# LECTURE NOTES ON

# NON-EQUILIBIUM FLUCTUATIONS

LUCA PELITI
SMRI (Italy)
`luca@peliti.org`

LECTURES GIVEN AT THE WORKSHOP "THEORETICAL METHODS FOR STUDYING NONEQUILIBRIUM FLUCTUATIONS", ORSAY (FRANCE), JUNE 8–9, 2017

© Luca Peliti, 2017

# Contents

# Preface

These are my notes for the three lectures that I gave during the Workshop "Theoretical methods for studying nonequilibrium fluctuations", held in Orsay (France), June 8–9, 2017. I am deeply grateful to the organizers, David Lacoste and Gatien Verley, for having organized this very interesting workshop. I could take part in the Workshop while visiting the PCT-Gulliver Laboratory at the École Supérieure de Physique et Chimie Industrielles in Paris, within the Labex CalTisPhysBio. I thank the Director of the Laboratory, Élie Raphaël, and the Director of the École, Jean-François Joanny, for their hospitality. I hope that these notes can be a useful reminder for those who have taken part in the lectures. I apologize for all typos and more serious mistakes that have crept in.

Paris, June 2017

## CHAPTER 1

# An introduction to stochastic thermodynamics and the thermodynamics of information

## 1.1 Equilibrium thermodynamics: A reminder

I start with a brief reminder of equilibrium thermodynamics, essentially for fixing notations.

Let us consider a *macroscopic* system S in a generic *macrostate* $\mathscr{X}$ described by few *thermodynamic variables* $\mathscr{X} = (X_0, X_1, \dots, X_r)$.[1] In these notes, I shall often consider $\mathscr{X}$ to be a macrostate *out of equilibrium*: it might be produced by applying some constraints on a thermodynamic system, and then by removing them. The *microstate* $x$ of S is described by a *very large* set of variables $x = (x_1, \dots, x_N)$, $N \sim 10^{23}$. The *"probability"* $W$ of $\mathscr{X}$ is (roughly) the number (or volume) of microscopic states that pertain to $\mathscr{X}$. We can then state the fundamental postulate of statistical mechanics as given by **Boltzmann's relation:**

> The *thermodynamic entropy* $S$ is given by
>
> $$S(\mathscr{X}) = k_\mathrm{B} \log W \qquad \text{(up to sub-extensive corrections)}, \tag{1.1}$$
>
> where $k_\mathrm{B} = 1.38 \cdot 10^{-23}\,\mathrm{J/K}$ is Boltzmann's constant.

Most relations of statistical mechanics follow from this expressions and the properties of $S$ in thermodynamics.

---

[1] I shall often use the parentheses notation for tuples, e.g., $A = (a_i)$ is understood as $A = (a_1, \dots, a_q)$, where $q$ is the length of $A$.

### Shannon and Gibbs entropies

Macroscopic states $\mathscr{X}$ are associated with a probability distribution $P^{\mathscr{X}} = (p_x)$ on the microstates, e.g., the *canonical distribution* $P^{\mathrm{eq}} = (p_x^{\mathrm{eq}})$, where

$$p_x^{\mathrm{eq}} := \exp\left[-(E_x - F)/k_{\mathrm{B}}T\right]. \tag{1.2}$$

One considers averages of observables of the form $\langle A\rangle_P := \sum_x p_x A_x$, where an **observable** $A_x$ is a function of the microstate.

The information content of a probability distribution (Shannon [1948]; Shannon and Weaver [1949]) is measured by the **Shannon entropy**:

$$\mathscr{S}[P] := -\sum_x p_x \log p_x. \tag{1.3}$$

The Shannon entropy is a measure of the uncertainty associated with the given distribution: the larger $\mathscr{S}[P]$, the less informative $P$ is. The link between Shannon's entropy $\mathscr{S}$ and the thermodynamical entropy $S$ is expressed by **Gibbs' relation:** If $P^{\mathrm{eq}}$ is an equilibrium probability distribution, one has $S = k_{\mathrm{B}}\mathscr{S}[P^{\mathrm{eq}}]$.

The probability distribution $P^{\mathrm{eq}}$ of a system in thermal equilibrium satisfies a *variational principle*. For instance, if$\langle E\rangle_P = \sum_x p_x E_x$ is fixed, then $\mathscr{S}[P^{\mathrm{eq}}] \geq \mathscr{S}[P]$, $\forall P$, where $P^{\mathrm{eq}}$ is the canonical distribution. By the same token, the Shannon entropy of a grand canonical distribution is larger than that of any distribution with the same values of the average energy and of the average number of particles.

### Free energy and information

We can actually derive a more general result by considering the equilibrium of a system in contact with a reservoir at temperature $T$. Let us assume that the system is initially in a constrained equilibrium macrostate $\mathscr{X}$ associated with a probability distribution $P$ over the microstates, for which the Shannon entropy is given by $\mathscr{S}[P]$. Define the **Kullback-Leibler divergence** $D_{\mathrm{KL}}[P\|Q]$ between two probability distributions $P = (p_x)$ and $Q = (q_x)$ by (cf. [Cover and Thomas, 1991, p.18–24])

$$D_{\mathrm{KL}}[P\|Q] := \sum_x p_x \log \frac{p_x}{q_x}. \tag{1.4}$$

One can then show that $D_{\mathrm{KL}}[P\|Q]$ has the following properties:

a) $D_{\mathrm{KL}}[P\|Q] \geq 0$;

b) $D_{\mathrm{KL}}[P\|Q] = 0$ only if $p_x = q_x$ “almost everywhere”. For discrete microstate spaces, this means $p_x = q_x$, $\forall x$. For continuous microstate spaces, the densities can differ only in regions whose probability vanishes anyway.

One has indeed, due to the fact that $-\log x$ is a convex function of $x$, and by Jensen's inequality (cf. A.1),

$$D_{\mathrm{KL}}[P\|Q] = -\sum_x p_x \log \frac{q_x}{p_x} \geq -\log \sum_x p_x \frac{q_x}{p_x} = -\log \sum_x q_x = -\log 1 = 0. \tag{1.5}$$

Note that the equality in (1.5) holds only if $q_x = p_x$ everywhere, since $\log x$ is a *strictly* concave function of its argument. Also note that in general $D_{\mathrm{KL}}[P\|Q] \neq D_{\mathrm{KL}}[Q\|P]$. For instance, if $q_x > 0$, $\forall x$, while $p_x = 0$ for some $x$, $D_{\mathrm{KL}}[P\|Q]$ is finite, but $D_{\mathrm{KL}}[Q\|P]$ diverges. We can now define the **availability** $\mathscr{F}[P]$ by

$$\mathscr{F}[P] = \langle E\rangle_P - T\mathscr{S}[P]. \tag{1.6}$$

For the equilibrium distribution at temperature $T$ one has $\mathscr{F}^{\mathrm{eq}} = \mathscr{F}[P^{\mathrm{eq}}] = \langle E\rangle^{\mathrm{eq}} - TS^{\mathrm{eq}} = F$ (Helmholtz free energy). We can then prove that for all probability distributions $P$ one has

$$\mathscr{F}^{\mathrm{eq}} \leq \mathscr{F}[P]. \tag{1.7}$$

One has indeed

$$\begin{aligned}
\Delta\mathscr{F} &:= \mathscr{F}[P] - \mathscr{F}[P^{\mathrm{eq}}] \\
&= \sum_x \left[\left(p_x E_x - p_x^{\mathrm{eq}} E_x\right) + k_{\mathrm{B}} T \left(p_x \log p_x - p_x^{\mathrm{eq}} \log p_x^{\mathrm{eq}}\right)\right] \\
&= \sum_x \left(p_x E_x + k_{\mathrm{B}} T \log p_x - F\right) \\
&= k_{\mathrm{B}} T \underbrace{\sum_x p_x \log \frac{p_x}{p_x^{\mathrm{eq}}}}_{\text{Kullback-Leibler divergence}} = k_{\mathrm{B}} T D_{\mathrm{KL}}[P\|P^{\mathrm{eq}}] \geq 0.
\end{aligned} \tag{1.8}$$

On the other hand, if $\langle E\rangle_P = \langle E\rangle_{P^{\mathrm{eq}}} = E$, we have

$$\Delta\mathscr{F} = -k_{\mathrm{B}} T \left(\mathscr{S}[P] - \mathscr{S}[P^{\mathrm{eq}}]\right) = k_{\mathrm{B}} T D_{\mathrm{KL}}[P\|P^{\mathrm{eq}}] \geq 0, \tag{1.9}$$

and therefore $\mathscr{S}[P] \leq \mathscr{S}[P^{\mathrm{eq}}]$ as announced.

If we now assume that the macrostate described by $P$ is such that one can define its thermodynamic entropy $S = k_{\mathrm{B}}\mathscr{S}[P]$. We can then evaluate the work $-W$ extracted[2] from the system in changing it from non-equilibrium to equilibrium:

$$-W = Q - \Delta E \quad \text{(first principle)} \leq T\Delta S - \Delta E \quad \text{(second principle)} = -\Delta\mathscr{F}. \tag{1.10}$$

Thus if a system in contact with a thermal reservoir is not in an equilibrium state, one can in principle bring it to equilibrium and extract a work $W$ in the process, which is proportional to the Kullback-Leibler divergence between the probability distribution of the initial state and that of the equilibrium state.

## 1.2 Stochastic thermodynamics: Manipulated systems

We now turn to the consideration of the thermodynamics of small systems. Since these systems are far from the thermodynamic limit, some of the properties which hold for ordinary (macroscopic) systems do not apply. In particular

---

[2] We consider as positive a work yielded to the system, and as negative a work extracted from it.

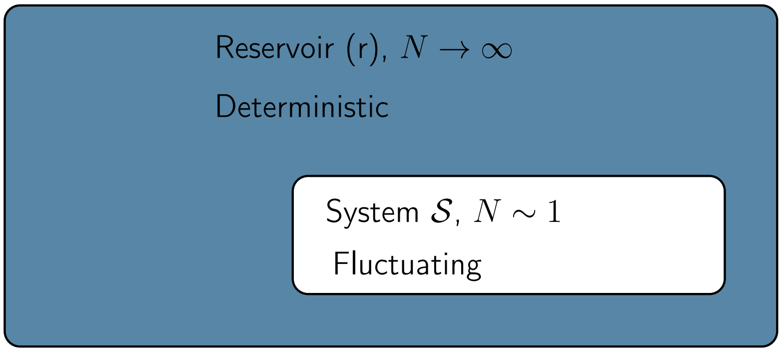


Figure 1.1: The conceptual scheme of small-system thermodynamics. A microscopic system $\mathscr{S}$, subject to fluctuations, is in contact with a large reservoir (r). The balance of thermodynamic quantities is evaluated on the reservoir, while only the physical balance (e.g., of energy) is evaluated at the level of the small system.

- Different ensembles are not equivalent;
- Fluctuations are relevant.

Stochastic thermodynamics is the thermodynamic of small systems out of equilibrium. We shall see that **microscopic reversibility** has strong implications for the behavior of these systems out of equilibrium. We shall also see why these implications, which in principle apply to systems of any size, are indeed relevant for small systems, which are characterized by large fluctuations.

## Small systems

Thermodynamics applies without qualifications to systems in the *thermodynamic limit*, $N \to \infty$, where $N$ is the number of units making the system. In this limit, fluctuations are negligible, and the average values of thermodynamic quantities, such as the work or heat exchanged with the surroundings, are equal to their typical values. When the system size becomes very small, eventually just a few molecules, fluctuations become important. In order to evaluate the implications of thermodynamics for small systems, we shall think that they are in contact with one (or more) reservoirs, which can be treated by usual thermodynamics. In this way, although we shall refrain from associating a non-equilibrium system with a well-defined thermodynamical entropy, we shall evaluate the change in entropy of the associated reservoir for the different realizations of a stochastic process that the (small) system of interest is undergoing. This is reminiscent of the conceptual scheme of quantum theory, in which the predictions of the microscopic thermodynamics (Schrödinger equation) can be interpreted in terms of the interactions of the small system with a macroscopic system, governed by classical mechanics. Ritort [2007] has provided a concise review of the field, and there is a more recent extensive review (Seifert [2012]).

## Markov evolution

We shall discuss a system represented by Markov chain in discrete time. The microstates are denoted by $x \in \{1,2,\dots,q\}$, and the time is described by a variable $t \in \{0,1,2,\dots\}$. The energy of the microstate $x$ is given by $E_x(\lambda)$, where $\lambda$ is an external parameter that can be manipulated. We assume that the system is in contact with a heat reservoir at temperature $T$. In this situation the equilibrium distribution is given by $P^{\text{eq}}(\lambda) = (p_x^{\text{eq}}(\lambda))$, where

$$p_x^{\text{eq}}(\lambda) = \mathrm{e}^{-(E_x(\lambda)-F(\lambda))/k_{\text{B}}T}; \tag{1.11}$$

$$F(\lambda) = -k_{\text{B}}T\log\sum_x \mathrm{e}^{-\beta E_x(\lambda)}. \tag{1.12}$$

The evolution of the microstates is described by a Markov chain, defined by the matrix $\mathsf{W}(\lambda) = (W_{x'x}(\lambda))$ which gives the conditional probability that if the system is in state $x$ at time $t$ it is in state $x'$ at time $t+1$. This probability depends on $\lambda$, but does not otherwise depend explicitly on time. By the law of total probabilities, the evolution of the probability $P(t) = (p_x(t))$ that the system is in microstate $x$ at time $t$ obeys the equation

$$p_x(t+1) = \sum_{x'} W_{xx'}(\lambda)p_x(t). \tag{1.13}$$

The equilibrium distribution $P^{\text{eq}}$ must be invariant under this dynamics. One has therefore the following condition linking $\mathsf{W}$ and $P^{\text{eq}}$:

$$\sum_{x'} W_{xx'}(\lambda)p_{x'}^{\text{eq}}(\lambda) = p_x^{\text{eq}}(\lambda). \tag{1.14}$$

However, for system obeying microscopic reversibility, the dynamics satisfies a stronger condition, namely that for any pair $(x,x')$ of microstates, the probability of a transition $x \longrightarrow x'$ must be equal to the probability of the reverse transition $x' \longrightarrow x$ at equilibrium. This condition is known as the **detailed-balance condition** and reads

$$W_{xx'}(\lambda)p_{x'}^{\text{eq}}(\lambda) = W_{x'x}(\lambda)p_x^{\text{eq}}(\lambda), \qquad \forall x \neq x'. \tag{1.15}$$

### Work and heat at equilibrium

Given a function $A_x$ of the microstate and the probability distribution $P = (p_x)$, the average $\langle A\rangle_P$ is defined by

$$\langle A\rangle_P = \sum_x A_x p_x. \tag{1.16}$$

In particular, if $P$ is the equilibrium distribution $P^{\text{eq}}(\lambda)$, the average will be denoted by $\langle\cdot\rangle_\lambda^{\text{eq}}$. The internal energy at equilibrium is thus given by

$$E(\lambda) = \langle E_x(\lambda)\rangle_\lambda^{\text{eq}} = \sum_x E_x(\lambda)\,\mathrm{e}^{-(E_x(\lambda)-F(\lambda))/k_{\text{B}}T}. \tag{1.17}$$

Let us now evaluate the change in the internal energy at equilibrium as $\lambda$ varies from, say, $\lambda$ to $\lambda+\mathrm{d}\lambda$. One has $E(\lambda+\mathrm{d}\lambda) = E(\lambda)+\mathrm{d}E(\lambda)$, with

$$\mathrm{d}E(\lambda) = \underbrace{\sum_x \mathrm{d}E_x(\lambda)\,p_x^{\text{eq}}(\lambda(t))}_{\mathrm{d}W} + \underbrace{\sum_x E_x(\lambda)\,\mathrm{d}p_x^{\text{eq}}(\lambda)}_{\mathrm{d}Q}. \tag{1.18}$$

Thus $\mathrm{d}E$ is split into two contributions: the first one, which depends on the change of the energies of the microstates, but not on the change of the probability distribution, can be identified with the work $\mathrm{d}W$ performed on the system. The second contribution involves the change in the occupation probability of the microstates. One can surmise that this change is due to the interaction of the system with the reservoir, and we can therefore identify this contribution with the heat $\mathrm{d}Q$ provided by the reservoir to the system. The change in the entropy due to the change in the probability distribution is given by

$$\mathrm{d}S^{\mathrm{eq}} = -k_{\mathrm{B}}\,\mathrm{d}\left(\sum_x p_x^{\mathrm{eq}}(\lambda)\log p_x^{\mathrm{eq}}(\lambda)\right) = -k_{\mathrm{B}}\sum_x \log p_x^{\mathrm{eq}}(\lambda)\,\mathrm{d}p_x^{\mathrm{eq}}(\lambda). \tag{1.19}$$

Using the expression (1.11) of the equilibrium distribution and taking into account that $\sum_x \mathrm{d}p_x^{\mathrm{eq}} = 0$ one sees that

$$\mathrm{d}S^{\mathrm{eq}} = -k_{\mathrm{B}}\sum_x \log p_x^{\mathrm{eq}}(\lambda)\,\mathrm{d}p_x^{\mathrm{eq}}(\lambda) = \frac{1}{T}\sum_x \left(E_x(\lambda) - F(\lambda)\right)\mathrm{d}p_x^{\mathrm{eq}}(\lambda) = \frac{\mathrm{d}Q}{T}. \tag{1.20}$$

## Manipulated system

Let us now assume that the system is manipulated by changing $\lambda$, according to a manipulation protocol given by $\boldsymbol{\lambda} = (\lambda(0), \lambda(1), \dots, \lambda(t_{\mathrm{f}}-1))$ from the initial time $t=0$ to the final time $t = t_{\mathrm{f}}$. During this interval, the system evolves along a path $\boldsymbol{x} = (x(0), x(1), x(2), \dots, x(t_{\mathrm{f}}))$. To be definite, we assume that transitions take place with the "old" probability matrix $\mathsf{W}(\lambda(t))$, and that then the energy of the states change to the new values $E_x(\lambda(t+1))$. The energy of the system changes during the evolution step from $t$ to $t+1$ by one of two possible mechanisms:

1. The system remains in a given state, say $x$, whose energy changes from $E_x(\lambda(t))$ to $E_x(\lambda(t+1))$. In this case the probability of the occupation of the state has not changed. We can thus identify the change in the system's energy as due to work, given by $\mathrm{d}\mathscr{W}_x = E_x(\lambda(t+1)) - E_x(\lambda(t))$.
2. The system jumps from $x(t)$ to a different state $x(t+1) = x'$. This jump is accompanied in general by a change in energy, which has to be provided by an interaction with the reservoir. We can then identify this energy change as due to heat exchange, with the heat given by $\mathrm{d}\mathscr{Q}_{x'x} = E_{x'}(\lambda(t)) - E_x(\lambda(t))$.

We can now evaluate the probability of a path $\boldsymbol{x}$, conditioned on its starting point $x(0)$. We obtain

$$\mathscr{P}[\boldsymbol{x}|x(0)] = W_{x(t_{\mathrm{f}}),x(t_{\mathrm{f}}-1)}(\lambda(t_{\mathrm{f}}-1))\cdots W_{x(2),x(1)}(\lambda(1))\,W_{x(1),x(0)}(\lambda(0)). \tag{1.21}$$

### Work and Heat along a path

We can thus evaluate separately the total work and the total heat along a path $\boldsymbol{x}$, given the manipulation protocol $\boldsymbol{\lambda}$. For the work $\mathscr{W}[\boldsymbol{x}]$ we obtain

$$\mathscr{W}[\boldsymbol{x}] = \sum_{t=0}^{t_{\mathrm{f}}-2} \mathscr{W}_{x(t+1)} = \sum_{t=0}^{t_{\mathrm{f}}-1}\left[E_{x(t+1)}(\lambda(t+1)) - E_{x(t)}(\lambda(t))\right]. \tag{1.22}$$

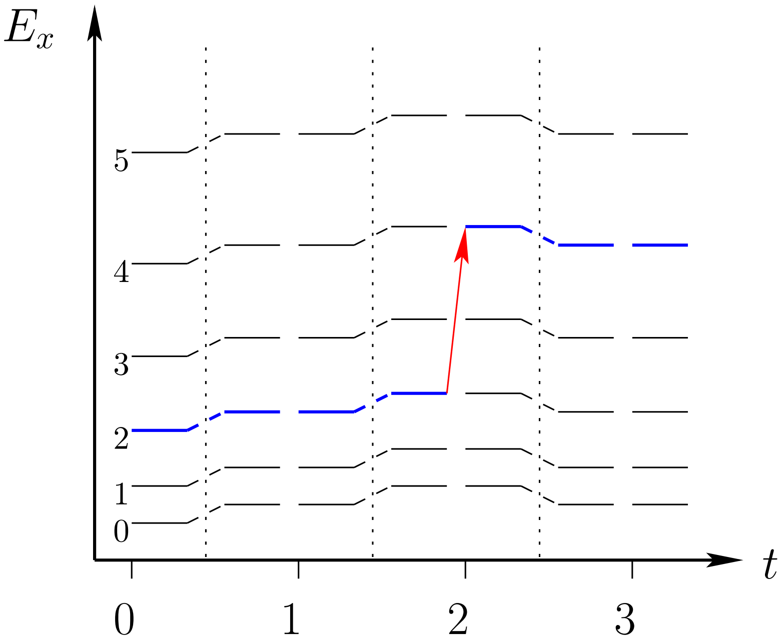


Figure 1.2: Schematic representation of a path followed by a manipulated system. The lines represent the energy levels of the different microstates of the system. Their height varies due to the manipulation of the parameter $\lambda$, which takes place at the times marked by a vertical dotted line. The blue lines follow the states actually occupied by the system, and the red arrow denotes a transition. Thus the system starts, in this case, in state 2 at time 0. The energy changes in going from $t = 0$ to $t = 1$, and then again in going from $t = 1$ to $t = 2$. At $t = 2$ the system undergoes a transition to state 4, which involves acquiring a quantity of heat $\mathcal{Q} = E_4(\lambda(2)) - E_2(\lambda(2))$ from the reservoir. The energy of this state changes (due to a change in the parameter $\lambda$) between $t = 2$ and $t = 3$. Thus the system's energy can change either because the energy of the state occupied by the system changes (work) or because the system jumps to a state with a different energy (heat).

Note that we assume that the system is allowed to make a transition at $t_\mathrm{f}$, but the manipulation stops at $t_\mathrm{f} - 1$. The heat $\mathcal{Q}$ is evaluated by summning the changes of energy due to changes in the microstate. One obtains

$$\mathcal{Q}[\boldsymbol{x}] = \sum_{t=0}^{t_\mathrm{f}-1} \left[E_{x(t+1)}(\lambda(t)) - E_{x(t)}(\lambda(t))\right]. \tag{1.23}$$

The total change in energy is given by

$$\Delta E = E_{x(t_\mathrm{f})}(\lambda(t_\mathrm{f} - 1)) - E_{x(0)}(\lambda(0)) = \mathcal{W}[\boldsymbol{x}] + \mathcal{Q}[\boldsymbol{x}]. \tag{1.24}$$

The work $\mathcal{W}$ and the heat $\mathcal{Q}$ both depend on the *whole* path, while $\Delta E$, since $E$ is a state function, depends only on initial and final states.

## Fluctuation relations: Crooks and Seifert

We assume that the states of the system are invariant under time reversal. Then to each path $\boldsymbol{x} = (x(0), x(1), \dots, x(t_\mathrm{f}-1), x(t_\mathrm{f}))$ we can associate the **time-reversed** path $\widehat{\boldsymbol{x}}$, defined by

$$\widehat{\boldsymbol{x}} := (x(t_\mathrm{f}), x(t_\mathrm{f}-1), \dots, x(1), x(0)); \tag{1.25}$$

$$\widehat{x}(t) := x(t_\mathrm{f}-t) = x(\widehat{t}). \tag{1.26}$$

We can likewise define the time-reversed protocol $\widehat{\boldsymbol{\lambda}}$, defined by

$$\widehat{\boldsymbol{\lambda}}: \qquad \widehat{\lambda}(t) := \lambda(t_\mathrm{f}-t) = \lambda(\widehat{t}); \qquad t = 1, \dots, t_\mathrm{f}. \tag{1.27}$$

Note the shift in the argument: $\boldsymbol{\lambda}$ is defined for $0 \le t \le t_\mathrm{f}-1$, while $\widehat{\lambda}$ is defined for $1 \le t \le t_\mathrm{f}$. We can then evaluate the conditional probability of the reversed path with the reversed protocol. Given the direct transition $x \longrightarrow x'$ for a given value of $\lambda$, the probability of the reverse transition $x' \longrightarrow x$ is simply given by $\widehat{W}_{xx'}(\lambda) = W_{xx'}(\lambda)$. Now, if *detailed balance* (1.15) holds, we have an explicit relation between these probabilities:

$$W_{xx'}(\lambda)\,\mathrm{e}^{-E_{x'}(\lambda)/k_\mathrm{B}T} = W_{x'x}(\lambda)\,\mathrm{e}^{-E_x(\lambda)/k_\mathrm{B}T}. \tag{1.28}$$

One has therefore

$$\widehat{W}_{xx'}(\lambda) = W_{x'x}(\lambda)\,\mathrm{e}^{-(E_x(\lambda)-E_{x'}(\lambda))/k_\mathrm{B}T}. \tag{1.29}$$

### Crooks' relation

A basic insight is reached by comparing the probability of a "forward" path $\boldsymbol{x}$ under the "forward" protocol $\boldsymbol{\lambda}$ with the probability of the reversed path $\widehat{\boldsymbol{x}}$ under the reversed protocol $\widehat{\boldsymbol{\lambda}}$. To ease the notation, we shall write the transition probabilities as functions of time, rather than via their dependence on $\lambda$:

$$W_{xx'}(\lambda(t)) = W_{xx'}(t). \tag{1.30}$$

We then have

$$\widehat{W}_{xx'}(t) = W_{x'x}(t_\mathrm{f}-t), \qquad t = 1, \dots t_\mathrm{f}. \tag{1.31}$$

Let us evaluate the ratio of the conditional probability of the "forward" path $\boldsymbol{x}$, given $x(0)$, under the "forward" protocol $\boldsymbol{\lambda}$, to that of the reverse path $\widehat{\boldsymbol{x}}$ given $\boldsymbol{x}(0)$, i.e., $x(t_\mathrm{f})$, under the reverse protocol $\widehat{\boldsymbol{\lambda}}$, which we denote by $\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}|\widehat{x}(0)]$. We obtain

$$\begin{aligned}
\frac{\mathscr{P}[\boldsymbol{x}|x(0)]}{\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}|\widehat{x}(0)]} &= \frac{W_{x(t_\mathrm{f})x(t_\mathrm{f}-1)}(t_\mathrm{f}-1)\cdots W_{x(2)x(1)}(1)\,W_{x(1)x(0)}(0)}{\widehat{W}_{\widehat{x}(t_\mathrm{f})\widehat{x}(t_\mathrm{f}-1)}(t_\mathrm{f})\cdots \widehat{W}_{\widehat{x}(2)\widehat{x}(1)}(2)\,\widehat{W}_{\widehat{x}(1)\widehat{x}(0)}(1)} \\
&= \frac{W_{x(t_\mathrm{f})x(t_\mathrm{f}-1)}(t_\mathrm{f}-1)}{W_{x(t_\mathrm{f}-1)x(t_\mathrm{f})}(t_\mathrm{f}-1)}\cdots\frac{W_{x(1)x(0)}(0)}{W_{x(0)x(1)}(0)} && \text{(by permutation)} \\
&= \mathrm{e}^{-\left(E_{x(t_\mathrm{f})}(t_\mathrm{f}-1)-E_{x(t_\mathrm{f}-1)}(t_\mathrm{f}-1)\right)/k_\mathrm{B}T}\cdots\mathrm{e}^{-(E_{x(1)}(0)-E_{x(0)}(0))/k_\mathrm{B}T} && \text{(detailed balance)} \\
&= \mathrm{e}^{-\mathscr{Q}[\boldsymbol{x}]/k_\mathrm{B}T}.
\end{aligned} \tag{1.32}$$

Thus this ratio of probabilities is related to the change in entropy of the reservoir. Since $\mathscr{Q} = -T\Delta S^{(\mathrm{r})}$ (assuming that the reservoir is always at equilibrium), we have

$$\frac{\mathscr{P}[\boldsymbol{x}|x(0)]}{\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}|\widehat{x}(0)]} = \mathrm{e}^{\Delta S^{(\mathrm{r})}[\boldsymbol{x}]/k_\mathrm{B}}. \tag{1.33}$$

This relation is fundamental in the theory of stochastic thermodynamics, and goes under the name of **Crook's relation**, since it was derived in Crooks [1999].

