

## Rigorous meaning of McLennan ensembles

Christian Maes<sup>1,a)</sup> and Karel Netočný<sup>2,b)</sup><sup>1</sup>*Instituut voor Theoretische Fysica, K.U.Leuven, Leuven 3001, Belgium*<sup>2</sup>*Institute of Physics, AS CR, Prague 8, 182 21 Czech Republic*(Received 5 November 2009; accepted 18 November 2009;  
published online 29 January 2010)

We analyze the exact meaning of expressions for nonequilibrium stationary distributions in terms of entropy changes. They were originally introduced by McLennan [“Statistical mechanics of the steady state,” *Phys. Rev.* **115**, 1405 (1959)] for mechanical systems close to equilibrium and more recent work by Komatsu and Nakagawa [“An expression for stationary distribution in nonequilibrium steady states,” *Phys. Rev. Lett.* **100**, 030601 (2008)] has shown their intimate relation to the transient fluctuation symmetry. Here we derive these distributions for jump and diffusion Markov processes and we clarify the order of the limits that take the system both to its stationary regime and to the close-to-equilibrium regime. In particular, we prove that it is exactly the (finite) transient component of the irreversible part of the entropy flux that corrects the Boltzmann distribution to first order in the driving. We add further connections with the notion of local equilibrium, with the Green–Kubo relation, and with a generalized expression for the stationary distribution in terms of a reference equilibrium process. © 2010 American Institute of Physics. [doi:10.1063/1.3274819]

### I. INTRODUCTION

According to McLennan,<sup>20,21</sup> the stationary density of an open mechanical system away but close to thermal equilibrium can be written in the modified Gibbs form,

$$\rho(x) \simeq \frac{1}{Z} e^{-\beta H(x)+W(x)}, \quad (1.1)$$

with the nonequilibrium correction  $W$  directly related to the entropy production or to the dissipation in the driven system. An essential feature of the formula, not quite visible yet, is that the distribution  $\rho$  is described in terms of macroscopic parameters only, such as external temperature and driving fields. It was also expected “that some formal advantages may be offered by an approach to nonequilibrium phenomena in which the Gibbs ensemble plays a more prominent role” (from the second paragraph in Ref. 20). Because of the suggested physical interpretation, this proposal opens the possibility to construct nonequilibrium statistical ensembles based on meaningful physical quantities, see also Refs. 11 and 12 for older and Ref. 22 for more recent work. However, (I.1) in Ref. 20 being just the result of a formal perturbation calculation (together with some projection techniques), there have remained a number of difficulties with the exact meaning of this proposal as well as with its scope of generality. We mention some of these problems, as are clarified in the present paper.

- (1) What entropy production does the correction term  $W$  represent? It can only be some transient component of the total entropy production as the total entropy production clearly diverges in the long-time (stationary) limit. Moreover, this divergence (equal to the steady entropy flux)

---

<sup>a)</sup>Electronic mail: christian.maes@fys.kuleuven.be.

<sup>b)</sup>Electronic mail: netocny@fzu.cz.

is, in fact, of order  $O(\varepsilon^2)$  in some “distance of equilibrium”  $\varepsilon$  since it comes from the product of thermodynamical forces and fluxes, both being  $O(\varepsilon)$ . The point will be that the transient irreversible part of the entropy production and its linear part are both finite: they coincide up to  $O(\varepsilon)$  and give a valid first-order correction to the Boltzmann distribution as in (1.1).

- (2) Is the proposal also valid on different levels of description than for mechanical systems? The formal perturbation approach as in Ref. 20 does not reveal the essence and the physical generality of the proposal. Here the insight comes from dynamical fluctuation theory: formula (1.1) is basically a consequence of the transient fluctuation symmetry.<sup>8</sup> In other words, it follows from the local detailed balance assumption which determines the time-antisymmetric structure of the space-time distribution in terms of the history-dependent entropy fluxes, cf. Refs. 13 and 14. We make that visible for Markov processes.
- (3) Can one go beyond close to equilibrium? As explained in Ref. 8 and further applied in Refs. 9 and 10, one can in principle obtain a formal perturbation series for the stationary distribution based on (all) the cumulants of the transient entropy production. An interpretation has been given for the second-order expansion where the divergences cancel out by a different way than explained in the present paper. We instead present a generalization involving the so-called dynamical activity or traffic, an advantage being that the evaluation is now from the start to be done under a reference equilibrium process.

We consider Markov processes of two types, jump and diffusion processes. Yet for simplicity we reserve Sec. II to Markov jump processes; the diffusion case is formally completely similar. We explain the relation with local equilibrium in Sec. II D. In Sec. III we make some specific remarks on the diffusion case and we also take there the opportunity to illustrate an alternative to the derivation in Sec. II B. As an application, the Green–Kubo relations are derived in Sec. III B. An illustration of the McLennan algorithm for an underdamped case is given in Sec. III C. We end, in Sec. IV, with a generalization away from equilibrium.

## II. MARKOV JUMP PROCESSES

After introducing some notation and basic concepts in the case of jump processes, we give our main result which is a rigorous version of the McLennan formula. A comparison to another approach and further remarks are added.

### A. Set-up and assumptions

Consider a continuous time Markov process  $x_t, t \geq 0$ , taking values in a finite state space  $\Omega \ni x, y, \dots$ . The transition rates are  $\lambda(x, y) \geq 0$  for jumps between the states  $x \rightarrow y$ . The Master equation for the probability  $\mu_t(x)$  of state  $x$  as function of time  $t$  is

$$\frac{d\mu_t(x)}{dt} = \sum_{y \neq x} \{\mu_t(y)\lambda(y, x) - \mu_t(x)\lambda(x, y)\}, \quad (2.1)$$

with some given initial law  $\mu_0 = \mu$  at time zero.

Physical input distinguishes between equilibrium and nonequilibrium dynamics. An equilibrium dynamics (with subscript 0) satisfies the condition of detailed balance, i.e.,

$$\rho_0(x)\lambda_0(x, y) = \rho_0(y)\lambda_0(y, x), \quad (2.2)$$

where  $\rho_0(x) \propto e^{-\beta U(x)}$  for some potential  $U$  and inverse temperature  $\beta$  is then stationary. This relation expresses the time reversibility of the stationary equilibrium process. For nonequilibrium systems, detailed balance (2.2) gets broken. An extension is known as the condition of local detailed balance. In terms of a potential  $U(x)$  and a work function (or driving)  $F(x, y) = -F(y, x)$ , the rates now satisfy

$$\lambda(x, y) = e^{\beta[F(x, y) + U(x) - U(y)]} \lambda(y, x), \quad (2.3)$$

where  $\beta \geq 0$  can still be interpreted as the inverse temperature of a reference reservoir, but that is not necessary except for setting the right units.

We can rewrite the local detailed balance condition (2.3) as

$$\rho_0(x) \lambda(x, y) = \gamma(x, y) e^{(\beta/2)F(x, y)} \quad (2.4)$$

for a symmetric  $\gamma(x, y) = \gamma(y, x)$ , which here is arbitrary.

