PAPER

Level 2 and level 2.5 large deviation functionals for systems with and without detailed balance

J Hoppenau, D Nickelsen and A Engel
University of Oldenburg, Institute of Physics, D-26111 Oldenburg, Germany
E-mail: johannes.hoppenau@uni-oldenburg.de

Keywords: large deviation theory, non-equilibrium steady states, stochastic thermodynamics

Abstract
Large deviation functions are an essential tool in the statistics of rare events. Often they can be obtained by contraction from a so-called level 2 or level 2.5 large deviation functional characterizing the empirical density and current of the underlying stochastic process. For Langevin systems obeying detailed balance, the explicit form of the level 2 functional has been known ever since the mathematical work of Donsker and Varadhan. We rederive the Donsker–Varadhan result using stochastic path-integrals. We than generalize the derivation to level 2.5 large deviation functionals for non-equilibrium steady states and elucidate the relation between the large deviation functionals and different notions of entropy production in stochastic thermodynamics. Finally, we discuss some aspects of the contractions to level 1 large deviation functions and illustrate our findings with examples.

1. Introduction

Thermodynamic quantities of small systems fluctuate measurably. Accordingly, the probability distributions of work, heat, and entropy are the focus of stochastic thermodynamics, for recent reviews see \cite{1, 2}. With the discovery of fluctuation theorems it became clear that the tails of these distributions are important for the characterization of small systems. Therefore, the statistics of rare events received an increased interest in the past decades.

The mathematical framework to address the statistics of rare events is large deviation theory \cite{3–5}. Of particular interest are quantities of the type

\[ a_T[x(t)] = \frac{1}{T} \int_{0}^{T} dt \, A(x(t)), \]  

(1.1)

sometimes called Brownian functionals \cite{6}, where \( x(t) \) denotes a \( d \)-dimensional stochastic process. If this process is ergodic with some stationary probability measure \( p_{st}(x) \), then, for large observation times \( T \), \( a_T \) will converge for almost all paths to its mean value

\[ \lim_{T \to \infty} a_T[x(\cdot)] = \langle a \rangle_{st} = \int dx \, p_{st}(x) A(x). \]  

(1.2)

To characterize the deviations of \( a_T \) from \( \langle a \rangle_{st} \), we write the distribution of \( a_T \) in the form

\[ p_T(a) \equiv p(a_T = a) = \exp \left\{ - T J(a) + o(T) \right\}, \]  

(1.3)

with the level 1 large deviation function

\[ J(a) = - \lim_{T \to \infty} \frac{1}{T} \ln p_T(a). \]  

(1.4)

If the limit exists, the random variable \( a_T \) is said to satisfy a large deviation principle \cite{5}.

The large deviation function \( J(a) \) contains the desired information about the statistics of \( a_T \). First of all, consistency requires \( J(\langle a \rangle_{st}) = 0 \) and \( J(a) \geq 0 \) for all \( a \). Taylor expansion around the minimum of \( J(a) \) up to second order yields a Gaussian probability distribution generalizing Einsteins theory of equilibrium fluctuations...
In addition, the complete function $J(a)$ characterizes the statistics of exponentially rare realizations $a_T$ that deviate substantially from $(a)_n$. Recent applications of large deviation functions to describe rare events in statistical mechanics can be found in $[8–14]$.

Different Brownian functionals deriving from the same stochastic process $x(t)$ have different large deviation functions. On the other hand, we may rewrite (1.1) as

$$a_T = \frac{1}{T} \int_0^T dt \, A(x(t)) \int dy \, \delta(y - x(s))$$

with the empirical density

$$\varrho_T(y; x(\cdot)) = \frac{1}{T} \int_0^T dt \, \delta(y - x(s)).$$

Due to its dependence on $x(\cdot)$ the empirical density is a random function. For an ergodic process we have

$$\lim_{T \to \infty} \varrho_T(y; x(\cdot)) = \rho_\text{st}(y).$$

If we assume that $\varrho_T$ by itself obeys a large deviation principle

$$P_T[\varrho(\cdot)] = P[\varrho_\text{st}(\cdot)] = \exp[-TI[\varrho(\cdot)] + o(T)],$$

with the large deviation functional

$$I[\varrho(\cdot)] = -\lim_{T \to \infty} \frac{1}{T} \ln P_T[\varrho(\cdot)],$$

all level 1 large deviation functions of Brownian functionals $a_T$ may be derived from $I[\varrho(\cdot)]$ by contraction:

$$I(a) = \min_{\varrho(\cdot)} \left\{ \int dy \, \lambda(y) \varrho(y) \right\}.$$

For this reason $I[\varrho(\cdot)]$ is known as level 2 large deviation functional.

To be more specific, let us consider as underlying stochastic process an overdamped Langevin dynamics

$$\dot{x} = f(x, t) + \frac{1}{\sqrt{\beta}} \xi(t),$$

with a force $f(x, t)$ and white noise $\xi(t)$ with correlation $\langle \xi_i(t) \xi_j(t') \rangle = 2 \delta_{ij} \delta(t - t')$. The dynamics may equivalently be described by the Fokker–Planck equation

$$\begin{align*}
\partial_t p(x, t) &= -\nabla \cdot f(x, t), \\
\dot{j}(x, t) &= f(x, t)p(x, t) - \frac{1}{\beta} \nabla p(x, t),
\end{align*}$$

where $p(x, t)$ is the probability density and $j(x, t)$ the probability current density for the variable $x$ at time $t$. Accordingly, we get for the stationary state with distribution $\rho_\text{st}$ and current $j_\text{st}$

$$\begin{align*}
0 &= -\nabla \cdot j_\text{st}(x), \\
j_\text{st}(x) &= f(x)p_\text{st}(x) - \frac{1}{\beta} \nabla p_\text{st}(x).
\end{align*}$$

If the stochastic process defined by (1.11) satisfies detailed balance, $j_\text{st}(x) \equiv 0$, Donsker and Varadhan derived quite some time ago $[15–18]$ the following explicit expression for the large deviation functional

$$I[\varrho(\cdot)] = \int dx \, \rho_\text{st}(x) \left[ \nabla \ln \frac{\varrho(x)}{\rho_\text{st}(x)} \right]^2,$$

Note $I[\varrho = \rho_\text{st}] = 0$ and $I[\varrho(\cdot)] \geq 0$ for all $\varrho(\cdot)$ as it should be.

This result is remarkable, yet due to the restriction to detailed balance it is limited to equilibrium situations. In stochastic thermodynamics non-equilibrium steady states (NESS) for which detailed balance is broken are of particular interest. It is hence natural to look for generalizations of the Donsker–Varadhan result to these more general forms of steady states.
In addition, in stochastic thermodynamics often quantities of the form
\begin{equation}
I = \frac{1}{T} \int_0^T dt \, \bar{x}(t) B(x(t)),
\end{equation}
like the entropy production, are of interest. Here the Stratonovich convention is used. Using the empirical current
\begin{equation}
\mu_T(y, x) = \frac{1}{T} \int_0^T dt \, \dot{x}(t) \delta(y - x(t))
\end{equation}
these quantities can be rewritten as
\begin{equation}
b_T(x) = \int dy \, B(y) \mu_T(y; x).
\end{equation}

Hence, if the joint large deviation functional
\begin{equation}
I[\varrho(\cdot), \mu(\cdot)] = -\lim_{T \to \infty} \frac{1}{T} \ln P_T[\varrho(\cdot), \mu(\cdot)]
\end{equation}
for the joint probability distribution \(P_T[\varrho(\cdot), \mu(\cdot)]\) for density and current together is known, the large deviation function of \(b_T\) can be derived. Since
\begin{equation}
\mu_T(y, x) \bigg| \varrho(y) = \mu(y)
\end{equation}
we have
\begin{equation}
I[\varrho(\cdot), \mu(\cdot)] = \min_{\mu(\cdot)} I[\varrho(\cdot), \mu(\cdot)].
\end{equation}

The level 2 large deviation functional can be obtained from \(I[\varrho(\cdot), \mu(\cdot)]\) by the contraction
\begin{equation}
I[\varrho(\cdot)] = \min_{\mu(\cdot)} I[\varrho(\cdot), \mu(\cdot)].
\end{equation}

A level 3 large deviation functional would be on trajectory level, \(I[x(\cdot)]\), from which by contraction \(I[\varrho(\cdot)]\) and \(I[\varrho(\cdot), \mu(\cdot)]\) can be derived. Since \(I[\varrho(\cdot), \mu(\cdot)]\) is an intermediate level between \(I[x(\cdot)]\) and \(I[\varrho(\cdot)]\), it was dubbed level 2.5 [14]. For the above contraction, however, no closed form for NESS is known. For the level 2.5 large deviation functionals, in contrast, closed expressions that are valid for equilibrium states and NESS, can be found in the mathematical and the mathematical physics literature [14, 19–22]. Closely related are the results of Bertini et al on large deviation functionals in macroscopic fluctuation theory [23, 24].