### Unconditional probabilities: Seifert's relation

To go from the conditional probabilities to the actual path probabilities it is sufficient to multiply by the occupation probabilities $p_{x_0}(0)$, $p_{x_\mathrm{f}}(t_\mathrm{f})$ at either end of the path. In this way we obtain

$$\mathscr{P}[\boldsymbol{x}] = \mathscr{P}[\boldsymbol{x}|x(0)]\,p^{(0)}_{x(0)}; \tag{1.34}$$

$$\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}] = \mathscr{P}[\widehat{\boldsymbol{x}}|\widehat{x(0)}]\,\widehat{p^{(0)}}_{\widehat{x(0)}} = \mathscr{P}[\widehat{\boldsymbol{x}}|\widehat{x(0)}]\,p^{(\mathrm{f})}_{x_\mathrm{f}}. \tag{1.35}$$

Thus we obtain

$$\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}] = \widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}|\widehat{x}(0)]\,p^{(\mathrm{f})}_{x_\mathrm{f}} = \mathscr{P}[\boldsymbol{x}]\mathrm{e}^{\mathscr{Q}[\boldsymbol{x}]/k_\mathrm{B}T-\left(\log p^{(0)}_{x(0)}-\log p^{(\mathrm{f})}_{x_\mathrm{f}}\right)}. \tag{1.36}$$

The second term in the exponential can be interpreted by introducing the notion of **stochastic entropy** (Crooks [1999]; Qian [2002]) via the relation

$$s_x(t) := -k_\mathrm{B}\log p_x(t). \tag{1.37}$$

We then have

$$\left(\log p^{(0)}_{x(0)} - \log p^{(\mathrm{f})}_{x_\mathrm{f}}\right) = \left(s_{x_\mathrm{f}}(t_\mathrm{f}) - s_{x(0)}(0)\right)/k_\mathrm{B} = \Delta s/k_\mathrm{B}. \tag{1.38}$$

This relation implies

$$\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}] = \mathscr{P}[\boldsymbol{x}]\,\mathrm{e}^{\mathscr{Q}[\boldsymbol{x}]/k_\mathrm{B}T-\Delta s/k_\mathrm{B}} = \mathscr{P}[\boldsymbol{x}]\,\mathrm{e}^{-\left(\Delta S^{(\mathrm{r})}+\Delta s\right)/k_\mathrm{B}}. \tag{1.39}$$

The exponent can be interpreted as the total **entropy production** associated with the path $\boldsymbol{x}$:

$$\Delta_\mathrm{i}S[\boldsymbol{x}] = \Delta S^{(\mathrm{r})}[\boldsymbol{x}] + \Delta s[\boldsymbol{x}]. \tag{1.40}$$

This appears as the increase of the entropy of the complex formed by the system under study and the reservoir with which it is in contact. Summing over $\boldsymbol{x}$, and taking into account that the probability $\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}]$ is normalized, and that to each path $\boldsymbol{x}$ corresponds one and only one reversed path $\widehat{\boldsymbol{x}}$, we obtain the following **integral fluctuation relation**:

$$\langle \mathrm{e}^{-\Delta_\mathrm{i}S[\boldsymbol{x}]/k_\mathrm{B}}\rangle = \sum_{\boldsymbol{x}}\mathscr{P}[\boldsymbol{x}]\,\mathrm{e}^{-\Delta_\mathrm{i}S[\boldsymbol{x}]/k_\mathrm{B}} = \sum_{\boldsymbol{x}}\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}] = 1. \tag{1.41}$$

By Jensen's inequality, we recover Clausius' expression of the second principle in the form

$$\langle \Delta_\mathrm{i}S[\boldsymbol{x}]\rangle \geq 0. \tag{1.42}$$

Relations (1.39,1.41) were derived by Seifert [2005]. In this way the total entropy production appears to be intimately related to the probability of distinguishing one path from its reverse.

## Jarzynski's relation

In the special case in which the system is driven out of an initial equilibrium state, i.e., $P^{(0)} = P^{\mathrm{eq}}$, we obtain

$$-\log p_{x(0)}(0) = \big(E_{x(0)}(\lambda_0) - F(\lambda_0)\big)/k_{\mathrm{B}}T; \tag{1.43}$$

$$-\log \widehat{p}_{x(0)}(0) = \big(E_{x_{\mathrm{f}}}(\lambda_{\mathrm{f}}) - F(\lambda_{\mathrm{f}})\big)/k_{\mathrm{B}}T; \tag{1.44}$$

$$\Delta s = [\Delta E - (F(\lambda_{\mathrm{f}}) - F(\lambda_0))]/k_{\mathrm{B}}T. \tag{1.45}$$

Thus the relation (1.39) takes on the form

$$\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}] = \mathscr{P}[\boldsymbol{x}]\,\mathrm{e}^{\mathscr{Q}[\boldsymbol{x}]/k_{\mathrm{B}}T - \Delta s/k_{\mathrm{B}}} = \mathscr{P}[\boldsymbol{x}]\,\mathrm{e}^{-[\mathscr{W}[\boldsymbol{x}] - (F(\lambda_{\mathrm{f}}) - F(\lambda_0))]/k_{\mathrm{B}}T}. \tag{1.46}$$

Therefore one has the integral relation

$$\langle \mathrm{e}^{-\mathscr{W}[\boldsymbol{x}]/k_{\mathrm{B}}T}\rangle = \mathrm{e}^{-(F(\lambda_{\mathrm{f}}) - F(\lambda_0))/k_{\mathrm{B}}T}. \tag{1.47}$$

This relation, derived by Jarzynski [1997], is at the origin of much research. It pointed out in particular on the possibility of extracting information about the free-energy differences of macrostates by performing non-equilibrium experiments.

### Elementary examples

In the present paragraph we see how Jarzynski's relation works out in the case of simple manipulating protocols.

- Let us consider a sudden change of the parameter $\lambda$:

  $$\lambda_0 \longrightarrow \lambda_{\mathrm{f}}. \tag{1.48}$$

  In a sudden change, the microstate of the system does not have the time to relax. Thus the change in energy is only due to the work, which is itself simply given by

  $$\mathscr{W}[\boldsymbol{x}] = E_{x(0)}(\lambda_{\mathrm{f}}) - E_{x(0)}(\lambda_0). \tag{1.49}$$

  Evaluating the average of $\mathrm{e}^{-\mathscr{W}/k_{\mathrm{B}}T}$ over the initial distribution $P^{\mathrm{eq}}(\lambda_0)$ we obtain

  $$\begin{aligned}\langle \mathrm{e}^{-\mathscr{W}/k_{\mathrm{B}}T}\rangle &= \sum_x \mathrm{e}^{-(E_x(\lambda_{\mathrm{f}}) - E_x(\lambda_0))/k_{\mathrm{B}}T}\,\mathrm{e}^{-(E_x(\lambda_0) - F(\lambda_0))/k_{\mathrm{B}}T}\\ &= \sum_x \mathrm{e}^{-(E_x(\lambda_{\mathrm{f}}) - F(\lambda_0))/k_{\mathrm{B}}T} = \mathrm{e}^{-(F(\lambda_{\mathrm{f}}) - F(\lambda_0))/k_{\mathrm{B}}T}.\end{aligned} \tag{1.50}$$

- On the opposite limit, let us assume that the change is so slow that the system has the time to relax to equilibrium for each value of the parameter $\lambda$. We can formalize this hypothesis by setting

  $$\lambda(t) = \ell(t/t_{\mathrm{f}}), \qquad \text{with } \ell(0) = \lambda_0 \quad \text{and } \ell(1) = \lambda_{\mathrm{f}}, \tag{1.51}$$

and letting $t_\mathrm{f} \to \infty$. We then have

$$\begin{aligned}
\mathcal{W}[\boldsymbol{x}] &= \sum_{t=0}^{t_\mathrm{f}-1} \left(E_{x(t+1)}(\lambda(t+1)) - E_{x(t+1)}(\lambda(t))\right) \\
&\simeq \sum_{t=0}^{t_\mathrm{f}-1} \left.\frac{\partial E_{x(t+1)}(\lambda)}{\partial \lambda}\right|_{\lambda(t)} \frac{\mathrm{d}\lambda}{\mathrm{d}t} \Delta t \\
&\simeq \sum_{t=0}^{t_\mathrm{f}-1} \left.\frac{\partial E_{x(t+1)}(\lambda)}{\partial \lambda}\right|_{\lambda=\ell(t/t_\mathrm{f})} \frac{1}{t_\mathrm{f}} \left.\frac{\mathrm{d}\ell}{\mathrm{d}x}\right|_{x=t/t_\mathrm{f}} \\
&= \int_0^{t_\mathrm{f}} \mathrm{d}t' \, \left\langle \frac{\partial E_{x(t')}(\lambda)}{\partial \lambda} \right\rangle_{\lambda=\lambda(t')} \dot{\lambda}(t') \\
&= \int_{\lambda_0}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \, \left\langle \frac{\partial E_x(\lambda)}{\partial \lambda} \right\rangle_\lambda = W^\mathrm{rev} = F(\lambda_\mathrm{f}) - F(\lambda_0). \qquad (1.52)
\end{aligned}$$

In going from the third to the fourth line we have taken advantage of the fact that the system is able to sample the equilibrium distribution for each value of $\lambda$. As a consequence, the work $\mathcal{W}$ is no more a fluctuating variable but is equal (with probability one) to the reversible work $W^\mathrm{rev}$, that, as we know, is equal to the free-energy difference.

### Dissipated work close to equilibrium

We can obtain an interesting relation for the dissipated work for the case in which the system is manipulated slowly. From Jensen's inequality we obtain

$$\langle \mathcal{W} \rangle \geq -k_\mathrm{B} T \log \langle \mathrm{e}^{-\mathcal{W}/k_\mathrm{B}T} \rangle. \qquad (1.53)$$

Therefore the dissipated work is non-negative on average:

$$W^\mathrm{diss} := \langle \mathcal{W} \rangle - \Delta F \geq 0. \qquad (1.54)$$

Assume $\mathcal{W} = \langle \mathcal{W} \rangle + \delta\mathcal{W}$, where $\delta\mathcal{W}$ is small and vanishes on average. This holds if $\lambda$ is manipulated slowly. We then obtain

$$\begin{aligned}
\mathrm{e}^{-\Delta F/k_\mathrm{B}T} &= \langle \mathrm{e}^{-\mathcal{W}/k_\mathrm{B}T} \rangle = \langle \mathrm{e}^{-(\langle \mathcal{W} \rangle + \delta\mathcal{W})/k_\mathrm{B}T} \rangle \simeq \mathrm{e}^{-\langle \mathcal{W} \rangle/k_\mathrm{B}T} \langle \mathrm{e}^{-\delta\mathcal{W}/k_\mathrm{B}T} \rangle \\
&= \mathrm{e}^{-\langle \mathcal{W} \rangle/k_\mathrm{B}T} \langle 1 - \delta\mathcal{W}/k_\mathrm{B}T + \frac{1}{2(k_\mathrm{B}T)^2} \delta\mathcal{W}^2 + \cdots \rangle \\
&\simeq \mathrm{e}^{-\langle \mathcal{W} \rangle/k_\mathrm{B}T} \left(1 + \frac{1}{2(k_\mathrm{B}T)^2} \langle \delta\mathcal{W}^2 \rangle + \cdots \right). \qquad (1.55)
\end{aligned}$$

Taking the log, if higher orders in $\delta\mathcal{W}$ can be neglected, we obtain

$$W^\mathrm{diss} \simeq \frac{1}{2k_\mathrm{B}T} \langle \delta\mathcal{W}^2 \rangle. \qquad (1.56)$$

This relation connects fluctuations of the work to the dissipated work close to equilibrium, and is therefore an instance of the Fluctuation-Dissipation Theorem, that we discuss in the next paragraph.

### Generalized Jarzynski relation and the Fluctuation-Dissipation Theorem

It is possible to generalize Jarzynski's relation by considering the average of a general function $A_x$ of the microstate $x$. Multiply both sides of (1.46) by $A_{x_\mathrm{f}}$, where $x_\mathrm{f}$ is the last state of the path $\boldsymbol{x}$. The argument of the exponential is nothing else as the dissipated fluctuating work

$$\mathscr{W}^{\mathrm{diss}}[\boldsymbol{x}] = \mathscr{W}[\boldsymbol{x}] - \Delta F. \tag{1.57}$$

Let us now take the sum over all paths $\boldsymbol{x}$. The left-hand side simply becomes the average of $A_x$ taken with the equilibrium distribution with the final value $\lambda_\mathrm{f}$ of $\lambda(t)$ in the protocol. On the right-hand side we obtain the average of $A_{x(t_\mathrm{f})}$ taken over all paths, weighted by the factor $\mathrm{e}^{-\mathscr{W}^{\mathrm{diss}}[\boldsymbol{x}]/k_\mathrm{B}T}$. We thus obtain the **generalized Jarzynski relation**

$$\langle A_{x(t_\mathrm{f})}\rangle_{\lambda_\mathrm{f}} = \langle A_{x(t_\mathrm{f})}\mathrm{e}^{-\mathscr{W}^{\mathrm{diss}}[\boldsymbol{x}]/k_\mathrm{B}T}\rangle_{\boldsymbol{\lambda}}. \tag{1.58}$$

Let us assume that the energy $E_x(\lambda)$ depends on $\lambda$ via the expression

$$E_x(\lambda) = E_x^{(0)} - \lambda B_x, \tag{1.59}$$

where $B_x$ is some function of the microstate. Then the work $\mathscr{W}$ has the expression (I use a continous-time formalism in this section)

$$\mathscr{W}[\boldsymbol{x}] = -\int_0^{t_\mathrm{f}} \mathrm{d}t\, \dot{\lambda}(t) B_{x(t)}. \tag{1.60}$$

Let the perturbation protocol $\boldsymbol{\lambda}$ satisfy $\lambda_0 = \lambda(t_\mathrm{f}) = 0$, with $\lambda(t)$ and its derivative small. We can then take the functional derivative of both sides of equation (1.58) with respect to $\lambda(t)$, and set $\boldsymbol{\lambda} \to 0$. The left-hand side vanishes. We thus obtain

$$0 = \left.\frac{\delta\langle A(t_\mathrm{f})\rangle}{\delta\lambda(t)}\right|_{\boldsymbol{\lambda}=\boldsymbol{0}} - \left.\langle A_{x(t_\mathrm{f})}\frac{\delta}{\delta\lambda(t)}\mathrm{e}^{-\mathscr{W}^{\mathrm{diss}}[\boldsymbol{x}]/k_\mathrm{B}T}\rangle_{\boldsymbol{\lambda}}\right|_{\boldsymbol{\lambda}=0}. \tag{1.61}$$

The first term on the rhs is equal to the **response function** $R(t_\mathrm{f}, t)$, defined by

$$R(t_\mathrm{f}, t) := \left.\frac{\delta\langle A(t_\mathrm{f})\rangle}{\delta\lambda(t)}\right|_{\boldsymbol{\lambda}=\boldsymbol{0}}. \tag{1.62}$$

The second term yields

$$\frac{1}{k_\mathrm{B}T}\frac{\mathrm{d}}{\mathrm{d}t}\langle A(t_\mathrm{f})B(t)\rangle. \tag{1.63}$$

We thus obtain the **Fluctuation-Dissipation theorem** (for a recent review, cf. Marini Bettolo Marconi et al. [2008]):

$$R(t_\mathrm{f}, t) = -\frac{1}{k_\mathrm{B}T}\frac{\mathrm{d}}{\mathrm{d}t_\mathrm{f}}\langle A(t_\mathrm{f})B(t)\rangle. \tag{1.64}$$

Thus Jarzynski's generalized equality (1.58) can be interpreted as an extension of the fluctuation-dissipation theorem far from equilibrium.

## Applications: Exploring the free-energy landscape

We can take advantage of Jarzynski's relation to obtain information on the dependence of the free energy of a system on an observable $M$ which acts as a collective coordinate.

- Let $M_x$ be a collective coordinate and $M^{\mathrm{eq}} := \langle M_x \rangle^{\mathrm{eq}}$ its equilibrium average.
- Assume that we can manipulate the system via an energy function which depends on the instantaneous value of $M$ and on a parameter $\lambda$:
$$U(M_x,\lambda) \longrightarrow E_x(\lambda) = E_x^{(0)} - U(M_x,\lambda). \tag{1.65}$$
- We wish to evaluate the free energy of the system when $M$ is constrained on some value:
$$F_0(M) := -k_{\mathrm{B}}T \log \sum_x \delta(M - M_x)\, \mathrm{e}^{-E_x^{(0)}/k_{\mathrm{B}}T}. \tag{1.66}$$
- Then the probability distribution of $M$ in the unperturbed system is given by
$$P^{\mathrm{eq}}(M) = \mathrm{e}^{-(F_0(M)-F_0)/k_{\mathrm{B}}T}. \tag{1.67}$$

**Work probability distribution**

To solve this problem, let us switch to a description of the system's evolution in continuous time. Define the evolution operator $\mathscr{L}_\lambda$ which governs the evolution of the instantaneous probability distribution $P(t) = (p_x(t))$:

$$\frac{\mathrm{d}p_x(t)}{\mathrm{d}t} = \left(\mathscr{L}_{\lambda(t)}\, P(t)\right)_x. \tag{1.68}$$

Here we have introduced the following notation: when a matrix $\mathscr{A} = (A_{xx'})$ $(x, x' = 1,\dots,q)$ is multiplied by a vector $Q = (q_x)$ $(x = 1,\dots,q)$, yielding a vector $V = (v_x)$, we set

$$v_x = \sum_{x'} A_{xx'} q_{x'} = (\mathscr{A}\, Q)_x. \tag{1.69}$$

The evolution operator can be expressed in terms of the transition rates $k_{x'x}$ that determine the condition probability per unit time that the system undergoes the transition $x \longrightarrow x'$ $(x \neq x')$. We have indeed the **master equation**

$$\frac{\mathrm{d}p_x(t)}{\mathrm{d}t} = \sum_{x'\,(\neq x)} \left[k_{xx'}(\lambda)\, p_{x'}(t) - k_{x'x}(\lambda)\, p_x(t)\right] = \left(\mathscr{L}_{\lambda(t)}\, P(t)\right)_x. \tag{1.70}$$

The evolution operator $\mathscr{L}_\lambda$ annihilates the equilibrium distribution $P^{\mathrm{eq}}(\lambda)$, so that $P^{\mathrm{eq}}(\lambda)$ remains stationary. One has indeed

$$k_{x'x}(\lambda)\, \mathrm{e}^{-E_x(\lambda)/k_{\mathrm{B}}T} = k_{xx'}(\lambda)\, \mathrm{e}^{-E_{x'}(\lambda)/k_{\mathrm{B}}T} \tag{1.71}$$

by the detailed-balance condition, which implies

$$\left(\mathscr{L}_\lambda\, P^{\mathrm{eq}}(\lambda)\right)_x = 0, \qquad \forall \lambda, x. \tag{1.72}$$

In this setting, a path $\boldsymbol{x}$ is defined by a sequence of states $x_k$ and of jump times $t_k$:

$$\boldsymbol{x} = ((x_0, t_0), (x_1, t_1), (x_2, t_2), \dots, (x_\mathrm{f}, t_{k_\mathrm{f}}), t_\mathrm{f}), \tag{1.73}$$

such that $x = x_k$ for $t_k \le t < t_{k+1}$, and we set $t_\mathrm{f} = t_{k_\mathrm{f}+1}$ as the last time (in which no jump occurs). The fluctuating work is then given by

$$\mathscr{W} = -\int_{t_0}^{t_\mathrm{f}} \mathrm{d}t'\, \dot{\lambda}(t')\, \frac{\partial U\big(M_{x(t')}, \lambda(t')\big)}{\partial \lambda}. \tag{1.74}$$

Let us evaluate the joint probability $\Phi(W,t) = (\phi_x(W,t))$ that at time $t$ the system is in state $x$ and the total work $\mathscr{W}(t)$ accumulated until that time is equal to $W$. This quantity obeys the following evolution equation:

$$\frac{\partial \phi_x(W,t)}{\partial t} = \big(\mathscr{L}_{\lambda(t)}\Phi\big)_x + \dot{\lambda}(t) \frac{\partial U(M_x, \lambda(t))}{\partial \lambda} \frac{\partial \phi_x}{\partial W}. \tag{1.75}$$

The first term on the right-hand side describes the evolution of the occupation probability of the states, and the second term describes the evolution of the accumulated work $\mathscr{W}$. This equation is difficult to handle, in particular because it becomes a partial differential equation for the $W$-dependence. But it can be simplified if one introduces the moment-generating function of the $W$ distribution. This function is defined by $\Psi(\mu, t) = (\psi_x(\mu, t))$, where

$$\psi_x(\mu, t) := \int \mathrm{d}W\, \mathrm{e}^{\mu W}\, \phi_x(W,t), \tag{1.76}$$

where $\mu$ is a parameter. We have indeed

$$\sum_x \left.\frac{\partial^k \psi_x(\mu,t)}{\partial \mu^k}\right|_{\mu=0} = \langle (\mathscr{W}(t))^k \rangle, \tag{1.77}$$

where the average is taken with respect to the instantaneous distribution $\Phi(W,t)$. It is a simple matter to obtain the evolution equation satisfied by $\Psi(\mu, t)$:

$$\frac{\partial \psi_x}{\partial t} = (\mathscr{L}_{\lambda(t)}\Psi)_x + \mu\, \dot{\lambda}(t) \frac{\partial U(M_x, \lambda(t))}{\partial \lambda} \psi_x(\mu, t). \tag{1.78}$$

In this equation $\mu$ appears as a parameter. It can be then separately solved for each value of $\mu$. In particular, for $\mu = -1/k_\mathrm{B}T$ one obtains

$$\psi_x(-1/k_\mathrm{B}T, t) = \mathrm{e}^{-(E_x(\lambda(t)) - F(\lambda(0)))/k_\mathrm{B}T}, \tag{1.79}$$

and therefore

$$\sum_x \psi_x(-1/k_\mathrm{B}T, t) = \langle \mathrm{e}^{-\mathscr{W}/k_\mathrm{B}T} \rangle = \mathrm{e}^{-(F(\lambda(t)) - F(\lambda(0)))/kt}. \tag{1.80}$$

In order to prove (1.79) let us define $\Gamma(t) = (\gamma_x(t))$, where

$$\gamma_x(t) = \mathrm{e}^{-(E_x(\lambda(t)) - F(\lambda(0)))/k_\mathrm{B}T}. \tag{1.81}$$

Then $\gamma_x(t)$ satisfies

$$\gamma_x(0) = p_x^{\mathrm{eq}}(\lambda(0)) = \Psi(x, -1/k_{\mathrm{B}}T, 0); \tag{1.82}$$

$$\frac{\partial \gamma_x}{\partial t} = -\frac{\dot{\lambda}}{k_{\mathrm{B}}T}\frac{\partial E_x(\lambda(t))}{\partial \lambda}\gamma_x(t) = \left(\mathscr{L}_{\lambda(t)}\,\Gamma(\lambda(t))\right)_x - \frac{\dot{\lambda}}{k_{\mathrm{B}}T}\frac{\partial H(x,\lambda(t))}{\partial \lambda}\gamma_x(t), \tag{1.83}$$

since $\Gamma(t) \propto P^{\mathrm{eq}}(\lambda(t))$ and $\mathscr{L}_\lambda P^{\mathrm{eq}}(\lambda) = 0$. Therefore

$$\Gamma(t) = \Psi(-1/k_{\mathrm{B}}T, t), \tag{1.84}$$

which we wished to prove. We can now multiply both sides of the expression (1.79) by $\delta(M - M_x)$ and sum over $x$. We obtain

$$\begin{aligned} \langle \delta(M - M_x)\mathrm{e}^{-\mathscr{W}/k_{\mathrm{B}}T}\rangle &= \sum_x \delta(M - M_x)\,\mathrm{e}^{-(E_x(\lambda(t)) - F_0)/k_{\mathrm{B}}T} \\ &= \exp\{-[F_0(M) - U(M,\lambda(t)) - F_0]/k_{\mathrm{B}}T\}. \end{aligned} \tag{1.85}$$

Multiplying both sides by $\mathrm{e}^{U(M,\lambda(t))/k_{\mathrm{B}}T}$ we finally obtain

$$\mathrm{e}^{U(M,\lambda(t))/k_{\mathrm{B}}T}\,\langle \delta(M - M_x)\mathrm{e}^{-\mathscr{W}/k_{\mathrm{B}}T}\rangle = \mathrm{e}^{-[F_0(M) - F_0]/k_{\mathrm{B}}T}. \tag{1.86}$$

This relation has been derived by Crooks [1998] and by Hummer and Szabo [2001]. It has been exploited to evaluate the free energy differences of some microscopic system manipulated via optical tweezers and other methods, as we shall see later.

## Dissipation and irreversibility

Let us rewrite equation (1.46) in the form

$$\mathrm{e}^{(\mathscr{W}[\boldsymbol{x}] - \Delta F)/k_{\mathrm{B}}T} = \frac{\mathscr{P}[\boldsymbol{x}]}{\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}]}. \tag{1.87}$$

Taking the log and averaging, we obtain

$$W^{\mathrm{diss}} = \langle \mathscr{W}\rangle - \Delta F = k_{\mathrm{B}}T \underbrace{\sum_{\boldsymbol{x}} \mathscr{P}[\boldsymbol{x}] \log \frac{\mathscr{P}[\boldsymbol{x}]}{\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}]}}_{D_{\mathrm{KL}}[\mathscr{P}\|\widehat{\mathscr{P}}]}. \tag{1.88}$$

Therefore $W^{\mathrm{diss}}$ measures the Kullback-Leibler divergence of $\mathscr{P}[\boldsymbol{x}]$ from $\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}]$. This result was derived by Kawai et al. [2007]. Let use now define the work distribution $P(W)$ for the forward process, and $\widehat{P}(W)$ for the reversed process by

$$P(W) = \sum_{\boldsymbol{x}} \delta(W - \mathscr{W}[\boldsymbol{x}])\,\mathscr{P}[\boldsymbol{x}]; \tag{1.89}$$

$$\widehat{P}(W) = \sum_{\boldsymbol{x}} \delta(W - \mathscr{W}[\boldsymbol{x}])\,\widehat{\mathscr{P}}[\boldsymbol{x}]. \tag{1.90}$$

We then obtain from (1.87)

$$W^{\mathrm{diss}} = k_{\mathrm{B}}T \int \mathrm{d}W\, P(W) \log \frac{P(W)}{\widehat{P}(-W)} = k_{\mathrm{B}}T D_{\mathrm{KL}}[P(W)\|\widehat{P}(-W)]. \tag{1.91}$$

Therefore the whole divergence of $\mathscr{P}[\boldsymbol{x}]$ from $\widehat{\mathscr{P}}[\widehat{\boldsymbol{x}}]$ is captured by the work distribution.

## Generalized Clausius Inequality

Let us recall the definition of availability for a system in a probability distribution $P$ in contact with a heat reservoir at temperature $T$, We have

$$\mathscr{F}[P] := \langle E\rangle_P - T\mathscr{S}[P] = F + k_\mathrm{B}T D_\mathrm{KL}[P\|P^\mathrm{eq}]. \tag{1.92}$$

On the other hand, we have the following expressions for the average energy and the average (fluctuating) entropy:

$$\langle E\rangle_P = \sum_x E_x p_x; \tag{1.93}$$

$$\mathscr{S}[P] = -k_\mathrm{B}\sum_x p_x \log p_x. \tag{1.94}$$

If the system undergoes a manipulation, and it follows a path $\boldsymbol{x}$, the fluctuating heat $\mathscr{Q}$ and the fluctuating work $\mathscr{W}$ are related to the change in energy $\Delta E$ by the relation (1.24), that we can rewrite as follows:

$$\mathscr{Q}[\boldsymbol{x}] = \Delta E - \mathscr{W}[\boldsymbol{x}]. \tag{1.95}$$

Taking the average, we obtain

$$\langle \mathscr{Q}[\boldsymbol{x}]/k_\mathrm{B}T - \Delta s/k_\mathrm{B}\rangle = \frac{1}{k_\mathrm{B}T}\left[\cancel{\langle\Delta E\rangle} - \langle\mathscr{W}\rangle - \cancel{\langle\Delta E\rangle} + \Delta F + \Delta\mathscr{F}^{(\mathrm{f})} - \Delta\mathscr{F}^{(\mathrm{i})}\right], \tag{1.96}$$

where $\Delta\mathscr{F}$ is the difference of availability associated with the instantaneous probability distribution $P$ and the instantaneous value of $\lambda$ from the corresponding equilibrium free energy:

$$\Delta\mathscr{F}(\lambda, P) = \langle E_x(\lambda)\rangle_P - T\mathscr{S}[P]. \tag{1.97}$$

Seifert's integral relation intimates that

$$\langle \mathrm{e}^{-\mathscr{Q}[\boldsymbol{x}]/k_\mathrm{B}T + \Delta s/k_\mathrm{B}}\rangle = 1. \tag{1.98}$$

Thus, by Jensen's inequality we obtain

$$\langle -\mathscr{Q}[\boldsymbol{x}]/k_\mathrm{B}T + \Delta s/k_\mathrm{B}\rangle = -\frac{1}{k_\mathrm{B}T}\left[\langle\mathscr{W}\rangle - \Delta F - \Delta\mathscr{F}^{(\mathrm{f})} + \Delta\mathscr{F}^{(\mathrm{i})}\right] \le 0. \tag{1.99}$$

Denoting by $W$ the average work provided to the system during the transformation, we obtain

$$W - \Delta F \ge \Delta\mathscr{F}^{(\mathrm{f})} - \Delta\mathscr{F}^{(\mathrm{i})}. \tag{1.100}$$

This result generalizes Clausius' inequality to systems out of equilibrium. It was derived by Sagawa and Ueda [2010] and in a slightly different form by Esposito and van den Broeck [2011].

