JOURNAL OF MATHEMATICAL PHYSICS **48**, 053306 (2007)

Minimum entropy production principle from a dynamical fluctuation law

Christian Maes[a)]
Instituut voor Theoretische Fysica, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium

Karel Netočný[b)]
Institute of Physics, Academy of Sciences of the Czech Republic, 18221 Prague, Czech Republic

(Received 22 January 2007; accepted 19 April 2007; published online 31 May 2007)

The minimum entropy production principle provides an approximative variational characterization of close-to-equilibrium stationary states, both for macroscopic systems and for stochastic models. Analyzing the fluctuations of the empirical distribution of occupation times for a class of Markov processes, we identify the entropy production as the large deviation rate function, up to leading order when expanding around a detailed balance dynamics. In that way, the minimum entropy production principle is recognized as a consequence of the structure of dynamical fluctuations, and its approximate character gets an explanation. We also discuss the subtlety emerging when applying the principle to systems whose degrees of freedom change sign under kinematical time reversal. © *2007 American Institute of Physics.* [DOI: 10.1063/1.2738753]

I. INTRODUCTION

Over the past century many attempts have been made to give a variational characterization of nonequilibrium conditions. The motivation was often found in the successes of variational methods in mechanics and in equilibrium statistical thermodynamics. Many so called *ab initio* methods in solid state physics have a variational character. For nonequilibrium purposes the best known but also widely criticized variational principle, that of the minimum entropy production principle (MinEP), goes back to the work of Prigogine.[18] In the present paper we will restrict to the version of the MinEP for Markov processes as was first described and proven by Klein and Meijer for some specific Markov models, see also Refs. 9, 5, and 15.

As has been clear since a long time, the MinEP is only valid in some approximation.[8] Without doubt, one restriction is that the system must be close to equilibrium, allowing only for a small breaking of the detailed balance condition; that is often referred to as the regime of irreversible thermodynamics.[6] Yet, the situation is more subtle and there have appeared examples in the literature violating the MinEP even close to equilibrium.[10,11] The situation is even more complicated and downright controversial when dealing with examples of macroscopic physics, where both positions and velocities mix and things appear to depend on the level of coarse graining.

At any rate and because of the enormous advantages of variational characterizations, there has been a continued interest in the nature of Prigogine's MinEP. It remains therefore very interesting to see if the principle can be understood not only by direct verification as was done in Refs. 9, 5, and 15 but also from the context of fluctuation theory. After all, also in equilibrium statistical physics there is an intimate relation between the variational principle characterizing equilibrium and the structure of equilibrium fluctuations. The very reason why thermodynamic equilibrium is

[a)]Electronic mail: christian.maes@fys.kuleuven.be
[b)]Electronic mail: netocny@fzu.cz

characterized by maximum entropy or, depending on the context, by minimum Helmholtz or Gibbs free energy is exactly because these thermodynamic potentials also appear as rate functions in the exponents governing equilibrium probabilities.

We show in this paper that a relation exists between the MinEP and the structure of steady state fluctuations for Markov processes. Our main finding is that the entropy production naturally emerges when analyzing the fluctuations of occupation times, first studied in the general context of the theory of large deviations by Donsker and Varadhan.[4] We show that in the close-to-equilibrium regime and when the state variable is even under time reversal, the Donsker-Varadhan (DV) functional coincides to leading order with the entropy production rate. When the state variables are odd under time reversal, such as for the electric current in the famous counterexample of Refs. 10 and 11, that affine relation between entropy production and the DV functional is no longer valid. It remains of course generally true that the variational principle associated with the DV functional is a valid generalization of the MinEP. Yet, a useful scheme for the computation of the DV functional for processes far from equilibrium remains an open problem.

The structure of the paper is as follows. In Sec. II we present a brief introduction to the large deviation theory of occupation times. In the mathematical details we often restrict ourselves to the case of continuous time and irreducible Markov processes on a finite state space. Many of the arguments have, however, a larger validity. For example, for a detailed balanced dynamics the DV functional can be computed explicitly; we review that result in Sec. III with a simple proof that is not restricted to finite state spaces.

Our main result follows from a perturbative evaluation of the DV functional close to equilibrium and is contained in Sec. IV, first on a formal and general level and then rigorously for finite state space. In Sec. V we explain how and when the leading order of the DV functional is related to the physical entropy production. That relation is formulated in our main Theorem 5.1.

We end with a variety of remarks and conclusions in Sec. VII. In particular, we briefly explain there the situation for Landauer's counterexample.[10,11]

II. LARGE DEVIATIONS FOR THE OCCUPATION TIMES

Suppose that $(X_t)_{t\geqslant 0}$ is a stationary ergodic Markov process. For most of what follows, we do not need to specify whether it is a jump process or a diffusion process, and on what space. Yet, it is sufficiently instructive and mathematically nontrivial to keep in mind a Markov process on a finite space which is irreducible. We are interested in the fraction of time that X_t spends in some set A of states. Formally, we define the empirical distribution p_T as

$$p_T(A) = \frac{1}{T}\int_0^T \mathrm{d}t\,\delta_{X_t\in A} \tag{2.1}$$

($\delta_{X_t\in A}=1$ if $X_t\in A$ and zero otherwise). As we assume a unique stationary measure ρ, we have that almost surely $p_T(A)\to\rho(A)$ as $T\uparrow+\infty$, by ergodicity. Yet there are fluctuations around that average and we can ask how big they are. That is a subject in the theory of large deviations and the answer is given by the asymptotic formula

$$\boldsymbol{P}^T[p_T\simeq\mu]\simeq\exp[-TI(\mu)]$$

that has to be understood in a logarithmic sense after taking the limit $T\uparrow\infty$. The rate function I has been found by Donsker and Varadhan[4,2,7,3] in the form

$$I(\mu) = \sup_{g>0} -\left\langle \frac{Lg}{g}\right\rangle_\mu, \tag{2.2}$$

where L is the generator of the Markov process and $\langle\cdot\rangle_\mu$ denotes the expectation under the measure μ. For a finite state space Ω,

$$Lg(x) = \sum_{y \in \Omega} k(x,y)[g(y) - g(x)], \quad x \in \Omega,$$

where $k(x,y) \geqslant 0$ is the rate for the transition $x \to y$.

The DV functional is always non-negative, $I(\mu) \geqslant 0$, and the equality takes place if and only if $\mu = \rho$ is the invariant measure. Further, it becomes infinite on measures that are not absolutely continuous with respect to ρ. For the precise mathematical formulation we refer to Refs. 2, 7, 4, and 3.

