

# An update on nonequilibrium linear response

M Baiesi<sup>1,2</sup> and C Maes<sup>3</sup>

<sup>1</sup> Dipartimento di Fisica ed Astronomia, Università di Padova, Via Marzolo 8,  
I-35131 Padova, Italy

<sup>2</sup> INFN, Sezione di Padova, Via Marzolo 8, I-35131 Padova, Italy

E-mail: [baiesi@pd.infn.it](mailto:baiesi@pd.infn.it)

<sup>3</sup> Instituut voor Theoretische Fysica, K.U. Leuven, Celestijnenlaad 200D, Leuven,  
Belgium

E-mail: [christian.maes@fys.kuleuven.be](mailto:christian.maes@fys.kuleuven.be)

**Abstract.** The unique fluctuation-dissipation theorem for equilibrium stands in contrast with the wide variety of nonequilibrium linear response formulæ. Their most traditional approach is “analytic”, which, in the absence of detailed balance, introduces the logarithm of the stationary probability density as observable. The theory of dynamical systems offers an alternative with a formula that continues to work when the stationary distribution is not smooth. We show that this method works equally well for stochastic dynamics, and we illustrate it with a numerical example for the perturbation of circadian cycles. A second “probabilistic” approach starts from dynamical ensembles and expands the probability weights on path space. This line suggests new physical questions, as we meet the *frenetic* contribution to linear response, and the relevance of the change in dynamical activity in the relaxation to a (new) nonequilibrium condition.

PACS numbers: 05.70.Ln, 05.40.-a, 05.20.-y

## 1. Introduction

An open system in contact with a large environment is in stationary equilibrium for a given reduced scale of description when at each moment its condition realizes minimal free energy. The total entropy production is then zero and the evolution is time-reversal invariant. Small disturbances break the stationary behavior and after some time an eventually new equilibrium is established mainly via dissipative effects. The comparison with the original equilibrium is the domain of linear response: since about 60 years now [1, 2], for classical, quantum, open, and closed systems basically the same formula relates the response to a small perturbation with an equilibrium correlation function. This formula (later called Kubo formula) exactly picks up the physical interpretation in terms of dissipation, hence the name fluctuation-dissipation theorem [3].

Response theory continues making sense for stimuli to nonequilibrium systems, where entropy is being produced already before the perturbation. Over the years various types of linear response formulæ have indeed been obtained for nonequilibria, including rather diverse but specific models such as in climatology and for glassy or coarsening dynamics. For better orientation the present paper identifies some common traits between the various approaches to put them in a more unified framework. Suppose for example that one wants a numerical code for predicting the response of an out-of-equilibrium systems, i.e., without actually perturbing the system. Without a clear picture of the available theoretical results, the choice of the method, if any, would be limited and possibly suboptimal. To answer requests like these, we aim at classifying the features of several “extended fluctuation-dissipation theorems” for nonequilibrium, highlight briefly their strengths and weaknesses, and the eventual relations between them.

Reviews and classifications of nonequilibrium response already exist. The paper by Marini Bettolo Marconi, Puglisi, Rondoni, and Vulpiani has reviewed response in statistical physics in the light of recent fluctuation relations [4]. Seifert and Speck have introduced a classification of some fluctuation-dissipation theorems into three classes [77]. Chetrite and Gupta present a more mathematical view [6]. The present work includes the case of deterministic dynamics and attempts a concise classification of linear response for nonequilibrium. We will conclude that there are three main classes. The first class of formulæ can be derived from the Kubo-Agarwal formula, which itself starts from a Dyson-expansion of the perturbed Markov semi-group. A second approach originates in the theory of dynamical systems and can be applied when the stationary distribution is not smooth; it gives rise to a numerical algorithm which in fact also applies for certain stochastic dynamics, as we show. The third class is much more probabilistic and treats noise as an important observable in the linear response. The synthesis is there provided by introducing the excess in dynamical activity in a second “frenetic” contribution to the traditional Kubo fluctuation-dissipation formula. In that sense, the name fluctuation–*dissipation* theorem (even “extended”) is not fully suitable for nonequilibrium systems, as e.g. their return to stationary nonequilibrium is not

uniquely characterized by dissipation [7]. On the mathematical side, our classification is to work either with generators (for Markov evolutions) or to work with weights on path space (mostly still limited to classical dynamical ensembles). Referring to those, we call it the analytic approach (Section 2) versus the probabilistic approach (Section 3). Remarks complementing those in previous sections, and conclusions, are found in Section 4.