## 1.3 Stochastic thermodynamics: Systems out of equilibrium

### Steady states out of equilibrium

We have so far considered systems for which the transition matrix W satisfies detailed balance. Let us now consider what happens if this condition is violated. In this case there is a cycle $x_0 \longrightarrow x_1 \longrightarrow \cdots \longrightarrow x_n \longrightarrow x_0$ such that

$$\frac{W_{x_0 x_n} \cdots W_{x_2 x_1} W_{x_1 x_0}}{W_{x_0 x_1} \cdots W_{x_{n-1} x_n} W_{x_n x_0}} \neq 1. \tag{1.101}$$

Let us consider, e.g., a system with three states, whose energies are given by $E_x$, $x \in \{0,1,2\}$. We assume that the system is in contact with two reservoirs: one, at temperature $T_0$, for the transitions $0 \rightleftharpoons 1$ and $1 \rightleftharpoons 2$, and one at temperature $T_1$ for the transition $0 \rightleftharpoons 2$. Thus we have

$$\frac{W_{ij}}{W_{ji}} = \mathrm{e}^{-(E_j - E_i)/k_\mathrm{B} T_0}, \qquad \text{for } \{i,j\} = \{0,1\} \text{ or } \{i,j\} = \{1,2\}; \tag{1.102}$$

$$\frac{W_{20}}{W_{02}} = \mathrm{e}^{-(E_2 - E_0)/k_\mathrm{B} T_1}. \tag{1.103}$$

We have indeed

$$\frac{W_{02} W_{21} W_{10}}{W_{01} W_{12} W_{20}} = \exp\left[-(E_2 - E_0)\left(\frac{1}{k_\mathrm{B} T_0} - \frac{1}{k_\mathrm{B} T_1}\right)\right] \neq 1. \tag{1.104}$$

It can be shown under mild hypotheses that in this situation there is a unique probability distribution $P^\mathrm{ss} = (p_x^\mathrm{ss})$ over the state which is left invariant by the evolution, i.e., which satisfies

$$\sum_{x'} W_{xx'} p_{x'}^\mathrm{ss} = p_x^\mathrm{ss}. \tag{1.105}$$

However, in the steady state, there is a non-vanishing probability flux between some states, i.e., one has, for some pairs $(x, x')$,

$$J_{xx'}^\mathrm{ss} := W_{xx'} p_{x'}^\mathrm{ss} - W_{x'x} p_x^\mathrm{ss} \neq 0. \tag{1.106}$$

In our simple case, it is still true that

$$\frac{W_{xx'}}{W_{x'x}} = \mathrm{e}^{\Delta S_{xx'}^{(\mathrm{r})}/k_\mathrm{B}}. \tag{1.107}$$

However, the concerned reservoir is different for different transitions. If for each pair of states $(x, x')$ for which a transition is allowed the concerned reservoir is unique, we can proceed pretty much along the same lines as for systems obeying detailed balance, since the relation (1.33) still holds, provided $\Delta S^{(\mathrm{r})}$ is interpreted as the entropy change of all reservoirs with which the system is in contact. A more general situation is one in which the same transition can be triggered by coupling with different reservoirs (cf. Esposito [2012]). We shall discuss this situation later on. To ease the notation, in this section we shall use units in which Boltzmann's constant $k_\mathrm{B} = 1$.

## Fluctuation relations

From (1.107) we obtain Crooks' relation in the form

$$\frac{\mathscr{P}[\boldsymbol{x}|x(0)]}{\mathscr{P}[\widehat{\boldsymbol{x}}|\widehat{x}(0)]} = \prod_{t=0}^{t_\mathrm{f}-1} \frac{W_{x(t+1)x(t)}}{W_{x(t)x(t+1)}} = \mathrm{e}^{\Delta S^{(\mathrm{r})}[\boldsymbol{x}]}. \tag{1.108}$$

### Evans-Searles fluctuation relation

Multiplying this relation by $p^{\mathrm{ss}}_{x_0}/p^{\mathrm{ss}}_{x\mathrm{f}}$ and taking the log we obtain

$$\log \frac{\mathscr{P}^{\mathrm{ss}}[\boldsymbol{x}]}{\mathscr{P}^{\mathrm{ss}}[\widehat{\boldsymbol{x}}]} = \Delta S^{(\mathrm{r})}[\boldsymbol{x}] + \Delta S^{\mathscr{S}} = \Delta S^{\mathrm{tot}}, \tag{1.109}$$

where $\Delta S^{\mathrm{tot}}$ is the total entropy production. Summing over all paths $\boldsymbol{x}$ with a given value of $\Delta S^{\mathrm{tot}}$ yields the **fluctuation relation** (Evans and Searles [1994]):

$$\frac{P^{\mathrm{ss}}(\Delta S^{\mathrm{tot}})}{P^{\mathrm{ss}}(-\Delta S^{\mathrm{tot}})} = \mathrm{e}^{\Delta S^{\mathrm{tot}}}. \tag{1.110}$$

This obviously implies $\langle \Delta S^{\mathrm{tot}} \rangle \geq 0$ by Jensen's inequality, but leaves open the possibility that $\Delta S^{\mathrm{tot}} < 0$ for some realizations of the process. This behavior has been, rather unfortunately, called "transient violation of the second law" (Evans et al. [1993]). However, there is no violation of the second law, since the *average* of the entropy production is always positive. It is important to realize that entropy production is measured in units of $k_\mathrm{B}$. Thus if the considered system is too large, or the duration of the path is too long, the probability of observing negative values of the entropy production becomes extremely small.

## Distribution function for the entropy

Let us now evaluate the joint probability $\Phi_x(S,t)$ distribution of the state $x$, and of the total entropy $S$ exchanged with the reservoirs up to time $t$. This quantity obeys the evolution equation

$$\Delta\Phi_x(S,t+1) = W_{x,x}\Phi_x(S,t) + \sum_{x'(\neq x)} W_{xx'}\Phi_{x'}(S-\Delta S_{xx'},t) - \Phi_x(S,t). \tag{1.111}$$

Let us remind of the identity

$$f(x-\Delta) = \mathrm{e}^{-\Delta\partial/\partial x} f(x), \tag{1.112}$$

which holds for each analytic function $f(x)$ and can be proved by Taylor expanding both sides of the relation. Applying this identity the previous relation assumes the form

$$\Delta\Phi_x(S,t+1) = \sum_{x'(\neq x)} \left[ W_{xx'}\mathrm{e}^{-\Delta S_{xx'}\partial/\partial S}\Phi(x',S,t) - W_{x'x}\Phi(x,S,t)\right], \tag{1.113}$$

where the entropy $\Delta S_{xx'}$ yielded to the reservoirs is given by (1.107) which we can rewrite as

$$\Delta S_{xx'} = \log \frac{W_{xx'}}{W_{x'x}}. \tag{1.114}$$

We now define the moment-generating function $\Psi_x(\mu, t)$ by

$$\Psi_x(\mu, t) := \int \mathrm{d}S \, \mathrm{e}^{\mu S} \, \Phi_x(S, t). \tag{1.115}$$

Then the dependence on $\Delta S_{xx'}$ in equation (1.113) transforms according to

$$\mathrm{e}^{-\Delta S_{xx'} \partial/\partial S} \longrightarrow \mathrm{e}^{-\mu \Delta S_{xx'}} = \left( \frac{W_{xx'}}{W_{x'x}} \right)^{\mu}. \tag{1.116}$$

As a consequence, $\Psi_x(\mu, t)$ satisfies the evolution equation

$$\Psi_x(\mu, t+1) = \sum_{x'} W_{xx'} \left( \frac{W_{xx'}}{W_{x'x}} \right)^{\mu} \Psi_{x'}(\mu, t) = \big( \mathsf{T}(\mu) \cdot \Psi \big)_x . \tag{1.117}$$

This result was obtained by Lebowitz and Spohn [1999]. Thus, if $\Lambda_{\mathrm{max}}(\mu)$ is the largest eigenvalue of $\mathsf{T}(\mu)$, we have, for large values of $t$,

$$\Psi_x(\mu, t) \propto \mathrm{e}^{-t\psi(\mu)}, \tag{1.118}$$

where

$$\psi(\lambda) = -\log \Lambda_{\mathrm{max}}(\mu). \tag{1.119}$$

In this way $\psi(\mu)$ appears as a large-deviation function (cf. Ellis [1985], Touchette [2009]) for the rate of entropy yielded to the reservoir in a time interval of duration $t$. Indeed, if we define the time-averaged entropy rate $\Sigma$ by

$$\Sigma = \frac{S}{t}, \tag{1.120}$$

the probability distribution function of $\Sigma$ satisfies a large-deviation principle. In fact, evaluating the inverse Laplace transform of $\Psi(\mu, t)$ for large values of $t$ by the saddle-point method, we obtain the following estimate of the the probability density $P_t(\Sigma)$ that the time-averaged value of the entropy yielded to the reservoir during a time $t$ is equal to $\Sigma$:

$$P_t(\Sigma) \propto \mathrm{e}^{-tI(\Sigma)}, \tag{1.121}$$

where the **rate function** $I(\Sigma)$ is the Legendre transform of $\psi(\mu)$:

$$I(\Sigma) = \min_{\mu} \big( \psi(\mu) + \mu \Sigma \big) \tag{1.122}$$

### The Gallavotti-Cohen symmetry

Let us evaluate $\mathsf{T}(\mu-1)$. We obtain

$$T_{xx'}(\mu-1) = W_{xx'}\left(\frac{W_{xx'}}{W_{x'x}}\right)^{\mu-1} = W_{x'x}\left(\frac{W_{x'x}}{W_{xx'}}\right)^{-\mu} = T_{x'x}(-\mu). \tag{1.123}$$

Thus $\mathsf{T}(\mu-1)$ and $\mathsf{T}(-\mu)$ are one the transposed of the other and have therefore the same spectrum. Thus their largest eigenvalues coincide, and one has

$$\psi(-\mu) = \psi(\mu-1). \tag{1.124}$$

Taking the Legendre transform, we obtain the following symmetry for the rate equation:

$$I(-\Sigma) = I(\Sigma) + \Sigma, \tag{1.125}$$

that can be interpreted as

$$\frac{P_t(\Sigma)}{P_t(-\Sigma)} = \mathrm{e}^{t\Sigma}. \tag{1.126}$$

This relation is known as the **Gallavotti-Cohen fluctuation theorem**, since it was derived by them in the context of hamiltonian dynamics (with some rather strong ergodicity conditions) (Gallavotti and Cohen [1995a], Gallavotti and Cohen [1995b]). As a consequence, the relation (1.124) will be called the Gallavotti-Cohen symmetry. Gallavotti [1996] showed that this relation implies, for states close to equilibrium, both the Fluctuation-Dissipation theorem and the reciprocity law of kinetic coefficients derived by Onsager.

Note that the fluctuation relation (1.110) holds for *finite times*, starting from the steady state, provided one takes for $\Delta S^{\mathrm{tot}}$ the total change in entropy of the system and the reservoirs. However, since the change in entropy of the system is bounded, while the change of the entropy in the reservoirs grows with time, we have for large values of $t$

$$\Delta S^{\mathrm{tot}} \simeq \Delta S^{(\mathrm{r})}. \tag{1.127}$$

One has however to take into account the fact that the probability of observing negative entropy production vanishes exponentially as $t$ grows.

### Transitions along different channels

Let us consider a situation in which the transition form a state $j$ to a state $i$ can take place with different modalities, identified by the variable $\alpha$: the transition rate with modality $\alpha$ is thus given by $W_{ij}^{\alpha}$, and the total transition rate is given by

$$W_{ij}^{\mathrm{tot}} = \sum_{\alpha} W_{ij}^{\alpha}. \tag{1.128}$$

The entropy exchange associated with the transition in modality $\alpha$ is given by

$$\Delta S_{ij}^{\alpha} = \log\frac{W_{ij}^{\alpha}}{W_{ji}^{\alpha}} \tag{1.129}$$

and the expected exchange rate associated with a transition is given by

$$\langle \dot{S}_{ij} \rangle = \sum_{\alpha} W_{ij}^{\alpha} \log \frac{W_{ij}^{\alpha}}{W_{ji}^{\alpha}}. \tag{1.130}$$

On the other hand, if we neglect the difference between different modalities, we have

$$\Delta S_{ij}^{\mathrm{tot}} = \log \frac{W_{ij}^{\mathrm{tot}}}{W_{ji}^{\mathrm{tot}}}, \tag{1.131}$$

yielding an exchange rate

$$\langle \dot{S}_{ij}^{\mathrm{tot}} \rangle = W_{ij}^{\mathrm{tot}} \log \frac{W_{ij}^{\mathrm{tot}}}{W_{ji}^{\mathrm{tot}}}. \tag{1.132}$$

Let us set

$$W_{ij}^{\alpha} = W_{ij}^{\mathrm{tot}} p_{ij}^{\alpha}, \tag{1.133}$$

so that $\sum_{\alpha} p_{ij}^{\alpha} = 1$. Then

$$\langle \dot{S}_{ij} \rangle = W_{ij}^{\mathrm{tot}} \log \frac{W_{ij}^{\mathrm{tot}}}{W_{ji}^{\mathrm{tot}}} + W_{ij}^{\mathrm{tot}} \sum_{\alpha} p_{ij}^{\alpha} \log \frac{p_{ij}^{\alpha}}{p_{ji}^{\alpha}}. \tag{1.134}$$

The last term is always positive. Thus (as it is natural to expect) neglecting the different modalities yields an underestimation of the entropy production. This is relevant, e.g., when nonequilibrium is generated by connecting the system with different thermal or chemical reservoirs.

## Adiabatic and non-adiabatic entropy production

In a stationary system without detailed balance a certain amount of dissipation is needed to maintain the steady state. Let us consider the entropy $\Delta S_{x'x}$ yielded to the environment upon a transition $x \longrightarrow x'$. We can write this quantity in the following way, introducing the instantaneous probability distribution $p_x$:

$$\Delta S_{xx'} = \log \frac{W_{xx'}}{W_{x'x}} = \underbrace{\log \frac{W_{xx'} p_{x'}}{W_{x'x} p_x}}_{\Delta_{\mathrm{i}} S_{xx'}} \underbrace{- \log \frac{p_{x'}}{p_x}}_{\Delta \mathscr{S}_{xx'}^{(\mathsf{S})}}. \tag{1.135}$$

We recognize in the first term the entropy production $\Delta_{\mathrm{i}} S$ (cf. equation (1.40)), while the second term is the change in the system's fluctuating entropy. Let us now evaluate the average of the first term as a function of the instantaneous probability distribution $P$. We obtain

$$\langle \Delta_{\mathrm{i}} S \rangle = \frac{1}{2} \sum_{xx'} \underbrace{\left( W_{xx'} p_{x'} - W_{x'x} p_x \right)}_{J_{xx'}} \underbrace{\log \frac{W_{xx'} p_{x'}}{W_{x'x} p_x}}_{X_{xx'}}. \tag{1.136}$$

The quantity $J_{xx'}$ is the net probability current from $x'$ to $x$ in the state considered. It appears multiplied by a "force" $X_{xx'}$ given by

$$X_{xx'} := \log \frac{W_{xx'} p_{x'}}{W_{x'x} p_x} = \underbrace{\log \frac{W_{xx'} p^{\mathrm{ss}}_{x'}}{W_{x'x} p^{\mathrm{ss}}_x}}_{X^{(\mathrm{ad.})}_{xx'}} + \underbrace{\log \frac{p_{x'} p^{\mathrm{ss}}_x}{p_x p^{\mathrm{ss}}_{x'}}}_{X^{(\mathrm{n.ad.})}_{xx'}}, \tag{1.137}$$

where we have introduced the steady-state probability distribution $P^{\mathrm{ss}} = (p^{\mathrm{ss}}_x)$. The force appears therefore split into an adiabatic contribution (the first term) and a non-adiabatic one (the second term). Note that the non-adiabatic term vanishes if the system is in the steady state. It is then easy to show that the average entropy production is split into two contributions of the form $\sum J X$, which are separately non-negative on average (Esposito and Van den Broeck [2010]):

$$\langle \Delta_{\mathrm{i}} S \rangle = \underbrace{\frac{1}{2} \sum_{xx'} J_{xx'} X^{(\mathrm{ad.})}_{xx'}}_{\Delta_{\mathrm{i}} S^{(\mathrm{ad.})} \geq 0} + \underbrace{\frac{1}{2} \sum_{xx'} J_{xx'} X^{(\mathrm{n.ad.})}_{xx'}}_{\Delta_{\mathrm{i}} S^{(\mathrm{n.ad.})} \geq 0}. \tag{1.138}$$

The adiabatic contribution is also called "housekeeping heat", because it is the entropy generated in order to maintain the system in its non-equilibrium steady state.

The adiabatic contribution satisfies an integral fluctuation relation, from which the fact that it is non-negative on average follows by means of Jensen's inequality. Let us define $\varphi_{x'x}$ by

$$\varphi_{x'x} := \mathrm{Prob}^{\mathrm{ss}}(x \longrightarrow x') = W_{x'x} p^{\mathrm{ss}}(x). \tag{1.139}$$

We then have

$$\varphi_{x'x} \geq 0, \quad \forall x, x'; \qquad \sum_{x'x} \varphi_{x'x} = 1, \quad \forall x. \tag{1.140}$$

Therefore $(\varphi_{x'x})$ is a probability distribution. Define moreover

$$\widehat{\varphi}_{xx'} = \varphi_{x'x}, \tag{1.141}$$

which corresponds to a probability distribution, too. We then have the integral fluctuation theorems (Speck and Seifert [2005], Esposito and Van den Broeck [2010]):

$$\langle \mathrm{e}^{-\Delta_{\mathrm{i}} S^{(\mathrm{ad})}} \rangle = \sum_{xx'} \frac{\varphi_{xx'}}{\varphi_{x'x}} \varphi_{x'x} = \sum_{xx'} \varphi_{xx'} = 1; \tag{1.142}$$

$$\langle \mathrm{e}^{-\Delta_{\mathrm{i}} S^{(\mathrm{n.ad.})}} \rangle = \sum_{xx'} W_{xx'} p_{x'} \frac{p_x p^{\mathrm{ss}}_{x'}}{p_{x'} p^{\mathrm{ss}}_x} = \sum_x p_x = 1. \tag{1.143}$$

These relation imply, by Jensen's inequality,

$$\langle \Delta_{\mathrm{i}} S^{(\mathrm{ad.})} \rangle \geq 0; \tag{1.144}$$

$$\langle \Delta_{\mathrm{i}} S^{(\mathrm{n.ad})} \rangle \geq 0. \tag{1.145}$$

## The Hatano-Sasa relation

The Hatano-Sasa relation, that we are now going to discuss, is a relation that applies to *general non-equilibrium systems* when their dynamics is manipulated. Therefore it does not require that the underlying dynamics satisfies microscopic reversibility. The only formal requirement is that the dynamics is *ergodic*, i.e., that the invariant probability distribution of the dynamics, when the manipulation parameter is fixed, is unique and attainable (this requires that there are no absorbing states and some further rather mild conditions). Let us consider a system with transition probabilities $\mathsf{W}(\lambda) = W_{xx'}(\lambda)$ which depend on a parameter $\lambda$, that can be manipulated according to a protocol $\boldsymbol{\lambda} = (\lambda(t))$. For each given value of $\lambda$ there is a steady-state probability distribution $P^{\mathrm{ss}}(\lambda) = (p_x^{\mathrm{ss}}(\lambda))$ on the microstates, which satisfies

$$p_x^{\mathrm{ss}}(\lambda) = \sum_{x'} W_{xx'}(\lambda)\, p_{x'}^{\mathrm{ss}}(\lambda). \tag{1.146}$$

Let us fix the protocol $\boldsymbol{\lambda}$. Then the conditional probability $\mathscr{P}[\boldsymbol{x}|x(0)]$, given its initial state $x(0)$, is given by

$$\mathscr{P}[\boldsymbol{x}|x(0)] = \prod_{t=0}^{t_{\mathrm{f}}-1} W_{x(t+1)x(t)}(\lambda(t)) \tag{1.147}$$

The probability of a path $\boldsymbol{x}$, assuming that the initial state is described by the steady-state distribution $P^{\mathrm{ss}}(\lambda_0)$ corresponding to the initial value of $\lambda$, is given by

$$\mathscr{P}[\boldsymbol{x}] = \mathscr{P}[\boldsymbol{x}|x(0)]\, p_{x(0)}^{\mathrm{ss}}(\lambda_0). \tag{1.148}$$

Let us define the quantity $\phi_x(\lambda)$ by

$$\phi_x(\lambda) := -\log p_x^{\mathrm{ss}}(\lambda). \tag{1.149}$$

This appears like a fluctuating entropy in the steady state of the dynamics. If the system undergoes a transition $x \longrightarrow x'$, we have of course

$$\Delta\phi_{x'x}(\lambda) := \phi_{x'}(\lambda) - \phi_x(\lambda). \tag{1.150}$$

Let us evaluate the sum of $\Delta\phi$ over a path $\boldsymbol{x}$, when the system is manipulated according to a protocol $\boldsymbol{\lambda}$:

$$\begin{aligned}
\Phi[\boldsymbol{x}] &:= \sum_{t=0}^{t_{\mathrm{f}}-1} \Delta\phi_{x(t+1)x(t)}(\lambda(t)) = \sum_{t=0}^{t_{\mathrm{f}}-1} \left[\phi_{x(t+1)}(\lambda(t)) - \phi_{x(t)}(\lambda(t))\right] \\
&= \phi_{x(t_{\mathrm{f}})}(\lambda(t_{\mathrm{f}})) - \sum_{t=0}^{t_{\mathrm{f}}-1} \left[\phi_{x(t+1)}(\lambda(t+1)) - \phi_{x(t+1)}(\lambda(t))\right] - \phi_{x(0)}(\lambda(0)) \\
&= \phi_{x(t_{\mathrm{f}})}(\lambda(t_{\mathrm{f}})) - \phi_{x(0)}(\lambda(0)) - \mathscr{A}[\boldsymbol{x}],
\end{aligned} \tag{1.151}$$

where we have defined $\mathscr{A}[\boldsymbol{x}]$ by

$$\mathscr{A}[\boldsymbol{x}] := \sum_{t=0}^{t_{\mathrm{f}}-1} \left[\phi_{x(t+1)}(\lambda(t+1)) - \phi_{x(t+1)}(\lambda(t))\right]. \tag{1.152}$$

### The Hatano-Sasa relation

Let us define $\Psi_x(t_\mathrm{f})$, for all paths $\boldsymbol{x}$ taking place between $t=0$ and $t=t_\mathrm{f}$ by the relation

$$\Psi_x(t_\mathrm{f}) := \sum_{\boldsymbol{x}:\, x(t)=x} \mathrm{e}^{-\mathscr{A}[\boldsymbol{x}]}\mathscr{P}[\boldsymbol{x}], \tag{1.153}$$

where $\mathscr{A}$ is defined by (1.152). We can show that $\Psi_x(t)$ has the expression

$$\Psi_x(t) = p_x^{\mathrm{ss}}(\lambda(t)). \tag{1.154}$$

One has indeed

$$\Psi_x(t{=}0) = p_x^{\mathrm{ss}}(\lambda_0). \tag{1.155}$$

One then has, assuming (1.154) holds for $t-1$,

$$\begin{aligned}\Psi_x(t) &= \sum_{x'} W_{xx'}(\lambda(t))\frac{p_{x'}^{\mathrm{ss}}(\lambda(t))}{p_{x'}^{\mathrm{ss}}(\lambda(t-1))}p_{x'}^{\mathrm{ss}}(\lambda(t-1)) = \sum_{x'} W_{xx'}(\lambda(t))p_{x'}^{\mathrm{ss}}(\lambda(t))\\ &= p_x^{\mathrm{ss}}(\lambda(t)).\end{aligned} \tag{1.156}$$

Summing over $\boldsymbol{x}$, one obtains a relation similar to Seifert's for $\mathscr{A}$:

$$\langle \mathrm{e}^{-\mathscr{A}}\rangle = 1. \tag{1.157}$$

This is the Hatano and Sasa [2001] relation. The quantity $\mathscr{A}$ can be connected, via (1.151), to the non-adiabatic entropy production defined in (1.138). Prost et al. [2009] have exploited this relation to obtain an analog of the Fluctuation-Dissipation theorem for general non-equilibrium systems.

## Biased ensembles

We have found in several cases the need to evaluate averages of the form

$$\psi(\mu,t) := \langle \mathrm{e}^{\mu\mathscr{Q}}\rangle = \sum_{\boldsymbol{x}}\mathscr{P}[\boldsymbol{x}]\mathrm{e}^{\mu\mathscr{Q}[\boldsymbol{x}]}, \tag{1.158}$$

where $\mathscr{Q}[\boldsymbol{x}]$ is some functional of the path $\boldsymbol{x}$. We now try and discuss the ways in which these averages can be evaluated. Let us first remark that, by taking the derivative of $\psi(\mu,t)$ with respect to $\mu$, we obtain

$$\frac{\partial\psi(\mu,t)}{\partial\mu} = \langle\mathscr{Q}\rangle_\mu\,\psi(\mu,t), \tag{1.159}$$

where we have defined the **biased average**

$$\langle\cdots\rangle_\mu := \sum_{\boldsymbol{x}}\cdots\frac{\mathrm{e}^{\mu\mathscr{Q}[\boldsymbol{x}]}}{\psi(\mu,t)}\mathscr{P}[\boldsymbol{x}]. \tag{1.160}$$

Integrating (1.159) we obtain

$$\psi(\mu,t) = \exp\left[\int_0^\mu \mathrm{d}\mu'\;\langle\mathscr{Q}\rangle_{\mu'}\right]. \tag{1.161}$$

Thus the problem boils down to the evaluation of the biased average $\langle\mathcal{Q}\rangle_\mu$. The difficulty lies in the fact that in its expression (1.160) appears the same function $\psi(\lambda,t)$ that we wish to evaluate. However, it only appears as a normalizing factor. Let us introduce an arbitrary (positive-valued) functional $\Pi[\boldsymbol{x}]$ of the path $\boldsymbol{x}$. Then one can express $\langle\mathcal{Q}\rangle_\mu$ in the following way:

$$\langle\mathcal{Q}\rangle_\mu = \frac{\sum_{\boldsymbol{x}}(\mathcal{Q}[\boldsymbol{x}]/\Pi[\boldsymbol{x}])\Pi[\boldsymbol{x}]\mathscr{P}[\boldsymbol{x}]\mathrm{e}^{\mu\mathcal{Q}[\boldsymbol{x}]}}{\sum_{\boldsymbol{x}}(1/\Pi[\boldsymbol{x}])\Pi[\boldsymbol{x}]\mathscr{P}[\boldsymbol{x}]\mathrm{e}^{\mu\mathcal{Q}[\boldsymbol{x}]}} = \frac{\langle\mathcal{Q}/\Pi\rangle_{\mu,\Pi}}{\langle 1/\Pi\rangle_{\mu,\Pi}}, \tag{1.162}$$

where we have defined

$$\langle\cdots\rangle_{\mu,\Pi} = \sum_{\boldsymbol{x}}(\cdots)\Pi[\boldsymbol{x}]\mathscr{P}[\boldsymbol{x}]\mathrm{e}^{\mu\mathcal{Q}[\boldsymbol{x}]}. \tag{1.163}$$

To be more concrete, let us consider a stationary non-equilibrium Markov process, with an initial distribution $P^{(0)} = p_x^{(0)}$. Then the path probability $\mathscr{P}[\boldsymbol{x}]$ is given by

$$\mathscr{P}[\boldsymbol{x}] = W_{x_N,x_{N-1}}W_{x_N,x_{N-1}}\cdots W_{x_1,x_0}p_x^{(0)}. \tag{1.164}$$

We also assume that we wish to evaluate (1.158) where $\mathcal{Q}[\boldsymbol{x}]$ is the total entropy exchange with the reservoir over a certain period of time of duration, say, $t$. Then $\Psi(\mu,t)$ satisfies the Lebowitz-Spohn equation of motion (1.117), which does not conserve probability. Let us define a **biased process** satisfying the evolution equation

$$\Psi_x(\mu,t+1) = \sum_{x'}\widetilde{W}_{xx'}\Psi_{x'}(\mu,t), \tag{1.165}$$

where the transition rates $\widetilde{W}_{xx'}$ are defined by

$$\begin{aligned}\widetilde{W}_{xx'} &:= W_{xx'}\left(\frac{W_{x'x}}{W_{xx'}}\right)^\mu \quad \text{for } x\neq x';\\ \widetilde{W}_{xx} &:= 1-\sum_{x'(\neq x)}\widetilde{W}_{x'x}.\end{aligned} \tag{1.166}$$

These rates differ from those of the Lebowitz-Spohn equation by the diagonal term. However, they now conserve probability and thus generate a *bona fide* Markov process. Let us now introduce the functional $\Pi[\boldsymbol{x}]$, defined by

$$\Pi[\boldsymbol{x}] := \prod_{k=1}^{M}\Pi_{x_k,x_{k-1}}(t_k), \tag{1.167}$$

where

$$\Pi_{xx'}(t) := \begin{cases}1, & \text{if } x\neq x';\\ \\ \widetilde{W}_{xx}(t)/W_{xx}(t), & \text{if } x = x'.\end{cases} \tag{1.168}$$

We can then define the **biased weight** of the paths $\boldsymbol{x}$:

$$\begin{aligned}\mathscr{P}_\lambda[\boldsymbol{x}] &:= \mathscr{P}[\boldsymbol{x}]\Pi[\boldsymbol{x}]\,\mathrm{e}^{\mu\mathcal{Q}[\boldsymbol{x}]} &(1.169)\\ &= \widetilde{W}_{x_N,x_{N-1}}\widetilde{W}_{x_{N-1},x_{N-2}}\cdots\widetilde{W}_{x_1,x_0}p_{x_0}^{(0)}. &(1.170)\end{aligned}$$

Therefore generating the paths via the process defined by (1.165) and weighting them according to equation (1.162) we can evaluate $\langle \mathcal{Q} \rangle_\mu$ and, by integrating over $\mu$, $\Psi(\mu, t)$. The large-deviation function can then be obtained in the large-time limit:

$$g(\mu) = \lim_{t\to\infty} -\frac{1}{t} \log \Psi(\mu, t). \tag{1.171}$$

This method has been applied, e.g., to the asymmetric exclusion process (ASEP) by Imparato and Peliti [2007]. The problem with this technique is that the biased process tends to sample poorly the full distribution. However, one can improve the statistics by monitoring the behavior of $1/\Pi[\boldsymbol{x}]$, trying to obtain a good compromise between sampling and reweighting. The problem can also be circumvented by using a cloning approach, introduced by Giardinà et al. [2006], which is described in detail in Vivien Lecomte's lectures at this Workshop.

## 1.4 Thermodynamics of Information

There is an excellent recent review on the topic: Parrondo et al. [2015]. A number of important papers are collected and discussed in Leff and Rex [2002]. The connections between thermodynamics and information predate the birth of information theory by several decades, since when Maxwell, and then Boltzmann and Gibbs, emphasized the statistical nature of the second law.

### Maxwell's demon

In his *Theory of Heat* (1871), J. C. Maxwell illustrated the statistical nature of the second principle by this very famous argument:

> Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower molecules to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

Thus if a demon has a way to ascertain the energy of a single molecule, it can use this information to reduce the entropy of the vessel, driving it away from equilibrium. It would then be in principle possible to exploit this disequilibrium to extract work.

**Szilard's engine**

Szilárd [1929] proposed an intriguing conceptual experiment that made the connection between the second law and information more precise. (The paper is translated in Szilard [2006] and can be found in Leff and Rex [2002].)