The Donsker–Varadhan result (1.15) may be re-written as
\begin{equation}
I[\varrho(\cdot)] = \frac{\beta}{4} \int dx \, \varrho(x) \left( \frac{j_{b}(x)}{\varrho(x)} \right)^2,
\end{equation}
with the probability current \(j_{b}(x)\) deriving from the empirical density, i.e. replacing \(\varrho\) with \(\varrho\) in the (1.12b). From the definitions (1.12b) and (1.13b) we find
\begin{equation}
\frac{j_{b}(x)}{\varrho(x)} = \frac{j_{a}(x)}{p_a(x)} - \frac{1}{\beta} \nabla \ln \frac{\varrho(x)}{p_a(x)},
\end{equation}
and since \(j_{a} = 0\) for a system with detailed balance, (1.15) and (1.23) coincide.

The explicit result for \(I[\varrho(\cdot), \mu(\cdot)]\) was found by Maes et al [20, 25] to be of the form
\begin{equation}
I[\varrho(\cdot), \mu(\cdot)] = \frac{\beta}{4} \int dx \, \varrho(x) \left( \frac{\mu(x)}{\varrho(x)} - \frac{j_{a}(x)}{\varrho(x)} \right)^2,
\end{equation}
in nice generalization of (1.23).

The aim of the present paper is twofold. We first rederive in sections 2 and 3 respectively the central results (1.15) and (1.25) using stochastic path-integrals. We hope that these derivations make the results more easily accessible to a physicist’s readership than the more mathematically minded original expositions [15–18, 20]. We then also elucidate the relation between the large deviation functional for the empirical density in stationary states of Langevin systems, the structure of the corresponding Fokker–Planck equation, and in section 4 the entropy production in a NESS. In this section we show that \(I[\varrho(\cdot), \mu(\cdot)]\) can be split into two contributions \(I_a[\mu(\cdot)]\) and \(I_{na}[\varrho(\cdot), \mu(\cdot)]\) linked to adiabatic and non-adiabatic entropy productions \(\dot{S}_a\) and \(\dot{S}_{na}\) [26–29]. Moreover, we consider in some detail the contraction of \(I[\varrho(\cdot), \mu(\cdot)]\) to \(I[\varrho(\cdot)]\) and \(I[\mu(\cdot)]\) respectively and exemplarily discuss the contraction to \(\int \dot{s}_a\) where the adiabatic entropy production \(s_a\) is a typical instance of \(b_T\) in (1.18). Finally, in section 5, we illustrate our findings with some examples.
2. Systems with detailed balance

For systems with detailed balance the external force derives from a potential, \( f(x) = -\nabla V(x) \), and the stationary current is zero, \( \dot{J}_n(x) = 0 \). The stationary distribution is the equilibrium distribution \( p_n(x) = 1/Z \exp(-\beta V(x)) \) with partition sum \( Z \). We can hence write the force as

\[
f(x) = \frac{1}{\beta} \nabla \ln p_n(x)
\]  

(2.1)

and may define the dynamics by fixing \( p_n \) instead of \( f \).

To derive the Donsker–Varadhan result (1.15) for the large deviation functional \( I[\varphi(t)] \), we write \( P_T[\varphi(t)] \) as the probability transformation of the probability density \( P[x(t)] \) for observing a trajectory \( x(t) \):

\[
P_T[\varphi(t)] = \int Dx(x)P(x(t)] \delta[\varphi(t) - \varphi_T(x(t))].
\]  

(2.2)

Using the integral representation of the \( \delta \)-functional, \( P_T[\varphi(t)] \) may be written as

\[
P_T[\varphi(t)] = \int Dq(q)\exp \left\{ -T \int dy \varphi(y)q(y) \right\} Q_T[q(t)],
\]  

(2.3)

with the cumulant generating function

\[
Q_T[q] = \int Dx(x)P_T[x(0)] \exp \left\{ -T \int dy \varphi_T(y; x(t))q(y) \right\}
\]  

\[
= \int Dx(x)P_T[x(0)] \exp \left\{ -\int_0^T dt q(x(t)) \right\}.
\]  

(2.4)

Like \( P_T[\varphi(t)] \), for large \( T \), \( Q_T[q(t)] \) can be written in a large deviation form:

\[
Q_T[q] = \exp \left\{ -T\lambda[q] + o(T) \right\}
\]  

(2.5)

with \( \lambda[q] = - \lim_{T \to \infty} 1/T \ln Q_T[q] \). Using this form of \( Q_T[q] \) as well as (1.8) in (2.3) gives

\[
\exp \left\{ -T I[\varphi(t)] + o(T) \right\} = \int Dq(q) \exp \left\{ -T \left[ \lambda[q] - \int dx \varphi(x)q(x) \right] + o(T) \right\}.
\]  

(2.6)

For large \( T \) the integral is dominated by its saddle-point and we find

\[
I[\varphi(t)] = \min_{q(t)} \left( \lambda[q] - \int dx \varphi(x)q(x) \right).
\]  

(2.7)

Hence, \( -\lambda[q(t)] \) is the Legendre–Fenchel transform of \( -I[\varphi(t)] \). Note that this course of action corresponds to the application of the Gärtner–Ellis theorem [5, 30, 31]. Our strategy will be to determine \( \lambda[q(t)] \) first and than solve the minimization problem (2.7) to obtain \( I[\varphi(t)] \).

The probability density of a trajectory is given by

\[
P_T[x(\cdot)] = p(x_0)P[x(\cdot)|x_0],
\]  

(2.8)

where \( p(x_0) \) is the probability density of the initial point \( x_0 = x(0) \) and \( P[x(\cdot)|x_0] \) the conditional probability density of the trajectory \( x(\cdot) \),

\[
P[x(t)|x_0] = \exp \left\{ -\int_0^T dt \left[ \frac{\beta}{4} (x(t) + \nabla V(x(t)))^2 - \frac{1}{2} \Delta V(x(t)) \right] \right\}.
\]  

(2.9)

Expanding the square in the exponent, the mixed term

\[
\int_0^T dt \frac{\beta}{2} \nabla V(x(t)) = \frac{\beta}{2} [V(x(T)) - V(x_0)]
\]  

(2.10)

gives rise to a boundary term that, for large \( T \), contributes to the \( o(T) \) terms only. We may hence write

\[
P[x(t)|x_0] = \exp \left\{ -\int_0^T dt \left[ \frac{\beta}{4} k^2(t) + U(x(t)) \right] + o(T) \right\}
\]  

(2.11)

with \( U(x) = \beta/4 [\nabla V(x)]^2 - \frac{1}{2} \Delta V(x) \).
Inserting (2.11) into (2.4) gives
\[
Q_T [q(y)] = \int dx_0 \, p(x_0) \int dx_T \, G_q(x_T, T|x_0, 0)
\]
(2.12)
with
\[
G_q(x_T, T|x_0, 0) = \int_{(x_0, 0)}^{(x_T, T)} \mathcal{D}x(\cdot) \exp \left[ -\frac{\beta}{4} x^2(t) + U(x(t)) + q(x(t)) \right].
\]
(2.13)
Even for simple choices of \( V(x) \) this path-integral can not be solved for general \( q(\cdot) \). Using the Feynman–Kac formula we may, however, obtain information on its large \( T \) behavior from the differential equation
\[
\partial_t G_q(x, t|x_0, 0) = L_q G_q(x, t|x_0, 0)
\]
(2.14)
for the time evolution of \( G_q(x, t|x_0, 0) \). The initial condition is \( G_q(x, 0|x_0, 0) = \delta(x - x_0) \) and \( L_q \) is the so-called tilted generator of the Fokker–Planck dynamics
\[
L_q = \frac{1}{\beta} \Delta - U(x) - q(x).
\]
(2.15)
To determine \( \lambda[q(\cdot)] \), we express the solution of (2.14) in terms of the eigenvectors and eigenfunctions of \( L_q \). We denote the right eigenvectors by \( \phi^r_n \) and the corresponding eigenvalues by \( \lambda^r_n \):
\[
L_q \phi^r_n = \lambda^r_n \phi^r_n.
\]
(2.16)
Owing to the symmetry of \( L_q \), the left eigenvectors are simply given by \( \phi^l_n = \phi^r_n \ast \):
\[
L_q \phi^l_n = \lambda^l_n \phi^l_n \ast,
\]
(2.17)
where \( \ast \) denotes the complex conjugate. At this point it is worth noting, that—although the generator of the Fokker–Planck dynamics is Hermitian—\( L_q \) is only Hermitian if \( q \in \mathbb{R} \). Assuming that \( L_q \) has a complete set of left and right eigenvectors\(^1\), i.e.
\[
\sum_n \phi^r_n(x) \phi^l_n(y) = \delta(x - y),
\]
(2.18)
we can formally write the solution of the differential equation (2.14) for arbitrary \( q(\cdot) \) as
\[
G_q(x_T, T|x_0, 0) = e^{T L_q} \delta(x_T - x_0) = e^{T L_q} \sum_n \phi^r_n(x_T) \phi^l_n(x_0)
\]
\[
= \sum_n e^{T \lambda^l_n} \phi^r_n(x_T) \phi^l_n(x_0).
\]
(2.19)
Inserting the above expression for \( G_q(x_T, T|x_0, 0) \) into (2.12) we get
\[
Q_T [q(y)] = \sum_n e^{T \lambda^l_n} \int dx_0 \, p(x_0) \phi^l_n(x_0) \int dx_T \phi^r_n(x_T).
\]
(2.20)
Comparing with (2.5), we find for \( T \to \infty \)
\[
\lambda[q(\cdot)] = -\lambda^l_q,
\]
(2.21)
where \( \lambda^l_q \) is the eigenvalue with the largest real part. For simplicity, we denote in the following the right eigenvector corresponding to this eigenvalue by \( \phi_q^r \). The two integrals in (2.20) contribute to the \( o(T) \) term in (2.5) only. Hence we do not need to know their value and can carry on without actually knowing the eigenvectors \( \phi_q^r \).

