrm{er}(t) = \int \rho_\mathrm{m}(x, t) h_\mathrm{er}(x, t) \mathrm{d}x. \tag{4.5}$$

In information theoretic terms, this is the entropy of the symbol emitted conditioned on the micro-state of the memory at time t. It satisfies [CT06]:

$$0 \leq I_\mathrm{er}(t) \leq H. \tag{4.6}$$

The second inequality above means that knowing the micro-state of the memory at time t of the erasure process on average reduces our uncertainty about the symbol emitted. The amount of information still contained in the memory at time t of the erasure process can be defined as the reduction in uncertainty about the symbol that was stored upon knowing the position of the particle at time t:

$$I(t) = H - I_\mathrm{er}(t). \tag{4.7}$$

This is just the *mutual information* between the position of the particle and the symbol originally recorded, see equation (3.22) page 34. It is a measure of how much information the position of the particle at time t can still provide about the symbol originally stored [CT06].

As we would intuitively expect, $I_\mathrm{er}(t_\mathrm{rec}) = 0$, i.e. at the beginning of the erasure process, no information is yet erased. Hence, at the beginning of the erasure process, the amount of information contained the memory is $I(t_\mathrm{rec}) = H$, the Shannon entropy of the probability distribution $\{P_k\}$. Thus, H can be seen as the amount of information that has to be recorded. At the end of the process, knowing the position of the particle does not reduce our uncertainty about the symbol that had been emitted and $I_{\mathrm{er}}(t_{\mathrm{f}}) = H$. We can safely say that at that point the memory does not contain anymore information about the symbol originally emitted. We have $I(t_{\mathrm{f}}) = 0$, which is consistent with this statement.

Applying equation (3.39) to the present situation, we can rewrite $I(t)$ in the following form:

$$I(t) = \sum_k P_k D[\rho_k(t) \| \rho_{\mathrm{m}}(t)]. \tag{4.8}$$

In fact, $\rho_k(x,t)$ is the distribution of x conditioned on the symbol recorded, $\rho_{\mathrm{m}}(x,t)$ is the marginal distribution of x and P_k is the marginal distribution of the symbol recorded. The quantity $D[\rho_k(t) \| \rho_{\mathrm{m}}(t)]$ is the relative entropy of two distributions obeying the same master equations. It is decreasing in time (see [CT06] page 34). Hence, the amount of information contained in the memory does decrease during the erasure process. This is due to the fact that the erasure protocol does not depend on k, so that during the erasure process, all the $\rho_k(x,t)$, and hence $\rho_{\mathrm{m}}(x,t)$, obey the same master equations.

Due to practical limitations, it might not be possible to prepare the memory in completely non overlapping states. Nevertheless, in this case it is still possible to record *some* information and it is possible to quantify the maximum amount $I_{\max}$ of information that be be stored:

$$I_{\max} = I(t_{\mathrm{rec}}) = \sum_k P_k D[\varphi_k \| \varphi_{\mathrm{m}}] \leq H, \tag{4.9}$$

where $\varphi_{\mathrm{m}}(x) = \sum_k P_k \varphi_k(x)$ was introduced in equation (4.3).

For any distribution $\tilde{\rho}(x)$ we have the identity

$$\sum_k P_k D[\rho_k(t) \| \tilde{\rho}] = I(t) + D[\rho_{\mathrm{m}}(t) \| \tilde{\rho}]. \tag{4.10}$$

In particular, for $\tilde{\rho}(x) = \rho_{\lambda(t)}(x)$, we get the following result:

$$\sum_k P_k D[\rho_k(t) \| \rho_{\lambda(t)}] = I(t) + D[\rho_{\mathrm{m}}(t) \| \rho_{\lambda(t)}] \geq I(t). \tag{4.11}$$

This result means that as long as the memory contains some information, it is out of equilibrium. Moreover, its distance to equilibrium is greater than the information contained. In particular, for time $t = t_{\mathrm{rec}}$, i.e. just at the end of the recording process, equation (4.11) implies:

$$\sum_k P_k D[\varphi_k \| \rho_{\lambda(t_{\mathrm{rec}})}] \geq H. \tag{4.12}$$

In section 4.1.2, we already observed that in order to record some information, the device serving as a memory has to be driven out of equilibrium. Inequality (4.12), and more generally equation (4.11) quantify the minimum amount of "non equilibriumness" needed. The average distance to equilibrium measured through the relative entropy to the equilibrium distribution has to be greater than the amount of information one wishes to record.

4.2 Thermodynamics of the processes

4.2.1 Entropy production

The recording process can in principle be performed reversibly in the sense that it can be performed with an arbitrarily small amount of entropy production. Hence we now focus on the entropy produced during the erasure process.

The rate of entropy production at time t of the erasure process, assuming that the symbol α_k was stored is, according to equation (3.87):

$$\dot{S}_k^{\mathrm{tot}} = -\left.\frac{\partial}{\partial t} D[\rho_k(t)\|\rho_{\lambda(t)}]\right|_{\lambda(t)} \tag{4.13}$$

Hence, using equation (4.11), we get that the average rate of entropy production reads:

$$\dot{S}_{\mathrm{tot}} = \sum_k P_k \dot{S}_k^{\mathrm{tot}} = -\left.\frac{\partial I}{\partial t}\right|_{\lambda(t)} - \left.\frac{\partial}{\partial t} D[\rho_{\mathrm{m}}(t)\|\rho_{\lambda(t)}]\right|_{\lambda(t)}. \tag{4.14}$$

Since the information content $I(t)$ does not explicitly depend on the control parameter $\lambda(t)$, we have:

$$\left.\frac{\partial I}{\partial t}\right|_{\lambda(t)} = \frac{\mathrm{d}I}{\mathrm{d}t}(t) = \dot{I}(t) \leq 0. \tag{4.15}$$

This quantity is the instantaneous rate of variation of the information content of the memory. In the previous paragraph, we showed that the information is a decreasing function of time. The quantity $-\dot{I}(t) \geq 0$ is the *instantaneous rate of information erasure*. Furthermore, since the distribution $\rho_{\mathrm{m}}(x,t)$ satisfies the same equation as the $\rho_k(x,t)$'s, the second term in the right hand side of equation (4.14) is non negative:

$$-\left.\frac{\partial}{\partial t} D[\rho_{\mathrm{m}}(t)\|\rho_{\lambda(t)}]\right|_{\lambda(t)} = \sigma(t) \geq 0. \tag{4.16}$$

This quantity is the rate at which entropy would be produced if the memory had been prepared in the state $\rho_{\mathrm{m}}(x,t)$ and evolved according to the protocol $\lambda(t)$. Combining equations (4.14), (4.15) and (4.16) leads us to the following result relating the rate of entropy production to the rate of information erasure:

$$\dot{S}_{\mathrm{tot}}(t) = -\dot{I}(t) + \sigma(t) \geq -\dot{I}(t). \tag{4.17}$$

Equation (4.17) above is a precise and general statement about the thermodynamic costs of information erasure. It is a generalization of Landauer's principle [Lan61]. In words, it states that during an erasure process, the instantaneous rate of entropy production is bounded from below by the instantaneous rate of information erasure.

Equation (4.17) shows that the entropy production rate has two non negative contributions $-\dot{I}(t)$ and $\sigma(t)$. The first contribution $-\dot{I}(t)$ comes from the destruction of the correlations between the microscopic state of the memory and the symbol that was emitted by the source. This contribution is positive as long as the information content is non zero. The second contribution $\sigma(t)$ comes from the relaxation of the marginal distribution $\rho_{\mathrm{m}}(x,t)$ towards the equilibrium state $\rho_{\lambda(t)}(x)$. It vanishes if and only if the two distributions are identical, $\rho_{\mathrm{m}}(x,t) = \rho_{\lambda(t)}(x)$. Hence, if one wants to minimize the entropy produced by an erasure process, one should ensure that at any time the marginal distribution is the equilibrium distribution. Such a process could be qualified as a "quasi-static erasure". To summarize, the erasure process is composed of two processes: (i) the convergence of the different $\rho_k(x,t)$'s towards each other, and hence towards the marginal distribution $\rho_{\mathrm{m}}(x,t)$, and (ii) the relaxation of the marginal distribution towards the equilibrium state $\rho_{\lambda(t)}(x)$. Each of these processes contribute to the entropy production.

Integrating equation (4.17) over the whole erasure process yields following lower bound to the total amount of entropy produced during this process:

$$\Delta S_{\mathrm{tot}} \geq H, \tag{4.18}$$

where $H = -\sum_k P_k \log P_k$ is the Shannon entropy of the source. Equation (4.18) above is a generalization of Landauer's principle to the situation where the information to be erased is not necessarily binary and the device serving as a memory is arbitrary. If it was not possible to prepare the memory in non-overlapping states (i.e. if the $\varphi_k(x)$ do overlap), then this lower bound is reduced:

$$\Delta S_{\mathrm{tot}} \geq I_{\mathrm{max}}, \tag{4.19}$$

where $I_{\mathrm{max}} \leq H$ is the maximal amount of information recorded and is given by equation (4.9). More generally, the entropy produced between times t and t' of the erasure process is bound from below by:

$$\Delta S_{\mathrm{tot}}(t,t') \geq I(t) - I(t'), \tag{4.20}$$

i.e. by the decrease in information between times t and t'.

4.2.2 Work performed

Recording process

The lower bound to the work performed during the recording process is given by the variation of the non equilibrium free energy of the memory along the process. Let $F_0 = F_{\lambda_0}[\varphi_0]$ be the initial free energy of the memory and $F_k^{\mathrm{rec}} = F_{\lambda_{\mathrm{rec}}}[\varphi_k]$ the free energy of the memory when α_k is recorded. The amount W_k^{rec} of work performed when recording the symbol α_k is bound from below by:

$$W_k^{\mathrm{rec}} \geq F_k^{\mathrm{rec}} - F_0, \tag{4.21}$$

The expected amount $W_{\mathrm{rec}} = \sum_k P_k W_k^{\mathrm{rec}}$ of work to perform for the recording process is then bounded from below by:

$$W_{\mathrm{rec}} \geq \sum_k P_k \left(F_k^{\mathrm{rec}} - F_0\right). \tag{4.22}$$

Noting that

$$\sum_k P_k E_{\lambda_{\mathrm{rec}}}[\varphi_k] = E_{\lambda_{\mathrm{rec}}}[\varphi_{\mathrm{m}}] \tag{4.23}$$

$$\sum_k P_k S[\varphi_k] = S[\varphi_{\mathrm{m}}] - H, \tag{4.24}$$

where $\varphi_{\mathrm{m}}(x) = \sum_k P_k \varphi_k(x)$ was already introduced, and recording the definition of the non equilibrium free energy (3.75) we get the following relation:

$$\sum_k P_k F_k^{\mathrm{rec}} = F_{\mathrm{m}}^{\mathrm{rec}} + TH, \tag{4.25}$$

where $F_{\mathrm{m}}^{\mathrm{rec}} = F_{\lambda_{\mathrm{rec}}}[\varphi_{\mathrm{m}}]$ would be the free energy of the memory if we had prepared it in the state $\varphi_{\mathrm{m}}(x)$ and T is the temperature of the heat bath.

Combining equations (4.25) and (4.22), we obtain:

$$W_{\mathrm{rec}} \geq F_{\mathrm{m}}^{\mathrm{rec}} - F_0 + TH. \tag{4.26}$$

The quantity $F_{\mathrm{m}}^{\mathrm{rec}} - F_0$ is the minimum amount of work that we would have to provide in order to prepare the memory in state $\varphi_{\mathrm{m}}(x)$. Hence, in order to record the symbol that appeared, i.e. in order to prepare the memory in state $\varphi_k(x)$ with probability P_k, one has to perform an extra amount of TH of work than to prepare the memory in the state $\varphi_{\mathrm{m}}(x)$. However, this extra amount of work is not dissipated. It could in principle be retrieved by reversing the process, i.e. by driving the memory back to the standard state $\varphi_0(x)$ *by a protocol depending on k.*

If the $\varphi_k(x)$ overlap, then one should simply replace H by $I_{\max}$ in equations (4.24), (4.25), and (4.26).