In general, the DV functional (2.2) is not so simple to compute explicitly, the main problem of course being to find the maximizer g. An important case where the solution has been known and is explicit is a reversible (or detailed balance) dynamics. These basic facts are reviewed in the next section, see also Refs. 4, 7, and 3. The rest of the paper is then devoted to identifying the leading term in the DV functional for a dynamics breaking the detailed balance.

III. DETAILED BALANCE DYNAMICS

Suppose that for any pair of real-valued functions ϕ and ψ,

$$\langle \phi(x_0)\psi(x_\tau)\rangle_\rho = \langle \phi(x_\tau)\psi(x_0)\rangle_\rho, \tag{3.1}$$

where $\langle\cdot\rangle_\rho$ is the expectation under the stationary Markov process. The corresponding symmetry of the generator can be obtained under

$$\lim_{\tau \downarrow 0} \frac{1}{\tau}\langle \phi(x_0)[\psi(x_\tau) - \psi(x_0)]\rangle_\rho = \langle \phi L\psi\rangle_\rho.$$

Theorem 3.1: *Under condition (3.1), the DV functional is*

$$I(\mu) = -\langle \sqrt{f} L \sqrt{f}\rangle_\rho, \tag{3.2}$$

where $f = d\mu/d\rho$ is the density of μ with respect to the reversible measure ρ. If μ is not absolutely continuous with respect to ρ then $I(\mu) = +\infty$.

Remark 3.1: One recognizes the Dirichlet form $\mathcal{D}(g,g) = -\langle gLg\rangle_\rho$, which is related to the spectral gap Δ by

$$\Delta = \inf_{g:\langle g\rangle_\rho = 0} \frac{\mathcal{D}(g,g)}{\langle g^2\rangle_\rho}. \tag{3.3}$$

As a consequence, one has the bound

$$I(\mu) = \mathcal{D}(\sqrt{f},\sqrt{f}) = \mathcal{D}(\sqrt{f} - \langle\sqrt{f}\rangle_\rho, \sqrt{f} - \langle\sqrt{f}\rangle_\rho) \geqslant \Delta[\langle f\rangle_\rho - \langle\sqrt{f}\rangle_\rho^2] = \Delta[1 - \langle\sqrt{f}\rangle_\rho^2]. \tag{3.4}$$

Proof: A standard proof for finite state space can be found, e.g., in Refs. 3 and 4. Here we present a simple variant of that argument that works for a general (detailed balanced) Markov process.

From Eq. (2.2),

$$I(\mu) = \sup_{g>0} \lim_{\tau \downarrow 0} \frac{1}{\tau}\left[1 - \left\langle \frac{e^{\tau L} g}{g}\right\rangle_\mu\right] = \sup_{g>0} \lim_{\tau \downarrow 0} \frac{1}{\tau}\left[1 - \left\langle \frac{f(x_0)g(x_\tau)}{g(x_0)}\right\rangle_\rho\right]. \tag{3.5}$$

Using reversibility (3.1) we subsequently get

$$\left\langle \frac{f(x_0)g(x_\tau)}{g(x_0)} \right\rangle_\rho = \frac{1}{2}\left\langle \frac{f(x_0)g(x_\tau)}{g(x_0)} + \frac{f(x_\tau)g(x_0)}{g(x_\tau)} \right\rangle_\rho = \frac{1}{2}\left\langle \left(\sqrt{\frac{f(x_0)g(x_\tau)}{g(x_0)}} - \sqrt{\frac{f(x_\tau)g(x_0)}{g(x_\tau)}} \right)^2 \right\rangle_\rho + \langle \sqrt{f(x_0)f(x_\tau)} \rangle_\rho \geqslant \langle \sqrt{f(x_0)f(x_\tau)} \rangle_\rho = \langle \sqrt{f} e^{\tau L} \sqrt{f} \rangle_\rho, \tag{3.6}$$

which is an optimal lower bound since the equality is attained if (and only if for an irreducible dynamics) $g \propto \sqrt{f}$. Hence,

$$I(\mu) = \lim_{\tau \downarrow 0} \frac{1}{\tau}[1 - \langle \sqrt{f} e^{\tau L} \sqrt{f} \rangle_\rho] = -\langle \sqrt{f} L \sqrt{f} \rangle_\rho \tag{3.7}$$

as claimed. □

IV. PERTURBATIVE EVALUATION OF THE DONSKER-VARADHAN FUNCTIONAL

A. Formal derivation

Fix a reference detailed balance dynamics with generator L_0 and with reference measure ρ^0, as in Sec. III. For a measure μ we write $f = \mathrm{d}\mu/\mathrm{d}\rho^0$ for its density with respect to ρ^0. A simple computation gives

$$\delta \left\langle \frac{f}{g} L g \right\rangle_{\rho^0} = \left\langle -\frac{f}{g^2} \delta g L g + \frac{f}{g} L \delta g \right\rangle_{\rho^0} = \left\langle \left(-\frac{f}{g^2} L g + L^+ \frac{f}{g} \right) \delta g \right\rangle_{\rho^0}, \tag{4.1}$$

where the adjoint L^+ is defined by

$$\langle \phi L \psi \rangle_{\rho^0} = \langle \psi L^+ \phi \rangle_{\rho^0} \tag{4.2}$$

on real functions. Hence, searching for the maximizer g^* of Eq. (2.2) normalized to $\langle g^* \rangle_{\rho^0} = 1$, we need to solve the equation

$$\frac{f}{g^{*2}} L g^* = L^+ \frac{f}{g^*}. \tag{4.3}$$

Note that for $L = L_0 = L_0^+$ that equation has a solution $g^* = \sqrt{f}/\langle \sqrt{f} \rangle_{\rho^0}$, in agreement with the conclusions of Sec. III.

