# Some recent developments in non-equilibrium statistical physics

K MALLICK

Institut de Physique Théorique, CEA Saclay, 91191 Gif-sur-Yvette, France
E-mail: kirone.mallick@cea.fr

**Abstract.** We first recall the laws of classical thermodynamics and the fundamental principles of statistical mechanics and emphasize the fact that the fluctuations of a system in macroscopic equilibrium, such as Brownian motion, can be explained by statistical mechanics and not by thermodynamics. In the vicinity of equilibrium, the susceptibility of a system to an infinitesimal external perturbation is related to the amplitude of the fluctuations at equilibrium (Einstein's relation) and exhibits a symmetry discovered by Onsager. We shall then focus on the mathematical description of systems out of equilibrium using Markovian dynamics. This will allow us to present some remarkable relations derived during the last decade and valid arbitrarily far from equilibrium: the Gallavotti–Cohen fluctuation theorem and Jarzynski's non-equilibrium work identities. These recent results will be illustrated by applying them to simple systems such as the Brownian ratchet model for molecular motors and the asymmetric exclusion process which is a basic example of a driven lattice gas.

## 1. Introduction

Equilibrium statistical mechanics tells us that the probability of a microstate of a system in equilibrium with a thermal reservoir is given by the Boltzmann–Gibbs law: if $\mathcal{H}$ is the Hamiltonian of the system, the probability distribution over the configuration space is proportional to $\exp(-\beta\mathcal{H})$ where $\beta$ is the inverse of temperature. This canonical prescription is the starting point for studying a system in thermodynamic equilibrium: it has provided a firm microscopic basis for the laws of classical thermodynamics, has allowed us to describe different states of matter (from liquid crystals to superfluids), and has led to a deep understanding of phase transitions that culminated in the renormalization group theory.

For a system out of equilibrium, the probability of a given microstate evolves constantly with time. In the long time limit such a system may reach a stationary state in which the probability measure over the configuration space converges to a well-defined and constant distribution. If the system carries macroscopic stationary currents that represent interactions and exchanges of matter or energy between

internal parts of the system or with the external world, this stationary distribution is generically not given by the canonical Boltzmann–Gibbs law. At present, there exists no theory that can predict the stationary state of a system far from equilibrium from the knowledge of the microscopic interactions of its elementary constituents amongst themselves and with the environment, and from the set of dynamical rules that governs its evolution. The search for general features and laws of non-equilibrium statistical mechanics is indeed a central topic of the current research.

For systems close to thermodynamic equilibrium, linear response theory yields the fluctuation–dissipation relations and the Onsager reciprocity relations; but these relations are not satisfied far from equilibrium. However, recently, some general results, that remain valid for systems arbitrarily far from equilibrium, have been derived. In particular, two remarkable identities, the Gallavotti–Cohen fluctuation theorem and Jarzynski's work relation, have inspired many research works during the last decade.

The outline of this paper is as follows: In §2, we briefly review the principles of thermodynamics and discuss their statistical basis. We emphasize that thermodynamics considers average values and that fluctuations can be accounted for only by statistical mechanics. Near equilibrium, we give an example of the fluctuation–dissipation relation by discussing Einstein's formula for Brownian motion that allowed Jean Perrin to determine the Avogadro number. In §3, we describe the probabilistic approach to non-equilibrium dynamics by treating the case of Markovian systems. Two such processes of great importance are the random walk and stochastic dynamics. We outline the derivation of Einstein's formula by Langevin, who was the first to introduce a stochastic force in Newton's equation. We also explain that a fundamental relation satisfied by systems at thermodynamic equilibrium is the detailed balance. We review some basic properties of detailed balance, explain their connection to time-reversal invariance, hermiticity and to Onsager's reciprocity relations. In §4, we first discuss large deviation functions and apply this concept to the asymmetric exclusion process (ASEP), a simple driven lattice gas that plays the role of a paradigm in the field of non-equilibrium statistical physics. The example of ASEP allows us to illustrate a general connection between large deviations and cumulants of a probability distribution. We then present a generalization of detailed balance which holds for a large class of systems out of equilibrium. Thanks to the generalized detailed balance, we derive a version of the Gallavotti–Cohen fluctuation theorem. This theorem is applied to an elementary model of molecular motors and is shown to yield both Einstein's and Onsager's classical relations. We end this review by discussing Jarzynski's work theorem and some of its applications.

## 2. Thermodynamics: A short review

Thermodynamics describes macroscopic properties of matter (solid, fluid, radiation etc.) in terms of a small number of macroscopic observables (such as pressure, volume, mass, temperature) assuming that these properties do not vary with time. The laws of thermodynamics allow us to derive some general relations amongst

these properties irrespective of the structure of matter at the atomic scale. Indeed, the two principles were established during 19th century before the dawn of atomic physics.

Thermodynamics can be viewed as the science of energy conversions. In order to establish a correct balance, two guiding principles must be respected:
(i) all forms of energy involved must be identified correctly and accounted for;
(ii) different forms of energy are not equivalent. Some energy conversions are free of cost while others come with a fee and need compensation (according to Clausius).

Thermodynamics is one of the most elegant branches of physics, but it is also notoriously difficult. This feature has been perfectly emphasized by Onsager: "As in other kinds of bookkeeping, the trickiest questions that arise in the application of thermodynamics deal with the proper identification and classification of the entries; the arithmetics is straightforward" (Onsager, 1967).

2.1 *The two principles of thermodynamics*

We shall start by reviewing some elementary conversion problems. The simplest example is the conversion of mechanical energy into different forms (kinetic and potential); a ball that falls from a height $h$ reaches the ground with velocity $v^2 = 2gh$ where $g \sim 9.8$ m/s$^2$ is the acceleration due to gravity. This is the content of the celebrated experiments that Galileo is said to have performed from Pisa's leaning tower (see figure 1). In this elementary calculation, the friction of air has been neglected: this conversion of potential energy into kinetic energy occurs without a fee (conservation of the total mechanical energy). The free fall of a body can also be used to perform a work $\mathcal{W}$ (e.g. by attaching it to a pulley), the value of which is given by

$$\mathcal{W} = \mathcal{E}_{\text{initial}} - \mathcal{E}_{\text{final}} = -\Delta\mathcal{E} = mgh\,,\tag{1}$$

where $\mathcal{E}$ represents the potential energy.

*Exercises:*

1. What is the order of magnitude of the maximal height that a pole-vaulter can jump? Recall that the highest speed a human being can reach is roughly 10 m/s.
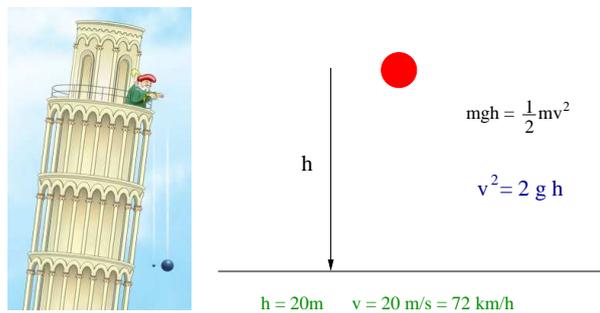


**Figure 1.** The legendary experiment of Galileo in Pisa and its schematic representation.
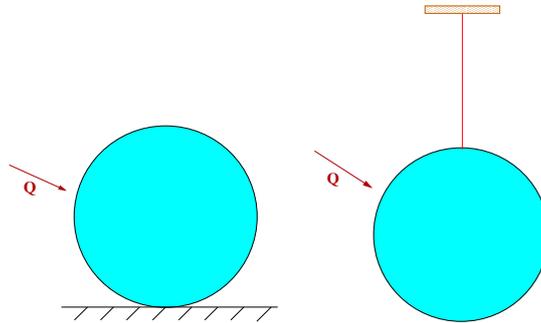
**Figure 2.** The two-spheres puzzle.

2. Describe the cyclic energy conversions in a pendulum. Draw the curves for potential energy and kinetic energy with respect to time.

The above processes are assumed to be free of dissipation and can be described in purely mechanical terms. In particular, they are perfectly reversible with time: for example, the motion of a pendulum clock allows us to measure time by slicing it into periods of constant duration but it does not tell us what the direction of time is: a movie displaying the oscillations of a dissipation-less pendulum can be run backwards in time without anybody noticing it.

In reality, some dissipation is always present: a ball bouncing on the ground looses energy at each shock and stops after a few bounces. Of course, energy as a whole is conserved because heat is evolved. This is the content of the First Principle of Thermodynamics: Heat must be taken into account when doing energy bookkeeping. The work of James Joule established that work and heat are two ways of exchanging energy between a system and its environment. This lead to the *First Principle of Thermodynamics*:

$$\Delta E = W + Q \quad \text{(The energy of the Universe is constant)}. \tag{2}$$

The energy $E$ in this equation is now the total internal energy of the system.

*Exercise*:

Energy balance problems can sometimes be rather subtle. Consider two perfectly identical spheres at the same temperature and made of the same material (figure 2). One sphere lies on the ground whereas the other is hanging, attached by a rigid thread. The same quantity of heat $Q$ is given to each sphere. Which sphere will be hotter? (we suppose that there is no heat transfer from a sphere to its environment, i.e., ground, air, thread etc.).

When dissipation is present, time-reversibility at macroscopic scale is lost. Projecting the movie of a ball bouncing on the ground backwards in time would display an impossible process: the ball would appear to bounce spontaneously higher and higher by absorbing heat from the ground. Such a process satisfies the first law of thermodynamics but would clearly never happen in reality.

In short, some processes are possible whereas others are not. It can be a very difficult task to detect the hidden flaw in some highly involved mechanisms. How
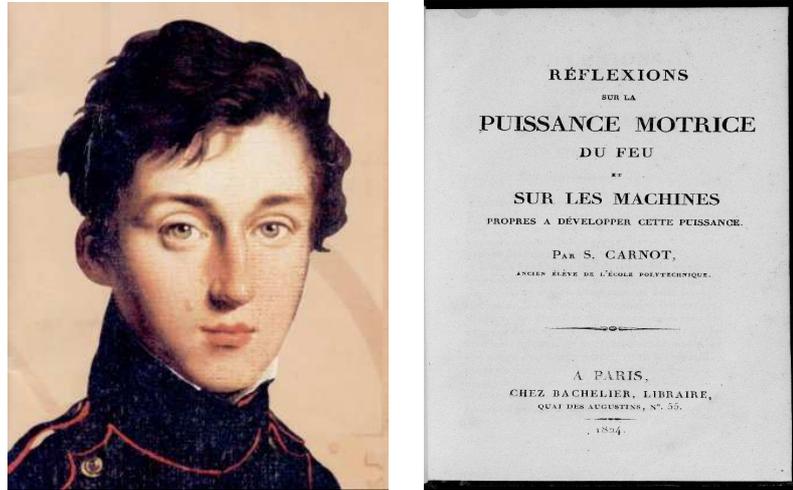
**Figure 3.** Sadi Carnot (1796–1832).

can one discriminate between possible and impossible processes? The solution to this problem is provided by the Second Law of Thermodynamics, elaborated by Carnot (1824) (figure 3), Clausius (1850) and Kelvin (1851). Two classical formulations of the second principle are [1]:

- *Clausius formulation*: No process is possible whose **sole** result is the transfer of heat from a cooler body to a hotter body.
- *Kelvin–Planck formulation*: No process is possible whose **sole** result is the absorption of heat from a reservoir and the conversion of that heat into work.

The Clausius and the Kelvin–Planck formulations present two elementary, universal, archetypical forbidden processes. These two statements can be shown to be equivalent and they cover all possible cases; they provide answer to the problem of deciding whether a given process is possible or not: by suitable couplings and mappings, an impossible process can always be transformed into a 'machine' that violates Kelvin–Planck's or Clausius' statement. At this stage, thermodynamics acquires a logical beauty akin to that of classical geometry. This elegant structure is perfectly demonstrated in the classical books of Fermi [2], Pippard [3] and in the recent textbook of Don Lemons [4].