Erasure process

Unlike the recording process, the erasure process is necessarily producing entropy. Hence in addition to the reversible work, which is given by the variations of the free energy, the work performed also contains some irreversible contribution, the work *dissipated* proportional to the entropy production. If the symbol α_k was recorded, the work performed per unit time at time t of the erasure process is given by:

$$\dot{W}_k^{\mathrm{er}}(t) = \dot{F}_k(t) + T\dot{S}_k^{\mathrm{tot}}, \tag{4.27}$$

where $F_k(t) = F_{\lambda(t)}[\rho_k(t)]$ is the instantaneous free energy of the memory at time t of the erasure process. Averaging over α_k and taking equation (4.17) into account, we obtain:

$$\dot{W}_{\mathrm{er}}(t) \geq \sum_k P_k \dot{F}_k(t) - T\dot{I}(t), \tag{4.28}$$

where $\dot{W}_{\mathrm{er}} = \sum_k P_k \dot{W}_k^{\mathrm{er}}$. This relation is just the work counterpart of equation (4.17). It states that during the erasure process the rate of work dissipation is bounded from below by $-T\dot{I}$, i.e. by the temperature times the rate of information erasure. The total amount of work performed during the erasure process satisfies:

$$W_{\mathrm{er}} \geq F_0 - \sum_k P_k F_k^{\mathrm{rec}} + TH, \tag{4.29}$$

where F_0 is the final free energy of the memory. Again, equation (4.29) is the work counterpart of equation (4.18). It states that the total work dissipated during the erasure process is at least TH.

Equation (4.24) can be generalized as follows at any time t of the erasure process:

$$\sum_k P_k S[\rho_k(t)] = S[\rho_{\mathrm{m}}(t)] - I(t). \tag{4.30}$$

As a consequence, the free energy of the memory satisfies a relation similar to equation (4.25) at any time t of the erasure process:

$$\sum_k P_k F_k(t) = F_{\mathrm{m}}(t) + TI(t), \tag{4.31}$$

where $F_{\mathrm{m}}(t) = F_{\lambda(t)}[\rho_{\mathrm{m}}(t)]$. Inserting this relation in equation (4.28) above yields:

$$\dot{W}_{\mathrm{er}}(t) \geq \dot{F}_{\mathrm{m}}(t). \tag{4.32}$$

Integrated over the whole erasure process, the minimum amount of work performed during the erasure process is given by:

$$W_{\mathrm{er}} \geq F_0 - F_{\mathrm{m}}^{\mathrm{rec}}. \tag{4.33}$$

In other words, the minimum amount of work to perform is given by the variations of the free energy of the marginal distribution $\rho_{\mathrm{m}}(x,t)$.

Summing equations (4.22) and (4.29), or (4.26) and (4.33), we obtain for the total amount of work performed during the recording and erasure cycle:

$$W_{\mathrm{tot}} = W_{\mathrm{rec}} + W_{\mathrm{er}} \geq TH. \tag{4.34}$$

This work is truly lost since the memory is in the same state at the beginning and at the end of the cycle. It got dissipated in form of heat to the heat bath. If we hadn't performed the process cyclically, i.e. if the initial and final states of the memory were different, then equation (4.34) would have to include some difference in free energy between the final and the initial state of the memory.

4.3 Discussion

The work performed during the erasure process behaves as if the memory had been prepared in the state φ_{m}. However, the recording process on average necessitates more work than needed to prepare the memory in the state φ_{m}, see equation (4.26). It could seem that when we "forget" which symbol was recorded at the end of the recording process, the state of the memory "collapses" from φ_k to φ_{m}. This collapse would be accompanied by an average decrease in free energy exactly corresponding to the information that was recorded (and got lost with the symbol):

$$\sum_k P_k F_k^{\mathrm{rec}} - F_{\mathrm{m}}^{\mathrm{rec}} = T\left(S[\varphi_{\mathrm{m}}] - \sum_k P_k S[\varphi_k]\right) = TH. \tag{4.35}$$

This would explain the energetic loss happening in this recording and erasure cycle.

If this was the case, then the information contained in the memory would instantaneously vanish at the very moment we forget the symbol, and an amount of entropy of

$$\Delta S_{\mathrm{tot}}^{\mathrm{forget}} = H \tag{4.36}$$

would be instantaneously produced. Apart of being unphysical, this would be in contradiction with previous discussions. Moreover, such a collapse of the φ_k into φ_{m} would not be described by the evolution equations. At time t of the erasure process, the state of the memory is $\rho_k(x,t)$ with probability P_k and not $\rho_{\mathrm{m}}(x,t)$. The macrostate of the memory obeys the master equations even between the recording and the erasure process and hence no collapse occurs since such a collapse would not be described by master equations. Instead, as argued in the preceding, for every k, $\rho_k(x,t)$ smoothly converges towards $\rho_\mathrm{m}(x,t)$.

However, when we "forget" the symbol recorded, *our* information about the micro-state of the memory suddenly changes. Hence, the probability distribution *we assign* to the micro-states of the memory changes from φ_k to φ_m. And all along the erasure process we can only optimize the protocol with respect to $\rho_\mathrm{m}(x,t)$. Hence it is natural that the minimum amount of work we have to provide is given by the variations of the free energy of this distribution.

Assume that at time t' we suddenly remember the symbol α_k that was recorded. By remembering which symbol was recorded, we suddenly recover some information about the micro-state of the memory and we can reassign them the right distribution $\rho_k(x,t)$. We can make use of that information by selecting a different protocol depending on the symbol that was recorded. Hence, between times t_rec and t' when we do not know which symbol was recorded, it seems to us that the free energy of the memory is $F_\mathrm{m}(t)$ and the minimum amount we have to perform is equal to $F_\mathrm{m}(t') - F_\mathrm{m}^\mathrm{rec}$. At time t', when we suddenly recover the symbol that was stored, we can reassign the "right" free energy $F_k(t')$ to the memory and it will seem to us that we have suddenly "gained" an amount

$$\sum_k P_k F_k(t') - F_\mathrm{m}(t') = TI(t') \leq TH \tag{4.37}$$

of free energy on average. This quantity is smaller than the free energy we had lost by "forgetting" the symbol recorded at time t_rec, see equation (4.35).

During this alternative process, we cannot say that we erased information about the symbol that appeared in the first place since we made use of this information from time t' to t_f. Nevertheless, we did loose something. What we lost is information *about the micro-state of the memory.* By "forgetting" the symbol recorded at time t_rec, we lost an amount H of information about the micro-state of the memory and by "remembering" the symbol, we recovered an amount $I(t') \leq H$ of information about the micro-state of the memory. So between times t_rec and t' we lost $H - I(t')$. This very quantity is also the minimum amount of entropy produced between t_rec and t':

$$\Delta S_\mathrm{tot}(t') \geq H - I(t'), \tag{4.38}$$

obtained by setting $t = t_\mathrm{rec}$ in equation (4.20).

All along the erasure process, the symbol recorded contains some information about the micro-state of the memory. But as this information is not used, it is erased because of the relaxation towards equilibrium, i.e. *by the heat bath.* This relaxation is accompanied by an amount of entropy production which is linked to the amount of information lost. As a consequence, the information the micro-state of the memory contains about the symbol initially emitted is erased as well.

4.4 A simple example

Let us now illustrate the previous results on the most simple system that one can imagine: Recording and erasing the result of a binary random variable on a two states memory. Let $\alpha_1 = \mathrm{h}$ (for "head") and $\alpha_2 = \mathrm{t}$ (for "tail") be the two possible results of a binary random variable. The probability for α_1 to appear is $P_1 = P$ and the probability for α_2 to appear is $P_2 = 1 - P$. Hence, the amount of information we wish to record is the Shannon entropy of the distribution $(P, 1 - P)$:

$$H(P) = -P \log P - (1 - P) \log(1 - P). \tag{4.39}$$

4.4.1 The memory as a two states system

The system we use as the memory is a two states system in contact with a heat bath at inverse temperature β. Let "Left" and "Right" label these two states and let E_L and E_R be their respective energies. At any time, the memory can make stochastic transitions between these two states. The energy needed to make a transition is provided by the heat bath in form of heat. We assume that the probability per unit time to make a transition has a Kramer's form: The probability w_RL per unit time to jump from the left state to the right state is then

$$w_\mathrm{RL} = \frac{1}{\tau} \exp\left(\beta E_\mathrm{L}\right), \tag{4.40}$$

and similarly, the probability per unit w_LR time to make a transition from the right state to the left state is given by:

$$w_\mathrm{LR} = \frac{1}{\tau} \exp\left(\beta E_\mathrm{R}\right), \tag{4.41}$$

where $\tau \propto \exp\left(-\beta V\right)$ is a time linked to the heigh of the potential barrier V that needs to be overcome to make a transition.

The macroscopic state of the memory is given by the occupation probabilities p_L and p_R of the left and right states. They evolve in time according to the following master equations:

$$\begin{aligned} \dot{p}_\mathrm{L} &= -w_\mathrm{RL}\, p_\mathrm{L} + w_\mathrm{LR}\, p_\mathrm{R} \\ \dot{p}_\mathrm{R} &= +w_\mathrm{RL}\, p_\mathrm{L} - w_\mathrm{LR}\, p_\mathrm{R}. \end{aligned} \tag{4.42}$$

The equilibrium state is given by the canonical distribution:

$$\begin{aligned} p_{\mathrm{L}}^{\mathrm{eq}} &= \frac{e^{-\beta E_{\mathrm{L}}}}{Z} \\ p_{\mathrm{R}}^{\mathrm{eq}} &= \frac{e^{-\beta E_{\mathrm{R}}}}{Z}, \end{aligned} \tag{4.43}$$

where $Z = \exp(-\beta E_{\mathrm{L}}) + \exp(-\beta E_{\mathrm{R}})$ is the partition function. For fixed value of E_{L}, E_{R} and τ, any initial distribution over the two micro-states of the memory will relax exponentially within a characteristic time τ towards the canonical distribution (4.43). Normalization requires $p_{\mathrm{R}} + p_{\mathrm{L}} = 1$ and hence the macroscopic state of the memory is fully specified by the probability p_{L} that its micro-state is the left state. To simplify the notations, we set $p_{\mathrm{L}} = p$ and in the following, when we speak about the macro-state of the memory, we mean the probability p that it occupies its left micro-state. Consequently, all the lower case p or q will refer to the probability for the memory to occupy the left micro-state.

We assume that we can control the energy difference $\Delta E = E_{\mathrm{R}} - E_{\mathrm{L}}$ between the two micro-states. Hence, through ΔE we can control the macroscopic state p of the memory. Again, for simplicity, we assume that $E_{\mathrm{L}} + E_{\mathrm{R}} = 0$. Hence we set $E_{\mathrm{L}} = E$ and $E_{\mathrm{R}} = -E$ and our control parameter is E. The equilibrium macro-state of the memory is then given by $p_{\mathrm{eq}}(E) = \exp(-\beta E)/Z(E)$ and the partition function by $Z(E) = 2\cosh(\beta E)$. The master equations (4.42) then simplify to:

$$\dot{p} = -\frac{1}{\tau}\left(p - p_{\mathrm{eq}}(E(t))\right), \tag{4.44}$$

where the time dependence of E gives the protocol.