Next, for a close-to-equilibrium dynamics and for small fluctuations we expand L, f, and g in power series

$$L^\epsilon = L_0 + \epsilon L_1 + \epsilon^2 L_2 + \cdots, \tag{4.4}$$

$$f^\epsilon = 1 + \epsilon f_1 + \epsilon^2 f_2 + \cdots, \tag{4.5}$$

$$g^\epsilon = 1 + \epsilon g_1 + \epsilon^2 g_2 + \cdots \tag{4.6}$$

and solve Eq. (4.3) perturbatively. Up to order ϵ it yields

$$2 L_0 g_1^* = L_0 f_1 + L_1^+ 1, \tag{4.7}$$

which is to be solved under the normalization constraint $\langle g_1^* \rangle_{\rho^0} = 0$. That can be achieved as follows. Writing $\mathrm{d}\rho^\epsilon/\mathrm{d}\rho^0 = h^\epsilon$ for the density of the (presumably unique for small ϵ) stationary measure under L^ϵ with respect to the reference reversible measure, the stationary equation $\rho^\epsilon L^\epsilon = 0$ can be equivalently written as $(L^\epsilon)^+ h^\epsilon = 0$. Expanding again $h^\epsilon = 1 + \epsilon h_1 + \cdots$, we find that h_1 verifies

$$L_0 h_1 = -L_1^+ 1 \tag{4.8}$$

and, by definition, $\langle h_1 \rangle_{\rho^0} = 0$. As a consequence, $g_1^* = (f_1 - h_1)/2$ is a solution of Eq. (4.7). Provided that g^* is, in fact, a global maximum, the DV functional (2.2) becomes, up to leading order,

$$I^\epsilon(\mu^\epsilon) = -\frac{\epsilon^2}{4}\langle f_1 L_0 f_1 - h_1 L_0 h_1 + 2L_1 f_1 - 2L_1 h_1 \rangle_{\rho^0} + o(\epsilon^2) = -\left\langle \sqrt{\frac{f^\epsilon}{h^\epsilon}} L^\epsilon \sqrt{\frac{f^\epsilon}{h^\epsilon}} \right\rangle_{\rho^0} + o(\epsilon^2). \tag{4.9}$$

The functional I^ϵ itself obviously also depends on ϵ as from Eq. (4.4); we are dealing with a dynamics close to a reference reversible dynamics. Observe that, since $f^\epsilon d\rho^\epsilon = h^\epsilon d\mu^\epsilon$, the leading term in the DV functional (4.9) (always for small deviations from equilibrium) resembles the DV functional (3.2) for the case of detailed balance. In Eq. (4.9) that leading term is now of order ϵ^2.

B. Rigorous result

The above formal perturbative argument can be justified on a mathematically precise level. In the present section we refine the above reasoning by restricting ourselves to the framework of continuous time Markov dynamics with a finite state space. Note that many of the standard nonequilibrium examples of stochastic lattice gases or interacting particle systems on a finite graph are thus included.[5,12,20] Observe also that some precision or justification is indeed needed, as one can otherwise construct counterexamples to the results that will follow. Other "infinite" or "continuous" models including diffusion processes still require additional estimates for a proper mathematical treatment that we are not giving here though; we will comment on one important example in Sec. VII. On the whole and perhaps surprisingly, even only to first order around equilibrium, a general and mathematically precise identification of the DV functional does not appear easy.

We fix a finite state space Ω, which will serve as vertex set for irreducible directed graphs, respectively, with rates $k^0(x,y) \geqslant 0$ (reference detailed balance) and with rates $k^\epsilon(x,y) \geqslant 0$ (perturbation) between the states $x \to y$. We assume that the reference rates $k^0(x,y)$ define an ergodic Markov process with the stationary distribution $\rho^0 > 0$ and such that $\rho^0(x)k^0(x,y) = \rho^0(y)k^0(y,x)$, sufficient for the reversibility in Eq. (3.1). The perturbed rates $k^\epsilon(x,y)$ defined for $|\epsilon| \leqslant \epsilon_0$ with some $\epsilon_0 > 0$ are assumed to be a smooth modification of the $k^0(x,y)$. For small enough ϵ the perturbed dynamics is hence ergodic too, with a unique invariant distribution $\rho^\epsilon > 0$ which is a smooth modification of ρ^0.

The modified dynamics has the generator

$$L^\epsilon g(x) = \sum_{x,y \neq x} k^\epsilon(x,y)[g(y) - g(x)]. \tag{4.10}$$

We further denote

$$M_{+1} = \{g > 0; \langle g \rangle_{\rho^0} = 1\}, \quad M_{+1}^\delta = \{g \in M_{+1}; g(x) \geqslant \delta,\ x \in \Omega\},$$

and we consider the functional

$$J_f^\epsilon(g) = \sum_{x,y \neq x} \rho^0(x) f(x) k^\epsilon(x,y) \left[1 - \frac{g(y)}{g(x)}\right], \tag{4.11}$$

for $f \in M_{+1}^\delta$ on $g \in M_{+1}$.

Proposition 4.1: *Suppose that* $f \in M_{+1}^\delta$ *for some* $\delta > 0$. *For all sufficiently small* $|\epsilon|$, *the functional* J_f^ϵ *has a unique maximizer* $g^{*\epsilon}(f)$ *in* M_{+1}, *and* $g^{*\epsilon}(f) \xrightarrow{\epsilon \downarrow 0} \sqrt{f}/\langle \sqrt{f} \rangle_{\rho^0}$, *uniformly in* M_{+1}^δ.

Theorem 4.1: *If* μ^ϵ *is a smooth deformation of* $\mu^0 = \rho^0$, *then the DV functional* $I^\epsilon(\mu^\epsilon)$ *has a*

Taylor expansion in ϵ *around* $\epsilon=0$, *with leading term*

$$I^{\epsilon}(\mu^{\epsilon}) = -\left\langle \sqrt{\frac{d\mu^{\epsilon}}{d\rho^{\epsilon}}} L^{\epsilon} \sqrt{\frac{d\mu^{\epsilon}}{d\rho^{\epsilon}}} \right\rangle_{\rho^0} + o(\epsilon^2). \tag{4.12}$$

The proofs are postponed to Sec. VI.

V. RELATION WITH ENTROPY PRODUCTION

We proceed with the *physical* interpretation of formula (4.12) for the DV functional. It will turn out that Eq. (4.12) equals the excess of entropy production with respect to the stationary entropy production. Clearly, to explain, we need some physical context for the dynamics itself. However, in order to avoid relying solely on concrete examples, we can start from the quite general observation that the physical entropy production as a variable on path space is measuring the breaking of time-reversal symmetry. That has been argued for at various places, see, e.g., Refs. 15 and 13 and references therein. When the distribution at time zero is given by μ, then the entropy production over the time interval $[0,\tau]$ is just the relative entropy of the path-space distribution $\boldsymbol{P}_{\mu}^{\tau}$ for the process started from μ, with respect to its time reversal:

$$\dot{S}^{\tau}(\mu) = \left\langle \log \frac{\mathrm{d}\boldsymbol{P}_{\mu}^{\tau}}{\mathrm{d}\boldsymbol{P}_{\mu_{\tau}}^{\tau}\Theta} \right\rangle_{\mu}, \tag{5.1}$$