## 2. Analytic approach

The framework we consider is that of Markov dynamics (no memory) on regions of  $\mathbb{R}^d$ . That includes deterministic dynamical systems, being essentially first order in time. It also includes jump processes as described with Master equations, but we use here the language of diffusion because it reduces naturally to that of deterministic dynamical systems in the limit of zero noise, realized here with a state-independent diffusion constant  $D \rightarrow 0$ . For even much greater simplicity we choose overdamped diffusions where the velocity field

$$\begin{aligned} \dot{x}_t = v(x_t) \quad \text{is given by} \\ v(x_t) = F(x_t) + \sqrt{2D} \xi_t \end{aligned} \tag{1}$$

for standard white noise  $\xi_t$  [8]. In Section 3.2 we also treat (underdamped or inertial) Langevin processes. In all these cases it makes sense to speak about the so called backward generator  $L$ , working on observables (Heisenberg picture). The expectation of an observable  $A$  at time  $t \geq 0$  is then given by

$$\langle A(x_t) \rangle = \int \rho(dx) (e^{tL} A)(x) \tag{2}$$

when at time zero the states are distributed with probability  $\rho$ , possibly singular with respect to the reference volume element  $dx$  on state space. We can also abbreviate that as  $d\langle A(t) \rangle/dt = \langle LA(t) \rangle$ . When no confusion can arise, we continue to write  $A(t) = A(x_t)$  for the (most often random) value of the observable at time  $t \geq 0$ , and  $\langle B(0)A(t) \rangle$  for the time-correlation between observable  $B$  at time zero and observable  $A$  at time  $t$ .

For the overdamped diffusion (1) the generator is

$$(LA)(x) = (F \cdot \nabla A)(x) + D \Delta A(x) \tag{3}$$

A perturbation changes the drift  $F \rightarrow F^h \equiv F + hF_1$  where, for simplicity, we avoid inserting a time-dependence in the small amplitude  $h$ , applied at all times  $t > 0$ . This leads to a perturbed backward generator  $L^h \equiv L + hL_1$  with  $L_1 A = F_1 \cdot \nabla A$ . The change in expectations at times  $t$  with respect to what we had for the unperturbed dynamics follows from (2):

$$\langle A(t) \rangle^h - \langle A(t) \rangle = \int \rho(dx) (e^{tL^h} - e^{tL})A(x) \tag{4}$$

To linear order in  $h$ ,

$$e^{tL^h} - e^{tL} = \int_0^t e^{sL} (L^h - L) e^{(t-s)L} ds + \mathcal{O}(h^2) \quad (5)$$

yielding a linear response

$$\begin{aligned} \left. \frac{d\langle A(t) \rangle^h}{dh} \right|_{h=0} &\equiv \lim_{h \rightarrow 0} \frac{\langle A(t) \rangle^h - \langle A(t) \rangle}{h} \\ &\equiv \chi(t) = \int_0^t ds R(t, s) \end{aligned} \quad (6)$$

including a susceptibility  $\chi(t)$  as the integration of a response function of the form

$$\boxed{R(t, s) = \int \rho(dx) (e^{sL} L_1 e^{(t-s)L} A)(x)} \quad (7)$$

More generally, when applying a time-dependent perturbation  $h_s$  at time  $s > 0$  we also have

$$\left. \frac{\delta \langle A(t) \rangle^h}{\delta h_s} \right|_{h=0} = R(t, s), \quad s < t \quad (8)$$

We mostly restrict ourselves to the case where  $\rho$  is stationary: the response depends only on the time difference  $\tau = t - s$  and we can write

$$\boxed{R(\tau) = \int \rho(dx) L_1 e^{\tau L} A(x)} \quad (9)$$

as the central object of study for the linear response of stationary Markov evolutions within an analytic approach<sup>‡</sup>.