Still, \( \lambda[q(\cdot)] \) is hard to get for general \( q(\cdot) \). However, we do not need an explicit expression. The Euler–Lagrange equation corresponding to the minimization problem (2.7) is of the form
\[
\left. \frac{\delta \lambda[q(\cdot)]}{\delta q(x)} \right|_{q = \tilde{q}} = \varrho(x),
\]
(2.22)
where \( \varrho(\cdot) \) denotes the \( q(\cdot) \) that minimizes the rhs of (2.7). Now, from standard first order perturbation theory, we have

\(^1\) For a treatment without this assumption see appendix.
\[
\lambda[q(\cdot) + \delta q(\cdot)] - \lambda[q(\cdot)] = -\int dx \, \phi_q(x)(L_q + \delta_q - L_q)\phi_q(x) + O(\|\delta q\|^2)
\]
(2.23)

which leads to

\[
\phi_q^2(x) = \varphi(x).
\]
(2.24)

Since \(\varphi(x)\) is positive, \(\phi_q\) is real and the left and right eigenvectors of \(L_q\) that corresponds to the largest eigenvalue \(\lambda_q^0\) are the same.

To finally determine \(I[\varphi(\cdot)]\), we plug the minimizing \(\phi_q = \sqrt{\varphi}\) into (2.7) and find

\[
I[\varphi(\cdot)] = \lambda[q(\cdot)] - \int dx \, \sqrt{\varphi(x)} [L_q + q(x)]\sqrt{\varphi(x)}
\]
\[
= -\int dx \, \sqrt{\varphi(x)} \left[ \frac{1}{\beta} \Delta - U(x) \right] \sqrt{\varphi(x)}.
\]
(2.25)

Fortunately, the so far undetermined \(q(\cdot)\) drops out and we arrive at the result (1.15) found by Donsker and Varadhan:

\[
I[\varphi(\cdot)] = -\int dx \, \sqrt{\varphi(x)} \left[ \frac{1}{\beta} \Delta - \frac{\beta}{4} (\nabla V(x))^2 + \frac{1}{2} \Delta V(x) \right] \sqrt{\varphi(x)}
\]
\[
= \int dx \left[ \frac{1}{\beta} (\nabla \sqrt{\varphi(x)})^2 + \frac{\beta}{4} (\nabla V(x))^2 \varphi(x) + \frac{1}{2} \nabla V(x) \nabla \varphi(x) \right]
\]
\[
= \frac{\beta}{4} \int dx \, \varphi(x) \left[ \nabla \frac{\varphi(x)}{\beta} + \nabla V(x) \right]^2
\]
\[
= \frac{\beta}{4} \int dx \, \varphi(x) \left[ \frac{1}{\beta} \nabla \ln \frac{\varphi(x)}{p_n(x)} \right]^2.
\]
(2.26)

Being normalized, \(\varphi(x)\) tends to zero for \(|x| \to \infty\) and the boundary terms in the integration by parts do not contribute.

### 3. Systems without detailed balance

We now consider systems in which the force \(f(x)\) does not derive from a potential. Consequently, it is \(j_{st}(x) \neq 0\) and detailed balance is violated. Nevertheless, we can still relate the force to the stationary distribution \(p_{st}\) and the stationary current \(j_{st}\) by

\[
f(x) = \frac{j_{st}(x)}{p_{st}(x)} + \frac{1}{\beta} \nabla \ln p_{st}(x).
\]
(3.1)

As it is a hard problem to find \(p_{st}\) and \(j_{st}\) from (1.13a) and (1.13b) for a given force \(f\), it often is convenient to follow the reverse strategy and fix \(p_{st}\) and \(j_{st}\) and then determine the associated force \(f\) from the above equation.

In this way, \(p_{st}\) defines the conservative and \(j_{st}\) the non-conservative part of \(f\).

In the following analysis we closely follow the lines in section 2. The essential difference to the case with detailed balance becomes apparent by inspecting the probability density of a trajectory

\[
P_x = \exp \left\{ -\int_0^T dt \left[ \frac{\beta}{4} (\dot{x}(t) - f(x(t)))^2 + \frac{1}{2} \nabla f(x(t)) \right] \right\}.
\]
(3.2)

The mixed term \(\int_0^T dt \dot{x}(t)f(x(t))\) is no longer negligible in the large \(T\) limit but may be of order \(T\). Therefore, the simplification applied in (2.11) can not be used here. This entails a more complicated form of the tilted generator which in turn implies that there is no such simple relation between the right and left eigenvectors as (2.17).

The fact that we now have \(j_{st}\) in addition to \(p_{st}\) to define the dynamics, motivates to include the empirical current \(\mu\) defined in (1.17) into our considerations and to investigate the form of the joint probability density \(P_{f}[\varphi(\cdot), \mu(\cdot)]\). The corresponding large deviation form reads
The cumulant generating function of $P_T[\rho(\cdot), \mu(\cdot)]$ now depends on two functions and is given by

$$Q_T[q(\cdot), k(\cdot)] = \int \mathcal{D}x(x) P_T[x(\cdot)] \exp \left[ -T \int dy \ q(y) \rho(y; x(\cdot)) - T \int dy \ k(y) \mu(y; x(\cdot)) \right]$$

$$= \int \mathcal{D}x(x) P_T[x(\cdot)] \exp \left[ -\int_0^T dt \ [q(x(t)) + \dot{x}(t)k(x(t))] \right].$$

The corresponding large deviation functional $\lambda[q(\cdot), k(\cdot)]$ is defined by

$$Q_T[q(\cdot), k(\cdot)] = \exp \left[ -T \lambda[q(\cdot), k(\cdot)] + o(T) \right].$$

At this point, it is convenient to substitute $q(\cdot)$ and $k(\cdot)$ by two new functions $\eta(\cdot)$ and $\gamma(\cdot)$ such that

$$q(\gamma(x), \eta(x)) = -\frac{1}{2} \nabla \gamma(x) - \frac{\beta}{2} f(x) \gamma(x) + \frac{\beta}{4} \gamma^2(x) + \eta(x),$$

$$k(\gamma(x), \eta(x)) = \frac{\beta}{2} \gamma(x).$$

In analogy with (2.12) we have

$$Q_T[\gamma(\cdot), \eta(\cdot)] = \int dx_0 \ p(x_0) \int dx_T \ G_{\gamma,\eta}(x_T, T|x_0, 0),$$

where now

$$G_{\gamma,\eta}(x_T, T|x_0, 0) = \int_{(x_0,0)}^{(x_T,T)} \mathcal{D}x(x) \exp \left\{ -\int_0^T dt \left[ \frac{\beta}{4} [\dot{x}(t) - f(x(t)) + \gamma(x(t))]^2 + \frac{1}{2} \nabla [f(x(t)) - \gamma(x(t))] + \eta(x(t))] \right\}.$$

Here, we have used (3.2) and (3.4). The time evolution of $G_{\gamma,\eta}(x_T, T|x_0, 0)$ is determined by

$$\partial_t G_{\gamma,\eta}(x, t|x_0, 0) = L_{\gamma,\eta} G_{\gamma,\eta}(x, t|x_0, 0)$$

with the tilted generator

$$L_{\gamma,\eta} = \nabla \left[ -f + \gamma + \frac{1}{\beta} \nabla \right] - \eta.$$
To find $I[\varrho(\cdot), \mu(\cdot)]$ from $\lambda[\gamma(\cdot), \eta(\cdot)]$ we have to solve the minimization problem

$$I[\varrho(\cdot), \mu(\cdot)] = \min_{q(\cdot), k(\cdot)} \left\{ \lambda[q(\cdot), k(\cdot)] - \int dx \varrho(x)q(x) - \int dx \mu(x)k(x) \right\}$$

$$= \min_{\gamma(\cdot), \eta(\cdot)} \left\{ \lambda[\gamma(\cdot), \eta(\cdot)] - \int dx \left\{ \frac{\beta}{2} \gamma(x) \left[ \frac{\nabla \varrho(x)}{\beta} + \mu(x) - f(x) \varrho(x) \right] + \frac{\beta}{4} \gamma^2(x) \varrho(x) + \eta(x) \varrho(x) \right\} \right\}.$$  

(3.18)

Here, we integrated by parts to change $(\nabla \gamma) \varrho$ into $-\gamma \nabla \varrho$, where, due to the normalizability of $q(x)$, the boundary terms do not contribute.

