1
Large deviation approach to nonequilibrium systems

Hugo Touchette and Rosemary J. Harris

Abstract. The theory of large deviations has been applied successfully in the last 30 years or so to study the properties of equilibrium systems and to put the foundations of equilibrium statistical mechanics on a clearer and more rigorous footing. A similar approach has been followed more recently for nonequilibrium systems, especially in the context of interacting particle systems. We review here the basis of this approach, emphasizing the similarities and differences that exist between the application of large deviation theory for studying equilibrium systems on the one hand and nonequilibrium systems on the other. Of particular importance are the notions of macroscopic, hydrodynamic, and long-time limits, which are analogues of the equilibrium thermodynamic limit, and the notion of statistical ensembles which can be generalized to nonequilibrium systems. For the purpose of illustrating our discussion, we focus on applications to Markov processes, in particular to simple random walks.

1.1
Introduction

Nonequilibrium systems are being increasingly studied using methods borrowed from the mathematical theory of large deviations, as developed in the 60s and 70s by Donsker and Varadhan, and Freidlin and Wentzell (see [1–3] for historical references). Indeed the central concepts and quantities of this theory – e.g., the large deviation principle, rate functions, generating functions, etc. – have now entered the standard jargon of driven nonequilibrium systems modelled as discrete- or continuous-time Markov processes (see, e.g., [3–5]).

With hindsight, one can argue that this evolution, although relatively recent, was to be expected: large deviation theory has been used successfully in equilibrium statistical mechanics for well over 30 years [3, 6–8], and so it is not surprising that this success finds its way into nonequilibrium statistical mechanics. However, there is more, in that the two scenarios – equilibrium and nonequilibrium – share many ideas, concepts and even a theoretical structure which happen to find a clear and precise expression in the language of large deviations. It is natural, therefore, to see this language being used for both theories.

By viewing equilibrium statistical mechanics from the point of view of large deviation theory, one gets a clear sense, for example, of why there is a Legendre

transform in thermodynamics connecting the entropy and the free energy, when this Legendre transform is valid, how equilibrium states relate to the notions of concentration and typicality, and how these states arise out of variational principles such as the maximum entropy principle or the minimum free energy principle. Similar ideas and results of nonequilibrium statistical mechanics are also made clear by viewing them through the prism of large deviation theory. In addition, one sees, as mentioned, essential similarities between equilibrium and nonequilibrium.

Our goal in this chapter is to explain these points and to illustrate them with simple examples, mainly involving continuous-time Markov processes. We start in the next section by recalling the basis and essential concepts of equilibrium statistical mechanics, and by discussing their analogues in nonequilibrium statistical mechanics. Among these, we mention the notions of statistical ensemble, stationarity, typicality, fluctuations, and scaling limit (e.g., thermodynamic limit, hydrodynamic or macroscopic limit, long-time limit). In Sec. 1.3, we re-express these concepts in the language of large deviations to define them in a precise, mathematical way and to emphasize the theoretical structure that underlies both equilibrium and nonequilibrium statistical mechanics. We illustrate this structure with a variety of applications in Sec. 1.4 and then end in Sec. 1.5 with some concluding remarks and open problems.

Much of the large deviation content explored in the present contribution can be found with more details in [3]. Here we focus on discussing the common goals, concepts, and results of equilibrium and nonequilibrium statistical mechanics, rather than providing a complete review of either subject, and on proposing a clear approach to studying nonequilibrium systems which parallels that used for studying equilibrium systems. We draw inspiration in doing so from the works of Oono [9–11], Eyink [12–14], and Maes *et al.* [15–18], among others, which show the emergence of similar ideas and views as early as the late 80s.

1.2 From equilibrium to nonequilibrium systems

Before we discuss how large deviation concepts enter in equilibrium and nonequilibrium statistical mechanics, we recall in this section the basis of each theory and emphasize some concepts shared by both. Of these, the most important to keep in mind is the concept of *typicality*, connected mathematically to the Law of Large Numbers and the concentration of probability distributions.

1.2.1 Equilibrium systems

The goal of *equilibrium* statistical mechanics, as is well known, is to explain and predict the emergence of macroscopic equilibrium states of systems composed of many particles by treating their microscopic states in a probabilistic way. The main properties of equilibrium states are that they are stationary in time, they are stable

against small perturbations, and are described by only a few variables, i.e., they are low-dimensional *macro*states compared to the high-dimensional *micro*states used for describing the many-particle system at the microscopic level.

Equilibrium states are also defined with respect to a given macrostate. Physicists often say that a system is *at equilibrium* or is *in a state of equilibrium*, but what is meant, to be more precise, is that the system has, say, an equilibrium energy or an equilibrium magnetization. Thus an equilibrium state is a particular state or a value of a macrostate or a collection of macrostates. This is different from saying that a system *is an equilibrium system*. We shall give a definition of the latter concept later, after describing the analogue of an equilibrium state for nonequilibrium systems.

For now, let us recall how equilibrium states are modelled in statistical mechanics. The basic ingredients are well known. Consider a system of N particles, which for simplicity we take to be a classical system, and let the sequence $\omega = (\omega_1, \omega_2, \ldots, \omega_N)$ denote the microscopic configuration or *microstate* of the system, where ω_i is the state of the ith particle. To study the statistical properties of this system, we consider a *prior* probability distribution $P(\omega)$, interpreted as the stationary distribution of the microscopic dynamics. The state space of one particle is denoted by Λ, so that $P(\omega)$ is a probability distribution over the N-particle space $\Lambda_N = \Lambda^N$.

Next, we consider a *macrostate* M_N, corresponding mathematically to a function $M_N(\omega)$ of the microstates, and proceed to compute the probability distribution of this random variable obtained under $P(\omega)$ using

$$P(M_N = m) = \int_{\Lambda_N} \delta(M_N(\omega) - m)\, P(\omega)\, d\omega. \tag{1.1}$$

If $P(\omega)$ is a valid model of an equilibrium system and the macrostate M_N is chosen properly, then what should be observed is that $P(M_N)$ is concentrated around certain highly probable values and that this concentration gets more pronounced as N gets larger. It is these *most probable* or *typical* values (points or states) of M_N that we call *equilibrium states* of M_N.

Mathematically, the concentration of $P(M_N)$ is akin to a Law of Large Numbers, in the sense that there exist sets B of values of M_N such that

$$\lim_{N\to\infty} P(M_N \in B) = 1 \qquad \text{and} \qquad \lim_{N\to\infty} P(M_N \notin B) = 0. \tag{1.2}$$

The smallest set B having this property corresponds to the set of equilibrium values of M_N or, more loosely, the set of equilibrium states of the system (as defined with respect to M_N).

We shall see in the next section that an essential property of the concentration of $P(M_N)$ on B is that it is *exponential* as a function of N (or, more generally, the volume of the system), which means that the probability that M_N deviates from one of its equilibrium values is exponentially small with the system size, N. Physically, these deviations are termed *fluctuations*, and so we say that the probability of fluctuations from equilibrium states is exponentially small with N.

The exponential concentration of $P(M_N)$ explains why large deviation theory is used in equilibrium statistical mechanics. Physically, it is also the reason why equilibrium states correspond to *typical* values of M_N and not, as often claimed, to *average* values of M_N. The fundamental property of equilibrium states is indeed that they do not fluctuate, or at least appear not to at the macroscopic level, so the fact that M_N has a well-defined average value is obviously not enough: M_N must converge in probability to some typical values.

The reason why average values are used in statistical mechanics is arguably that they are conceptually simpler than typical (or concentration) values and that, if M_N has a unique typical value, then its average is the same as its typical value (we say in this case that M_N is *self-averaging*). However, the use of averages is somewhat misleading as it detracts us from the essential property of equilibrium states, which is again that these states arise probabilistically from the concentration of a probability distribution. Equilibrium states are first and foremost *typical* states arising from the scaling limit that is the *thermodynamic limit* [19, 20].

This is an important point which leads us to the discussion of extensivity versus intensivity. Physically, it should be clear that the total energy H_N of an N-body system with short-range interactions does not concentrate in the thermodynamic limit, simply because the energy is *extensive* for such a system, and so diverges with N. As a result, one cannot formally say that the system has an *equilibrium energy* in the thermodynamic limit. Rather, the correct macrostate having an equilibrium value, i.e., the one that concentrates in the thermodynamic limit, is the energy per particle or mean energy $h_N = H_N/N$ which is an *intensive* quantity.

To explain this point more clearly, let us consider a sum

$$S_N = \sum_{i=1}^{N} X_i \tag{1.3}$$

of N independent and identically distributed random variables $X_1, \ldots, X_N$. If the mean $\langle X_1 \rangle = \mu$ of these random variables is finite, then as $N \to \infty$ the distribution of S_N/N concentrates to the mean in such a way that

$$\lim_{N\to\infty} P(|S_N/N - \mu| > \epsilon) = 0 \tag{1.4}$$

for all $\epsilon > 0$, in accordance with the Law of Large Numbers. The point to note about this result, which is equivalent to the second limit shown in (1.2), is that it holds for the mean sum S_N/N and not for the sum S_N: the distribution of $P(S_N)$ does not concentrate; in fact, it flattens as $N \to \infty$. Similarly, it is easy to show that the distribution of S_N/N^α flattens for all $\alpha \in (0, 1)$ and concentrates in a trivial way at zero for all $\alpha > 1$. Hence, the only normalization of a sum S_N of independent (or near independent) random variables with finite mean that yields a non-trivial concentration point is S_N/N.