The second principle was put on a quantitative basis by Clausius who introduced, in 1851, the entropy state function which measures the degree of irreversibility of a process. This is expressed by Clausius' inequality, which becomes an equality if and only if the process is reversible

$$S_2 - S_1 \geq \int_{1 \to 2} \frac{\partial Q}{T}. \tag{3}$$

A consequence of this inequality, when applied to the Universe considered as a whole, is that the entropy of the Universe is increasing. The Clausius or Kelvin–Planck statements of the second law can be reformulated in a more formal way:
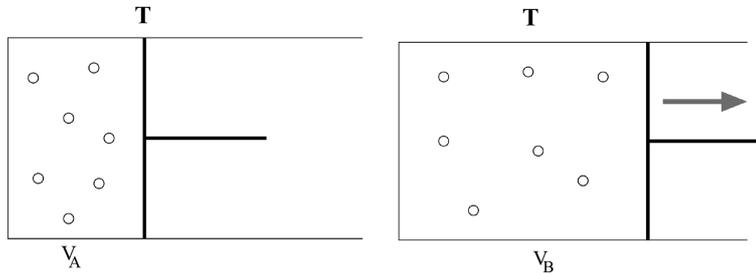
**Figure 4.** An illustration of the maximum work relation on a simple piston–gas system.

any process that would result in a decrease of the entropy of an isolated system is impossible.

Although energy is a familiar concept that plays a prominent role in many processes, one should never forget entropy that secretly drives many phenomena observed in daily life, for example the melting of ice. Ultimately, thermodynamic effects are due to the interplay of energy and entropy: A thermal system seeks to minimize its energy while maximizing its entropy at the same time. The subtle balance between these two state functions is encoded in the free energy $F$ that plays a fundamental role in statistical physics:

$$F = E - TS. \tag{4}$$

The interpretation of free energy as maximum available work is classical. Consider a system that evolves from a state A to a state B, both at a temperature $T$ equal to that of the environment (see figure 4 as an example). Suppose that the system exchanges heat only with its environment. Then, the decrease of free energy represents the maximum available work at the given temperature $T$:

$$\mathcal{W}_{\text{Avail}} \leq F_{\text{initial}} - F_{\text{final}} = -\Delta F. \tag{5}$$

The equality is valid when the process is reversible. Comparing with eq. (1) which is purely mechanical (with no heat transfer) we observe that the role of potential energy is now played by the thermodynamic potential $F$, and that the equality is replaced by an inequality because of dissipative effects.

Equivalently, the work $W$ we perform on the system is the opposite of the work $\mathcal{W}_{\text{Avail}}$ available from the system. Thus, we have

$$W \geq F_{\text{B}} - F_{\text{A}} = \Delta F. \tag{6}$$

Here again the equality occurs only for reversible processes. In general, because of irreversibility, we have to perform a work $W$ that is strictly greater than the free-energy variation. However, in the last section of this review, we shall explain that there exists an exact identity, known as Jarzynski's relation, which is valid for both reversible and irreversible processes. The classic inequality (6) is a consequence of Jarzynski's relation.

## 2.2 *Molecular theory of heat and the framework of statistical mechanics*

The works of Maxwell, Boltzmann and Gibbs have led to a statistical interpretation of entropy and to the foundations of statistical mechanics which encompasses classical thermodynamics and provides its microscopic basis.

The phase space of a macroscopic system that contains typically $10^{23}$ molecules is huge. Our senses and our measurement devices are sensitive only to a few global and averaged properties. In fact, at the microscopic scale, a system, even at equilibrium, evolves continuously from one microstate to another but most of the time these different microstates are perceived by us as being the same macroscopic state. We have indeed access only to extremely coarse-grained sense-data and a tremendous number of microstates are compatible with this data. Thus, coarse-graining from the microscale to the macroscale implies a huge loss of information: entropy quantifies this fact.

More precisely, consider an isolated system of macroscopic volume $V$ with total energy $E$. We suppose that all microscopic states of the system having energy $E$ are equiprobable: this assumption, known as the 'microcanonical probability ensemble' is a foundation-stone of statistical mechanics. (This can be proved rigorously for systems that display sufficiently strong ergodicity.) The thermodynamic entropy $S$ of a microcanonical system enumerates the total number of such microscopic states. If $\Omega(E, V)$ represents this total number, then the celebrated Boltzmann formula (figure 5) reads

$$S = k \log \Omega, \quad \text{with } k \simeq 1.38 \times 10^{-23}. \tag{7}$$

The determination of entropy is fundamentally a counting, combinatorial problem. Of course, there are many refinements to be studied in a course in statistical mechanics (see e.g. [5]). In particular, for a system at a given temperature $T$ (i.e. a system in contact with a thermal reservoir at temperature $T$), the probability of observing a microscopic configuration $\mathcal{C}$ of energy $E(\mathcal{C})$ is given by the Boltzmann–Gibbs canonical law:

$$P_{\text{eq}}(\mathcal{C}) = \frac{\mathrm{e}^{-E(\mathcal{C})/kT}}{Z}. \tag{8}$$

The partition function $Z$ (Zustandsumme) which insures that all probabilities sum up to 1 (normalization) is given by

$$Z = \sum_{\mathcal{C}} \mathrm{e}^{-E(\mathcal{C})/kT} = \sum_{E} \Omega(E) \mathrm{e}^{-E/kT}. \tag{9}$$

The canonical law, which implies a probabilistic description of the microscopic structure of a thermal system, can be deduced logically from Boltzmann's formula [5]. The framework of statistical mechanics is laid out by the following relation, deduced from eqs (4), (7)–(9) and which links the free energy with the partition function:

$$F = -kT \log Z. \tag{10}$$

*K Mallick*



**Figure 5.** Ludwig Boltzmann (1844–1906). The celebrated formula for the entropy is written on Boltzmann's grave in Vienna.

Similarly, the probabilistic expression of the entropy for a system at temperature $T$ is given by

$$S = -k \sum_{\mathcal{C}} P_{\text{eq}}(\mathcal{C}) \log P_{\text{eq}}(\mathcal{C}) . \qquad (11)$$

The precise logical order under which statistical mechanics is presented depends on the books and the lecturer's tastes. Some relations can be taken as fundamental and others as mathematical consequences or vice-versa. The important fact is that statistical mechanics gives us a systematic procedure to analyse systems at thermal equilibrium:

- Describe the microstates of the system and find a suitable microscopic Hamiltonian.
- Calculate $Z$ and deduce the free energy $F$.
- Derive from $F$ the thermodynamic properties of the system such as its phase diagram.

Of course, applying this well-defined program to a given problem can be incredibly difficult. Nobody knows how to calculate $Z$ for the three-dimensional Ising model.

2.3 *Near equilibrium: Brownian motion and linear response*

Equilibrium is a dynamical concept: a system in thermal equilibrium keeps on evolving from one microstate to another even if it appears to our imperfect senses to remain in the same macrostate. Thermodynamics deals only with averaged values: it cannot account for microscopic fluctuations. Though these fluctuations
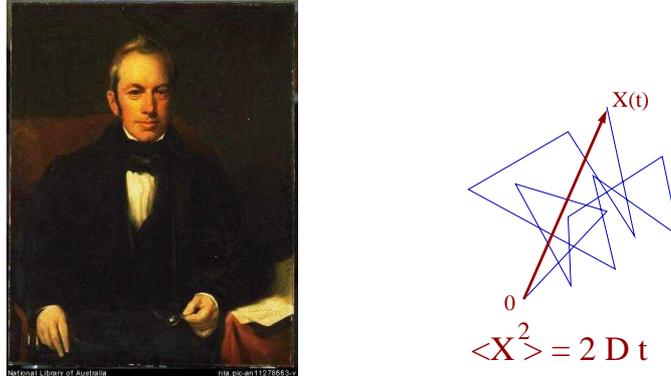
**Figure 6.** Robert Brown (1773–1858) and a sketch of Brownian motion.

are usually very minute (of relative order of $10^{-11}$ for a system containing one mole of matter) they can be detected either by using measuring devices which are becoming finer and finer or by studying very small systems. Statistical mechanics allows us to calculate the probability distributions of observables (and not only their averages) and perfectly describes the thermal fluctuations.

The paradigm for thermal fluctuations is Brownian motion discovered by Robert Brown (figure 6) who observed through a microscope, the perpetual, restless, giggling of a pollen grain in water. This phenomenon is the signature, at our scale, of the granular, discontinuous, structure of matter. It is the experimental footprint of the existence of atoms.

The theory of Brownian motion was elaborated by Albert Einstein, in 1905. The Brownian particle (for example, a grain of pollen) is restlessly shaken by random shocks with the molecules of water. Because of these shocks, the pollen grain undergoes an erratic motion and diffuses with time around its original position: although the average position of the Brownian particle does not change with time (just because of the isotropy of space) the quadratic average (i.e., the variance) of the position grows linearly with time:

$$\langle X^2(t)\rangle = 2Dt. \tag{12}$$

For a spherical particle of radius $a$, immersed in a liquid of viscosity $\eta$ at (absolute) temperature $T$, the diffusion constant $D$ is, given according to Einstein, by

$$D = \frac{RT}{6\pi\eta a\mathcal{N}}, \tag{13}$$

where $\mathcal{N} \simeq 6.02 \times 10^{23}$ is the Avogadro number and $R \simeq 8.31$ J·K$^{-1}$ is the perfect gas constant. This formula is extraordinary in the sense that it relates observables $D$, $T$, $\eta$ and $a$, which are all macroscopic to the number $\mathcal{N}$ of atoms in a mole of matter. This relation allowed Jean Perrin to weigh experimentally an atom of hydrogen (as he himself stated in his book 'The Atoms'); indeed $1/\mathcal{N}$ is roughly equal to the mass of one atom of hydrogen in grams. In his experiments, Perrin used small latex spheres with $a \sim 0.1$ $\mu$m, immersed in water ($\eta = 10^{-3}$ kg$^{-1}$ s$^{-1}$)

at temperature $T = 300$ K. The typical value of $D$ is then $10^{-12}$ m$^2$/s, i.e., the Brownian particles diffuse about one micrometre in one second. All these values, though not strictly macroscopic, were observable with an optical microscope at the beginning of the twentieth century and are much larger than the atomic scale. The theory of Brownian motion and its experimental verification established beyond any doubt the existence of atoms, considered previously to be a mere hypothesis.

Einstein's formula (13) can be interpreted as the simplest manifestation of the fluctuation–dissipation relation: consider that the pollen grain of size $a$, immersed in water, is subjected to a small drag force $f_{\text{ext}}$ (suppose for example that it is being pulled by an external operator). Because of this force, the pollen acquires a velocity $v$, and is subjected to a frictional force $-\gamma v$ because of the viscosity $\eta$ of the surrounding water. The friction coefficient $\gamma$ was calculated by Stokes at linear order in the velocity and it is given by $\gamma = 6\pi\eta a$ (assuming the pollen to be a perfect sphere). Balancing the drag force with the frictional force leads to the limiting speed:

$$v_\infty = \sigma f_{\text{ext}} \quad \text{with } \sigma = \frac{1}{6\pi\eta a} \,. \tag{14}$$

The susceptibility $\sigma$ measures the linear response to the external drive $f_{\text{ext}}$. Using this concept of susceptibility, Einstein's relation can be rewritten as

$$D = kT\sigma, \tag{15}$$

where $k = R/\mathcal{N}$ is the Boltzmann's constant. In other words, fluctuations at equilibrium, quantified by $D$, are proportional to the susceptibility $\sigma$ which quantifies the linear response to a small external perturbation that drives the system out of equilibrium. This relation

$$\text{Susceptibility (linear response)} \propto \text{Equilibrium fluctuations}$$

can be further generalized to the dynamic response of a system to a small time-varying external drive (this is known as Kubo formula).

There are many good books and articles on Brownian motion and linear response. Some useful references are [5–8].

### 2.4 *Feynman's ratchet and pawl*

One reason why Brownian motion was so troublesome to the 19th century physicists was (apart from the fact that they could not find a suitable explanation for it) that the pollen grain was undergoing a kind of perpetual motion even while remaining in contact with a single heat source (the water bath). Moreover, one could conceive a Gedanken experiment in which this perpetual motion could be coupled to a mechanical rectifier such as a wheel allowed to rotate only in one direction. Thus, when the Brownian particle would move in one direction, say eastwards, the wheel would rotate whereas it would stay still if the particle moved westwards. This is in essence the celebrated ratchet and pawl model discussed by Feynman in Chap. 46 of his *Lectures on Physics*, Volume 1 [9]. Thus, the second principle would be

in trouble because this rectified motion of the wheel could be used to extract some work from a single heat source.