Analogously to the minimization in $q(\cdot)$ for the detailed balance case, the minimization in $\eta(\cdot)$ leads in the present case to

$$\psi^*_{\gamma,\eta}(x) \phi_{\gamma,\eta}(x) = \varrho(x).$$

(3.19)

For the optimization in $\gamma(\cdot)$ it is convenient to use the relation

$$\lambda[\gamma(\cdot), \eta(\cdot)] = -\int dx \left[ L^+_{\gamma,\eta} \psi_{\gamma,\eta}(x) \right] \phi_{\gamma,\eta}(x).$$

(3.20)

With $L^+_{\gamma,\eta} - L^+_{\gamma,\eta} = -\delta \gamma^* \nabla$ and making the substitution $\phi_{\gamma,\eta} = \varrho / \psi^*_{\gamma,\eta}$, the minimization in $\gamma(\cdot)$ leads to

$$\nabla \psi^*_{\gamma,\eta} = \psi^*_{\gamma,\eta} \frac{\beta}{2} A,$$

(3.21)

where

$$A := \frac{1}{\beta} \frac{\nabla \varrho}{\varrho} + \frac{\mu}{\varrho} - f + \dot{\gamma}.$$  

(3.22)

We are now able to put everything together and get from (3.18)

$$I[\varrho(\cdot), \mu(\cdot)] = \int dx \left\{ -\left[ L^+_{\gamma,\eta} \psi_{\gamma,\eta} \right] \phi_{\gamma,\eta} - \frac{\beta}{2} \gamma \varrho A + \frac{\beta}{4} \varrho \gamma^2 - \dot{\eta} \varrho \right\}$$

$$= \int dx \left\{ \frac{\varrho}{\psi^*_{\gamma,\eta}} \left[ A - \frac{\nabla \varrho}{\beta \varrho} - \frac{\mu}{\varrho} - \frac{1}{\beta} \nabla \varrho A - \frac{\beta}{2} \psi^*_{\gamma,\eta} - \frac{\beta}{2} \gamma \varrho A + \frac{\beta}{4} \varrho \gamma^2 \right] \right\}$$

$$= \int dx \frac{\beta}{2} \left\{ \frac{\varrho}{\beta} \nabla^2 - \frac{\nabla \varrho}{\beta} + \mu \right\} A - \frac{\varrho}{\beta} \nabla \varrho A - \frac{1}{2} \varrho \nabla^2 - \gamma \varrho A + \frac{1}{2} \varrho \gamma^2 \right\}$$

$$= \frac{\beta}{4} \int dx \varrho \left[ \frac{\nabla \varrho}{\beta \varrho} + \frac{\mu}{\varrho} - f \right] - \int dx \mu \nabla \ln \psi^*_{\gamma,\eta}.$$  

(3.23)

Unlike the detailed balance case, the dependence on $\dot{\gamma}$ and $\dot{\eta}$ does not vanish completely but survives in the last term involving $\psi^*_{\gamma,\eta}$. However, this term contributes to the $o(T)$ terms in (3.5) only. To prove this, we first integrate by parts:

$$\int dx \mu(x) \nabla \ln \psi^*_{\gamma,\eta}(x) = -\int dx \ln \psi^*_{\gamma,\eta}(x) \nabla \mu(x).$$

(3.24)

Next we recall the definition (1.17) of $\mu$ for an arbitrary trajectory $\{y(\cdot)\}$ and find

$$\nabla \mu(x) = \frac{1}{T} \int_0^T dt \left. y \nabla \delta(x - y(t)) \right|_{y=0} = \frac{1}{T} \left[ \delta(x - y(0)) - \delta(x - y(T)) \right]$$

(3.25)

which in turn implies

$$\int dx \mu \nabla \ln \psi^*_{\gamma,\eta} = \frac{1}{T} \left[ \ln \psi^*_{\gamma,\eta}(y(T)) - \ln \psi^*_{\gamma,\eta}(y(0)) \right] = o(T).$$

(3.26)

We are thus left with

$$I[\varrho(\cdot), \mu(\cdot)] = \frac{\beta}{4} \int dx \varrho(x) \left[ \frac{1}{\beta} \nabla \ln \varrho(x) + \frac{\mu(x)}{\varrho(x)} - f(x) \right]^2.$$  

(3.27)
\[ I[\varrho(\cdot), \mu(\cdot)] = \int dx \varrho(x) \left[ \frac{\mu(x)}{\varrho(x)} - \frac{j_a(x)}{p_a(x)} \right] \cdot \nabla \ln \frac{\varrho(x)}{p_a(x)} + \frac{1}{\beta} \nabla \ln \frac{\varrho(x)}{p_a(x)} \right]^2. \] (3.28)

Upon repeated integrations by parts, the mixed term of the square can be shown to vanish

\[ I_{\text{mix}}[\varrho(\cdot), \mu(\cdot)] = \int dx \varrho(x) \left[ \frac{\mu(x)}{\varrho(x)} - \frac{j_a(x)}{p_a(x)} \right] \cdot \nabla \ln \frac{\varrho(x)}{p_a(x)} = 0. \] (3.29)

The integral over the boundary in the last line must vanish due to the normalizability of the involved distributions. Thus, we arrive with the final result

\[ I[\varrho(\cdot), \mu(\cdot)] = \frac{\beta}{4} \int dx \varrho(x) \left[ \frac{\mu(x)}{\varrho(x)} - \frac{j_a(x)}{p_a(x)} \right] \cdot \nabla \ln \frac{\varrho(x)}{p_a(x)} + \frac{1}{\beta} \nabla \ln \frac{\varrho(x)}{p_a(x)} \right]^2. \] (3.30)

As for the Donsker–Varadhan result, by substituting the current \( j_p \) from (1.24) into (3.28), we obtain the compact form \[14, 20\]

\[ I[\varrho(\cdot), \mu(\cdot)] = \frac{\beta}{4} \int dx \varrho(x) \left[ \frac{\mu(x)}{\varrho(x)} - j_p(x) \right]^2. \] (3.31)

4. Discussion

The large deviation functionals derived in the previous section characterize the distributions of empirical density and empirical current. They acquire a more intuitive meaning due to their relation to different forms of entropy production. In the first part of the discussion we establish these connections. We then consider the question how to obtain the large deviation functionals \( I[\varrho(\cdot)] \) of the empirical density alone and \( I[\mu(\cdot)] \) of the empirical current alone for a non-equilibrium steady state. Finally, we discuss the large deviation function \( f(s_a) \) of the adiabatic entropy production \( s_a[\mu(\cdot)] \).

4.1. Entropy production

We start with the simpler case of systems obeying detailed balance. The stationary current is zero, \( j_a(x) \equiv 0 \), and therefore, the dynamics is uniquely defined by the stationary distribution \( p_a \) alone. If the system is initially out of equilibrium in a state \( p = p_a \), it relaxes to equilibrium, and entropy is being produced with a rate \[2\]

\[ S_I[\varrho(\cdot)] = \beta \int dx p(x) \left[ \frac{j_p(x)}{p(x)} \right]^2. \] (4.1)

Here, \( j_p(x) \) denotes the current corresponding to \( p(x) \) according to

\[ j_p(x) = -\frac{1}{\beta} p(x) \nabla \ln \frac{p(x)}{p_a(x)}. \] (4.2)

Comparison with the Donsker–Varadhan result (1.23) reveals \[20, 25\]

\[ I[\varrho(\cdot)] = \frac{1}{4} S_I[\varrho(\cdot)], \] (4.3)
i.e., the large deviation functional $I[\varphi(\cdot)]$ is, up to a constant factor, nothing but the entropy production of the system when being in the state $\varphi(\cdot)$ instead of the equilibrium state $p_a$. For large $T$, the entropy production hence quantifies how unlikely a deviation of the empirical density from the true equilibrium distribution is.

