Elements of a unified framework for response formulae

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Elements of a unified framework for response formulae

Matteo Colangeli[1]**, Valerio Lucarini**[2,3]

[1] Dipartimento di Matematica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

[2] Meteorologisches Institut, University of Hamburg

KlimaCampus, Grindelberg 7, 20144 Hamburg, Germany

[3] Department of Mathematics and Statistics

University of Reading, Reading, RG6 6AX, UK

E-mail: colangeli@calvino.polito.it, valerio.lucarini@zmaw.de

Abstract.   We provide a physical interpretation of the first and second order terms occurring in Ruelle's response formalism. We show that entropy fluxes play a major role in determining the response of the system to perturbations. Along this line, we show that our framework allows one to recover a wealth of previous results of response theory in both deterministic and stochastic contexts. In particular, we are able to shed light on the crosslinks between the dynamical systems approach á la Ruelle and large deviations methods.

1. Introduction

A crucial endeavour of statistical mechanics concerns the description of the response of a system to an external force perturbing the baseline dynamics, in terms of physically accessible quantities. The wealth of results collected in this long-standing research field has been organized within the general framework of response theory. Namely, response theory represents a variety of technical tools enabling one to compute the change, when the perturbation is added, in the expectation value of an observable from the knowledge of the invariant measure of the unperturbed dynamics and of the structure of the perturbation. One first attempt in this direction dates back to the seminal work of Kubo [1], who introduced a perturbative theory tackling the derivation of response formulae for Hamiltonian thermostatted systems subjected to an external force. Kubo's derivation unveiled, in particular, one version of the celebrated Fluctuation-Dissipation Theorem (FDT) [2, 3], which establishes a conceptually rich and a practically useful connection between the (linear) response of a system to external perturbations and the equilibrium fluctuations of suitably defined observables [4]. As discussed in Ref. [5], the chance of bridging the external forcing with the fluctuations computed along the unperturbed dynamics mostly relies on the assumption that the invariant measure of the unperturbed system is smooth. Therefore, when considering chaotic dissipative dynamical systems, equipped with an invariant measure with support on a strange attractor, the FDT no longer holds, in general, as shown by Ruelle in Ref. [6]. To first order in the perturbation, in fact, the response formula results, namely, from the sum of two terms. The first one can be cast into a correlation function evaluated with respect to the unperturbed measure along unstable and neutral manifolds, and can be regarded as the natural nonequilibrium extension of the correlation function occurring in equilibrium theory. The second term, instead, involves the computation of the statistical properties of the dynamics along the stable manifold: this has no counterpart in equilibrium. The foreseen breaking of the FDT for dissipative deterministic dynamics rules out, in principle, the possibility of interpreting the response solely from the knowledge of the unperturbed steady state dynamics,

thus hindering the investigation of such systems which are ubiquitous in nature. Note that, just to take a relevant example in a geophysical setting, the validity of a FDT would allow, in principle, to compute forced climate change from the investigation of natural climate variability. Nonetheless, few examples have been reported in the literature, witnessing that the FDT seemingly survive also beyond equilibrium [7, 8].

The problem of lack of smoothness of the invariant measure can be circumvented through different routes. One may, in first instance, introduce some noise on top of the deterministic dissipative dynamics, so as to mimic, say, the effect of round-off errors in numerical algorithms or the presence of unresolved scales [9]. The latter perspective is adopted, in particular, in the approach pursued by Zwanzig [10] in his projection operator formalism, in which the dynamics of the macroscopic, physically relevant, variables is triggered also by a term, typically regarded as noise, which echoes the intrinsic coupling with the neglected, more microscopic, degrees of freedom. Hence, the introduction of a small amount of noise, which can be motivated on physical grounds, allows one to restore the invoked smoothness of the invariant measure, thus making the FDT still applicable [11, 12].

Nevertheless, even without adding noise, some recent findings on the FDT for dissipative dynamical systems [13] confirm that the link between response and fluctuations computed with respect to the unperturbed invariant measure can be safely restored for most of systems of interest in Physics. The reason is that, in statistical mechanics, one typically deals with projected dynamics, and these are associated with regular probability distributions in the corresponding lower dimensional spaces. A major focus of this paper is on the derivation of a general formalism able to encompass former results obtained for deterministic as well as stochastic dynamics. We aim, in particular, at providing a more straightforward interpretation of the response formalism in terms of observables of clearer physical relevance. In this direction, in a recent work [14] focusing on the stochastic dynamics of open mesoscopic systems, it was shown that the response of a generic observable can be cast, at the various orders, in terms of correlation functions featuring two main time-extensive quantities: the (excess) entropy flux, which can be understood within the standard thermodynamic

framework, and a novel term, called dynamical activity, which still lacks a conclusive physical interpretation. This work represents a step forward along that direction: we take the point of view of deterministic mechanics and provide some of the terms appearing in Ruelle's original response formula with a physical content. To this aim, as discussed above, we assume the existence of a smooth invariant measure, and discuss the resulting response formula up to the second order. We thus show that, at the first order, a prominent role is played by the entropy production, hence recovering former results obtained for deterministic [15] as well as stochastic dynamics [16]. At the second order, the expectation value of a given observable is ruled by two time-symmetric terms. The first term is related to the first correction to the average entropy production of the system [17], whereas the second term keeps track of the phase space fluctuations of the entropy production. Although an obvious correspondence with time-symmetric dynamical activity, introduced in [14] to deal with stochastic diffusive dynamics, is still missing, the formalism developed below allows one to pave the a promising bridge between different methods used in response theory.

The work is organized as follows.

In Section 2 we present a rewriting of the linear response introduced in Ref. [17], and we evidence the role the so-called dissipative flux as being the relevant observable entering the linear order response formalism.

In Section 3 we show how our results can be used to ascertain and interpret the linear response to perturbations in various deterministic settings.

The case of stochastic dynamics is analyzed in Section 4, in which we show that our framework is consistent with the response formulae obtained from large deviations methods.

In Section 5 we present some results pertaining to the second order terms and shed light on the onset of two different quantities.

Conclusions are drawn in Sec. 6.