The same observation applies to equilibrium states. The reason again why equilibrium states appear stable at the macroscopic level is that the fluctuations around these states are very improbable and become the more unlikely the bigger the system gets. Mathematically, the way to make sense of this observation is to consider

random variables that have the property of concentrating in the thermodynamic limit $N \to \infty$. From this point of view, the total energy H_N is not a "good" macrostate to consider because it does not concentrate in this limit. The same goes, similarly to S_N, for the macrostates $H_N/\sqrt{N}$ or H_N/N^2: the distribution of the former flattens, whereas the distribution of the latter concentrates trivially to zero. We get a non-trivial concentration only for H_N/N.

This at least is generally true for short-range interacting systems. For long-range interacting systems, such as gravitational systems or mean-field systems, the "good" energy macrostate to consider might be H_N/N^2 or, more generally, H_N/N^α with $\alpha \geq 2$ [21]. The choice of α will depend on the system considered, but the requirement again is that the distribution of the macrostate that is studied should concentrate when $N \to \infty$. From this point of view, one might have to consider different thermodynamic (or scaling) limits and macrostates in order to correctly describe the equilibrium states of those systems.

1.2.2 Nonequilibrium systems

The study of nonequilibrium systems is conceptually more difficult than that of equilibrium systems because one is interested in describing not only the stationary fluctuations, as is done for equilibrium systems, but also the fluctuating dynamics of the system arising in time and the fluctuations of macrostates or observables integrated over time. Thus, in addition to considering the number of particles present in a system (or its volume), one also needs to consider the evolution of that system in time. This implies that different scaling limits may be taken depending on the system and macrostate or observable studied.

For definiteness, we consider here nonequilibrium systems modelled by Markovian processes. To simplify the presentation of these models, we assume for now that the stochastic evolution takes place in discrete time (although continuous-time models will also be discussed in the following sections). In this case, the *microstate* ω that represented before the configuration of an equilibrium system at a fixed (yet unspecified) instant of time is now a complete trajectory $\omega = \{\omega_i\}_{i=1}^n$ consisting of n timesteps. The assumption that the process is Markovian then amounts to assuming that the prior distribution $P(\omega)$ can be decomposed according to a Markov chain

$$P(\omega) = P(\omega_1)P(\omega_2|\omega_1)\dots P(\omega_n|\omega_{n-1}), \tag{1.5}$$

with initial distribution $P(\omega_1)$ and transition matrix elements $P(\omega_i|\omega_{i-1})$, which, in most cases, are assumed to be time-homogeneous (i.e., time-independent). This form of prior is, from a pragmatic point of view, our stochastic model for ω from which all distributions of macrostates or observables are computed. Thus, so far, the formalism is abstractly the same as for equilibrium systems: a system is described by its microstate ω and a probability distribution $P(\omega)$ on the space of microstates. What changes for nonequilibrium systems is the interpretation of ω as a time-trajectory of a system of one or more particles.

This difference allows us to consider many types of macrostates or observables. For example, one can consider a *fixed-time* or *static* observable $M(\omega_i)$ which is a function of the state of the system at a specific timestep i. One can also consider *dynamic* observables of the form

$$M_n(\omega) = \frac{1}{n}\sum_{i=1}^{n} f(\omega_i), \tag{1.6}$$

referred to mathematically as *additive observables* or *additive functionals* [3], which involve states at different times. Another type of *dynamic* observable, which arises in the context of particle currents, is

$$M_n(\omega) = \frac{1}{n}\sum_{i=1}^{n-1} f(\omega_i, \omega_{i+1}). \tag{1.7}$$

Whatever the observable chosen, the goal when studying nonequilibrium systems is to compute the probability distribution $P(M = m)$ of a given observable M starting from the prior distribution $P(\omega)$ defining our model of that system, and to see whether this distribution concentrates, in some scaling limit, over specific values of M. The scaling limit that needs to be considered depends on the system and observable chosen: it can be the infinite-volume limit $N \to \infty$ for a fixed-time observable $M(\omega_i)$, the infinite-time limit $n \to \infty$ for additive or current-like observables, or a combination of these two limits if the latter observables involve many particles. Other limits can also be conceived, e.g., the small-noise limit of dynamical systems perturbed by noise, the continuous-time limit of discrete-time systems, or the continuous-space limit of discrete-space systems [3].

Some of these limits will be explored in the following sections. The essential point to note is that these *scaling* or *hydrodynamic limits* are expected to give rise to a concentration of the probability distribution $P(M)$ similar to the one discussed for equilibrium systems, and thus to the emergence of typical states for the system or observable studied.

1.2.3 Equilibrium versus nonequilibrium systems

So far we have not attempted to distinguish equilibrium from nonequilibrium systems in any precise way other than to hint that the distribution $P(\omega)$ describes a "static" random variable in the case of equilibrium systems and a "dynamic" random variable, i.e., a *stochastic process*, in the case of nonequilibrium systems. But what makes a system an equilibrium or a nonequilibrium system?

To answer this question in a mathematical way, we need to consider the stochastic time evolution of a system and study how the prior distribution $P(\omega)$ of its complete trajectory ω behaves when the time ordering of ω is reversed. To be more precise, consider the discrete-time trajectory $\omega = (\omega_1, \omega_2, \ldots, \omega_{n-1}, \omega_n)$ containing n timesteps, and define the *time-reversed* trajectory ω^R associated with ω

as the trajectory obtained by re-ordering the states of ω in reverse order, i.e., $\omega^R = (\omega_n, \omega_{n-1}, \ldots, \omega_2, \omega_1)$. Then we say that the system modelled by $P(\omega)$ is an *equilibrium system* if $P(\omega) = P(\omega^R)$ for all ω [4].[1)] If this condition is not satisfied, then we say that the system is a *nonequilibrium system*. Mathematically, this condition is equivalent to the notion of *reversibility* or *detailed balance*, stated here at the level of complete trajectories rather than the more usual level of transition rates. Thus, the stochastic dynamics of an equilibrium system satisfy detailed balance, whereas those of a nonequilibrium system do not.

The rationale behind this definition is that equilibrium prior distributions, such as the microcanonical and canonical distributions, are stationary distributions of stochastic dynamics verifying detailed balance, and that the very notion of detailed balance captures the physical observation that equilibrium systems are systems in which fluctuations arise with no "preferred" direction in time. Nonequilibrium systems, by contrast, have a stochastic dynamics that is not symmetric under time reversal. This does not mean that nonequilibrium systems do not have stationary distributions; they often do, but the form of these distributions is generally much more complicated than equilibrium distributions.

1.3 Elements of large deviation theory

We show in this section how the mathematical theory of large deviations makes precise the observation that probability distributions of macrostates concentrate *exponentially* with some scaling parameter (e.g., number of particles, volume, integration time, noise power, etc.). This exponential concentration is the source, for equilibrium systems, of the Legendre transform connecting the entropy and the free energy, and, therefore, of the Legendre structure of thermodynamics. For nonequilibrium systems, it also gives rise to a Legendre transform between quantities that are the nonequilibrium analogues of the entropy and the free energy.

1.3.1 General results

To explain the central ideas and results of large deviation theory, we first consider a general random variable or macrostate A_n indexed by the parameter n which can, for example, be the number of particles or the number of timesteps.

The starting point of large deviation theory is the observation that the probability distribution $P(A_n)$ of A_n is, for many random variables of interest, decaying to zero exponentially fast with n. The exponential decay is in general not exact; rather, what often happens is that the dominant term in the expression of $P(A_n)$

1) For simplicity, we assume that the ω's themselves have even parity under time-reversal.

is a decaying exponential with n, so that we can write

$$P(A_n = a) \approx e^{-nI(a)}, \tag{1.8}$$

with $I(a)$ the rate of decay. When $P(A_n)$ has this form, we say that $P(A_n)$ or A_n satisfies a *large deviation principle* (LDP). To be more precise, we say that $P(A_n)$ or A_n satisfies an LDP if the limit

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(A_n = a) = I(a) \tag{1.9}$$

exists. The decay function $I(a)$ defined by this limit is called the *rate function*. The factor n in the exponential is called the *speed* of the LDP [2, 3, 8].

The interest in large deviations arises because many random variables and stochastic processes satisfy such a principle (although not all do; see, e.g., [3]). The goal of large deviation theory, in this context, is to provide methods for proving that a given random variable or process satisfies an LDP and for obtaining the rate function controlling the rate of decay of the LDP.