For systems without detailed balance it has been shown \cite{26-29} that the entropy production rate $S_t$ may be subdivided into two contributions, the so-called adiabatic and non-adiabatic parts, $S_t = S_a + S_{\text{na}}$, where

$$\dot{S}_a[p(\cdot)] = \beta \int \mathrm{d}x \ p(x) \left[ \frac{j_a(x)}{p_a(x)} \right]^2,$$

$$\dot{S}_{\text{na}}[p(\cdot)] = \beta \int \mathrm{d}x \ p(x) \left[ \frac{j_{\text{na}}(x)}{p(x)} - \frac{j_a(x)}{p_a(x)} \right]^2.$$

The non-adiabatic part $S_{\text{na}}$ is the generalization of (4.1) and describes the entropic cost to relax to the stationary state. Correspondingly, it vanishes for $j_a(x) = j_a(x)$ and $p(x) = p_a(x)$. The adiabatic part $S_a$, on the other hand, remains non-zero even in the steady state and characterizes the dissipation necessary to maintain stationarity away from equilibrium. The two contributions therefore address the two basic mechanisms that commonly go along with non-equilibrium situations: driving and out of equilibrium boundary conditions respectively \cite{26-29}.

Using (1.24) we may write the large deviation functional (3.31) in the form

$$I[\varphi(\cdot), \mu(\cdot)] = \frac{\beta}{4} \int \mathrm{d}x \ \varphi(x) \left[ \left( \frac{\mu(x)}{\varphi(x)} - \frac{j_a(x)}{p_a(x)} \right)^2 + \left( \frac{j_a(y)}{\varphi(y)} - \frac{j_a(y)}{p_a(y)} \right)^2 \right],$$

which suggests to split it into the two contributions

$$I_a[\varphi(\cdot), \mu(\cdot)] = \frac{\beta}{4} \int \mathrm{d}x \ \varphi(x) \left[ \left( \frac{\mu(x)}{\varphi(x)} - \frac{j_a(x)}{p_a(x)} \right)^2 \right],$$

$$I_{\text{na}}[\varphi(\cdot)] = \frac{\beta}{4} \int \mathrm{d}x \ \varphi(x) \left[ \left( \frac{j_{\text{na}}(y)}{\varphi(y)} - \frac{j_a(y)}{p_a(y)} \right)^2 \right],$$

such that

$$I[\varphi(\cdot), \mu(\cdot)] = I_a[\varphi(\cdot), \mu(\cdot)] + I_{\text{na}}[\varphi(\cdot)].$$

Comparing (4.7b) and (4.5) we find in analogy to (4.3)

$$I_{\text{na}}[\varphi(\cdot)] = \frac{1}{4} \dot{S}_{\text{na}}[\varphi(\cdot)].$$

The second term in $I[\varphi(\cdot), \mu(\cdot)]$, $I_{\text{na}}[\varphi(\cdot)]$, is hence (up to a constant factor) equal to the non-adiabatic part of the entropy production that accounts for deviations of $\varphi$ from $p_a$. Accordingly, it characterizes the difference between empirical and true stationary distribution and is independent of $\mu$.

The correspondence between $I_a[\varphi(\cdot), \mu(\cdot)]$ and $S_a[p(\cdot)]$ as defined in (4.4) is not quite as close. The reason is twofold. Firstly, $S_a[p(\cdot)]$ is a functional of $p$ alone and has no dependence on a current. Secondly, $I_a[\varphi(\cdot), \mu(\cdot)]$ must be zero for $\varphi = p_a$ and $\mu = j_a$, whereas $S_a[p(\cdot)]$ has to remain non-zero even in the steady state. To connect $S_a[p(\cdot)]$ with $I_a[\varphi(\cdot), \mu(\cdot)]$, we have hence (in addition to the prefactor $1/4$) to make the replacement $j_a/p_a \to j_a/p_a - \mu/\varphi$. The first contribution to $I[\varphi(\cdot), \mu(\cdot)]$, $I_a[\varphi(\cdot), \mu(\cdot)]$, hence measures the distance between the empirical and the true stationary current. Evaluated at $\mu = 0$ it gives (up to a constant factor) the adiabatic entropy production in state $\varphi$.

We note in passing that the above expressions may be formally simplified by introducing the local mean velocity $v = j/p$, see, e.g., \cite{2}, giving rise to $\nu = \mu/\varphi, v_a = j_a/\varphi, v_a = j_a/p_a$, and the large deviation functionals $I[\varphi(\cdot), \nu(\cdot)], I_a[\varphi(\cdot), \nu(\cdot)]$, and $I_{\text{na}}[\varphi(\cdot)]$. We then find, e.g., from $I_a[\varphi(\cdot), \nu(\cdot)] = 0$ immediately $\nu = v_a$, completely independent of $\varphi$.

4.2. Contractions

The large deviation functional $I[\varphi(\cdot), \mu(\cdot)]$ characterizes the joint probability distribution $P[\varphi(\cdot), \mu(\cdot)]$ for empirical density and empirical current. In many situations one is interested in the distribution of either the density or the current alone. These distributions are obtained by marginalization

$$P[\varphi(\cdot)] = \int \mathcal{D}\mu(\cdot) P[\varphi(\cdot), \mu(\cdot)], \quad P[\mu(\cdot)] = \int \mathcal{D}\varphi(\cdot) P[\varphi(\cdot), \mu(\cdot)].$$

If the involved probability distributions obey large deviation principles, marginalization naturally transforms into contraction \cite{5}, i.e.
\[ I[\varrho(\cdot)] = \min_{\mu} I[\varrho(\cdot), \mu(\cdot)], \quad I[\mu(\cdot)] = \min_{\varrho} I[\varrho(\cdot), \mu(\cdot)]. \] (4.11)

We start with the determination of \( I[\varrho(\cdot)] \). To find \( I[\varrho(\cdot)] = I[\varrho(\cdot), \mu(\cdot)] \), we have to determine the optimal current \( \bar{\mu}(\varrho) \) that minimizes \( I[\varrho(\cdot), \mu(\cdot)] \) for every empirical distribution \( \varrho \). Since \( I[\varrho(\cdot), \mu(\cdot)] \) depends on \( \mu \) solely via \( I_{\mu}[\varrho(\cdot), \mu(\cdot)] \), we need to minimize \( I_{\mu} \) only. In view of (3.24)–(3.26), this minimization has to be done under the constraint \( \nabla \cdot \mu = 0 \):

\[ I[\varrho(\cdot)] = \min_{\mu}[\nabla \mu = 0] I[\varrho(\cdot), \mu(\cdot)]. \] (4.12)

Including the constraint with a Lagrange multiplier field \( \kappa(x) \) yields the Euler–Lagrange equation

\[ \varrho(x) \left( \frac{\bar{\mu}(x)}{\varrho(x)} - \frac{j_{st}(x)}{P_a(x)} \right) \frac{1}{\varrho(x)} = \frac{\nabla \kappa(x)}{\varrho(x)} = 0. \] (4.13)

The optimal current can hence be determined from the equations

\[ \bar{\mu}(x; \varrho) = \frac{\varrho(x)}{P_a(x)} j_{st}(x) + \varrho(x) \nabla \kappa(x), \] (4.14a)

\[ \Delta \kappa(x) = -\frac{j_{st}(x)}{P_a(x)} \cdot \nabla \ln \frac{\varrho(x)}{P_a(x)} - \nabla \kappa(x) \cdot \nabla \ln \varrho(x), \] (4.14b)

where the second line follows from \( \nabla \cdot \bar{\mu} = 0 \). The resulting large deviation functional reads

\[ I[\varrho(\cdot)] = I_{\text{sw}}[\varrho(\cdot)] + \frac{\beta}{4} \int dx \varrho(x) \left[ \nabla \kappa(x) \right]^2. \] (4.15)

For systems with detailed balance we have \( j_{st}(x) \equiv 0 \), and \( \kappa \equiv 0 \) solves (4.14b). From (4.15) we then recover the Donsker–Varadhan result in the form (1.23). Remarkably, in this case the optimal current is zero for all \( \varrho \).

For \( j_{st}(x) \neq 0 \) the determination of \( I[\varrho(\cdot)] \) remains a challenging problem.

We continue with the determination of \( I[\mu(\cdot)] \). Now we need the optimal density \( \bar{\varrho} \) that minimizes \( I[\varrho(\cdot), \mu(\cdot)] \):

\[ I[\mu(\cdot)] = \min_{\varrho} I[\varrho(\cdot), \mu(\cdot)] = I[\bar{\varrho}(\cdot), \mu(\cdot)]. \] (4.16)

The Euler–Lagrange equation for \( \bar{\varrho} \) reads

\[ \left( \frac{\bar{\mu}}{\bar{\varrho}} \right)^2 - \left( \frac{j_{st}}{P_a} \right)^2 + \frac{1}{\beta^2} \left( \nabla \ln \bar{\varrho} \right)^2 = \kappa, \] (4.17)

where \( \kappa \) is the Lagrange parameter to be determined from the constraint \( \int \varrho(x, \kappa) dx = 1 \). Solving the above Euler–Lagrange equation analytically is known to be difficult.