We can now turn to the instantaneous rate of entropy production. Let us first introduce the relative entropy between two distributions $(p, 1-p)$ and $(q, 1-q)$ over a binary random variable:

$$D(p\|q) = p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q}. \tag{4.45}$$

This is a particular case of equation (3.23). The quantity $D(p\|q)$ is non negative and it is zero if and only if $p = q$. The instantaneous rate of entropy production $\dot{S}_{\mathrm{tot}}$ given by equation (3.87) reads in this simple case:

$$\dot{S}_{\mathrm{tot}}(t) = -\dot{p}(t)\frac{\partial}{\partial p}D(p(t)\|p_{\mathrm{eq}}(E(t))). \tag{4.46}$$

where $p(t)$ is the solution of the master equation (4.44) and $E(t)$ is given by the protocol.

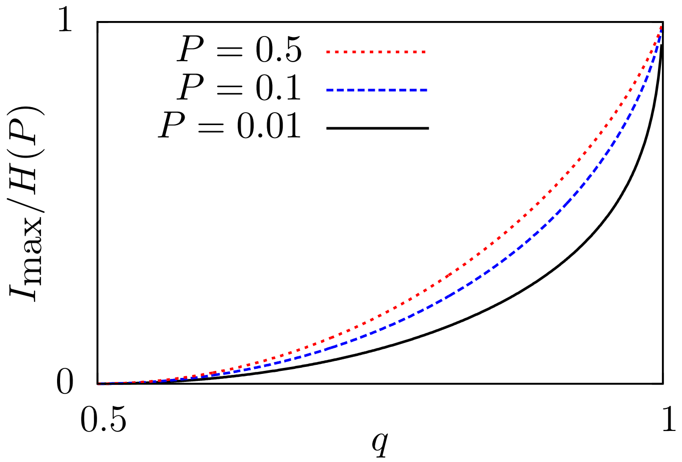

Figure 4.1: $I_{\max}/H(P)$ as a function of q for a symmetric memory: $q_1 = 1 - q_2 = q$.

4.4.2 Information stored in the memory

Now that we have described the system serving as a memory we have to choose the states encoding α_1 and α_2. Let $(q_1, 1-q_1)$ encode α_1 and $(q_2, 1-q_2)$ encode α_2. The maximum amount of information that we can store in the memory in this setup is given by:

$$I_{\max} = P_1 D(q_1 \| q_{\mathrm{m}}) + P_2 D(q_2 \| q_{\mathrm{m}}), \tag{4.47}$$

where $q_{\mathrm{m}} = P_1 q_1 + P_2 q_2$ is the marginal probability for the memory to occupy the left micro-state at the end of the recording process. The maximum amount of information stored in the memory satisfies:

$$0 \leq I_{\max} \leq H(P). \tag{4.48}$$

It is zero if and only if $q_1 = q_2$ and it is equal to $H(P)$, the amount of information emitted by the source, if and only if the two distributions $(q_1, 1-q_1)$ and $(q_2, 1-q_2)$ do not overlap. This happens when $q_1 = 1$ and $q_2 = 0$ (and vice versa). In this case, recording α_1 means to confine the memory in its left micro-state and recording α_2 means to confine the memory in its right micro-state. However, this might be difficult to implement in practice since it would necessitate to apply an infinite energy difference between the two micro-states.

If $q_1 = 1 - q_2 = q$, then we speak about a *symmetric memory*. On figure 4.1 we plotted $I_{\max}/H(P)$ as a function of q for a symmetric memory for three values of P. As we just said, $I_{\max} = 0$ for $q = 1/2$ and $I_{\max} = H(P)$ for $q = 1$.

4.4.3 The erasure processes

The erasure protocol

For the illustration, we assume that we were able to prepare the memory in the two non overlapping macro-states $q_1 = 1$ and $q_2 = 0$. Moreover, we assume that at the beginning of the erasure process, the two micro-states of the memory have the same energy $E_{\mathrm{rec}} = 0$. We propose the following erasure process: From time $t = 0$ to time $t = t_{\mathrm{f}} = 20\tau$, the parameter E is driven from 0 to the final value $E_{\mathrm{f}} = 5/\beta$ according to

$$E(t) = \frac{t}{t_{\mathrm{f}}} E_{\mathrm{f}}. \tag{4.49}$$

Moreover, we set $P_1 = P = 0.3$. The values of parameters P, q_1, q_2, E_{rec}, E_{f} and t_{f} are arbitrary and the particular values used here were chosen to best illustrate what happens during the erasure.

State of the memory during the erasure

Let $p_1(t)$ (resp. $p_2(t)$) be the macro-state of the memory at time t of the erasure process if the symbol α_1 (resp. α_2) was recorded. In other word, $p_1(t)$ (resp. $p_2(t)$) is the solution to the master equation (4.44) with the protocol given by (4.49) and with $q_1 = 1$ (resp. $q_2 = 0$) as initial condition. We introduce:

$$p_{\mathrm{m}}(t) = P_1 p_1(t) + P_2 p_2(t), \tag{4.50}$$

the marginal probability for the memory to occupy the left micro-state at time t of the erasure process.

On figure (4.2), we plotted the time evolution of p_1, p_2, p_{m} and of the equilibrium state p_{eq} during the erasure process. The functions $p_1(t)$ and $p_2(t)$ quickly converge towards each other, and hence towards $p_{\mathrm{m}}(t)$, on a time scale of order τ.

Inferring the symbol stored during the erasure

If the memory is found in the left micro-state at time t of the erasure process, then the probability that α_1 was stored is given by Bayes' rule (4.2):

$$P(\alpha_1|\mathrm{L};t) = \frac{p_1(t)}{p_{\mathrm{m}}(t)} P_1. \tag{4.51}$$

Similarly, if the memory is found in the right micro-state at time t, then the probability that the α_2 was stored is:

$$P(\alpha_2|\mathrm{R};t) = \frac{1 - p_2(t)}{1 - p_{\mathrm{m}}(t)} P_2. \tag{4.52}$$

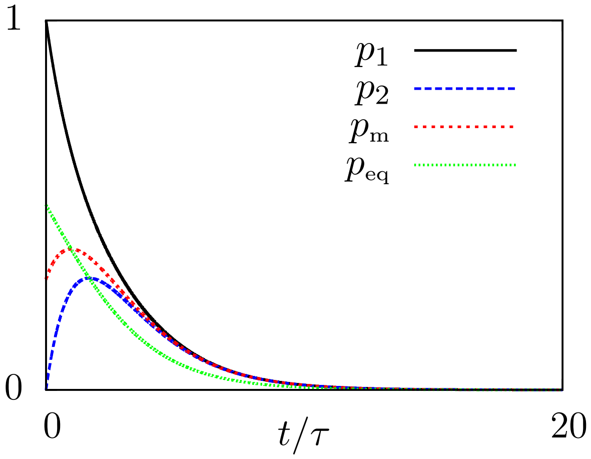

Figure 4.2: State of the memory as a function of time during the erasure process.

On figure (4.3) we plotted the time evolution of these quantities along the erasure process.

At time $t = 0$, both of them are equal to one. This is due to the fact that we chose non overlapping states to encode α_1 and α_2. In other words, at time $t = 0$, the micro-state of the memory contains all the information about the symbol recorded. After the beginning of the erasure process, $P(\alpha_1|\mathrm{L})$ and $P(\alpha_2|\mathrm{R})$ both drop below one. In fact, if the memory is in the left micro-state at some time $t > 0$ of the erasure process, it might be that α_1 was recorded, but it might also be that α_2 was recorded and that the memory made a transition from the right to the left micro-state. At the end of the erasure process, $P(\alpha_1|\mathrm{L})$ and $P(\alpha_2|\mathrm{R})$ respectively converge towards P_1 and P_2, the a priori probabilities that α_1, respectively α_2 was recorded. At that point, knowing the micro-state of the memory does not provide any information about the symbol that was recorded.

Information erasure and entropy production

The distance of the memory to equilibrium at time t of the erasure process is given by:

$$D(t) = P_1 D(p_1(t)\|p_{\mathrm{eq}}(E(t))) + P_2 D(p_2(t)\|p_{\mathrm{eq}}(E(t))). \tag{4.53}$$

The amount of information contained in the memory reads:

$$I(t) = P_1 D(p_1(t)\|p_{\mathrm{m}}(t)) + P_2 D(p_2(t)\|p_{\mathrm{m}}(t)). \tag{4.54}$$

On figure 4.4, we plotted the time evolution of these two quantities at the beginning of the erasure process. Clearly, the information content quickly decreases to

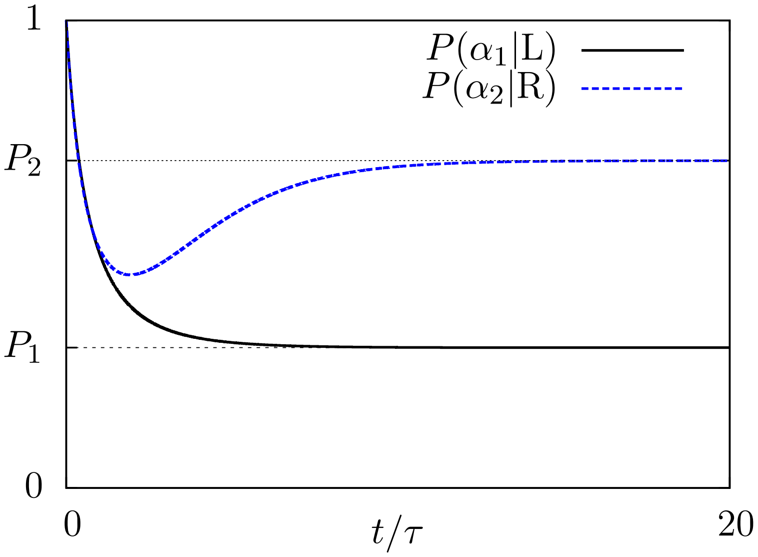

Figure 4.3: Conditional probabilities of the symbols given the micro-state of the memory along the erasure process. For $t = 0$ both $P(\alpha_1|\mathrm{R})$ and $P(\alpha_2|\mathrm{R})$ are equal to 1: Knowing the micro-state of the memory unambiguously yields the symbol recorded. For $t \gg \tau$, $P(\alpha_1|\mathrm{R}) = P_1$ and $P(\alpha_2|\mathrm{R}) = P_2$, i.e. knowing the micro-state of the memory does not provide any information about the symbol that was stored.

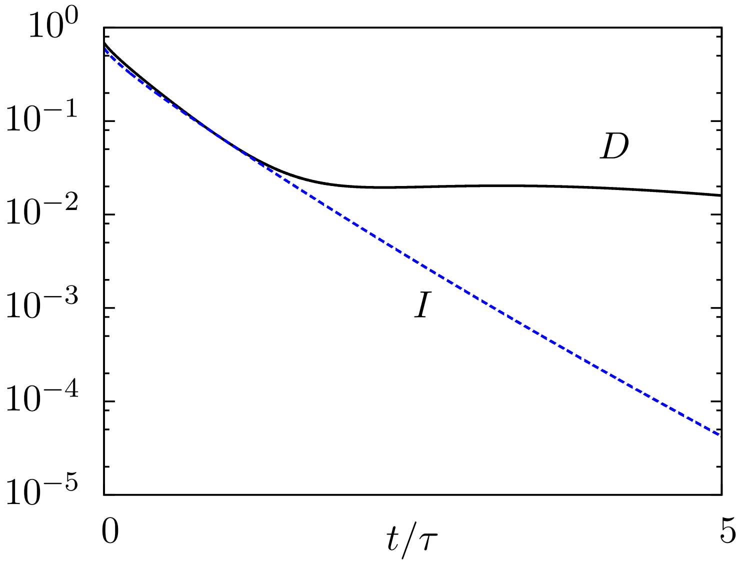

Figure 4.4: Distance to equilibrium and information as a function or time.