- Consider a closed cylindrical vessel of volume $V$, in contact with a heat reservoir at temperature $T$.

- Let the vessel contain an ideal gas, made of a single molecule. The molecule is initially free to wander in the whole cylinder.

- Assume that the vessel can be separated into two possibly unequal sections of volumes $V_{\mathrm{L}}$ (left) and $V_{\mathrm{R}} = V - V_{\mathrm{L}}$ (right) respectively by inserting a partition that can be moved left or right in the cylinder, thus becoming a piston.

- Let an agent initially insert the piston, and then ascertain by some means if the molecule is in the left (L) or in the right (R) section of the cylinder.

- The following procedure depends on the case:

    - If the particle is in L, let the piston be slowly moved to the right, until it reaches the end of the cylinder. The molecule will keep bouncing on the piston, yielding a work $-W$ which is given by the usual expression for an isothermal expansion:

      $$-W_{\mathrm{L}} = \int_{V_{\mathrm{L}}}^{V} p\,\mathrm{d}V = k_{\mathrm{B}} T \log \frac{V}{V_{\mathrm{L}}}. \tag{1.172}$$

      This will take place with a probability $p_{\mathrm{L}} = V_{\mathrm{L}}/V$.

    - By the same token, if the particle is on the right, the piston will be moved to the left, until it reaches the end of the cylinder. One thus obtains a work $-W_{\mathrm{R}} = k_{\mathrm{B}} T \log(V/V_{\mathrm{R}})$. This will take place with a probability $p_{\mathrm{R}} = V_{\mathrm{R}}/V$.

- At the end of the expansion, the partition is removed and the cylinder returns to the initial state. The cycle can then be repeated.

- In this way, the agent can extract from the reservoir a work $-W$ which is equal to

  $$-W = k_{\mathrm{B}} T \left[ p_{\mathrm{L}} \log \frac{V}{V_{\mathrm{L}}} + p_{\mathrm{R}} \log \frac{V}{V_{\mathrm{R}}} \right] = -k_{\mathrm{B}} T \sum_{\ell \in \{\mathrm{L},\mathrm{R}\}} p_\ell \log p_\ell = k_{\mathrm{B}} T \mathscr{S}[P] \tag{1.173}$$

  per cycle. The maximum value of this expression is obtained when the two sections have equal volume, and $-W = k_{\mathrm{B}} T \log 2$.

In this way information about the location of the particle in the cylinder is turned into work extraction. This idea was generalized by Bennett and popularized by Feynman, by considering a machine in which a very long sequence of microscopic vessels, each containing a single molecule in a *known* L or R location, are exploited to extract work by a reservoir ([Feynman, 1996, p.146–147]).

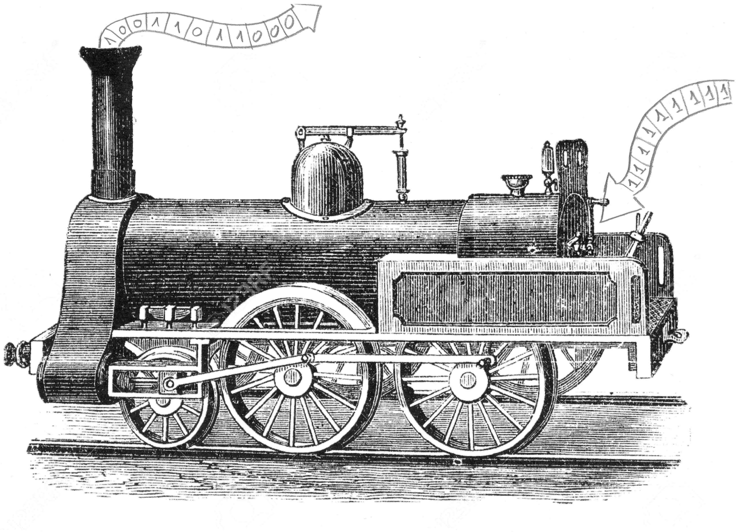


Figure 1.3: The Bennett-Feynman information-fueled engine. A tape containing a large number of Szilard cylinders, with the molecule in state 1 (R), is fed into the machine. Inside, each cylinder undergoes Szilard's manipulation, and work $-W = k_{\mathrm{B}} T \log 2$ is extracted on average for each site in the tape. At the end of the manipulation the location of the molecule in the cylinder is randomized. The net result is that the information contained in the tape is randomized, and work $-W = k_{\mathrm{B}} T \log 2$ for each tape location is extracted.

## Landauer's principle

How can this result be reconciled with the second principle? The Second Principle, in Lord Kelvin's words (Thomson [1851]), intimates that

> it is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects.

Yet the Szilard engine appears to accomplish exactly that. We have to consider the effects of the manipulation on the total system, made by the particle and the measurement device. Let us assume, e.g., that the two variables, i.e., the location $x \in \{\mathrm{L}, \mathrm{R}\}$ of the particle and its measured location $y \in \{\mathrm{L}, \mathrm{R}\}$, are initially independent and each distributed with equal probabilities on the two possible states. Before the measurement, the total system can be in any of the four states (L,L), (L,R), (R,L) or (R,R) with equal probability. The Shannon entropy of this distribution is equal to $\log 4 = 2 \log 2$. After the measurement (assumed error-free) only the states (L,L) or (R,R) are available. Therefore the entropy of the total system (with this initial condition) has been reduced to $\log 2$ by the measurement, and this cannot be achieved without providing some work to the system, equal at least to $-T\,\Delta S = k_{\mathrm{B}} T \log 2$. By the same token, after the measurement and the manipulation have taken place, the states of the particle and of the measurement devices are uncorrelated, and the total system can find itself in any of the four states mentioned above.

Let us consider a mechanical system evolving by Hamiltonian dynamics. Let the initial condition be described by a distribution in a region $\Gamma_0$ of phase space, of volume $\Omega_0$. If the system is manipulated, but prevented from exchanging heat with its surroundings, the volume $\Omega_t$ of the evolved region $\Gamma_t$ will be equal to the initial one by Liouville's theorem. If on the contrary we find that the phase-space volume has been reduced, we can conclude that the system has exchanged heat with its surrounding, yielding them a quantity $\mathscr{Q}$ of heat such that the entropy increase of the surroundings more than compensates its own entropy loss. This argument is weak because the concept of entropy is well established only for thermodynamic systems at equilibrium. However, here we suppose to be measuring only the entropy increase (or decrease!) of the heat bath, that is supposed to be always at equilibrium, and for which therefore the thermodynamic notion of entropy applies.

To be definite, let $X$ denote the possible microstates of a system at equilibrium at temperature $T$.[3] The energy of microstate $x$ is given by $E_x$. Let the system be connected to a measuring apparatus with states $Y$, all of equal energy (for simplicity). The measurement is a manipulation that introduces correlations between the states of $X$ and $Y$. These correlations are measured by the **mutual information** $I(X;Y)$ defined by

$$I(X;Y) = \sum_{x,y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x) P_Y(y)}. \tag{1.174}$$

Since this expression has the form of the Kullback-Leibler divergence between the joint probability distribution $P_{XY}$ of $X$ and $Y$ and the product $P_X P_Y$ of the marginal distributions, it is immediate to see that it is non-negative, and vanishes only if $X$ and $Y$ are independent. Moreover, it is symmetrical upon the exchange of $X$ and $Y$. Now, if the system and the measuring apparatus are not coupled, the equilibrium distribution of the $(X,Y)$ system factorizes into $P_X^{\mathrm{eq}} P_Y^{\mathrm{eq}}$. Let us assume that the measurement yields a joint probability distribution $P_{XY}$ with some mutual dependence. If the measurement does not perturb the state of $X$, the marginal distribution $P_X$ remains invariant. We can also assume that the measuring apparatus is optimized so that the marginal $P_Y(y)$ after the measurement is equal to the initial one. (It is left as an exercise to check what happens if this is not true.) Let us evaluate the availability $\mathscr{F}$ after the measurement: we obtain

$$\mathscr{F} = \langle E \rangle_{P_{XY}} + k_{\mathrm{B}} T \sum_{x,y} P_{XY}(x,y) \log P_{XY}(x,y), \tag{1.175}$$

while the equilibrium free energy, by the factorization of the probabilities, is given by

$$\mathscr{F}^{\mathrm{eq}} = \langle E \rangle_{P_X^{\mathrm{eq}}} + k_{\mathrm{B}} T \sum_{x,y} P_X^{\mathrm{eq}}(x) P_Y^{\mathrm{eq}}(y) \log \left( P_X^{\mathrm{eq}}(x) P_Y^{\mathrm{eq}}(y) \right). \tag{1.176}$$

---

[3] Much of this discussion I gathered from TE Ouldridge (Imperial College, London) in the Luxembourg Workshop on Chemical Reaction Networks.

In our hypotheses, $\langle E\rangle_{P_{XY}} = \langle E\rangle_{P_X}$. Thus

$$\mathscr{F} - \mathscr{F}^{\mathrm{eq}} = k_{\mathrm{B}}T\sum_{x,y} P_{XY}(x,y)\log\frac{P_{XY}(x,y)}{P_X^{\mathrm{eq}}(x)P_Y^{\mathrm{eq}}(y)} = k_{\mathrm{B}}T\,I(X;Y). \tag{1.177}$$

According to (1.8), the maximum work that can be extracted from the relaxation of the system to equilibrium is given by $\mathscr{F} - \mathscr{F}^{\mathrm{eq}}$. Thus we obtain that a measurement allows one to extract from the reservoir a work $-\mathscr{W}$ whose average satisfies

$$-W = -\langle \mathscr{W}\rangle \le k_{\mathrm{B}}T\,I(X;Y). \tag{1.178}$$

To reconcile this result with the second principle we have to admit that during the manipulation corresponding to the measurement a quantity of heat equal at least to $W$ has been provided to the surrounding, increasing the entropy of the reservoir by at least $k_{\mathrm{B}}I(X;Y)$. Thus the operation of the Szilárd engine can be summarized as follows:

1. Let us start from the situation in which the partition is placed in the middle of the cylinder, but no measurement has yet been done. The Shannon entropy of the $(X,Y)$ system is equal to $2\log 2$.

2. The measurement is performed, and now one has $y = x \in \{\mathrm{L},\mathrm{R}\}$. The Shannon entropy of the system is equal to $\log 2$. Therefore an amount of work equal at least to $k_{\mathrm{B}}T\log 2$ has been performed on the system, and passed as heat to the reservoir.

3. By performing the free expansion, an amount of work no larger than $k_{\mathrm{B}}T\log 2$ is gleaned from the reservoir.

4. The partition is placed again in the middle of the cylinder, and the first state is recovered, since the correlation between $X$ and $Y$ is now broken.

The working of the Szilard engine can be analyzed from the point of view of stochastic thermodynamics. This applies more generally to the ways of exploiting the information on a thermodynamic system obtained by measurement to extract work from a heat reservoir.

- Let us consider a system at equilibrium, whose states are distributed by the “a priori” probability $P_X^{\mathrm{eq}} = (p_x^{\mathrm{eq}})$.

- A measurement apparatus yields a signal $M$ correlated with the state of the system. Thus, having received the signal $m$ the distribution of the system is given by the conditional $P_m = (p_{x|m})$.

- Let us denote by $P_{X|M} = (p_{x,m})$ the joint probability of $X$ and $M$. We have $p_{x,m} = p_{x|m}p_m$.

- Assume that the internal energy $\langle E\rangle_{P_X}$ does not vary upon the measurement.

- We can now evaluate the change in the availability before and after the measurement:

$$\mathscr{F}^{(\mathrm{f})} - \mathscr{F}^{(\mathrm{i})} = k_{\mathrm{B}} T \sum_{x,m} p_{x,m} \log \frac{p_{x,m}}{p_x p_m} \tag{1.179}$$

$$= k_{\mathrm{B}} T \underbrace{\sum_{x,m} p_{x,m} \log \frac{p_{x|m}}{p_x}}_{I(X;M)}. \tag{1.180}$$

  Here $I(X;M)$ is the mutual information between the system $X$ and the measurement $M$.

- By exploiting the relation (1.10) between the differences in availability and the work $W$ to be performed on the system we obtain the inequality

$$W \geq -k_{\mathrm{B}} T \, I(X;M). \tag{1.181}$$

One can also consider the reverse process, in which a system is manipulated in order to be brought to a reference state (i.e., by reducing its entropy). This allows us to obtain Landauer's bound in the context of stochastic thermodynamics.

- Assume again that $\langle \Delta E \rangle$ does not vary upon the manipulation.
- Consider, e.g., a two-state memory: $P^{(\mathrm{i})} = (p_0, p_1)$ which is brought to a reference state: $P^{(\mathrm{i})} \longrightarrow P^{(\mathrm{f})} = \delta_{x0}$.
- Evaluate the change in availability: $\Delta\mathscr{F}^{(\mathrm{f})} - \Delta\mathscr{F}^{(\mathrm{i})} = -TS^{(\mathrm{f})} + TS^{(\mathrm{i})} = k_{\mathrm{B}} T \mathscr{H}[P^{(\mathrm{i})}]$, where $\mathscr{S}[P^{(\mathrm{i})}]$ is the Shannon entropy of the initial distribution.
- We obtain therefore $W \geq k_{\mathrm{B}} T \mathscr{H}[P^{(\mathrm{i})}]$, which corresponds to Landauer's bound.

Historically this analysis followed a rather wandering path. Szilárd [1929] hinted that the demon could be exorcised if the measurement of the particle's position required dissipation, and Brillouin [1949] made the argument more explicit by considering detailed measurement protocols. On the other hand, it was argued that a dissipationless measurement could be performed *provided* the measuring apparatus was in a "standard state" before the measurement. This does not contradict what we said above, because in this case the change of state of $Y$ does not change the entropy of the $X \cup Y$ system. But to produce the measuring apparatus in the reference state, an entropy-reduction manipulation should take place. Landauer [1961] pointed out that in this context it was necessary to spend some work to erase the former state of the measuring apparatus (or of the demon's memory). If the demon is considered as a mechanical device operating from the measurement's input it is not necessary to explicitly consider the demon's memory, but only the state of the measuring apparatus. In the context of the thermodynamics of computation, Bennett [2003] generalized it in the following way:

> Any *logically irreversible* manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding *entropy increase* in non-information bearing degrees of freedom of the information processing apparatus or its environment.

This statement is referred to as **Landauer's principle.** Sagawa and Ueda [2010] pointed out that the entropy increase mentioned by Bennett can take place as well in the measurement as in the erasure state. However, if we only consider the coupling between the measuring apparatus and the system, and treat the demon as a purely mechanical model, the erasure step is unnecessary. The only point is that after the cycle is finished, and the partition is reset, the correlation between the particle and the measuring system is broken and has to be reestablished.

## Irreversible reset

An experimental demonstration of Landauer's principle has been carried out by Bérut et al. [2012]. In this experiment, a colloidal particle is trapped in a bistable potential created by rapidly switching a laster trap between two neighboring locations. The height of the potential barrier can be controlled by varying the laser intensity, and a uniform potential drop (tilt) can be produced by slowly moving the underlying fluid. Then it is possible to device a protocol such that, irrespectively on whether the particle is initially on state L or R, it ends up eventually in state 1, thus providing a "reference state" for further manipulations. The potential $V_\lambda(x)$ felt by the particle can be obtained by sampling the histogram $P(x)$ of its positions for each state of the laser, via the relation $p_\lambda(x) \propto \mathrm{e}^{-V_\lambda(x)/k_\mathrm{B}T}$. In this way, the work $\mathscr{W}$ performed on the particle on a trajectory $\boldsymbol{x} = (x(t))$can be evaluated via the expression

$$\mathscr{W}[\boldsymbol{x}] = \int_{t_0}^{t_\mathrm{f}} \mathrm{d}t\, \dot{\lambda}(t)\, \frac{\partial V_{\lambda(t)}(x(t))}{\partial \lambda}. \tag{1.182}$$

The experiment shows that, in the chosen protocol, the Landauer bound can be approached when the manipulation is carried out sufficiently slowly.

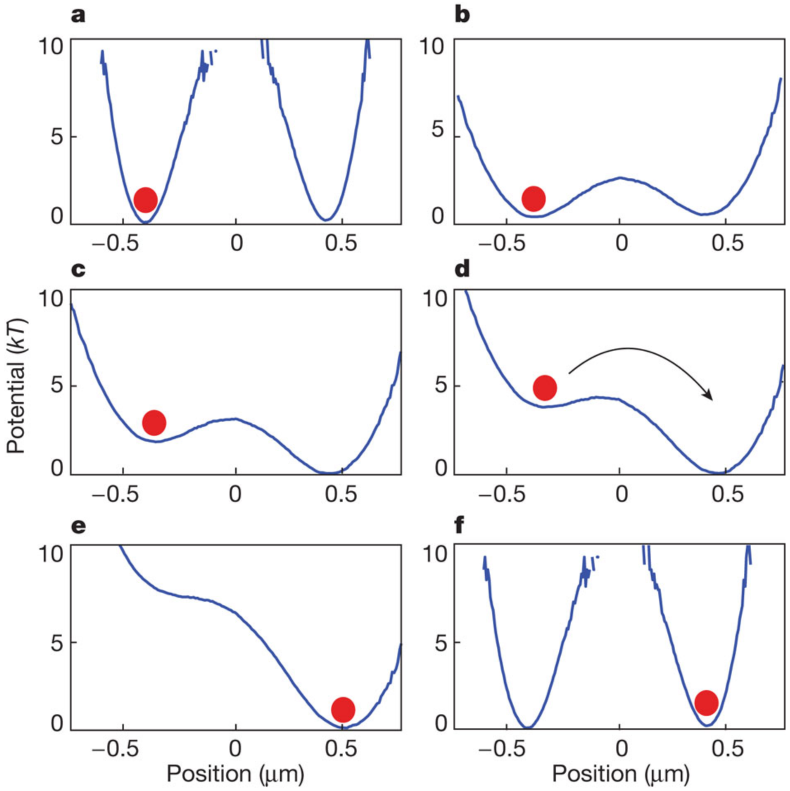


Figure 1.4: The erasure protocol used in the experiment by Bérut et al. [2012]. One bit of information stored in a bistable potential is erased by first lowering the central barrier and then applying a tilting force. In the figures, we represent the transition from the initial state, 0 (left-hand well), to the final state, 1 (right-hand well). We do not show the obvious 1 ⟶ 1 transition. Indeed the procedure is such that irrespective of the initial state, the final state of the particle is always 1. The potential curves shown are those measured in the experiment. From the original publication.

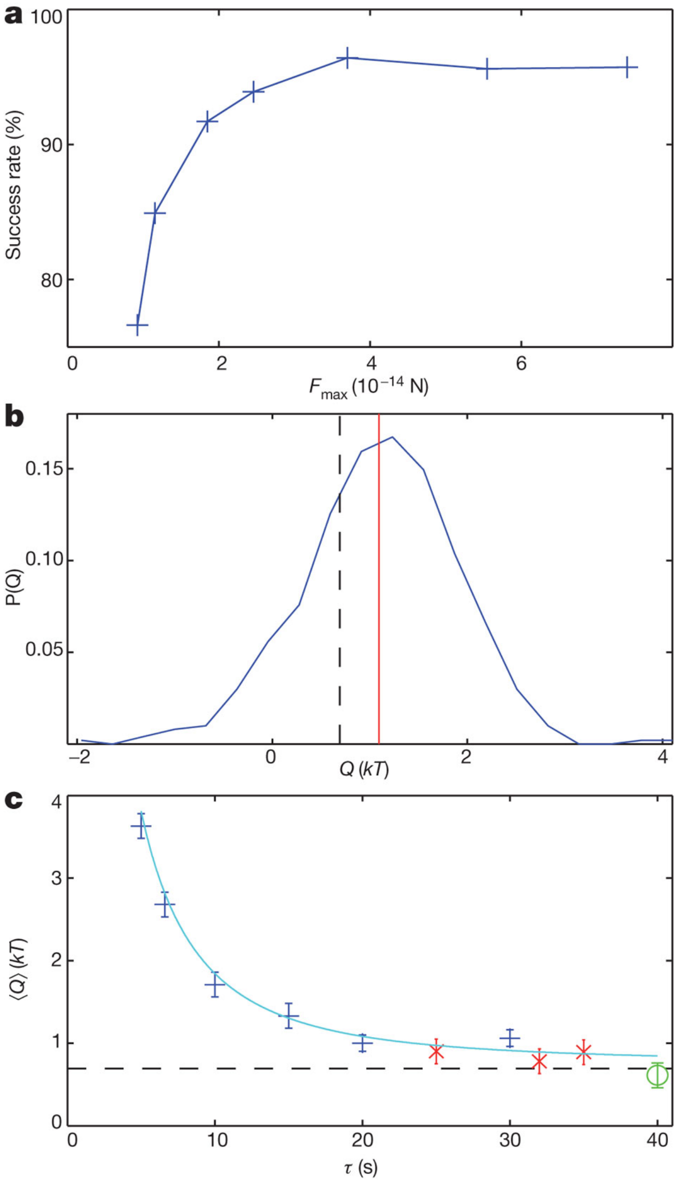


Figure 1.5: a, Success rate of the erasure cycle as a function of the maximum tilt amplitude, $F_{\max}$, for constant $F_{\max}\tau$. b, Approach to the Landauer bound. Heat distribution $P(Q)$ for transition $0 \longrightarrow 1$ with $t = 25$s and $F_{\max} = 1.89 \cdot 10^{-14}$N. The solid vertical line indicates the mean dissipated heat, $\langle Q \rangle$, and the dashed vertical line marks the Landauer bound, $\langle Q \rangle_{\text{Landauer}}$. c, Mean dissipated heat for an erasure cycle as a function of protocol duration, $\tau$, measured for three different success rates, $r$: plus signs, $r \geq 0.90$; crosses, $r \geq 0.85$; circles, $r \geq 0.75$. The horizontal dashed line is the Landauer limit. The continuous line is the fit with the function $[A\exp(-t/\tau_K)+1]B/\tau$, where $\tau_K$ is the Kramers time for the low barrier. Error bars, 1 s.d. From Bérut et al. [2012].

CHAPTER 2

# Applications of Stochastic Thermodynamics to living systems

## 2.1 Introduction

I cannot find a better justification for discussing the relevance of stochastic thermodynamics in living processes that the observations made in Phillips and Quake [2006] on the energy scale relevant for the tiny objects that make up the machinery of life. The relevant thermal energy scale is $k_\mathrm{B}T$ which, at 300 K is about $4.1\cdot 10^{-21}$ J. The typical length scale is the nanometer, and therefore we expect the typical force to be of the order of a few picoNewton. Interestingly, the energies of a number of processes become comparable with $k_\mathrm{B}T$ (and with each other) at the nanometer scale, as detailed in Figure 2.1. On the other hand, the energies of chemical bonds (of the order of the eV$\simeq 1.6\cdot 10^{-19}$ J) are somehow larger, allowing for the stability of chemical complexes. These observation imply that it is not possible to understand the workings of the molecular machinery of life without taking into account fluctuations.

### A rotary motor: ATP synthase

A number of essential cell mechanisms are dedicated to **free-energy transduction** (cf. Hill [1989]), i.e., the transformation of available free energy from one form to another, which is possibly more suitable for the cell's workings. For example, in most eukaryotes the free energy received either from light or from the decomposition of organic or inorganic compounds is transduced into the difference of an electro-chemical potential across a membrane. This difference can then be put to work for several processes. One very important example of this mechanisms is provided by ATP synthase

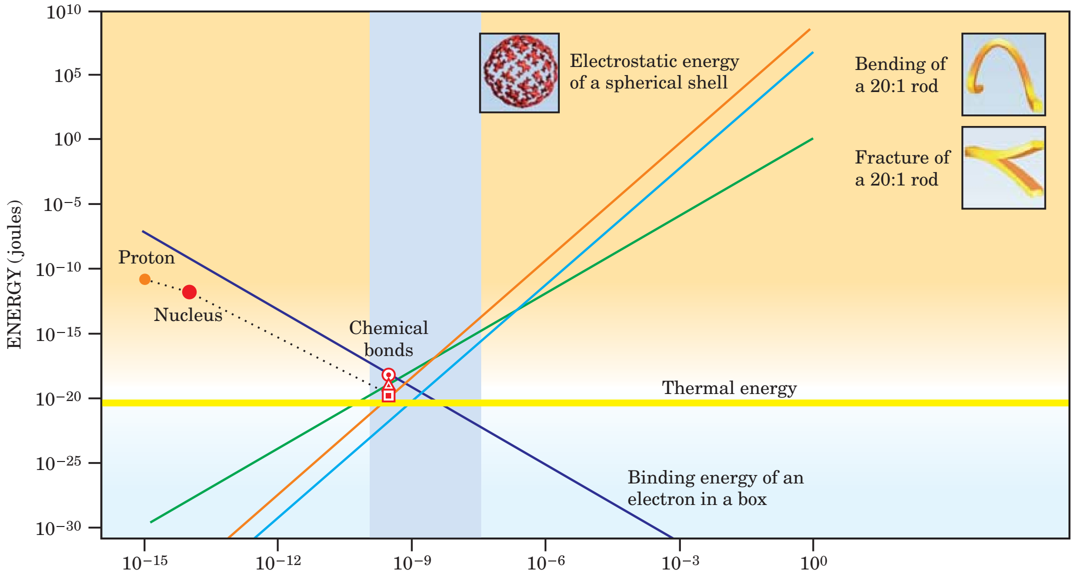


Figure 2.1: The confluence of energy scales is illustrated in this graph, which shows how thermal, chemical, mechanical, and electrostatic energies associated with an object scale with size. As the characteristic object size approaches that at which molecular machines operate (shaded), all the energies converge. The horizontal line shows the thermal energy scale $k_B T$ which, of course, does not depend on an object's size. We estimate binding energy (purple) by considering an electron in a box; for comparison, the graph shows measured binding energies for hydrogen bonds (square), phosphate groups in ATP (triangle), and covalent bonds (circle), along with characteristic energies for nuclear and subatomic particles. In estimating the bending energy (blue), we took an elastic rod with an aspect ratio of 20:1 bent into a semicircular arc, and to compute the fracture energy (green) we estimated the energy in chemical bonds in a longitudinal cross section of the rod. The electrostatic energy (orange) was obtained for a spherical protein with singly charged amino acids of specified size distributed on the surface. From Phillips and Quake [2006].

(cf. Figure 2.2), an enzyme which exploits proton imbalance across the mitochondrial membrane to produce ATP. Remarkably, this protein is composed to two main units, called $F_0$ and $F_1$, which together form a kind of rotary motor. More specifically, the insoluble $F_0$ unit remains attached to the membrane, while the $F_1$ unit is hydrophilic and responsible for hydrolyzing ATP. The workings of this protein have been thoroughly investigated via single-molecule experiments (Yasuda et al. [1998]). In particular, the experimenters were able to show that the enzyme can reverse its function, consuming ATP while pushing protons across the membrane. Following the movement of a single $F_1$ unit, they exhibited its functioning by 120° rotation steps, which also occasionally contains some backward steps (cf. Figure 2.3).

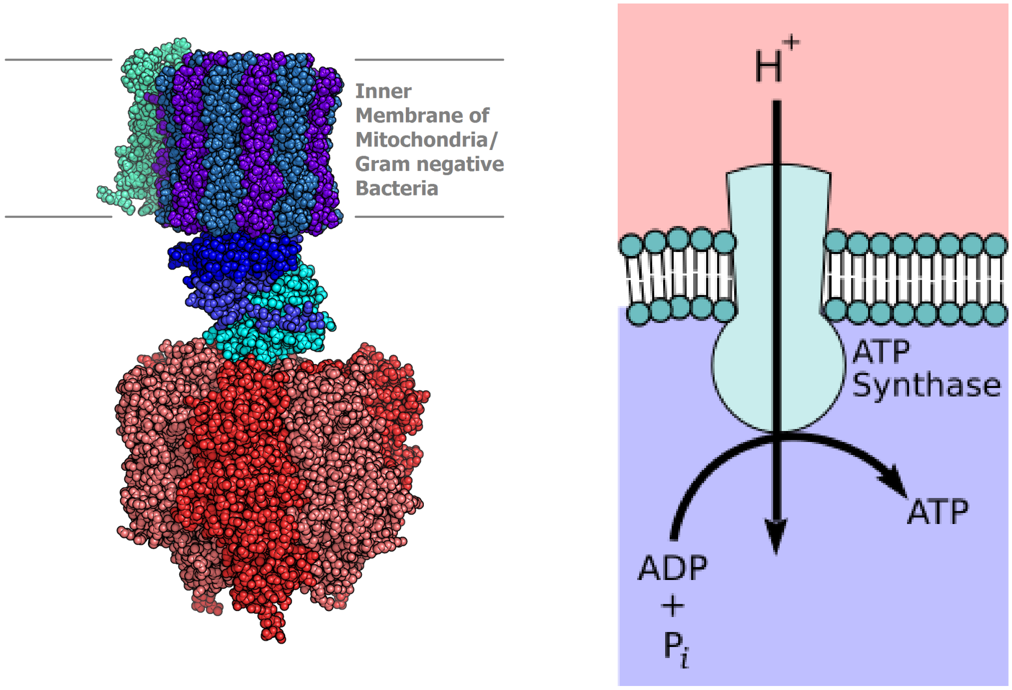


Figure 2.2: ATP synthase is an enzyme that creates the energy storage molecule adenosine triphosphate (ATP). Left: Molecular model of ATP synthase determined by X-ray crystallography. The blue part is the $F_0$ unit, while the red part is the $F_1$ one. Right: Depiction of the operation of ATP synthase using the chemiosmotic proton gradient to power ATP synthesis through oxidative phosphorylation. From *Wikipedia*.

### Nonequilibrium processes in the cell

It is not too far-fetched to describe the inner workings of the cell as an intricate network of far-from-equilibrium chemical reactions. These chemical reactions take part in the free energy transduction and also take advantage of free-energy imbalances in handling different kinds of information, e.g., by using transport to enhance spatial localization of some reactions, or by using template-directed polymerization for the transcription of DNA into RNA, or of mRNA into polypeptides, or for the synthesis of compounds useful for the cell, like ATP. All these reactions involve free-energy imbalances of a few $k_\mathrm{B}T$'s and objects of the size of a few nanometers at most. Thus fluctuations are strong, and may be stronger for reaction which involve small number of molecules. It is interesting therefore to investigate what stochastic thermodynamics has to tell us on chemical reactions.