where $(\Theta\omega)_t=\omega_{\tau-t}$ is the time reversal of the trajectory ω over the interval $[0,\tau]$, and μ_{τ} is the evolved distribution at time τ, i.e., the solution of the Master equation $\mathrm{d}\mu_t/\mathrm{d}t=\mu_t L$, $\mu_0=\mu$. Since the process is Markovian, the mean entropy production can be written as $\dot{S}^{\tau}(\mu)=\int_0^{\tau}\sigma(\mu_t)\mathrm{d}t$, where

$$\sigma(\mu) = \lim_{\tau\downarrow 0} \frac{\dot{S}^{\tau}(\mu)}{\tau} \tag{5.2}$$

is the mean entropy production rate. Taken as a functional on distributions μ, Eq. (5.2) defines the crucial quantity to be discussed in the present section. In particular, we can evaluate it under the same conditions as for Theorem 4.1. It means that we evaluate the entropy production rate in μ^{ϵ} and that we have a dynamics that is close to equilibrium, indicated by changing the notation σ to σ^{ϵ}. The main result of the paper is then summarized in the following general and remarkable relation.

Theorem 5.1: *Under the conditions of Theorem 4.1*,

$$I^{\epsilon}(\mu^{\epsilon}) = \frac{1}{4}[\sigma^{\epsilon}(\mu^{\epsilon}) - \sigma^{\epsilon}(\rho^{\epsilon})] + o(\epsilon^2).$$

Before we give the proof of that theorem, we briefly remind the reader of the physical context of entropy production, at least within the limited setup of Markov jump processes. We refer to Refs. 19, 17, 5, 15, and 9 for additional material.

A. Entropy production in Markov jump processes

For the Markov jump processes of Sec. IV B the entropy production rate (5.2) becomes

$$\sigma(\mu) = \sum_{x,y\neq x} \mu(x)k(x,y)\log\frac{\mu(x)k(x,y)}{\mu(y)k(y,x)}. \tag{5.3}$$

In the case of detailed balance, $\rho(x)k(x,y)=\rho(y)k(y,x)$, it is easily verified that $\sigma(\mu)$ is the time derivative of the relative entropy:

$$\sigma(\mu) = \sum_x \log \frac{\mu(x)}{\rho(x)} \sum_{y\neq x} [\mu(x)k(x,y) - \mu(y)\rho(y,x)] = -\sum_x \log \frac{\mu(x)}{\rho(x)} \left. \frac{\mathrm{d}\mu_t(x)}{\mathrm{d}t} \right|_{t=0}$$

$$= -\frac{\mathrm{d}}{\mathrm{d}t} S(\mu_t|\rho)|_{t=0}, \quad S(\mu|\rho) = \sum_x \mu(x)\log \frac{\mu(x)}{\rho(x)}. \tag{5.4}$$

When there is a driving away from equilibrium, there is some mean entropy production even in the stationary regime. To be specific, assume that each state x is given an energy $E(x)$ and that the transition $x \leftrightarrow y$ is possible thanks to the interaction with a heat reservoir at inverse temperature $\beta(x,y)=\beta(y,x)$. The rates are taken to satisfy the local detailed balance condition

$$\frac{k(x,y)}{k(y,x)} = \mathrm{e}^{\beta(x,y)[E(x)-E(y)]}. \tag{5.5}$$

For a motivation, see Refs. 5 and 15. As a reference we have the Boltzmann-Gibbs distribution $\rho(x) \propto \mathrm{e}^{-\beta E(x)}$, with β some reference inverse temperature.

Entropy production rate (5.3) can be split into a contribution which is associated with the system and can be written as the time derivative of some entropy function, and a part measuring the change of entropy in the environment, i.e.,

$$\sigma(\mu) = \sigma_S(\mu) + \sigma_R(\mu).$$

For the system part we take, compare with Eq. (5.4),

$$\sigma_S(\mu) = \sum_{x,y\neq x} \mu(x)k(x,y)\log \frac{\mu(x)\rho(y)}{\mu(y)\rho(x)} = -\frac{\mathrm{d}}{\mathrm{d}t} S(\mu_t|\rho)|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} [S(\mu_t) - \beta\langle E\rangle_{\mu_t}]|_{t=0}, \tag{5.6}$$

with $S(\mu)=-\Sigma_x \mu(x)\log \mu(x)$ the Shannon entropy and $\langle E\rangle_\mu = \Sigma_x \mu(x)E(x)$ the mean energy. Hence, $\sigma_S(\mu)$ is recognized as $(-\beta$ times) the rate of change in the free energy.

The environment part is then

$$\sigma_R(\mu) = \sum_{x,y\neq x} \mu(x)k(x,y)\log \frac{\rho(x)k(x,y)}{\rho(y)k(y,x)}$$

$$= \frac{1}{2} \sum_{x,y\neq x} [\beta(x,y) - \beta][E(x) - E(y)][\mu(x)k(x,y) - \mu(y)k(y,x)]$$

$$= \frac{1}{2} \sum_{x,y\neq x} [\beta(x,y) - \beta]\langle J_E(x,y)\rangle_\mu, \tag{5.7}$$

where $\langle J_E(x,y)\rangle_\mu = \langle J_E(y,x)\rangle_\mu$ is the mean energy transfer, or heat, to the reservoir associated with the transitions $x \leftrightarrow y$. In other words, $\sigma_R(\mu)$ is the change of entropy in the environment plus the term

$$\beta \sum_x E(x)[\mu(x)k(x,y) - \mu(y)k(y,x)] = \beta \frac{\mathrm{d}}{\mathrm{d}t} \langle E\rangle_{\mu_t}|_{t=0},$$

which is just the counterterm we have subtracted from the system part (5.6).

B. Proof of Theorem 5.1

Following our general strategy, we compute Eq. (5.2) by a perturbation expansion around a reference detailed balanced dynamics. Again, the expansion is mathematically fully justified for a finite state space, at least under the conditions of the theorem.