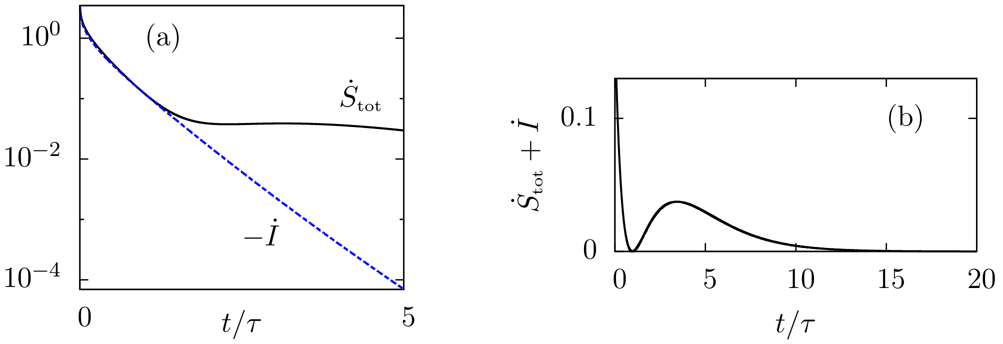

Figure 4.5: (a) Time evolution of the rates of entropy production, $\dot{S}_{\rm tot}$, and of information erasure $-\dot{I}$. (b) Difference between the two rates along the erasure process. The rates are given in units of τ^{-1}. As expected, the rate of entropy production is not less than the rate at which information is erased all along the process. At time $t_0 \simeq 0.97\tau$, the two rates are equal. This is due to the fact that at this time, the marginal distribution is equal to the equilibrium one, $p_{\rm m}(t_0) = p_{\rm eq}(E(t_0))$, as can be seen on figure 4.2.

zero. Moreover, the distance to equilibrium is always greater than the amount of information contained in the memory.

The rate of entropy production at time t of the erasure process is given by equation (4.46) averaged over the two possible scenarios:

$$\dot{S}_{\rm tot}(t) = -P_1\dot{p}_1(t)\frac{\partial}{\partial p_1}D(p_1(t)\|p_{\rm eq}(E(t))) - P_2\dot{p}_2(t)\frac{\partial}{\partial p_2}D(p_2(t)\|p_{\rm eq}(E(t))). \quad (4.55)$$

The rates of entropy production and information erasure are plotted on panel (a) of figure 4.5 and their difference on panel (b) of the same figure. As expected, the rate of entropy production is greater than the rate of information erasure. The total amount of entropy produced by the erasure until time t is just the integral of (4.55) between 0 and t:

$$\Delta S_{\rm tot}(t) = \int_0^t \dot{S}_{\rm tot}(t')\mathrm{d}t'. \quad (4.56)$$

On figure 4.6, we plotted the time evolution of the total amount of entropy production and of the amount of information erased $I_{\rm max} - I(t)$ along the erasure process. As expected, at any time, the amount of entropy produced is greater than the amount of information erased.

In the example treated here, we focused on the recording and the erasure processes only. However, in a practical situation, one would probably like to store the information for a finite amount of time before erasing it. This can be done by increasing τ, i.e. increasing the height of the potential barrier V between the two micro-states of the memory. In fact, the information decreases significantly only on a time scale

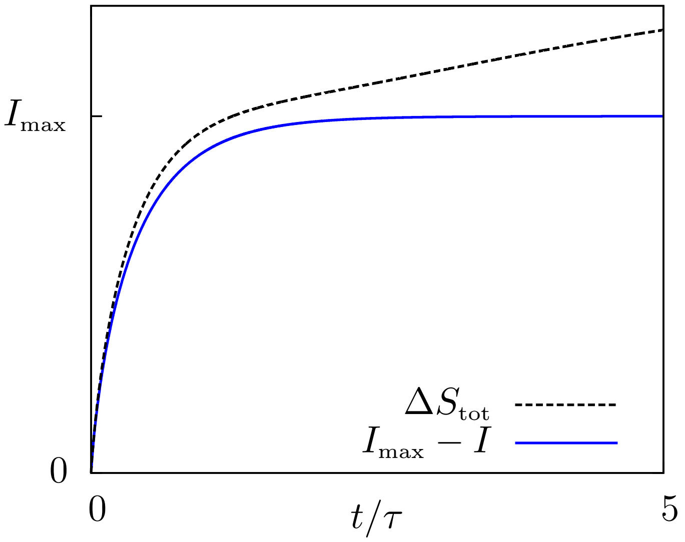

Figure 4.6: Entropy produced and information erased during the erasure process.

of order τ. Hence it is possible to store information for $t \ll \tau$.

4.5 Conclusion

In this chapter, stochastic thermodynamics was used to study one of the simplest information processing operations, namely the recording and the erasure of information. The two main results of this chapter are: i) The erasure of information is irreversible and the degree of irreversibility is directly related to the amount of information erased, see equation (4.17); ii) in order to contain some information, the device serving as a memory needs to be out of equilibrium, see equation (4.11).

The first step was to define what it means to record, store, and erase information. In other words, we had to find the requirements that should be satisfied by any physical implementation of these processes. The main requirement concerning the recording process is that, once the information is recorded, it should be possible to manipulate the memory without knowing what was recorded. In practice, this requirement implies that, at the end of the recording process, the value λ_{rec} of the control parameter should be independent of the symbol α_k that was recorded. In the same way, the erasure process should be performed in a way that does not depend on the information stored. In other words, during the erasure process, the time dependence of the control parameter should not depend on the symbol α_k originally recorded.

Next, we used stochastic thermodynamics to model the memory and to implement processes satisfying the aforementioned requirements. Finally, we had to identify the amount of information still present in the memory at any intermediate time of the erasure process. It was identified as the reduction in our uncertainty about the symbol that was originally recorded upon knowing the microscopic state of the memory at some intermediate time. In other words, the amount of information contained in the memory is the mutual information between the microscopic state of the memory and the symbol recorded.

The results presented in this chapter have their dual counterpart in the previous chapter. In fact, in the previous chapter, we saw that when a system is out of equilibrium, it means that we have more information about its micro-state than if the system was in equilibrium. In this chapter, we saw that in order to contain some information, the system serving as a memory needs to be out of equilibrium. Similarly, in the previous chapter, the rate of entropy production was shown to be equal to the rate at which we loose information about the micro-state of the system. In this chapter, we showed that when erasing information, one produces entropy at a rate at least equal to the rate at which information is erased. While in the previous chapter we saw how information theory can be relevant in order to understand non-equilibrium isothermal thermodynamics, in this chapter we showed how this very theory is relevant to understand the physics of information processing.

5 Acquisition of information

In section 3.3, page 50, we saw that if one is able to obtain information about the micro-state of a thermodynamic system, then this information can be converted into useful work. The acquisition and/or processing of this information must therefore involve costs that are at least high enough to compensate this information-to-work conversion.

The aim of this chapter is to investigate the process of obtaining information about the micro-state of a system. A minimal model for a measurement device is developed and the process of measuring the micro-state of another system is described. The measurement process is irreversible and the entropy produced is greater than the information obtained. The "reversible limit" where the entropy production is equal to the information obtained is reached when the two quantities are zero, i.e. when nothing happens.

5.1 Physical modelling of a measurement device

5.1.1 Measurement and information

Consider a system $\mathcal{S}$ in equilibrium with a heat bath at inverse temperature β. Let us assume for simplicity that the phase space of $\mathcal{S}$ is discrete. Let $p_{\mathcal{S}}(x)$ be the equilibrium probability that $\mathcal{S}$ occupies micro-state x. At some point, we perform a measurement of the micro-state of $\mathcal{S}$. The measurement is characterized by the conditional probability $p(y|x)$ to observe outcome y given that $\mathcal{S}$ is in micro-state x. This conditional distribution contains all the information about the measurement precision and the measurement errors. Typically, one could consider that a measurement was successful if $y = x$ and erroneous in the other cases. An error free measurement would be achieved if $p(y|x) = \delta_{x,y} = 1$ if $y = x$ and 0 otherwise. In general, a measurement is error free if the measurement outcome is a deterministic function of the micro-state of $\mathcal{S}$.

The amount of information about the micro-state of $\mathcal{S}$ provided by the measurement is given by the mutual information:

$$I = \sum_x p_{\mathcal{S}}(x) \sum_y p(y|x) \log \frac{p(y|x)}{p_{\mathcal{M}}(y)}, \tag{5.1}$$

where $p_{\mathcal{M}}(y) = \sum_x p(y|x)p_{\mathcal{S}}(x)$ is the marginal probability to observe the measurement outcome y. The information provided by the measurement can be exploited in order to transform heat into work in a cyclic isothermal process, as we mentioned in paragraph 3.3 page 50. The maximum amount of work extracted that way is TI.

Such a process is not possible without further entropy production. In fact, this process would lead the destruction of up to I of entropy. Hence, the thermodynamics tells us that the acquisition and/or processing of an amount I of information should be accompanied by the production of a corresponding amount of entropy. In the following, we develop a minimal model for a physical measurement device able to regularly provide an amount I of information. We compute the entropy produced in a cyclic measurement process and show that it is greater than I.

5.1.2 The measurement device

Any measurement device should satisfy a certain number of requirements. It should be a physical system subject to the laws of thermodynamics. In particular, measurement errors should include at least thermal fluctuations. The measurement device should be driven by the quantity to be measured, i.e. by the micro-state of the system $\mathcal{S}$. Finally, we assume that the measurement is "ideal" in that there is no back action of the measurement device on the system. This is certainly an idealization and it would be of great interest to investigate situations where it is not the case.

In order to meet these requirements, we assume that the measurement device is a thermodynamic system $\mathcal{M}$ in contact with a heat bath (not necessarily at the same temperature as $\mathcal{S}$). The energy $E_{\mathcal{M}}(y|x)$ of a micro-state y of $\mathcal{M}$ depends on the micro-state x of $\mathcal{S}$. Furthermore, we assume that the relaxation time $\tau_{\mathcal{M}}$ of $\mathcal{M}$ is much smaller than the time between two transitions of $\mathcal{S}$. That way we are sure that the measurement outcome y only depends on the quantity x to be measured. Hence, during the measurement, the micro-states of $\mathcal{M}$ are distributed according to the canonical distribution (2.60):

$$p(y|x) = \frac{\exp\left(-\beta_{\mathcal{M}} E_{\mathcal{M}}(y|x)\right)}{Z_{\mathcal{M}}(x)}, \tag{5.2}$$

where $Z_{\mathcal{M}}(x) = \sum_y \exp\left(-\beta E_{\mathcal{M}}(y|x)\right)$ is the partition function of the measurement device when $\mathcal{S}$ is in micro-state x and $\beta_{\mathcal{M}}$ is the inverse temperature of the heat bath $\mathcal{M}$ is in contact with. Finally, we assume that there is no direct back action of $\mathcal{M}$ on $\mathcal{S}$.

5.2 The measurement process

5.2.1 Information acquisition

The micro-state x of $\mathcal{S}$ plays the role of a control parameter for $\mathcal{M}$. If we let $\mathcal{M}$ in contact with $\mathcal{S}$, then every time $\mathcal{S}$ makes a transition form x to x', the energy levels of $\mathcal{M}$ are instantly modified from $E_{\mathcal{M}}(y|x)$ to $E_{\mathcal{M}}(y|x')$. We assume that $\mathcal{M}$ is in contact with a work source which provides the work needed to change the value of its energy levels. Before $\mathcal{S}$ has the time to make another transition, $\mathcal{M}$ relaxes from $p(y|x)$, which becomes non equilibrium, towards the new equilibrium distribution $p(y|x')$. This relaxation is accompanied by the production of an amount $\Delta S_{\text{tot}}(x, x')$ of entropy given by equation (3.88) on page 49, which, in our case, reads:

$$\Delta S_{\text{tot}}(x, x') = \sum_y p(y|x) \log \frac{p(y|x)}{p(y|x')}. \tag{5.3}$$

Hence, if $\mathcal{M}$ is left in contact with $\mathcal{S}$, it will produce entropy every time $\mathcal{S}$ makes a transition. This is the price to pay to continuously monitor the micro-state of $\mathcal{S}$. It can be shown that the joint system formed by $\mathcal{S}$ and $\mathcal{M}$ will never reach equilibrium because of the lack of back action of $\mathcal{M}$ on $\mathcal{S}$. However this is outside the scope of the present work.