However, as discussed below (1.18), we stress that \( I[\mu(\cdot)] \) is particularly relevant for contractions to large deviation functions of entropy productions. Consider for example the adiabatic entropy production \( s_s[x(\cdot)] \) on the trajectory level [2, 32]

\[ s_s[x(\cdot)] = \beta \int_0^T dt \dot{x}(t) \cdot \frac{j_{st}(x(t))}{P_a(x(t))}. \] (4.18)

Note that as opposed to the spatial average \( \bar{S}_s \) in (4.4), \( s_s[x(\cdot)] \) is a fluctuating quantity and satisfies a fluctuation theorem, see [2, 10] for details. Identifying \( B(y) = j_{st}(y)/P_a(y) \) in (1.18), we can rewrite \( s_s[x(\cdot)] \) as

\[ s_s[\mu(\cdot)] = \beta T \int dy \mu(y) \cdot \frac{j_{st}(y)}{P_a(y)}. \] (4.19)

We now use \( s_s[\mu(\cdot)] \) as an example to illustrate the contraction to large deviation functions. Having performed the contraction in (4.16) to obtain \( I[\mu(\cdot)] \), the subsequent contraction

\[ J(s_s) = \min_{\mu} \left| \nabla \mu = 0, s_s = s_s[\mu] \right| I[\mu(\cdot)] \] (4.20)

would yield the large deviation function \( J(s_s) \) for the adiabatic entropy production. But, due to the nonlinearity of the Euler–Lagrange equation (4.17) for the contraction to \( I[\mu(\cdot)] \), analytical progress in finding \( J(s_s) \) is notoriously difficult. Solving the Euler–Lagrange equation (4.17) numerically, however, is feasible, but the resulting parametrical dependence of \( \bar{\rho} \) and \( \kappa \) on \( \mu \) hinders the formulation of the Euler–Lagrange equation for the above contraction and complicates the determination of \( J(s_s) \) even in a numerical treatment.

An alternative strategy, where some progress can be made, is to flip the order of contractions and first perform...
and the Lagrange parameter \( r \) and \( \mu \) from \( s_a = s_a[\mu] \) from (4.20). The optimal current hence reads
\[
\bar{\mu}(x) = (1 - \kappa_2) \frac{\varphi(x)}{p_a(x)} \bar{j}_a(x) + \varphi(x) \nabla \kappa_1(x)
\] (4.23)
which substituted into \( I_s[\varphi(\cdot), \mu(\cdot)] \) yields
\[
I_s[\varphi(\cdot), s_a] = \frac{\beta}{4} \int \! dx \, \varphi(x) \left[ \nabla \kappa_1(x) - \kappa_2 \frac{\bar{j}_a(x)}{p_a(x)} \right]^2.
\] (4.24)
The Lagrange function \( \kappa_1 \) and the Lagrange parameter \( \kappa_2 \) have to be determined from the equations for the respective constraints,
\[
0 = \Delta \kappa_1(x) + \nabla \kappa_1(x) \cdot \nabla \ln \varphi(x) + (1 - \kappa_2) \frac{\bar{j}_a(x)}{p_a(x)} \cdot \nabla \ln \varphi(x)
\] (4.25a)
\[
s_a = T \beta \int \! dy \, \varphi(y) \frac{j_a(y)}{p_a(y)} \cdot \left[ (1 - \kappa_2) \frac{j_a(y)}{p_a(y)} + \nabla \kappa_1(y) \right].
\] (4.25b)
Substituting the empirical adiabatic entropy production rate \( \dot{S}_a[\varphi(\cdot)] \) from (4.4), we can rewrite (4.25b) as
\[
\frac{s_a}{T} = \frac{\beta}{\kappa_2} \left[ \dot{S}_a[\varphi(\cdot)] - \frac{s_a}{T} \right] + \frac{T \beta}{\kappa_2} \int \! dy \, \varphi(x) \left[ (\nabla \kappa_1)^2 - \kappa_2 \nabla \kappa_1 \cdot \nabla \kappa_1 \right]
\] (4.26)
and find from substituting (4.23) and (4.27)
\[
I_s[\varphi(\cdot), s_a] = \frac{\kappa_2}{4} \left( \dot{S}_a[\varphi(\cdot)] - \frac{s_a}{T} \right) = \frac{\beta}{4} \int \! dy \, \varphi(x) \left[ \frac{\bar{\mu} - \bar{j}_a}{p_a} \right]
\] (4.27)
where the second line follows after two integrations by parts. The remaining quantities to be determined are \( \kappa_1(x; s_a) \) and \( \kappa_2(s_a) \) from (4.25a) and (4.25b), and \( \varphi(\kappa) \) from (4.17) where \( \mu = \bar{\mu}(s_a) \) needs to be substituted from (4.23) and \( \kappa \) follows from the constraint \( \int \! \varphi(x) \kappa \, dx = 1 \). Hence, by first contracting to \( I[\varphi(\cdot), s_a] \) and then to \( I[\mu(\cdot)] \) and then to \( I(s_a) \), all necessary equations to determine \( I(s_a) \) are known and a numerical treatment is possible.

5. Illustration

To illustrate our findings, we first consider a simple example without detailed balance given by the force
\[
f(r) = -ar e_r + bre_\phi
\] (5.1)
with the constants \( a > 0 \) and \( b \) and \( x \) being parameterized with \( r \) and \( \phi \) in two-dimensional polar coordinates. The stationary distribution and the stationary current are
\[
p_a(r) = \frac{\beta a}{2\pi} \exp \left[ -\beta \Phi(r) \right], \quad \Phi(r) = \frac{a}{2} r^2,
\] (5.2a)
\[
j_a(r) = p_a(r) v_a(r), \quad v_a(r) = bre_\phi
\] (5.2b)
and we can thus also write
\[
f(r) = \frac{1}{\beta} \frac{\partial}{\partial r} \ln p_a(r) e_r + \frac{j_a(r)}{p_a(r)},
\] (5.3)
\[
= -\partial_r \Phi(r) e_r + v_a(r).
\] (5.4)
The conservative force \( -\partial_r \Phi(r) \) points to the origin and keeps the dynamics bounded, the non-conservative force \( v_a(r) \) generates a circular current and violates the detailed balance condition.

For a numerical illustration, we fix the parameters to be \( a = 1, \ b = 2, \ T = 500, \ \beta = 2 \) and numerically solve the Langevin equation (1.11) with \( f(r) \) from (5.3) using \( 10^5 \) timesteps to obtain an ensemble of \( 10^5 \).
trajectories \( x(\cdot) \). The initial values \( x(0) \) are drawn from the stationary distribution (5.2a) in order to save the relaxation time and have the system in the NESS the whole time. To each trajectory, we determine the empirical density \( \rho (\cdot) \) and current \( \mu (\cdot) \) from the definitions (1.6) and (1.17) and calculate from that the values of \( I_\alpha \rho (\cdot), \mu (\cdot) \) and \( I_\alpha | \rho (\cdot), \mu (\cdot) | \) using (4.7a) and (4.7b).

In figure 1 we show a typical and a rare realization of \( \phi \), and in figure 2 a typical and a rare realization of \( \mu \). The typical realizations are selected by picking from the ensemble the \( \phi \) and \( \mu \) with the minimum value of \( I_\alpha \rho (\cdot), \mu (\cdot) \) respectively, and correspondingly, the rare realizations are qualified by the maximum value of \( I_\alpha \rho (\cdot), \mu (\cdot) \) within the ensemble. Note that in order to select the typical and rare realization of \( \mu \), we only need \( I_\alpha \rho (\cdot), \mu (\cdot) \). We see that the typical realization of \( \phi \) is close to \( p_{st} \), whereas the rare realization clearly deviates from \( p_{st} \). In contrast, both realizations of \( \mu \) are close to \( j_{st} \), implying that \( I_\alpha \rho (\cdot), \mu (\cdot) \) is much sharper with respect to \( \mu \) than to \( \phi \), that is, the system predominantly fluctuates with regard to position and barely with regard to direction of movement.

To consider a numerical example in which fluctuations of \( \mu \) play a role, we keep the stationary distribution (5.2a) but alter \( \nu_\alpha \),

\[
\nu_\alpha(r, \varphi) = br^3 \sin(\varphi) e_x + \frac{b}{\beta a} \left( r^2 + \frac{2}{\beta a} \frac{\cos(\varphi)}{r} \right) e_x,
\]

ensuring that \( \nabla \cdot (p_{st}(r) \nu_\alpha(r, \varphi)) = 0 \) still holds. To obtain an ensemble of \( \phi \) and \( \mu \) for this system, we repeat the simulation of the Langevin equation with the new force \( f(r, \varphi) \) according to (5.3) and \( \nu_\alpha(r, \varphi) \) from (5.5).
In figures 3 and 4 we show the typical and rare realizations of $\rho$ and $m$. Again, we see that the typical $\rho$ and $m$ are close to $p_{st}$ and $j_{st}$ respectively and the rare $\rho$ substantially deviates from $p_{st}$, but now also the rare $m$ clearly deviates from $j_{st}$. Hence, for this example, the system fluctuates in both position and direction of movement.