2. Entropy production and Linear Response Theory

We start by concisely recalling Ruelle's approach to response theory in deterministic dynamical systems [6, 17, 18]. Let $(\mathcal{U}, S_o^t, \mu_o)$ be a dynamical system, with $\mathcal{U}$ denoting a compact phase space, $S_o^t : \mathcal{U} \to \mathcal{U}$ a one-parameter group of diffeomorphisms and μ_o the invariant natural measure. Let also $x \in \mathcal{U}$ denote a generic phase space point $x = (q_1, ..., q_N, p_1, ..., p_N)$. We assume, from here onwards, that the measure μ_o is absolutely continuous with respect to the Lebesgue measure, i.e. it is equipped with an invariant density $\rho_o(x)$ such that

$$\mu_o(dx) = \rho_o(x)dx \quad . \tag{1}$$

We consider, then, the effect of adding a small (possibly time-depedent) perturbation $f_t(x)$ at time $t = 0$, which induces the following structure of the equations of dynamics:

$$\dot{x}_t = F(x_t) + f_t(x_t) \quad , \tag{2}$$

where we used the shorthand notation $x_t = x(t)$, and F denotes the drift of the unperturbed dynamics. In the following, when convenient, we will also split the perturbation as: $f_t(x) = h_t X(x)$, where h_t is a time modulation of the phase space function $X(x)$. Because the perturbation is small, one may follow a perturbative approach to express the change in measure μ_o induced by the perturbation itself. Following Ruelle's arguments [17, 19], the response in a generic observable $B : \mathcal{U} \to \mathbb{R}$ may be expressed as a perturbative expansion:

$$\langle B(t) \rangle^h = \langle B(t) \rangle^o + \sum_{n=1}^{\infty} \langle \delta B(t) \rangle_n^h$$

where the superscript "h" on the lhs of the above equality is meant to recall that the average is computed wrt the perturbed density, whereas, on the rhs, the term $\langle B(t) \rangle^o = \int dx_0 \rho_o(x_0) B(x_t)$ denotes the expectation value of $B(t)$ wrt the unperturbed density. The terms $\langle \delta B(t) \rangle_n^h$ attain the formal structure:

$$\langle \delta B(t) \rangle_n^h = \int_{-\infty}^{+\infty} ds_n ... \int_{-\infty}^{+\infty} ds_1 G^{(n)}(s_1, ..., s_n) h_{t-s_1} ... h_{t-s_n} \quad . \tag{3}$$

The nth order Green function, in particular, can be read off explicitly:

$$G^{(n)}(s_1, ..., s_n) = \int dx \rho_o(x) \theta(s_1) ... \theta(s_n - s_{n-1}) \times$$

$$\times \Lambda\Pi(s_n - s_{n-1})...\Lambda\Pi(s_2 - s_1)\Lambda\Pi(s_1)B(x) \quad , \tag{4}$$

where $\theta(t)$ denotes the heaviside step function, and with:

$$\Lambda\Phi = X(x) \cdot \frac{\partial}{\partial x}\Phi \qquad \text{and} \qquad \Pi(t)\Phi = \Phi \circ S_o^t \quad . \tag{5}$$

At the linear order, Eq. (3) simplifies into:

$$\langle \delta B(t) \rangle_1^h = \int_0^{+\infty} G^{(1)}(t-s)h_s ds = \int_0^t R(t-s)h_s ds \tag{6}$$

where $R(t)$ denotes the so-called *response function* and $G^{(1)}(t) = \theta(t)R(t)$. We recall that, by using the Kramers-Kronig relations [20], the principle of causality boils down, in the frequency domain, to the following alternative relations between the Fourier transforms $\chi(\nu) = \mathcal{F}(G(t))$ and $\hat{R}(\nu) = \mathcal{F}(R(t))$:

$$\hat{R}(\nu) = 2\mathrm{Re}\{\chi(\nu)\} \quad \text{or} \quad \hat{R}(\nu) = 2\mathrm{Im}\{\chi(\nu)\} \quad , \tag{7}$$

depending on whether the response function is, respectively, even or odd under the time-reversal, cf. [2, 19]. A simple calculation shows, then, that, at the first order, the expansion (3) leads to the familiar FDT:

$$\langle \delta B(t) \rangle_1^h = \int_0^t h_s ds \int \rho_o(x_0)X(x_0)\left(\frac{\partial}{\partial x_0}B(x_{t-s})\right)dx_0 =$$

$$= \int_0^t ds \int \sigma_s(x_s)B(x_t)\rho_o(x_0)dx_0 = \langle B(x_t)S(\omega)\rangle^o \quad , \tag{8}$$

where we introduced the *dissipative flux* $\sigma_s(x_s)$:

$$\sigma_s(x_s) = h_s\gamma(x_s) \quad , \tag{9}$$

with

$$\gamma(x) = -\frac{1}{\rho_o(x)}\left[\frac{\partial}{\partial x} \cdot (X(x)\rho_o(x))\right] \quad . \tag{10}$$

In Eq. (8) we have denoted by $S(\omega)$ the integral of σ_s over the path $\omega = (x_s, s \in [0,t])$ started from x_0: $S(\omega)$ is a phase function which, under suitable assumptions (e.g., when the perturbation is conservative, cf. Sec. 3), can be associated to the total entropy produced along the path. Moreover, from Eqs. (6) and (8), one also obtains the formal expression for the response function:

$$R(t-s) = \langle B(x_t)\gamma(x_s)\rangle^o \quad . \tag{11}$$

The phase function $\sigma_s(x_s)$, introduced in (9), can be split, using Eq. (10), as follows:

$$\sigma_s(x_s) = \underbrace{-\frac{\partial}{\partial x_s} \cdot f_s(x_s)}_{\sigma_A} + \underbrace{f_s(x_s) \cdot \frac{\partial}{\partial x_s}\left(-\log \rho_o(x_s)\right)}_{\sigma_B} \quad . \tag{12}$$

While the term σ_A, in Eq. (12), corresponds to a purely dissipative contribution, induced by nonconservative perturbations leading to dissipative dynamics, the term σ_B, in turn, which contains the function $(-\log \rho_o)$ (referred to, in the literature, as the *information potential*, cf. Sec. 4.2), is related to the total entropy produced by the perturbation and released into the environment, regardless of whether the perturbation is conservative of not (this is, actually, the term originally introduced in Kubo's theory [2]). Noticeably, the expression in Eq. (12) recovers the original Agarwal formula [21], which hence corresponds to the linear order contribution of Ruelle's formal expansion. As also outlined in [22], the use of Eq. (12) is typically hindered by the lack of the knowledge of the reference density ρ_o and, possibly, of the details of the perturbed dynamics triggered by the phase function $X(x)$. Nevertheless, the Agarwal formula (12) has enjoyed a growing popularity in the literature and was, since then, derived following different routes, cf. [3, 7, 23, 24, 25, 26].