In the present paper we are concerned with the close-to-equilibrium regime where  $F$  is small. To make it precise, we parametrize the distance to equilibrium explicitly by assuming that  $\gamma(x, y) = \gamma_\varepsilon(x, y)$  and  $F(x, y) = F_\varepsilon(x, y)$  [and hence also  $\lambda(x, y) = \lambda_\varepsilon(x, y)$ ] depend on a parameter  $\varepsilon \in [0, \varepsilon_0]$ . With no loss of generality we let  $F_\varepsilon(x, y) = \varepsilon F_1(x, y)$ . As becomes obvious later, the  $\varepsilon$ -dependence of  $\gamma_\varepsilon(x, y)$  is irrelevant for the first-order calculations.

We also consider the probability of trajectories, or rather, how to obtain probability densities in path space. For this we start with a probability law  $\mu$  at time zero for the Markov process and write  $\mathcal{P}_\mu$  for its path-space distribution over a time interval  $[0, T]$ . That has a density with respect to the corresponding stationary equilibrium process  $\mathcal{P}_0$  with rates  $\lambda_0(x, y)$  and starting from  $\rho_0$ , explicitly given by the Girsanov formula,

$$\frac{d\mathcal{P}_\mu}{d\mathcal{P}_0}(\omega) = \frac{\mu(x_0)}{\rho_0(x_0)} \exp \left\{ - \int_0^T (\xi(x_t) - \xi_0(x_t)) dt + \sum_{0 < t \leq T} \log \frac{\lambda(x_{t-}, x_t)}{\lambda_0(x_{t-}, x_t)} \right\}, \quad (2.5)$$

where  $\omega = (x_t)_{t=0}^T$ ,  $x_t \in \Omega$ , is a piecewise constant right-continuous trajectory, with the escape rates,

$$\xi(x) = \sum_{y \neq x} \lambda(x, y),$$

and with the last sum in the exponent being over the jump times  $t$  where the state changes from  $x_{t-}$  to  $x_t$ . Mathematical details are found in, e.g., Appendix 2 of Ref. 7.

From (2.5), the path-space action  $A$  in

$$d\mathcal{P}_\mu(\omega) = d\mathcal{P}_0(\omega) \frac{\mu(x_0)}{\rho_0(x_0)} e^{-A(\omega)} \quad (2.6)$$

equals

$$A(\omega) = \exp \left\{ \int_0^T (\xi(x_t) - \xi_0(x_t)) dt - \sum_{0 < t \leq T} \log \frac{\lambda(x_{t-}, x_t)}{\lambda_0(x_{t-}, x_t)} \right\}.$$

As a result, its time-antisymmetric part is

$$S_{\text{IRR}}^T(\omega) = A(\theta\omega) - A(\omega) = \beta \sum_{0 < t \leq T} F(x_{t-}, x_t), \quad (2.7)$$

where the time reversal is the right-continuous modification of  $\theta\omega = (x_{T-t})_{t=0}^T$  for any  $\omega = (x_t)_{t=0}^T$ . We have used that the first integral in the exponent of (2.5) is time symmetric and that (2.2)–(2.4) combine to produce the forcing in the sum over jump times. We recognize the resulting  $S_{\text{IRR}}^T(\omega)$  as the “irreversible” part in the entropy flux as function of the path  $\omega$ . Note that the pathwise relation (2.7) between the time-reversal symmetry breaking and the entropy flux is a consequence of condition (2.3). For general arguments, see, e.g., Ref. 18.

The mean value of that irreversible part of the entropy production is obtained by taking the average of (2.7) with respect to our process, using its Markov property,

$$\begin{aligned}
\langle S_{\text{IRR}}^T \rangle_\mu &= \int d\mathcal{P}_\mu(\omega) S_{\text{IRR}}^T(\omega) = \int_0^T dt \lim_{\tau \downarrow 0} \frac{1}{\tau} \langle S_{\text{IRR}}^\tau \rangle_{\mu_t} \\
&= \beta \int_0^T dt \sum_x \mu_t(x) \sum_{y \neq x} \lambda_\varepsilon(x, y) F_\varepsilon(x, y) \\
&= \beta \int_0^T dt \left\langle \sum_{y \neq x} \lambda_\varepsilon(x_t, y) F_\varepsilon(x_t, y) \right\rangle_\mu.
\end{aligned} \tag{2.8}$$

Hence, for fixed  $T$  we have to first order in  $\varepsilon$ ,

$$\langle S_{\text{IRR}}^T \rangle_\mu = \varepsilon \beta \int_0^T dt \langle w_1(x_t) \rangle_\mu^0 + O(\varepsilon^2), \tag{2.9}$$

where the averaging  $\langle \cdot \rangle_\mu^0$  is now over the equilibrium reference process started from  $\mu$ , and

$$w_1(x) = \sum_{y \neq x} \lambda_0(x, y) F_1(x, y) \tag{2.10}$$

is the linear term in the mean entropy flux when at state  $x$ .

## B. McLennan formula

To be explicit about the various dependencies, we write  $\rho_T^\varepsilon$  for the  $\varepsilon$ -dependent solution at time  $T$  to the master equation (2.1), started from the equilibrium law  $\mu_0 = \rho_0$ . The smoothness of the deformation is assumed uniformly in time  $T$ , and we write  $\rho_\varepsilon = \lim_T \rho_T^\varepsilon$ . We also denote the stationary entropy flux by  $\sigma_\varepsilon$ ; it is given as

$$\begin{aligned}
\sigma_\varepsilon &= \frac{1}{T} \langle S_{\text{IRR}}^T \rangle_{\rho_\varepsilon} \\
&= \beta \sum_x \rho_\varepsilon(x) \sum_{y \neq x} \lambda_\varepsilon(x, y) F_\varepsilon(x, y) \\
&= \frac{\beta}{2} \sum_{x, y} [\rho_\varepsilon(x) \lambda_\varepsilon(x, y) - \rho_\varepsilon(y) \lambda_\varepsilon(y, x)] F_\varepsilon(x, y) \\
&= \frac{\varepsilon \beta}{2} \sum_{x, y} \gamma_0(x, y) \left[ \frac{\rho_\varepsilon(x)}{\rho_0(x)} - \frac{\rho_\varepsilon(y)}{\rho_0(y)} \right] F_1(x, y) + o(\varepsilon^2)
\end{aligned} \tag{2.11}$$

independently of time span  $T$ . We finally recall the linear term  $w_1$  from (2.10).