In order for the pollen grain to cause rotation of the wheel in the Gedanken experiment at a perceptible rate, this wheel must be very small. However, all bodies are subjected to thermal fluctuations which typically are inversely proportional to their size. This universal character of thermal fluctuations leads to the resolution of the paradox: the one-way wheel is also subjected to intrinsic thermal fluctuations which cause it to move in the forbidden direction. A precise calculation, see e.g. [10], shows that the two effects (the rotation of the wheel by the Brownian particle vs. the spontaneous motion in the forbidden direction) perfectly compensate each other and no net rotation of the wheel occurs: 'the second law is saved'.

This concept of rectification of thermal fluctuations will be useful in non-equilibrium situations and will provide us a basic model for molecular motors in biological cells which are to be described later.

## 3. Mathematical description of non-equilibrium dynamics

### 3.1 *Markovian dynamics*

An efficient way to describe systems out of equilibrium is to use a probabilistic approach that goes back to Einstein's 1905 paper and to Smoluchowsky's work at the same period. The idea is to write an evolution equation for the probability $P_t(\mathcal{C})$ for the system to be in the microstate (or configuration) $\mathcal{C}$ at time $t$. In order to achieve such a description, one has to:

(1) Enumerate the microstates $\{\mathcal{C}_1, \mathcal{C}_2, \ldots\}$ of the system. These microstates can form a discrete or a continuous set depending on the problem studied.

(2) Specify the transition rates between two configurations. An important and common assumption is that these rates do not depend on the previous history of the system but only on the configuration $\mathcal{C}$ at time $t$ and on the target configuration $\mathcal{C}'$ at time $t + \mathrm{d}t$: this is the Markovian hypothesis which amounts to neglecting short time correlations. Thus, one must specify which transitions $\mathcal{C} \to \mathcal{C}'$ are allowed between time $t$ and $t + \mathrm{d}t$ and give their probability $M(\mathcal{C}', \mathcal{C})\mathrm{d}t$.

The evolution of the system can thus be viewed as a path on the configurations network (figure 7) and the time evolution of $P_t(\mathcal{C})$ is obtained by writing a balance equation in this network between the incoming flux and the outgoing flux at each configuration:

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t(\mathcal{C}) = \sum_{\mathcal{C}' \neq \mathcal{C}} M(\mathcal{C}, \mathcal{C}') P_t(\mathcal{C}') - \left\{ \sum_{\mathcal{C}' \neq \mathcal{C}} M(\mathcal{C}', \mathcal{C}) \right\} P_t(\mathcal{C}). \tag{16}$$

This rather simple looking equation is nothing but the master equation for non-equilibrium statistical mechanics which plays, in this field, a role analogous to Newton's equation in mechanics or Schrödinger's equation in quantum mechanics. The master (or Markov) equation (16) is a linear equation in the probability vector $P_t$ whose components are given by $\{P_t(\mathcal{C})\}$. It is more convenient to write it in a matrix form
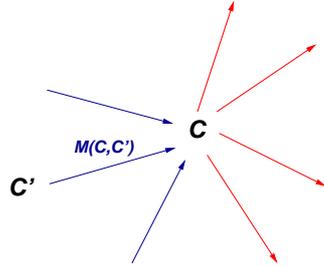
**Figure 7.** A network representation of the transition rates in a Markovian dynamics.

$$\frac{\mathrm{d}P_t}{\mathrm{d}t} = M \cdot P_t, \tag{17}$$

where the Markov operator $M$ has matrix elements given by $M(\mathcal{C}, \mathcal{C}')$ for $\mathcal{C} \neq \mathcal{C}'$ and its diagonal terms are defined as

$$M(\mathcal{C}, \mathcal{C}) = - \sum_{\mathcal{C}' \neq \mathcal{C}} M(\mathcal{C}', \mathcal{C}) \,. \tag{18}$$

The Markov operator $M$ fully encodes the microscopic dynamics of the system in this probabilistic approach which already implies a level of coarse-graining. Many physical properties of non-equilibrium systems are obtained by investigating the characteristics of $M$ such as its symmetries or the various types it can assume etc. Markov operators have the following basic properties:

(i) Non-diagonal terms are non-negative.

(ii) Diagonal terms are negative.

(iii) Equation (18) implies that the sum of the elements of $M$ over any column vanishes. This in turn implies the conservation of total probability, i.e., $\sum_{\mathcal{C}} P_t(\mathcal{C})$ is constant.

(iv) A vector $P_\infty$ in kernel of $M$ is a stationary state, i.e.,

$$\frac{\mathrm{d}P_\infty}{\mathrm{d}t} = 0 \,. \tag{19}$$

Under rather general assumptions, it is possible to prove the Perron–Frobenius theorem [7] which ensures that: (i) the kernel of $M$ is non-degenerate and therefore the stationary state is unique; (ii) all other eigenvalues of $M$ have strictly negative real parts; the inverse of these real parts correspond to the intrinsic relaxation times of the system towards its stationary state (the imaginary parts characterize the oscillations during relaxation).

*Remark.* We shall limit our discussion to the Markovian case. The analysis of non-Markovian systems is an active field which requires specific techniques. It often happens, however, that a non-Markovian problem can be imbedded into a larger Markovian system by taking into account supplementary degrees of freedom.
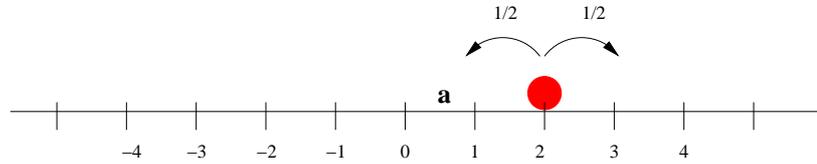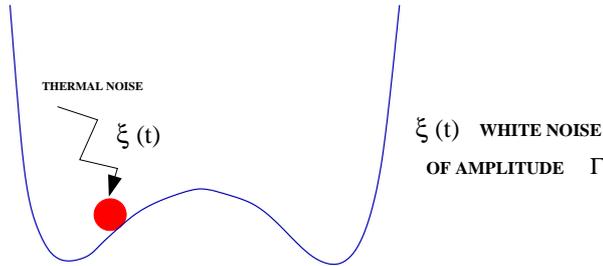
**Figure 8.** A one-dimensional random walk.



**Figure 9.** A Langevin particle in a double-well potential subject to a random thermal noise of amplitude $\Gamma$.

### 3.2 *Two examples*

In this subsection, we recall two archetypical models of Markovian dynamics: the random walk on a discrete lattice and the Langevin equation describing a particle in a confining potential and subject to a random thermal noise.

(i) The discrete random walk (figure 8):

The random walker, the paradigm of statistical physics, is used to explain diffusive behaviour (a classic reference is [6]). We consider a random walker on a one-dimensional discrete lattice with spacing $a$ where the walker can jump from a site to its two neighbours with probability $\mathrm{d}t$ between time $t$ and $t+\mathrm{d}t$. The master equation for the random walker is given by

$$\frac{\mathrm{d}P_t(na)}{\mathrm{d}t} = P_t((n+1)a) + P_t((n-1)a) - 2P_t(na), \qquad (20)$$

and the corresponding Markov operator is nothing but the discrete Laplacian:

$$M = \begin{pmatrix} \ddots & \ddots & \ddots & \ddots & & \\ 0 & 1 & -2 & 1 & 0 & \\ 0 & 0 & 1 & -2 & 1 & 0 \\ & 0 & 0 & 1 & -2 & 1 & \ddots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

(ii) A particle subject to a random force (figure (9):

In the second example (see [6,7]) we describe a particle at position $x$ in continuous space subject to a deterministic force that derives from a potential $\mathcal{U}(x)$, a frictional force proportional to its velocity, and a random force $\xi(t)$ that models thermal noise. Following Langevin, the motion of such a particle is described by Newton's equation with a stochastic contribution $\xi(t)$:

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\gamma\frac{\mathrm{d}x}{\mathrm{d}t} - \nabla\mathcal{U}(x) + \xi(t)\,. \qquad (21)$$

The crux of the matter lies in the correct description of the random noise $\xi(t)$. A usual model is to take $\xi$ to be a Gaussian white noise of amplitude $\Gamma$. Then, by definition, the random variable $(\xi(t)\sqrt{\mathrm{d}t})$ has a Gaussian probability distribution of variance $2\Gamma$. Besides, $\xi(t)$ and $\xi(t')$ are uncorrelated for $t \neq t'$. (This definition works well if time is discretized as in numerical simulations. A correct mathematical definition would require a non-trivial limiting procedure in order to explain properly what is meant by $\sqrt{\mathrm{d}t}$.)

The configuration space of the Langevin particle is the continuous phase space $(x, v)$, $x$ being the position and $v$ the velocity. The distribution $P_t(x, v)$, which represents the probability of being at $x$ with velocity $v$, satisfies an evolution equation that can be derived rigorously from the Langevin equation (21). The phase space being continuous, the master equation (16) takes the form of a partial differential equation known as the Fokker–Planck equation:

$$\frac{\mathrm{d}P_t}{\mathrm{d}t} = -\frac{\partial}{\partial x}\{vP_t\} + \frac{\partial}{\partial v}\left\{\frac{\gamma v + \nabla U}{m}P_t\right\} + \frac{\Gamma}{m^2}\frac{\partial^2 P_t}{\partial v^2} = \mathcal{L}_{\mathrm{FP}}\cdot P_t. \qquad (22)$$

The role of the Markov matrix $M$ is now played by the differential Fokker–Planck operator $\mathcal{L}_{\mathrm{FP}}$. It is important to remark that the amplitude $\Gamma$ of the Gaussian white noise $\xi$ appears as the coefficient of the second-order derivative. We also emphasize that Fokker–Planck equation is a deterministic equation (in contrast to the Langevin equation) because the noise has been integrated out. Here again the operator $\mathcal{L}_{\mathrm{FP}}$ fully encodes the dynamics of the system. From the mathematical point of view, however, things have become more difficult: working with differential operators requires functional analysis on infinite dimensional spaces. For instance, the statement of the Perron–Frobenius theorem requires more stringent assumptions and its proof is much more involved than in the case where the Markov matrix $M$ is finite-dimensional.

### 3.3 *Connection to thermodynamics*

In the probabilistic approach of non-equilibrium dynamics that we have described above, we made no mention of equilibrium statistical mechanics. Up till now, the discussion has been purely mathematical. However, this description has to be related to thermal physics. The connection with thermodynamics is obtained by imposing the condition that the stationary state of the Markovian dynamics $P_\infty$ (eq. (19)) be given by the Boltzmann–Gibbs canonical formula (8) which describes a system in thermal equilibrium:

$$P_\infty(\mathcal{C}) = P_{\text{eq}}(\mathcal{C}) = \frac{\mathrm{e}^{-E(\mathcal{C})/kT}}{Z} \, . \tag{23}$$

This matching condition with thermodynamics imposes a global constraint on the rates in the Markov operator:

$$\sum_{\mathcal{C}' \neq \mathcal{C}} M(\mathcal{C}, \mathcal{C}')\mathrm{e}^{-E(\mathcal{C}')/kT} = \mathrm{e}^{-E(\mathcal{C})/kT} \left\{ \sum_{\mathcal{C}' \neq \mathcal{C}} M(\mathcal{C}', \mathcal{C}) \right\} . \tag{24}$$

This condition implies that the system described by the stochastic dynamics reaches ultimately a state of thermodynamic equilibrium. For systems far from equilibrium with a non-thermodynamic stationary state, this relation does not hold.