Moreover, we also observe that the observable $S(\omega)$ formally resembles the structure of the so-called *Dissipation Function* introduced in Ref. [15] (cf. Eq. (9) therein), and for which transient and steady state Fluctuation Relations have been proven under rather general conditions [27].

Thus, Eq. (12) suggests that, to first order, the linear response can be cast as an equilibrium time correlation function between the chosen observable and $\sigma_s(x_s)$, which includes the two different aforementioned source terms σ_A and σ_B. For weakly perturbed dynamical systems, the proposal of expressing the response function as a correlation between the chosen observable and the dissipative flux σ_s can be already traced back to the seminal works of McLennan [28] and Zubarev[29]. We remark, moreover, that this general framework is not restricted to deviations from an *equilibrium* reference state only, but holds for any steady state equipped with a smooth invariant density $\rho_o(x)$. In particular, as it will be discussed in Sec. 3, our approach allows one to make a bridge with the results outlined in Ref. [13], where

the authors considered a reference invariant SRB measure equipped with a smooth marginal probability density, resulting from the projection of the full SRB measure along the direction of the initial impulsive perturbation. We also point out that the response formalism introduced above, based on the use of the operators (5), is prone to be also used in the set-up of stochastic dynamics [17]: the application of the method to diffusive systems is deferred to Sec. 4.

It is worth mentioning two basic properties characterizing the dissipative flux $\sigma_s(x_s)$.

First, note that the expectation $\langle\sigma_s(x_s)\rangle^o$, computed with the unperturbed density, vanishes:

$$\int \rho_o(x)\sigma_s(x)dx = -\int dx \left[\frac{\partial}{\partial x} \cdot (X(x)\rho_o(x))\right] = 0 \quad . \tag{13}$$

The relation (13) allows one, hence, to interpret the observable $\sigma_s(x_s)$ as an "excess" dissipative flux of the perturbed process with respect to the unperturbed one. In more physical terms, $\sigma_s(x_s)$ can be thus regarded as the surplus of entropy production, due to the perturbation, with respect to the "housekeeping" heat flux needed to maintain the steady state [30]. Notice that to ensure that the response function be integrable, one typically requires the correlations to decay sufficiently fast [17].

Next, let us consider the time-reversal symmetry property of $\sigma_s(x_s)$. To this aim, following Roberts *et al.* [31], we define the involution G as:

$$D_G \cdot f_t = -f_t \circ G \quad , \tag{14}$$

with $G \circ G = 1$, where D_G denotes the Jacobian matrix of G ‡.

A dynamical system is said to be *reversible* if there exists an involution $G : \mathcal{U} \to \mathcal{U}$ fulfilling Eq. (14) (i.e. it reverses the direction of time). Using (14) one thus finds:

$$\sigma(Gx) = -\sigma(x) \quad , \tag{15}$$

where we assumed the unperturbed density to be invariant under the involution G, $\rho_o(x) = \rho_o(Gx)$ §. Equation (15) indicates that the phase function $\sigma_t(x)$ is, as expected, odd under

‡ It is worth recalling that, in the discrete time case, the dynamical flow is replaced by a mapping $M : \mathcal{U} \to \mathcal{U}$, and Eq. (14), correspondingly, takes the form: $G \circ M \circ G = M^{-1}$ [32, 33].

§ The property of invariance of ρ_o under the involution G was called, in Ref. [33], as *phase space detailed balance*, because it reduces, via a suitable projection onto the space of stochastic dynamics, to the *detailed balance* relation, which stands as the hallmark of an equilibrium stochastic dynamics [14].

time-reversal, and so is, therefore, its time integrated value $S(\omega)$, over the path ω [32]:

$$S(G\omega) = -S(\omega)$$

3. Deterministic dynamics

The formalism developed in Sec. 2 can be applied to a wealth of different physical situations, in which one considers the effect of adding a small perturbation to a reference dynamics, enjoying a steady state. It is worth remarking that, in order to apply the formalism introduced above, one merely requires that the reference steady state be equipped with a density. In this Section we will investigate the application of the formalism described above in deterministic dynamical systems, enjoying either an equilibrium or a nonequilibrium steady states. We will restrict ourselves to the linear case.

3.1. Hamiltonian dynamics

It is instructive to address the case in which the reference microscopic dynamics is Hamiltonian and the steady state is an equilibrium one. Thus, we consider the effect of adding a small conservative perturbation, to be expressed as the gradient of a scalar potential function $V(x)$. This corresponds to replacing the Hamiltonian $H_0(x)$ as:

$$H_0(x) \to H_0(x) - h_t V(x) \quad .$$

Therefore, in Eq. (2), one sets $F(x) = S\nabla H_0(x)$, where S denotes the symplectic matrix, and $f_t(x) = -h_t S\nabla V(x)$. Next, by taking, for simplicity, the canonical distribution as the reference equilibrium density, i.e. $\rho_o(x) = Z^{-1}e^{-\beta H_0(x)}$ (where Z denotes the canonical partition function and $\beta = 1/k_B T$ is the inverse temperature), a straightforward calculation reveals that the term σ_A in (12) vanishes and the term σ_B attains the form:

$$\sigma_s(x_s) = \beta h_s \dot{V}(x_s) \quad , \tag{16}$$

where we introduced the shorthand notation $\dot{B}(x_t) = dB(x_t)/dt$. Thus, by inserting (16) in (8), one finds:

$$\langle \delta B(t)\rangle_1^h = \beta \int \rho(x_0)dx_0 B(x_t) \int_0^t ds \dot{V}(x_s)h_s$$

$$= \beta \int \rho(x_0)dx_0 B(x_t) \left[(V(x_t)h_t - V(x_0)h_0) - \int_0^t ds V(x_s)\dot{h}_s \right]$$

$$= \beta \langle S(\omega)B(x_t) \rangle^o \quad . \tag{17}$$

Therefore, for Hamiltonian dynamics, Eq. (11) reduces to the classical Green-Kubo form [2, 3, 30]:

$$R(t - s) = \beta \frac{d}{ds} \langle B(x_t)V(x_s) \rangle^o \quad . \tag{18}$$

In (17), the term $(V(x_t)h_t - V(x_0)h_0)$ corresponds to the extra change of energy in the environment due to the perturbation, whereas $\int_0^t ds V(s)\dot{h}_s$ is the work done by the perturbation. Therefore, the case of conservative perturbations shows that the linear response term can be effectively cast into an equilibrium correlation between the given observable and the total entropy produced, over the path, by the perturbation and released into the environment, cf. [22, 30]. Let us consider a reference Hamiltonian given by $H_0 = \sum_{i=1}^{N} p_i^2/(2m)$ (with m denoting the mass of the particles), which, after the addiction of a time independent, homogeneous, electric field E, takes the form $H_0 \to H_0 - \sum_{i=1}^{N} \kappa_i(q_i \cdot E)$, with κ_i denoting the charge of the i-th particle. The application of Eq. (10) leads, thus, to the standard Kubo expression $\sigma(x) = \beta E \cdot J$, where the electric current J takes the form $J = \sum_{i=1}^{N} (p_i/m \cdot \kappa_i)$.