**Theorem II.1:** *Suppose that the equilibrium process (2.2) is irreducible. The following limiting identities are verified:*

$$\lim_{T \uparrow +\infty} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \log \frac{\rho_T^\varepsilon(x)}{\rho_0(x)} = \lim_{\varepsilon \rightarrow 0} \lim_{T \uparrow +\infty} \frac{1}{\varepsilon} \log \frac{\rho_T^\varepsilon(x)}{\rho_0(x)} \tag{2.12}$$

$$= -\beta \int_0^{+\infty} dt \langle w_1(x_t) \rangle_x^0. \tag{2.13}$$

Moreover,

$$\varepsilon \beta \int_0^{+\infty} dt \langle w_1(x_t) \rangle_x^0 = \lim_{T \uparrow +\infty} [\langle S_{\text{IRR}}^T \rangle_x - \sigma_\varepsilon T] + O(\varepsilon^2). \tag{2.14}$$

*Remark II.2:* According to the above result, the stationary distribution has the form

$$\rho_\varepsilon(x) = \rho_0(x) \exp \left\{ -\varepsilon \beta \int_0^{+\infty} dt \langle w_1(x_t) \rangle_x^0 + O(\varepsilon^2) \right\} = \rho_0(x) \exp \left\{ \lim_{T \uparrow +\infty} [\sigma_\varepsilon T - \langle S_{\text{IRR}}^T \rangle_x] + O(\varepsilon^2) \right\}, \quad (2.15)$$

which is consistent with the original McLennan's proposal (1.1) in the sense that it identifies the correction term  $W$  as the transient part of the "irreversible" entropy production for the process started from state  $x$ . Notice that this transient part is  $O(\varepsilon)$ , in contrast to the stationary entropy production rate which is  $O(\varepsilon^2)$ ; the latter being also the leading order of the long-time divergence that needs to be removed. Hence, loosely speaking, the McLennan proposal is all correct for close-to-equilibrium processes provided that the divergence present in higher orders in  $\varepsilon$  is killed by a suitable counterterm.

*Proof:* The first equality (2.12) follows from the irreducibility of the reference equilibrium process. The  $\varepsilon$ -dependent process is obtained by its smooth deformation, see (2.3) and (2.4). Therefore,  $\rho_T^\varepsilon \rightarrow \rho_\varepsilon$  uniformly in  $\varepsilon \in [0, \varepsilon_1]$  with some  $0 < \varepsilon_1 \leq \varepsilon_0$ .

In order to prove the equality (2.13), the point of departure is the transient fluctuation symmetry. The formula (2.6) obviously implies

$$d\mathcal{P}_{\rho_0}(\omega) = d\mathcal{P}_0(\omega) e^{-A(\omega)}$$

when starting (in the left-hand side) the nonequilibrium process from the equilibrium law  $\rho_0$ . Since the equilibrium process  $\mathcal{P}_0$  is time-reversal invariant, we have for (2.7)

$$S_{\text{IRR}}^T(\omega) = \log \frac{d\mathcal{P}_{\rho_0}}{d\mathcal{P}_{\rho_0}^\theta}(\omega)$$

and hence, for all functions  $f$  on path space,

$$\langle f \rangle_{\rho_0} = \langle f \theta \exp(-S_{\text{IRR}}^T) \rangle_{\rho_0}. \quad (2.16)$$

That is an exact (for all finite times  $T$ ) fluctuation symmetry. Take in (2.16)  $f(\omega) = \delta_{x_T=x}$ , the Kronecker-delta function is equal to 1 if the trajectory ends up at state  $x$  and zero otherwise. We get

$$\rho_T^\varepsilon(x) = \rho_0(x) \langle e^{-S_{\text{IRR}}^T} \rangle_x, \quad (2.17)$$

where the right-hand side averages over the nonequilibrium process started from the state  $x$ . We substitute (2.7) and we use that the Poisson number of jumps has all exponential moments, to expand the exponential in (2.17). Using (2.8), it is then easy to verify that

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \log \langle e^{-S_{\text{IRR}}^T} \rangle_x = -\beta \left\langle \int_0^T dt w_1(x_t) \right\rangle_x^0, \quad (2.18)$$

which has a limit as  $T \uparrow +\infty$  since, uniformly in the initial  $x$ , the equilibrium process relaxes exponentially fast to its stationary law for which  $\sum_x \rho_0(x) w_1(x) = 0$  by (2.2). That proves formula (2.13).

We now turn to (2.14). For all initial laws  $\mu$ ,

$$\langle S_{\text{IRR}}^T \rangle_\mu - \sigma_\varepsilon T = \beta \int_0^T dt \sum_x [\mu_t(x) - \rho_\varepsilon(x)] \sum_{y \neq x} \lambda_\varepsilon(x, y) F_\varepsilon(x, y)$$

has a limit  $T \uparrow +\infty$ . Since all terms are uniformly bounded, we can here freely exchange the limits  $T \uparrow +\infty$  and  $\varepsilon \downarrow 0$ . The  $T$ -limit has leading order in  $\varepsilon$ , for  $\varepsilon \downarrow 0$ ,

$$\varepsilon\beta \int_0^{+\infty} dt \sum_x [\mu_t^0(x) - \rho_0(x)] \sum_{y \neq x} \lambda_0(x,y) F_1(x,y) = \varepsilon\beta \int_0^{+\infty} dt \sum_x \mu_t^0(x) \sum_{y \neq x} \lambda_0(x,y) F_1(x,y) \quad (2.19)$$

because  $F_1(x,y)$  is antisymmetric and  $\rho_0(x)\lambda_0(x,y)$  is symmetric under  $x \leftrightarrow y$ . That proves (2.14). ■

### C. Variations

What is mostly new about the above arguments is the point of departure (2.16), which naturally links McLennan's correction to the "irreversible" entropy production, cf. Ref. 8. There are other schemes that do not have this advantage; still they can be applied to obtain systematic corrections to the equilibrium distribution. We discuss one such a formulation and we apply it to obtain a different perturbation scheme not having an interpretation in terms of the entropy production.

By using the fundamental theorem of calculus, the stationary measure  $\rho_\varepsilon = \rho$  can be obtained in the form

$$\rho(x) = \rho_0(x) + \int_0^{+\infty} dt e^{tL^*} L^* \rho_0(x) \quad (2.20)$$

for initial law  $\rho_0$  and with  $L^* \mu(x) = \sum_{y \neq x} [\lambda(y,x)\mu(y) - \lambda(x,y)\mu(x)]$  the forward generator of the Markov jump process under consideration. Then, from (2.4),

$$L^* \rho_0(x) = \rho_0(x)h(x), \quad h(x) = \sum_{y \neq x} \lambda(x,y)[e^{-\beta F(x,y)} - 1]$$

is of order  $\varepsilon$ , and

$$e^{tL_0^*}[\rho_0 h](x) = \rho_0(x)e^{tL_0}h(x)$$

with  $L_0$  the backward equilibrium generator. Substituting to (2.20), we get the expression

$$\rho_\varepsilon(x) = \rho_0(x) \left[ 1 + \int_0^{+\infty} dt e^{tL_0} h(x) \right] + O(\varepsilon^2). \quad (2.21)$$

Since

$$h(x) = -\varepsilon\beta \sum_{y \neq x} \lambda_0(x,y) F_1(x,y) + O(\varepsilon^2),$$

we have rederived formula (2.13).