We split the entropy production rate similarly as in the previous section, taking now the invariant distribution ρ^0 corresponding to $\epsilon=0$ as the reference: starting from Eq. (5.2),

$$\sigma(\mu)=\lim_{\tau\downarrow 0}\frac{1}{\tau}\left\langle \log\frac{\mathrm{d}\mu}{\mathrm{d}\rho^0}(\omega_0)-\log\frac{\mathrm{d}\mu_\tau}{\mathrm{d}\rho^0}(\omega_\tau)+\log\frac{\mathrm{d}\boldsymbol{P}^\tau_{\rho^0}}{\mathrm{d}\boldsymbol{P}^\tau_{\rho^0}\Theta}\right\rangle_\mu=\lim_{\tau\downarrow 0}\frac{1}{\tau}\left[\left\langle \log\frac{\mathrm{d}\mu}{\mathrm{d}\rho^0}\right\rangle_\mu-\left\langle \log\frac{\mathrm{d}\mu_\tau}{\mathrm{d}\rho^0}\right\rangle_{\mu_\tau}\right]+\lim_{\tau\downarrow 0}\frac{1}{\tau}\left\langle \log\frac{\mathrm{d}\boldsymbol{P}^\tau_{\rho^0}}{\mathrm{d}\boldsymbol{P}^\tau_{\rho^0}\Theta}\right\rangle_\mu. \tag{5.8}$$

The first term is the limit

$$\sigma_S(\mu)=\lim_{\tau\downarrow 0}\frac{1}{\tau}\left[\left\langle \log\frac{\mathrm{d}\mu}{\mathrm{d}\rho^0}\right\rangle_\mu-\left\langle \log\frac{\mathrm{d}\mu}{\mathrm{d}\rho^0}\right\rangle_{\mu_\tau}-\left\langle \frac{\mathrm{d}\mu_\tau}{\mathrm{d}\mu}\log\frac{\mathrm{d}\mu_\tau}{\mathrm{d}\mu}\right\rangle_\mu\right]$$
$$=-\langle L\log f\rangle_\mu-\left\langle \frac{\mathrm{d}\mu L}{\mathrm{d}\mu}\right\rangle_\mu=-\langle L\log f\rangle_\mu. \tag{5.9}$$

Expanding both $\mu\equiv\mu^\epsilon$ and $L\equiv L^\epsilon$ as in Eqs. (4.4) and (4.5), we get

$$\sigma^\epsilon_S(\mu^\epsilon)=-\langle f^\epsilon L^\epsilon\log f^\epsilon\rangle_{\rho^0}=-\epsilon^2\langle f_1L_0f_1+L_1f_1\rangle_{\rho^0}+o(\epsilon^2). \tag{5.10}$$

Similarly, for the second term in Eq. (5.8), now denoted by σ^ϵ_R, we have

$$\sigma^\epsilon_R(\mu^\epsilon)=\lim_{\tau\downarrow 0}\frac{1}{\tau}\left\langle \log\frac{\mathrm{d}\boldsymbol{P}^{\tau,\epsilon}_{\rho^0}}{\mathrm{d}\boldsymbol{P}^{\tau,0}_{\rho^0}}(\omega)-\log\frac{\mathrm{d}\boldsymbol{P}^{\tau,\epsilon}_{\rho^0}}{\mathrm{d}\boldsymbol{P}^{\tau,0}_{\rho^0}}(\Theta\omega)\right\rangle^\epsilon_{\mu^\epsilon}$$
$$=\lim_{\tau\downarrow 0}\frac{1}{\tau}\left\langle \frac{\mathrm{d}\mu^\epsilon}{\mathrm{d}\rho^0}(\omega_0)\frac{\mathrm{d}\boldsymbol{P}^{\tau,\epsilon}_{\rho^0}}{\mathrm{d}\boldsymbol{P}^{\tau,0}_{\rho^0}}(\omega)\left\{\frac{\mathrm{d}\boldsymbol{P}^{\tau,\epsilon}_{\rho^0}}{\mathrm{d}\boldsymbol{P}^{\tau,0}_{\rho^0}}(\omega)-\frac{\mathrm{d}\boldsymbol{P}^{\tau,\epsilon}_{\rho^0}}{\mathrm{d}\boldsymbol{P}^{\tau,0}_{\rho^0}}(\Theta\omega)\right\}\right\rangle^0_{\rho^0}+o(\epsilon^2)$$
$$=\lim_{\tau\downarrow 0}\frac{1}{\tau}\left\langle \frac{\mathrm{d}\boldsymbol{P}^{\tau,\epsilon}_{\rho^0}}{\mathrm{d}\boldsymbol{P}^{\tau,0}_{\rho^0}}(\omega)\left\{\frac{\mathrm{d}\mu^\epsilon}{\mathrm{d}\rho^0}(\omega_0)-\frac{\mathrm{d}\mu^\epsilon}{\mathrm{d}\rho^0}(\omega_\tau)\right\}\right\rangle^0_{\rho^0}+\lim_{\tau\downarrow 0}\frac{1}{2\tau}\left\langle\left\{\frac{\mathrm{d}\boldsymbol{P}^{\tau,\epsilon}_{\rho^0}}{\mathrm{d}\boldsymbol{P}^{\tau,0}_{\rho^0}}(\omega)-\frac{\mathrm{d}\boldsymbol{P}^{\tau,\epsilon}_{\rho^0}}{\mathrm{d}\boldsymbol{P}^{\tau,0}_{\rho^0}}(\Theta\omega)\right\}^2\right\rangle^0_{\rho^0}$$
$$+o(\epsilon^2)=-\langle L^\epsilon f^\epsilon\rangle_{\rho^0}+\Delta^\epsilon+o(\epsilon^2)=-\epsilon^2\langle L_1f_1\rangle_{\rho^0}+\Delta^\epsilon+o(\epsilon^2), \tag{5.11}$$

where $\boldsymbol{P}^{\tau,0}_{\rho^0}$ and $\langle\cdot\rangle^0_{\rho^0}$ refer to the path-space distribution under the reference detailed balance dynamics ($\epsilon=0$) started from ρ^0. The term Δ^ϵ is simply independent of f^ϵ. All in all we have found, up to leading order,

$$\sigma^\epsilon(\mu^\epsilon)=-\epsilon^2\langle f_1L_0f_1+2L_1f_1\rangle_{\rho^0}+\Delta^\epsilon+o(\epsilon^2). \tag{5.12}$$

Comparing with the result Eq. (4.9) or (4.12) finishes the proof.