A similar reasoning applies to systems with continuous states and in particular to stochastic systems described by a Langevin equation and the corresponding Fokker–Planck equation. Hence, a particle in a potential $\mathcal{U}$ and subject to a thermal noise at temperature $T$ will ultimately reach thermal equilibrium with canonical distribution given by

$$P_{\text{eq}}(x, v) = \frac{1}{Z}\mathrm{e}^{-\frac{1/2mv^2 + \mathcal{U}(x)}{kT}} \, . \tag{25}$$

This canonical Boltzmann–Gibbs law must be the stationary solution of the Fokker–Planck equation (22). The condition that the expression (25) is in the kernel of the Fokker–Planck operator $\mathcal{L}_{\text{FP}}$ fixes the value of noise-amplitude $\Gamma$ as a function of the temperature $T$ and the friction coefficient $\gamma$

$$\Gamma = \gamma kT \, . \tag{26}$$

Using this relation in the case of free Brownian motion (for which $\mathcal{U} = 0$), the Langevin and the Fokker–Planck equations reduce to

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\gamma v + \xi(t) \quad \text{and} \quad \frac{\mathrm{d}P_t}{\mathrm{d}t} = -\frac{\partial}{\partial x}\{vP_t\} + \frac{\gamma}{m}\frac{\partial}{\partial v}(vP_t) + \frac{\gamma kT}{m^2}\frac{\partial^2 P_t}{\partial v^2} \, . \tag{27}$$

The Fokker–Planck equation implies that the Brownian particle diffuses linearly with time and allows us to recover Einstein's relation (multiply each term of this equation by $x^2$ and integrate the RHS by parts; repeat the operation by multiplying by $xv$ and then by $v^2$. This leads to a closed system of three linear equations for the second moments $\langle x^2 \rangle$, $\langle v^2 \rangle$ and $\langle xv \rangle$). We thus obtain

$$\langle x^2 \rangle = 2Dt \quad \text{with} \quad D = \frac{kT}{\gamma} \quad \text{and} \quad \gamma = 6\pi\eta a. \tag{28}$$

This line of reasoning follows in spirit the Langevin's derivation of Einstein's relation.
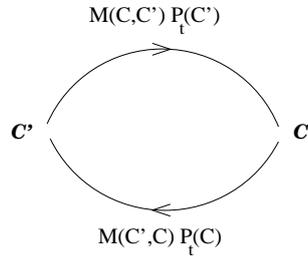
**Figure 10.** The local current between two configurations in a Markovian system.

3.4 *Time-reversal invariance and detailed balance*

The master equation (16) can be rewritten as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}P_t(\mathcal{C}) = \sum_{\mathcal{C}'} \{M(\mathcal{C},\mathcal{C}')\,P_t(\mathcal{C}') - M(\mathcal{C}',\mathcal{C})\,P_t(\mathcal{C})\} = \sum_{\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}')\,, \quad (29)$$

where $J_t(\mathcal{C},\mathcal{C}')$ represents the net local probability current flowing between the configurations $\mathcal{C}'$ and $\mathcal{C}$ at time $t$;

$$J_t(\mathcal{C},\mathcal{C}') = M(\mathcal{C},\mathcal{C}')\,P_t(\mathcal{C}') - M(\mathcal{C}',\mathcal{C})\,P_t(\mathcal{C})\,. \tag{30}$$

The local current encodes a balance of ingoing and outgoing currents between $\mathcal{C}'$ and $\mathcal{C}$ that is represented graphically in figure 10.

The local current is antisymmetric with respect to the configurations:

$$J_t(\mathcal{C},\mathcal{C}') = -J_t(\mathcal{C}',\mathcal{C})\,. \tag{31}$$

Besides, total probability conservation leads to

$$\sum_{\mathcal{C},\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}') = 0\,. \tag{32}$$

This relation is valid for any time $t$.

When $t \to \infty$, the system reaches a steady state with a stationary probability density. This implies that the total current flowing to $\mathcal{C}$ compensates exactly the total current flowing from $\mathcal{C}$:

$$\sum_{\mathcal{C}'} J_\infty(\mathcal{C},\mathcal{C}') = 0\,. \tag{33}$$

This global balance condition is nothing but a reformulation of the stationarity condition (19).

Taking into account the fact that the microscopic dynamics of the system which is represented by an effective Markovian model is Hamiltonian and that Hamiltonian dynamics is in general time-reversible, Onsager derived the following, much stronger, constraint:

$$M(\mathcal{C}, \mathcal{C}')P_{\text{eq}}(\mathcal{C}') = M(\mathcal{C}', \mathcal{C})P_{\text{eq}}(\mathcal{C}). \tag{34}$$

This remarkable relation, known as detailed balance, is a consequence of the time-reversal symmetry of the microscopic dynamics of the system. Detailed balance implies that the net local current $J_t(\mathcal{C}, \mathcal{C}')$ between two configurations $\mathcal{C}$ and $\mathcal{C}'$ vanishes at thermodynamic equilibrium. This very strong condition implies of course the vanishing of the global currents (eq. (33)). The detailed balance condition is a fundamental dynamic property of equilibrium systems that goes beyond the laws of classical thermodynamics. (Magnetic fields and overall rotation effects have been excluded from our discussion for the sake of simplicity. A modified version of detailed balance allows to take into account these effects; see [7] for a derivation.)

There is an elegant way to express detailed balance as a Hermiticity property of the Markov operator $M$. Indeed, detailed balance (34) can be rewritten as

$$P_{\text{eq}}(\mathcal{C})^{-1/2}M(\mathcal{C}, \mathcal{C}')P_{\text{eq}}(\mathcal{C}')^{1/2} = P_{\text{eq}}(\mathcal{C}')^{-1/2}M(\mathcal{C}', \mathcal{C})P_{\text{eq}}(\mathcal{C})^{1/2}. \tag{35}$$

Introducing now the diagonal matrix

$$Q = \begin{pmatrix} \ddots & & \\ & P_{\text{eq}}(\mathcal{C})^{1/2} & \\ & & \ddots \end{pmatrix}, \tag{36}$$

we observe that detailed balance implies that the operator $\mathcal{W} = Q^{-1}MQ$ is Hermitian:

$$\mathcal{W} = \mathcal{W}^{\dagger}. \tag{37}$$

A system in thermodynamic equilibrium satisfies detailed balance. Conversely, if the steady state of a system breaks the detailed balance condition then this steady state is necessarily a non-equilibrium stationary state (NESS). Hence, from a technical point of view, violating detailed balance ensures that the system is out of equilibrium. Moreover, the mathematical description of systems far from equilibrium typically involves non-Hermitian evolution operators.

### 3.5 *Physical interpretation of detailed balance*

We have said above that detailed balance was derived by Onsager as a consequence of the time reversibility of the underlying Hamiltonian dynamics. Here, we shall prove the converse of this assertion: the equilibrium state of a system which satisfies the detailed balance relation (34) is necessarily invariant by time reversal.

We first recall two properties of any Markovian dynamics which will be useful in the following:

(1) The probability of remaining in $\mathcal{C}$ during a time interval $\tau$ is given by

$$\lim_{\mathrm{d}t \to 0} (1 + M(\mathcal{C}, \mathcal{C})\mathrm{d}t)^{\tau/\mathrm{d}t} = e^{M(\mathcal{C}, \mathcal{C})\tau}. \tag{38}$$

(2) The probability of going from $\mathcal{C}$ to $\mathcal{C}'$ during $\mathrm{d}t$ is: $M(\mathcal{C}', \mathcal{C})\mathrm{d}t$.
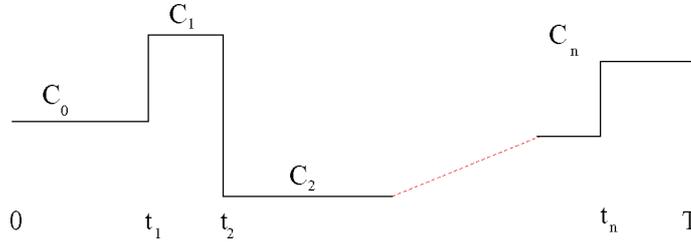
**Figure 11.** A typical trajectory with discrete jumps in a Markovian dynamics.

Let us study a system in its equilibrium state. As we have already emphasized, macroscopic equilibrium is in fact a dynamical state in which the system continuously evolves between microscopic configurations. The frequency with which a given configuration is visited is proportional to its stationary probability. We consider a trajectory $\mathcal{C}(t)$ of the system between time $t = 0$ and a final observation time $T$ (see figure 11). We suppose that at $t = 0$ the system is in configuration $\mathcal{C}_0$, that it jumps from $\mathcal{C}_0$ to $\mathcal{C}_1$ at time $t_1 > 0$ and remains in $\mathcal{C}_1$ till time $t_2$; at $t_2 > t_1$, it jumps from $\mathcal{C}_1$ to $\mathcal{C}_2$ and remains in $\mathcal{C}_2$ till $t_3$ etc. More generally, the system jumps from $\mathcal{C}_k$ to $\mathcal{C}_{k+1}$ at time $t_{k+1}$, for $k = 1, \ldots, n$. The final jump from configuration $\mathcal{C}_{n-1}$ to $\mathcal{C}_n$ occurs at $t_n$ and the system remains in $\mathcal{C}_n$ till the final time $T$. What is the probability $\Pr\{\mathcal{C}(t)\}$ of observing the trajectory $\mathcal{C}(t)$? Using recursively the two properties recalled above we obtain:

$$\Pr\{\mathcal{C}(t)\} = \mathrm{e}^{M(\mathcal{C}_n,\mathcal{C}_n)(T-t_n)} M(\mathcal{C}_n,\mathcal{C}_{n-1})\mathrm{d}t_n \, \mathrm{e}^{M(\mathcal{C}_{n-1},\mathcal{C}_{n-1})(t_n-t_{n-1})} \cdots$$
$$\mathrm{e}^{M(\mathcal{C}_2,\mathcal{C}_2)(t_3-t_2)} M(\mathcal{C}_2,\mathcal{C}_1)\mathrm{d}t_2 \, \mathrm{e}^{M(\mathcal{C}_1,\mathcal{C}_1)(t_2-t_1)}$$
$$\times M(\mathcal{C}_1,\mathcal{C}_0)\mathrm{d}t_1 \, \mathrm{e}^{M(\mathcal{C}_0,\mathcal{C}_0)t_1} P_{\mathrm{eq}}(\mathcal{C}_0). \tag{39}$$

We now calculate the probability of observing the time-reversed trajectory $\hat{\mathcal{C}}(t) = \mathcal{C}(T-t)$ (see figure 12). The system starts at $t = 0$ in configuration $\mathcal{C}_n$ and remains in that configuration till the time $T - t_n$ at which it jumps into $\mathcal{C}_{n-1}$. The next jump from $\mathcal{C}_{n-1}$ to $\mathcal{C}_{n-2}$ occurs at time $T - t_{n-1}$. More generally, the system jumps from $\mathcal{C}_k$ to $\mathcal{C}_{k-1}$ at time $T - t_k$, for $k = n, n-1, \ldots, 1$. At date $T - t_1$, the system reaches the configuration $\mathcal{C}_0$ and remains in it till the final time $T$. The probability of this trajectory is given by

$$\Pr\{\hat{\mathcal{C}}(t)\} = \mathrm{e}^{M(\mathcal{C}_0,\mathcal{C}_0)t_1} M(\mathcal{C}_0,\mathcal{C}_1)\mathrm{d}t_1 \, \mathrm{e}^{M(\mathcal{C}_1,\mathcal{C}_1)(t_2-t_1)} \cdots$$
$$\mathrm{e}^{M(\mathcal{C}_{n-1},\mathcal{C}_{n-1})(t_n-t_{n-1})} M(\mathcal{C}_{n-1},\mathcal{C}_n)\mathrm{d}t_n$$
$$\times \mathrm{e}^{M(\mathcal{C}_n,\mathcal{C}_n)(T-t_n)} P_{\mathrm{eq}}(\mathcal{C}_n). \tag{40}$$

The ratio of the probability of observing a given trajectory (39) to the probability of the time-reversed trajectory (40) is thus given by

$$\frac{\Pr\{\mathcal{C}(t)\}}{\Pr\{\hat{\mathcal{C}}(t)\}} = \frac{M(\mathcal{C}_n,\mathcal{C}_{n-1})M(\mathcal{C}_{n-1},\mathcal{C}_{n-2})\ldots M(\mathcal{C}_1,\mathcal{C}_0)\, P_{\mathrm{eq}}(\mathcal{C}_0)}{M(\mathcal{C}_0,\mathcal{C}_1)M(\mathcal{C}_1,\mathcal{C}_2)\ldots M(\mathcal{C}_{n-1},\mathcal{C}_n)\, P_{\mathrm{eq}}(\mathcal{C}_n)}. \tag{41}$$

If, in the numerator of this expression, we use recursively the detailed balance condition (34)
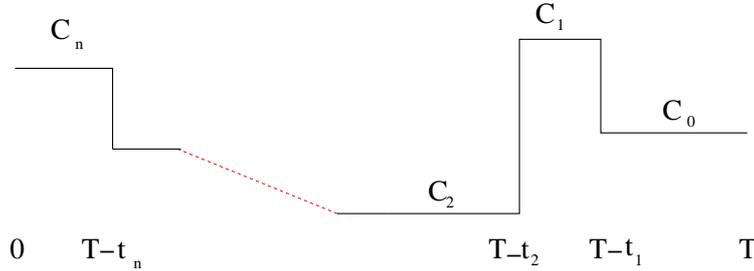
**Figure 12.** The time-reversed trajectory of the trajectory drawn in figure 11.

$$M(\mathcal{C}_{k+1},\mathcal{C}_k)P_{\text{eq}}(\mathcal{C}_k) = P_{\text{eq}}(\mathcal{C}_{k+1})M(\mathcal{C}_k,\mathcal{C}_{k+1}) \quad \text{for} \quad k = 0, 1, \ldots, n-1,$$

we find that

$$\frac{\Pr\{\mathcal{C}(t)\}}{\Pr\{\hat{\mathcal{C}}(t)\}} = 1 \,. \tag{42}$$

We have thus shown that detailed balance implies that the dynamics in the stationary state is time reversible.