## 2.2 Enzymes and molecular motors

In this section we consider a few chemical reactions, of increasing complexity, from the point of view of stochastic thermodynamics. The theory was developed in large

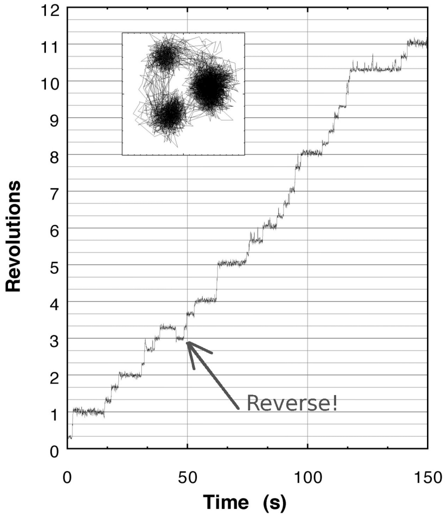


Figure 2.3: Stepwise rotation of the $F_1$ unit of the ATP synthase at low ATP concentration. The inset shows the angular position of the unit as inferred from the position of a long actin filament attached to the unit. Note the backward step in the early part of the track. From Yasuda et al. [1998].

part by Gaspard [2004].

## Isomerization

Let us consider an isomerization reaction, in which a chemical A is transformed in a chemical B:

$$\mathrm{A} \underset{k^-}{\overset{k^+}{\rightleftharpoons}} \mathrm{B}. \tag{2.1}$$

We denote by [X] the concentration of chemical $\mathrm{X} \in \{\mathrm{A}, \mathrm{B}\}$. We then have, e.g.,

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{A}] = -k^+[\mathrm{A}] + k^-[\mathrm{B}], \tag{2.2}$$

and therefore, at equilibrium,

$$\frac{[\mathrm{B}]}{[\mathrm{A}]} = \frac{k^+}{k^-}. \tag{2.3}$$

On the other hand, if the difference in (free) energy between a B molecule and an A molecule[1] is equal to $\epsilon$, we must have, at temperature $T$,

$$\frac{[\mathrm{B}]}{[\mathrm{A}]} = \mathrm{e}^{-\epsilon/k_\mathrm{B}T}. \tag{2.4}$$

---

[1] I speak of “free energy”, since the molecules could have some internal degrees of freedom, and therefore $\epsilon$ could acquire a temperature dependence arising from a trace over them. I explicitly rule out, however, dependence on concentration arising from entropy of mixing, etc.

Therefore the reaction scheme (2.1) is compatible with thermodynamic equilibrium only if

$$\frac{k^+}{k^-} = \mathrm{e}^{-\epsilon/k_\mathrm{B}T}. \tag{2.5}$$

This corresponds to the de Donder relation which relates the forward and backward reaction steps of a chemical reaction to the affinity $A$, i.e., to the change in free energy associated with the extent of reaction. This relation was fist derived by Wegscheider [1901], and de Donder [1936] exploited it to introduce the concept of the chemical affinity of a reaction.

Given the microscopic reaction rates $k^+$ and $k^-$, we can write down the master equation for the isomerization reaction of a system containing a total of $N$ A+B molecules. If we denote by $n$ the number of A molecules, the rates of the transitions $n \longrightarrow n \pm 1$ are given by

$$\begin{aligned} n &\longrightarrow n-1: \qquad k^- n; \\ n &\longrightarrow n+1: \qquad k^+(N-n). \end{aligned} \tag{2.6}$$

Denoting by $p_n(t)$ the probability that there are $n$ A molecules at time $t$, we have

$$\frac{\mathrm{d}p_n}{\mathrm{d}t} = \sum_{n'\,(\neq n)} \left[W_{nn'}p_{n'}(t) - W_{n'n}p_n(t)\right], \tag{2.7}$$

where the transition rates $W_{n'n}$ are given by:

$$W_{n'n} = k^+ n\,\delta_{n',n-1} + k^-\,(N-n)\,\delta_{n',n+1}. \tag{2.8}$$

We have therefore, e.g., for $n' = n+1$,

$$\frac{W_{n+1,n}}{W_{n,n+1}} = \frac{k^+ n}{k^-(N-n-1)} \simeq \mathrm{e}^{-\Delta G/k_\mathrm{B}T}, \tag{2.9}$$

where $\Delta G = G(n+1) - G(n)$, in which

$$G(n) = n\log n + (N-n)\left[\epsilon + \log(N-n)\right], \tag{2.10}$$

is the Gibbs free energy of a state with $n$ A particles, taking into account the entropy of mixing. This is the connection which gives the "macroscopic" rates in term of the "microscopic" ones appearing in (2.5).

## "Passive" catalysis

Let us now consider the case in which the isomerization is catalyzed by an enzyme E. We assume that the enzyme forms with either A or B a complex E$^*$, which can then decompose in either A+E or B+E. This can be represented by the scheme shown in figure 2.2. We obtain therefore the reaction scheme

$$\mathrm{A}+\mathrm{E} \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} \mathrm{E}^* \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} \mathrm{B}+\mathrm{E}, \tag{2.11}$$

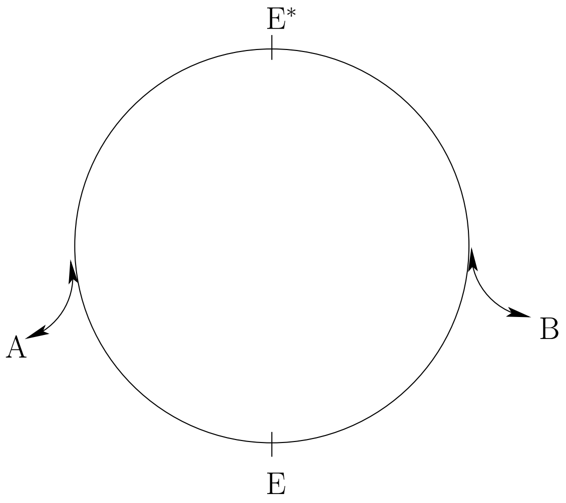


Figure 2.4: Reaction scheme of “passive” catalysis. The unbound enzyme E can bind either an A or a B molecule, going to the bound state $E^*$. It can then release either of them. All reactions are considered reversible.

where $\mathrm{E}^*$ represents the bound enzyme. The condition of thermodynamic equilibrium require

$$\frac{k_1^+ k_2^+}{k_1^- k_2^-} = \mathrm{e}^{-\epsilon/k_\mathrm{B}T}. \tag{2.12}$$

We can parametrize the reaction constants in this way:

1. Two overall attempt frequencies $\omega_1$, $\omega_2$;
2. An energy barrier $\epsilon^*$ for the formation of the complex $\mathrm{E}^*$.

Then we have $k_1^+ = \omega_1\,\mathrm{e}^{-\epsilon^*/k_\mathrm{B}T}$, $k_1^- = \omega_1$, etc., and

$$\frac{k_1^+}{k_1^-} = \mathrm{e}^{-\epsilon^*/k_\mathrm{B}T}; \tag{2.13}$$

$$\frac{k_2^+}{k_2^-} = \mathrm{e}^{-(\epsilon-\epsilon^*)k_\mathrm{B}T}. \tag{2.14}$$

The relation (2.5) then straightforwardly follows.

We can now discuss the evolution of the catalysis by introducing an integer $n$ measuring the number of B molecules synthesized since an arbitrary starting point. This number can be negative. We also assign an index to represent the state of the enzyme: 0 if it is free, and 1 if it is bound. We then have a master equation for the catalysis:

$$\frac{\mathrm{d}p_{0n}}{\mathrm{d}t} = k_1^- p_{1n} + k_2^+ p_{1,n-1} - \left(k_1^+[\mathrm{A}] + k_2^-[\mathrm{B}]\right) p_{0n}; \tag{2.15}$$

$$\frac{\mathrm{d}p_{1n}}{\mathrm{d}t} = k_1^+[\mathrm{A}]\, p_{0n} + k_2^-[\mathrm{B}]\, p_{0,n+1} - \left(k_1^- + k_2^+\right) p_{1n}. \tag{2.16}$$

The dependence on the concentration is related of course to the rate at which an enzyme molecule meets either an A or a B molecule. It is useful to simplify the equation by introducing the effective rate constants $\tilde{k}$, defined by

$$\tilde{k}_1^- = k_1^-; \qquad \tilde{k}_1^+ = k_1^+[\mathrm{A}]; \tag{2.17}$$

$$\tilde{k}_2^- = k_2^-[\mathrm{B}]; \qquad \tilde{k}_2^+ = k_2^+. \tag{2.18}$$

We can now write down an evolution equation for the probability that the enzyme molecule is free or bound. Let us denote by $P = \sum_n p_{1n}$ the probability that it is bound. We then have

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \left(\tilde{k}_1^+ + \tilde{k}_2^-\right)(1-P) - \left(\tilde{k}_1^- + \tilde{k}_2^+\right)P. \tag{2.19}$$

The steady-state solution at fixed A and B concentration is

$$P^{\mathrm{ss}} = \left(1 + \frac{\tilde{k}_1^- + \tilde{k}_2^+}{\tilde{k}_1^+ + \tilde{k}_2^-}\right)^{-1}. \tag{2.20}$$

At equilibrium, when $[\mathrm{B}] = [\mathrm{A}]\,\mathrm{e}^{-\epsilon/k_\mathrm{B}T}$, with the above parametrization of the rate constants, we have

$$P^{\mathrm{eq}} = \frac{[\mathrm{A}]\,\mathrm{e}^{-\epsilon^*/k_\mathrm{B}T}}{1 + [\mathrm{A}]\,\mathrm{e}^{-\epsilon^*/k_\mathrm{B}T}}, \tag{2.21}$$

which is consistent with the Langmuir isotherm.

### Large deviations

Let us now look at the large deviations of the catalysis rate, in a steady state in which the concentration of A and B is fixed. We define the generating function $\Psi_\alpha(\lambda, t)$ by

$$\Psi_\alpha(\lambda, t) = \sum_{n=-\infty}^{+\infty} \mathrm{e}^{\lambda(n+\alpha/2)} p_{\alpha,n}(t), \tag{2.22}$$

where $\alpha \in \{0, 1\}$ and the sum runs over all integers. In this way we have assigned by convention a half-integer value of $n$ to states where the enzyme is bound. Then, letting $z = \mathrm{e}^{\lambda/2}$, from eqs.(2.15,2.16) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \Psi_0 \\ \Psi_1 \end{pmatrix} = \mathsf{L}(z)\begin{pmatrix} \Psi_0 \\ \Psi_1 \end{pmatrix}, \tag{2.23}$$

where the evolution matrix $\mathsf{L}(z)$ is given by

$$\mathsf{L}(z) = \begin{pmatrix} -\left(\tilde{k}_1^+ + \tilde{k}_2^-\right), & z^{-1}\tilde{k}_1^- + z\tilde{k}_2^+ \\ z\tilde{k}_1^+ + z^{-1}\tilde{k}_2^-, & -(\tilde{k}_1^- + \tilde{k}_2^+) \end{pmatrix}. \tag{2.24}$$

Let us remark that $z$ only appears in the off-diagonal elements, and determines the eigenvalues of $\mathsf{L}(z)$ only via their product

$$\Pi(z) = z^2\tilde{k}_1^+\tilde{k}_2^+ + \tilde{k}_1^-\tilde{k}_1^+ + \tilde{k}_2^-\tilde{k}_2^+ + z^{-2}\tilde{k}_1^-\tilde{k}_2^-. \tag{2.25}$$

Thus, for a value $\bar{z}$ of the variable which exchanges the values of the first and last term, the eigenvalues of $\mathsf{L}(z)$ remain the same. We have therefore

$$\bar{z}^2 \tilde{k}_1^+ \tilde{k}_2^+ = z^{-2} \tilde{k}_1^- \tilde{k}_2^-. \tag{2.26}$$

Taking into account the relation (2.12), the definition of the effective rates (2.17,2.18), and the definition of the free-energy change per step

$$\Delta G = \epsilon - k_{\mathrm{B}} T \ln \frac{[\mathrm{A}]}{[\mathrm{B}]}, \tag{2.27}$$

this symmetry corresponds to

$$\lambda \longrightarrow \bar{\lambda} = \Delta G / k_{\mathrm{B}} T - \lambda. \tag{2.28}$$

In particular, for the equilibrium case, where $\Delta G = 0$, we obtain the symmetry with respect to a sign change for $\lambda$.

Let us denote by $\theta(\lambda)$ the leading eigenvalue of $\mathsf{L}$. Then for long times $t$ we have

$$\Psi \sim \mathrm{e}^{t\theta(\lambda)}, \tag{2.29}$$

while we have just seen that

$$\theta(\lambda) = \theta(\Delta G / k_{\mathrm{B}} T - \lambda). \tag{2.30}$$

This corresponds to the Gallavotti-Cohen symmetry. To interpret this symmetry, let us evaluate the probability $p_n(t) = \sum_\alpha p_{\alpha,n}(t)$ by inverting the Laplace transformation defined in equation (2.22). We have, for large values of $t$,

$$p_n(t) \simeq \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}\lambda}{2\pi \mathrm{i}}\, \mathrm{e}^{-\lambda n + t\theta(\lambda)}. \tag{2.31}$$

We can focus on values of $n \simeq J t$, where $J$ is some average production rate. This corresponds to looking at large deviations (of order $t$) in the number $n$ of produced B particles. We can then evaluate the integral via the saddle-point method, obtaining an expression for the probability $P_t(J) = p_{Jt}(t)$:

$$P_t(J) \simeq \mathrm{e}^{-t\psi(J)}, \tag{2.32}$$

where

$$\psi(J) = \min_\lambda [\lambda J - \theta(\lambda)]. \tag{2.33}$$

Let us assume that this minimum is reached for $\lambda^*$ satisfying

$$\theta'(\lambda^*) = J. \tag{2.34}$$

Then if $\lambda^*$ corresponds to a given value of $J$, $\Delta G / k_{\mathrm{B}} T - \lambda^*$ corresponds to the opposite value of $J$. We have therefore

$$P_t(-J) = \mathrm{e}^{-J t \Delta G / k_{\mathrm{B}} T} P_t(J), \tag{2.35}$$

where the total entropy produced in an interval of duration $t$ appears in the argument of the exponential.

The explicit expression of $\theta(\lambda)$ is the following:

$$\begin{aligned}\theta(\lambda) = -\frac{1}{2}\Big[&[\mathrm{A}]\,k_1^+ + k_1^- + k_2^+ + [\mathrm{B}]\,k_2^- \\ &+ \sqrt{\left(\tilde{k}_1^+ + \tilde{k}_1^- + \tilde{k}_2^+ + \tilde{k}_2^-\right)^2 + 4\left(1-\mathrm{e}^{-\lambda}\right)\left(\tilde{k}_1^+\tilde{k}_2^+\mathrm{e}^{\lambda} - \tilde{k}_1^-\tilde{k}_2^-\right)}\,\Big].\end{aligned} \tag{2.36}$$

One can check that the expression under square root is always positive. Its asymptotic behavior is as follows:

$$\theta(\lambda) \simeq \begin{cases} \sqrt{\tilde{k}_1^+\tilde{k}_2^+}\,\mathrm{e}^{\lambda/2}, & \text{for } \lambda \to +\infty; \\ \sqrt{\tilde{k}_1^-\tilde{k}_2^-}\,\mathrm{e}^{-\lambda/2}, & \text{for } \lambda \to -\infty. \end{cases} \tag{2.37}$$

We can then exploit this result to evaluate the effective isomerization rate $J$, i.e., the average number of $\mathrm{A} \longrightarrow \mathrm{B}$ transitions per unit time, via the relation (2.34). We obtain

$$J = \frac{\tilde{k}_1^+\tilde{k}_2^+ - \tilde{k}_1^-\tilde{k}_2^-}{\tilde{k}_1^+ + \tilde{k}_1^- + \tilde{k}_2^+ + \tilde{k}_2^-}. \tag{2.38}$$

It is clear that $J$ vanishes at equilibrium, when eq. (2.12) is satisfied. In a similar way one can obtain the “diffusion coefficient” for the motion along the $n$ axis. Since the derivative $\mathrm{d}\theta/\mathrm{d}\lambda$ can take any real value for $|\lambda|$ large enough, the possible values of $J$ are also unbounded. By plugging in this behavior into (2.33) we obtain, for large $|J|$,

$$\psi(J) \simeq \frac{1}{2}\,|J|. \tag{2.39}$$

Thus there are exponential tails in the distribution of the production rate.

### Entropy production

We now consider the distribution of the entropy production. In the transition from $(\alpha, n)$ to $(\bar{\alpha}, n')$ ($\bar{\alpha} = 1-\alpha$, $n' \in \{n \pm 1, n\}$), the system is given an entropy contribution (in units $k_{\mathrm{B}}$) given by

$$\Delta S((\bar{\alpha}, n'), (\alpha, n)) = \log \frac{W_{(\bar{\alpha},n'),(\alpha,n)}}{W_{(\alpha,n),(\bar{\alpha},n')}}. \tag{2.40}$$

We have

$$W_{(1,n),(0,n)} = \tilde{k}_1^+; \tag{2.41a}$$
$$W_{(1,n-1),(0,n)} = \tilde{k}_2^-; \tag{2.41b}$$
$$W_{(0,n),(1,n)} = \tilde{k}_1^-; \tag{2.41c}$$
$$W_{(0,n+1),(1,n)} = \tilde{k}_2^+. \tag{2.41d}$$

Let us denote by $(\Phi_\alpha(\lambda, t))$ the generating function of the entropy exchanged, conditioned by the state $\alpha$ of the enzyme. We then have

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\Phi_0\\ \Phi_1\end{pmatrix} = \mathsf{M}(\lambda)\begin{pmatrix}\Phi_0\\ \Phi_1\end{pmatrix}, \tag{2.42}$$

where

$$(\mathsf{M}(\lambda))_{00} = -\tilde{k}_1^+ - \tilde{k}_2^-[\mathrm{B}]; \tag{2.43}$$

$$(\mathsf{M}(\lambda))_{01} = \tilde{k}_1^-\left(\frac{\tilde{k}_1^-}{\tilde{k}_1^+}\right)^\lambda + \tilde{k}_2^+\left(\frac{\tilde{k}_2^+}{\tilde{k}_2^-}\right)^\lambda; \tag{2.44}$$

$$(\mathsf{M}(\lambda))_{10} = \tilde{k}_1^+\left(\frac{\tilde{k}_1^+}{\tilde{k}_1^-}\right)^\lambda + k_2^-[\mathrm{B}]\left(\frac{\tilde{k}_2^-}{\tilde{k}_2^+}\right)^\lambda; \tag{2.45}$$

$$(\mathsf{M}(\lambda))_{11} = -\tilde{k}_1^- - \tilde{k}_2^+. \tag{2.46}$$

The evident symmetry

$$\mathsf{M}(1-\lambda) = \mathsf{M}^\dagger(\lambda) \tag{2.47}$$

again yields the Gallavotti-Cohen relation.

One can check that the leading eigenvalue of $\mathsf{M}(\lambda)$ vanishes at equilibrium as a consequence of (2.12). Indeed the eigenvalues depend on $\lambda$ via the product of the off-diagonal elements, which is given by

$$\Pi_1(\lambda) = \tilde{k}_1^+\tilde{k}_1^- + \tilde{k}_2^-\tilde{k}_2^+ + \tilde{k}_1^+\tilde{k}_2^+\left(\frac{\tilde{k}_1^+\tilde{k}_2^+}{\tilde{k}_1^-\tilde{k}_2^-}\right)^\lambda + \tilde{k}_1^-\tilde{k}_2^-\left(\frac{\tilde{k}_1^-\tilde{k}_2^-}{\tilde{k}_1^+\tilde{k}_2^+}\right)^\lambda, \tag{2.48}$$

and we have, because of (2.12) and of the definitions (2.17,2.18) of the $\tilde{k}$'s in terms of the "bare" rates $k$,

$$\frac{\tilde{k}_1^+\tilde{k}_2^+}{\tilde{k}_1^-\tilde{k}_2^-} = \frac{[\mathrm{A}]\,[\mathrm{B}]^{\mathrm{eq}}}{[\mathrm{A}]^{\mathrm{eq}}[\mathrm{B}]}, \tag{2.49}$$

which is equal to 1 at equilibrium. Thus at equilibrium the eigenvalues of $\mathsf{M}(\lambda)$ are independent of $\lambda$.

## Misfolding

We now consider the case in which the enzyme can make errors, i.e., yield (starting from A) a "misfolded" product C rather than the "correct" one B. The reaction scheme is shown in figure 2.5. The right-hand cycle corresponds to the folding cycle shown in figure 2.2, while the left-hand cycle corresponds to misfolding. The rate equations read:

$$\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = k_1^-\,[\mathrm{E}^*] - k_1^+\,[\mathrm{E}]\,[\mathrm{A}]; \tag{2.50a}$$

$$\frac{\mathrm{d}[\mathrm{B}]}{\mathrm{d}t} = k_2^+\,[\mathrm{E}^*] - k_2^-\,[\mathrm{E}]\,[\mathrm{B}]; \tag{2.50b}$$

$$\frac{\mathrm{d}[\mathrm{C}]}{\mathrm{d}t} = k_3^+\,[\mathrm{E}^*] - k_3^-\,[\mathrm{E}]\,[\mathrm{C}]; \tag{2.50c}$$

$$\frac{\mathrm{d}[\mathrm{E}^*]}{\mathrm{d}t} = \left(k_1^+\,[\mathrm{A}] + k_2^-\,[\mathrm{B}] + k_3^-\,[\mathrm{C}]\right)[\mathrm{E}] - \left(k_1^- + k_2^- + k_3^-\right)[\mathrm{E}^*]; \tag{2.50d}$$

$$\frac{\mathrm{d}[\mathrm{E}]}{\mathrm{d}t} = \left(k_1^- + k_2^- + k_3^-\right)[\mathrm{E}^*] - \left(k_1^+\,[\mathrm{A}] + k_2^-\,[\mathrm{B}] + k_3^-\,[\mathrm{C}]\right)[\mathrm{E}]. \tag{2.50e}$$

In this case we need to consider a two-dimensional state space $(n,m)$, where $n$ is the number of B molecules and $m$ the number of C molecules produced since an arbitrary time. Let us denote by $p_0(n,m,t)$ the probability of having the free enzyme in the state $(n,m)$, and by $p_1(n,m,t)$ the probability that it is bound, in state $(n,m)$. These probabilities satisfy the master equation

$$\begin{aligned}\frac{\mathrm{d}p_0(n,m,t)}{\mathrm{d}t} &= k_1^-\,p_1(n,m,t) + k_2^+\,p_1(n-1,m,t) + k_3^+\,p_1(n,m-1,t) + \\ &\quad - \left(k_1^+\,[\mathrm{A}] + k_2^-\,[\mathrm{B}] + k_3^-\,[\mathrm{C}]\right)p_0(n,m,t);\end{aligned} \tag{2.51a}$$

$$\begin{aligned}\frac{\mathrm{d}p_1(n,m,t)}{\mathrm{d}t} &= k_1^+\,[\mathrm{A}]\,p_0(n,m,t) + k_2^-\,[\mathrm{B}]\,p_0(n+1,m,t) + k_3^-\,[\mathrm{C}]\,p_0(n,m+1,t) \\ &\quad - \left(k_1^- + k_2^+ + k_3^+\right)p_1(n,m,t).\end{aligned} \tag{2.51b}$$

The “right” cycle $\mathscr{C}_\mathrm{B}$ involving B corresponds to the enzyme-assisted folding we have seen before, and therefore the relations (2.12) still apply. The “left” cycle $\mathscr{C}_\mathrm{C}$ involving C introduces an extra reaction $\mathrm{E}^* \longrightarrow \mathrm{E} + \mathrm{C}$, with two new rates, $k_3^\pm$, which satisfy

$$\frac{k_3^+}{k_3^-} = \mathrm{e}^{-(\epsilon'-\epsilon^*)}, \tag{2.52}$$

where $\epsilon'$ is the (free) energy per molecule of C, measured in units of $k_\mathrm{B}T$. We can thus parametrize the transition rate as before, introducing the (free) energy $\epsilon'$ for the state C of the molecule, and an attempt frequency $\omega_3$:

$$k_3^+ = \omega_3\,\mathrm{e}^{-(\epsilon'-\epsilon^*)}; \qquad k_3^- = \omega_3. \tag{2.53}$$

We make no assumption on $\epsilon'$: C can be more or less stable than the “correct” fold B. In this case the discrimination between “right” and “wrong” product is only due to the attempt frequency $\omega_3$, which we can assume to be sufficiently smaller than $\omega_2$.

We can now introduce the effective reaction rates $\tilde{k}$, by complementing (2.17,2.18) by

$$\tilde{k}_3^- = k_3^-\,[\mathrm{C}]; \qquad \tilde{k}_3^+ = k_3^+. \tag{2.54}$$

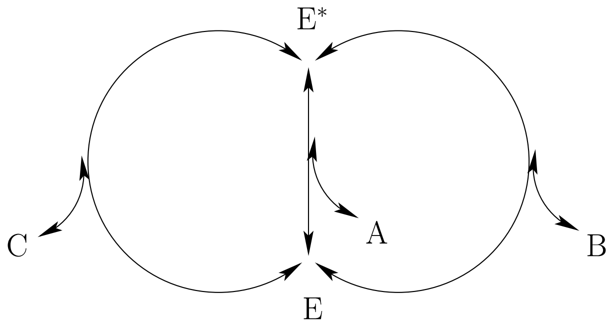


Figure 2.5: The cycles involved in the misfolding reaction. The free enzyme E can bind an A, B or C molecule, and go to the bound state E*. It then produces an A, B or C molecule, with different rates.