Among these methods, let us mention two that are especially useful. The first is known as the *Gärtner-Ellis Theorem* [8] and proceeds by calculating the following function:

$$\lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln\langle e^{nkA_n}\rangle, \tag{1.10}$$

known as the *scaled cumulant generating function* (SCGF). The statement of the Gärtner-Ellis Theorem, in its simplified form, is that, if $\lambda(k)$ is differentiable for all $k \in \mathbb{R}$, then A_n satisfies an LDP with rate function given by the *Legendre-Fenchel transform* of $\lambda(k)$:

$$I(a) = \max_{k\in\mathbb{R}}\{ka - \lambda(k)\}. \tag{1.11}$$

In many physical applications, $\lambda(k)$ is actually differentiable *and* strictly convex, and in this case the Legendre-Fenchel transform reduces to the better-known *Legendre transform*, written as

$$I(a) = k(a)a - \lambda(k(a)), \tag{1.12}$$

with $k(a)$ the unique solution of $\lambda'(k) = a$.

The Gärtner-Ellis Theorem is useful in practice because it bypasses the direct calculation of $P(A_n)$. By calculating the SCGF of A_n and by checking that this function is differentiable, we instantly prove that $P(A_n)$ satisfies an LDP and obtain the rate function controlling the concentration of $P(A_n)$ in the limit $n \to \infty$.

For certain random variables, $\lambda(k)$ can be calculated but is not differentiable; see [3] for examples. In this case, it can be proved that the Legendre-Fenchel transform of $\lambda(k)$ yields only the convex envelope of $I(a)$ [3]. To obtain the full rate function, one may use other methods, such as the *contraction principle* [2, 3, 8, 22]. The

basis of this method is to express A_n as a function $f(B_n)$ of some random variable B_n satisfying an LDP with rate function $J(b)$, i.e.,

$$P(B_n = b) \approx e^{-nJ(b)}. \tag{1.13}$$

If such a random variable and function can be found, then the contraction principle states that A_n also satisfies an LDP with rate function given by

$$I(a) = \min_{b:f(b)=a} J(b). \tag{1.14}$$

The minimization appearing in this result is a natural consequence of the approximation method known as Laplace's principle [2, 3, 8] as applied to the integral

$$P(A_n = a) = \int_{b:f(b)=a} P(B_n = b)\, db. \tag{1.15}$$

Assuming that the probability distribution $P(B_n)$ decays exponentially with n, this integral is dominated by the largest exponential term such that $f(b) = a$, which means that we can write

$$P(A_n = a) \approx \exp\left(-n \min_{b:f(b)=a} J(b)\right) \tag{1.16}$$

with sub-exponential correction factors in n. Thus we see that $P(A_n)$ satisfies an LDP with rate function given by Eq. (1.14).

The name "contraction" arises in the context of this result from the fact that the function f can in general be a many-to-one function, in which case we are "contracting" the fluctuations of B_n down to the fluctuations of A_n in such a way that the probability of the fluctuation $A_n = a$ is the probability of the most probable (yet exponentially improbable) fluctuation $B_n = b$ leading to $A_n = a$.

This interplay between the appearance of exponentially small terms in integrals and the possibility to approximate these integrals by their largest term using Laplace's principle also explains the appearance of the "$\max$" in the Gärtner-Ellis Theorem and the Legendre transform connecting rate functions and SCGFs [3]. In this sense, large deviation theory can be thought of as a "calculus" of exponentially decaying probability distributions, connecting the properties of integrals such as $\langle e^{nkA_n} \rangle$, which are exponential in n, with the exponential properties of $P(A_n)$ itself.

1.3.2 Equilibrium large deviations

The application of the results stated above to equilibrium systems is straightforward. For definiteness, we consider a general macrostate $M_N(\omega)$ involving N particles and study its probability distribution $P_\beta(M_N)$ in the canonical ensemble defined by the prior probability distribution

$$P_\beta(\omega) = \frac{e^{-\beta H_N(\omega)}}{Z_N(\beta)}, \qquad Z_N(\beta) = \int_{\Lambda_N} e^{-\beta H_N(\omega)}\, d\omega, \tag{1.17}$$

where H_N is the Hamiltonian of the system considered.

If $P_\beta(M_N)$ satisfies an LDP, then the limit

$$\lim_{N\to\infty} -\frac{1}{N}\ln P(M_N = m) = I_\beta(m) \tag{1.18}$$

exists and defines the rate function $I_\beta(m)$ of M_N in the canonical ensemble at fixed inverse temperature β. The SCGF associated with this rate function is

$$\lambda_\beta(k) = \lim_{N\to\infty} \frac{1}{N}\ln\langle e^{NkM_N}\rangle_\beta, \tag{1.19}$$

where

$$\langle e^{NkM_N}\rangle_\beta = \int_{\Lambda_N} e^{NkM_N(\omega)} P_\beta(\omega)\, d\omega. \tag{1.20}$$

If $\lambda_\beta(k)$ is differentiable in k, then by the Gärtner-Ellis Theorem we have that $I_\beta(m)$ is the Legendre-Fenchel transform of $\lambda_\beta(k)$.

The connection between the LDP of M_N and its equilibrium values comes directly by writing the LDP informally in the form

$$P_\beta(M_N = m) \approx e^{-NI_\beta(m)}. \tag{1.21}$$

Since rate functions are always positive, this result shows that $P_\beta(M_N)$ decays exponentially fast with N except at points where $I_\beta(m)$ vanishes. As noted before, these points must correspond to the points where $P_\beta(M_N)$ concentrates in the limit $N \to \infty$, and so to the equilibrium values of M_N. Mathematically, we therefore define the set $\mathcal{E}_\beta$ of equilibrium values of M_N in the canonical ensemble as the set of global minima and zeros of the rate function $I_\beta(m)$:

$$\mathcal{E}_\beta = \{m : I_\beta(m) = 0\}. \tag{1.22}$$

A similar definition can be given for the equilibrium values of M_N in the microcanonical ensemble or any other ensemble by replacing $P_\beta(\omega)$ with the prior probability distribution defining these ensembles; see Sec. 5.3 of [3].

The rate function describes of course not only the equilibrium states but also the fluctuations around these states. In particular, if the rate function $I_\beta(m)$ admits a Taylor expansion of the form

$$I_\beta(m) = a(m - m^*)^2 + O(|m - m^*|^3) \tag{1.23}$$

around a given equilibrium value m^*, then the *small* fluctuations of M_N around m^* are Gaussian-distributed. The rate function, however, is rarely an exact parabola which means that the *larger* fluctuations of M_N away from m^* are in general not Gaussian-distributed. Their distribution is determined by the rate function $I_\beta(m)$ and depends on the system studied.

This explains the word "large" in large deviation: contrary to the Central Limit Theorem, which gives only information about the distribution of random variables

around their mean, large deviation theory gives information about this distribution near the mean but also away from the mean – i.e., it gives information about both the *small* and the *large* fluctuations or deviations of random variables. From this point of view, large deviation theory can be thought of as generalizing both the Law of Large Numbers and the Central Limit Theorem.

These observations are valid for any random variable. A more specific connection with equilibrium systems can be established by studying the large deviations of the mean energy $h_N = H_N/N$ with respect to the uniform distribution $P(\omega) = 1/|\Lambda_N|$. In this case, the integral

$$P(h_N = u) = \int_{\Lambda_N} \delta(h_N(\omega) - u)\, P(\omega)\, d\omega \tag{1.24}$$

is, up to a multiplicative constant, the density of states giving the number of microstates having a given mean energy $h_N(\omega) = u$. For most (if not all) equilibrium systems, the density of states is known to grow exponentially with N (or, more generally, the volume) and this directly implies an LDP for $P(h_N = u)$, which we write as

$$P(h_N = u) \approx e^{Ns(u)}. \tag{1.25}$$

In this form, it is clear that the function $s(u)$ obtained with the limit

$$s(u) = \lim_{N\to\infty} \frac{1}{N} \ln P(h_N = u), \tag{1.26}$$

is the thermodynamic entropy associated with h_N. To be more precise, it is the entropy of h_N, as usually calculated from the density of states, minus an unimportant additive constant; see Sec. 5.2 of [3].

To complete the connection with thermodynamics, note that the generating function

$$\langle e^{Nkh_N} \rangle = \int_{\Lambda_N} e^{Nkh_N(\omega)}\, P(\omega)\, d\omega \tag{1.27}$$

can be interpreted as the canonical partition function, whereas the SCGF of h_N, defined as

$$\lambda(k) = \lim_{N\to\infty} \frac{1}{N} \ln\langle e^{Nkh_N} \rangle, \tag{1.28}$$

can be seen as the analogue of the free energy function of the canonical ensemble. To be more precise, re-define the partition function $Z_N(\beta)$ by including the prior uniform distribution $P(\omega)$ in the integral over Λ_N:

$$Z_N(\beta) = \int_{\Lambda_N} e^{-\beta H_N(\omega)}\, P(\omega)\, d\omega \tag{1.29}$$

and define the free energy[2)] by

$$\varphi(\beta) = \lim_{N\to\infty} -\frac{1}{N} \ln Z_N(\beta). \tag{1.30}$$

2) In thermodynamics, the free energy is more commonly defined with an additional factor $1/\beta$ in front of the logarithm.