### 3.6 *Entropy production in Markovian systems*

By analogy with eq. (11) which gives an expression of the entropy in the canonical ensemble of statistical mechanics, it is possible to define formally a time-dependent 'entropy' function for any Markovian system [7]:

$$S(t) = -\sum_{\mathcal{C}} P_t(\mathcal{C}) \log P_t(\mathcal{C}) \,. \tag{43}$$

Using the Markov equation (29) in terms of the local currents, the time derivative of this function is given by

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = -\sum_{\mathcal{C}} \frac{\mathrm{d}P_t(\mathcal{C})}{\mathrm{d}t}\left(\log P_t(\mathcal{C}) + 1\right) = -\sum_{\mathcal{C},\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}') \log P_t(\mathcal{C})$$

$$= \sum_{\mathcal{C},\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}') \log P_t(\mathcal{C}') \,, \tag{44}$$

where we have used the global conservation (32). The last equality is obtained by using the antisymmetry of the local currents (31) and by exchanging the role of the dummy variables $\mathcal{C}$ and $\mathcal{C}'$. The expression for the time derivative of $S(t)$ can be written in a more elegant manner by taking the half-sum of the last two equalities:

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = \frac{1}{2}\sum_{\mathcal{C},\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}') \log \frac{P_t(\mathcal{C}')}{P_t(\mathcal{C})} \,. \tag{45}$$
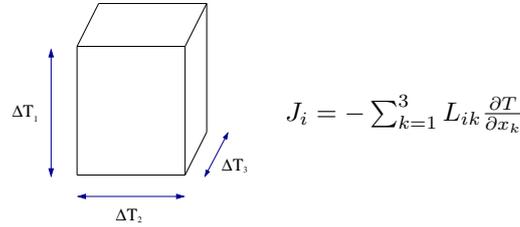
Transforming this expression, we obtain

$$J_i = -\sum_{k=1}^{3} L_{ik} \frac{\partial T}{\partial x_k}$$

**Figure 13.** Illustration of Onsager's relations for heat flow in an anisotropic crystal.

$$
\begin{aligned}
\frac{\mathrm{d}S(t)}{\mathrm{d}t} &= \frac{1}{2} \sum_{\mathcal{C},\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}') \left( \log \frac{P_t(\mathcal{C}')}{P_t(\mathcal{C})} + \log \frac{M(\mathcal{C},\mathcal{C}')}{M(\mathcal{C}',\mathcal{C})} - \log \frac{M(\mathcal{C},\mathcal{C}')}{M(\mathcal{C}',\mathcal{C})} \right) \\
&= \frac{1}{2} \sum_{\mathcal{C},\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}') \log \frac{M(\mathcal{C},\mathcal{C}')P_t(\mathcal{C}')}{M(\mathcal{C}',\mathcal{C})P_t(\mathcal{C})} \\
&\quad - \frac{1}{2} \sum_{\mathcal{C},\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}') \log \frac{M(\mathcal{C},\mathcal{C}')}{M(\mathcal{C}',\mathcal{C})} \\
&\equiv \frac{d_i S}{\mathrm{d}t} + \frac{d_e S}{\mathrm{d}t} ,
\end{aligned}
\tag{46}
$$

where the first expression in the last equality is called the entropy production term, also denoted by $\sigma_i$, and the second term the entropy flux. The entropy production can be proved, using convexity relations, to be always positive, $\sigma_i \geq 0$, whereas the entropy flux can be positive or negative. At equilibrium, $\sigma_i$ vanishes identically because of detailed balance (eq. (34)). In the vicinity of equilibrium the linear response theory can be used to show that $\sigma_i$ decreases towards 0 when the system relaxes to equilibrium. For a system in a non-equilibrium stationary state that does not satisfy detailed balance, $\sigma_i$ does not vanish but entropy production and entropy flux compensate each other exactly.

### 3.7 *Onsager's reciprocity relations*

Another fundamental consequence of time reversibility is the reciprocity relation also derived by Onsager (1931). We shall briefly present here a special case of these relations.

Consider an anisotropic crystal in the form of a rectangular parallelepiped and apply temperature gradients along the axes of this crystal (see figure 13). At linear order, the heat-current $J_i$ along the direction $i$ is given by

$$J_i = -\sum_{k=1}^{3} L_{ik} \frac{\partial T}{\partial x_k} \qquad \text{(Fourier's law)}.$$

The conductivity tensor **L** is always a symmetric tensor even when the symmetry of the crystal does not demand it:

$$L_{ik} = L_{ki}. \tag{47}$$

Such reciprocity relations are true for more general coupled irreversible processes. In particular, they allow to put on rigorous grounds the theory of thermoelectric effects that cannot be accounted for by classical thermodynamics alone [1].

## 4. Far from equilibrium: Fluctuation theorems and work identities

Systems in a non-equilibrium stationary state typically exhibit a non-vanishing macroscopic current $J$ (e.g., a current of particles, or a heat flux). Therefore, time reversal and detailed balance are broken in the mathematical description of the system. The stationary state, which is in the kernel of the Markov operator $M$, is in general not given by a Boltzmann–Gibbs law (which satisfies detailed balance). In fact, there is no general rule at present which would allow us to calculate the stationary state knowing the external constraints applied to the system. As opposed to the case of thermal equilibrium, there is no general theory of non-equilibrium statistical mechanics. One could imagine that by including the macroscopic currents along with the usual thermodynamic variables (such as temperature, density, pressure) one could construct some generalized potentials the minimization of which gives the stationary state. But we do not know if such potentials exist. Some general results on systems far from equilibrium have however been found in the last decade. We shall review some of these results in this section.

### 4.1 *Probabilities of rare events and large deviations*

The concept of a large-deviation function will be illustrated by the following example.

Let $\epsilon_1, \ldots, \epsilon_N$ be $N$ binary variables where $\epsilon_k = \pm 1$ with probability $1/2$ for $k = 1, \ldots, N$. Suppose that $\epsilon_k$'s are independent and identically distributed. Their sum is denoted by $S_N = \sum_1^N \epsilon_k$. We recall:

(1) The law of large numbers implies that $S_N/N \to 0$ (almost surely).
(2) The central limit theorem implies that $S_N/\sqrt{N}$ becomes a Gaussian variable of unit variance.

We now quantify the probability that $S_N/N$ takes a non-typical value $r$, with $-1 < r < 1$. One can show (using the Stirling formula) that in the large $N$ limit

$$\Pr\left(\frac{S_N}{N} = r\right) \sim \mathrm{e}^{-N\,\Phi(r)}, \tag{48}$$

where $\Phi(r)$ is given by

$$\Phi(r) = \frac{1+r}{2}\log\left(\frac{1+r}{2}\right) + \frac{1-r}{2}\log\left(\frac{1-r}{2}\right) + \log 2\,. \tag{49}$$

Mean density $\rho_0 = \frac{N}{V}$.
In the volume $1 \ll v \ll V$
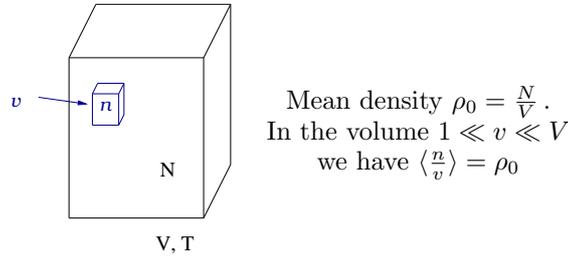we have $\langle \frac{n}{v} \rangle = \rho_0$

**Figure 14.** Illustration for the large deviation of local density in a box filled with a perfect gas.

The function $\Phi(r)$ is called a large deviation function. Because of the law of large numbers, we know that $\Pr(\frac{S_N}{N} = 0)$ tends to 1 when $N \to \infty$. The large deviation function must therefore vanish at $r = 0$, which is indeed the case.

Large deviations are a useful and well-known tool in probability. Their relation with thermodynamics can be clarified by the following example of density fluctuations in a gas [11]. Suppose that $N$ molecules of a perfect gas occupy a macroscopic vessel of volume $V$. The gas is at thermodynamic equilibrium at temperature $T$. The mean density is given by $\rho_0 = N/V$. Consider a smaller volume $v \ll V$ in the vessel, with $v$ large enough to contain a large number of particles (see figure 14). Let $n$ be the number of gas molecules in $v$: $n$ is a random variable that fluctuates with time (or from sample to sample if we adopt the canonical ensemble approach). Of course, we have $\langle n/v \rangle = \rho_0$: the local density fluctuates around $\rho_0$ with typical fluctuations of the order of $\sqrt{v/V}$. For a given $\rho$, the probability of observing a local density equal to $\rho$ in $v$ assumes the typical large deviation behaviour:

$$\Pr\left(\frac{n}{v} = \rho\right) \sim \mathrm{e}^{-v\,\Phi(\rho)} \quad \text{with} \quad \Phi(\rho_0) = 0. \tag{50}$$

This large deviation function can be calculated to be

$$\Phi(\rho) = f(\rho, T) - f(\rho_0, T) - (\rho - \rho_0)\frac{\partial f}{\partial \rho_0},$$

where $f = -\frac{1}{V} \log Z(V, N, T)$ is the free energy per unit volume (in units of $kT$). Here, the large deviation function that quantifies density fluctuations is nothing but the thermodynamic free energy (but for some simple terms). This example provides an interpretation of the free energy in purely probabilistic terms. It also suggests that large deviation functions should play an important role in statistical mechanics in general and may be good candidates for non-equilibrium 'potentials'.

### 4.2 *A paradigm for non-equilibrium physics: The asymmetric exclusion process*

An appropriate route to explore and discover the characteristics of a general theory for systems out of equilibrium is through the study of mathematical models. These models should be simple enough to allow exact mathematical analysis, but at the
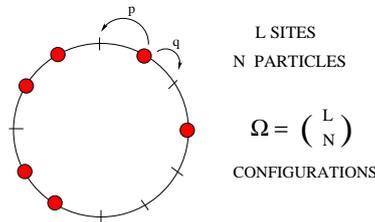
**Figure 15.** The asymmetric exclusion process on a periodic ring.

same time they must exhibit a rich phenomenology and must be versatile enough to be of physical significance. In the theory of phase transitions such a corner-stone was provided by the Ising model. It is natural to expect that a well-crafted dynamical version of the Ising model would play a key role in the development of non-equilibrium statistical mechanics. Driven lattice gases (for a review, see [12,13]) provide such models; they represent particles hopping on a lattice and interacting through hard-core exclusion. The particles are subjected to an external field that induces a bias in the hopping rates resulting in a stationary macroscopic current in the system. Due to this current, the microscopic detailed balance condition is violated and the system is generically in a non-equilibrium stationary state.

The asymmetric simple exclusion process (ASEP) in one dimension is a special case of a driven lattice gas. This stochastic particle system was introduced simultaneously as a biophysical model for protein synthesis on RNA and as a purely mathematical tool for the study of interaction of Markov processes in the mathematical literature. Subsequently, the ASEP has been used to study a wide range of physical phenomena: hopping conductivity in solid electrolytes, transport of macromolecules through thin vessels, reptation of polymer in a gel, surface growth, sequence alignment and molecular motors. Variants of ASEP has also been used as models for traffic flow that have been implemented in big cities such as Geneva.