We are not interested in continuously following the transitions of $\mathcal{S}$. We just wish to get some information about the micro-state of $\mathcal{S}$ at one particular moment and then exploit that information in a feed back process. In order to minimize the entropy production, we assume that we are able to "separate" $\mathcal{M}$ from $\mathcal{S}$. When $\mathcal{M}$ is separated from $\mathcal{S}$, its energy levels are just left as they are and do not change anymore. When we want to measure the micro-state of $\mathcal{S}$, we let $\mathcal{M}$ in contact with $\mathcal{S}$ for a time τ_{cont} which is such that:

- The probability that $\mathcal{S}$ makes a transition during τ_{cont} is vanishingly small.
- The measurement device has the time to relax towards equilibrium: $\tau_{\text{cont}} \gg \tau_{\mathcal{M}}$.

That way, $\mathcal{S}$ is left unchanged and at the end of the contact, the micro-state of $\mathcal{M}$ is correlated with the micro-state of $\mathcal{S}$. The mutual information between the micro-states of $\mathcal{M}$ and $\mathcal{S}$ is I given by equation (5.1). Once this information is obtained, it can be used in a cyclic feed-back process. At the end of this measurement and feed-back process, the micro-states of $\mathcal{M}$ and $\mathcal{S}$ are independent and we can perform a new cycle.

Before the contact, the micro-states of $\mathcal{M}$ are distributed according to $p(y|x_0)$ where x_0 is the micro-state $\mathcal{S}$ was in *during the previous measurement*. As a consequence, $\mathcal{M}$ is still correlated with the micro-state $\mathcal{S}$ was in during the previous measurement. In other words, $\mathcal{M}$ still contains some information about the previous micro-state of $\mathcal{S}$. This information was already used in the previous feed-back process and it cannot be used again, but it is still here. At the end of the contact, the micro-states of $\mathcal{M}$ are distributed according to the new equilibrium distribution $p(y|x_1)$, where x_1 is the current micro-state of $\mathcal{S}$. At that moment, the micro-state of $\mathcal{M}$ is correlated with the current micro-state of $\mathcal{S}$ and has no more information about the micro-state $\mathcal{S}$ was in during the previous cycle. The information about the past micro-state of $\mathcal{S}$ has been replaced by information about its current micro-state.

Let us summarize the cycle of measurement and feed-back.

1. At the beginning of the cycle, the energy of a micro-state y of $\mathcal{M}$ is $E_{\mathcal{M}}(y|x_0)$ where x_0 is the micro-state occupied by $\mathcal{S}$ during the previous measurement cycle. The probability that $\mathcal{M}$ occupies y is $p(y|x_0)$ which is also the equilibrium distribution with the energies $E_{\mathcal{M}}(y|x_0)$ given by equation (5.2). The probability that $\mathcal{S}$ was in micro-state x_0 during the previous measurement is $p_{\mathcal{S}}(x_0)$.

2. $\mathcal{M}$ is put in contact with $\mathcal{S}$. Its energies are instantaneously changed to $E_{\mathcal{M}}(y|x_1)$, where x_1 is the micro-state currently occupied by $\mathcal{S}$. The probability that $\mathcal{S}$ occupies the micro-state x_1 is $p_{\mathcal{S}}(x_1)$.

3. $\mathcal{M}$ relaxes towards the new equilibrium $p(y|x_1)$. The amount of entropy produced by the relaxation is $\Delta S_{\text{tot}}(x_0, x_1)$ given by equation (5.3). At the end of the relaxation, the information about the previous micro-state of $\mathcal{S}$ is replaced by information about its current micro-state.

4. $\mathcal{M}$ is separated from $\mathcal{S}$ so that it is not anymore affected by possible transitions of $\mathcal{S}$. The total duration of the contact was so short that $\mathcal{S}$ didn't make any transition during that time period.

5. A cyclic process which depends on the micro-state y of $\mathcal{M}$ is performed on $\mathcal{S}$. Such a process enables to convert up to TI of heat at temperature T into work. At the end of that process, $\mathcal{M}$ is still correlated with the micro-state $\mathcal{S}$ was in during step 3.

At the end of the cycle, the micro-state of $\mathcal{M}$ is uncorrelated with the micro-state of $\mathcal{S}$ and a new cycle of measurement and feed-back can start.

5.2.2 Entropy produced

Entropy is produced during step 3 when $\mathcal{M}$ relaxes towards the new equilibrium. The amount of entropy produced if $\mathcal{S}$ was in micro-state x_0 during the previous measurement and is in micro-state x_1 during the current one is $\Delta S_{\text{tot}}(x_0, x_1)$ given by equation (5.3). Since x_0 and x_1 are independent, the average amount of entropy produced is:

$$\Delta S_{\text{tot}} = \sum_{x_0,x_1} p_{\mathcal{S}}(x_0) p_{\mathcal{S}}(x_1) \Delta S_{\text{tot}}(x_0,x_1) = \sum_{x_0,x_1} p_{\mathcal{S}}(x_0) p_{\mathcal{S}}(x_1) \sum_y p(y|x_0) \log \frac{p(y|x_0)}{p(y|x_1)}. \tag{5.4}$$

We would like to compare this quantity to the amount I of information obtained by one measurement event.

Using the trivial identity

$$\log \frac{p(y|x_0)}{p(y|x_1)} = \log \frac{p(y|x_0)}{p_{\mathcal{M}}(y)} + \log \frac{p_{\mathcal{M}}(y)}{p(y|x_1)} \tag{5.5}$$

in the right hand side of equation (5.4), we can isolate two different contributions to ΔS_{tot}:

$$\Delta S_{\text{tot}} = \Delta S^1_{\text{tot}} + \Delta S^2_{\text{tot}}, \tag{5.6}$$

where

$$\Delta S^1_{\text{tot}} = \sum_{x_0} p_{\mathcal{S}}(x_0) \left(\sum_{x_1} p_{\mathcal{S}}(x_1) \right) \sum_y p(y|x_0) \log \frac{p(y|x_0)}{p_{\mathcal{M}}(y)} \tag{5.7}$$

and

$$\Delta S^2_{\text{tot}} = \sum_{x_1,y} p_{\mathcal{S}}(x_1) \left(\sum_{x_0} p(y|x_0) p_{\mathcal{S}}(x_0) \right) \log \frac{p_{\mathcal{M}}(y)}{p(y|x_1)}. \tag{5.8}$$

In the expressions above, $p_{\mathcal{M}}(y) = \sum_x p(y|x) p_{\mathcal{S}}(x)$ is the marginal probability that $\mathcal{M}$ is in state y. Let us now analyze these two contributions separately.

Since, $\sum_{x_1} p_{\mathcal{S}}(x_1) = 1$, the first contribution ΔS^1_{tot} is equal to:

$$\Delta S^1_{\text{tot}} = \sum_x p_{\mathcal{S}}(x_0) \sum_y p(y|x_0) \log \frac{p(y|x_0)}{p_{\mathcal{M}}(y)} = I. \tag{5.9}$$

This quantity is the mutual information between the micro-state of $\mathcal{S}$ during the previous measurement and the micro-state of $\mathcal{M}$ just before the contact. Furthermore, the quantity

$$\Delta S^1_{\text{tot}}(x) = \sum_y p(y|x) \log \frac{p(y|x)}{p_{\mathcal{M}}(y)}. \tag{5.10}$$

is the amount of entropy that would be produced if $\mathcal{M}$ relaxed from $p(y|x)$ to $p_{\mathcal{M}}(y)$, i.e. if the energies of $\mathcal{M}$ were instantaneously changed from $E_{\mathcal{M}}(y|x)$ to some $E^0_{\mathcal{M}}(y)$ such that the $p_{\mathcal{M}}(y)$ would be equilibrium with respect to $E^0_{\mathcal{M}}(y)$ and $\mathcal{M}$ would be let to relax towards this new equilibrium. Hence, the quantity (5.9) is the average amount of entropy produced if the energy levels of $\mathcal{M}$ were equal to $E_{\mathcal{M}}(y|x)$ with probability $p_{\mathcal{S}}(x)$ and they were instantaneously changed to $E^0_{\mathcal{M}}(y)$.

Similarly, taking into account that $\sum_{x_0} p(y|x_0)p_{\mathcal{S}}(x_0) = p_{\mathcal{M}}(y)$, the second contribution ΔS^2_{tot} can be rewritten as:

$$\Delta S^2_{\text{tot}} = \sum_{x_1} p_{\mathcal{S}}(x_1) \sum_y p_{\mathcal{M}}(y) \log \frac{p_{\mathcal{M}}(y)}{p(y|x_1)}. \tag{5.11}$$

The quantity

$$\Delta S^2_{\text{tot}}(x) = \sum_y p_{\mathcal{M}}(y) \log \frac{p_{\mathcal{M}}(y)}{p(y|x)} \geq 0. \tag{5.12}$$

is non negative and it is zero if and only if $p(y|x_1) = p_{\mathcal{M}}(y)$ because it is a Kullback-Leibler distance. Hence ΔS^2_{tot} is zero if and only if y and x_1 are independent, i.e. if the measurement does not provide any information. The quantity $\Delta S^2_{\text{tot}}(x_1)$ is the amount of entropy that would be produced in the relaxation from $p_{\mathcal{M}}(y)$ to $p(y|x_1)$, i.e. in the process were $\mathcal{M}$ starts in $p_{\mathcal{M}}(y)$ and its energy levels are instantaneously changed to $E_{\mathcal{M}}(y|x_1)$.

To summarize, the entropy production (5.4) is thus composed of two non negative contributions:

$$\Delta S_{\text{tot}} = \Delta S^1_{\text{tot}} + \Delta S^2_{\text{tot}}, \tag{5.13}$$

where

$$\Delta S^1_{\text{tot}} = \sum_{x_0} p_{\mathcal{S}}(x_0) \Delta S^1_{\text{tot}}(x_0) = I \tag{5.14}$$

is the information gained by the measurement and

$$\Delta S^2_{\text{tot}} = \sum_{x_1} p_{\mathcal{S}}(x_1) \Delta S^2_{\text{tot}}(x_1). \tag{5.15}$$

Hence, the average amount of entropy produced by the measurement device per cycle is strictly greater than the amount of information available at each cycle:

$$\Delta S_{\text{tot}} - I = \Delta S^2_{\text{tot}} \geq 0. \tag{5.16}$$

There is equality $\Delta S_{\text{tot}} = I$ if and only if $\Delta S^2_{\text{tot}} = 0$ which only happens if the information itself vanishes, i.e. $I = 0$.

5.2.3 Decorrelation, re-correlation

The measurement process described above is a particular case of random driving. The energy levels of $\mathcal{M}$, $E_{\mathcal{M}}(y|x)$ depend on the value x. At each time step, a new value of the control parameter x is randomly drawn from a distribution $p_{\mathcal{S}}(x)$ independently of its previous value. Then, $\mathcal{M}$ relaxes towards the new equilibrium

$p(y|x)$. At each time step, as $\mathcal{M}$ relaxes, it decorrelates from the previous value of x and correlates to the new one.