Then, it is easy to see that $\varphi(\beta) = -\lambda(k)$ with $k = -\beta$. From this connection, it is also easy to see that the Gärtner-Ellis Theorem implies that, if $\varphi(\beta)$ is differentiable, then

$$s(u) = \min_{\beta \in \mathbb{R}} \{\beta u - \varphi(\beta)\}. \tag{1.31}$$

This Legendre-Fenchel transform, with its inverse transform expressing $\varphi(\beta)$ in terms of $s(u)$ (see Sec. 5.2 of [3]), is the formal expression of the Legendre transform appearing in thermodynamics. The large deviation derivation of this transform makes it clear that it is valid under specific mathematical conditions (viz., the differentiability of φ) and that it arises because of the exponential nature of both $P(h_N)$ and $Z_N(\beta)$ and the Laplace's principle linking these two functions in the thermodynamic limit. In this sense, the Legendre transform of thermodynamics does not arise out of any physical requirement – it is a consequence of the large deviation structure of statistical mechanics, which appears in the thermodynamic limit.

1.3.3 Nonequilibrium large deviations

Let us now consider a nonequilibrium macrostate or observable $M_n(\omega)$ involving n timesteps of a Markov process $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ described by the transition matrix elements $P(\omega_i|\omega_{i-1})$. The large deviation properties of M_n can be studied similarly as done for equilibrium macrostates by calculating the SCGF $\lambda(k)$ associated with M_n and by obtaining the rate function of M_n as the Legendre-Fenchel transform of $\lambda(k)$ provided that $\lambda(k)$ is differentiable.

The Markov structure of the process underlying M_n can be used here to obtain more explicit expressions for $\lambda(k)$. In the case of the additive observable shown in Eq. (1.6), for example, we have

$$\lambda(k) = \ln \zeta(\mathbf{P}_k), \tag{1.32}$$

$\zeta(\mathbf{P}_k)$ being the largest eigenvalue of the so-called *tilted* transition matrix $\mathbf{P}_k$ with elements

$$P_k(\omega_i|\omega_{i-1}) = P(\omega_i|\omega_{i-1})e^{kf(\omega_i)}. \tag{1.33}$$

For the current-like observable M_n shown in Eq. (1.7), we have the same result but with $\mathbf{P}_k$ now given by

$$P_k(\omega_i|\omega_{i-1}) = P(\omega_i|\omega_{i-1})e^{kf(\omega_i,\omega_{i-1})}. \tag{1.34}$$

These results are valid if the state-space of the Markov chain is bounded. For unbounded state-spaces, $\lambda(k)$ is not necessarily given by the logarithm of the dominant eigenvalue of $\mathbf{P}_k$. We shall see a related example in Sec. 1.4.4.

For Markov processes evolving continuously in time, the above statements translate into the following ones. For an additive functional of the form

$$M_T(\omega) = \frac{1}{T}\int_0^T f(\omega_t)\,dt, \tag{1.35}$$

the SCGF $\lambda(k)$, calculated in the limit $T \to \infty$, is given by the largest eigenvalue of the *tilted generator*,

$$G_k(\omega', \omega) = G(\omega', \omega) + kf(\omega)\delta_{\omega',\omega}, \tag{1.36}$$

with $G(\omega', \omega)$ the elements of the generator of the original process. Note the absence of the logarithm here, as we are dealing with the generator, not the transition matrix. For current-like observables having the form

$$M_T(\omega) = \frac{1}{T}\sum_{i=0}^{N(T)-1} f(\omega_{t_i}, \omega_{t_{i+1}}) \tag{1.37}$$

where the sum is over the random transitions between states happening at times $\{t_0, t_1, \ldots, t_{N(T)-1}\}$, $\lambda(k)$ is given instead by the largest eigenvalue of a tilted generator with elements

$$G_k(\omega', \omega) = G(\omega', \omega)e^{kf(\omega,\omega')}. \tag{1.38}$$

Observables involving other scaling limits, in addition to the time or particle number limits, can be treated in a similar way, as discussed for example in Sec. 1.4.5. In all cases, the LDPs that we obtain give us information about the fluctuations of the observable of interest similar to the information obtained from equilibrium LDPs. In particular, the global minima and zeros of the rate function determine the typical values of that observable, which are physically interpreted as typical steady states or hydrodynamic states or equations, depending on the observable studied. Then, as for random variables in general, the shape of the rate function around its minima determines the behaviour of the small and large fluctuations of the observable around its typical values.

With the knowledge of the rate function of an observable, it is possible for example to determine whether this observable satisfies the so-called fluctuation relation symmetry. Consider, to be specific, an observable M_T integrated over the time T and assume that M_T satisfies an LDP with rate function $I(m)$. We say that M_T satisfies a *Gallavotti-Cohen-type fluctuation relation* if

$$\frac{P(M_T = m)}{P(M_T = -m)} \approx e^{Tcm}, \tag{1.39}$$

with c a positive constant. This means that the positive fluctuations of M_T are exponentially more probable than negative fluctuations of equal magnitude. In large deviation terms, it is easy to see that a sufficient condition for having this result is that $I(m)$ satisfy the following symmetry relation:

$$I(-m) - I(m) = cm. \tag{1.40}$$

In terms of the SCGF, we have equivalently

$$\lambda(k) = \lambda(-k - c). \tag{1.41}$$

The next section includes an example of a very simple Markov process having this fluctuation symmetry. For other more complicated examples, see, e.g., [3, 23].

1.4 Applications to nonequilibrium systems

In this section we aim to illustrate, more concretely, how the large deviation formalism of the preceding section can be applied to nonequilibrium systems as introduced in Sec. 1.2.2. We concentrate on Markov processes in continuous time, providing a detailed pedagogical treatment for toy models of random walkers and indicating how the same techniques can be applied to more complicated (and hence more interesting) models of interacting particles. Among other sources we draw here on the comprehensive review of Derrida [4] in which further details of many-particle applications can be found.

1.4.1 Random walkers in discrete and continuous time

To fix ideas, let us start by analysing perhaps the simplest possible model – a random walker in discrete space. Specifically, we consider a particle moving on a one-dimensional lattice of L sites with, for now, periodic boundary conditions. The microstate of the model is the particle's position $\omega \in \{1, 2, \ldots, L\}$ and, in discrete-time, the probabilities to move between those positions $P(\omega_i|\omega_{i-1})$ are contained in the transition matrix $\mathbf{P}$. We assume that the particle has probability p per timestep, with $0 < p < 1$, to move in the clockwise direction and probability q per timestep, with $0 < q \leq 1 - p$, to move in the anti-clockwise direction. Note that, if $p + q < 1$, the particle also has a finite probability to remain stationary in a given timestep. The position of the particle on the state space $\{1, 2, \ldots, L\}$ is thus a Markov chain with transition matrix

$$\mathbf{P} = \begin{pmatrix} 1-p-q & p & 0 & \ldots & q \\ q & 1-p-q & p & \ldots & 0 \\ 0 & q & 1-p-q & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p & 0 & 0 & \ldots & 1-p-q \end{pmatrix}. \tag{1.42}$$

It is a simple exercise to show that this Markov chain has a limiting (stationary) distribution with probability $1/L$ for the particle to be found on any given site – an intuitively obvious result! Slightly more interesting, from the nonequilibrium point of view, is the particle current $J_n(\omega)$ which we define as the net number

of clockwise jumps made by the particle *per timestep*. This is a function of the form (1.7) with

$$f(\omega_i, \omega_{i+1}) = \delta_{\omega_i+1,\omega_{i+1}} - \delta_{\omega_i-1,\omega_{i+1}}. \tag{1.43}$$

A straightforward calculation shows that the mean stationary current is given by $\langle J_n \rangle = p - q$. We shall see in Sec. 1.4.3 that this corresponds to the concentration point of the probability distribution $P(J_n = j)$ which satisfies an LDP. For now, note that there is an obvious qualitative difference between the case of $p = q$ (zero mean current) and the case of $p \neq q$ (non-zero mean current). Mathematically, this difference is just the distinction between a reversible and a non-reversible Markov chain, in the sense of detailed balance. As explained in Sec. 1.2.3, we identify the former case with an equilibrium system and the latter with a nonequilibrium system.

The continuous-time version of this random walk can be understood by associating a physical time increment Δt with each discrete timestep (where above we implicitly assumed $\Delta t = 1$), setting the hopping probabilities per timestep to $p\Delta t$ and $q\Delta t$ and then taking the limit $\Delta t \to 0$. Formally, our particle then remains at a given site for an exponentially distributed waiting time, with mean $1/(p+q)$, before moving clockwise, with probability $p/(p+q)$, or anti-clockwise, with probability $q/(p+q)$. Note, in particular, that p and q are now interpreted as rates rather than probabilities and can each be greater than unity. The infinitesimal generator corresponding to this process is

$$\mathbf{G} = \begin{pmatrix} -p-q & p & 0 & \dots & q \\ q & -p-q & p & \dots & 0 \\ 0 & q & -p-q & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p & 0 & 0 & \dots & -p-q \end{pmatrix}. \tag{1.44}$$

Unsurprisingly, a picture similar to the discrete-time case emerges: the process has a stationary state which has mean density $1/L$ on each site and mean current $p - q$. Following the previous sections, we shall be interested next in deriving such mean values as concentration points of LDPs. To be specific, we shall illustrate the general discussion below with explicit calculations related to the continuous-time random walk model and various modifications of it. We note that, as mentioned in Sec. 1.2.2, the appropriate scaling limit for which an LDP holds depends on the observable we wish to consider.