The scheme (2.3) and (2.4) in combination with the  $\varepsilon$ -dependence according to  $F(x,y) = \varepsilon F_1(x,y)$  is a special way of breaking the detailed balance condition. One easily encounters other physically relevant mechanisms. For example, suppose that for the reference equilibrium dynamics transitions  $x \rightleftharpoons y$  between specific states are forbidden. We may imagine two uncoupled equilibrium systems. The nonequilibrium dynamics could introduce a small coupling with, for example,

$$\lambda(x,y) = \varepsilon k(x,y), \quad \lambda(y,x) = \varepsilon k(y,x)$$

for these specific transitions. We do not longer enjoy then the absolute continuity of the nonequilibrium process with respect to the equilibrium reference and the relations (2.5) and (2.6) break down. The entropy production as a function on path space does not depend on  $\varepsilon$ . Nevertheless, from (2.20) we can still compute the linear correction to the reference law  $\rho_0$ . It is exactly of the form (2.21) but with

$$h(x) = \varepsilon \sum_{y \leftrightarrow x} \left[ \frac{\rho_0(y)}{\rho_0(x)} k(y, x) - k(x, y) \right] = \varepsilon \sum_{y \leftrightarrow x} k(x, y) [e^{-\phi(x, y)} - 1] \quad (2.22)$$

and

$$\phi(x, y) = \log \frac{\rho_0(x) k(x, y)}{\rho_0(y) k(y, x)}, \quad (2.23)$$

where the sum is over the *for the equilibrium dynamics* forbidden transitions from state  $x$ . The difference with nonequilibrium perturbations where one adds a small driving in a local detailed balance condition, as in (2.3) and (2.4), is manifest. The correction to equilibrium is not of the form of an entropy flux. In other words, not all perturbations from equilibrium, even physical ones, lead to the same type of correction to the equilibrium distribution: the specific McLennan correction in terms of the irreversible entropy flux arises (only) by the change from detailed balance to local detailed balance by inserting some small driving.

### D. Example: Boundary driven lattice gas

The following example makes the above considerations and formulas more concrete. We also take the opportunity to explain the relation with local equilibrium.

We consider a lattice gas on the sites  $\{-N, \dots, 0\}$ , where the configurations  $x, y$  indicate the vacancy or the presence  $x(i) = 0, 1$  of a particle at each site  $i$ . The dynamics distinguishes two ways of updating.

We concentrate on the case where there is a bulk conservation law as in Kawasaki dynamics. To be specific we choose the bulk transitions as an exchange of nearest neighbor occupancy: we write  $x^{i, i+1}$  for the configuration that equals  $x$  except that the occupations at sites  $i$  and  $i+1$  are interchanged. Then,

$$\lambda(x, y) = a_i \exp\left(-\frac{\beta}{2}\{U(y) - U(x)\}\right) \quad \text{when } y = x^{i, i+1}, \quad i = -N, \dots, -1. \quad (2.24)$$

At the left and right boundaries, there is a birth and death process: with  $y^i$  the configuration for which at site  $i$  the occupation has been inverted,

$$\lambda(x, y) = \exp\left(-\frac{\beta}{2}\{U(y) - U(x)\}\right) \exp\left(-\frac{\beta b_i}{2}(2x(i) - 1)\right) \quad \text{when } y = x^i, \quad i = -N, 0,$$

with  $b_i$  playing the role of chemical potential of left and right reservoirs.

Suppose now that  $a_i = 1$ ,  $b_{-N} = 0$ ,  $b_0 = \varepsilon$  which is a close-to-equilibrium system in the sense of Sec. II A; the equilibrium law is  $\rho_0(x) \propto \exp[-\beta U(x)]$ , reached for  $\varepsilon = 0$ . The expression (2.10) becomes

$$w_1(x) = \exp\left(-\frac{\beta}{2}\{U(x^0) - U(x)\}\right) [2x(0) - 1]. \quad (2.25)$$

On the other hand, consider the function  $g(x) = \sum_{i=-N}^0 v_i x(i)$  for some profile  $v_i$ . Then,

$$\begin{aligned} L_0 g(x) &= \sum_{i=-N}^{-1} [v_i - v_{i+1}] j_i(x) + v_0 \exp[-\beta\{U(x^0) - U(x)\}] [2x(0) - 1] \\ &\quad + v_{-N} \exp[-\beta\{U(x^{-N}) - U(x)\}] [2x(-N) - 1], \end{aligned} \quad (2.26)$$

with systematic currents over the bonds  $(i, i+1)$ ,

$$j_i(x) = \lambda(x, x^{i,i+1})[x(i+1) - x(i)],$$

as they appear in the continuity equation for local particle number. Choose  $v_i = 1 + i/N$  which makes  $v_0 = 1$ ,  $v_{-N} = 0$ , and  $v_{i+1} - v_i = 1/N$ . Comparing (2.25) with (2.26) yields

$$L_0 g(x) = -\frac{1}{N} \sum_{i=-N}^{-1} j_i(x) + w_1(x). \quad (2.27)$$

For the McLennan form (2.13) we must take the time integral of (2.25) so that

$$\rho_\varepsilon(x) \propto \rho_0(x) \exp \left\{ \varepsilon \beta \sum_{i=-N}^0 v_i x(i) \right\} \exp \left\{ -\frac{\varepsilon \beta}{N} \sum_{i=-N}^{-1} \int_0^{+\infty} dt e^{tL_0} j_i(x) \right\}. \quad (2.28)$$

This expression (2.28) for the approximate stationary distribution is of the form of local equilibrium for the conserved quantity (particle number) containing the linear profile  $v_i$  for the (local) chemical potential. The remaining integral,

$$\varepsilon \beta \int_0^{+\infty} dt \frac{1}{N} \sum_{i=-N}^{-1} \langle j_i(x_t) \rangle_x^0,$$

makes mathematical sense because the local currents die out exponentially fast for the equilibrium dynamics. It was also discussed in the same context as formula (3.49) in Ref. 6. As  $N$  gets large, the average current gets even smaller for each fixed time. It appears, without proof, that for boundary driven spatially extended systems the McLennan regime close to equilibrium can also be reached by taking  $N$  large, for fixed chemical potential difference, while that is not included in formulations such as (2.13).

### III. DIFFUSION PROCESSES

In this section we treat Markov diffusion processes and we give an alternative derivation of the main result.

#### A. General argument

We consider the class of  $d$ -dimensional inhomogeneous Itô diffusions,

$$dx_t = \{ \chi(x_t) [F(x_t) - \nabla U(x_t)] + \nabla \cdot D(x_t) \} dt + \sqrt{2D(x_t)} dB_t, \quad (3.1)$$

defined on a torus (i.e., we assume periodic boundary conditions). We use the notation  $\nabla \cdot D$  for the vector with components  $\sum_k \partial_l D_{kl}$  and assume  $D = \chi / \beta$ ; the latter being the Einstein relation as a variant of the local detailed balance condition for diffusion. The  $d$ -dimensional vector  $dB_t$  has independent standard Gaussian white noise components. We assume the fields  $F(x)$ ,  $U(x)$ , and  $\chi(x)$  are smooth and the matrices  $\chi(x)$  are symmetric and strictly positive at all points  $x$ .