The product of the transition rates taken along $\mathscr{C}_{\mathrm{B}}$ clockwise, divided by the corresponding product taken counterclockwise, satisfies (2.12). Substituting the expressions for the effective transition rates, we obtain

$$\frac{\tilde{k}_1^+ \tilde{k}_2^+}{\tilde{k}_1^- \tilde{k}_2^-} = \mathrm{e}^{-\epsilon} \frac{[\mathrm{A}]}{[\mathrm{B}]}. \tag{2.55}$$

In the same way, for the cycle $\mathscr{C}_{\mathrm{C}}$ we obtain

$$\frac{\tilde{k}_1^+ \tilde{k}_3^+}{\tilde{k}_1^- \tilde{k}_3^-} = \mathrm{e}^{-\epsilon'} \frac{[\mathrm{A}]}{[\mathrm{C}]}. \tag{2.56}$$

**Large deviations**

Let us rewrite the master equation satisfied by $p_\alpha(n,m,t)$ by means of the effective reaction rates:

$$\begin{aligned}\frac{\mathrm{d}p_0(n,m,t)}{\mathrm{d}t} &= \tilde{k}_1^- \, p_1(n,m,t) + \tilde{k}_2^+ \, p_1(n-1,m,t) + \tilde{k}_3^+ \, p_1(n,m-1,t) + \\ &\quad - \left(\tilde{k}_1^+ + \tilde{k}_2^- + \tilde{k}_3^-\right) p_0(n,m,t); \qquad (2.57\text{a}) \\ \frac{\mathrm{d}p_1(n,m,t)}{\mathrm{d}t} &= \tilde{k}_1^+ \, p_0(n,m,t) + \tilde{k}_2^- \, p_0(n+1,m,t) + \tilde{k}_3^- \, p_0(n,m+1,t) \\ &\quad - \left(\tilde{k}_1^- + \tilde{k}_2^+ + \tilde{k}_3^+\right) p_1(n,m,t). \qquad (2.57\text{b})\end{aligned}$$

We can then define the generating function by introducing

$$\lambda = (\lambda_1, \lambda_2), \tag{2.58}$$

and setting, for $\alpha \in \{0,1\}$,

$$\Psi_\alpha(\lambda,t) = \sum_{n,m} \mathrm{e}^{\lambda\cdot(n+\alpha\frac{1}{2},\, m+\alpha\frac{1}{2})} \, p_\alpha(n,m,t), \tag{2.59}$$

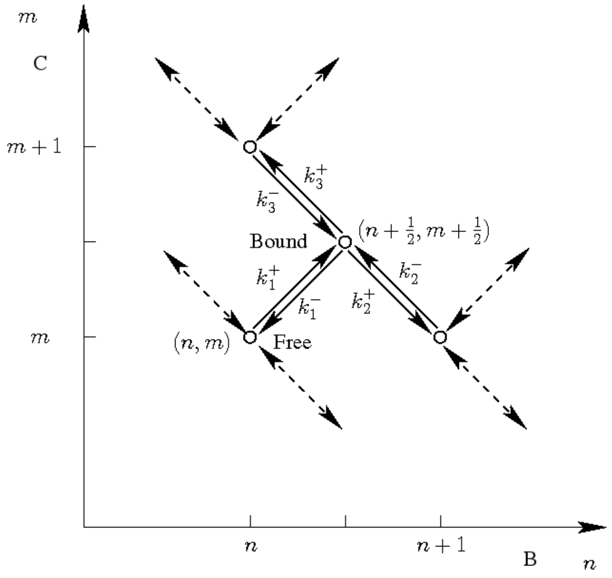


Figure 2.6: Reaction scheme for isomerization with misfolding. The free enzyme E can bind an A, B or C molecule, and go to the bound state E$^*$. It can then release an A, B or C molecule, with different rates. The states $(n,m)$, with $n$ and $m$ both integers, correspond to the free enzyme. If a molecule of B is produced, $n$ increases by 1. If a molecule of C is produced, $m$ increases by 1.

where the sum runs over all integer values of $(n,m)$. The generating function satisfies the evolution equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\Psi_0\\ \Psi_1\end{pmatrix} = \mathscr{L}_2(\lambda)\begin{pmatrix}\Psi_0\\ \Psi_1\end{pmatrix} \tag{2.60}$$

where, having defined $z_i = \mathrm{e}^{\lambda_i/2}$, $i \in \{1,2\}$, the matrix $\mathscr{L}_2(\lambda)$ is defined by

$$(\mathscr{L}_2(\lambda))_{00} = -\left(\tilde{k}_1^+ + \tilde{k}_2^- + \tilde{k}_3^-\right); \tag{2.61a}$$

$$(\mathscr{L}_2(\lambda))_{01} = \tilde{k}_1^- / (z_1 z_2) + \tilde{k}_2^+ (z_1/z_2) + \tilde{k}_3^+ (z_2/z_1); \tag{2.61b}$$

$$(\mathscr{L}_2(\lambda))_{10} = \tilde{k}_1^+ z_1 z_2 + \tilde{k}_2^- (z_2/z_1) + \tilde{k}_3^- (z_1/z_2); \tag{2.61c}$$

$$(\mathscr{L}_2(\lambda))_{11} = -\left(\tilde{k}_1^- + \tilde{k}_2^+ + \tilde{k}_3^+\right). \tag{2.61d}$$

The leading eigenvalue of $\mathscr{L}_2(\lambda)$ is given by

$$\begin{aligned}\theta(z_1,z_2) = \tfrac{1}{2}\Big\{ &-\left(\tilde{k}_1^+ + \tilde{k}_1^- + \tilde{k}_2^+ + \tilde{k}_2^- + \tilde{k}_3^+ + \tilde{k}_3^-\right)\\ &+ \Big[\left(\tilde{k}_1^+ + \tilde{k}_1^- + \tilde{k}_2^+ + \tilde{k}_2^- + \tilde{k}_3^+ + \tilde{k}_3^-\right)^2\\ &+ 4\Big(\tilde{k}_1^+\tilde{k}_2^+(-1+z_1^2) + \tilde{k}_1^-\tilde{k}_2^-(z_1^{-2}-1) + \tilde{k}_3^-\tilde{k}_1^-(z_2^{-2}-1)\\ &+ \tilde{k}_1^+\tilde{k}_3^+(-1+z_2^2) + \tilde{k}_2^-\tilde{k}_3^+\left(-1+\frac{z_2^2}{z_1^2}\right) + \tilde{k}_3^-\tilde{k}_2^+\left(\frac{z_1^2}{z_2}-1\right)\Big)\Big]^{1/2}\Big\}.\end{aligned} \tag{2.62}$$

One can check that this expression is invariant with respect to the transformation $(z_1, z_2) \longrightarrow (\bar{z}_1, \bar{z}_2)$, where

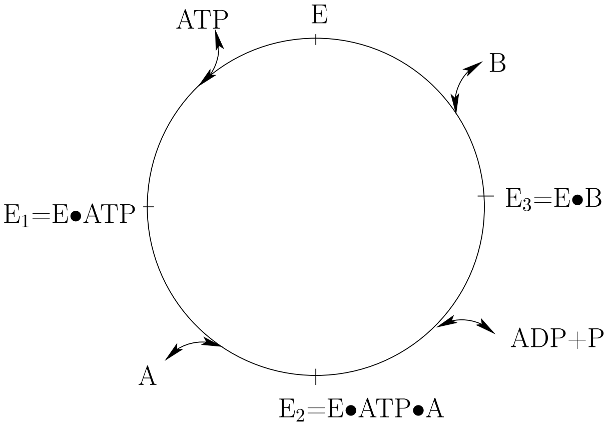
$$\bar{z}_1^2 = \frac{\tilde{k}_1^- \tilde{k}_2^-}{\tilde{k}_1^+ \tilde{k}_2^+ z_1^2} = \frac{1}{z_1^2} \mathrm{e}^{\epsilon} \frac{[\mathrm{B}]}{[\mathrm{A}]}; \tag{2.63a}$$

$$\bar{z}_2^2 = \frac{\tilde{k}_1^- \tilde{k}_3^-}{\tilde{k}_1^+ \tilde{k}_3^+ z_2^2} = \frac{1}{z_2^2} \mathrm{e}^{\epsilon'} \frac{[\mathrm{C}]}{[\mathrm{A}]}. \tag{2.63b}$$

This corresponds to *two* symmetries similar to (2.28), which must be applied at the same time.

## Active catalysis

Let us now consider the situation in which the folding A ⟶ B is accompanied by the hydrolysis of an ATP molecule. We can describe the reaction by the following scheme, where all reactions are considered reversible:



We then have the following rate equations:

$$\frac{\mathrm{d}[\mathrm{E}]}{\mathrm{d}t} = k_4^+ [\mathrm{E}_3] + k_1^- [\mathrm{E}_1] - k_1^+ [\mathrm{E}][\mathrm{ATP}] - k_4^- [\mathrm{E}][\mathrm{B}]; \tag{2.64a}$$

$$\frac{\mathrm{d}[\mathrm{E}_1]}{\mathrm{d}t} = k_1^+ [\mathrm{E}][\mathrm{ATP}] + k_2^- [\mathrm{E}_2] - k_1^- [\mathrm{E}_1] - k_2^+ [\mathrm{E}_1][\mathrm{A}]; \tag{2.64b}$$

$$\frac{\mathrm{d}[\mathrm{E}_2]}{\mathrm{d}t} = k_2^+ [\mathrm{E}_1][\mathrm{A}] + k_3^- [\mathrm{E}_3][\mathrm{ADP}][\mathrm{P}] - (k_2^- + k_3^+)[\mathrm{E}_2]; \tag{2.64c}$$

$$\frac{\mathrm{d}[\mathrm{E}_3]}{\mathrm{d}t} = k_3^+ [\mathrm{E}_2] + k_4^- [\mathrm{E}][\mathrm{B}] - (k_4^+ + k_3^- [\mathrm{ADP}][\mathrm{P}])[\mathrm{E}_3]. \tag{2.64d}$$

Detailed balance intimates, for each step $i$ of the reaction ($i = 1, \ldots, 4$) the relations

$$\frac{k_i^+}{k_i^-} = \mathrm{e}^{-\Delta\epsilon_i / k_\mathrm{B} T}, \tag{2.65}$$

where $\Delta\epsilon_i$ is the free-energy difference between the complex enzyme+reactants at the beginning and at the end of the reaction step. Thus the kinetic rates satisfy the following identity:

$$\frac{k_1^+ k_2^+ k_3^+ k_4^+}{k_1^- k_2^- k_3^- k_4^-} = \mathrm{e}^{-(\Delta\epsilon_{\mathrm{A}}+\Delta\epsilon_{\mathrm{ATP}})/k_{\mathrm{B}}T}, \tag{2.66}$$

where

$$\Delta\epsilon_{\mathrm{A}} = \epsilon_{\mathrm{B}} - \epsilon_{\mathrm{A}}; \tag{2.67}$$

$$\Delta\epsilon_{\mathrm{ATP}} = \epsilon_{\mathrm{ADP}} + \epsilon_{\mathrm{P}} - \epsilon_{\mathrm{ATP}}. \tag{2.68}$$

We can consider a steady state in which the reactant concentrations are kept fixed. Then the system is traversed by a current $J_{\mathrm{A}} = k_2^+[\mathrm{E}_1][\mathrm{A}] - k_2^-[\mathrm{E}_2]$ of A molecules per unit time which fold into B, while consuming $J_{\mathrm{ATP}} = k_1^+[\mathrm{E}][\mathrm{ATP}] - k_1^-[\mathrm{E}_1]$ ATP molecules per unit time.

Thus we have a state variable $\alpha$ which can take four values: $\alpha \in \{0,1,2,3\}$, and two chemical variables, i.e., $i$ for the A molecules, and $j$ for the ATP molecules. As above, we count “a half” for bound particles. The kinetic equations carry on to the following master equation for $P_\alpha(i,j)$:

$$\begin{aligned}
\frac{\partial P_0(i,j,t)}{\partial t} &= k_4^+ P_3(i-\tfrac{1}{2},j,t) + k_1^- P_1(i,j+\tfrac{1}{2},t)\\
&\quad - \left(k_1^+[\mathrm{ATP}] + k_4^-[\mathrm{B}]\right) P_0(i,j,t); && (2.69\mathrm{a})\\
\frac{\partial P_1(i,j+\tfrac{1}{2},t)}{\partial t} &= k_1^+[\mathrm{ATP}] P_0(i,j,t) + k_2^- P_2(i+\tfrac{1}{2},j+\tfrac{1}{2})\\
&\quad - \left(k_1^- + k_2^+[\mathrm{A}]\right) P_1(i,j+\tfrac{1}{2},t); && (2.69\mathrm{b})\\
\frac{\partial P_2(i+\tfrac{1}{2},j+\tfrac{1}{2},t)}{\partial t} &= k_2^+[\mathrm{A}] P_1(i,j+\tfrac{1}{2},t) + k_3^-[\mathrm{ADP}][\mathrm{P}] P_3(i+\tfrac{1}{2},j+1)\\
&\quad - \left(k_2^- + k_3^+\right) P_2(i+\tfrac{1}{2},j+\tfrac{1}{2},t); && (2.69\mathrm{c})\\
\frac{\partial P_3(i+\tfrac{1}{2},j,t)}{\partial t} &= k_3^+ P_2(i+\tfrac{1}{2},j-\tfrac{1}{2},t) + k_4^-[\mathrm{B}] P_0(i+1,j,t)\\
&\quad - \left(k_4^- + k_3^+[\mathrm{ADP}][\mathrm{P}]\right) P_3(i+\tfrac{1}{2},j,t). && (2.69\mathrm{d})
\end{aligned}$$

As before, we can incorporate the concentrations by defining effective rates, e.g.,

$$\begin{aligned}
\tilde{k}_1^+ &= k_1^+[\mathrm{ATP}]; && (2.70\mathrm{a})\\
\tilde{k}_2^+ &= k_2^+[\mathrm{A}]; && (2.70\mathrm{b})\\
\tilde{k}_3^- &= k_3^-[\mathrm{ADP}][\mathrm{P}]; && (2.70\mathrm{c})\\
\tilde{k}_4^- &= k_4^-[\mathrm{B}], && (2.70\mathrm{d})
\end{aligned}$$

Thus we can define the generating function

$$\Psi_\alpha(\boldsymbol{\lambda},t) = \sum_{i_1,i_2=-\infty}^{+\infty} P_\alpha(i_1,i_2,t)\,\mathrm{e}^{\lambda_1 i_1 + \lambda_2 i_2}, \tag{2.71}$$

depending on $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, and where the sum runs over integers and half-integers. It satisfies the evolution equations

$$\frac{\partial \Psi_0}{\partial t} = \tilde{k}_4^+ \mathrm{e}^{\lambda_1/2} \Psi_3 + \tilde{k}_1^+ \mathrm{e}^{-\lambda_2/2} \Psi_1 - \left(\tilde{k}_1^+ + \tilde{k}_4^-\right) \Psi_0; \tag{2.72a}$$

$$\frac{\partial \Psi_1}{\partial t} = \tilde{k}_1^+ \mathrm{e}^{\lambda_2/2} \Psi_0 + \tilde{k}_2^- \mathrm{e}^{-\lambda_1/2} \Psi_2 - \left(\tilde{k}_1^- + \tilde{k}_2^+\right) \Psi_1; \tag{2.72b}$$

$$\frac{\partial \Psi_2}{\partial t} = k_2^+ \mathrm{e}^{\lambda_1/2} \Psi_1 + k_3^- \mathrm{e}^{-\lambda_2/2} \Psi_3 - \left(\tilde{k}_2^- + \tilde{k}_3^+\right) \Psi_2; \tag{2.72c}$$

$$\frac{\partial \Psi_3}{\partial t} = \tilde{k}_3^+ \mathrm{e}^{\lambda_2/2} \Psi_2 + \tilde{k}_4^- \mathrm{e}^{-\lambda_1/2} \Psi_0 - \left(\tilde{k}_4^- + \tilde{k}_3^+\right) \Psi_3. \tag{2.72d}$$

Introducing the vector $\boldsymbol{\Psi} = (\Psi_\alpha)$ and the shorthands

$$z_1 = \mathrm{e}^{\lambda_1/2}; \tag{2.73a}$$

$$z_2 = \mathrm{e}^{\lambda_2/2}; \tag{2.73b}$$

this equation can be written in the form

$$\frac{\partial \boldsymbol{\Psi}}{\partial t} = \mathsf{M}(\boldsymbol{z})\, \boldsymbol{\Psi}, \tag{2.74}$$

where the matrix $\mathsf{M}(\boldsymbol{z})$ is given by

$$\mathsf{M}(\boldsymbol{z}) = \begin{pmatrix} -\left(\tilde{k}_1^+ + \tilde{k}_4^-\right), & \tilde{k}_1^- / z_2, & 0, & \tilde{k}_4^+ z_1 \\ \tilde{k}_1^+ z_2, & -\left(\tilde{k}_1^- + \tilde{k}_2^+\right), & \tilde{k}_2^- / z_1, & 0 \\ 0, & \tilde{k}_2^+ z_1, & -\left(\tilde{k}_2^- + \tilde{k}_3^+\right), & \tilde{k}_3^- / z_2 \\ \tilde{k}_4^- / z_1, & 0, & \tilde{k}_3^+ z_2, & -\left(\tilde{k}_4^- + \tilde{k}_3^+\right) \end{pmatrix}. \tag{2.75}$$

The matrix $\mathsf{M}(\boldsymbol{z})$ exhibits the symmetry

$$\mathsf{M}(\overline{\boldsymbol{z}}) = \mathsf{Q}^{-1} \mathsf{M}^\dagger(\boldsymbol{z}) \mathsf{Q}, \tag{2.76}$$

where

$$\overline{z}_1^2 = \frac{\tilde{k}_2^- \tilde{k}_4^-}{\tilde{k}_2^+ \tilde{k}_4^+} \frac{1}{z_1^2}; \tag{2.77a}$$

$$\overline{z}_2^2 = \frac{\tilde{k}_1^- \tilde{k}_3^-}{\tilde{k}_1^+ \tilde{k}_3^+} \frac{1}{z_2^2}, \tag{2.77b}$$

and $\mathsf{Q}$ is a diagonal matrix, $\mathrm{diag}(q_1, q_2, q_3, q_4)$,with

$$q_1 = k_2^- \sqrt{\frac{k_1^- k_3^+ k_3^-}{k_1^+}}; \tag{2.78a}$$

$$q_2 = k_2^- k_3^-; \tag{2.78b}$$

$$q_3 = k_3^- \sqrt{\frac{k_2^+ k_2^- k_4^-}{k_4^+}}; \tag{2.78c}$$

$$q_4 = \sqrt{\frac{k_1^- k_2^+ k_3^+ k_3^- k_4^-}{k_1^+ k_4^+}}. \tag{2.78d}$$

Therefore the leading eigenvalue $\theta(\boldsymbol{z})$ of $\mathsf{M}(\boldsymbol{z})$ satisfies the Gallavotti-Cohen symmetry

$$\theta(\overline{\boldsymbol{z}}) = \theta(\boldsymbol{z}), \tag{2.79}$$

where the transformation can be expressed in terms of the affinity as before. Note that because there is only one independent cycle in this reaction scheme, one only has a single symmetry transformation.

## 2.3 Fluctuation relations

We have seen that Crooks relation leads to equation (1.86), which allows one, in principle, to evaluate the free-energy landscape as a function of a collective coordinate $M$ by nonequilibrium measurements. This relation has been exploited in particular to obtain the free energy difference between a folded and unfolded RNA hairpin by Collin et al. [2005], and the similar difference for a polypeptide domain by Imparato et al. [2008].

### RNA hairpin pulling experiments

In Collin et al. [2005], the system was a small interfering (si)RNA hairpin that targets the mRNA of one receptor of the HIV virus, and then unfolds irreversibly, with a dissipated work of the order of $6k_\mathrm{B}T$. Thus it was expected that the distribution of work $W$ upon folding and refolding will be sufficiently wide to allow for the application of the fluctuation relation in the form

$$\widehat{P}(-W) = \mathrm{e}^{-(W-\Delta F)/k_\mathrm{B}T}\, P(W), \tag{2.80}$$

where $P(W)$ corresponds to unfolding and $\widehat{P}(W)$ to refolding. The hairpin was connected to two DNA handles which allowed them to attach to ssDNA-coated latex spheres which were then pulled by optical tweezers. The force is obtained by measuring the displacement of the sphere from the focus of the laser (with a know intensity of the beam) and the work can in this case be approximated quite well by the force-time displacement integral

$$W = \sum_i \Delta x_i\, F_i. \tag{2.81}$$

It is consequence of equation (2.80) that the two distributions $P(W)$ and $\widehat{P}(-W)$ cross at $W = \Delta F$:

$$\widehat{P}(-W^*) = P(W^*) \quad \Rightarrow \quad W^* = \Delta F. \tag{2.82}$$

If the work distribution is wide enough for this crossing to take place where the number of points with $W \simeq W^*$ is large for both probability distribution, one can be quite confident in using this relation for the evaluation of $\Delta F$. The result of the unfolding and folding experiments are shown in figure 2.7. In this way one can explore the free-energy landscape for the secondary and tertiary structure of small RNA's.

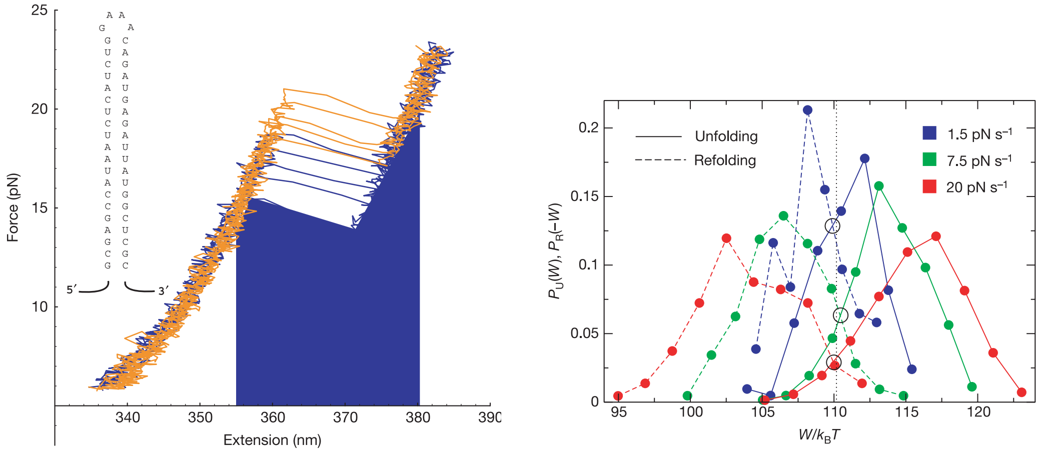


Figure 2.7: Left: Force–extension curves for RNA hairpin unfolding and refolding. The stochasticity of the unfolding and refolding process is characterized by a distribution of unfolding or refolding work trajectories. Five unfolding (orange) and refolding (blue) force– extension curves for the RNA hairpin are shown (loading rate of 7.5 pN/s). The blue area under the curve represents the work returned to the machine as the molecule switches from the unfolded to the folded state. The RNA sequence is shown as an inset. Right: Test of the CFT using an RNA hairpin. Work distributions for RNA unfolding (continuous lines) and refolding (dashed lines). Negative work, $P_R(-W)$, is plotted for refolding. Statistics: 130 pulls and three molecules ($r = 1.5\,\text{pN/s}$), 380 pulls and four molecules ($r = 7.5\,\text{pN/s}$), 700 pulls and three molecules ($r = 20.0\,\text{pN/s}$), for a total of ten separate experiments. Work values were binned into about ten equally spaced intervals. Unfolding and refolding distributions at different speeds show a common crossing around $\Delta G = 110.3\,k_BT$. From Collin et al. [2005].

## Unfolding a large protein

Imparato et al. [2008] have studied the mechanical unfolding of a protein composed of eight repeats of the Ig27 domain of the large protein titin. The unfolding was performed via manipulation by means of an atomic-force microscope (AFM). In single-molecule force spectroscopy experiments exploiting an AFM, the molecule deposited onto the substrate is grabbed by the AFM probe. Mechanical unfolding of the protein is thus induced by moving the probe away from the substrate with a constant velocity $v$ (linear protocol). As the probe is retracted, the force exerted on the molecule increases until the molecule suddenly unfolds. If the molecule is composed of multiple domains, as in the present case, further retraction causes the extension of successive folded domains, until all the domains are unfolded and the protein-tip interaction is broken. The corresponding force-distance curve has a typical sawtooth structure, as shown in figure 2.8 (Bottom left). The energy landscape of the polypeptide can be evaluated by exploiting equation (1.86), which we rewrite in the form

$$\langle \delta(\ell - \ell_t)\,\mathrm{e}^{-W_t/k_BT} \rangle_t\, \mathrm{e}^{U(\ell, z(t))} = \exp\left[-F_0(\ell)/k_BT\right]/Z_0, \tag{2.83}$$

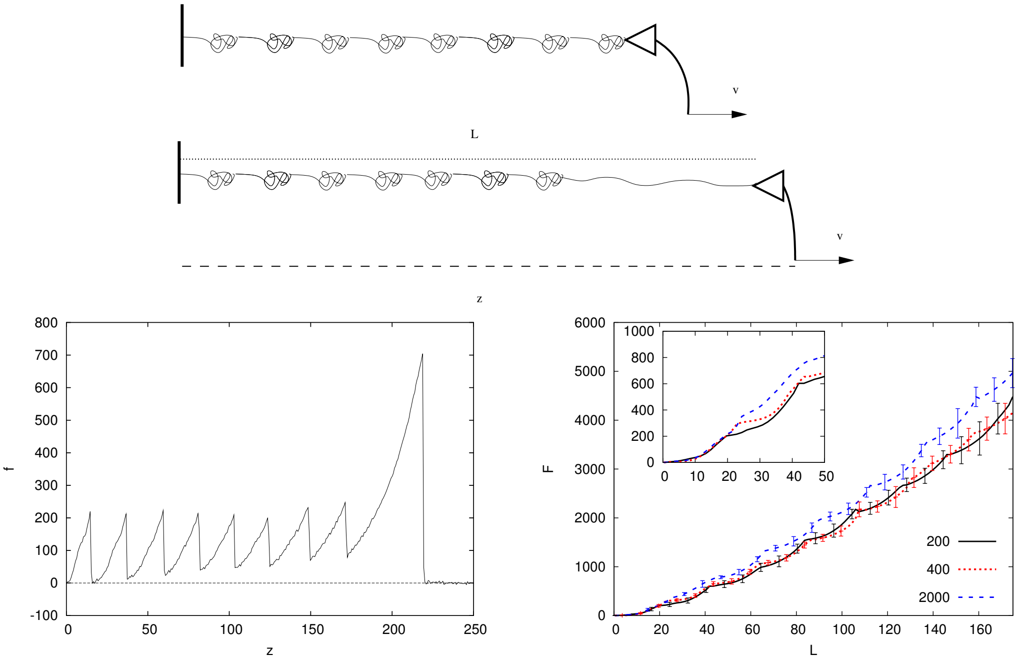


Figure 2.8: Top: Cartoon of the protein made by eight repeats of the Ig27 domain of titin, attached to the tip on an AFM. Bottom: Left: Force extension curve of the polyprotein composed of eight repeats of the Ig27 domain. The typical distance between two consecutive peaks ranges between 22 and 26 nm. Right: Free-energy landscape $F_0$ as a function of the molecular elongation $\ell$, for different values of the probe velocity, $v = 200$, 400, 2000 nm/s: the number of unfolding trajectories considered is 66, 35, 29, respectively. The error bars are obtained by using the jackknife approach for data resampling: for each value of $v$ subsamples of the total number of trajectories are considered, and for each subsample the free-energy landscape is evaluated. One obtains thus a set of curves $F_0(\ell)$, and the error bars are obtained, for each value of $\ell$, as the standard deviation of the mean with respect to this set. The bars are mutually shifted for clarity's sake. Inset: Plot of $F$ in the small-$\ell$ range. From Imparato et al. [2008].

where $z(t)$ is the position of the probe at time $t$, $\ell$ the elongation of the molecule, $W_t$ the work accumulated up to time $t$, $Z_0$ is the partition function of the unperturbed system, $U(\ell, z(t)) = (k/2)(z(t) - \ell)^2$ is the external potential associated with the tip of the AFM, and the average is taken with respect to all realizations of the process up to time $t$. By this method one can evaluate the free energy $F_0(\ell)$ as a function of the elongation $\ell$. From figure 2.8 one can see that the results can be trusted provided the manipulation protocol is not too fast.

CHAPTER 3

# Accuracy-speed-dissipation trade-off

## 3.1 The Problem

Several core processes in the cell are related to information handling. For example, transcription is the process by which single-stranded RNA is polymerized on a DNA template. The transcript is than exploited in different ways, e.g., mRNA acts itself as template for translation, the process in which ribosomes synthesize polypeptides which are then further processed to make proteins. In DNA replication, a ssDNA strand is polymerized using a pre-existing DNA strand as a template. In these processes, the possibility of a mismatch between the template and the newly incorporated unit of the growing polymer is always open. Template-free processes are also relevant in information handling: for example, the actual translation from the codon (made of three consecutive nucleotides) to an amino acid in the growing polypetide chain during translation is encoded by the complex tRNA-amino acid. The complexation of tRNA is performed by special enzymes, called aminoacylases, which must bind on the one hand with a tRNA carrying the anticodon in tis proper place, and on the other hand with the correct amino acid, and then bind the two molecules together. This process is also prone to error, in particular since there is no natural thermodynamical preference (higher affinity) of a given amino acid for its corresponding tRNA.

While these processes all involve some probability for errors, it is observed that the actual observed error rates in physiological conditions are much smaller than expected on the basis of thermodynamics alone in most life forms. For example, in the transcription process the discrimination between correct matching of the nucleotide in the growing RNA strand and the one in the DNA template is expected to be due to the correspondence in hydrogen binding. Since the energy $\epsilon$ involved in hydrogen bonding in water is of the order of $2.5\,k_{\mathrm{B}}T$, one would expect error rates of the order of $\mathrm{e}^{-\epsilon/k_{\mathrm{B}}T} \sim 10^{-2}$. In human DNA replication error rates are of the order of $10^{-10}$ (interestingly, the error rates apparently are inversely correlated with genome length, so that there are usually roughly the same number of errors per replication, at least for eukaryotes). In RNA transcription, error rates are of the order of $10^{-4} \div 10^{-5}$, and in translation the error rates are somewhat larger ($10^{-3} \div 10^{4}$), but always well below the error rates expected on the basis of equilibrium thermodynamics.

A solution in principle to the puzzle of the existence of these small error rates was independently found by Ninio [1975b] (cf. also Ninio [1975a]) and Hopfield [1974]. They conjectured that discrimination would be enhanced if the polymerization reaction allowed for additional steps, each of which could discriminate between correct and incorrect substrate and allowed for the release of the incorrect substrate by an irreversible (ATP-fueled) step. This mechanism was called **kinetic proofreading** by Hopfield, and the term is remained. A slightly different mechanism, also involving an ATP-fueled step, was proposed by Bennett [1979]. More recently, several detailed mechanisms for kinetic proofreading have been proposed (see, e.g., Voliotis et al. [2009], Murugan et al. [2012], Murugan et al. [2014], Peliti and Rao [2015]), while the theory has been developed by Andrieux and Gaspard [2008] (see also Andrieux and Gaspard [2009]) and by Sartori and Pigolotti [2013] (cf. also Sartori and Pigolotti [2015]). In particular, since incrementing the accuracy beyond its equilibrium value involves bringing the system out of equilibrium, and therefore dissipation, it is interesting to evaluate the performance of the different proofreading schemes from both point of view, i.e., error rate and dissipation, to which it is also reasonable to add speed, i.e., the amount of processing steps that can be performed in a unit time. If improving the performance on one of these aspects involves degrading it on one or both of the others, one speaks of tradeoffs. It is interesting therefore to see if and how the thermodynamics of small systems allows us to identify fundamental tradeoffs in these functions.