4.2.1 *Basic properties of the ASEP*: We first describe the exclusion process on a periodic one-dimensional lattice with $L$ sites (sites $i$ and $L + i$ are identical). A lattice site cannot be occupied by more than one particle. The state of a site $i$ ($1 \leq i \leq L$) is characterized by a Boolean number $\tau_i = 0$ or 1 according as the site $i$ is empty or occupied. A configuration $\mathcal{C}$ can be represented by the sequence $(\tau_1, \tau_2, \ldots, \tau_L)$. The system evolves with time according to the following stochastic rule (see figure 15): a particle on a site $i$ at time $t$ jumps, in the interval between $t$ and $t + dt$, with probability $p \, dt$ to the neighbouring site $i + 1$ if this site is empty (exclusion rule) and with probability $q \, dt$ to the site $i - 1$ if this site is empty. The jump rates $p$ and $q$ are normalized so that $p + q = 1$. The dynamics conserves the total number of particles $N$. In the steady state on a periodic ring, all configurations of the ASEP are equiprobable: the stationary measure is flat. A particularly interesting case is the totally asymmetric exclusion process (TASEP), where the jumps are totally biased in one direction ($p = 1, q = 0$ or $p = 0, q = 1$).

Another interesting variant of the ASEP is the model with open boundaries which can be viewed as a driven lattice gas in contact with two reservoirs. In the bulk, the dynamics of the ASEP is the same as in the periodic case; at the boundaries
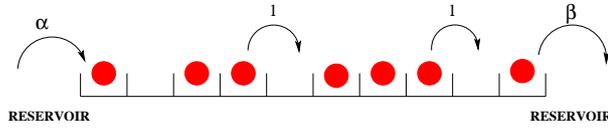
**Figure 16.** The totally asymmetric exclusion process in an open chain. If $L$ is the size of the system, there are $2^L$ different configurations.

particles can enter or leave the system with various input/output rates. The totally asymmetric version of this model is illustrated in figure 16: particles in the bulk jump to right with rate 1; a particle can be added with rate $\alpha$ if the site 1 is empty; a particle can be removed with rate $\beta$ if the site L is occupied. The open ASEP undergoes boundary-induced phase transitions: In the large system size limit, the expressions for the current and the density profile become non-analytic for certain values of the input/output rates. This leads to a phase diagram with three main regions: a low density phase (when typically the input rates are small and the output rates are large), a high density phase (when the input rates are large and the output rates small), and a maximal current phase (in which the bulk density is $1/2$ regardless of the boundary rates).

The ASEP is a kind of minimal model for non-equilibrium statistical mechanics. The exclusion constraint (at most 1 particle per site) plays the role of a hard-core repulsion; this coupling between particles turns the ASEP into a genuine non-trivial $N$-body problem. The asymmetry of the hopping rates mimics an external driving field and the input/output rates at the boundaries represent the chemical potentials of the reservoirs: both the external field and the reservoir drive a macroscopic current into the system, which manifests itself as a violation of detailed balance. The ASEP is a purely Markovian process (there is no underlying Hamiltonian): if we call $P_t(\mathcal{C})$ the probability of configuration $\mathcal{C}$ at time $t$ the time evolution of $P_t(\mathcal{C})$ is given by the master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}P_t(\mathcal{C}) = \sum_{\mathcal{C}'} M(\mathcal{C},\mathcal{C}')P_t(\mathcal{C}')$$
$$= \sum_{\mathcal{C}'} (M_0(\mathcal{C},\mathcal{C}') + M_1(\mathcal{C},\mathcal{C}') + M_{-1}(\mathcal{C},\mathcal{C}'))P_t(\mathcal{C}'). \tag{51}$$

The Markov matrix $M$ encodes the dynamics of the exclusion process: the non-diagonal element $M_1(\mathcal{C},\mathcal{C}')$ represents the transition rate from configuration $\mathcal{C}'$ to $\mathcal{C}$ when a particle hops in the forward (i.e., anti-clockwise) direction, and the non-diagonal element $M_{-1}(\mathcal{C},\mathcal{C}')$ represents the transition rate from configuration $\mathcal{C}'$ to $\mathcal{C}$ when a particle hops in the backward (i.e., clockwise) direction. The diagonal term $M_0(\mathcal{C},\mathcal{C}) = -\sum_{\mathcal{C}'\neq\mathcal{C}} (M_1(\mathcal{C}',\mathcal{C}) + M_{-1}(\mathcal{C}',\mathcal{C}))$ represents the exit rate from the configuration $\mathcal{C}$.

The asymmetric simple exclusion process is rich enough to exhibit a complex phenomenology but is simple enough to allow exact solutions in one dimension. In particular, many problems about the ASEP can be solved by Bethe ansatz (for reviews, see [11,14,15]). The Bethe ansatz, which plays a key role in the study of low-dimensional quantum systems, can be applied here, because ASEP is equivalent

to a non-Hermitian spin chain. For example, the Markov matrix of the ASEP on the periodic lattice of size $L$ is given by

$$M = \sum_{i=1}^{L} \left( pS_i^- S_{i+1}^+ + qS_i^+ S_{i+1}^- + \frac{1}{4} S_i^z S_{i+1}^z \right) - \frac{L}{4} , \tag{52}$$

where $S^+, S^-$ and $S^z$ are Pauli matrices (we recall that the site $L+1$ is the same as the site 1). The Markov matrix is therefore expressed as a spin chain on the periodic lattice. For $p = q = 1/2$, the Markov matrix $M$ is identical to the antiferromagnetic Heisenberg XXX spin chain which was exactly solved by Hans Bethe (1931). But for $p \neq q$, this spin chain is non-Hermitian: here again we observe that non-Hermiticity is related to non-equilibrium and to the breakdown of detailed balance. The Markov matrix has eigenvalues, $E$, such that $M\psi = E\psi$ which are usually complex numbers. The ground state corresponds to $E = 0$ and all excited states, that correspond to relaxation times, are such that $\Re(E) < 0$ (because of the Perron–Frobenius theorem).

4.2.2 *Large deviations of the current in ASEP*: The ASEP is an ideal tool for theorists to explain the concepts and to test general claims about non-equilibrium systems. It is one of the rare models on which analytical calculations can be performed. We shall use it here to illustrate the idea of large deviation function for the total current in the system and show that it is simply related to the generating function of cumulants of the current. (Besides, there is at present a large experimental interest in measuring higher cumulants of the current in low-dimensional quantum systems and in this context too, the ASEP appears to be a useful model.)

We consider the ASEP with $N$ particles on a ring of size $L$ and call $Y_t$ the total distance covered by all the particles on the ring between time 0 and $t$. What is the statistics of $Y_t$? The mean value of $Y_t$ is the average total current $J$, and the current fluctuations define a diffusion constant $D$. Both these quantities can be calculated exactly and in the case of the totally asymmetric exclusion process, $p = 1, q = 0$, the expressions obtained are rather simple.

$$J = \lim_{t \to \infty} \frac{\langle Y_t \rangle}{t} = \frac{N(L-N)}{L-1} \quad \text{(mean total current)}$$

$$D = \lim_{t \to \infty} \frac{\langle Y_t^2 \rangle - \langle Y_t \rangle^2}{t} = \frac{LN(L-N)}{(L-1)(2L-1)} \frac{C_{2L}^{2N}}{\left( C_L^N \right)^2} \quad \text{(fluctuations)}.$$

The full statistics of $Y_t$ is encoded in the generating function $\langle \mathrm{e}^{\gamma Y_t} \rangle$:

$$\frac{1}{t} \log \langle \mathrm{e}^{\gamma Y_t} \rangle = J\gamma + D\frac{\gamma^2}{2!} + C_3 \frac{\gamma^3}{3!} \dots , \tag{53}$$

and for $t \to \infty$, it can be shown that

$$\langle \mathrm{e}^{\gamma Y_t} \rangle \simeq \mathrm{e}^{\mathcal{E}(\gamma)t} , \tag{54}$$

where the function $\mathcal{E}(\gamma)$ can be expressed as the solution of an eigenvalue problem and can be calculated exactly by Bethe ansatz (for reviews, see [11,14,15]).

In the long time limit, $Y_t/t$ will converge to the average current $J$, but it is also interesting to know the distribution of $Y_t/t$. Again, this distribution assumes the generic large deviation form:

$$\Pr\left(\frac{Y_t}{t} = j\right) \sim e^{-t\Phi(j)} \quad \text{with } \Phi(j) = 0 \text{ for } j = J, \tag{55}$$

where $\Phi(j)$ is the large deviation function of the total current. From eqs (54) and (55), we obtain

$$\left\langle e^{\gamma Y_t}\right\rangle = \int \Pr(Y_t) e^{\gamma Y_t} dY_t = t \int \Pr\left(\frac{Y_t}{t} = j\right) e^{\gamma t j} dj \sim \int e^{\gamma t j - t\Phi(j)}. \tag{56}$$

Using the saddle-point method, we deduce that

$$\mathcal{E}(\gamma) = \max_j \left\{\gamma j - \Phi(j)\right\}. \tag{57}$$

We have thus shown that the large deviation function $\Phi(j)$ and the cumulant generating function $\mathcal{E}(\gamma)$ are Legendre transforms of each other. This relation is generally valid and will be used later in the derivation of the Gallavotti–Cohen theorem.

### 4.3 *Generalized detailed balance*

We have seen that the violation of detailed balance is the source of macroscopic currents which maintain the system far from equilibrium. In the ASEP case, this violation is due to two different factors: (i) the existence of an external driving force that pushes the particles in a given direction; (ii) the presence of reservoirs of unequal chemical potential (or temperature) that generates a current. The second case is particularly important to model the interaction of a system with its environment and the fluxes that are induced by this interaction. For a system connected to reservoirs, there often exists a relation which plays a role similar to that of detailed balance and implies some fundamental properties of the stationary state. This relation is called generalized detailed balance. We shall discuss it in the case of a discrete Markovian system [16] (such as ASEP) which can undergo an elementary transition between two configurations during the interval $(t, t + dt)$. We shall suppose that we are studying an observable $Y_t$ which varies by $y$ at each elementary transition (for example, for ASEP, $Y_t$ can be the total current). For each elementary transition, we can specify how $Y_t$ changes:

$$\mathcal{C} \to \mathcal{C}' \text{ and } Y_t \to Y_t + y \quad \text{with probability } M_y(\mathcal{C}', \mathcal{C}) dt. \tag{58}$$

By time reversal, the transition occurs from $\mathcal{C}' \to \mathcal{C}$. Assuming that $y$ is odd (i.e., it changes its sign), we have $Y_t \to Y_t - y$. Finally, we suppose that there exists a constant $\gamma_0$ such that transition rates satisfy the generalized detailed balance condition:
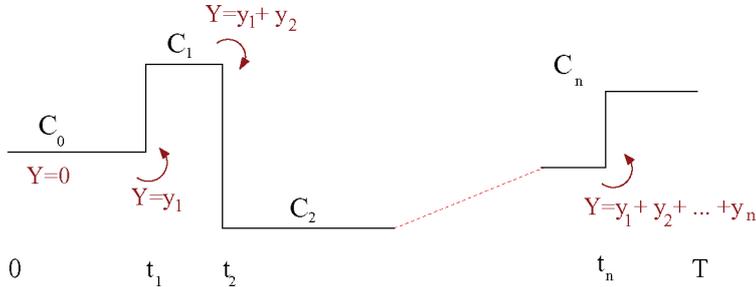
**Figure 17.** A trajectory in a Markovian system taking into account the variation of the observable $Y_t$ at each jump.

$$M_{+y}(\mathcal{C}', \mathcal{C}) P_{\text{stat}}(\mathcal{C}) = M_{-y}(\mathcal{C}, \mathcal{C}') \, \mathrm{e}^{\gamma_0 y} \, P_{\text{stat}}(\mathcal{C}'). \tag{59}$$

For $\gamma_0 = 0$, usual detailed balance is recovered. This relation holds, under general assumptions, for a system in contact with reservoirs that drive it out of equilibrium (in fact, it can be shown to be a consequence of usual detailed balance for the global model obtained by taking into account the system plus the reservoirs).