To each distribution with density  $\mu$  there is associated current density,

$$j_\mu = \chi(F - \nabla U)\mu - D \nabla \mu, \quad (3.2)$$

and the stationarity of the law  $\mu = \rho$  is equivalent with the condition  $\nabla \cdot j_\rho = 0$ . The  $j_\mu$  gives the expected profile of the “real” particle current at given density  $\mu$  in the sense that for any smooth function  $f$ ,



$$\left\langle \int_0^T f(x_t) \circ dx_t \right\rangle_\mu = \int_0^T dt \int f(x) j_{\mu_t}(x) dx. \quad (3.3)$$

The left-hand side is the average of a Stratonovich-stochastic integral under the diffusion process started at time  $t=0$  from density  $\mu$ . In particular, the instantaneous mean work of the force  $F$  is then

$$W(\mu) = \int F \cdot j_\mu dx = \int w \mu dx \quad (3.4)$$

for

$$w = F \cdot \chi F - \chi F \cdot \nabla U + \nabla \cdot (DF). \quad (3.5)$$

In order to check the McLennan proposal we isolate the linear order in (3.4). We take the case  $F=0$ ,  $\rho_0 = \exp[-\beta U]/Z > 0$ , as the equilibrium reference and we expand with small parameter  $\varepsilon$ ,

$$F = \varepsilon F_1 + \dots,$$

$$w = \varepsilon w_1 + \dots,$$

$$\rho = \rho_\varepsilon = \rho_0(1 + \varepsilon h_1 + \dots) \quad (3.6)$$

assuming smooth behavior around  $\varepsilon=0$ . From (3.5),

$$w_1 = \nabla \cdot (DF_1) - \chi F_1 \cdot \nabla U = \frac{\nabla \cdot (\rho_0 \chi F_1)}{\beta \rho_0}, \quad (3.7)$$

the linear term in the mean work performed by the force  $F$ . It turns out, as in (2.13) and in agreement with McLennan's proposal, that  $h_1$  is given in terms of the mean total work performed on the particle started from  $x$  under the reference dynamics.

**Theorem III.1:** *Suppose the process (3.1) converges exponentially fast and uniformly in initial states to its unique stationary probability distribution with smooth density  $\rho$  around  $\rho_0$  as in (3.6). Then,*

$$h_1(x) = -\beta \left\langle \int_0^\infty w_1(x_t) dt \right\rangle_x^0. \quad (3.8)$$

*Proof:* The current (3.2) can be rewritten in terms of the reference equilibrium density,

$$j_\mu = \mu \chi (F - \nabla U) - D \nabla \mu = \mu \chi F - \rho_0 D \nabla \left( \frac{\mu}{\rho_0} \right), \quad (3.9)$$

so that the stationarity condition  $\nabla \cdot j_\rho = 0$  implies

$$\begin{aligned} 0 &= \nabla \cdot \left[ \rho_0 D \nabla \left( \frac{\rho}{\rho_0} \right) - \rho \chi F \right] \\ &= \rho_0 D \nabla \cdot \nabla \left( \frac{\rho}{\rho_0} \right) + \rho_0 (\nabla \cdot D) \cdot \nabla \left( \frac{\rho}{\rho_0} \right) - \rho_0 \chi \nabla U \cdot \nabla \left( \frac{\rho}{\rho_0} \right) - \nabla \cdot (\rho \chi F) \\ &= \rho_0 L_0 \left( \frac{\rho}{\rho_0} \right) - \nabla \cdot (\rho \chi F) \end{aligned} \quad (3.10)$$

with  $L_0$  the backward reference equilibrium generator. To linear order, that gives

$$L_0 h_1 = \beta w_1 \quad (3.11)$$

with  $w_1$  from (3.7).

As the relaxation to equilibrium is fast enough,  $L_0$  can be inverted on the space,

$$\Omega^\perp = \left\{ g: \int g \rho_0 dx = 0 \right\} \quad (3.12)$$

and since  $w_1 \in \Omega^\perp$ , the stationary solution must have first order,

$$h_1 = \beta L_0^{-1} w_1 \quad (3.13)$$

or

$$h_1(x) = -\beta \int_0^\infty (e^{tL_0} w_1)(x) dt = -\beta \left\langle \int_0^\infty w_1(x_t) dt \right\rangle_x^0 \quad (3.14)$$

as required. ■

One might be tempted to write

$$\frac{\rho_\varepsilon}{\rho_0}(x) \stackrel{?}{=} 1 - \beta \left\langle \int_0^\infty w(x_t) dt \right\rangle_x + O(\varepsilon^2) \quad (3.15)$$

(i.e., with the *full* work and possibly under the nonequilibrium measure) but this is only formally true in the sense that the  $O(\varepsilon)$  terms on both sides are equal; however, the second term on the right-hand side generally diverges now because of its  $O(\varepsilon^2)$  part. Another way to see that is by observing that  $w \notin \Omega^\perp$ , in general, in which case  $L_0^{-1}w$  does not exist.

Still, one can proceed similarly as in the case of jump processes and add a suitable counter-term on the right-hand side of (3.15). All that explains what is actually the rigorous meaning of McLennan's proposal: to correctly describe the first-order correction, one is not allowed to deal with the *full* transient entropy production unless the divergences coming from the high-order corrections are removed. For safe first-order calculations one needs to take the linear part of the entropy production functional only, as done above in Theorem III.1.

## B. Green–Kubo relations

An expression for the close-to-equilibrium stationary density obviously yields information about the stationary current in linear response around equilibrium. That provides another derivation of the well-known Green–Kubo relations between the current and equilibrium time correlations.

We use the same notation as in Sec. III A but this time we need to indicate the dependence on the driving force  $F$ , e.g., the linear part of the work (3.7) is now written as

$$w_1^F = \frac{\nabla \cdot (\rho_0 \chi F_1)}{\beta \rho_0}. \quad (3.16)$$

By expanding the mean stationary current (3.2) ( $\mu = \rho$ ) in powers of  $\varepsilon$ ,  $j_\rho^F = \varepsilon j_1^F + \dots$ , the leading term has the form

$$j_1^F = \rho_0 \chi [F_1 - \nabla \cdot (L_0^{-1} w_1^F)], \quad (3.17)$$

where we have substituted the McLennan form. Suppose  $G = \varepsilon G_1 + \dots$  is another smooth field. Then we have, up to order  $\varepsilon^2$ , that  $\int G \cdot j_\rho^F dx = \varepsilon^2 \int G_1 \cdot j_1^F dx + o(\varepsilon^2)$  and

$$\int G_1 \cdot j_1^F dx = \int \rho_0 [G_1 \cdot \chi F_1 + \beta w_1^G L_0^{-1} w_1^F] dx. \quad (3.18)$$

Since  $L_0$  is self-adjoint with respect to the scalar product  $(f, g) = \int \rho_0 \bar{f} g dx$  (and  $L_0^{-1}$  is therefore self-adjoint on  $\Omega^\perp$ ) and as the matrix  $\chi$  is symmetric, we immediately get the Onsager reciprocity relations in the form

$$\int G_1 \cdot j_1^F dx = \int F_1 \cdot j_1^G dx. \quad (3.19)$$

Note that for  $F=G$  the formula (3.18) gives the leading (=second order) term in the expansion for the stationary instantaneous mean work  $W^F(\rho^F)$ , see (3.4), whereas for  $F \neq G$  it corresponds to the “interference” contribution when the driving fields are added,

$$W^{F+G}(\rho^{F+G}) - W^F(\rho^F) - W^G(\rho^G) = 2\varepsilon^2 \int G_1 \cdot j_1^F dx + o(\varepsilon^2). \quad (3.20)$$

**Theorem III.2:** *Under the same assumptions as in Theorem III.1,*

$$\int G_1 \cdot j_1^F dx = \int dx \rho_0 [G_1 \cdot \chi F_1 + \beta w_1^G L_0^{-1} w_1^F] = \lim_{T \uparrow +\infty} \frac{\beta}{2T} \left\langle \int_0^T G_1(x_t) \circ dx_t \int_0^T F_1(x_s) \circ dx_s \right\rangle^0, \quad (3.21)$$

where  $\circ$  indicates the Stratonovich integration (incorporating the scalar product) and the last expectation is under the equilibrium process.