In the preceding paragraph, we showed that the average amount of entropy produced is the sum of two contributions. Each of these contributions is equal to the amount of entropy that would be produced in two different processes. In fact, $\Delta S^1_{\text{tot}}(x)$ is the entropy produced by $\mathcal{M}$ in the relaxation from $p(y|x)$ to $p_{\mathcal{M}}(y)$. Such a relaxation is obtained if x is instantaneously varied to a deterministic value $\tilde{x}$ which is such that $E_{\mathcal{M}}(y|\tilde{x}) = E^0_{\mathcal{M}}(y)$ independently of the initial value of x. In this process, just before $\mathcal{M}$ starts to relax, its micro-state is correlated with the initial value of the control parameter x. At the end of the process there are no more correlations since the final value of the control parameter is deterministic. Hence, the sole effect of this process was to destroy the initial correlations between y and x. Furthermore, we showed that the average amount of entropy produced is precisely the mutual information I between the initial micro-state of $\mathcal{M}$ and the initial value of the control parameter x, see equation (5.14). This result is reminiscent of the main result of the previous chapter: Information is erased and as a consequence, a corresponding amount of entropy is produced.

The quantity $\Delta S^2_{\text{tot}}(x)$ on the other hand is equal to the entropy that would be produced in the relaxation from $p_{\mathcal{M}}(y)$ to $p(y|x)$. This relaxation occurs if the control parameter is instantaneously driven from $\tilde{x}$ to x, i.e. if the energies of $\mathcal{M}$ are instantaneously driven from $E_{\mathcal{M}}(y|\tilde{x}) = E^0_{\mathcal{M}}(y)$ to $E_{\mathcal{M}}(y|x)$. If x is chosen randomly according to the distribution $p_{\mathcal{S}}(x)$, then at the end of the relaxation, the micro-state of $\mathcal{M}$ is correlated to this new value of the control parameter and their mutual information is I given by equation (5.1). Since the initial value of the control parameter is deterministic, the establishment of this correlation is the only effect of this process. In the limit of an error-free measurement, i.e. $p(y|x) = \delta_{x,y}$, ΔS^2_{tot} diverges. However, it is possible to find cases where ΔS^2_{tot} is smaller than the information I obtained[1].

The total entropy produced during the measurement event is the same as the entropy that would be produced if the two processes of decorrelation and re-correlation just described were operated separately. In the following, we analyze these processes of decorrelation and re-correlation in more details on a simple example.

[1] Here is an example of situation where $\Delta S^2_{\text{tot}} < I$. Imagine that $\mathcal{S}$ and $\mathcal{M}$ can occupy two micro-states, labelled by "L" and "R" as in the example described in section 5.3 below. Assume that $p_{\mathcal{S}}(\text{L}) = 1 - 10^{-4}$ (hence $p_{\mathcal{S}}(\text{R}) = 1 - p_{\mathcal{S}}(\text{L}) = 10^{-4}$), and $p(\text{L}|\text{L}) = 10^{-4}$ and $p(\text{L}|\text{R}) = 1 - 10^{-2}$. Then $p_{\mathcal{M}}(\text{L}) = p(\text{L}|\text{L})p_{\mathcal{S}}(\text{L}) + p(\text{L}|\text{R})p_{\mathcal{S}}(\text{R}) \approx 2 \cdot 10^{-4}$, see equation (5.18). One can check that for these numbers, $I \approx 8.6 \cdot 10^{-4}$ whereas $\Delta S^2_{\text{tot}} \approx 5.0 \cdot 10^{-4}$.

5.3 Example

5.3.1 Measuring the state of a two states system

As in the previous chapter, we now consider the most simple example in order to illustrate the results of the previous section. Let $\mathcal{S}$ be a system that can occupy two micro-states labelled by “L” and “R” with respective energies $E_{\mathcal{S}}(\mathrm{R})$ and $E_{\mathcal{S}}(\mathrm{L})$. For simplicity, we assume $E_{\mathcal{S}}(\mathrm{R}) = E_{\mathcal{S}}(\mathrm{L})$, so that at equilibrium the two states are equally probable: $p_{\mathcal{S}}(\mathrm{L}) = p_{\mathcal{S}}(\mathrm{R}) = 1/2$. The Shannon entropy of $p_{\mathcal{S}}$, quantifying the information we miss about the micro-state of $\mathcal{S}$, is equal to $\log 2$.

Our aim is to measure the state of $\mathcal{S}$. Let p be the probability of a successful measurement. The probability of measurement error is then $1-p$. The conditional probability $p(y|x)$ to observe the measurement outcome $y \in \{\mathrm{L},\mathrm{R}\}$ given that $\mathcal{S}$ is in micro-state $x \in \{\mathrm{L},\mathrm{R}\}$ is:

$$p(y|x) = \begin{cases} p & \text{if} \quad y = x \\ 1-p & \text{if} \quad y \neq x. \end{cases} \tag{5.17}$$

The marginal probability to observe $y = \mathrm{L}$ as a measurement outcome is given by:

$$p_{\mathcal{M}}(\mathrm{L}) = p(\mathrm{L}|\mathrm{L})p_{\mathcal{S}}(\mathrm{L}) + p(\mathrm{L}|\mathrm{R})p_{\mathcal{S}}(\mathrm{R}). \tag{5.18}$$

Inserting (5.17) in the expression above yields:

$$p_{\mathcal{M}}(\mathrm{L}) = \frac{p}{2} + \frac{1-p}{2} = \frac{1}{2}. \tag{5.19}$$

As a consequence, the marginal probability to observe $y = \mathrm{R}$ as a measurement outcome is equal to $p_{\mathcal{M}}(\mathrm{R}) = 1 - p_{\mathcal{M}}(\mathrm{L}) = 1/2$. The mutual information between the measurement outcome and the micro-state of $\mathcal{S}$ given by equation (5.1) becomes in this case:

$$I(p) = p \log \frac{p}{1/2} + (1-p) \log \frac{1-p}{1/2} \leq \log 2. \tag{5.20}$$

This quantity is zero for $p = 1/2$. In this case, x and y are completely independent. For $p = 1$, $I(1) = \log 2$, the maximum amount of information we can obtain about the micro-state of $\mathcal{S}$. In fact, in this case, knowing the measurement outcome we are able to infer the micro-state of x with certainty.

Since there are two possible measurement outcomes, the measurement device $\mathcal{M}$ should be able to occupy at least two micro-states. Its energy levels should be such that $p(y|x)$ is the equilibrium distribution. Let $E_{\mathrm{suc}} = E_{\mathcal{M}}(\mathrm{R}|\mathrm{R}) = E_{\mathcal{M}}(\mathrm{L}|\mathrm{L})$ be the energy of the micro-state of $\mathcal{M}$ corresponding to a successful measurement and $E_{\mathrm{err}} = E_{\mathcal{M}}(\mathrm{L}|\mathrm{R}) = E_{\mathcal{M}}(\mathrm{R}|\mathrm{L})$ be the energy of the micro-state of $\mathcal{M}$ corresponding to an erroneous measurement. For $p(y|x)$ to be the equilibrium distribution of $\mathcal{M}$, we need:

$$\Delta E = E_{\text{err}} - E_{\text{suc}} = T_{\mathcal{M}} \log \frac{p}{1-p}, \tag{5.21}$$

where $T_{\mathcal{M}} = 1/\beta_{\mathcal{M}}$ is the temperature of the heat bath $\mathcal{M}$ is in contact with. Remark that here as in the previous chapter, if we want an error free measurement, $p = 1$, we need an infinite energy difference between the two micro-states.

5.3.2 Entropy produced by the measurement

As in the general case, we consider the following cycle. Initially, $\mathcal{S}$ and $\mathcal{M}$ are uncorrelated and the energies of the measurement device are as we left them at the end of the previous measurement. They are equal to $E(y|\text{L})$ or $E(y|\text{R})$ with equal probability $1/2$ depending on whether $\mathcal{S}$ was in micro-state L or R during the previous measurement. At some point, we put $\mathcal{M}$ in contact with $\mathcal{S}$ and the energy levels of $\mathcal{M}$ instantaneously adapt to the new micro-state of $\mathcal{S}$. Here there are two possibilities: Either the micro-state of $\mathcal{S}$ is the same as during the previous measurement, in which case the energies of $\mathcal{M}$ do not change, or the micro-state of $\mathcal{S}$ is not the same, in which case the energies of the micro-states of $\mathcal{M}$ are simply exchanged. Each of the two possibilities is equally probable.

The amount of entropy produced in the case where the micro-state of $\mathcal{S}$ is different during the two consecutive cycles is:

$$\Delta S^0_{\text{tot}}(p) = p \log \frac{p}{1-p} + (1-p) \log \frac{1-p}{p}. \tag{5.22}$$

In the case where the micro-states of $\mathcal{S}$ are the same during the two consecutive measurement, no entropy is produced. The two scenarios being equally probable, the average amount of entropy produced during one measurement event is half of the quantity above:

$$\Delta S_{\text{tot}}(p) = \frac{1}{2}\Delta S^0_{\text{tot}}(p) = \frac{1}{2}\left(p \log \frac{p}{1-p} + (1-p) \log \frac{1-p}{p}\right). \tag{5.23}$$

This quantity can be decomposed into two non negative parts:

$$\Delta S_{\text{tot}}(p) = \Delta S^1_{\text{tot}}(p) + \Delta S^2_{\text{tot}}(p), \tag{5.24}$$

where $\Delta S^1_{\text{tot}}(p) = I(p)$ is the information provided by the measurement, given by equation (5.20) and $\Delta S^2_{\text{tot}}(p)$ is given by:

$$\Delta S^2_{\text{tot}}(p) = \frac{1}{2}\left(\log \frac{1/2}{p} + \log \frac{1/2}{1-p}\right). \tag{5.25}$$

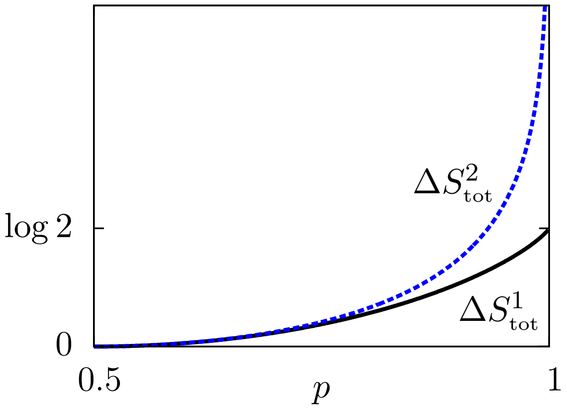

Figure 5.1: The two contributions ΔS^1_{tot} and ΔS^2_{tot} to the total entropy produced by the measurement device during one measurement event as a function of the probability of a successful measurement p. The first contribution is equal to the mutual information (5.20) $\Delta S^1_{\text{tot}} = I(p)$ and as we can see, the second contribution satisfies $\Delta S^2_{\text{tot}} \geq I(p)$ with equality if and only if the two are zero.

Both $\Delta S^1_{\text{tot}}(p)$ and $\Delta S^1_{\text{tot}}(p)$ are non negative and are zero if and only if $p = 1/2$. In this case, the measurement device is not touched and the measurement provides no information. Moreover, as can be seen on figure 5.1, $\Delta S^2_{\text{tot}}(p) \geq I(p)$ with equality if and only $p = 1/2$, i.e. if $\Delta S^1_{\text{tot}}(p) = \Delta S^2_{\text{tot}}(p) = 0$.

As already discussed in paragraph 5.2.3, page 82, $\Delta S^1_{\text{tot}}(p)$ can be seen as the amount of entropy produced because of erasure of the information about the micro-state of $\mathcal{S}$ during the previous measurement and $\Delta S^2_{\text{tot}}(p)$ as the entropy produced because of the acquisition of new information about the current micro-state of $\mathcal{S}$. In this particular example, the amount of entropy $\Delta S^2_{\text{tot}}(p)$ produced because of the acquisition of new information is greater than the amount of information $I(p)$ acquired.