3.2. Dissipative dynamics

Let us consider, next, the case of a dissipative steady state dynamics, perturbed by a small impulsive perturbation, which modifies the initial condition as $x_0 \to x_0 + \delta x_0$. Therefore, by setting $f_s = \delta(s)\delta x_0$, one finds:

$$\sigma_s(x_s) = -\delta(s) \left(\frac{\partial \log \rho(x_s)}{\partial x_s} \cdot \delta x_0 \right) \quad , \tag{19}$$

which leads to the response formula outlined in Ref. [13, 34]. One natural objection to the expression (19) points to the fact that the invariant measure of a chaotic dissipative system μ_o is singular with respect to the Lebesgue measure, and is typically supported on a fractal attractor. This may, hence, prevent the application of the approach outlined in Sec. 2 to dissipative systems. Indeed, the standard FDT ensures that the statistical features of a

perturbation are related to the statistical properties of the unperturbed system, but that cannot be the case, in general, in dissipative systems. The reason is that, given an initial state x_0 on the attractor and a generic impulsive perturbation δx_0, the perturbed initial state $\tilde{x}_0 = x_0 + \delta x_0$ and its time evolution may lie outside the support of the measure, hence their statistical properties cannot be expressed by μ, which attributes vanishing probability to such states. Nevertheless, in Ref. [13] it was shown that a different route is possible to compute the response, which is worth recalling shortly here. By denoting as $\rho_o(x_0; \delta x_0) = \rho_o(x_0 - \delta x_0)$ the perturbed initial density, and by $W(x_0, 0 \to x_t, t)$ the transition probability determined by the dynamics, one may express the response of the coordinate x^i as:

$$\langle \delta x^i(t) \rangle^h = \int \int x_t^i \left[\rho_o(x_0 - \delta x_0) - \rho_o(x_0) \right] W(x_0, 0 \to x_t, t) dx_0 dx_t \quad .(20)$$

By also assuming, for sake of simplicity, that all components of δx_0 vanish, except the i-th component, one finds that Eq. (20) can be written as:

$$\langle \delta x^i(t) \rangle^h = \int x_t^i (\widetilde{\rho}_t(x^i; \delta x_0) - \widetilde{\rho}_o(x_t^i)) dx_t^i \quad , \tag{21}$$

where $\widetilde{\rho}_o(x_t^i)$ and $\widetilde{\rho}_t$ are the marginal probability distributions defined by:

$$\widetilde{\rho}_o(x_t^i) = \int \rho_o(x_t) \prod_{j \neq i} dx_t^j \quad , \qquad \widetilde{\rho}_t(x_t^i; \delta x_0) = \int \rho_t(x_0; \delta x_0) \prod_{j \neq i} dx_t^j \quad .$$

Thus, since projected singular measures are expected to be smooth [13, 35], especially if the dimension of the projected space is sensibly smaller than that of the original space, one finds that a FDT, written in the form of Eq. (8), can typically be extended to a large fraction of the dissipative deterministic systems of interest in Physics.

4. Stochastic diffusions and large deviations

Let us now turn our attention to stochastic dynamics. In general, the presence of noise allows one to characterize the steady state dynamics, even in presence of dissipation, by regular probability densities, thus overcoming the typical difficulties encountered in deterministic dynamical systems. A detailed analysis of the response formulae valid for Markovian Langevin-type stochastic differential equations is given in Ref. [36], where an expansion

formally resembling the structure of Eqs. (4) and (5) is developed. In particular, in Ref. [36], Ruelle shows that, in the zero noise limit, the various terms of the expansion reproduce the corresponding order terms pertaining to the deterministic dynamics discussed in Sec. 2. This occurs because, under suitable assumptions, the SRB states are stable under small random perturbations [37, 38]. It is therefore tempting to use the formalism of Sec. 2 to describe stochastic models amenable to an analytical solution, so as to compare our results with those obtained using other methods, e.g. the path-integral formalism described in Refs. [14, 22]. We will thus focus, first, on stochastic diffusion processes described by overdamped Langevin equations, in which one disregards inertial effects, thus letting forces be proportional to velocities rather than to accelerations [30, 39]. These processes correspond to the high damping limits of the underdamped, or intertial, stochastic dynamics, whose analysis is deferred to Sec. 4.3. Let us start considering overdamped diffusion processes for the state $x \in \mathbb{R}^n$, defined in the Itô sense by:

$$\dot{x}_t = \chi \cdot [F(x_t) + f_t(x_t)] + \nabla \cdot D(x_t) + \sqrt{2D(x_t)}\,\xi_t \quad , \tag{22}$$

where ξ_t denotes a standard white noise and f_t denotes the perturbation to the reference dynamics. The mobility χ and the diffusion constant D are strictly positive (symmetric) $n \times n$-matrices, which, provided that the system is in contact with a thermostat at inverse temperature $\beta > 0$, are connected via the well-known Einstein relation $\chi = \beta D$. The force F denotes the drift of the reference, unperturbed dynamics, and can be expressed as:

$$F = F_{nc} - \nabla U \quad , \tag{23}$$

where F_{nc} denotes a nonconservative force pulling the reference dynamics out of equilibrium, while U is the energy of the system. The Fokker-Planck equation for the time dependent density ρ_t, relative to the diffusion process described by (22), reads

$$\frac{\partial \rho_t}{\partial t}(x_t) = -\nabla \cdot j_\rho, \quad \text{with} \ \ j_\rho = [\chi(F + f_t)\rho_t(x_t) - \frac{\chi}{\beta}\nabla \rho_t(x_t)] \quad , \tag{24}$$

where j_ρ denotes the probability current [40]. Rather than attempting a direct analytical solution of Eq. (24), one may tackle the analysis of Eq. (22) from the rather different standpoint of large deviations theory [14, 41]. We shorty recapitulate the main steps of the