It should furthermore be interesting to perform the contractions discussed in the previous section. Due to the non-linearity of the Euler–Lagrange equations, however, analytical contractions are, even for the simple example (5.3), barely possible. We regard a numerical treatment of the contractions in section 4.2 as beyond the scope of this paper. Instead, we use the parametrization

$$\rho(x, y) = \rho_{st}(x, y), \quad \mu(x, y) = \mu_{st}(x, y)$$

for illustrative purposes. The empirical density is chosen such that for $\sigma = 1$ and $\beta = 2$, the system fluctuates in both position and direction of movement.

For this functional choice of $\rho$ and $\mu$, and $p_{st}$ and $j_{st}$ from (5.2a) and (5.2b), the rate function $I = I_{st} + I_{na}$ can be calculated analytically:
As expected, it is \( I[\sigma] = \sigma_{st}, c = b \) = 0.

First, we consider the contraction to \( I[\rho] \). In terms of the parametrization \((5.6a)\), we are to calculate \( I[\rho] \) by plugging in the optimal \( c \) minimizing \( I[\sigma, c] \). From \( I_a[\sigma, c] \) above we immediately find that \( \tilde{c}(\sigma) \equiv b \) and we are left with the non-adiabatic part

\[
I[\sigma] = I_a[\sigma, \tilde{c} = b] + I_{na}[\sigma] = I_{na}[\sigma].
\]

Next, to illustrate the contraction to \( I[\mu] \), we determine \( I[c] \). By minimizing \( I[\sigma, c] \) in \( \sigma \) we find the optimal \( \sigma \) to be

\[
\tilde{\sigma}(c)^2 = \frac{\sigma_{st}^2}{\sqrt{\beta^2 \sigma_{st}^2(b - c)^2 + 1}}.
\]

In figure 5 we plot \( I[\sigma, c] = I_a[\sigma, c] + I_{na}[\sigma] \) from \((5.7a)\) and \((5.7b)\) together with the line \( \tilde{\sigma}(c) \) from \((5.9)\) that minimizes \( I[\sigma, c] \) depending on the value of \( c \). The dependency of \( \tilde{\sigma}(c) \) on \( c \) implies that the optimal empirical density is not independent of the empirical current, or, in other words, having sampled an empirical current different from the stationary current, it is likely to also have sampled an empirical density deviating from the stationary distribution. Thus, we illustrate with this example that we can not just set \( I[\rho(c)] = I[\rho(c), \mu(c) \equiv I[\rho(c)] \) but have to perform the contraction \( I[\rho(c)] = I[\rho(c), \mu(c) \equiv I[\rho(c)] \) as done in section 4.2 to derive \( I[\rho(c)] \) for systems without detailed balance. To proceed with the contraction to \( I[c] \), we plug \( \sigma = \tilde{\sigma}(c) \) in \((5.7a)\) and \((5.7b)\) and get

\[
I_a[c] = I_a[\sigma = \tilde{\sigma}(c), c] = \frac{(b - c)^2}{2\sqrt{(b - c)^2 + 1}},
\]

\[
I_{na}[c] = I_{na}[\sigma = \tilde{\sigma}(c)] = \frac{\sqrt{(b - c)^2 + 1} - 1)^2}{2\sqrt{(b - c)^2 + 1}}.
\]
\[ I[c] = I_b[c] + I_m[c] = \sqrt{(b - c)^2 + 1} - 1. \]  
\[ (5.12) \]

Clearly, both functionals \( I_b[c] \) and \( I_m[c] \) are individually zero for \( b = c \).

Finally, we derive by contraction the large deviation function \( J(s_a) \) for this example. From (4.19) follows that
\[ s_a(b) = 2\beta T\sigma^2 b. \]  
\[ (5.13) \]

Substitution into (5.7a) gives in agreement with the first term in (4.27)
\[ I_a[\sigma, s_a] = \frac{(2\beta T\sigma^2 b^2 - s_a)^2}{8\beta T^2\sigma^2 b^2} = \left( \frac{S_a(\sigma) - s_a}{4S_a(\sigma)} \right)^2, \]  
\[ (5.14) \]

where in the second line we plugged in \( S_a(\sigma) = 2\beta\sigma^2b^2 \) from (4.4). In the last step we optimize \( I[\sigma, s_a] = I_a[\sigma, s_a] + I_{na}[\sigma] \) in \( \sigma \),
\[ \sigma(s_a)^2 = \sigma^2 \left( \frac{\bar{\sigma}^2}{S_a^2 + 4b^2} \right), \]  
\[ (5.15) \]

and find by substitution into \( I[\sigma, s_a] \) the desired large deviation function
\[ J(s_a) = I[\sigma(s_a), s_a] = \frac{1}{2} \left( S_a + \frac{4b^2}{S_a} \right) \left( \frac{s_a}{T} + \frac{4b^2}{S_a} \right) - \frac{1}{2} \left( \frac{s_a}{T} + \frac{4b^2}{S_a} \right). \]  
\[ (5.16) \]

As expected, it is \( J(s_a = TS_a) = 0 \), and we see that \( p(s_a) \) has exponential tails.

6. Conclusion

In the present paper we discussed a derivation of level 2 and level 2.5 large deviation functionals for Langevin systems using stochastic path-integrals. Quite naturally, our approach parallels in part the more mathematical one [14–18]; we also use a variant of a cumulant generating functional in (2.4) and (3.4), equations (2.7) and (3.18) are tantamount to the Gärtnert–Ellis theorem [30, 31], and the operators (2.15) and (3.11) are related to the tilted generators of the stochastic dynamics considered. Nevertheless, we hope that our derivation sheds some new light on the results discussed and may be more accessible to a less mathematically inclined reader.

We started with a rederivation of the classical Donsker–Varadhan result for the level 2 large deviation functional \( I[\varrho(\cdot)] \) for Langevin systems that obey detailed balance. The derivation makes use of a simple relation between the left and right eigenfunctions of the tilted Fokker–Planck operator. Due to the lack of such a relation in cases without detailed balance, the generalization of our derivation to NESS is not immediate. Similarly to previous studies [19, 20], the introduction of the empirical current \( \mu(\cdot) \) and consideration of \( I[\varrho(\cdot), \mu(\cdot)] \) instead of \( I[\varrho(\cdot)] \) provided the missing relation between the eigenfunction sets and allowed us to rederive the closed form of \( I[\varrho(\cdot), \mu(\cdot)] \) found by Maes et al [20] along other lines.

In our derivation it was important to note that the definition of the empirical current implies a vanishing divergence, \( \nabla \mu = 0 \), except for the start and end point of the underlying trajectory. This fact constitutes an argument why \( I[\varrho(\cdot), \mu(\cdot)] \) is to be considered for divergenceless \( \mu \) only, as also discussed in [20].

Just as the irreversible entropy production may be split into an adiabatic and a non-adiabatic part for systems violating detailed balance, the large deviation functional \( I[\varrho(\cdot), \mu(\cdot)] \) can be written as the sum of two non-negative contributions \( I_a[\varrho(\cdot), \mu(\cdot)] \) and \( I_{na}[\varrho(\cdot)] \), in direct correspondence to the respective entropy production rates. The empirical current \( \mu(\cdot) \) only enters the adiabatic part \( I_a[\varrho(\cdot), \mu(\cdot)] \).

The ‘master’ functional \( I[\varrho(\cdot), \mu(\cdot)] \) is meant as a starting point to derive other large deviation functionals and functions by contraction. However, apparently these contractions can rarely be done explicitly. Contraction to \( I[\varrho(\cdot)] \) is possible for detailed balance systems and reproduces the Donsker–Varadhan result with a vanishing optimal current \( \bar{\mu} \) for all \( \varrho \). Large deviation functions of entropy productions follow from contractions of \( I[\mu(\cdot)] \), so the large deviation functional for the empirical current alone is of great interest. The contraction from \( I[\varrho(\cdot), \mu(\cdot)] \) to \( I[\mu(\cdot)] \), however, is notoriously involved. For a numerical determination we have provided a set of equations for the large deviation function of the adiabatic entropy production, \( J(s_a) \), as an example. Carrying out the numerics for \( J(s_a) \) and analytical progress in the determination of \( I[\mu(\cdot)] \) are subjects for future efforts.
Acknowledgments

Financial support from DFG under EN278/9-1 is gratefully acknowledged. We thank Marcel Kahlen for fruitful discussions and Hugo Touchette for many helpful remarks.

Appendix. Degenerate eigenvalues

In general the operator \( L \) is not Hermitian and it is hence not guaranteed that \( L \) has a complete set of eigenfunctions. Here, we argue that the result that \( \lambda(Lq(\cdot)) \) can be identified with the negative of the largest eigenvalue of the operator \( Lq(\cdot) \) (and \( \gamma(\cdot), \eta(\cdot) \) respectively) also holds if \( L \) is not diagonalizable. For notational simplicity we omit the dependency of \( L \) on \( q(\cdot), g(\cdot) \) and \( h(\cdot) \) and use the Einstein summation convention.