VI. PROOF OF PROPOSITION 4.1 AND OF THEOREM 4.1

Let $0<\delta<1$ be given and fix μ by giving $f=\mathrm{d}\mu/\mathrm{d}\rho^0\in M^\delta_{+1}$. In order to localize the maximizer $g^{*\epsilon}(f)$ of the functional J^ϵ_f, we decompose the set M_{+1} as follows. Given $\alpha,\beta>0$ such that $\alpha+\beta<\sqrt{\delta}$ we introduce

$$N^\alpha_{+1}(\mu)=\left\{g\in M_{+1};\left|g(x)-\frac{\sqrt{f(x)}}{\langle\sqrt{f}\rangle_{\rho^0}}\right|<\alpha,\quad x\in\Omega\right\}. \tag{6.1}$$

Obviously, $N^\alpha_{+1}(\mu)\subset M^\beta_{+1}$, and writing $[M^\beta_{+1}]^c=M_{+1}\setminus M^\beta_{+1}$ we have the disjoint decomposition

$$M_{+1}=N^\alpha_{+1}(\mu)\cup[M^\beta_{+1}\setminus N^\alpha_{+1}(\mu)]\cup[M^\beta_{+1}]^c. \tag{6.2}$$

In what follows we are going to prove that, choosing ϵ,α,β small enough, the functional J^ϵ_f takes its maximum inside $N^\alpha_{+1}(\mu)$ and that is unique by a local convexity argument.

We start with a lemma that follows immediately from the assumptions. Recall that the state space is assumed finite; let $|\Omega|=N$.

Lemma 6.1: *There is an irreducible graph G with vertex set Ω and for which over all edges (x,y), $\rho^0(x)k^0(x,y)\geqslant\gamma$ for some $\gamma>0$. Moreover, $k^\epsilon(x,y)\geqslant\frac{1}{2}k^0(x,y)$ for all sufficiently small $|\epsilon|>0$.*

The next lemma states that when g is outside M_{+1}^β, then $J_f^\epsilon(g)$ can be made very negative $(\beta\downarrow 0)$.

Lemma 6.2: *For all sufficiently small $|\epsilon|$ and for all $g\in[M_{+1}^\beta]^c$,*

$$J_f^\epsilon(g)\leqslant C-\frac{1}{2}\gamma\delta\beta^{-1/(N-1)}, \tag{6.3}$$

with C a constant independent of f, g, and ϵ.

Proof: From the previous lemma,

$$\begin{aligned}J_f^\epsilon(g)&=\sum_{x,y\neq x}\rho^0(x)f(x)k^\epsilon(x,y)\left[1-\frac{g(y)}{g(x)}\right]\\&\leqslant\max_x\sum_{y\neq x}k^\epsilon(x,y)-\frac{1}{2}\sum_{(x,y)\in G}\rho^0(x)k^0(x,y)\left[\frac{f(x)g(y)}{g(x)}+\frac{f(y)g(x)}{g(y)}\right]\\&\leqslant C-\frac{1}{2}\gamma\delta\sum_{(x,y)\in G}\left[\frac{g(y)}{g(x)}+\frac{g(x)}{g(y)}\right]\end{aligned} \tag{6.4}$$

for a suitable C and ϵ small enough. Since $g\in[M_{+1}^\beta]^c$, there is $\bar{x}\in\Omega$ such that $g(\bar{x})<\beta$. Hence there exists a pair $(x,y)\in G$ such that either $g(y)\geqslant\beta^{-1/(N-1)}g(x)$ or $g(x)\geqslant\beta^{-1/(N-1)}g(y)$. To see that, assume that this is not true and denote by $l(x)$ the length of the shortest path in G connecting $\bar{x}$ and x. Then, using $l(x)\leqslant N-1$,

$$\langle g\rangle_{\rho^0}\leqslant\max_x g(x)\leqslant g(\bar{x})\max_x\beta^{-l(x)/(N-1)}\leqslant\beta^{-1}g(\bar{x})<1, \tag{6.5}$$

which is a contradiction. ■

Now comes the statement that the maximum is also outside $M_{+1}^\beta\backslash N_{+1}^\alpha(\mu)$. Use the shorthand $g_0=\sqrt{f}/\langle\sqrt{f}\rangle_{\rho^0}$.

Lemma 6.3. *For all sufficiently small $|\epsilon|>0$ we have $J_f^\epsilon(g)<J_f^\epsilon(g_0)$ whenever $g\in M_{+1}^\beta\backslash N_{+1}^\alpha(\mu)$.*

Proof: As clear from the proof of Theorem 3.1, g_0 is the unique maximizer of J_f^0 in M_{+1} due to the irreducibility assumption. Using that $M_{+1}^\beta\backslash N_{+1}^\alpha(\mu)$ is a compact set (in the Euclidean metric, say),

$$\sup_{g\in M_{+1}^\beta\backslash N_{+1}^\alpha(\mu)}J_f^0(g)<J_f^0(g_0). \tag{6.6}$$

By the continuity of $J_f^\epsilon(g)$ at $\epsilon=0$ which is uniform in $g\in M_{+1}^\beta$, we can choose $|\epsilon|>0$ sufficiently small so that Eq. (6.6) extends to

$$\sup_{g\in M_{+1}^\beta\backslash N_{+1}^\alpha(\mu)}J_f^\epsilon(g)<J_f^\epsilon(g_0). \tag{6.7}$$

Hence, the lemma follows. ■

Lemma 6.4. *There is $\alpha>0$ such that for all sufficiently small $|\epsilon|>0$, J_f^ϵ is a strictly concave function in $N_{+1}^\alpha(\mu)$.*

Proof: For any $\psi:\Omega\to\mathbb{R}$, a direct computation yields

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right|_{t=0} J_f^0(g_0+t\psi) = -\sum_{x,y\neq x} \rho^0(x)k^0(x,y)\left[\psi(x)\left(\frac{f(y)}{f(x)}\right)^{1/4} - \psi(y)\left(\frac{f(x)}{f(y)}\right)^{1/4}\right]^2. \tag{6.8}$$

Since the term in the square bracket is strictly positive unless $\psi(x)/\psi(y)=\sqrt{f(x)/f(y)}$, and using Lemma 6.1, the right-hand side in Eq. (6.8) is strictly negative unless $\psi \propto \sqrt{f}$. In particular, it implies that $\mathrm{d}^2 J_f^0(g_0+t\psi)/\mathrm{d}t^2(t=0)$ is a strictly negative quadratic form on the linear subspace defined by $\langle\psi\rangle_{\rho^0}=0$.

By continuity, it first extends to the strict negativity of the quadratic form $\mathrm{d}^2 J_f^0(g+t\psi)/\mathrm{d}t^2(t=0)$ on the same linear subspace and for all $g \in N_{+1}^{\alpha}(\mu)$ with some $\alpha>0$. Finally, it implies the strict negativity of $\mathrm{d}^2 J_f^{\epsilon}(g+t\psi)/\mathrm{d}t^2(t=0)$ on $\langle\psi\rangle_{\rho^0}=0$ for all $g \in N_{+1}^{\alpha}(\mu)$ and for all sufficiently small $|\epsilon|>0$.