## Reaction schemes

Here I briefly describe the class of models considered by Sartori and Pigolotti [2015]. We consider a reaction in which an enzyme catalyzes a reaction involving a substrate s which can be right (r) or wrong (w). The reaction is represented by a Markov chain with $n$ states, $i \in \{1,\ldots,n\}$. The scheme can be represented by the diagram in figure 3.1. Conventionally, $j=0$ represent the binding of a new substrate to the enzyme, and $j=n+1$ corresponds to the release of the product. Thus the system is described by the master equation

$$\frac{\mathrm{d}p_i^{\mathrm{s}}}{\mathrm{d}t} = \sum_{j=0}^{n+1} \left[k_{ij}^{\mathrm{s}} p_j^{\mathrm{s}} - k_{ji}^{\mathrm{s}} p_i^{\mathrm{s}}\right] = \sum_j J_{ij}. \tag{3.1}$$

However, the states $j=0$ and $j=n+1$ both correspond to the empty enzyme. Let us suppose that the steady-state probability of this state is 0. The net production rate

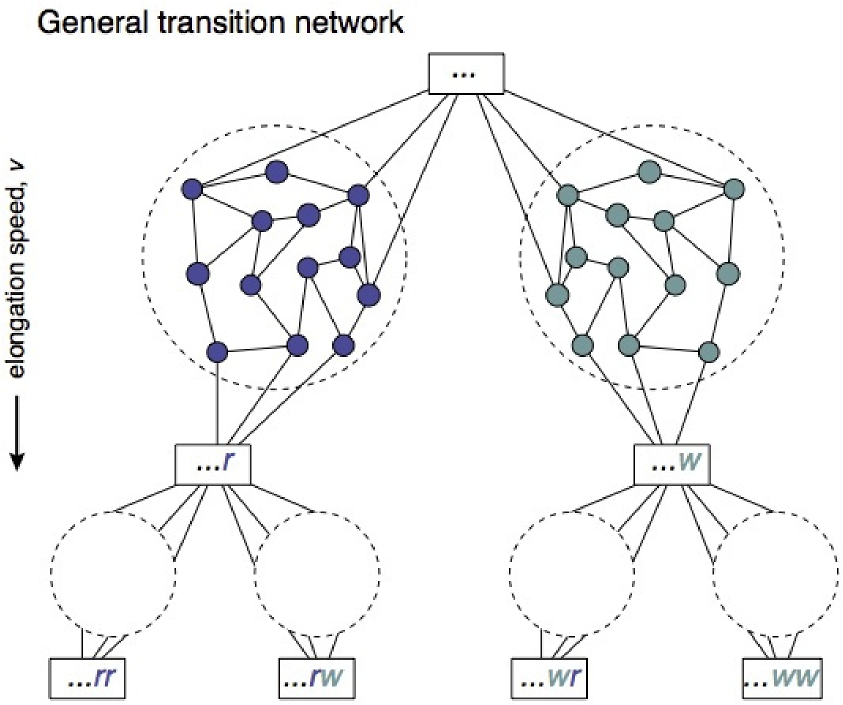


Figure 3.1: State space of the template-assisted polymerization model. Monomer incorporation occurs via a network of intermediate states represented inside the dashed circles. The two colors distinguish networks leading to incorporation of right and wrong monomers. The structure is repeated in a tree-shaped structure as the polymer grows by addition of more and more monomers. From Sartori and Pigolotti [2015].

can be split into the r and the w channel:

$$J = J^{\mathrm{r}} + J^{\mathrm{w}} = \sum_{i=1}^{n} \left[ k^{\mathrm{r}}_{n+1,i} p^{\mathrm{r}}_i - k^{\mathrm{r}}_{i,n+1} p^{\mathrm{r}}_{n+1} \right] + \sum_{i=1}^{n} \left[ k^{\mathrm{w}}_{n+1,i} p^{\mathrm{w}}_i - k^{\mathrm{w}}_{i,n+1} p^{\mathrm{w}}_{n+1} \right]. \tag{3.2}$$

This must be equated with the out-flux from $i = 0$. We can make the ansatz

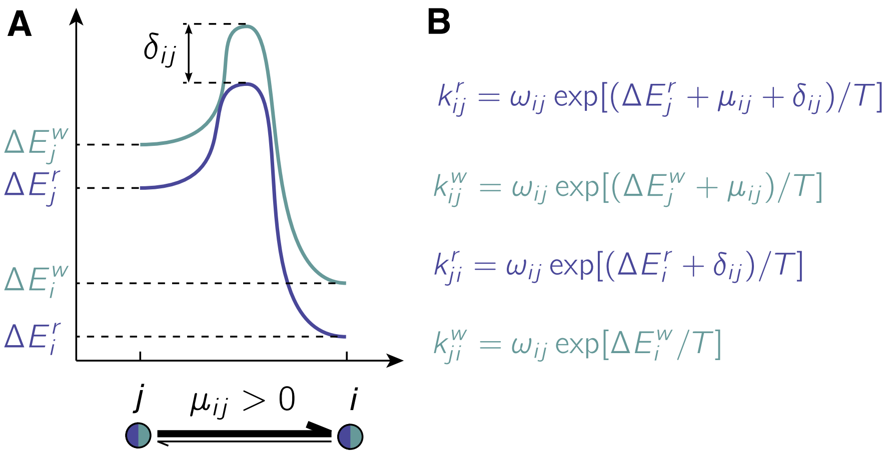
$$J^{\mathrm{w}} = \xi \, J^{\mathrm{out}}. \tag{3.3}$$

This defines the error rate $\xi$, while $J^{\mathrm{out}}$ can be identified with the total yield of the reaction on a single enzyme.

The entropy production is given by the usual expression. In the steady state (denoted by ss) it boils down to

$$\dot{S}_{\mathrm{tot}} = \sum_{<ij>} \sum_{\mathrm{ss}} J^{\mathrm{ss}}_{ij} \log \frac{k^{\mathrm{ss}}_{ij} p^{\mathrm{ss}}_j}{k^{\mathrm{ss}}_{ji} p^{\mathrm{ss}}_i}. \tag{3.4}$$

We can parametrize the rates $k_{ij}^{(\mathrm{r,w})}$, $k_{ji}^{(\mathrm{r,w})}$ in terms of an attempt rate $\omega_{ij}$, which is the same for all the processes for the $(ij)$ link one is considering, a difference $\delta_{ij}$ in the height of the barrier that the wrong substrate has to overcome in going from $i$ to $j$ (or viceversa) with respect to that the right substrate has to overcome ($\delta_{ij} > 0$, so that the wrong substrate has fewer chances to overcome the barrier), and the thermodynamic contributions. These can be split into the energy difference $\Delta E_i^{(\mathrm{r,w})}$ of the enzyme in state $i$ with respect to a reference state, and the contribution $\mu_{ij}$ of the chemical driving, due, e.g., to the hydrolysis or the synthesis of, say, ATP, in going from $j$ to $i$. This is made clearer by the following scheme, which is also due to Sartori and Pigolotti [2015].



We thus obtain the following parametrization of the reaction rates:

$$k_{ij}^{\mathrm{r}} = \omega_{ij} \exp\left[(\Delta E_j^{\mathrm{r}} + \mu_{ij} + \delta_{ij})/k_{\mathrm{B}}T\right]; \tag{3.5}$$

$$k_{ij}^{\mathrm{w}} = \omega_{ij} \exp\left[(\Delta E_j^{\mathrm{w}} + \mu_{ij})/k_{\mathrm{B}}T\right]; \tag{3.6}$$

$$k_{ji}^{\mathrm{r}} = \omega_{ij} \exp\left[(\Delta E_i^{\mathrm{r}} + \delta_{ij})/k_{\mathrm{B}}T\right]; \tag{3.7}$$

$$k_{ji}^{\mathrm{w}} = \omega_{ij} \exp\left[\Delta E_i^{\mathrm{w}}/k_{\mathrm{B}}T\right]. \tag{3.8}$$

We can then evaluate the entropy production in the steady state:

$$\dot{S}_{\mathrm{tot}} = \sum_{<ij>} \sum_{\mathrm{s}} J_{ij}^{(\mathrm{s}),ss} \frac{\mu_{ij}}{T} + \sum_{<ij>} \sum_{\mathrm{s}} J_{ij}^{(\mathrm{s}),\mathrm{ss}} \log\left(\frac{p_j^{(\mathrm{s}),\mathrm{ss}}}{p_i^{(\mathrm{s}),\mathrm{ss}}}\right) + \sum_{<ij>} \sum_{\mathrm{s}} J_{ij}^{(\mathrm{s}),\mathrm{ss}} \frac{\Delta E_j^{(\mathrm{s})} - \Delta E_i^{(\mathrm{s})}}{T}. \tag{3.9}$$

Except for the first term, all other term vanish when summed over the internal links, due to flux conservation in the steady state. Only the links which lead outside of the network yield a contribution, and these can be expressed in terms of the rates of right or wrong incorporation. Thus one obtains

$$\begin{aligned} T\dot{S}_{\mathrm{tot}} &= \sum_{<ij>} \sum_{\mathrm{s}} J_{ij}^{(\mathrm{s}),ss} \mu_{ij} - \xi J \left[\log \xi + \Delta E^{(\mathrm{w})}\right] - \xi J^{\mathrm{ss}} \left[\log(1-\xi)\Delta E^{(\mathrm{r})}\right] \\ &= J^{\mathrm{ss}} \left[\Delta W - \Delta F - k_{\mathrm{B}} T D_{\mathrm{KL}}(\xi \| \xi_{\mathrm{eq}})\right]. \end{aligned} \tag{3.10}$$

To evaluate the working of the system, it is convenient to assume that the final product is released with a rate $F$ (equal for both right and wrong substrate) when the final state is reached. We can then evaluate the entropy production and flow in the steady state according to the following expressions:

$$\dot{S}_{\mathrm{i}} := \frac{1}{2}\sum_{i\neq j}{}'\sum_{\mathrm{s}}\left(k_{ij}^{(\mathrm{s})}p_j^{(\mathrm{s}),\mathrm{ss}} - k_{ji}^{(\mathrm{s})}p_i^{(\mathrm{s}),\mathrm{ss}}\right)\log\frac{k_{ij}^{(\mathrm{s})}p_j^{(\mathrm{s}),\mathrm{ss}}}{k_{ji}p_i^{(\mathrm{s}),\mathrm{ss}}}; \tag{3.11}$$

$$\dot{S}_{\mathrm{e}} := -\frac{1}{2}\sum_{i\neq j}{}'\sum_{\mathrm{s}}\left(k_{ij}^{(\mathrm{s})}p_j^{(\mathrm{s}),\mathrm{ss}} - k_{ji}^{(\mathrm{s})}p_i^{(\mathrm{s}),\mathrm{ss}}\right)\log\frac{k_{ij}^{(\mathrm{s})}}{k_{ji}^{(\mathrm{s})}}. \tag{3.12}$$

Since in the steady state the entropy of the system does not change with time, we have

$$-\frac{\mathrm{d}}{\mathrm{d}t}\sum_k p_k\log p_k = \dot{S}_{\mathrm{i}} + \dot{S}_{\mathrm{e}} \underbrace{-\dot{S}_F}_{\text{catalysis}} = 0, \tag{3.13}$$

where the have emphasized that the entropy flow due to the final catalysis step is negative in order for the reaction to have a positive flux of the products. We can then evaluate the mean step duration (i.e., the mean time needed to release one unit of the product) by

$$\tau := \left(\text{total catalysis rate: } J^{(\mathrm{r})} + J^{(\mathrm{w})}\right)^{-1}. \tag{3.14}$$

The entropy production per step is then given by multiplying the total entropy production rate by $\tau$:

$$\Delta_{\mathrm{i}}S := \tau\,\dot{S}_{\mathrm{i}} \geq 0. \tag{3.15}$$

One can similarly evaluate the entropy flow per unit product:

$$\Delta_{\mathrm{e}}S := \tau\,\dot{S}_{\mathrm{e}} \leq 0. \tag{3.16}$$

The free-energy dissipation in the final catalysis step is given by

$$\Delta S_F := \tau\,\dot{S}_F = \tau F\sum_{\mathrm{s}}\bar{p}_{\text{final state(s)}}\,\Delta\mu_{\text{final step}}. \tag{3.17}$$

Because of the relation (3.13), we can define the **efficiency** $\eta$ by

$$\eta := \Delta S_F/\Delta_{\mathrm{e}}S = 1 + \Delta_{\mathrm{i}}S/\Delta_{\mathrm{e}}S, \tag{3.18}$$

which satisfies, since $\Delta_{\mathrm{e}}S \leq 0$ and $\Delta S_F \geq 0$,

$$0 \leq \eta \leq 1. \tag{3.19}$$

Let us take into account (3.10), and the following relations:

$$D_{\mathrm{KL}}(\xi\|\xi_{\mathrm{eq}}) = \xi\log\frac{\xi}{\xi_{\mathrm{eq}}} + (1-\xi)\log\frac{1-\xi}{1-\xi_{\mathrm{eq}}} \approx \xi_{\mathrm{eq}} \ll 1, \tag{3.20}$$

$$T\,\Delta S^{(\mathrm{w})} = \Delta W^{(\mathrm{w})} - \Delta F - T\log\frac{\xi}{\xi_{\mathrm{eq}}} \geq 0. \tag{3.21}$$

We can thus obtain the following expression of the error rate:

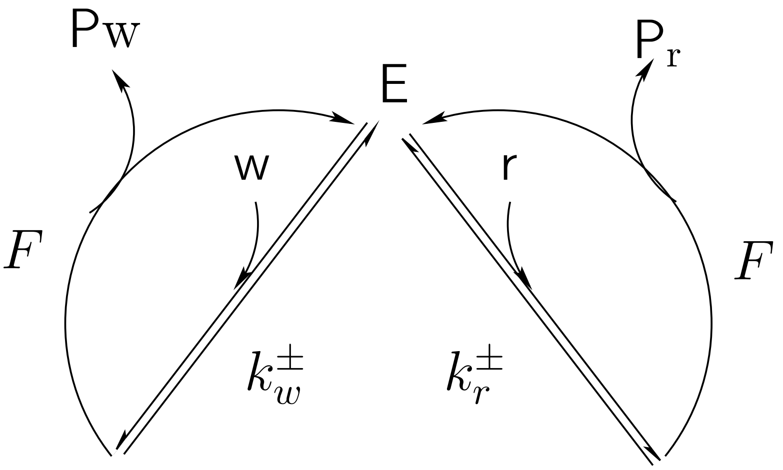
$$\xi = \xi_{\mathrm{eq}}\exp\left[-\Delta S^{(\mathrm{w})} + (\Delta W^{(\mathrm{w})} - \Delta F)/k_{\mathrm{B}}T\right]. \tag{3.22}$$

This relation highlights that reducing the error below its equilibrium value with a fixed work budget requires some entropy production, i.e., dissipation.

## 3.2 Models

### Kinetic and energetic discrimination

Let us consider a simple Michaelis-Menten process, described by the scheme

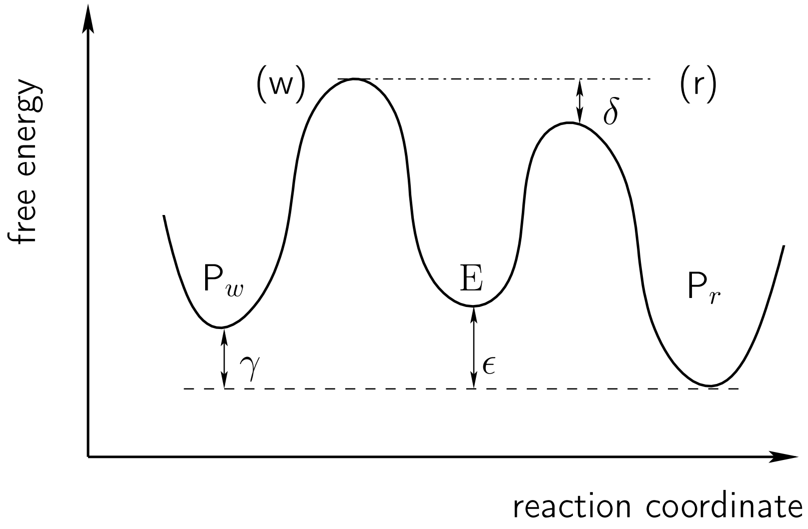


Here E denotes the empty enzyme, and the rates are given by

$$k_\mathrm{r}^+ = \omega \mathrm{e}^{\delta+\epsilon}, \qquad k_\mathrm{r}^- = \omega \mathrm{e}^{\delta}; \tag{3.23}$$

$$k_\mathrm{w}^+ = \omega \mathrm{e}^{\epsilon}, \qquad k_\mathrm{w}^- = \omega \mathrm{e}^{\gamma}. \tag{3.24}$$

The parametrization of the rates can be physically interpreted in terms of the free-energy landscape represented in the following scheme (energy units in $k_\mathrm{B}T$, scheme after Sartori and Pigolotti [2013]):



Thus $\delta$ parametrized the difference in barrier height that the substrate-enzyme complex must overcome to go to the product state, while $\gamma$ parametrizes the free-energy difference of the final product. One can then straightforwardly evaluate the steady state distribution and the error rate, obtaining

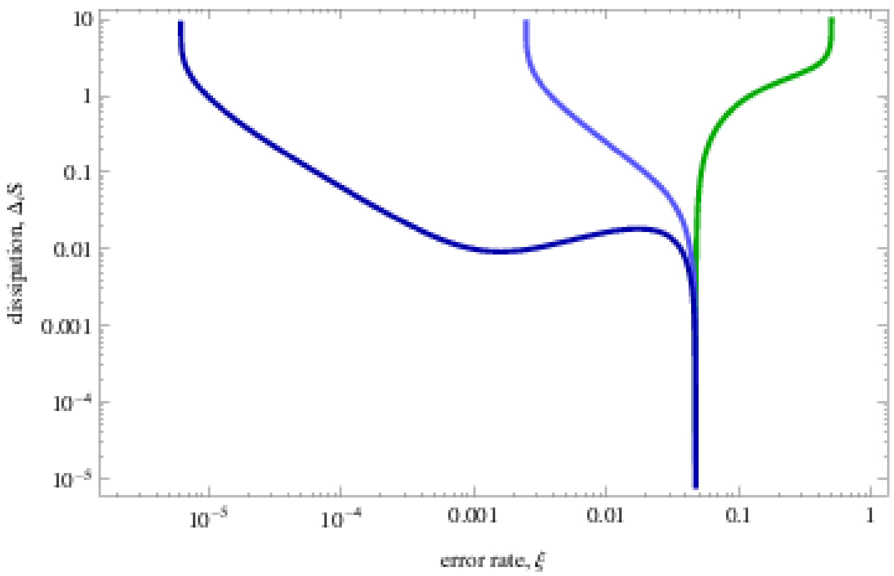
$$\xi = \frac{F\bar{p}_{\mathrm{w}}}{F\bar{p}_{\mathrm{w}} + F\bar{p}_{\mathrm{r}}} = \frac{\mathrm{e}^{\delta}\omega + F}{(\mathrm{e}^{\gamma}+1)\,\mathrm{e}^{\delta}\omega + \left(\mathrm{e}^{\delta}+1\right)F} \geq \frac{1}{\mathrm{e}^{\max\{\delta,\gamma\}}+1} \simeq \mathrm{e}^{-\max\{\delta,\gamma\}}. \tag{3.25}$$

We can therefore distinguish two discrimination regimes:

**Energetic discrimination:** $\gamma > \delta$. In this regime, the minimum error rate $\xi_{\min}$ is reached close to equilibrium ($F \to 0$).

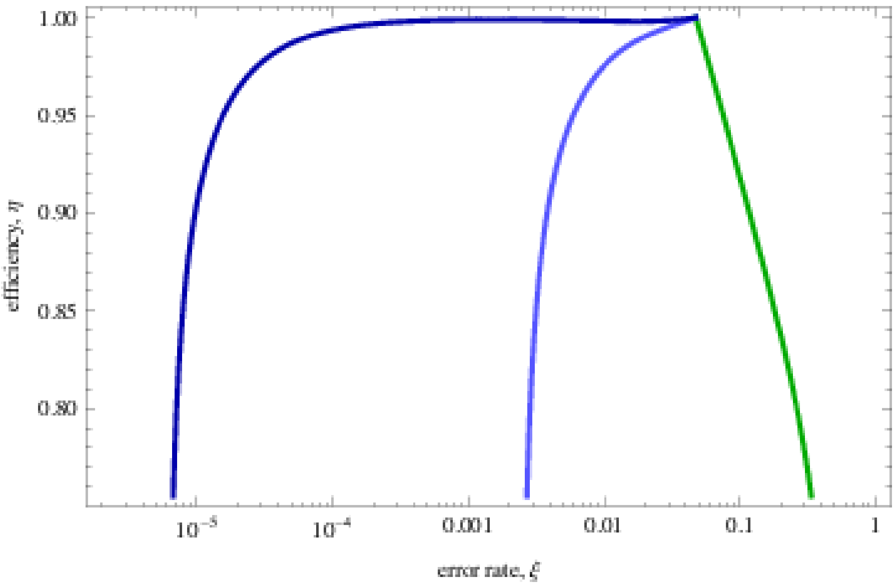
**Kinetic discrimination:** $\gamma < \delta$; In this regime, $\xi_{\min}$ is reached for $F \to \infty$.

The behavior of the dissipation as a function of the error rate for the different regimes is shown in the following plot.

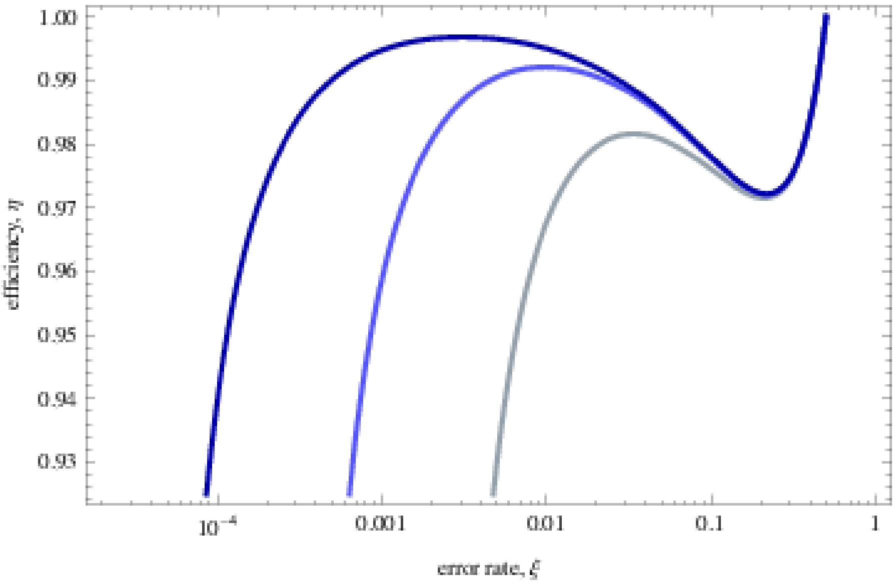


The entropy production per unit product $\Delta_{\mathrm{i}}S$ is plotted as a function of the error rate $\xi$ for the simple Michaelis-Menten model with $\gamma = 3$, $\epsilon = 10$, $\omega = 1$ and $\delta = 0$ (green), $\delta = 6$ (light blue), $\delta = 12$ (dark blue). Note that the blue lines (for which $\delta > \gamma$, corresponding to kinetic discrimination) yield a monotonically decreasing error rate as $\Delta_{\mathrm{i}}S$ grows. The green line corresponds to energetic discrimination, and the error rate increases as $\Delta_{\mathrm{i}}S$ grows. From Peliti and Rao [2015].

We can also evaluate the efficiency $\eta$ as a function of the error rate $\xi$, obtaining the following plot, also from Peliti and Rao [2015].



Here the efficiency $\eta$ (defined in (3.18)) is plotted vs. the error rate $\xi$ for the simple Michaelis-Menten model. The parameters are as in the previous figure. Note that the efficiency decreases with the error rate for kinetic discrimination (blue lines), while it increases for energetic discrimination (green line). To understand in greater detail what happens for the kinetic discrimination regime in the intermediary error-rate range, one can blow up the plateau which appears for the blue curves in the previous figure, obtaining the following plot (note the change of scales).



Efficiency-error trade-off in the purely kinetic regime of discrimination, for the simple Michaelis-Menten model with $\gamma = 3$, $\epsilon = 10$, $\omega = 1$ and $\delta = 6$ (grey), $\delta = 8$ (light blue),

$\delta = 10$ (dark blue). From Peliti and Rao [2015]. Note the presence of an efficiency maximum for intermediate values of the error rate. In contrast, we have seen that the entropy production $\Delta \mathrm{i} S$ monotonically increases as the error rate gets smaller.

To make these considerations more concrete, let us consider the working regimes of two different DNA polymerases.

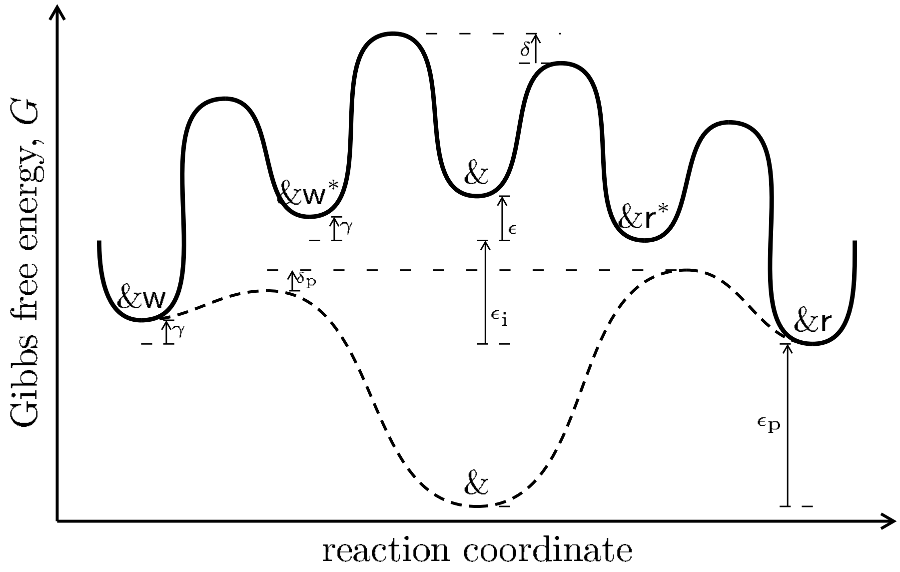
- T7 Phage DNA polymerase: Estimates from Tsai and Johnson [2006] yield $\eta \sim 10^{-2} \div 10^{-6}$, $v \sim 5 \div 250$bps, thus $\epsilon \sim 5$, $\gamma \sim 14$, $\delta \sim 8$: apparently this polymerase works in the energetic regime.

- Pol$\gamma$ (human) DNA polymerase: Estimates by Andrieux and Gaspard [2008] yield $\eta \sim 10^{-3} \div 10^{-5}$, with $\epsilon \sim 5$ this yields $\delta \sim 11$, $\gamma \sim 0$, which would correspond to the kinetic regime.

## The Ninio-Hopfield model

In this context, proofreading appears as the opening of a pathway that allows for the release of the substrate before reaching the product release step. This is summarized by the following diagram, in which the empty enzyme is denoted by &. Note the appearance of an intermediate step &s* (with $\mathrm{s} \in \{\mathrm{r}, \mathrm{w}\}$), from which the system can proceed further to the product release step or backward to the empty-enzyme state. The final complexed state, &s, can also reach the empty-enzyme state by the pathway denoted by (3), which involves a large decrease in the Gibbs free energy, and is therefore essentially irreversible (this step is represented by the broken line in the free-energy diagram).

$$\begin{tikzcd}
\& \arrow[r, shift left, "(1)\quad k_{\mathrm{s}}^{+}"] \arrow[r, leftarrow, shift right, "k_{\mathrm{s}}^{-}"'] \arrow[rr, bend right=40, Leftrightarrow, "K_{\mathrm{s}}^{-} \qquad (3) \qquad K_{\mathrm{s}}^{+}"'] & \&\mathrm{s}^{*} \arrow[r, shift left, "(2)\quad h_{\mathrm{s}}^{+}"] \arrow[r, leftarrow, shift right, "h_{\mathrm{s}}^{-}"'] & \&\mathrm{s} \arrow[r, "F"] & \mathrm{S}
\end{tikzcd}$$



With this scheme, the reaction rates can be parametrized as follows:

$$k_r^+ = \omega e^{\delta+\epsilon}, \qquad k_r^- = \omega e^{\delta}, \tag{3.26}$$

$$k_w^+ = \omega e^{\epsilon}, \qquad k_w^- = \omega e^{\gamma}, \tag{3.27}$$

$$h_r^+ = \omega_i e^{\epsilon_i}, \qquad h_r^- = \omega_i, \tag{3.28}$$

$$h_w^+ = \omega_i e^{\epsilon_i}, \qquad h_w^- = \omega_i, \tag{3.29}$$

$$K_r^+ = \omega_p e^{-\delta_p}, \qquad K_r^- = \omega_p e^{\epsilon_p-\delta_p}, \tag{3.30}$$

$$K_w^+ = \omega_p, \qquad K_w^- = \omega_p e^{\epsilon_p+\gamma}. \tag{3.31}$$

Here the label "p" denotes the proofreading pathway. The results for this model are shown in figure 3.2.

We can conclude that kinetic and energetic discrimination regimes can cooperate in the proofreading pathway, reducing the error rate. Moreover, faster, more dissipative and more efficient process obtains when the kinetic discrimination predominates on the first pathway. The minimum error rate is given by $\xi_{\min} = e^{-2\gamma}$ (Hopfield [1974]):

$$\xi_{\min} \simeq e^{-(\max(\gamma,\delta)+\gamma+\delta_p)} \tag{3.32}$$

Let us point out that the minimum error rate is always achieved in the vanishing catalysis rate ($F \to 0$) limit. In this regime all the free energy provided to the system is spent to proofread at the expense of the process speed. Hence, both the mean step time and the dissipation diverge when one approaches the minimum error rate.