We assume, that \( G(x, t|x_0, 0) \) can be written as linear combination of \( N \) orthogonal functions \( f_n(x) \):

\[
G(x, t|x_0, 0) = g_n(t|x_0, 0)f_n(x)
\]

(A.1)

with

\[
g_n(t|x_0, 0) = \int dx f_n^*(x)G(x, 0|x_0, 0).
\]

(A.2)

Using this form of \( G \), the differential equations (2.14) and (3.10) turn into

\[
\partial_t g_n(t) = L_{m,n}g_n(t)
\]

(A.3)

with the matrix

\[
L_{m,n} = \int dx f_m^*(x)Lf_n(x).
\]

(A.4)

The formal solution of this differential equation is given by

\[
g_n(T|x_0, 0) = (e^{TT})_{m,n}g_n(0|x_0, 0)
\]

(A.5)

with the initial condition \( g_n(0|x_0, 0) = \int dx f_n^*(x)\delta(x - x_0) = f_n^*(x_0) \). Inserting this into (2.12) or (3.8) gives for the cumulant generating function

\[
Q_T = \int dx_T \int dx_0 p(x_0)f_n(x_T)(e^{TT})_{m,n}f_m^*(x_0).
\]

(A.6)

Let us now introduce a similarity transformation \( L' = HJH^{-1} \). For the matrix exponential, this leads to

\[
e^{tt} = H e^{TT} H^{-1}.
\]

If \( L' \) has \( N \) different eigenvalues \( \lambda_m \), we chose the columns of \( H \) to be the eigenvectors of \( L' \). Then \( J \) is a diagonal matrix with the eigenvalues on the diagonal. This leads to \( (e^{TT})_{m,n} = \delta_{m,n}e^{\lambda_m} \). If we put this into (A.6) and further identify \( \phi^\nu = f_m(x)H_{m,n} \) and \( \psi^\nu = H_{n,m}f_n^*(x) \) we recover (2.20) and (3.16) respectively.

If \( L' \) has only \( k < N \) different eigenvalues \( \lambda_m \) that are of algebraic multiplicity \( s_m \), we chose the columns \( Y_{(1),1}\) \( \ldots \) \( Y_{(1),s_1} \) \( \ldots \) \( Y_{(k),1}\) \( \ldots \) \( Y_{(k),s_k} \) of \( H \) to be the generalized eigenvectors of \( L' \), i.e. \( Y_{(m,i)} \in \ker(L - \lambda_m id)^i \). In analogy, we denote the rows of \( H^{-1} \) by \( z_{(1),1}\) \( \ldots \) \( z_{(1),s_1}\) \( \ldots \) \( z_{(k),1}\) \( \ldots \) \( z_{(k),s_k} \). This choice of \( H \) allows \( J \) to be of Jordan normal form:

\[
J = \left( \begin{array}{ccc}
\hat{h} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \hat{h}
\end{array} \right)
\]

(A.7)

where the Jordan blocks \( f_m \) are \( s_m \times s_m \) matrices, with \( \lambda_m \) on the diagonals and ones on the superdiagonals. For the exponential, this gives

\[
e^J = \left( \begin{array}{ccc}
e^{\hat{h}} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & e^{\hat{h}}
\end{array} \right)
\]

(A.8)

where the block matrices are given by

\[
(e^{\hat{h}})_{ij} = \begin{cases} 
0 & \text{if } i > j, \\
\lambda^{i-j} & \text{otherwise}.
\end{cases}
\]

(A.9)
Inserting all this into (A.6) gives

\[ Q_T = \sum_{n=1}^{k} c_{n} \int dx T \int dx_0 \pi_{0}(x_0) \sum_{m=1}^{N} \sum_{l=0}^{M} \phi_{n,m}(x_T) \frac{T}{l!} \psi_{n,m+l}(x_0), \]

(A.10)

with \( \phi_{n,m} = f_j(x_T) f_{j_l}(n,m) \) and \( \psi_{n,m+l} = z_{n,m+l} f_{j_l}(x_0) \). Hence, like in the case with \( N \) different eigenvalues, for large \( t \), \( Q_T \) is dominated by \( e^{\lambda_0 t} \), where \( \lambda_0 \) is the eigenvalue with the largest real part and the integrals over \( x_T \) and \( x_0 \) are part of the \( \alpha(T) \) term in (2.5) and (3.5).

References

[1] Jarzynski C 2011 Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale Ann. Rev. Condens. Matter Phys. 2 329–51
[2] Seifert U 2012 Stochastic thermodynamics, fluctuation theorems and molecular machines Rep. Prog. Phys. 75 126001
[3] Hollander F D 2008 Large Deviations (Providence, RI: American Mathematical Society)
[4] Derrida B 2007 Non-equilibrium steady states: fluctuations and large deviations of the density and of the current J. Stat. Mech. P07023–P07023
[5] Touchette H 2009 The large deviation approach to statistical mechanics Phys. Rep. 478 1–69
[6] Majumdar S N 2005 Brownian functionals in physics and computer science Curr. Sci. 89 2076–92
[7] Landau L D and Lifshitz E M 1980 Statistical Physics (Course of Theoretical Physics vol 5) (Oxford: Butterworth–Heinemann)
[8] Lebowitz J L and Spohn H 1999 A Gallavotti-Cohen type symmetry in the large deviation functional for stochastic dynamics J. Stat. Phys. 95 333–65
[9] Bodineau T and Derrida B 2007 Cumulants and large deviations of the current through non-equilibrium steady states C. R. Phys. 8 540–55
[10] Speck T, Engel A and Seifert U 2012 Large deviation function for the entropy production: optimal trajectory and role of fluctuations J. Stat. Mech. 2012 P12001
[11] Nemoz T and Sasa S 2014 Computation of large deviation statistics via iterative measurement–and–feedback procedure Phys. Rev. Lett. 112 090602
[12] Verley G, Willaert T, Van den Broeck C and Esposito M 2014 The unlikely Carnot efficiency Nat. Commun. 5 4721
[13] Harbola U, Van den Broeck C and Lindeberg K 2014 Large deviation function and fluctuation theorem for classical particle transport Phys. Rev. E 89 012114
[14] Barato A C and Chetrite R 2015 A formal view on level 2.5 large deviations and fluctuation relations J. Stat. Phys. 44 18
[15] Donsker M D and Varadhan S R 1975 Asymptotic evaluation of certain Markov process expectations for large time: I Commun. Pure Appl. Math. 28 1–47
[16] Donsker M D and Varadhan S R 1975 Asymptotic evaluation of certain Markov process expectations for large time: II Commun. Pure Appl. Math. 28 279–301
[17] Donsker M D and Varadhan S R 1983 Asymptotic evaluation of certain Markov process expectations for large time: III Commun. Pure Appl. Math. 29 389–461
[18] Donsker M D and Varadhan S R 1983 Asymptotic evaluation of certain Markov process expectations for large time: IV Commun. Pure Appl. Math. 36 183–212
[19] Kesten G and Maller J 1993 Large deviations for a general class of random vectors IEEE Trans. Inf. Theory 39 10561057
[20] Maes C, Netočný K and Wynants B 2008 Steady state statistics of driven diffusions Physica A 387 2675–89
[21] Chernyak V, Chertkov S, Malinin S and Teodorescu R 2009 Non-equilibrium thermodynamics and topology of currents J. Stat. Phys. 137 109–47
[22] Bertini L, Faggionato A and Gabrielli D 2015 Large deviations of the empirical flow for continuous time Markov chains Ann. Inst. Henri Poincaré Probab. Stat. 51 867–900
[23] Bertini L, De Sole A and Gabrielli D 2002 Macroscopic fluctuation theory for stationary non-equilibrium states J. Stat. Phys. 107 635–75
[24] Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2009 Current fluctuations in stochastic lattice gases Phys. Rev. Lett. 94 030601
[25] Wynants B 2010 Structure of nonequilibrium fluctuations dissipation and activity PhD Thesis Katholieke Universiteit Leuven
[26] Esposito M, Harbola U and Mukamel S 2007 Entropy fluctuation theorems in driven open systems: application to electron counting statistics Phys. Rev. E 76 1–7
[27] Esposito M and Van den Broeck C 2010 Three detailed fluctuation theorems Phys. Rev. Lett. 104 090601
[28] Esposito M and Van den Broeck C 2010 Three faces of the second law: I Master equation formulation Phys. Rev. E 82 011143
[29] Van den Broeck C and Esposito M 2010 Three faces of the second law: II Fokker–Planck formulation Phys. Rev. E 82 011144
[30] Gärtner J 1977 On large deviations from the invariant measure Theory Probab. Appl. 22 24–39
[31] Ellis R S 1984 Large deviations for a general class of random vectors Ann. Probab. 12 1–2
[32] Speck T and Seifert U 2005 Integral fluctuation theorems for the housekeeping heat J. Phys. A: Math. Gen. 38 L581–8