Proof of Theorem 4.1: Pick some $\alpha>0$ such that Lemma 6.4 holds, and fix $\beta>0$ to satisfy $J_{\mu}^{\epsilon}(g_0) > C-\frac{1}{2}\gamma\delta\beta^{-1/(N-1)}$ for all $|\epsilon|>0$ small enough. Then by Lemmas 6.2 and 6.3, the maximizer of J_{μ}^{ϵ} exists and is localized in N_{+1}^{α}; moreover it is unique by Lemma 6.4. As the α can be chosen arbitrarily small (and observe that $\alpha \downarrow 0$ drives $\epsilon \downarrow 0$), one also has $g^{*\epsilon}(\mu) \xrightarrow{\epsilon\downarrow 0} g_0$.

If $k^{\epsilon}(x,y)$, $x\neq y \in \Omega$, are all differentiable then the maximizer $g^{*\epsilon}$ coincides with a solution of Eq. (4.3) in the domain N_{+1}^{α}; recall that the latter necessarily exists and is unique. The perturbative calculation in Sec. IV A then follows by an application of the inverse mapping theorem. ■

VII. CONCLUSIONS: MINIMUM ENTROPY PRODUCTION PRINCIPLE

A. Summary

Our analysis goes beyond merely checking MinEP; rather, it enables to view it as a consequence of a dynamical variant of Einstein's formula for equilibrium fluctuations. In simple terms, Theorem (5.1) reads that the probability for the empirical distribution p_T to coincide with some $\mu^{\epsilon}=\rho^{\epsilon}+O(\epsilon)$ has the following generic structure:

$$\boldsymbol{P}^{T,\epsilon}(p_T \simeq \mu^{\epsilon}) \propto \mathrm{e}^{-(T/4)[\sigma^{\epsilon}(\mu^{\epsilon})+o(\epsilon^2)]+o(T)}. \tag{7.1}$$

By ergodicity, the maximal probability is obtained for $p_T=\rho^{\epsilon}$. According to the above it is also obtained by minimizing the entropy production. Hence, the MinEP emerges as an immediate consequence of the structure of dynamical fluctuations. Moreover, its approximate status is also understood since the relation between the entropy production and the true DV functional is restricted to the leading order of expansion around equilibrium. A systematic perturbation expansion of $I^{\epsilon}(\mu^{\epsilon})$ would provide corrections to that principle; we will not discuss that issue now. Some further remarks end the paper.

B. Remarks

What has been said so far about the entropy production is subject to one further *physical* condition: that the Markov process describes the dynamics of time-reversal symmetric variables. Only then are Eq. (5.2) or (5.3) correct expressions for the entropy production rate. Yet, certain observables, e.g., momentum or magnetic field, cannot be expected in detailed balance in the sense of Eq. (3.1) even at a closed system dynamics. Instead, a symmetry under time reversal can only be seen when also the sign of these so called time-reversal odd observables is changed. A deeper reason why such a generalization is needed is that the fundamental equations of motion are often second order in time. For processes on variables that are odd under time reversal the above analysis needs a modification (see also the next remark).

We give an example of a Gaussian Markov diffusion process (X_t). Suppose a Langevin dynamics of the form

$$dX_t = (\mathcal{E} - \gamma X_t)dt + \sqrt{\frac{2\gamma}{\beta}}dW_t, \tag{7.2}$$

with standard Wiener process W_t. The force $\mathcal{E}$ is constant and $\gamma > 0$ is some friction coefficient. For scalar $X_t \in \mathbb{R}$ the process is detailed balanced in the sense of Eq. (3.1) with respect to $\rho(dx) \propto \exp[-\beta/2(x-(\mathcal{E}/\gamma))^2]dx$, a Gibbs distribution for inverse temperature β. From Theorem 3.1 one easily computes the corresponding DV functional to be

$$I(\mu) = \frac{\gamma}{4\beta}\left\langle \frac{(f')^2}{f} \right\rangle_\rho, \quad f = \frac{d\mu}{d\rho}. \tag{7.3}$$

Is that equal to the entropy production? It now depends on whether X_t is even or odd under time reversal.

Assume first that X_t models the position of an overdamped oscillator. That is an even variable and the detailed balance (3.1) is verified; the stationary process is in equilibrium. The entropy production is found most easily from Eq. (5.4):

$$\sigma(\mu) = -\frac{d}{dt} S(\mu_t|\rho)|_{t=0} = \frac{\gamma}{\beta}\left\langle \frac{(f')^2}{f} \right\rangle_\rho. \tag{7.4}$$

Hence, we get $\sigma(\mu) = I(\mu)/4$, consistent with our general result.

Alternatively, suppose now that $X_t \equiv V_t$ is instead the fluctuating velocity of a Langevin particle dragged by force $\mathcal{E}$. Although Eq. (3.1) remains valid, it no longer expresses time-reversal invariance since the kinematical time reversal (changing the sign of the velocity) is not applied. Furthermore, if $\mathcal{E} \neq 0$, then $\langle \phi(v_0)\psi(v_\tau)\rangle_\rho \neq \langle \phi(-v_\tau)\psi(-v_0)\rangle_\rho$ breaking even a (generalized) reversibility. In particular, there is for $\mathcal{E} \neq 0$ a nonzero stationary entropy production. That mean entropy production can be obtained by the methods of Ref. 16 in the form

$$\sigma(\mu) = \frac{\gamma}{\beta}\left\langle \frac{1}{f}\left(f' + \frac{\beta\mathcal{E}}{\gamma}f\right)^2 \right\rangle_\rho \tag{7.5}$$

and is different from Eq. (7.4). Using that $\langle f' \rangle_\rho = \beta\langle v - \mathcal{E}/\gamma \rangle_\mu$ and $\sigma(\rho) = \beta\mathcal{E}^2/\gamma$, we obtain the following modification of Theorem 5.1:

$$I(\mu) = \frac{1}{4}[\sigma(\mu) + \sigma(\rho) - 2\beta\mathcal{E}\langle v \rangle_\mu]. \tag{7.6}$$

In particular, the stationary distribution is now found as a minimizer of the functional $\sigma(\mu) - 2\beta\mathcal{E}\langle v \rangle_\mu$. Equivalently, since $\sigma(\rho) = \beta\mathcal{E}\langle v \rangle_\rho$, the stationary measure is now characterized by a (constrained) *maximum* entropy production principle

$$\max_\mu \{\sigma(\mu) | \sigma(\mu) = \beta\mathcal{E}\langle v \rangle_\mu\} = \sigma(\rho). \tag{7.7}$$

The above also provides an explanation for the counterexample to MinEP given by Landauer.[10,11] There one considers an electrical circuit with resistance R, inductance L, and voltage source $\mathcal{E}$ in series. The physical entropy production is $\hat{\sigma}(j) = \beta R j^2$, corresponding to the Joule heat caused by the current j through the resistance R. Apparently, the stationary current $j^* = \mathcal{E}/R$ does not coincide with the minimum of the entropy production.