*Remark III.3:* The middle term in (3.21) is what follows from applying the McLennan formula to the mean current close to equilibrium. The equality (3.21) then yields the linear response formula for the close-to-equilibrium stationary current in terms of current-current time correlations, its right-hand side. The result can be formally summarized as saying that

$$j_\rho^F(x) = \int \mathcal{R}(x, y) F(y) dy + o(\varepsilon), \quad (3.22)$$

with a symmetric response function  $\mathcal{R}(x, y) = \mathcal{R}(y, x)$  given by

$$\mathcal{R}(x, y) = \lim_{T \uparrow +\infty} \frac{\beta}{2T} \langle J_T(x) J_T(y) \rangle^0 \quad (3.23)$$

and

$$J_T(x) = \int_0^T \delta(x_t - x) \circ dx_t \quad (3.24)$$

is the time-integrated empirical current density.

*Proof:* As the first equality in (3.21) is formula (3.18), we only need to prove the second equality there.

Using the standard relation between the Stratonovich and the Itô integrals, each integration along the equilibrium process [corresponding to  $F=0$  in (3.1)] on the right-hand side of (3.18) can be computed as follows:

$$\begin{aligned}
\int_0^T F_1 \circ dx_t &= \int_0^T F_1 \cdot dx_t + \int_0^T \sum_{kl} D_{kl} \frac{\partial F_{1,l}}{\partial x_k} dt \\
&= \int_0^T [-F_1 \cdot \chi \nabla U + \nabla \cdot (DF_1)] dt + \int_0^T F_1 \cdot (2D)^{1/2} dB_t \\
&= \int_0^T w_1^F dt + \int_0^T F_1 \cdot (2D)^{1/2} dB_t.
\end{aligned} \tag{3.25}$$

We must substitute (3.25) into the right-hand side of (3.21). We start with the cross terms, one of which is

$$\begin{aligned}
\left\langle \int_0^T w_1^G dt \int_0^T F_1 \cdot (2D)^{1/2} dB_s \right\rangle &= \left\langle \int_0^T w_1^G dt \int_0^t F_1 \cdot (2D)^{1/2} dB_s \right\rangle^0 \\
&= - \left\langle \int_0^T w_1^G dt \left[ \int_t^T F_1 \cdot (2D)^{1/2} dB_s + 2 \int_t^T w_1^F ds \right] \right\rangle^0 \\
&= -2 \left\langle \int_0^T w_1^G dt \int_t^T w_1^F ds \right\rangle^0.
\end{aligned} \tag{3.26}$$

In the second equality we have applied the time reversal of the last equality in (3.25), using that the Stratonovich integral is time antisymmetric and that the equilibrium process is time-reversal symmetric. Analogously,

$$\left\langle \int_0^T w_1^F dt \int_0^T G_1 \cdot (2D)^{1/2} dB_s \right\rangle^0 = -2 \left\langle \int_0^T w_1^F dt \int_t^T w_1^G ds \right\rangle^0 = -2 \left\langle \int_0^T w_1^G dt \int_0^t w_1^F ds \right\rangle^0. \tag{3.27}$$

Hence, both cross terms together give

$$-2 \left\langle \int_0^T w_1^G dt \int_0^T w_1^F ds \right\rangle^0. \tag{3.28}$$

As a consequence, the correlation is then

$$\begin{aligned}
&\left\langle \int_0^T G_1(x_t) \circ dx_t \int_0^T F_1(x_s) \circ dx_s \right\rangle^0 \\
&= 2 \left\langle \int_0^T G_1 \cdot DF_1(x_t) dt \right\rangle^0 - \left\langle \int_0^T w_1^G(x_t) dt \int_0^T w_1^F(x_s) ds \right\rangle^0 \\
&= 2T \langle G_1 \cdot DF_1 \rangle^0 - \int_{\alpha T}^{(1-\alpha)T} dt \int_{-t}^{T-t} ds \langle w_1^G(x_t) w_1^F(x_s) \rangle^0 + O(\alpha T) \\
&= 2T \langle G_1 \cdot DF_1 \rangle^0 - (1-2\alpha)T \int_{-\infty}^{+\infty} ds \langle w_1^G(x_0) w_1^F(x_s) \rangle^0 + O(\alpha T) + O(Te^{-\kappa \alpha T})
\end{aligned}$$

for an arbitrary  $0 < \alpha < 1/2$  and with  $\kappa$  the rate of exponential decay of the time correlations. Dividing by  $2T$  and taking the limits  $T \uparrow +\infty$  and then  $\alpha \downarrow 0$ , one finally obtains

$$\lim_{T \uparrow +\infty} \frac{1}{2T} \left\langle \int_0^T G_1(x_t) \circ dx_t \int_0^T F_1(x_s) \circ dx_s \right\rangle^0 = \langle G_1 \cdot DF_1 \rangle^0 - \int_0^{+\infty} ds \langle w_1^G e^{sL_0} w_1^F \rangle^0. \quad (3.29)$$

*Remark III.4:* The same proof applies for jump processes. One then has the analogs  $w_1^F = \sum_y \lambda_0(x, y) F(x, y)$ ,  $F(x, y) = -F(y, x)$ , and  $j_1^F = \gamma(x, y) [F(x, y) + h^F(x) - h^F(y)]$ ,  $\gamma(x, y) = \rho_0(x) \lambda_0(x, y) = \gamma(y, x)$ ,  $h^F(x) = L_0^{-1} w_1^F(x)$ ,  $\rho^F(x) = \rho_0(x) [1 + h^F(x)]$ . Theorem III.2 applies equally under the condition of uniform exponential relaxation. ■

### C. Even and odd variables: Example

For models whose configurations do not transform trivially under time reversal, the above construction requires a generalization. If the involution  $\pi$ ,  $\pi^2 = \mathbf{1}$ , is the kinematical time reversal on the state space, the detailed balance condition takes the generalized form,

$$L_0^+ = \pi L_0 \pi, \quad \rho_0 \pi = \rho_0, \quad (3.30)$$

for the adjoint  $L_0^+$  in the sense  $\int \bar{f}(L_0 g) \rho_0 dx = \int \overline{(L_0^+ f)} g \rho_0 dx$ . Typical examples are dissipative mechanical systems, e.g., underdamped diffusion processes, with states given by both position and momentum variables for which  $\pi$  turns the sign of the momentum. Models of heat conduction are a prime example, and the close-to-equilibrium analysis would imply Fourier's law—if indeed the equilibrium time correlation functions can be controlled in the Green–Kubo formula, which remains highly nontrivial, see, e.g., Refs. 1 and 4.