1.4.2
Large deviation principle for density profiles

A general interacting particle system has a state space consisting of all possible particle configurations and, on a coarse-grained scale, one is often interested in the probability of observing a particular fixed-time density profile in space. This leads

to the concept of a density function LDP which can be straightforwardly extended from equilibrium to nonequilibrium; see e.g., [4].

The appropriate scaling limit expressing such an LDP is the infinite volume (and infinite particle number) limit. To be precise, for a system defined on a lattice of linear size L in d dimensions, one considers taking the thermodynamic limit $L \to \infty$ whilst rescaling the coordinates $\mathbf{r}$ to $\mathbf{x} = \mathbf{r}/L$. The probability of seeing a given density profile $\rho(\mathbf{x})$ is then expected to obey

$$P[\rho(\mathbf{x})] \approx \exp\left\{-L^d \mathcal{F}[\rho(\mathbf{x})]\right\} \tag{1.45}$$

as $L \to \infty$. This is a *functional LDP*, as P and $\mathcal{F}$ are both functionals of $\rho(\mathbf{x})$.[3] The use of the square brackets emphasizes this point.

For equilibrium systems with short-range interactions, the form of the large deviation rate functional $\mathcal{F}$ is obtained from the knowledge of $f(\rho)$, the free energy per site, as

$$\mathcal{F}[\rho(\mathbf{x})] = \int_0^1 \left[f(\rho(\mathbf{x})) - f(\rho^*) - (\rho(\mathbf{x}) - \rho^*) f'(\rho^*)\right] \, d\mathbf{x}, \tag{1.46}$$

where ρ^* is just the mean number of particles per site [4]. We see here that $\mathcal{F} = 0$ for the uniform density profile $\rho(x) = \rho^*$. In other words, as expected, ρ^* is the typical density about which the probability distribution concentrates in the large volume limit.

To illustrate these results, let us consider a collection of independent random walkers in one dimension, discussed as a special case in [24]. To facilitate later generalizations, we now choose to work with open boundaries rather than periodic boundary conditions. Specifically, we modify our set-up to consider L sites coupled to left and right boundary reservoirs with densities ρ_L and ρ_R respectively, so that particles are input from the left reservoir with rate $p\rho_L$ and from the right reservoir with rate $q\rho_R$. In the bulk, and for exiting the system, each particle independently has the dynamics of the single random walker defined in Sec. 1.4.1 above.

For equal reservoir densities, $\rho_L = \rho_R = \rho^*$, it is a relatively simple exercise to show that for any p, q, the number of particles on each site is a Poisson distribution with mean ρ^*. The corresponding free energy (see e.g., [24]) leads to a large deviation functional of the form

$$\mathcal{F}[\rho(\mathbf{x})] = \int_0^1 \left[\rho^* - \rho(x) + \rho(x) \ln \frac{\rho(x)}{\rho^*}\right] \, dx. \tag{1.47}$$

A Taylor expansion of the integrand readily demonstrates Gaussian fluctuations about $\rho(x) = \rho^*$. Note that, in this special case of equal reservoir densities, the form of $\mathcal{F}$ is the same for $p = q$ and $p \neq q$: it is a local, convex, functional as is typically the case for equilibrium. Furthermore, an ensemble equivalence argument

3) $\mathcal{F}$ is the large deviation analogue of the Ginzburg-Landau free energy expressed as a function of the particle density in the grand-canonical ensemble.

suggests that in the $L \to \infty$ limit, the density large deviation functional would be the same for periodic boundary conditions with a fixed mean density ρ^*.

For interacting particle systems with *non-equal* reservoir densities (i.e., boundary driving) the situation is much more interesting. In particular, the density large deviation functional is generically expected to have a non-local structure reflecting the long-range spatial correlations characteristic of nonequilibrium. This is seen for example, in the case of the symmetric simple exclusion process [25, 26] which can be treated analytically by using the well-known matrix product ansatz [27] and an associated additivity property. In the corresponding asymmetric simple exclusion process, $\mathcal{F}$ is non-convex for some parameters indicating a phase transition [28, 29]. The form of $\mathcal{F}$ in certain models can be obtained by utilising the macroscopic fluctuation theory of Bertini *et al.* [5, 30] to which we shall return in Sec. 1.4.5.

1.4.3
Large deviation principle for current fluctuations

The presence of non-zero currents is a generic feature of nonequilibrium stationary states. For Markov processes, one generically finds that a given current J_T (e.g., the net number of particles hopping between two lattice sites) time-averaged over the interval $[0, T]$ obeys a large deviation principle with speed T, i.e.,

$$P(J_T = j) \approx e^{-TI(j)}. \tag{1.48}$$

The relevant scaling limit here is the long-time, $T \to \infty$, limit. Although this may be combined with an infinite volume limit (as in Sec. 1.4.5 below), there is particular interest in current fluctuations in *small systems*, e.g., trapped colloidal particles (see the chapter of this book by Reid *et al.*) or single molecule biological experiments (see the contribution by Alemany *et al.*). In this spirit, we use this section to illustrate the calculation and properties of $I(j)$ for a single random walker. For a treatment of general Markov diffusions, see [16, 31].

As follows from the general discussion in Sec. 1.3.3, the rate function $I(j)$ can be obtained as the Legendre-Fenchel transform of the SCGF $\lambda(k)$ of J_T, which, for a continuous-time process with finite state-space, is given by the principal eigenvalue $\zeta(k)$ of a tilted generator. For example, for the single-particle random walker on a ring with a current J_T defined, as in Sec. 1.4.1, to be the net number of clockwise jumps the particle makes per unit time, then we need the principal eigenvalue of the matrix

$$\mathbf{G}(k) = \begin{pmatrix} -p-q & pe^{k} & 0 & \dots & qe^{-k} \\ qe^{-k} & -p-q & pe^{k} & \dots & 0 \\ 0 & qe^{-k} & -p-q & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ pe^{k} & 0 & 0 & \dots & -p-q \end{pmatrix}. \tag{1.49}$$

It is easy to show that the normalized vector $(1/L, 1/L, \dots, 1/L)$ is a left-

eigenvector of this matrix with eigenvalue

$$-p(1-e^{k})-q(1-e^{-k}) \tag{1.50}$$

and an appeal to Perron-Frobenius theory [32] supports the assertion that this is the desired principal eigenvalue $\zeta(k)$ which is equal to the SCGF $\lambda(k)$. From here it is a straightforward, albeit tedious, exercise to calculate $I(j)$:

$$I(j)=\max_{k}\{kj-\lambda(k)\}=p+q-\sqrt{j^2+4pq}+j\ln\frac{j+\sqrt{j^2+4pq}}{2p}. \tag{1.51}$$

Note that this result can also be obtained by arguing that the clockwise jumps form a Poisson process with rate p, whereas the anti-clockwise jumps form a Poisson process with rate q. Considering the long-time limit of the Poisson process, we hence have separate large deviation functions for the clockwise current J_+ and the anti-clockwise current J_- with respective rate functions

$$I(j_+)=p-j_++j_+\ln\frac{j_+}{p},\qquad I(j_-)=q-j_-+j_-\ln\frac{j_-}{q}. \tag{1.52}$$

The rate function for the net current $J_T=J_+-J_-$ can then be obtained by the method of contraction discussed in Sec. 1.3.1.

Notice that the rate function for the current of the random walker obeys the fluctuation relation symmetry

$$I(-j)-I(j)=cj \tag{1.53}$$

with $c=\ln(p/q)$, so that

$$\frac{P(J_T=j)}{P(J_T=-j)}=e^{Tcj}. \tag{1.54}$$

This result is a simple example of a fluctuation relation of the Gallavotti-Cohen type [33–35]. As mentioned in Sec. 1.3.3, this result can also be expressed as the SCGF property $\lambda(k)=\lambda(-k-c)$ which derives ultimately from a straightforwardly-verified symmetry of the tilted generator:

$$\mathbf{G}(k)^{T}=\mathbf{G}(-k-c) \tag{1.55}$$

where T denotes here transpose (not time). Note that for other currents (e.g., counting the jumps across just a single bond), $\mathbf{G}(k)^{T}$ is no longer equal to $\mathbf{G}(-k-c)$ but, under quite general conditions, is related to it by a similarity transform so that the two matrices have identical eigenvalues and the symmetry (1.54) still holds. The diagonal change-of-basis matrix in the similarity transform is related to current boundary terms which, for finite state-space, are irrelevant in the long-time limit. The plethora of different finite-time fluctuation relations can be associated with different choices for these boundary terms; see, e.g., [23, 36] and elsewhere in this volume.