To understand the situation, we embed the network dynamics in a stochastic process by combining Kirchhoff's second law with the Johnson-Nyquist noise voltage on the resistance to get the equation

$$\mathrm{d}j_t = \frac{1}{L}(\mathcal{E} - Rj_t)\mathrm{d}t + \sqrt{\frac{2R}{\beta L^2}}\mathrm{d}W_t \tag{7.8}$$

(the Nyquist prefactor for the noise being determined from the fluctuation-dissipation relation). That is a linear Langevin equation of the form (7.2) for the current which is odd under time reversal. Hence, the conclusion of the previous remark applies and, in particular, both Theorem 5.1 and MinEP are no longer valid. Yet, we can obtain the correct variational principle from the DV functional.

Consider indeed the functional

$$\bar{I}(\bar{j}) = \inf_{\mu}\{I(\mu)|\langle j\rangle_\mu = \bar{j}\}, \tag{7.9}$$

which, by the contraction principle, is the large deviation rate function for the empirical average $j_T = 1/T\int_0^T j_t \mathrm{d}t$ as $T\uparrow +\infty$,

$$\boldsymbol{P}[j_T \simeq \bar{j}] \propto \exp[-T\bar{I}(\bar{j})]. \tag{7.10}$$

Here it is easy to compute from Eq. (7.3):

$$\bar{I}(\bar{j}) = \frac{\beta R}{4}\left(\bar{j} - \frac{\mathcal{E}}{R}\right)^2 \tag{7.11}$$

and that is then also the corrected variational functional to consider. For other examples and for further details we refer to Ref. 1.

Our result as formulated in Theorem 5.1 is no longer valid if we are away from the perturbation regime and the assumptions are not verified. As an example, consider again a Markov dynamics on a finite state space Ω and let μ be a distribution supported in some $\Omega_0 \subsetneq \Omega$, i.e., $\mu(x)=0$ for all $x \in \Omega\backslash\Omega_0$. As one immediately checks from Eq. (5.3), $\sigma(\mu)=+\infty$ whenever there are some $x \in \Omega_0$, $y \in \Omega\backslash\Omega_0$ such that $\mu(x)k(x,y) \neq 0$. On the other hand, the DV functional is bounded: $I(\mu) \leqslant \max_x \Sigma_{y\neq x} k(x,y)$.

The DV theory is not restricted to the time averages in the sense of Eq. (2.1). More generally, one can study fluctuations along a discrete sequence of observations with a time interval τ between the observations.[14] The time averages are then of the form

$$\frac{\tau}{T}(F(X_\tau) + F(X_{2\tau}) + \cdots + F(X_{T-\tau}) + F(X_T))$$

and we are concerned with their large deviations along the limit $T=n\tau\uparrow +\infty$. For every τ there is a rate function $I_\tau(\mu)$. The case (2.2) corresponds to $\lim_{\tau\downarrow 0} I_\tau(\mu)/\tau = I(\mu)$. Obviously, one can investigate the close-to-equilibrium behavior for every one of these cases and, in principle, one obtains for each of them a variational principle.

ACKNOWLEDGMENTS

One of the authors (K.N.) is grateful to the Instituut voor Theoretische Fysica, Katholieke Universiteit Leuven for kind hospitality, and acknowledges the support from the Grant Agency of the Czech Republic (Grant No. 202/07/0404).

[1] Bruers, S., Maes, C., and Netočný, K., e-print arXiv:cond-mat/0701035.
[2] Demboo, A., and Zeitouni, O., *Large Deviation Techniques and Applications* (Jones and Barlett, Boston, 1993).
[3] Deuschel, J.-D., and Stroock, D. W., *Large Deviations*, Pure and Applied Mathematics Vol. 137 (Academic, Boston, 1989).
[4] Donsker, M. D., and Varadhan, S. R., "Asymptotic evaluation of certain Markov process expectations for large time, I.," Commun. Pure Appl. Math. **28**, 1–47 (1975).
[5] Eyink, G., Lebowitz, J. L., and Spohn, H., in *Chaos, Soviet-American Perspectives in Nonlinear Science*, edited by D. Campbell (American Institute of Physics, New York, 1990), pp. 367–391.
[6] de Groot, S. R., and Mazur, P., *Non-equilibrium Thermodynamics* (North-Holland, Amsterdam, 1969).

[7] den Hollander, F., *Large Deviations* (Field Institute Monographs, Providence, RI, 2000).
[8] Jaynes, E. T., "The minimum entropy production principle," Annu. Rev. Phys. Chem. **31**, 579–601 (1980).
[9] Klein, M. J., and Meijer, P. H. E., "Principle of minimum entropy production," Phys. Rev. **96**, 250–255 (1954).
[10] Landauer, R., "Inadequacy of entropy and entropy derivatives in characterizing the steady state," Phys. Rev. A **12**, 636–638 (1975).
[11] Landauer, R., "Stability and entropy production in electrical circuits," J. Stat. Phys. **13**, 1–16 (1975).
[12] Liggett, T. M., *Interacting Particle Systems* (Springer, Berlin, 1985).
[13] Maes, C., "On the origin and the use of fluctuation relations for the entropy," Séminaire Poincaré **2**, 29–62 (2003).
[14] Maes, C., and Netočný, K., e-print arXiv:cond-mat/0612525.
[15] Maes, C., and Netočný, K., "Time-reversal and entropy," J. Stat. Phys. **110**, 269–310 (2003).
[16] Maes, C., Netočný, K., and Verschuere, M., "Heat conduction networks," J. Stat. Phys. **111**, 1219–1244 (2003).
[17] Maes, C., Redig, F., and Van Moffaert, A., "On the definition of entropy production via examples," J. Math. Phys. **41**, 1528–1554 (2000).
[18] Prigogine, I., *Introduction to Non-Equilibrium Thermodynamics* (Wiley-Interscience, New York, 1962).
[19] Jiang, D.-Q., Qian, M., and Qian, M.-P., *Mathematical Theory of Nonequilibrium Steady States*, Lecture Notes in Mathematics Vol. 833 (Springer, Berlin, 2004).
[20] Spohn, H., *Large Scale Dynamics of Interacting Particles* (Springer, Heidelberg, 1991).