Since the McLennan form is fundamentally a consequence of the transient fluctuation symmetry, see Sec. II B, the above arguments need only minor changes. Instead of repeating the whole derivation, we restrict ourselves to a simple example that elucidates essential points.

We consider the model of a linear *RLC*-circuit with two resistors in series and one external voltage ( $E$ ); cf. Sec. 3.3 in Ref. 2. The two independent free variables are the potential  $U$  (even) over the first resistance  $R_1$  and the current  $I$  (odd) through the second resistance  $R_2$ . To have a nontrivial transient regime, an inductance  $L$  is added in series with the resistors and also a capacitance  $C$  is connected in parallel with the first resistor. The environment is at inverse temperature  $\beta$ .

The stochastic dynamics for  $(U, I)$  is given by Kirchoff's laws combined with the Johnson–Nyquist theory according to which resistances  $R_{1,2}$  immersed in a thermal environment are source of an extra random voltage; these can be modeled as independent Brownian motions with variance  $R_{1,2} \beta^{-1}$  per unit time. Altogether,

$$CdU_t = \left( I_t - \frac{U_t}{R_1} \right) dt + (\beta R_1)^{-1/2} dB_{1,t},$$

$$LdI_t = (E - R_2 I_t - U_t) dt + (\beta R_2)^{-1/2} dB_{2,t}. \quad (3.31)$$

The reference equilibrium dynamics corresponds to  $E=0$ , for which the stationary density reads  $\rho_0(U, I) \propto \exp\{-(CU^2 + LI^2)/2\}$  and the detailed balance condition (3.30) is verified with  $\pi(U, I) = (U, -I)$ . In the driven case,  $E \neq 0$ , and by the linearity of the example it is again easy to compute the stationary density  $\rho_E$ . Instead we use this example to illustrate and to verify the McLennan proposal (1.1).

The goal is to get  $\rho_E$  from physically identifying the transient entropy flux. The irreversible part of the dissipation here equals  $\beta$  times the work done by the battery as a function of the trajectory  $\omega = \{(U_t, I_t), t \in [0, T]\}$ . Specifically,

$$S_{\text{IRR}}^T(\omega) = \beta E \int_0^T I_t dt. \quad (3.32)$$

Of course, when in doubt, one can compute it also by the general algorithm as the time-antisymmetric part of the logarithmic density of the path measure with respect to its time reversal, cf. (2.5)–(2.10). By the linearity and homogeneity of the system, the source voltage  $E$  takes over the driving parameter  $\varepsilon$  of the previous analysis and, using the notation from there, the expected irreversible part of the dissipation is

$$\langle S_{\text{IRR}}^T \rangle_\mu = E \int_0^T \langle w_1(I_t) \rangle_\mu^0 dt, \quad w_1 = \beta I, \quad (3.33)$$

cf. (2.9) and (2.10) or (3.4) and (3.5). It is now easy to check that the stationary density satisfies the identities

$$\begin{aligned} \log \frac{\rho_E}{\rho_0}(U, I) &= -E \left\langle \int_0^{+\infty} w_1(x_t) dt \right\rangle_{U, I}^{0,+} + O(E^2) \\ &= \beta E (L_0^+)^{-1} I + O(E^2) \\ &= \beta E \left( \frac{R_1 C}{R_1 + R_2} U + \frac{L}{R_1 + R_2} I \right) + O(E^2), \end{aligned} \quad (3.34)$$

where the equilibrium dynamics on the first line runs according to the generator  $L_0^+$ , i.e., backward in time. Since it is identical to  $\pi L_0 \pi$ , it only generates sign changes in the computation—that is, the only point in which the previous analysis must be modified. The result of the McLennan theory thus gives  $\rho_E$  correctly as the Gaussian density with the same covariance matrix as  $\rho_0$  but with nonzero averages,

$$\langle I \rangle = \frac{E}{R_1 + R_2}, \quad \langle U \rangle = \frac{R_1 E}{R_1 + R_2}. \quad (3.35)$$

The point is that even in cases where computations would be more involved, the McLennan formula is in terms of a physical quantity that can often be written down without the need to go much into further details of the model.

#### IV. GENERALIZATION BEYOND CLOSE TO EQUILIBRIUM

It is natural to ask whether similar representations of the stationary distribution remain valid also beyond close to equilibrium. The correction to quadratic order has been systematically explored in Refs. 8–10, within the program of steady state thermodynamics. Here we add a general expression, (4.2) below, from which a cumulant expansion around equilibrium could be started in principle.

We look back at (2.5) and (2.6) that we now write as

$$d\mathcal{P}_\mu(\omega) = d\mathcal{P}_\mu^0(\omega) e^{-A(\omega)},$$

where the equilibrium reference  $\mathcal{P}_\mu^0$  is starting from the law  $\mu$ . We decompose the action  $A$  into a time-antisymmetric and a time-symmetric part,

$$S = A\theta - A, \quad \mathcal{T} = A\theta + A,$$

where we have abbreviated (2.7) to  $S$ . From (2.6), the time-symmetric part  $\mathcal{T}$  is the time-integrated excess in escape rates for jump processes. It is more generally related to the dynamical activity in the process. We have also called it traffic in the context of dynamical fluctuation theory.<sup>15,17</sup> We thus have

$$d\mathcal{P}_\mu = d\mathcal{P}_\mu^0 \exp\left(\frac{S-T}{2}\right), \quad d\mathcal{P}_{\rho_0} \circ \theta = d\mathcal{P}_0 \exp\left(-\frac{S+T}{2}\right), \quad (4.1)$$

where the second identity uses the time-reversal invariance of the equilibrium process (started at  $\rho_0$ ). The integrated form of the second identity reads

$$\langle f(\omega) \rangle_{\rho_0} = \langle f(\theta\omega) e^{-(1/2)(S(\omega)+T(\omega))} \rangle_0,$$

with the rightmost expectation over the equilibrium process. The left expectation is for the non-equilibrium process starting in  $\rho_0$  so that, taking functions  $f(\omega) = f(\omega_t)$  of the state at a single time  $t$ , we get information about the approach to the nonequilibrium stationary density. For the finite state space  $\Omega$ , taking  $f(\omega) = \delta_{x_T=x}$ , we get

$$\text{Prob}(x_T = x) = \rho_T^\varepsilon(x) = \rho_0(x) \langle e^{-(1/2)(S+T)} \rangle_x^0, \quad x \in \Omega, \quad (4.2)$$

which is an exact formula for the nonequilibrium density at time  $T$ , no matter how far from equilibrium, entirely in terms of the reference equilibrium dynamics starting at  $x$ . Recall that the exponent in the average is extensive in (large) time  $T$  while also of (small) order  $\varepsilon$  around equilibrium. Together with the normalization condition,