We shall see with an example below that, for infinite state space, the boundary terms may become relevant. In passing, we also note that whilst the fluctuation theorem can be elegantly expressed as a property of the large deviation rate function, the existence of a large deviation principle is not, as sometimes believed, a *necessary* prerequisite for the existence of a fluctuation relation of form (1.54). A simple counter-example is provided by a random walker with right and left hopping rates increasing in time as $p \times t$ and $q \times t$ respectively. It is easy to show that such a system has no stationary state (the mean current increases indefinitely), but since the ratio of rates of right and left steps is constant, a relation of the form (1.54) still holds.

1.4.4 Interacting particle systems: features and subtleties

Thus far, the explicit examples of this section have been concerned with single-particle random walks or non-interacting collections thereof. The same general formalism applies for interacting particle systems [37], although analytically tractable models are the exception rather than the rule. Paradigmatic examples include the symmetric and asymmetric simple exclusion processes, mentioned already in Sec. 1.4.2, and the zero-range process (ZRP) [38]. Among other results, the current large deviations in the open-boundary asymmetric exclusion process have recently been calculated [39, 40]. Here we concentrate on the ZRP with open boundary conditions [41] (connected to so-called Jackson networks of queueing theory [42]) in order to exemplify subtleties arising in systems with unbounded state space.

For our purposes, it suffices to consider the ZRP on a one-dimensional open lattice, although related issues have also been examined for queuing models on more complicated geometries [43]. In one dimension, each site $l \in \{1, 2, \ldots, L\}$ contains an integer number of particles n_l which can hop to the nearest neighbour sites according to a continuous-time dynamics. Specifically, in the bulk the topmost particle on each site hops to the right (or to the left) with rate pw_n (respectively, qw_n) where w_n is a function of the number of particles n on the departure site. Particles are injected onto site 1 (or L) with rate α (respectively, δ) and extracted with rate γw_n (respectively, βw_n).

The properties of the model depend crucially on the function w_n. The choice $w_n \propto n$ corresponds to non-interacting particles, such as the random walkers considered above, whereas other forms represent an effective on-site attraction or repulsion. In particular, if w_n is bounded as $n \to \infty$, i.e.,

$$\lim_{n\to\infty} w_n = a < \infty, \tag{1.56}$$

then the model exhibits a condensation transition where, for some choices of boundary rates, particles "pile up" indefinitely at one or more sites and there is no stationary state. For boundary rates outside this regime, the stationary state of the model is a product measure characterized by a site-dependent fugacity. The mean current across each bond (i.e., between each pair of sites) depends on the rates p,

q, α, β, γ, δ, but not explicitly on w_n. However, the form of w_n determines the relationship between the fugacity and the particle density and also the relationship between α and δ and effective reservoir densities at the boundaries.

To calculate the fluctuations around the stationary-state current, we can try to look, as above, for the principal eigenvalue of a tilted generator (which can be represented in terms of tensor products of matrices encoding the particle dynamics on each site). The form of this tilted generator will depend on the bond(s) across which we choose to measure the current. In the case where w_n is unbounded, in the sense that $w_n \to \infty$ as $n \to \infty$, then the spectrum of the tilted generator is always gapped and the Legendre transform of the principal eigenvalue, which can be explicitly calculated in terms of the transition rates, gives the large deviation rate function for all values of current. Furthermore, as might be expected, the principal eigenvalue, and hence the current fluctuations, are the same for currents across all bonds.

On the other hand, for w_n bounded, the spectrum of the tilted generator becomes gapless for some values of k and certain boundary terms can also diverge. Mathematically, this means that $\lambda(k)$ is no longer simply given by the principal eigenvalue. Physically, this possibility is related to the fact that, over long-but-finite timescales, an arbitrarily large number of particles can accumulate on each site. This manifests in the following properties of the current large deviation function:

- It is bond inhomogeneous so that the probability of seeing extreme current fluctuations depends on where the current is measured;
- It depends on the initial probability distribution of the system;
- It does not obey the Gallavotti-Cohen fluctuation relation for large current fluctuations.

All of these properties are seen even in the single site zero-range process, for which the complete spectrum of the tilted generator and the form of $\lambda(k)$ for all k can be explicitly obtained [36, 44]. A phase diagram in x–k space, where x is the fugacity characterizing the initial state, reveals that there are two types of "phase transition" in $\lambda(k)$ analogous to first-order and continuous transitions in equilibrium. At the former, $\lambda(k)$ has a non-differentiable point, while at the latter, it remains differentiable. An attentive reader may question how we can then obtain $I(j)$, since the Gärtner-Ellis Theorem of Sec. 1.3.1 requires differentiability of $\lambda(k)$ for all k. In fact, the Legendre-Fenchel transform yields the convex envelope of $I(j)$ which contains linear sections corresponding to the non-differentiable points of $\lambda(k)$. For the ZRP, one can argue on physical grounds that this *is* the correct form of $I(j)$ because the most probable way to realise an average current j in the linear regime involves a phase separation in time with the system spending part of its history in a state with one average current and part in a state with another average current. This argument relies on the system having only short-range correlations in time (just as the analogous Maxwell construction in equilibrium requires short-range correlations in space) so it might be expected to fail, for example, in non-Markovian systems.

1.4.5 Macroscopic fluctuation theory

Underlying the so-called *macroscopic fluctuation theory* is the concept of the hydrodynamic limit which describes the emergence of a deterministic coarse-grained description from stochastic microscopic rules [45]. Recall from Secs. 1.2 and 1.3 that such a non-fluctuating macroscopic state, corresponding to the concentration point of some probability distribution, is expected to be given by the zero of a large deviation rate function. In this subsection, we sketch the approach of Bertini *et al.* [5, 30] for calculating this rate function.

We are interested here in systems with particle conservation in the bulk and a key ingredient is the functional form of the dependence of the instantaneous local current on the density. The correct scaling required so that the *joint* distribution of current and density profiles concentrates in the limit $L \to \infty$ depends on the form of this current-density relationship. Specifically, we focus our attention here on *diffusive* processes for which the relevant macroscopic coordinates are $\mathbf{x} = \mathbf{r}/L$ and $\tau = t/L^2$. This class of systems includes symmetric and weakly-asymmetric versions of both the exclusion process and the zero-range process, but not their asymmetric counterparts for which *Euler scaling* $\tau = t/L$ is needed. To illustrate loosely the procedure involved in taking the hydrodynamic limit, we now return to our favourite example of random walkers.

Consider a collection of non-interacting particles on a one-dimensional lattice with boundary reservoirs ρ_L and ρ_R, as in Sec. 1.4.2, and a weak asymmetry in the bulk hopping dynamics, viz., $p = 1/2 + E/2L$ and $q = 1/2 - E/2L$. The starting point for the hydrodynamic description is to consider the mean current between two neighbouring lattice sites, l and $l+1$, say. In terms of the site densities (mean occupation numbers) ρ_l and ρ_{l+1}, the mean current is

$$\langle J_{l,l+1} \rangle = p\rho_l - q\rho_{l+1}, \tag{1.57}$$

which yields a lattice continuity equation of the form

$$\frac{\partial \rho_l(t)}{\partial t} = \langle J_{l-1,l} \rangle - \langle J_{l,l+1} \rangle = [p\rho_{l-1}(t) - q\rho_l(t)] - [p\rho_l(t) - q\rho_{l+1}(t)]. \tag{1.58}$$

Now writing $t = L^2\tau$ and $l = xL$ we have

$$\frac{1}{L^2}\frac{\partial \rho(x,\tau)}{\partial \tau} = \left[p\rho(x - \tfrac{1}{L}, \tau) - q\rho(x,\tau)\right] - \left[p\rho(x,\tau) - q\rho(x + \tfrac{1}{L}, \tau)\right]. \tag{1.59}$$

Then assuming *local stationarity* and carrying out a Taylor expansion to second order, yields

$$\frac{\partial \rho(x,\tau)}{\partial \tau} = -\frac{\partial}{\partial x}\left[E\rho - \frac{1}{2}\frac{\partial \rho}{\partial x}\right] \equiv -\frac{\partial}{\partial x}\hat{J}(x,\tau) \tag{1.60}$$

with $\hat{J}(x,\tau)$ a rescaled current.