derivation, cf. Refs. [14, 16] and references therein for a more detailed discussion. The key idea is to determine the perturbed probability density through its embedding in the path-space distribution. That is, given the (random) paths $\omega = (x(s), s \in [0, t])$, one may connect the distribution P on paths starting from ρ_o and subjected to the perturbation f_t, with the reference distribution P^o concerning paths starting from ρ_o and undergoing the reference dynamics, via the formula:

$$P(\omega) = e^{-\mathcal{A}(\omega)} P^o(\omega) \quad . \tag{25}$$

The relation (25) defines the action $\mathcal{A}(\omega)$, which is typically local in space-time and is, thus, similar to Hamiltonians or Lagrangians encountered in equilibrium statistical mechanics, see e.g. [42]. One can also verify that

$$\mathcal{A} = (\mathcal{T} - S)/2 \quad ,$$

where $\mathcal{T}(\omega)$ and $S(\omega)$ are path-dependent quantities corresponding, respectively, to the time-symmetric and time-antisymmetric components of the action. That is, defining the time-reversal operator g as:

$$g\omega = ((\pi x)_{t-s}, 0 \leq s \leq t) \quad , \tag{26}$$

(with πx equal to x except for flipping any other variable with negative parity under time-reversal), one can write:

$$S(\omega) = \mathcal{A}(g\omega) - \mathcal{A}(\omega) , \quad \mathcal{T}(\omega) = \mathcal{A}(g\omega) + \mathcal{A}(\omega) \quad . \tag{27}$$

The quantity $S(\omega)$, under the assumption of local detailed balance [43], is the entropy flux triggered by the perturbation and released into the environment [14]. On the other hand, the quantity $\mathcal{T}(\omega)$ is referred to in the literature as *the dynamical activity* or *traffic* [30, 39] and appears to be much more concerned with kinetics than it is embedded into thermodynamics. For example, in the set-up of Markov jump processes, $\mathcal{T}(\omega)$ is suitable to a physical interpretation: it measures how the escape rate from a trajectory ω changes when the perturbation f_t is added [30].

A simple calculation unveils the following general expression for the action pertaining to the process described by Eq. (22):

$$\mathcal{A}(\omega) = \frac{\beta}{2} \int_0^t ds \left[f_s \cdot \chi F + \nabla \cdot (D f_s) + \frac{1}{2} f_s \cdot \chi f_s \right] - \frac{\beta}{2} \int_0^t dx_s \circ f_s \tag{28}$$

where the last stochastic integral (with the $\circ$) is in the sense of Stratonovich. From (27) and (28), one can derive the following expressions for $S(\omega)$ and $\mathcal{T}(\omega)$:

$$S(\omega) = \beta \int_0^t dx_s \circ f_s \quad \text{and} \quad \mathcal{T}(\omega) = \mathcal{T}_1 + \mathcal{T}_2 \quad , \tag{29}$$

with

$$\mathcal{T}_1 = \beta \int_0^t ds \left[f_s \cdot \chi F + \nabla \cdot (D f_s) \right] \quad \text{and} \quad \mathcal{T}_2 = \frac{\beta}{2} \int_0^t ds f_s \cdot \chi f_s \quad . \tag{30}$$

If the chosen observable B is endowed with an even kinematical parity, the following linear response formula can be thus established [30]:

$$\langle \delta B(t) \rangle_1^h = \langle B(x_t) S(\omega) \rangle^o = - \langle B(x_0) S(\omega) \rangle^o = - \int dx_0 \rho_o(x_0) B(x_0) \langle S(\omega) \rangle_{x_0}^o \quad . \tag{31}$$

The expression (31), thus, inherits the structure of the response formula (8) obtained for deterministic systems. The quantity $\langle S \rangle_{x_0}^o$, in Eq. (31), denotes the conditional expectation of the entropy flux $S(\omega)$ over $[0, t]$ given that the path started from the state x_0. It can also be written as [14, 16]::

$$\langle S \rangle_{x_0}^o = \beta \int_0^t \langle w(x_s) \rangle_{x_0}^o ds \quad , \tag{32}$$

where $w(x_s)$ corresponds to the instantaneous (time-antisymmetric, random) work made by the perturbation f_t. The analysis of some specific models and examples comes next.

4.1. Expansion around detailed balance dynamics

An interesting example is obtained by considering overdamped diffusion processes whose reference dynamics is a (equilibrium) detailed balance dynamics [14], i.e. $F_{nc} = 0$ in Eq. (23). Let us also take the reference distribution of states to be the equilibrium one, $\rho_o(x) \propto e^{-\beta U(x)}$. To further simplify the problem, let us also assume the perturbation f to be time independent

and the matrices χ and D to be independent of x. The quantity $w(x_s)$ pertaining to this stochastic dynamics can be explicitly computed [14, 16]:

$$w(x) = \frac{\chi}{\beta}\nabla \cdot f - \chi f \cdot \nabla U \qquad (33)$$

By inserting the expression (33) into Eq. (32) and by further using Eq. (31), one readily obtains the FDT for the process under consideration. Alternatively, one may adopt the formalism detailed in Sec. (2) to derive the linear response. To this aim, by computing $\sigma_s(x_s)$, given in Eqs. (9) and (10), one immediately recovers the expression for $\langle w(x_s)\rangle^o_{x_0}$ in (32) hence recovering the above response formula.

It is also instructive to consider the case of a (conservative) perturbation changing the potential U into $U - h_t V$. For the case under consideration, the general response formula holds [30]:

$$R(t-s) = \frac{\beta}{2}\frac{d}{ds}\langle V(x_s)B(x_t)\rangle^o - \frac{\beta}{2}\langle LV(x_s)B(x_t)\rangle^o = -\beta\langle LV(x_s)B(x_t)\rangle^o \quad ,(34)$$

where we introduced the (backward) generator L of the process‖, defined as

$$L = -\chi\nabla U \cdot \nabla + \frac{\chi}{\beta}\nabla^2 \quad .$$

To derive the last equality in Eq. (34), one uses the time-reversal symmetry of the equilibrium correlations and the properties of the adjoint generator $L^* = L$: ¶

$$\langle L^*V(x_s)B(x_t)\rangle^o = \langle V(x_s)LB(x_t)\rangle^o = \frac{d}{dt}\langle V(x_s)B(x_t)\rangle^o = -\frac{d}{ds}\langle V(x_s)B(x_t)\rangle^o \quad ,(35)$$

for $s < t$. On the other hand, when turning back to the formalism of Sec. 2, it is not difficult to obtain the following expression for the quantity $\gamma(x)$ defined in (10):

$$\gamma(x) = \beta\chi\nabla U \cdot \nabla V - \chi\nabla^2 V \quad . \qquad (36)$$

Thus, by plugging Eq. (36) into Eq. (11), one immediately recovers Eq. (34).