5.3.3 Decorrelation, re-correlation

State of $\mathcal{M}$ during the measurement event

Let us resolve the infinitely fast relaxation of the measurement device happening during one measurement event. The time evolution of $\mathcal{M}$ is given by the master equations (4.42). The probability per unit time that $\mathcal{M}$ makes a transition from y to y' given that $\mathcal{S}$ is in micro-state x is given by:

$$w_{y'y}(x) = \frac{1}{\tau_{\mathcal{M}}} \exp\left(\beta E_{\mathcal{M}}(y|x)\right), \tag{5.26}$$

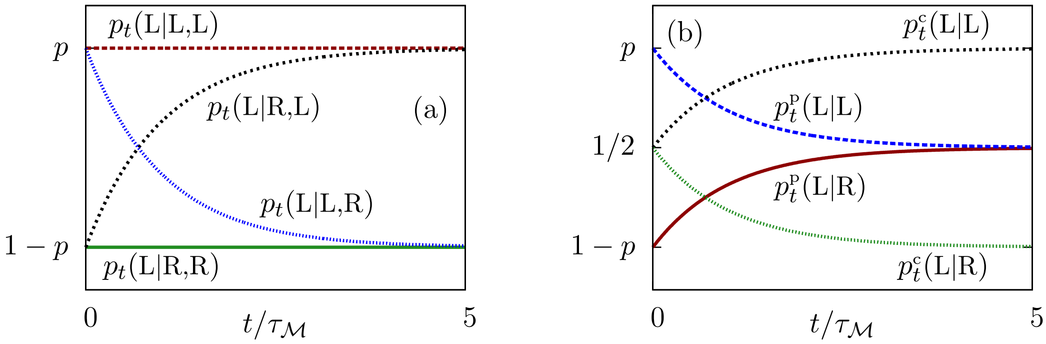

Figure 5.2: (a) Time evolution of the probability $p_t(\mathrm{L}|x_0, x_1)$ for $\mathcal{M}$ to be in micro-state 'L' conditioned on the micro-states x_0 and x_1 $\mathcal{S}$ during the previous and the current measurement cycle. There are four different possibilities. (b) Time evolution of the probability $p_t^{\mathrm{p}}(\mathrm{L}|x_0)$ for $\mathcal{M}$ to be in micro-state 'L' conditioned on the micro-state x_0 of $\mathcal{S}$ during the previous measurement cycle, and of the probability $p_t^{\mathrm{c}}(\mathrm{L}|x_1)$ of the same event, but conditioned on the micro-state x_1 of $\mathcal{S}$ during the current measurement cycle.

where $\tau_{\mathcal{M}}$ is the relaxation time of the measurement device. The probability $p_t(y)$ that $\mathcal{M}$ is in micro-state y at time t given that $\mathcal{S}$ is in micro-state x evolves according to the master equation:

$$\dot{p}_t(y) = \sum_{y'} w_{yy'}(x) p_t(y'). \tag{5.27}$$

For x fixed, this distribution relaxes exponentially fast towards the equilibrium distribution:

$$p_t(y) = \exp\left(-\frac{t}{\tau_{\mathcal{M}}}\right)(p_0(y) - p_{\mathrm{eq}}(y)) + p_{\mathrm{eq}}(y), \tag{5.28}$$

where $p_0(y)$ is the initial probability that $\mathcal{M}$ is in state y.

If $\mathcal{S}$ was in micro-state x_0 during the previous measurement and is in micro-state x_1 during the current one, then the initial distribution of the micro-states of $\mathcal{M}$ is $p(y|x_0)$ and the equilibrium (and hence final) distribution is $p(y|x_1)$. Hence, the probability that $\mathcal{M}$ is in micro-state y at time t is given by:

$$p_t(y|x_0, x_1) = \exp\left(-\frac{t}{\tau_{\mathcal{M}}}\right)(p(y|x_0) - p(y|x_1)) + p(y|x_1), \tag{5.29}$$

If $x_0 = x_1 = x$, then $p_t(y|x_0, x_1) = p(y|x)$ is constant and if $x_0 \neq x_1$, then $p_t(y|x_0, x_1)$ relaxes from $p(y|x_0)$ to $p(y|x_1)$ exponentially fast. The four different possibilities for the evolution of $p_t(\mathrm{L}|x_0, x_1)$ are plotted on figure 5.2 (a).

Information contained in $\mathcal{M}$

At time t $\tau_{\mathcal{M}}$ of this measurement event, $\mathcal{M}$ still has some information about the micro-state of $\mathcal{S}$ during the previous measurement and already has some information about the current micro-state of $\mathcal{S}$. We want to quantify these amounts of information and analyze their time evolution.

The amount of information $\mathcal{M}$ still has about the previous micro-state of $\mathcal{S}$ at time t is quantified by the mutual information I_t^{p} between the micro-state of $\mathcal{M}$ and the previous micro-state of $\mathcal{S}$:

$$I_t^{\mathrm{p}} = \sum_{x_0} p_{\mathcal{S}}(x_0) \sum_{y} p_t^{\mathrm{p}}(y|x_0) \log \frac{p_t^{\mathrm{p}}(y|x_0)}{p_t^{\mathcal{M}}(y)}, \tag{5.30}$$

where $p_t^{\mathrm{p}}(y|x_0) = \sum_{x_1} p_t(y|x_0, x_1) p_{\mathcal{S}}(x_1)$ is the conditional probability that $\mathcal{M}$ occupies micro-state y at time t given that $\mathcal{S}$ was in x_0 during the previous measurement, and $p_t^{\mathcal{M}}(y) = \sum_{x_0,x_1} p_t(y|x_0, x_1) p_{\mathcal{S}}(x_0) p_{\mathcal{S}}(x_1)$ is the marginal probability that $\mathcal{M}$ occupies micro-state y at time t. The latter is constant equal to $1/2$ because of the symmetries of the problem. Similarly, the amount of information I_t^{c} that $\mathcal{M}$ already has about the current micro-state of $\mathcal{S}$ at time t is given by:

$$I_t^{\mathrm{c}} = \sum_{x_1} p_{\mathcal{S}}(x_1) \sum_{y} p_t^{\mathrm{c}}(y|x_1) \log \frac{p_t^{\mathrm{c}}(y|x_1)}{p_t^{\mathcal{M}}(y)}, \tag{5.31}$$

where $p_t^{\mathrm{c}}(y|x_1) = \sum_{x_0} p_t(y|x_0, x_1) p_{\mathcal{S}}(x_0)$ is the conditional probability that $\mathcal{M}$ occupies micro-state y at time t given that the current micro-state of $\mathcal{S}$ is x_1.

The quantities $p_t^{\mathrm{p}}(\mathrm{L}|x_0)$ and $p_t^{\mathrm{c}}(\mathrm{L}|x_1)$ are plotted on figure 5.2 (b). The probabilities $p_t^{\mathrm{p}}(\mathrm{L}|x_0)$ conditioned on the previous micro-state of $\mathcal{S}$ start at p (for $x_0 = \mathrm{L}$) and $1 - p$ (for $x_0 = \mathrm{R}$) and converge towards $1/2$. The information $\mathcal{M}$ has about the previous micro-state of $\mathcal{S}$ decreases as they approach this value. On the other hand, the probabilities $p^{\mathrm{c}}(\mathrm{L}|x_1)$ conditioned on the current micro-state of $\mathcal{S}$ start at $1/2$ and converge towards p (for $x_1 = \mathrm{L}$) and $1 - p$ (for $x_1 = \mathrm{R}$). The information $\mathcal{M}$ has about the current micro-state of $\mathcal{S}$ increases as they converge. The time evolution of I_t^{p} and I_t^{c} is plotted on figure 5.3 (a). As expected, the information $\mathcal{M}$ has about the previous micro-state of $\mathcal{S}$ decreases from $I(p)$ to 0 while the information $\mathcal{M}$ has about the current one increases from 0 to $I(p)$. Figure 5.3 (b) shows that for this specific example, the rate of entropy $\dot{S}_{\mathrm{tot}}$ production is greater than the rate $\dot{I}_t^{\mathrm{c}}$ at which information is obtained about the current micro-state of $\mathcal{S}$ *plus* the rate $-\dot{I}_t^{\mathrm{p}}$ at which information about the previous micro-state of $\mathcal{S}$ is lost.

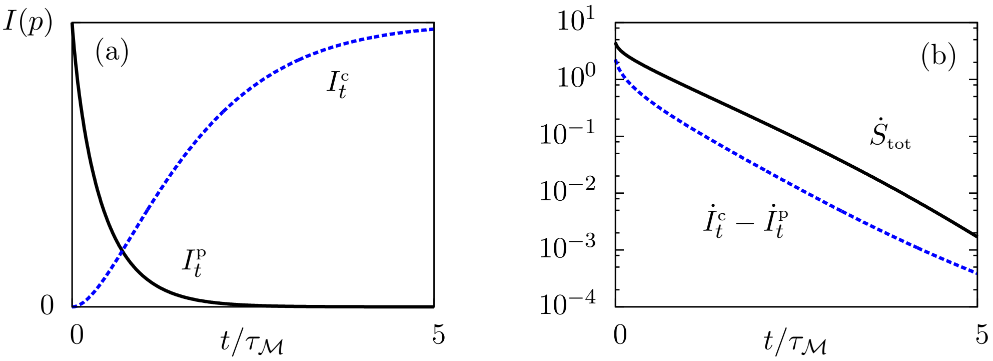

Figure 5.3: (a) Time evolution of the mutual information between the micro-state of $\mathcal{M}$ and the micro-state of $\mathcal{S}$ during the previous measurement, I_t^{p} (black solid line) and the current micro-state of $\mathcal{S}$, I_t^{c} (blue dashed line). (b) Rate of entropy production (black solid line) and total amount of information processed per unit time (blue dashed line). The latter is the sum of the rate at which information about the previous micro-state of $\mathcal{S}$ is lost and the rate at which new information about its current micro-state is gained. In this specific example, the entropy production is greater than this sum.

5.4 Conclusion

In this chapter, we tried to understand the thermodynamic costs associated with the acquisition of information about the microscopic state of a system in contact with a heat bath. The measurement device was modelled as physical system also in contact with a heat bath (not necessarily the same than the original system). The assumptions we made about the measurement device are: i) it should receive information from the original system, and ii) it should relax infinitely fast. Furthermore, we assumed no back action of the measurement device onto the system.

Under these assumptions, we showed that two processes occur simultaneously during a measurement cycle. On the one hand, the measurement device loses the information it had about the previous measurement cycle and, on the other hand, it gains information about the current cycle. Each of these processes is entropy producing. The entropy production due to the loss of information is equal to the amount of information that was lost in a way that is reminiscent of the previous chapter. The entropy production due to the gain of information, however, is not necessarily related to the amount of information gained. In the limit of an error-free measurement, though, it diverges.

The measurement process turned out to be a particular case of a random driving. In fact, the micro-state x of the system plays the role of a control parameter for the measurement device and it is randomly changed in a way such that to subsequent values are independent and that the measurement device has the time to relax towards the new (random) equilibrium. Hence, it would be interesting to investigate the entropy produced by a system subject to an arbitrary random driving and check whether the results obtained in this chapter are still valid. The thermodynamics of randomly driven system was already addressed in [SSBC12, BHS12, DE13], but no attempt was made to identify the amount of information gained and lost and to relate them to the entropy production.

The next step in order to understand the thermodynamics of measurement is to explicitly model the interaction between the system and the measurement device and to include back action of the measurement device onto the system. Such a model was developed in [SSBE13] for a particular system involving two coupled quantum dots. Interestingly, in this work, the rate of entropy production diverges in the error-free limit. However, there exist no general model yet.