Similarly, for general d-dimensional diffusive systems obeying Fick's Law at the macroscopic level, we expect

$$\hat{\mathbf{J}}(\mathbf{x},t) = \sigma[\rho(\mathbf{x},t)]\mathbf{E} - D[\rho(\mathbf{x},t)]\nabla\rho(\mathbf{x},t) \tag{1.61}$$

where $D[\rho(\mathbf{x},t)]$ is the diffusivity associated with the density profile $\rho(\mathbf{x},t)$ and $\sigma[\rho(\mathbf{x},t)]$ is the corresponding mobility. The hydrodynamic equation, describing the deterministic or macroscopic limit as $L \to \infty$, is

$$\frac{\partial\rho(\mathbf{x},\tau)}{\partial\tau} = -\nabla\cdot\hat{\mathbf{J}}(\mathbf{x},\tau). \tag{1.62}$$

Together, Eq. (1.61) and (1.62) only represent the macroscopic or typical behavior obtained in the hydrodynamic limit. To describe the fluctuations around this limit, let us now find the joint rate function of the density and current. Specializing to the case $\mathbf{E} = 0$, one observes (see, e.g., [4]) that for boundary reservoirs with equal density ρ^* the fluctuations of the microscopic current across each bond can be characterized by

$$\lim_{t\to\infty}\frac{\langle\mathbf{J}^2\rangle}{t} = \frac{\sigma(\rho^*)}{L}. \tag{1.63}$$

At the macroscopic level, this motivates adding to $\hat{\mathbf{J}}(\mathbf{x},\tau)$ a term representing Gaussian white noise with variance $\sigma[\rho(x,t)]$, which leads to an LDP for the joint probability of seeing a particular density and current profile having the form

$$P[\rho(\mathbf{x},\tau),\hat{\mathbf{J}}(\mathbf{x},\tau)] \sim \exp\left\{-L^d\int_0^T\int_0^1\frac{(\hat{\mathbf{J}}(\mathbf{x},\tau)+D[\rho(\mathbf{x},\tau)]\nabla\rho(\mathbf{x},\tau))^2}{2\sigma[\rho(\mathbf{x},\tau)]}\,d\mathbf{x}\,d\tau\right\} \tag{1.64}$$

with $\hat{\mathbf{J}}$ and ρ linked by the continuity equation (1.62).

In principle, from here one can use the tools of variational calculus to look for the optimal density profile $\rho(\mathbf{x},\tau)$ leading to a given final density $\rho(\mathbf{x})$. This contraction leads to an implicit expression for the density rate function $\mathcal{F}[\rho(\mathbf{x})]$ defined in Sec. 1.4.2. Finding explicit solutions is a difficult task, since in general the optimal profile is time-dependent. However, successful treatments along these lines include the one-dimensional zero-range and Kipnis-Marchioro-Presutti models [30, 46]; in [30] it is also shown that the approach is consistent with the independent results of Derrida *et al.* [25, 26] for the symmetric simple exclusion process.

It turns out to be easier to calculate the current large deviation function [47], in particular, if one assumes that the optimal profile leading to a particular current fluctuation is independent of time and that the optimal current is constant in space. In one dimension, the first assumption can be shown [48] to be equivalent to the additivity principle of Bodineau and Derrida [49], which is known to break down in systems with dynamical phase transitions; the second assumption is important

for higher-dimensional systems [50]. Under these conditions, the rate function for the (rescaled) current is then given by

$$I(\hat{\mathbf{J}}) = \min_{\rho(\mathbf{x})} \int_0^1 \frac{(\hat{\mathbf{J}} + D[\rho(\mathbf{x})]\nabla\rho(\mathbf{x}))^2}{2\sigma[\rho(\mathbf{x})]}\, d\mathbf{x}. \tag{1.65}$$

For any particular model, this integral must be minimized with $\rho(\mathbf{x})$ matched to the reservoir densities at the boundaries. This generically yields a current distribution with non-Gaussian tails even though the fluctuations themselves are locally Gaussian.

Various scenarios for the form of the macroscopic current large deviation function (including those indicating dynamical phase transitions) are discussed in [47]. The example models treated there include the one-dimensional zero-range process with w_n unbounded, the special case $w_n = n$ corresponding again to non-interacting random walkers. It has recently been pointed out that, if Eq. (1.65) holds, the optimal density profile is the same for all currents with the same magnitude $|\hat{\mathbf{J}}|$ leading to what has been dubbed an *isometric fluctuation relation* [51]. Macroscopic fluctuation theory has also been generalized to treat models with dissipated energy [52].

1.5 Final remarks

In this chapter we have merely skimmed the surface of the way in which large deviation theory can provide a framework for understanding existing results in the theory of nonequilibrium systems and probing for new ones. We conclude here with some pointers to other relevant work and ideas for future research directions.

Firstly, we note that the study of fluctuation theorems and relations, as briefly touched on in Secs. 1.3.3 and 1.4.3, is a vast subject which percolates through many of the contributions in this volume; for overviews, see for example the chapters by Spinney and Ford, Gaspard, and Rondoni and Jepps. In particular, we have not discussed here the important concept of entropy production and its distribution, which is central for many fluctuation relation statements and can often be expressed in a large deviation form.

A second topic worth mentioning is the extension of the concept of statistical ensemble, discussed for equilibrium systems in Sec. 1.3.2, to nonequilibrium systems. Just as one distinguishes equilibrium systems with fixed energy (microcanonical ensemble) from systems with the energy fixed on average via a conjugate Lagrange parameter (canonical ensemble), one can construct a microcanonical ensemble of trajectories for a nonequilibrium system that is constrained to realise a particular observable value (e.g., a particular current) or a canonical ensemble of trajectories that realise that constraint on average. An example of the former ensemble, obtained for the ASEP on a ring conditioned on enhanced flux, has recently been analyzed in [53]. A study of canonical-type nonequilibrium ensembles, which are also known as *biased ensembles*, can be found in the work of Sollich and Jack [54].

In the context of large deviation theory, these ensembles can be understood in terms of conditional LDPs and the so-called Gibbs conditioning principle [2].

Related to the topic of nonequilibrium ensembles is the issue of determining configurations or states giving rise to fluctuations. We can already get information about the most probable (typical) way to realise a given current fluctuation from the eigenvector corresponding to the principal eigenvalue of the tilted generator (1.38). More work is still needed to understand the properties and correlations of such current-carrying states and constrained states in general. The associated inverse problem of determining microscopic rates which are most likely to yield given macroscopic properties has been studied by Evans [55, 56] and Monthus [57].

Throughout this chapter we have assumed that non-equilibrium systems of interest are modelled by Markov processes. However, the memoryless property may be an inappropriate approximation for the description of many systems where long-range temporal correlations are known to be important; see e.g., [58] and references therein. Recent work to characterize the large deviation properties of certain classes of history-dependent models can be found in [59] and [60]. In particular, it is shown in [59] that modifying a continuous-time random walker (as introduced in Sec. 1.4.1) so that the hopping rates at time t depend on the average current up to time t can lead to an altered "speed" (i.e., power of time T) in the LDP for current. Fluctuation relations with the right-hand side of (1.39) replaced by $e^{T^\alpha cm}$ have also appeared in the context of anomalous dynamics; see the contribution by Klages *et al.* in this book. There is much scope for future work investigating many-particle non-Markovian processes and establishing a common framework for the results. In this regard, and in a more general way, we expect the large deviation formalism to continue playing an important role in quantifying nonequilibrium fluctuations in small systems.

Acknowledgments

HT is grateful for the hospitality of the École Normale Supérieure of Lyon and the National Institute for Theoretical Physics at Stellenbosch University, where parts of this chapter were written.

Bibliography

1 Deuschel, J.D. and Stroock, D.W. (1989) *Large Deviations*, Academic Press, Boston.

2 Dembo, A. and Zeitouni, O. (1998) *Large Deviations Techniques and Applications*, Springer, New York, 2nd edn..

3 Touchette, H. (2009) The large deviation approach to statistical mechanics. *Phys. Rep.*, **478** (1-3), 1–69.

4 Derrida, B. (2007) Non-equilibrium steady states: Fluctuations and large deviations of the density and of the current. *J. Stat. Mech.*, **2007** (07), P07 023.

5 Bertini, L., Sole, A.D., Gabrielli, D., Jona-Lasinio, G., and Landim, C. (2007) Stochastic interacting particle systems out of equilibrium. *J. Stat. Mech.*, **2007** (07), P07 014.

6 Ruelle, D. (1969) *Statistical Mechanics: Rigorous Results*, W. A. Benjamin,

Amsterdam.

7 Lanford, O.E. (1973) Entropy and equilibrium states in classical statistical mechanics, in *Statistical Mechanics and Mathematical Problems*, *Lecture Notes in Physics*, vol. 20 (ed. A. Lenard), Springer, Berlin, *Lecture Notes in Physics*, vol. 20, pp. 1–113.

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

9 Oono, Y. (1989) Large deviation and statistical physics. *Prog. Theoret. Phys. Suppl.*, **99**, 165–205.

10 Paniconi, M. and Oono, Y. (1997) Phenomenological framework for fluctuations around steady state. *Phys. Rev. E*, **55** (1), 176–188.

11 Oono, Y. and Paniconi, M. (1998) Steady state thermodynamics. *Prog. Theoret. Phys. Suppl.*, **130**, 29–44.

12 Eyink, G.L. (1990) Dissipation and large thermodynamic fluctuations. *J. Stat. Phys.*, **61** (3), 533–572.

13 Eyink, G.L. (1996) Action principle in nonequilibrium statistical dynamics. *Phys. Rev. E*, **54**, 3419–3435.

14 Eyink, G.L. (1998) Action principle in statistical dynamics. *Prog. Theoret. Phys. Suppl.*, **130**, 77–86.

15 Maes, C., Netočný, K., and Shergelashvili, B. (2006) A selection of nonequilibrium issues, in *Methods of Contemporary Mathematical Statistical Physics*, *Lecture Notes in Mathematics*, vol. 1970 (eds M. Biskup, A. Bovier, F. den Hollander, D. Ioffe, F. Martinelli, K. Netočný, and C. Toninelli), Springer, Berlin, pp. 1–60.