For example, consider ASEP on a ring: at each transition $\mathcal{C} \to \mathcal{C}'$ the total current $Y_t$ varies by $\pm 1$ as a particle jumps in the positive or negative direction, i.e., the increment $y$ satisfies $y = \pm 1$. A jump in the positive direction occurs with the probability $M_{+1}(\mathcal{C}', \mathcal{C}) = p$, whereas a jump in the negative direction occurs with probability $M_{-1}(\mathcal{C}', \mathcal{C}) = q$. We recall that the stationary probability is uniform: $P_{\text{stat}}(\mathcal{C}) = P_{\text{stat}}(\mathcal{C}') = \left( C_L^N \right)^{-1}$. Hence, we observe that generalized detailed balance holds in this case by so choosing $\gamma_0$ that $p = q \mathrm{e}^{\gamma_0}$, i.e., $\gamma_0 = \log(q/p)$. (But the usual detailed balance is not satisfied when $p \neq q$.)

### 4.4 *Generalized detailed balance, time reversal and the Gallavotti–Cohen symmetry*

We now investigate the relation between generalized detailed balance and time reversal [16]. We have to modify the calculations done in eqs (39)–(42) by taking into account all the factors of the type $\mathrm{e}^{\gamma_0 y}$ that appear at each jump in the ratio between the probabilities of forward and time-reversed trajectories.

Following the same steps as in §3.5, we finally obtain

$$\frac{\Pr\{\mathcal{C}(t)\}}{\Pr\{\hat{\mathcal{C}}(t)\}} = \mathrm{e}^{\gamma_0 Y\{\mathcal{C}(t)\}}, \tag{60}$$

where $Y\{\mathcal{C}(t)\} = y_1 + y_2 + \cdots + y_n$ is the total distance covered by all the particles when the system follows the trajectory $\mathcal{C}(t)$ between 0 and $t$.

We now recall that $Y$ is odd under time reversal and therefore we have $Y\{\hat{\mathcal{C}}(t)\} = -Y\{\mathcal{C}(t)\}$. Summing eq. (60) over all possible histories between time 0 and $t$, and taking $\gamma$ to be an arbitrary real number, we obtain

$$\sum_{\mathcal{C}(t)} \mathrm{e}^{(\gamma - \gamma_0) Y\{\mathcal{C}(t)\}} \Pr\{\mathcal{C}(t)\} = \sum_{\hat{\mathcal{C}}(t)} \mathrm{e}^{-\gamma Y\{\hat{\mathcal{C}}(t)\}} \Pr\{\hat{\mathcal{C}}(t)\}. \tag{61}$$

Because the relation between $\mathcal{C}(t)$ and $\hat{\mathcal{C}}(t)$ is one-to-one, we deduce that

$$\langle \mathrm{e}^{(\gamma-\gamma_0)Y_t} \rangle = \langle \mathrm{e}^{-\gamma Y_t} \rangle. \tag{62}$$

Using the fact that $\langle \mathrm{e}^{\gamma Y_t} \rangle \simeq \mathrm{e}^{\mathcal{E}(\gamma)t}$ (see eq. (54)), we obtain

$$\mathcal{E}(\gamma - \gamma_0) = \mathcal{E}(-\gamma). \tag{63}$$

By Legendre transform of this equation, the Gallavotti–Cohen fluctuation theorem is derived:

$$\Phi(j) = \Phi(-j) - \gamma_0 j. \tag{64}$$

From the definition of the large deviation function, the fluctuation theorem implies that in the long-time limit

$$\frac{\mathrm{Pr}\left(\frac{Y_t}{t} = j\right)}{\mathrm{Pr}\left(\frac{Y_t}{t} = -j\right)} \simeq \mathrm{e}^{\gamma_0 j t}. \tag{65}$$

This symmetry property of the large deviation function is valid far from equilibrium. This fact has been proved rigorously by various authors in many different contexts (chaotic systems, Markovian dynamics, Langevin dynamics etc.).

*Remark.* In the original works (see [16] and references therein), the authors studied the large deviation function for the entropy production $\sigma$. Recalling that the entropy flow is given by (46)

$$\frac{\mathrm{d}_e S}{\mathrm{d}t} = \frac{1}{2} \sum_{\mathcal{C},\mathcal{C}'} J_t(\mathcal{C},\mathcal{C}') \log \frac{M(\mathcal{C}',\mathcal{C})}{M(\mathcal{C},\mathcal{C}')},$$

one defines an entropy transfer for each jump as

$$y = \log \frac{M(\mathcal{C}',\mathcal{C})}{M(\mathcal{C},\mathcal{C}')}.$$

The increment of the entropy flow at each jump is thus given by $y$. A property similar to the generalized detailed balance is tautologically true for $y$:

$$M_y(\mathcal{C}',\mathcal{C}) = M_{-y}(\mathcal{C},\mathcal{C}')\mathrm{e}^{\gamma_0 y} \quad \text{with } \gamma_0 = 1.$$

This relation implies a fluctuation theorem, given by

$$\Phi(\sigma) - \Phi(-\sigma) = -\sigma,$$

where $\Phi(\sigma)$ is the large deviation function associated with entropy flow.

### 4.5 *Application to molecular motors*

In this section, we illustrate the Gallavotti–Cohen theorem by a model from cellular biophysics.
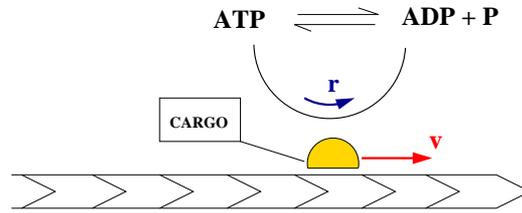
**Figure 18.** Schematic representation of a molecular motor: by hydrolysing ATP, the motor proceeds along the polar filament and carries a 'cargo' molecule.

A significant part of the eukaryotic cellular traffic relies on 'motor' proteins that move in a deterministic way along filaments similar in function to railway tracks or freeways (kinesins and dyneins move along tubulin filaments; myosins move along actin filaments). The filaments are periodic (of period ∼10 nm) and have a fairly rigid structure; they are also polar: a given motor always moves in the same direction. These motors appear in a variety of biological contexts: muscular contraction, cell division, cellular traffic, material transport along the axons of nerve cells etc.

Molecular motors move using the ratchet effect: they provide an example of rectification of Brownian motion (for reviews, see [17,18]). This rectification process relies on an external energy source, provided by ATP hydrolysis, that enables the motor to undergo transitions between different states and when these transitions break the detailed balance, a directed motion can set in (see figure 18) if the motor–filament interaction is asymmetric. In order to move, the motor consumes $r$ ATP fuel molecules per unit time, which are hydrolysed to ADP + P:

$$\mathrm{ATP} \rightleftharpoons \mathrm{ADP} + \mathrm{P}.$$

The relevant chemical potential is thus given by $\Delta\mu = \mu_{\mathrm{ATP}} - \mu_{\mathrm{ADP}} - \mu_{\mathrm{P}}$.

The principle of the motor is shown in figure 19 where the motor is represented by a small particle that can move in a one-dimensional space. At the initial time $t = 0$, the motor is trapped in one of the wells of a periodic asymmetric potential of period $a$. Between time 0 and $t_f$, the asymmetric potential is erased and the particle diffuses freely and isotropically at temperature $T$. At time $t_f$, the asymmetric potential is re-impressed, the motor slides down in the nearest potential valley and, because of damping, is trapped in one of the wells. The motor has maximal chance to end up in the same well where it was at time $t = 0$. However, it has a small probability to be trapped in the well located to the right and (because of the asymmetry of the potential) an even smaller probability to end up in the left well. In other words, because the potential is asymmetric, the motor has more chances to slide down towards the right: this leads on average to a net total current.

In general, the motor is subject to an external force $f_{\mathrm{ext}}$ which tilts the potential. Besides, when ATP is in excess, the chemical potential $\Delta\mu = \mu_{\mathrm{ATP}} - \mu_{\mathrm{ADP}} - \mu_{\mathrm{P}}$ becomes positive. A basic problem is then to determine the velocity of the motor $v(f_{\mathrm{ext}}, \Delta\mu)$ (mechanical current) and the ATP consumption rate $r(f_{\mathrm{ext}}, \Delta\mu)$
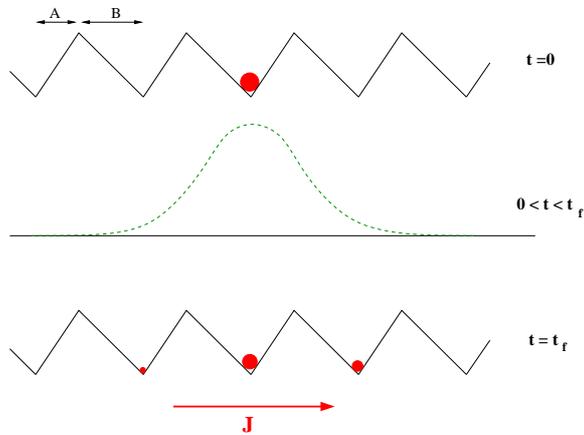
**Figure 19.** The principle of a Brownian ratchet: by inscribing and erasing periodically an asymmetric potential, a directed motion of the particle is induced. In this example, the potential is a saw-tooth function of period $a = A + B$. Since $B > A$, the slopes are different and the potential is asymmetric. The relative sizes of the probabilities of ending in one of the wells are represented by the sizes of the disks in the lowest picture. The right and left probabilities being different, this leads on average to a net total current $J$.

(chemical current) as functions of the external mechanical and chemical loads. In the following we shall write these external loads in dimensionless units. Using the period $a$ of the potential and the temperature $kT$, we identify $f_{\text{ext}}$ with $f_{\text{ext}}a/(kT)$, $\Delta\mu$ with $\Delta\mu/(kT)$, and the velocity $v$ with $v/a$.

We now describe various physical regimes that will illustrate the concepts presented in the previous sections, following the analysis of [19]:

(1) *At thermodynamic equilibrium:* We have $f_{\text{ext}} = 0$ and $\Delta\mu = 0$ and the two currents $v$ and $r$ vanish. The position of the motor and the number of ATP molecules exhibit Gaussian fluctuations, characterized by two diffusion constants $D_1$ and $D_2$.

(2) *Near equilibrium:* For small values of $f_{\text{ext}}$ and $\Delta\mu$, linear response theory holds and we can write (in dimensionless units)

$$v = L_{11}f_{\text{ext}} + L_{12}\Delta\mu,$$
$$r = L_{21}f_{\text{ext}} + L_{22}\Delta\mu.$$

Here, we can apply the classical results of Einstein and Onsager. Einstein's relations imply that the direct susceptibilities are identical to the diffusion constants

$$L_{11} = D_1 \quad \text{and} \quad L_{22} = D_2.$$

Onsager's reciprocity relation implies that the cross-susceptibilities are equal:

$$L_{12} = L_{21}.$$

Both these relations can be verified on explicit cases.

(3) *Far from equilibrium:* $v$ and $r$ are non-linear functions of the external loads $f_{\text{ext}}$ and $\Delta\mu$. If $X_t$ is the total displacement of the motor and $Q_t$ the total number

of ATP molecules consumed between 0 and $t$, one can show that in the long time limit

$$\langle e^{\gamma X_t + \lambda Q_t} \rangle \simeq e^{\mathcal{E}(\gamma, \lambda)t}. \tag{66}$$

The calculation of $\mathcal{E}(\gamma, \lambda)$, which can be done for simple models, gives access to the complete statistics of the mechanical and chemical currents in the system when $t \to \infty$. In particular, we have

$$v = \frac{\partial \mathcal{E}}{\partial \gamma}(0,0) \quad \text{and} \quad r = \frac{\partial \mathcal{E}}{\partial \lambda}(0,0). \tag{67}$$

The function $\mathcal{E}(\gamma, \lambda)$ satisfies the Gallavotti–Cohen symmetry:

$$\mathcal{E}(\gamma, \lambda) = \mathcal{E}(-\gamma - f_{\text{ext}}, -\lambda - \Delta\mu). \tag{68}$$

Using eq. (67), we find that

$$v = -\frac{\partial \mathcal{E}}{\partial \gamma}(-f_{\text{ext}}, -\Delta\mu) \quad \text{and} \quad r = -\frac{\partial \mathcal{E}}{\partial \lambda}(-f_{\text{ext}}, -\Delta\mu). \tag{69}$$

These identities are valid arbitrarily far from equilibrium.