6 Conclusion

Irreversibility and information are intimately related. The work presented here aimed at quantifying and exploiting this relation. The main result of this work is that whenever information is lost, there is irreversibility involved. The information loss could be intended as in chapter 4 or might be a by-product of some other process as in chapter 5.

In order to derive the main results of this thesis, it was first necessary to clarify the role of information in equilibrium and non-equilibrium thermodynamics. The thermodynamic description of a system misses a lot of information about the microscopic components of that system. We saw that this missing information is captured in the thermodynamic entropy of the system. Moreover, having more information about the microscopic state of a system than the specification of its thermodynamic state would provide amounts to say that the system is out of equilibrium. The assumption that a non-equilibrium system eventually relaxes towards equilibrium then implies that the extra information contained in the non-equilibrium state eventually gets lost.

The theory of stochastic thermodynamics successfully formalizes this idea in a way that is compatible with isothermal thermodynamics. This theory is the simplest extension of equilibrium statistical mechanics to non-equilibrium isothermal processes. It simply assumes that when a system is coupled to an equilibrium heat bath, it relaxes towards equilibrium according to a linear equation. Stochastic thermodynamics successfully generalizes various equilibrium state functions (like the entropy or the free energy) to non-equilibrium states and it provides an explicit expression for the entropy production. Using the tools of information theory allowed us to identify the entropy production of stochastic thermodynamics as the amount of information about the microscopic state of the system, that gets lost in the relaxation.

The theory of stochastic thermodynamics was then used to investigate the thermodynamics of information processing. We focused on two simple operations, namely the recording and erasure of information on a physical memory and the acquisition of information by a physical measurement device. The question was whether these operations can be performed in a reversible way, and if not, whether there is any link between the amount of entropy produced and the amount of information processed.

In chapter 4, we investigated the recording and the erasure of information on a physical memory [GK13]. We first had to clarify what it actually means to record and erase information. In other words, we had to find the requirements that should be obeyed by any physical implementation of the recording and erasure of information. We then used stochastic thermodynamics to implement processes meeting these requirements. Next, we identified the amount of information that is present in the memory all along the erasure process. Finally, we could show that the rate at which the information decreases during the erasure is a lower bound to the rate at which thermodynamic entropy is produced. Moreover, we showed that, as long as the memory contains some information, it is out of equilibrium to an extend linked with the amount of information still present.

In chapter 5, we developed a simple model for a measurement device and for the process of measuring the micro-state of some other system [GK11]. As in the preceding chapter, we tried to find the minimal requirements that any measurement device should satisfy. We then used stochastic thermodynamics to compute the entropy that is produced in a cyclic measurement process. We showed that the entropy production had two different non-negative contributions, one coming from the loss of information about the previous measurement cycle, and one due to the gain of information about the current cycle. These two contributions appear to behave differently. The former is equal to the amount of information that is lost, in a way reminiscent of chapter 5. The latter can be smaller than the information gained, but it diverges in the error-free limit.

The link between information gain or loss and irreversibility can be (and will be) exploited further at the fundamental level as well as at the applied level.

A fundamental issue that was not addressed in this work is the microscopic expression for the entropy production in adiabatic transformation. In fact, throughout this work, we only considered transformations operated on systems in contact with an equilibrium heat bath. The irreversibility always relied on the fact that any non-equilibrium system relaxes towards equilibrium through heat exchanges with the bath. During an adiabatic process, however, the system exchanges only work with its environment and its evolution is dictated by the Hamiltonian equations. There is currently no theory explaining entropy production based on Hamiltonian dynamics. The results of stochastic thermodynamics now help us to understand the difficulty. In fact, Hamiltonian dynamics conserves the Shannon entropy. In other words, the information we have about the microscopic state of the system does not change under Hamiltonian evolution. In order to understand the microscopic origin of entropy production, one could search for reasons why the information one has about the microscopic state of a thermally isolated system gets lost, despite Hamiltonian evolution.

On the other hand, the thermodynamics of information processing might find interesting applications in biological systems. In fact, biological systems are a preferred field of application of stochastic thermodynamics, since many biological process operate in solution, at a constant temperature. Moreover, the acquisition, the transfer, or the erasure of information are task commonly performed even by the simplest living organisms. The results obtained in this work might be a first step towards the understanding of the energetic costs involved in biological information processing.

Bibliography

[BAP$^+$12] Antoine Bérut, Artak Arakelyan, Artyom Petrosyan, Sergio Ciliberto, Raoul Dillenschneider, and Eric Lutz. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature*, 483(7388):187–189, March 2012.

[BHS12] A. C. Barato, D. Hartich, and U. Seifert. Information-theoretic vs. thermodynamic entropy production in autonomous sensory networks. *arXiv:1212.3186*, December 2012.

[Cal85] Herbert B. Callen. *Thermodynamics and introduction to Thermostatistics*. John Wiley & Sons, 2$^{\text{nd}}$ edition, August 1985.

[Cro98] Gavin E. Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *Journal of Statistical Physics*, 90(5-6):1481–1487, March 1998.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, July 2006.

[DBE13] Giovanni Diana, Baris G. Bagci, and Massimiliano Esposito. Finite-time erasing of information stored in fermionic bits. *Physical Review E*, 87(1):012111, January 2013.

[DE13] Giovanni Diana and Massimiliano Esposito. Mutual entropy-production and sensing in bipartite systems. arXiv e-print 1307.4728, July 2013.

[DL09] Raoul Dillenschneider and Eric Lutz. Memory erasure in small systems. *Physical Review Letters*, 102(21):210601, May 2009.

[ES12] Massimiliano Esposito and Gernot Schaller. Stochastic thermodynamics for "Maxwell demon" feedbacks. *EPL (Europhysics Letters)*, 99(3):30003, August 2012.

[Esp12] Massimiliano Esposito. Stochastic thermodynamics under coarse graining. *Physical Review E*, 85(4):041125, April 2012.

[EVdB10] Massimiliano Esposito and Christian Van den Broeck. Three faces of the second law. I. Master equation formulation. *Physical Review E*, 82(1):011143, July 2010.

[EVdB11] M. Esposito and C. Van den Broeck. Second law and Landauer principle far from equilibrium. *EPL (Europhysics Letters)*, 95(4):40004, August 2011.

[Gib02] J. Willard (Josiah Willard) Gibbs. *Elementary principles in statistical mechanics : developed with especial reference to the rational foundation of thermodynamics.* New York : C. Scribner, 1902.

[GK11] Léo Granger and Holger Kantz. Thermodynamic cost of measurements. *Physical Review E*, 84(6):061110, December 2011.

[GK13] Léo Granger and Holger Kantz. Differential Landauer's principle. *EPL (Europhysics Letters)*, 101(5):50004, March 2013.

[GS97] Bernard Gaveau and L.S. Schulman. A general framework for non-equilibrium phenomena: the master equation and its formal consequences. *Physics Letters A*, 229(6):347–353, June 1997.

[HITD10] H.-H. Hasegawa, J. Ishikawa, K. Takara, and D.J. Driebe. Generalization of the second law for a nonequilibrium initial state. *Physics Letters A*, 374(8):1001–1004, February 2010.

[HP11a] Jordan M Horowitz and Juan M R Parrondo. Designing optimal discrete-feedback thermodynamic engines. *New Journal of Physics*, 13(12):123019, December 2011.

[HP11b] Jordan M. Horowitz and Juan M. R. Parrondo. Thermodynamic reversibility in feedback processes. *EPL (Europhysics Letters)*, 95(1):10005, July 2011.

[Jay57] E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620–630, May 1957.

[Khi57] A. I. Khinchin. *Mathematical Foundations of Information Theory.* Dover Publications, Incorporated, 1957.

[Lan61] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183 –191, July 1961.

[Lan94] R. Landauer. Zig-zag path to understanding [physical limits of information handling]. In *Physics and Computation, 1994. PhysComp '94, Proceedings.* , pages 54 –59, November 1994.

[LR02] Harvey S Leff and Andrew F Rex. *Maxwell's demon 2 : entropy, classical and quantum information, computing.* Institute of Physics, Bristol, 2002.

[LY99] Elliott H. Lieb and Jakob Yngvason. The physics and mathematics of the second law of thermodynamics. *Physics Reports*, 310(1):1–96, March 1999.

[Max71] James Clerk Maxwell. *Theory of Heat.* Longmans, 1871.

[OLT$^+$12] Alexei O. Orlov, Craig S. Lent, Cameron C. Thorpe, Graham P. Boechler, and Gregory L. Snider. Experimental test of Landauer's principle at the sub-$k_{\mathrm{B}}T$ level. *Japanese Journal of Applied Physics*, 51:06FE10, 2012.

[Pie00] Barbara Piechocinska. Information erasure. *Physical Review A*, 61(6):062314, May 2000.

[Sch76] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern Physics*, 48(4):571–585, October 1976.

[Sei08] U. Seifert. Stochastic thermodynamics: principles and perspectives. *The European Physical Journal B*, 64(3-4):423–431, January 2008.

[Sei12] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12):126001, December 2012.

[Sha48] Claude E Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, July 1948.

[Shi95] Kousuke Shizume. Heat generation required by information erasure. *Physical Review E*, 52(4):3495–3499, October 1995.

[SSBC12] Susanne Still, David A. Sivak, Anthony J. Bell, and Gavin E. Crooks. Thermodynamics of prediction. *Physical Review Letters*, 109(12):120604, September 2012.

[SSBE13] Philipp Strasberg, Gernot Schaller, Tobias Brandes, and Massimiliano Esposito. Thermodynamics of a physical model implementing a Maxwell demon. *Physical Review Letters*, 110(4):040601, January 2013.

[SU09] Takahiro Sagawa and Masahito Ueda. Minimal energy cost for thermodynamic information processing: Measurement and information erasure. *Physical Review Letters*, 102(25):250602, June 2009.

[SU10] Takahiro Sagawa and Masahito Ueda. Generalized jarzynski equality under nonequilibrium feedback control. *Physical Review Letters*, 104(9):090602, March 2010.

[SU12a] Takahiro Sagawa and Masahito Ueda. Fluctuation theorem with information exchange: Role of correlations in stochastic thermodynamics. *Physical Review Letters*, 109(18):180602, November 2012.

[SU12b] Takahiro Sagawa and Masahito Ueda. Nonequilibrium thermodynamics of feedback control. *Physical Review E*, 85(2):021104, February 2012.

[SU13] Takahiro Sagawa and Masahito Ueda. Role of mutual information in entropy production under information exchanges. arXiv e-print 1307.6092, July 2013.

[Szi29] L. Szilard. Über die Entropieverminderung in einem thermodynamischen System bei eingriffen intelligenter Wesen. *Zeitschrift für Physik*, 53(11):840–856, 1929.

[THD10] K. Takara, H.-H. Hasegawa, and D.J. Driebe. Generalization of the second law for a transition between nonequilibrium states. *Physics Letters A*, 375(2):88–92, December 2010.

[Tho74] William Thomson. Kinetic theory of the dissipation of energy. *Nature*, 9:441–444, April 1874.

[Tho79] William Thomson. The sorting demon of Maxwell. *Proceedings of the Royal Institution*, 9:113, 1879.

[VdBE10] Christian Van den Broeck and Massimiliano Esposito. Three faces of the second law. II. Fokker-Planck formulation. *Physical Review E*, 82(1):011144, July 2010.

Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.
Die Arbeit wurde am Max-Planck-Institut für Physik komplexer Systeme in der Arbeitsgruppe „Nicht-lineare Dynamik und Zeitreihen Analyse“ angefertigt und von Prof. Dr. Holger Kantz betreut.
Ich erkenne die Promotionsordnung der Fakultät Mathematik und Naturwissenschaften der Technischen Universität Dresden vom 23.02.2011 an.

Léo Granger