The chosen examples, pertaining to equilibrium dynamics subjected to time-independent

‖ Hence, from the perspective of large deviations theory, the equilibrium FDT may also be equivalently cast into a *fluctuation-activity* relation, by exploiting the properties of the only time-symmetric term of the action.

¶ L^* is defined with the help of the stationary distribution ρ_o: for any two state functions f and g, L^* is such that $\int dx\rho_o(x)g(x)L^*f(x) = \int dx\rho_o(x)f(x)Lg(x)$. For detailed balance dynamics, in particular, one has $L^* = \pi L\pi$, where π flips the variables which are odd under time-reversal, cf. Eq. (26).

perturbations, corroborate, hence, the equivalence between the linear response formulae derived either from the deterministic Ruelle's expansion or from the above introduced large deviation formalism.

4.2. Nonequilibrium steady states

By setting $F_{nc} \neq 0$, in Eq. (22), one spoils the time-reversibility of the reference dynamics. Therefore, given enough time, the reference dynamics reaches a nonequilibrium steady state described by an invariant density $\rho_o(x)$. The latter is typically not known, nevertheless the approach traced in Sec. 3 allows one to obtain linear response formulae recovering the corresponding expressions obtained via the path-integral formulation outlined above. In the steady state, one can use the definition of the probability current given in Eq. (24), to define the *information potential* $\mathcal{I}_{\rho_o}$ [22, 44] as:

$$\mathcal{I}_{\rho_o} = -\frac{d \log \rho_o}{dx} = \frac{\beta}{\chi} u - \beta F \quad , \tag{37}$$

where $u \equiv j_{\rho_o}/\rho_o$ denotes a probability velocity. Therefore, using Eqs. (11) and (12), with Eq. (37), one finds the general response function for nonequilibrium overdamped diffusion processes:

$$R(t-s) = \chi \left\langle \left[-\frac{d}{dx_s} \cdot f(x_s) + \mathcal{I}_{\rho_o}(x_s) \cdot f(x_s) \right] B(x_t) \right\rangle^o \quad . \tag{38}$$

In particular, if the perturbation takes the gradient form $f = \nabla V$, an easy calculation yields:

$$R(t-s) = \beta \left\langle (u(x_s) \cdot \nabla V(x_s)) B(x_t) \right\rangle^o - \beta \left\langle LV(x_s) B(x_t) \right\rangle^o \quad , \tag{39}$$

with $L = \chi F \cdot \nabla + \chi/\beta \nabla^2$. Next, by using the relations (35), with

$$L^* = -\chi F \cdot \nabla + \frac{\chi}{\beta} \nabla^2 + 2 \frac{\chi}{\beta} \nabla (\log \rho_o) \cdot \nabla = L - 2u \cdot \nabla \quad ,$$

one can suitably transform Eq. (39) into the equivalent form:

$$R(t-s) = -\beta \left\langle (u(x_s) \cdot \nabla V(x_s)) B(x_t) \right\rangle^o + \beta \frac{d}{ds} \left\langle B(x_t) V(x_s) \right\rangle^o \quad , \tag{40}$$

which successfully recovers Eq. (26) of [22]. It is worth remarking that the function $u(x)$, in (37), is unknown in general. Nevertheless, Eq. (39) is noteworthy at a formal level, for it

shows that the response function can be expressed in terms of a suitable correlation function computed wrt to reference stationary density characterizing the nonequilibrium steady state. One also readily notices that Eq. (40) reconstructs the classical Kubo formula (18) when setting $F_{nc} = 0$ (i.e. $u = 0$) or when describing the response in a reference frame moving with drift velocity u, cf. Refs. [22, 45]. Thus, the fist correlation on the rhs of Eq. (40), including the function u, stands as the true nonequilibrium extension of the FDT to nonequilibrium steady states.

4.3. Inertial dynamics: a solvable example

In this paragraph, we apply the formalism detailed in Sec. 4 to tackle the description of underdamped diffusion processes, in which inertial effects are taken into account. We, thus, consider states $(q, p) = (q_1, ..., q_N, p_1, ..., p_N) \in \mathbb{R}^{2N}$ of positions and momenta of N particles, each of which is subjected to a viscous force $(-\nu_i p_i)$, and is coupled with its own heat bath, characterized by a standard white noise ξ_t, with diffusion coefficient D_i and inverse temperature $\beta_i = \nu_i / D_i$, with $i = 1, ..., N$. We will restrict our analysis to the so-called Ornstein-Uhlenbeck diffusion process, which is amenable to an analytical solution. The equations of the unperturbed dynamics read:

$$\dot{q}_i = p_i \quad ,$$
$$\dot{p}_i = -\frac{\partial U(q)}{\partial q_i} - \nu_i p_i + \sqrt{2D_i}\, \xi_t \quad , \tag{41}$$

with $U(q) = 1/2m\omega_0^2 \sum_{i=1}^{N} q_i^2$ denoting, here, the internal energy of the particle system. The Hamiltonian takes, hence, the structure $H_0(q, p) = \sum_{i=1}^{N} p_i^2/(2m) + U(q)$, and the unperturbed (equilibrium) dynamics enjoys an invariant density of the form [40]:

$$\rho_o(q, p) = \frac{1}{\mathcal{Z}} e^{-\beta H_0(q,p)} \quad , \tag{42}$$

with $\mathcal{Z}$ denoting a normalizing factor. We consider, here again, the effect of a perturbation changing the Hamiltonian as $H_0 \to H_0 - h_t V(q)$. By applying the formula (12) to the set of equations (41), and using Eq. (42), one thus obtains:

$$\gamma(q, p) = \frac{\beta}{m} \sum_{i=1}^{N} \nabla_{q_i} V \cdot p_i = \frac{\beta}{m} \nabla_q V \cdot p \quad , \tag{43}$$

where, in the last equality of Eq. (43), we made use of a compact notation. From Eq. (43), one thus obtains the desired response formula:

$$R(t-s) = \frac{\beta}{m} \langle (\nabla_q V(q_s) \cdot p_s) B(q_t, p_t) \rangle^o = \beta \frac{d}{ds} \langle B(q_t, p_t) V(q_s) \rangle^o \quad , \quad (44)$$

which is consistent with previous derivations discussed in Refs. [22, 46].