## The Murugan-Huse-Leibler model

One can consider a more general class of models, where the reaction network contains multiple proofreading cycles. A number of these models have been introduced by Murugan et al. [2012] (see also Murugan et al. [2014]), and are represented by schemes of the following form, where the broken lines represent an indefinite number of cycles:

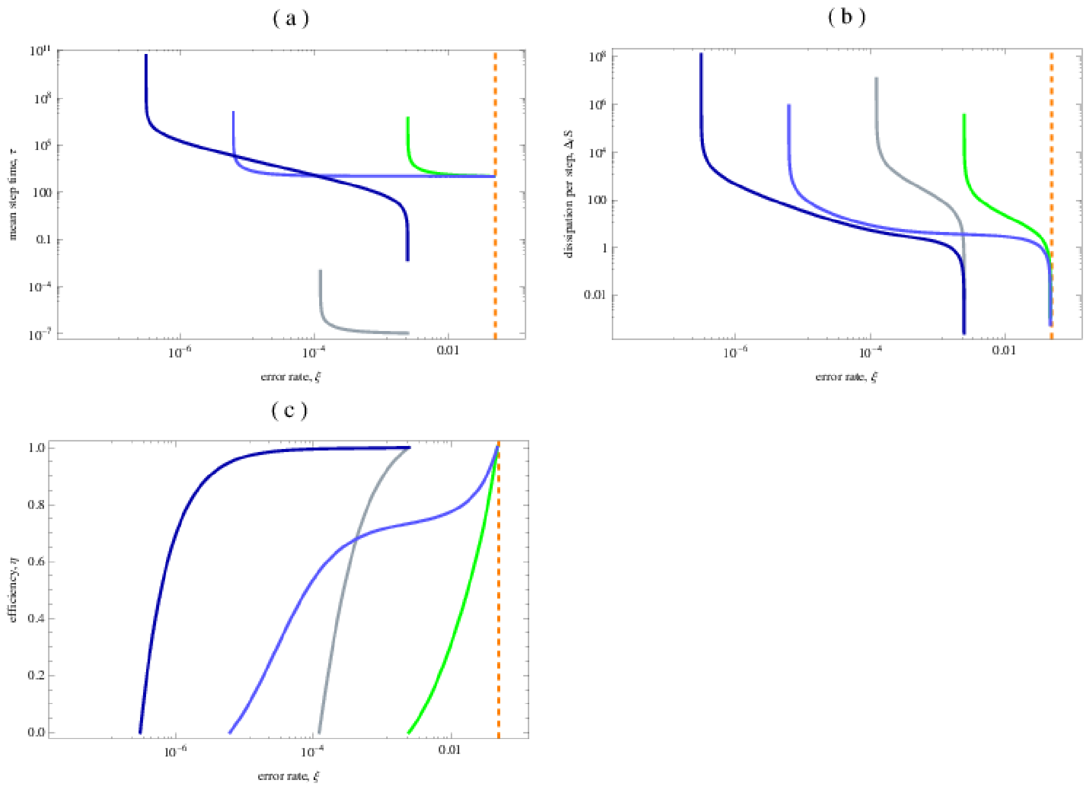
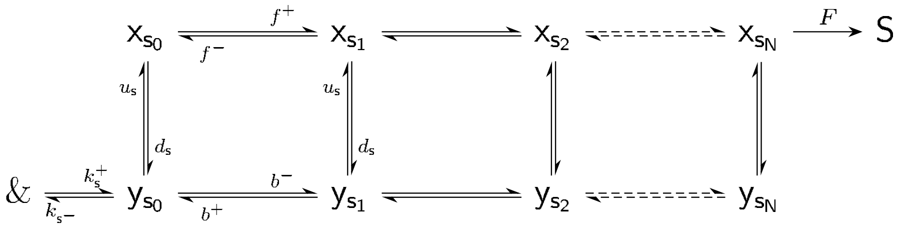


Figure 3.2: Mean step time (a), dissipation per step (b), and efficiency (c) versus the error-rate for the Hopfield model. We have chosen the following values for the coefficients: $\epsilon = 10$, $\omega = 1$, $\gamma = 3$ and $(\delta, \delta_\mathrm{p}) = (0,0)$ (green curve: purely energetic discrimination), $(\delta, \delta_\mathrm{p}) = (6,0)$ (light blue curve: kinetic discrimination on the first chemical pathway), $(\delta, \delta_\mathrm{p}) = (0,6)$ (blue curve: kinetic discrimination on the proofreading chemical pathway), $(\delta, \delta_\mathrm{p}) = (6,6)$ (dark blue curve: kinetic discrimination on both chemical pathways). The other constant, namely $\omega_\mathrm{i}$, $\omega_\mathrm{p}$, $\epsilon_\mathrm{i}$, $\epsilon_\mathrm{p}$, are those which minimize the error-rate function $\xi$. In this way, the minimum achievable error rate is recovered by the trade-offs. They have been obtained by numerically minimizing the error-rate function given the discrimination constants and the driving energy $\epsilon$, and are consistent with the values predicted in Hopfield [1974]. The dashed orange curve refers to the equilibrium error rate for $\gamma = 3$. Therefore, in the Hopfield model one consistently obtains smaller error rates than in the MM model for equal values of the discrimination constants. The light blue and blue curves highlight that kinetic and energetic discrimination regimes yield a resulting lower error rate only when they cooperate in the proofreading pathway. Remarkably, when the kinetic discrimination predominates on the first pathway the process becomes faster, more dissipative and more efficient. Finally, it is worth noting that high efficiency at lower error rates appears as a general trait of the kinetic discrimination for simple models, like the MM and the Hopfield ones. From Peliti and Rao [2015].



Any substrate chemical network is coupled with $N$ chemical forces of equal affinity $\mathscr{A}$. These forces act on the $N$ cycles of type $y_{s_n} \to x_{s_n} \to x_{s_{n+1}} \to y_{s_{n+1}} \to y_{s_n}$ and drive the chemical complex towards the final product state on the $x$-chain of reactions and towards the free enzyme state on the $y$-chain of reactions. The $x_{s_i} \rightleftharpoons y_{s_i}$ reactions discriminate the substrates either kinetically or energetically. The discrimination constants, denoted by $\gamma$ and $\delta$, and will be taken to be equal for any of the $N+1$ pathways $x_{s_i} \rightleftharpoons y_{s_i}$.

We will focus our analysis to the behavior of the chemical network for varying number of independent forces, and thus of cycles. For this purpose we neglect the discrimination performed on the first stage $\& \rightleftharpoons y_{s_1}$ and assume that the related rate constants are the same for both substrates: $k_r^+ = k_w^+ = \omega e^{\epsilon}$ and $k_r^- = k_w^- = \omega$.

Thus the rate constants related to the ladder part of the network are expressed by

$$u_r = \omega_u e^{\epsilon_U + \delta}, \qquad d_r = \omega_u e^{\delta}, \tag{3.33}$$

$$u_w = \omega_u e^{\epsilon_u}, \qquad d_w = \omega_u e^{\gamma}, \tag{3.34}$$

$$f^+ = \omega_f, \qquad f^- = \omega_f e^{-\epsilon_f}, \tag{3.35}$$

$$b^+ = \omega_b, \qquad b^- = \omega_b e^{-\epsilon_b}. \tag{3.36}$$

We can evaluate the error rate, the mean step rate and the efficiency by solving the steady-state dynamics, inverting the related error-rate function, and substituting the result in the relevant expressions, obtaining the results shown in fig. 3.3.

## Conclusions

Natural biological processes, like the enzyme assisted assembly processes that we have considered, must satisfy several, sometimes contradictory, requirements. The error rate should be lower than a threshold, the average duration of the process should not be too long, and the free-energy consumption should not be too large. It is likely that the detailed mechanisms of naturally occurring processes have evolved to optimize some combination of these quantities, whose expression depends on the requirements of the process itself. Clarifying the interplay between these requirements will constitute an important step towards formulating general thermodynamic principles limiting the accuracy of non-equilibrium information processing.

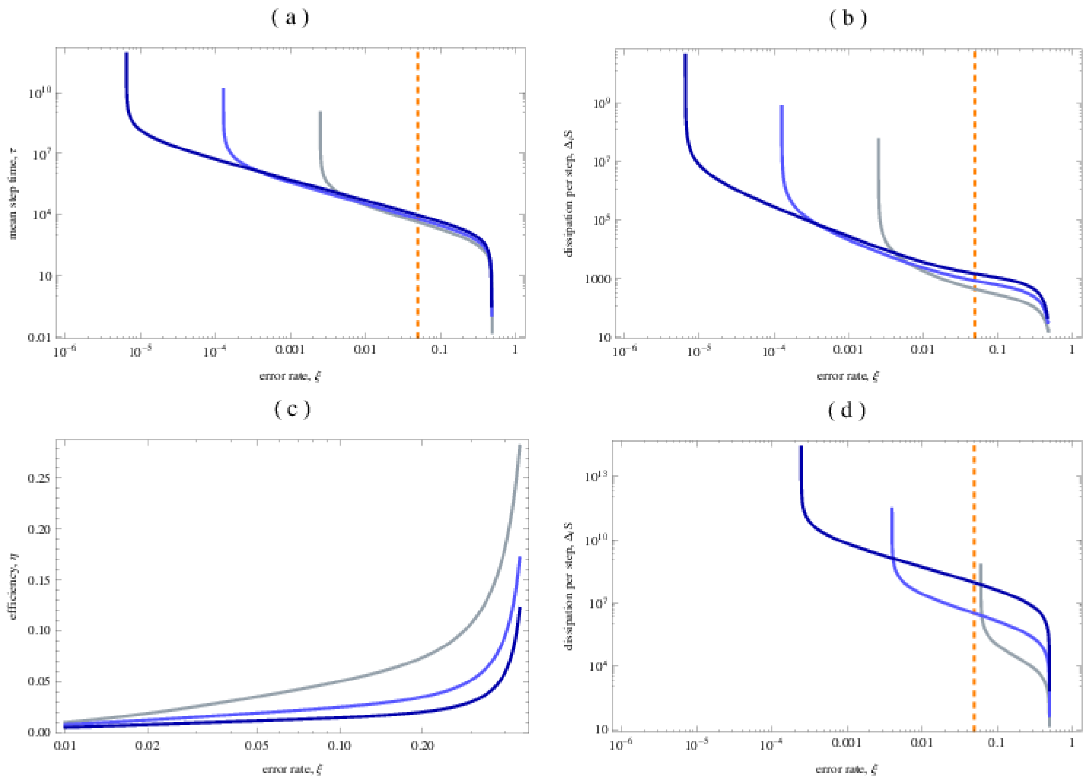


Figure 3.3: Mean step time (a), dissipation per step (b), and efficiency (c) versus the error rate for the MHL model in the purely energetic discrimination regime ($\gamma = 3$ and $\delta = 0$). The plot (d), instead, shows the dissipation per step for the same model in the purely kinetic discrimination regime ($\gamma = 0$ and $\delta = 3$). The curves are distinguished by the number of cycles in the network: light green and light purple curves, $N = 1$; green and purple curves, $N = 2$; dark green and dark purple curves, $N = 3$. The curves in (a–c) correspond to the following values of the parameters: $\gamma = 3$, $\epsilon_{\mathrm{u}} = 8$, $\epsilon_{\mathrm{f}} = 8$ and $\epsilon_{\mathrm{b}} = 8$, which represent the driving energy related to the reactions $x_{s_i} \to y_{s_i}$, $x_{s_i} \to x_{s_{i+1}}$ and $y_{s_{i+1}} \to y_{s_i}$, respectively. The other constant, $\omega_{\mathrm{u}}$, $\omega_{\mathrm{f}}$, $\omega_{\mathrm{b}}$, are those which minimize the minimal error rate $\xi_{\mathrm{min}}$, when the other parameters are kept fixed. The first two plots in (a) and (b) highlight the lowering of the error rate as the number of forces increase. However, both the mean step time and the entropy production increase and the system exhibits a progressively lower efficiency (c). Interestingly, in the kinetic regime (d), the system needs at least two forces in order to reduce the error rate. Let us observe that the model embedding just one cycle is very similar to Hopfield scheme, except for the non-discriminating pathways $\& \rightleftharpoons \mathrm{y}_{\mathrm{s}_0}$. We thus suppose that these last pathways, detaching the free enzyme state from the part of the network performing the proofreading, spoil the discrimination in the kinetic regime. Finally, for the plot (d) the numeric constants chosen are: $\omega = 1$, $\epsilon = 10$, $\delta = 3$. The values of the other coefficients are those which minimize the minimal error rate $\xi_{\mathrm{min}}$, and are given by $\epsilon_{\mathrm{u}} = 8$, $\epsilon_{\mathrm{f}} = 9$, $\epsilon_{\mathrm{b}} = 9$, $\omega_{\mathrm{u}} = 0.1$, $\omega_{\mathrm{f}} = 10$, $\omega_{\mathrm{b}} = 10$). From Peliti and Rao [2015].

APPENDIX

# A

# Complements

## A.1 Convex functions and Jensen's inequality

Many inequalities in information theory are a consequence of a general relation called *Jensen's inequality,* which holds for *convex functions.*

**Definition 1** *A function $f(x)$ is said to be* **convex** *over an interval $(a,b)$ if given any $x_0, x_1 \in (a,b)$ and $\alpha \in [0,1]$, one has*

$$f(\alpha x_0 + (1-\alpha)x_1) \le \alpha f(x_0) + (1-\alpha) f(x_1). \tag{A.1}$$

A function $f$ is said to be *concave* if $-f$ is convex. If the inequality is strict for any $x_0 \neq x_1$ and $\alpha \in (0,1)$, $f$ is said to be *strictly convex.*

**Theorem 1** *If the function $f(x)$ has a second derivative which is non-negative (positive) in $(a,b)$, then $f$ is convex (strictly convex).*

*Proof.* Given $x_0, x$ we have, by Taylor's expansion and by the mean-value theorem, for some $u$ between $x_0$ and $x$,

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(u)(x-x_0)^2, \tag{A.2}$$

where the last term is non-negative. Now let $x = \alpha x_0 + (1-\alpha)x_1$. We obtain

$$f(x_0) \ge f(x) + f'(x)(x_0 - x) = f(x) - f'(x)(1-\alpha)(x_1 - x_0). \tag{A.3}$$

Similarly we obtain

$$f(x_1) \ge f(x) + f'(x)(x_1 - x) = f(x) + f'(x)\alpha(x_1 - x_0). \tag{A.4}$$

Multiplying (A.3) by $\alpha$ and (A.4) by $1-\alpha$ and adding, we obtain (A.1). The strict inequality is derived along the same lines when $f''(x) > 0$, $\forall x \in (x_0, x_1)$. ■

The converse of this theorem also holds in the form:

**Theorem 2** *Let the convex function* $f(x)$ *be twice derivable in the interval* $[x_0, x_1]$. *Then* $f''(x) \geq 0$ *in the interval.*

*Proof.* By definition, we have, for any $x$ for which $f''(x)$ exists,

$$f''(x) = \lim_{h \to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}. \tag{A.5}$$

On the other hand, by convexity, one has

$$\frac{1}{2}\big(f(x+h) + f(x-h)\big) \geq f(x). \tag{A.6}$$

Thus the quantity on the right-hand side of the relation above is non-negative for all $h \neq 0$. In the limit we obtain $f''(x) \geq 0$. ■

One can see that the strict inequality need not hold, even if $f(x)$ is strictly convex, by considering the function $f(x) = x^4$.

We shall also need the following lemma:

**Theorem 3** *Let* $f(x)$ *be a convex function which is twice derivable in* $(a, b)$ *and let* $x_0 \in (a, b)$. *Then* $f(x) \geq f(x_0) + f'(x_0)(x - x_0)$ *for* $x \in (a, b)$. *The inequality is strict if* $f(x)$ *is strictly convex.*

*Proof.* By Taylor expansion and the mean-value theorem, we have, for some $u \in (x_0, x)$,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(u)(x - x_0)^2, \tag{A.7}$$

where the last term is non-negative. Thus we have proved the weak inequality. To prove the strict inequality, let us assume that $f(x)$ is strictly convex, and that $f(x) = f(x_0) + f'(x_0)(x - x_0)$ for some $x \neq x_0$. Then, by convexity, we have, for any $u \in (x_0, x)$,

$$f(u) \leq f(x)\frac{u - x_0}{x - x_0} + f(x_0)\frac{x - u}{x - x_0} = f'(x_0)(u - x_0) + f(x_0). \tag{A.8}$$

On the other hand, by the result we have just derived, we have

$$f(u) \geq f'(x_0)(u - x_0) + f(x_0). \tag{A.9}$$

This implies

$$f(u) = f(x)\frac{u - x_0}{x - x_0} + f(x_0)\frac{x - u}{x - x_0}, \tag{A.10}$$

which violates the strict convexity. ■

If $X$ is a random variable, let us denote by $\langle f \rangle$ the average of $f(x)$ according to the distribution of $X$. We then have

**Theorem 4 (Jensen's inequality)** *If $f$ is a convex function and $X$ is a random variable, then*

$$\langle f\rangle \geq f(\langle x\rangle). \tag{A.11}$$

*If $f$ is strictly convex, the equality in (A.11) implies that $X$ is a constant.*

*Proof.* Define $g(x) := f(\langle X\rangle) + f'(\langle X\rangle)(x - \langle X\rangle)$. Since $f(x)$ is convex, one has $f(x) \geq g(x)$, $\forall x$, with equality for $x = \langle X\rangle$. Moreover, if $f(x)$ is strictly convex, one has $f(x) > g(x)$, $\forall x \neq \langle X\rangle$. Since $g(x)$ is linear, $\langle g(X)\rangle = g(\langle X\rangle)$. Thus we have

$$\langle f\rangle \geq \langle g\rangle = g(\langle X\rangle) = f(\langle X\rangle). \tag{A.12}$$

The inequality is strict if $f(x)$ is strictly convex, unless $X$ assumes only one value. ■

## A.2 The basics of large-deviation theory

Let us consider a sequence $X_t$ of i.i.d. random variables, assuming the value $x_k$ (for, say, $k \in \{1,\dots q\}$) with probability $P = (p_k)$. Let us consider the behavior of the sum $M_n$ of the values of $X_t$ for $t = 1,\dots n$:

$$M_n := \sum_{t=1}^{n} X_t. \tag{A.13}$$

Then we know that $\langle M_n\rangle = n\langle X\rangle$ and $\sigma_n^2 := \langle (M_n - \langle M_n\rangle)^2\rangle = n\langle (X - \langle X\rangle)^2\rangle$, as long as $\langle X\rangle$ and $\langle X^2\rangle$ exist. Moreover, from the Central Limit theorem, we know that in this case the probability that $M_n$ exhibits a small deviation (of order $\sqrt{n}$) from its average $n\langle X\rangle$ behaves like a Gaussian:

$$P(M_n{=}M) \propto \exp\left[-\frac{(M - n\langle X\rangle)^2}{2\sigma_n^2}\right]. \tag{A.14}$$

The aim of large-deviation theory is to extend these results to cases in which $M_n$ deviates from its average by a quantity of order $n$. In our simple case, it can be shown that the distribution of $M_n$ satisfies a **large-deviation principle** for $n \to \infty$. The principle stipulates that the probability density of the mean $m := M_n/n$ behaves exponentially for large $n$. This implies that the following limit exists:

$$I(m) := -\lim_{n\to\infty} \frac{1}{n} \log P(M_n{=}nm). \tag{A.15}$$

The function $I(m)$ is called the **rate function**. Since the limit vanishes for $m = \langle X\rangle$, $I(m{=}\langle X\rangle) = 0$. The relation (A.15) is often written in the form

$$P_n(M) \asymp \mathrm{e}^{-nI(M/n)}. \tag{A.16}$$

It is instructive to evaluate the rate equation for our simple system. The probability distribution function $P_n(M)$ is given by

$$P_n(M) = \sum_{k_1\cdots k_n} \prod_{t=1}^{n} p_{k_t} \delta\left(\sum_{t=1}^{n} x_{k_t} - M\right). \tag{A.17}$$

This expression is awkward due to the delta function. We can however make it more compliant by introducing the integral representation of the delta function:

$$\delta(x) = \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}y}{2\pi\mathrm{i}}\, \mathrm{e}^{yx}. \tag{A.18}$$

We thus obtain

$$\begin{aligned} P_n(M) &= \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}y}{2\pi\mathrm{i}} \sum_{k_1\cdots k_n} \prod_{t=1}^{n} p_{k_t} \exp\left[ y \sum_{\tau=1}^{n} x_{k_\tau} - yM \right] \\ &= \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}y}{2\pi\mathrm{i}}\, \mathrm{e}^{yM} \prod_{t=1}^{n} \sum_{k_t} p_{k_t}\, \mathrm{e}^{-yx_{k_t}} = \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}y}{2\pi\mathrm{i}}\, \mathrm{e}^{-yM} \left\langle \mathrm{e}^{yx} \right\rangle^{n}. \end{aligned} \tag{A.19}$$

Setting $M = nm$ and defining $\psi(y) = -\log\langle \mathrm{e}^{yx}\rangle$ we can evaluate the integral by the saddle-point method:

$$P_n(M) = \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}y}{2\pi\mathrm{i}} \exp\left[-n\left(ym + \psi(y)\right)\right] \approx \mathrm{e}^{-nI(m)}, \tag{A.20}$$

where

$$I(m) := \min_{y} \left(\psi(y) + ym\right). \tag{A.21}$$

Thus, the rate function is obtained by the Legendre transform of the $\psi(y)$, which is the cumulant-generating function of the initial distribution $p_x$.

There is a second level of large-deviation theory, which deals with the asymptotic behavior of the empirical distributions. In our simple case, the empirical distribution $\Xi_n$ is a collection of $q$ random variables $\Xi_n = (\xi_{n,1}, \ldots \xi_{n,q})$ defined by

$$\xi_{n,\ell} := \frac{1}{n} \sum_{k=1}^{\ell} \delta_{x_k,\ell}. \tag{A.22}$$

Thus $\xi_{n,\ell}$ measures the frequency of the outcome $\ell$ over $n$ trials. By the law of large numbers we expect that $\lim_{n\to\infty} \xi_{n,\ell} = p_\ell$ with probability one. We are interested in evaluating the probability that $\Xi_n$ differs substantially from $P$ for large $n$. Given $\xi = (\xi_1, \ldots, \xi_q)$, the probability that $\Xi = \xi$ is given by the probability that $n_\ell = n\xi_\ell$ for $\ell = 1, \ldots, q$, where $n_\ell$ is the number of times that $X = \ell$ in $n$ trials. This probability is given by

$$P(\Xi{=}\xi) = \frac{n!}{(n\xi_1)!(n\xi_2)!\cdots(n\xi_q)!} p_1^{n\xi_1} \cdots p_q^{n\xi_q} \approx \exp\left[ -n \sum_{k=1}^{q} \xi_k \left(\log \xi_k - \log p_k\right) \right], \tag{A.23}$$

where we have used Stirling's formula. We recognize in the sum the Kullback-Leibler divergence between the empirical distribution $\Xi$ and the probability $P$:

$$D_{\mathrm{KL}}[\Xi \| P] = \sum_{k=1}^{n} \xi_k \log \frac{\xi_k}{p_k}. \tag{A.24}$$

Thus the probability of the empirical distribution satisfies a large-deviation principle, whose rate function is the Kullback-Leibler divergence between the empirical distribution itself and the event probability. The behavior of $\Xi_n$ for $n \to \infty$, as dictated by the law of large numbers, follows immediately. This result was derived by Boltzmann [1877]. In the words of Ellis [1999]:

> This fundamental calculation represents a revolutionary moment in human culture during which both statistical mechanics and the theory of large deviations were born.

# Bibliography

Andrieux, D. and Gaspard, P. (2008). Nonequilibrium generation of information in copolymerization processes. *Proc. Natl. Acad. Sci. USA*, 105:9516–9521.

Andrieux, D. and Gaspard, P. (2009). Molecular information processing in nonequilibrium copolymerizations. *Journal of Chemical Physics*, 130:014901.

Bennett, C. H. (1979). Dissipation-error tradeoff in proofreading. *Biosystems*, 11:85–91.

Bennett, C. H. (2003). Notes on Landauer's principle, reversible computation and Maxwell's demon. *Studies in History and Philosophy of Modern Physics*, 34:501–510.

Boltzmann, L. (1877). Uber die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respecktive den Sätzen über das Wärmegleichgewicht. *Wiener Berichte*, 2:373–435.

Brillouin, L. (1949). Life, thermodynamics and cybernetics. *American Scientist*, 37:554–568.

Bérut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., and Lutz, E. (2012). Experimental verification of Landauer's principle linking information and thermodynamics. *Nature*, 483:187–190.

Collin, D., Ritort, F., Jarzynski, C., Smith, S. B., Tinoco, I., and Bustamante, C. (2005). Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature*, 437:231–234.

Cover, T. M. and Thomas, J. A. (1991). *Elements of Information Theory*. Wiley Series in Telecommunications. Wiley.

Crooks, G. E. (1998). Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. *Journal of Statistical Physics*, 90:1481–1487.

Crooks, G. E. (1999). The entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60:2721–2728.

de Donder, T. (1936). *Thermodynamic Theory of Affinity: A Book of Principles*. Oxford U. P.

Ellis, R. E. (1999). The theory of large deviations: from Boltzmann's 1877 calculation to equilibrium macrostates in 2D turbulence. *Physica D*, 133:106–136.

Ellis, R. S. (1985). *Entropy, Large Deviations and Statistical Mechanics*. Springer.

Esposito, M. (2012). Stochastic thermodynamics under coarse-graining. *Phys. Rev. E*, page 041125.

Esposito, M. and Van den Broeck, C. (2010). Three detailed fluctuation theorems. *Phys. Rev. Lett.*, 104:090601.

Esposito, M. and van den Broeck, C. (2011). Second law and Landauer principle far from equilibrium. *EPL (Europhysics Letters)*, 95:40004.

Evans, D. J., Cohen, E., and Morriss, G. (1993). Probability of second law violations in shearing steady states. *Physical Review Letters*, 71:2401–2404.

Evans, D. J. and Searles, D. J. (1994). Equilibrium microstates which generate second law violating steady states. *Physical Review E*, 50:1645–1648.

Feynman, R. P. (1996). *Feynman Lectures on Computation*. Addison-Wesley. Edited by Hey, A.J. G. and Allen, R. W.

Gallavotti, G. (1996). Extension of Onsager's reciprocity to large fields and the chaotic hypothesis. *Physical Review Letters*, 77:4334–4337.

Gallavotti, G. and Cohen, E. G. D. (1995a). Dynamical ensembles in nonequilibrium statistical mechanics. *Physical Review Letters*, 74:2694–2697.

Gallavotti, G. and Cohen, E. G. D. (1995b). Dynamical ensembles in stationary states. *Journal of Statistical Physics*, 80:931–970.

Gaspard, P. (2004). Fluctuation theorem for nonequilibrium reactions. *Journal of Chemical Physics*, 120:8898–8905.

Giardinà, C., Kurchan, J., and Peliti, L. (2006). Direct evaluation of large-deviation functions. *Phys. Rev. Lett.*, 96:120603.

Hatano, T. and Sasa, S.-i. (2001). Steady state thermodynamics of Langevin systems. *Phys. Rev. Lett.*, 86:3463–3466.

Hill, T. L. (1989). *Free Energy Transduction and Biochemical Cycle Kinetics*. Dover.

Hopfield, J. J. (1974). Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. USA*, 71:4135–4139.

Hummer, G. and Szabo, A. (2001). Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl. Sci. USA*, 98:3658—-3661.

Imparato, A. and Peliti, L. (2007). The distribution function of entropy flow in stochastic systems. *J. Stat. Mech.*, page L02001.

Imparato, A., Sbrana, F., and Vassalli, M. (2008). Reconstructing the free-energy landscape of a polyprotein by single-molecule experiments. *EPL (Europhysics Letters)*, 82(5):58006.

Jarzynski, C. (1997). Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, pages 2690–2694.

Kawai, R., Parrondo, J., and Van den Broeck, C. (2007). Dissipation: the phase-space perspective. *Physical review letters*, 98:080602.

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

Lebowitz, J. L. and Spohn, H. (1999). A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics. *Journal of Statistical Physics*, 95(1):333–365.

Leff, H. and Rex, A. F., editors (2002). *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*. CRC Press.

Marini Bettolo Marconi, U., Puglisi, A., Rondoni, L., and Vulpiani, A. (2008). Fluctuation-dissipation: Response theory in statistical physics. *Physics Reports*, 461:111.

Murugan, A., Huse, D. H., and Leibler, S. (2012). Speed, dissipation, and error in kinetic proofreading. *Proc. Natl. Acad. Sci. USA*, 109:12034–12039.

Murugan, A., Huse, D. H., and Leibler, S. (2014). Discriminatory proofreading regimes in nonequilibrium systems. *Phys. Rev. X*, page 021016.

Ninio, J. (1975a). Kinetic amplification of enzyme discrimination. *Biochimie*, 57:587–595.

Ninio, J. (1975b). La précision dans la traduction génétique. In Sadron, C., editor, *École de Roscoff 1974. L'évolution des macromolécules biologiques*, pages 51–68. CNRS.

Parrondo, J. M. R., Horowitz, J. M., and Sagawa, T. (2015). Thermodynamics of information. *Nature Physics*, 11:131–139.

Peliti, L. and Rao, R. (2015). Thermodynamics of accuracy in kinetic proofreading: Dissipation and efficiency trade-offs. *Journal of Statistical Mechanics*, page P06001.

Phillips, R. and Quake, S. R. (2006). The biological frontier of physics. *Physics Today*, 59:38–43.

Prost, J., Joanny, J.-F., and Parrondo, J. M. R. (2009). Generalized fluctuation-dissipation theorem for steady-state systems. *Phys. Rev. Lett.*, 103:090601.

Qian, H. (2002). Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation. *Physical Review E*, 65:01602.

Ritort, F. (2007). The nonequilibrium thermodynamics of small systems. *C. R. Physique*, 8:528–539.

Sagawa, T. and Ueda, M. (2010). Generalized Jarzynski equality under nonequilibrium feedback control. *Phys. Rev. Lett.*, 104:090602.

Sartori, P. and Pigolotti, S. (2013). Kinetic versus energetic discrimination in biological copying. *Phys. Rev. Lett.*, 110:188101.

Sartori, P. and Pigolotti, S. (2015). Thermodynamics of error correction. *Phys. Rev. X*, 5:041039.

Seifert, U. (2005). Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.*, 95:040602.

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

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656.

Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. The University of Illinois Press, Urbana (Ill.).

Speck, T. and Seifert, U. (2005). Integral fluctuation theorem for the housekeeping heat. *J. Phys. A: Math. Gen.*, 38:L581–L588.

Szilard, L. (2006). On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behavioral Science*, 9:301–310.

Szilárd, L. (1929). Über die Entropieverminderung in einem thermodynamischen System bei Eingreffen intelligenter Wesen. *Zeitschrift für Physik*, 53:840–856.

Thomson, W. (1851). On the universal tendency in nature to the dissipation of mechanical energy. *Philosophical Magazine*, page 304.

Touchette, H. (2009). The large deviation approach to statistical mechanics. *Physics Reports*, pages 1–69.

Tsai, Y. C. and Johnson, K. A. (2006). A new paradigm for dna polymerase specificity. *Biochemistry*, 45:9675–9687.

Voliotis, M., Cohen, N., Molina-París, C., and Liverpool, T. B. (2009). Backtracking and proofreading in DNA transcription. *Phys. Rev. lett.*, 102:258101.

Wegscheider, R. (1901). Über simultane Gleichgewichte und die Beziehungen zwischen Thermodynamik und Reactionskinetik homogener Systeme. *Monatshefte für Chemie / Chemical Monthly*, page 849–906.

Yasuda, R., Hiroyuki, N., Kinosita, Jr., K., and Yoshida, M. (1998). $F_1$-ATPase is a highly efficient molecular motor that rotates with discrete 120° steps. *Cell*, 93:1117–1124.