$$\langle e^{1/2(S-T)} \rangle_\mu^0 = 1 \quad (4.3)$$

for all initial laws  $\mu$ , as follows from the first identity in (4.1), this can be taken as the starting point for systematic expansions. In the first order around equilibrium, the expected entropy flux  $S$  and the expected traffic  $T$  are identical. That explains how (2.17) and hence McLennan's formula follow from (4.2). That finally is why we call (IV.2) a generalization, still involving the irreversible entropy flux  $S$  but now also the traffic  $T$ . Note that all expectations in (4.2) and (4.3) are under the equilibrium process, in contrast to a slightly different construction suggested in Ref. 8 which only uses the variable entropy production but with respect to the full nonequilibrium dynamics. We see that difference, equilibrium versus nonequilibrium expectation, also when comparing (4.2) with (2.17) where  $S_{\text{IRR}}^T = S$ ,

$$\langle e^{-S_{\text{IRR}}^T} \rangle_x = \langle e^{-(1/2)(S+T)} \rangle_x^0.$$

Yet, starting from either of the two expressions, there remains the difficulty of writing down a general and physically meaningful expression of the stationary distribution  $\rho$  in a sufficiently explicit way, also because of generic nonlocal aspects in the relation between potential and stationary density, given a nonequilibrium driving.<sup>16</sup>

*Remark IV.1:* The expression (4.2) or the original McLennan formula (1.1) is perhaps related to what is argued to follow from a maximum entropy principle, cf. Refs. 3 and 5. Remark, however, that finding the correct constraints or observables for which to apply such a principle remains highly unclear. In particular, it appears that certainly beyond linear order around equilibrium, also the time-symmetric fluctuation sector must get involved, cf. Ref. 19.

## V. CONCLUSIONS

McLennan's formula gives a Gibbsian-like expression for the steady law close to equilibrium under the condition of local detailed balance. The correction to equilibrium involves the transient entropy flux. It was seen before in Ref. 8 how that arises from a transient fluctuation formula. We have added mathematical precision in the order of limits (time versus distance from equilibrium). We have shown how it relates to local equilibrium and to the Green-Kubo relations, and we have presented a generalization involving the dynamical activity.

In our opinion it remains important to attempt a thermodynamic interpretation of also higher order corrections to equilibrium.

## ACKNOWLEDGMENTS

The authors thank T. S. Komatsu, N. Nakagawa, S. Sasa, and H. Tasaki for very fruitful discussions. K.N. acknowledges the support from the Grant Agency of the Czech Republic (Grant No. 202/07/0404). C.M. benefits from the Belgian Interuniversity Attraction Poles Programme P6/02.

- <sup>1</sup>Bonetto, F., Lebowitz, J. L., and Rey-Bellet, L., in *Fourier's Law: A Challenge for Theorists, Mathematical Physics 2000*, edited by Fokas, A., Grigoryan, A., Kibble, T., and Zegarlinsky, B. (Imperial College Press, London, 2000), pp. 128–151.
- <sup>2</sup>Bruers, S., Maes, C., and Netočný, K., “On the validity of entropy production principles for linear electrical circuits,” *J. Stat. Phys.* **129**, 725 (2007).
- <sup>3</sup>Dewar, R., “Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states,” *J. Phys. A* **36**, 631 (2003).
- <sup>4</sup>Kundu, A., Dhar, A., and Narayan, O., “Green-Kubo formula for heat conduction in open systems,” *J. Stat. Mech.* **2009**, L03001.
- <sup>5</sup>Evans, R. M. L., “Rules for transition rates in nonequilibrium steady states,” *Phys. Rev. Lett.* **92**, 150601 (2004).
- <sup>6</sup>Eyink, G. L., Lebowitz, J. L., and Spohn, H., *Chaos/Xaoc. Soviet-American Perspectives on Nonlinear Science* (American Institute of Physics, New York, 1990), pp. 367–391.
- <sup>7</sup>Kipnis, C. and Landim, C., *Scaling Limits of Interacting Particle Systems* (Springer-Verlag, Berlin, 1998).
- <sup>8</sup>Komatsu, T. S. and Nakagawa, N., “An expression for stationary distribution in nonequilibrium steady states,” *Phys. Rev. Lett.* **100**, 030601 (2008).
- <sup>9</sup>Komatsu, T. S., Nakagawa, N., Sasa, S., and Tasaki, H., “Representation of nonequilibrium steady states in large mechanical systems,” *J. Stat. Phys.* **134**, 401 (2009).
- <sup>10</sup>Komatsu, T. S., Nakagawa, N., Sasa, S., and Tasaki, H., “Steady state thermodynamics for heat conduction—microscopic derivation,” *Phys. Rev. Lett.* **100**, 230602 (2008).
- <sup>11</sup>Lebowitz, J. L. and Bergmann, P. G., “Irreversible Gibbsian ensembles,” *Ann. Phys.* **1**, 1 (1957).
- <sup>12</sup>Lebowitz, J. L., “Stationary nonequilibrium Gibbsian ensembles,” *Phys. Rev.* **114**, 1192 (1959).
- <sup>13</sup>Lebowitz, J. and Spohn, H., “A Gallavotti–Cohen type symmetry in large deviation functional for stochastic dynamics,” *J. Stat. Phys.* **95**, 333 (1999).
- <sup>14</sup>Maes, C. and Netočný, K., “Time-reversal and entropy,” *J. Stat. Phys.* **110**, 269 (2003).
- <sup>15</sup>Maes, C. and Netočný, K., “Canonical structure of dynamical fluctuations in mesoscopic steady state,” *Europhys. Lett.* **82**, 30003 (2008).
- <sup>16</sup>Maes, C., Netočný, K., and Shergelashvili, B., “On the nonequilibrium relation between potential and stationary distribution for driven diffusion,” *Phys. Rev. E* **80**, 011121 (2009).
- <sup>17</sup>Maes, C., Netočný, K., and Wynants, B., “Steady state statistics of driven diffusions,” *Physica A* **387**, 2675 (2008).
- <sup>18</sup>Maes, C., Redig, F., and Van Moffaert, A., “On the definition of entropy production, via examples,” *J. Math. Phys.* **41**, 1528 (2000).
- <sup>19</sup>Maes, C. and van Wieren, M. H., “Time-symmetric fluctuations in nonequilibrium systems,” *Phys. Rev. Lett.* **96**, 240601 (2006).
- <sup>20</sup>McLennan, J. A., Jr., “Statistical mechanics of the steady state,” *Phys. Rev.* **115**, 1405 (1959).
- <sup>21</sup>McLennan, J. A., Jr., *Introduction to Nonequilibrium Statistical Mechanics* (Prentice-Hall, Englewood Cliffs, NJ, 1989).
- <sup>22</sup>Tasaki, S. and Matsui, T., *Fundamental Aspects of Quantum Physics* (Waseda University, Tokyo, 2001); *QP-PQ: Quantum Probability White Noise Analysis* (World Scientific, River Edge, NJ, 2003), Vol. 17, p. 100.