5. Higher order terms

In this Section, we concentrate on the structure of the second order contribution to the response formulae detailed in sec. 2. Our aim is, again, to provide a physical interpretation of the terms appearing in the formalism beyond the standard FDT, thus complementing the formal properties studied in Ref. [19] and setting the stage for a useful bridge with the large deviation method introduced in Sec. 4. At the second order, Eq. (3) attains the structure:

$$\langle \delta B(t) \rangle_2^h = \int_{-\infty}^{\infty} ds_1 \int_{-\infty}^{\infty} ds_2 G^{(2)}(s_1, s_2) h_{t-s_1} h_{t-s_2} \quad , \quad (45)$$

with

$$G^{(2)}(s_1, s_2) = \theta(s_1)\theta(s_2 - s_1) \int dx_0 \rho_o(x_0) X(x_0) \cdot \frac{\partial}{\partial x_0} \left[X(x_{s_2-s_1}) \cdot \frac{\partial}{\partial x_{s_2-s_1}} B(x_{s_2}) \right] (46)$$

If the reference dynamics is conservative, one finds:

$$\langle \delta B(t) \rangle_2^h = \int_0^{\infty} ds_1 h_{t-s_1} \int_{s_1}^{\infty} ds_2 h_{t-s_2} \times$$
$$\times \left\langle \left[\gamma^{(2)}(x_0, x_{s_2-s_1}) + \chi^{(2)}(x_0, x_{s_2-s_1}) \right] B(x_{s_2}) \right\rangle^o \quad , \quad (47)$$

where we introduced the second order terms $\gamma^{(2)}$ and $\chi^{(2)}$, defined respectively as:

$$\gamma^{(2)}(x_0, x_s) = \gamma(x_0)\gamma(x_s) \quad , \quad (48)$$

$$\chi^{(2)}(x_0, x_s) = -X(x_s)\frac{\partial \gamma(x_0)}{\partial x_s} \quad . \quad (49)$$

At variance with the dissipative flux $\sigma(x)$, which enters the linear response formula (8), one notices, here, the onset of the second order quantities, $\gamma^{(2)}$ and $\chi^{(2)}$, which are even under time-reversal, as it can be easily verified by using Eq. (14). The time integral of the expectation value of $\gamma^{(2)}$, in Eq. (48), which basically corresponds to the time correlation of the dissipative flux computed at time $t = 0$ and at $t = s$, yields the leading non-vanishing

contribution to the excess entropy production [17], cf. also Eq. (13). On the other hand, one also finds that the term (49) can be rewritten as:

$$\chi^{(2)}(x_0, x_s) = -X(x_s)\frac{\partial\gamma(x_0)}{\partial x_0}[D_{S_{\partial}^s}(x_0)]^{-1} \quad .$$

The $\chi^{(2)}$ term is responsible for describing the coupling between the perturbation and the gradient of the dissipative flux: its relevance stems, hence, from the presence of fluctuations of the observable $\gamma(x)$ in the phase space. In this perspective, more insight into the meaning of such term might be thus obtained by referring to the Fluctuation Theorem reported in [15, 27]. From the perspective of the large deviation approach of Sec. 4, the role of time-symmetric quantities becomes also similarly visible when going to the second order. In particular, in Ref. [14], it was shown that the second order term attains the structure:

$$\langle \delta B(t)\rangle_2^h = -\frac{1}{2}\langle B(x_t)S(\omega)\mathcal{T}_1(\omega)\rangle^o \quad , \tag{50}$$

featuring the combined contribution of both the (linear order) time-symmetric and time-antisymmetric components of the action. It is not entirely obvious to establish a neat correspondence between our second order results and those appearing in Eq. (50), mostly because the way the expansion is performed differs between the two methods when nonlinear terms are considered. Nevertheless, it is definitely worth attempting to shed light further on the deterministic interpretation of the dynamical activity term $\mathcal{T}$, whose role in statistical mechanics has been largely unnoticed so far.

6. Conclusions

The analysis of the response of statistical mechanical systems to external perturbations is of crucial relevance for both theoretical reasons and for devising numerical and laboratory experiments. Depending on whether the underlying dynamics is Hamiltonian or dissipative, deterministic or stochastic, a wealth of mathematical techniques have been introduced, in the literature, to obtain perturbative response formulae. While the various mathematical formalisms settle on firm grounds, and are often prone to an algorithmic implementation, an open question addresses the physical interpretation of the terms entering the perturbation

method. In this paper we have unveiled the onset of some recurrent structures occurring in different formalisms and commented on their thermodynamic, or kinetic, foundation.

We have taken as starting point the response formulae proposed by D. Ruelle, through the prism of chaotic dynamical systems theory. While our approach stems from the assumption that the unperturbed dynamics is endowed with a smooth invariant measure, we have also discussed the extension to dissipative deterministic dynamics, where the smoothness of the marginal density pertaining to the projected dynamics becomes crucial. Moreover, we have highlighted the role of the dissipative flux at the first and second order of expansion, for deterministic as well as stochastic diffusion processes. We could, hence, draw a promising line connecting our results to those developed within the set-up of large deviations theory. This is a first promising step of an ambitious program, which definitely calls for further investigation.

Acknowledgments

References

[1] R. Kubo, Statistical-mechanical theory of irreversible processes: I. General theory and simple applications to magnetic and conduction problems *J. Phys. Soc. Japan* **12** 570 (1957).

[2] R. Kubo, The fluctuation-dissipation theorem, *Rep. Prog. Phys.* **29**, 255 (1966).

[3] U. Marini Bettolo Marconi, A. Puglisi, L. Rondoni, A. Vulpiani, Fluctuation-Dissipation: Response Theory in Statistical Physics, *Phys. Rep.* **461**, 111 (2008).

[4] D. N. Zubarev, Nonequilibrium Statistical Thermodynamics, *Consultant Bureau, New York* (1974).

[5] M. Falcioni, A. Vulpiani, The relevance of chaos for the linear response theory, *Phys. A* **215**, 481 (1995).

[6] D. Ruelle, General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium, *Phys. Lett. A* **245**, 220 (1998).

[7] A. Gritsun, G. Branstator, Climate Response using a three-dimensional operator based on the fluctuation-dissipation theorem, *Journal of the Atmospheric Sciences* **64**, 2558 (2007).