16 Maes, C. and Netočný, K. (2008) Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. *Europhys. Lett.*, **82** (3).

17 Maes, C. and Netočný, K. (2007) Minimum entropy production principle from a dynamical fluctuation law. *J. Math. Phys.*, **48**, 053 306.

18 Maes, C. and Netočný, K. (2003) Time-reversal and entropy. *J. Stat. Phys.*, **110** (1), 269–310.

19 Goldstein, S. (2001) Boltzmann's approach to statistical mechanics, in *Chance in Physics: Foundations and Perspectives*, *Lecture Notes in Physics*, vol. 574 (eds J. Bricmont, D. Dürr, M.C. Galavotti, G.C. Ghirardi, F. Petruccione, and N. Zanghi), Springer, Berlin, *Lecture Notes in Physics*, vol. 574, pp. 39–54.

20 Goldstein, S. and Lebowitz, J.L. (2004) On the (Boltzmann) entropy of non-equilibrium systems. *Physica D*, **193** (1-4), 53–66.

21 Campa, A., Dauxois, T., and Ruffo, S. (2009) Statistical mechanics and dynamics of solvable models with long-range interactions. *Phys. Rep.*, **480** (3-6), 57–159.

22 Touchette, H. (2010) Methods for calculating nonconcave entropies. *J. Stat. Mech.*, **2010** (05), P05 008.

23 Harris, R.J. and Schütz, G.M. (2007) Fluctuation theorems for stochastic dynamics. *J. Stat. Mech.*, **2007** (07), P07 020.

24 Derrida, B. and Gerschenfeld, A. (2009) Current fluctuations in one dimensional diffusive systems with a step initial density profile. *J. Stat. Phys.*, **137** (5–6), 978–1000.

25 Derrida, B., Lebowitz, J.L., and Speer, E.R. (2001) Free energy functional for nonequilibrium systems: An exactly solvable case. *Phys. Rev. Lett.*, **87** (15), 150 601.

26 Derrida, B., Lebowitz, J.L., and Speer, E.R. (2002) Large deviation of the density profile in the steady state of the open symmetric simple exclusion process. *J. Stat. Phys.*, **107** (3–4), 599–634.

27 Derrida, B., Evans, M.R., Hakim, V., and Pasquier, V. (1993) Exact solution of a 1D asymmetric exclusion model using a matrix formulation. *J. Phys. A: Math. Gen.*, **26** (7), 1493–1517.

28 Derrida, B., Lebowitz, J.L., and Speer, E.R. (2002) Exact free energy functional for a driven diffusive open stationary nonequilibrium system. *Phys. Rev. Lett.*, **89** (3), 030 601.

29 Derrida, B., Lebowitz, J.L., and Speer, E.R. (2003) Exact large deviation functional of a stationary open driven diffusive system: The asymmetric exclusion process. *J. Stat. Phys.*, **110** (3–6), 775–810.

30 Bertini, L., Sole, A.D., Gabrielli, D., Jona-Lasinio, G., and Landim, C. (2002)

Macroscopic fluctuation theory for stationary non-equilibrium states. *J. Stat. Phys.*, **107** (3–4), 635–675.

31 Maes, C., Netočný, K., and Wynants, B. (2008) Steady state statistics of driven diffusions. *Physica A*, **387** (12), 2675–2689.

32 Minc, H. (1988) *Nonnegative Matrices*, Wiley, New York.

33 Evans, D.J. and Searles, D.J. (1994) Equilibrium microstates which generate second law violating steady states. *Phys. Rev. E*, **50** (2), 1645–1648.

34 Gallavotti, G. and Cohen, E.G.D. (1995) Dynamical ensembles in nonequilibrium statistical mechanics. *Phys. Rev. Lett.*, **74** (14), 2694–2697.

35 Lebowitz, J.L. and Spohn, H. (1999) A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.*, **95** (1–2), 333–365.

36 Rákos, A. and Harris, R.J. (2008) On the range of validity of the fluctuation theorem for stochastic Markovian dynamics. *J. Stat. Mech.*, **2005** (05), P05 005.

37 Liggett, T.M. (1985) *Interacting Particle Systems*, Springer, New York.

38 Evans, M.R. and Hanney, T. (2005) Nonequilibrium statistical mechanics of the zero-range process and related models. *J. Phys. A: Math. Gen.*, **38** (19), R195–R240.

39 Lazarescu, A. and Mallick, K. (2011) An exact formula for the statistics of the current in the TASEP with open boundaries. *J. Phys. A: Math. Theor.*, **44** (31), 315 001.

40 de Gier, J. and Essler, F.H.L. (2011) Large deviation function for the current in the open asymmetric simple exclusion process. *Phys. Rev. Lett.*, **107**, 010 602.

41 Levine, E., Mukamel, D., and Schütz, G.M. (2005) Zero-range process with open boundaries. *J. Stat. Phys.*, **120** (5–6), 759–778.

42 Jackson, J.R. (1957) Networks of waiting lines. *Oper. Res.*, **5** (4), 518–521.

43 Chernyak, V.Y., Chertkov, M., Goldberg, D.A., and Turitsyn, K. (2010) Non-equilibrium statistical physics of currents in queuing networks. *J. Stat. Phys.*, **140** (5), 819–845.

44 Harris, R.J., Rákos, A., and Schütz, G.M. (2006) Breakdown of Gallavotti-Cohen fluctuation theorem for stochastic dynamics. *Europhys. Lett.*, **75** (2), 227–233.

45 Kipnis, C. and Landim, C. (1999) *Scaling Limits of Interacting Particle Systems*, *Grundlehren der mathematischen Wissenschaften*, vol. 320, Springer, Berlin.

46 Bertini, L., Gabrielli, D., and Lebowitz, J.L. (2005) Large deviations for a stochastic model of heat flow. *J. Stat. Phys.*, **121** (5–6), 843–885.

47 Bertini, L., Sole, A.D., Gabrielli, D., Jona-Lasinio, G., and Landim, C. (2006) Non equilibrium current fluctuations in stochastic lattice gases. *J. Stat. Phys.*, **123** (2), 237–276.

48 Bertini, L., Sole, A.D., Gabrielli, D., Jona-Lasinio, G., and Landim, C. (2005) Current fluctuations in stochastic lattice gases. *Phys. Rev. Lett.*, **94** (3), 030 601.

49 Bodineau, T. and Derrida, B. (2004) Current fluctuations in non-equilibrium diffusive systems: an additivity principle. *Phys. Rev. Lett.*, **92** (18), 180 601.

50 Pérez-Espigares, C., del Pozo, J.J., Garrido, P.L., and Hurtado, P.I. (2011) *Nonequilibrium Statistical Physics Today: Proceedings of the 11th Granada Seminar on Computational and Statistical Physics*, American Institute of Physics, *AIP Conf. Proc.*, vol. 1332, chap. Large deviations of the current in a two-dimensional diffusive system, pp. 204–213.

51 Hurtado, P.I., Pérez-Espigares, C., del Pozo, J.J., and Garrido, P.L. (2011) Symmetries in fluctuations far from equilibrium. *Proc. Natl. Acad. Sci.*, **108** (19), 7704–7709.

52 Prados, A., Lasanta, A., and Hurtado, P.I. (2011) Large fluctuations in driven dissipative media. *Phys. Rev. Lett.*, **107**, 140 601.

53 Popkov, V., Schütz, G.M., and Simon, D. (2010) ASEP on a ring conditioned on enhanced flux. *J. Stat. Mech.*, **2010** (10), P10 007.

54 Jack, R.L. and Sollich, P. (2010) Large deviations and ensembles of trajectories in stochastic models. *Prog. Theor. Phys. Supplement*, **184**, 304–317.

55 Evans, R.M.L. (2004) Rules for transition

rates in nonequilibrium steady states. *Phys. Rev. Lett.*, **92** (15), 150 601.

56 Evans, R.M.L. (2005) Detailed balance has a counterpart in non-equilibrium steady states. *J. Phys. A: Math. Gen.*, **38** (2), 293–313.

57 Monthus, C. (2011) Non-equilibrium steady states: maximization of the Shannon entropy associated with the distribution of dynamical trajectories in the presence of constraints. *J. Stat. Mech.*, **2011** (03), P03 008.

58 Rangarajan, G. and Ding, M. (eds) (2003) *Processes with Long-Range Correlations: Theory and Applications*, *Lecture Notes in Physics*, vol. 621, Springer, Berlin.

59 Harris, R.J. and Touchette, H. (2009) Current fluctuations in stochastic systems with long-range memory. *J. Phys. A: Math. Theor.*, **42** (34), 342 001.

60 Maes, C., Netočný, K., and Wynants, B. (2009) Dynamical fluctuations for semi-Markov processes. *J. Phys. A: Math. Theor.*, **42** (36), 365 002.