To end this section, we show explicitly on the present model that, near equilibrium, the Gallavotti–Cohen fluctuation theorem yields Einstein and Onsager's relations. Developing the last equation for small values of $f_{\text{ext}}$ and $\Delta\mu$ (i.e., near equilibrium) we find

$$v = \frac{\partial^2 \mathcal{E}}{\partial \gamma^2}\Big|_{(0,0)} f_{\text{ext}} + \frac{\partial^2 \mathcal{E}}{\partial \gamma \partial \lambda}\Big|_{(0,0)} \Delta\mu, \tag{70}$$

$$r = \frac{\partial^2 \mathcal{E}}{\partial \lambda \partial \gamma}\Big|_{(0,0)} f_{\text{ext}} + \frac{\partial^2 \mathcal{E}}{\partial \lambda^2}\Big|_{(0,0)} \Delta\mu. \tag{71}$$

Identifying the linear response coefficients, we obtain

$$L_{11} = \frac{\partial^2 \mathcal{E}}{\partial \gamma^2}\Big|_{(0,0)} = \frac{\langle X_t^2 \rangle - \langle X_t \rangle^2}{t} = D_1 \quad \text{(Einstein)},$$

$$L_{12} = \frac{\partial^2 \mathcal{E}}{\partial \gamma \partial \lambda}\Big|_{(0,0)} = \frac{\partial^2 \mathcal{E}}{\partial \lambda \partial \gamma}\Big|_{(0,0)} = L_{21} \quad \text{(Onsager)},$$

$$L_{22} = \frac{\partial^2 \mathcal{E}}{\partial \lambda^2}\Big|_{(0,0)} = \frac{\langle Q_t^2 \rangle - \langle Q_t \rangle^2}{t} = D_2 \quad \text{(Einstein)}.$$

The Gallavotti–Cohen fluctuation theorem is a very strong property of systems far from equilibrium that satisfy generalized detailed balance. In particular, this theorem contains the classical relations of Einstein and Onsager, which are valid only in the vicinity of equilibrium. But it contains more information (it is a functional identity) and remains true far from equilibrium.

4.6 *Jarzynski's work theorem*

In this last section, we review another remarkable recent result in non-equilibrium statistical physics.

When discussing classical thermodynamics, we recalled (see eq. (6)), that the work performed on a system in contact with a heat reservoir at temperature $T$ satisfies the relation

$$\langle W \rangle \geq F_{\mathrm{B}} - F_{\mathrm{A}} = \Delta F \,, \tag{72}$$

where $F_{\mathrm{A}}$ is the free energy of the initial state and $F_{\mathrm{B}}$ that of the final state. We point out that in classical thermodynamics the work performed really means an average over many experiments (e.g., an operator pulling a piston enclosing a perfect gas from volume $V_{\mathrm{A}}$ to $V_{\mathrm{B}}$ (see figure 4). To emphasize this fact we have rewritten here eq. (6) with the notation $\langle W \rangle$ instead of simply $W$.

A decade ago, Christopher Jarzynski found that this classical inequality, well-known since the 19th century, can be deduced from an underlying remarkable identity valid for non-equilibrium systems. In the beginning, this identity was proved only for Hamiltonian systems, but Jarzynski (and others) have been extending its validity to more and more cases (such as Markovian dynamics or Langevin systems) and have verified it on exactly solvable models (see for example [20] and references therein; for an elegant proof of the Jarzynski theorem for Langevin dynamics using path integrals, see [21]). Experimental results [22] have also confirmed the Jarzynski relation which is now firmly established and is considered to be one of the few exact results in non-equilibrium statistical mechanics.

Jarzynski's identity states that

$$\langle \mathrm{e}^{-W/kT} \rangle = \mathrm{e}^{-\Delta F/kT}. \tag{73}$$

The average in this equation is taken over a non-equilibrium ensemble of individual trajectories of duration $\tau$ such that:

(i) At $t = 0$, the system is in equilibrium in the state A.

(ii) Between 0 and $\tau$, the operator acts on the system by changing a control parameter $V(t)$ according to a fixed well-defined protocol which does not have to be quasi-static and which drives the system far from equilibrium.

(iii) At $\tau$, the control parameter is fixed to a value $V_{\mathrm{B}}$. The system is not at equilibrium.

During the whole process, the system remains in contact with a heat bath at temperature $T$. After an infinite time, it will reach the equilibrium state B.

We emphasize that Jarzynski's identity connects data related to a non-equilibrium process (the exponential work average on the left-hand side of the identity) with thermodynamics (the free energy on the right-hand side).

*Remarks*

(1) From convexity (Jensen's inequality), we have

$$\langle \mathrm{e}^{-W/kT} \rangle \geq \mathrm{e}^{-\langle W \rangle/kT}.$$

Hence, Jarzynski's work theorem yields the classical inequality for the maximum available work.

(2) However, in order to have an equality, there must be individual trajectories that do not satisfy the classical inequality (72), i.e., there must be some realizations for which

$$W < \Delta F.$$

Such special realizations are called 'transient violations of the second principle'. It must be emphasized that the second principle is not violated because the second principle concerns averages and states that the average of the performed work is greater than the free energy difference, which remains true. The second principle does not say anything about individual behaviour. However, in thermodynamics, we are so used to the fact that individual measurements usually reflect the typical average behaviour that we forget the fact that these two quantities can be different.

(3) These 'transient violations' of the second principle can be quantified thanks to an identity (due to G Crooks) which is even more precise than Jarzynski's relation. Let $V^{\mathcal{F}}(t)$ be a protocol of duration $\tau$ that drives the system from $V_A$ to $V_B$ and let $V^{\mathcal{R}}(t) = V^{\mathcal{F}}(\tau - t)$ be the time-reversed protocol. It is then possible to measure the work done during the forward process and the work done during the reversed process. These quantities are both random variables with probability distributions $P^{\mathcal{F}}$ and $P^{\mathcal{R}}$, respectively. The following identity is satisfied by these probability distributions (Crooks, 1999):

$$\frac{P^{\mathcal{F}}(W)}{P^{\mathcal{R}}(-W)} = \mathrm{e}^{(W-\Delta F)/kT}. \tag{74}$$

Note the similitude between this identity and the fluctuation theorem in the form given in eq. (65). We note that Jarzynski's relation is a direct consequence of this equation:

$$\langle \mathrm{e}^{-W/kT} \rangle = \int \mathrm{e}^{-W/kT} P^{\mathcal{F}}(W) \mathrm{d}W$$

$$= \mathrm{e}^{-\Delta F/kT} \int P^{\mathcal{R}}(-W) \mathrm{d}W = \mathrm{e}^{-\Delta F/kT},$$

where, to derive the last equality we have used the fact that $P^{\mathcal{R}}$ is a normalized probability distribution.

We can also calculate the order of magnitude of the probability of a transient violation of amplitude $\zeta > 0$:

$$\mathrm{Prob}^{\mathcal{F}}(W < \Delta F - \zeta) = \int_{-\infty}^{\Delta F - \zeta} P^{\mathcal{F}}(W) \mathrm{d}W$$

$$= \int_{-\infty}^{\Delta F - \zeta} P^{\mathcal{R}}(-W) \mathrm{e}^{(W-\Delta F)/kT} \mathrm{d}W$$

$$= \int_{-\infty}^{0} P^{\mathcal{R}}(\zeta - \Delta F - v) \mathrm{e}^{v/kT} \mathrm{e}^{-\zeta/kT} \mathrm{d}v$$

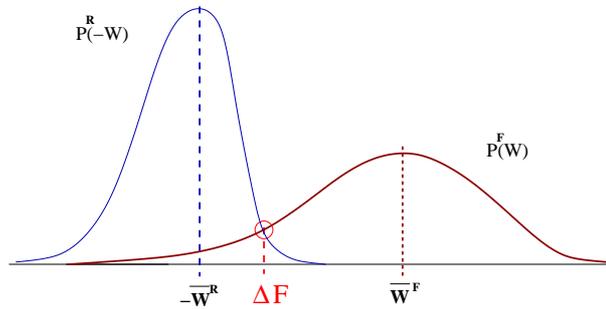$$\leq \mathrm{e}^{-\zeta/kT}. \tag{75}$$

**Figure 20.** Graphical representation of the Crooks' relation.

To derive the last inequality, we have used the fact that $e^{v/kT} \leq 1$ for $v \leq 0$ and also that $\int_{-\infty}^{0} P^{\mathcal{R}} dv \leq 1$. The probability of a violation of amplitude $\zeta > 0$ is exponentially small with $\zeta$, but again such violations are necessary to ensure the validity of Crooks' and Jarzynski's relations. We observe that for a transient violation to have non-vanishing probability, $\zeta$ must be of the order of $kT$. On the other hand, $\Delta F \sim NkT$, where $N$ is the number of degrees of freedom in the system, which usually are of the order of the Avogadro number. Therefore, transient violations are one part in $10^{23}$ for macroscopic systems: this is totally unobservable. One has to work with very small systems, such as biophysical objects, to have the chance to observe anything.

(4) Jarzynski and Crooks' identities allow to determine equilibrium free energy differences by doing non-equilibrium experiments. Many experimental results have been obtained using single-molecule manipulations [22]. The idea behind these experiments is shown in figure 20, which represents graphically the Crooks' relation (74). The forward work distribution and the distribution of minus the reversed work cross each other at $\Delta F$. These distributions can be measured and plotted using non-equilibrium experimental conditions. This provides a method for determining free energy variations, which characterize equilibrium states, through non-equilibrium data.

## 5. Conclusion

The field of non-equilibrium statistical mechanics has been undergoing fast progress in the last years. This subject has benefited from the cross-fertilization of diverse branches of physics and mathematics such as biophysics, dynamical systems, probability theory, integrable systems and non-Hermitian quantum mechanics. Nevertheless, a global theory, akin to thermodynamics, that would describe general non-equilibrium behaviour still remains to be found. My aim in this short review is to try to connect recent results with well-known knowledge. I also chose to focus on some simple examples that (hopefully) contain some universal features, rather than discussing the general framework. I hope that the reader will find here an incentive to work through more technical papers and to contribute to the subject.

## Acknowledgments

## References

[1] M W Zemansky and R H Dittman, *Heat and thermodynamics* (McGraw Hill Book Company, 1981)

[2] E Fermi, *Thermodynamics* (Dover, New York, 1956)

[3] A B Pippard, *The elements of classical thermodynamics* (Cambridge University Press, 1964)

[4] D Lemons, *Mere thermodynamics* (The John Hopkins University Press, Baltimore, 2009)

[5] F Reif, *Statistical physics: Berkeley physics course* (McGraw-Hill Book Company, 1967) Vol. 5

[6] S Chandrasekhar, *Rev. Mod. Phys.* **15**, 1 (1943)

[7] N Van Kampen, *Stochastic processes in physics and chemistry* (North Holland, Amsterdam, 1992)

[8] B Duplantier, *Brownian motion, 'diverse and undulating'*, arXiv:0705.1951

[9] *Feynman's Lectures in Physics*, Vol. 1, Chapter 46

[10] C Jarzynski and O Mazonka, *Phys. Rev.* **E59**, 6448 (1999)

[11] B Derrida, *J. Stat. Mech.* P07023 (2007)

[12] H Spohn, *Large scale dynamics of interacting particles* (Springer, Berlin, 1991)

[13] B Schmittmann and R K P Zia, Statistical mechanics of driven diffusive systems, in: *Phase transitions and critical phenomena* edited by C Domb and J L Lebowitz (Academic Press, San Diego, 1995) Vol. 17

[14] T Halpin-Healy and Y-C Zhang, *Phys. Rep.* **254**, 215 (1995)

[15] O Golinelli and K Mallick, *J. Phys.* **A39**, 12679 (2006)

[16] J L Lebowitz and H Spohn, *J. Stat. Phys.* **95**, 333 (1999)

[17] F Jüllicher, A Ajdari and J Prost, *Rev. Mod. Phys.* **69**, 1269 (1997)

[18] R Dean Astumian and P Hänggi, *Phys. Today* (Nov. 2002) p. 33

[19] A Lau, D Lacoste and K Mallick, *Phys. Rev. Lett.* **99**, 158102 (2007); *Phys. Rev.* **E78**, 011915 (2008)

[20] C Jarzynski, *Eur. Phys. J.* **B64**, 331 (2008)

[21] O Narayan and A Dhar, *J. Phys.* **A37**, 63 (2004)

[22] C Bustamante, J Liphardt and F Ritort, *Phys. Today* (July 2005) p. 43