[8] A. Gritsun, G. Branstator, and A. Majda, Climate response of linear and quadratic functionals using the fluctuation dissipation theorem, *J. Atmos. Sci.*, **65**, 2824 (2008).

[9] R. Abramov, A. Majda, Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems, *Nonlinearity*, **20**, 2793 (2007).

[10] R. Zwanzig, Nonequilibrium Statistical Mechanics, *Oxford University Press* (2001).

[11] J. Wouters, V. Lucarini, Disentangling multi-level systems: averaging, correlations and memory, *J. Stat. Mech.* (2012) P03003.

[12] J. Wouters, and V. Lucarini, Multi-level dynamical systems: Connecting the Ruelle response theory and the Mori-Zwanzig approach, *J. Stat. Phys.*, **151**, 850 (2013).

[13] M. Colangeli, L. Rondoni, A. Vulpiani, Fluctuation-dissipation relation for chaotic non-Hamiltonian systems, *J. Stat. Mech.* (2012) L04002.

[14] M. Colangeli, C. Maes, B. Wynants, A meaningful expansion around detailed balance, *J. Phys. A: Math. and Theor.* **44**, 095001 (2011).

[15] D. J. Evans, D. J. Searles, S. R. Williams, On the Fluctuation Theorem for the Dissipation Function and its connection with Response Theory, *J. Chem. Phys.* **128**, 014504 (2008).

[16] C. Maes, K. Netočný, Rigorous meaning of McLennan ensembles, *J. Math. Phys.* **51**, 015219 (2010).

[17] D. Ruelle, Smooth Dynamics and New Theoretical Ideas in Nonequilibrium Statistical Mechanics, *J. Stat. Phys.* **95**, 393 (1999).

[18] D. Ruelle, Gaps and New Ideas in our Understanding of Nonequilibrium, *Physica A* **263**, 540 (1999).

[19] V. Lucarini, M. Colangeli, Beyond the linear fluctuation-dissipation theorem: the role of causality, *J. Stat. Mech.* (2012) P05013.

[20] V. Lucarini V, J. J. Saarinen, K.-E. Peiponen, E. M. Vartiainen, Kramers-Kronig Relations in Optical Materials Research (Heidelberg, Springer, 2005).

[21] G. S. Agarwal, Fluctuation-dissipation theorems for systems in non-thermal equilibrium and applications, *Z. Phys.* **252**, 25 (1972).

[22] M. Baiesi, C. Maes, An update on the nonequilibrium linear response, *New Journal of Physics* **15** 013004 (2013).

[23] W. Weidlich, Fluctuation-dissipation theorem for a class of stationary open systems *Z. Phys.* **248** 234 (1971).

[24] G. N. Bochkov, Yu. E. Kuzolev, Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics: I. Generalized fluctuation-dissipation theorem, *Phys. A* **106**, 443 (1981).

[25] G. N. Bochkov, Yu. E. Kuzolev, Fluctuation-dissipation relations: achievements and misunderstandings, *Phys. Usp.* **56**, 6 (2013).

[26] P. Hänggi, H. Thomas, Stochastic processes: time-evolution, symmetries and linear response, *Phys. Rep.* **88**, 207 (1982).

[27] D. J. Searles, L. Rondoni, D. J. Evans, The Steady state Fluctuation Relation for the Dissipation Function, *J. Stat. Phys.* **128**, 1337 (2007).

[28] J. A. McLennan Jr., Statistical mechanics of the steady state, *Phys. Rev.* **115**, 1405 (1959).

[29] D. N. Zubarev, V. P. Kalashnikov, Extremal properties of the nonequilibrium statistical operator, *Theor. Math. Phys.* **1** 108 (1969).

[30] M. Baiesi, C. Maes, B. Wynants, Nonequilibrium linear response for Markov dynamics, I: jump processes and overdamped diffusions, *J. Stat. Phys.* **137**, 1094 (2009).

[31] J. A. G. Roberts, G. R. W. Quispel, Chaos and time-reversal symmetry. Order and chaos in reversible dynamical systems, *Phys. Rep.* **216**, 63 (1992).

[32] M. Colangeli, P. De Gregorio, R. Klages, L. Rondoni, Steady state fluctuation relation with discontinuous "time reversibility" and invariant measures, *J. Stat. Mech.* (2011) P04021.

[33] M. Colangeli, L. Rondoni, Equilibrium, fluctuation relations and transport for irreversible deterministic dynamics, *Physica D* **241**, 6 (2012).

[34] G. Boffetta, G. Lacorata, S. Musacchio, A. Vulpiani, Relaxation of finite perturbations: beyond the fluctuation dissipation relation, *Chaos* **13**, 3 (2003).

[35] F. Bonetto, A. Kupiainen, J.L. Lebowitz, Absolute continuity of projected SRB measures of coupled Arnold cat map lattices, *Ergod. Th. Dynam. Sys.* **25**, 59 (2005).

[36] D. Ruelle, Nonequilibrium statistical mechanics near equilibrium: computing higher order terms, *Nonlinearity* **11**, 5 (1998).

[37] W. Cowieson, L.S. Young, SRB measures as zero-noise limits, *Ergodic Theory and Dynamical Systems* **25**, 1115 (2005).

[38] Yu. Kifer, *Random perturbations of dynamical systems* (Birkhäuser, Boston, 1988).

[39] C. Maes, K. Netočný, B. Wynants Steady state statistics of driven diffusions, *Physica A* **387**, 2675 (2008).

[40] H. Risken, *The Fokker-Planck Equation*, 2nd edn. (Springer Berlin, 1989).

[41] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications (Springer-Verlag, New York, 1998).

[42] C. Maes, On the origin and the use of fluctuation relations for the entropy, *Séminaire Poincaré* **2**, 29 (2003).

[43] S. Katz, J. L. Lebowitz, and H. Spohn, Phase transitions in stationary nonequilibrium states of model lattice systems, *Phys. Rev. B*, **28** 1655 (1983).

[44] J. Prost, J. F. Joanny, J. M. Parrondo, Generalized fluctuation-dissipation theorem for steady-state systems, *Phys. Rev. Lett.* **103**, 090601 (2009).

[45] R. Chetrite, K. Gawędzki, Eulerian and Lagrangian pictures of non-equilibrium diffusions, *J. Stat. Phys.* **137**, 890 (2009).

[46] M. Baiesi, E. Boksenbojm, C. Maes, B. Wynants, Nonequilibrium Linear Response for Markov Dynamics, II: Inertial Dynamics, *J. Stat. Phys.* **139**, 492 (2010).