The direct reading of the right-hand side of (9), further discussed under Section 2.2, is that  $e^{\tau L} A$  evolves the observable  $A$  for a time  $\tau$  and then  $L_1$  acts on the result to evaluate it in state  $x$ . However, we start in the next section with the more frequent approach of acting on  $\rho$ .

### 2.1. Acting on probabilities

In this section we focus on manipulations with the stationary probability distribution  $\rho$ . The basic step from (9) is partial integration, which means that it is assumed here that  $\rho$  has a smooth density with respect to the reference volume element,  $\rho(dx) = \rho(x) dx$ . In many cases that appears to be a reasonable physical assumption when the level of description is mesoscopic to macroscopic, independent of whether the system is driven or not (see [9]).

<sup>‡</sup> We can of course also use (9) for discrete jump processes. For each pair  $x, y$  of states, the perturbation enters as a modification of jump rates  $w(x, y) \rightarrow w(x, y) + h w_1(x, y)$ . The matrix  $L_1$  to be put in (9) has elements  $w_1(x, y)$  if  $x \neq y$ , and  $w_1(x, x) = -\sum_y w_1(x, y)$ .

2.1.1. *Kubo–Agarwal formula* Assuming a smooth density  $\rho$  we have that (9) can be rewritten as

$$R(\tau) = \int dx \rho(x) \frac{\mathcal{L}_1 \rho}{\rho}(x) e^{\tau L} A(x) \quad (10)$$

where the adjoint  $\mathcal{L}_1$  is defined by  $\int dx (\mathcal{L}_1 \rho)(x) A(x) = \int dx \rho(x) (L_1 A)(x)$ . Adjoints are forward generators of the time-evolution on densities, as appears e.g. in Master equations. For the diffusion (3) the adjoint of  $L$  is the operator of the Fokker-Planck equation ( $\partial_t \rho = \mathcal{L} \rho$ , see [8])

$$\mathcal{L} \rho = -\nabla \cdot (F \rho) + \Delta(D \rho) \quad (11)$$

and  $\mathcal{L}_1 \rho = -\nabla \cdot (F_1 \rho)$  so that (10) takes the form

$$R(\tau) = -\langle [\nabla \cdot F_1(0) + F_1(0) \cdot \nabla \log \rho(0)] A(\tau) \rangle \quad (12)$$

which is a specific realization of

$$\boxed{R(\tau) = \left\langle \frac{\mathcal{L}_1 \rho}{\rho}(0) A(\tau) \right\rangle} = \langle B(0) A(\tau) \rangle \quad (13)$$

with *observable*  $B(x) = \frac{\mathcal{L}_1 \rho}{\rho}(x)$ . Note that in general the stationary expectation  $\langle B \rangle = 0$  because  $\int dx \mathcal{L}_1 \rho(x) = 0$  from the normalization of  $\rho$ . Applications of that Agarwal formula [10] in practice meet the difficulty of needing to know the density  $\rho$  (which is usually not available) and the details of the dynamics for  $\mathcal{L}_1$ . It is thus a result (from partial integration) on a formal level.

Formula (13) is associated to equilibrium, see formula (2.5) in [11]. Of course we have only used that  $\rho$  is smooth. It is easy to verify that in the case of detailed balance with conservative forces  $F = -\nabla U$  at temperature  $D = 1/\beta$  we obtain the Kubo formula [2] for the linear response relation in equilibrium. Indeed, say for (3) with perturbation  $F_1 = -\nabla V$  around equilibrium  $\rho \propto \exp[-\beta U]$ , we get

$$\frac{\mathcal{L}_1 \rho}{\rho} = -\beta \nabla V \cdot \nabla U + \Delta V = \beta L V \quad (14)$$

so that

$$\left\langle \frac{\mathcal{L}_1 \rho}{\rho}(0) A(\tau) \right\rangle = \beta \langle L V(0) A(\tau) \rangle = \beta \frac{d}{d\tau} \langle V(0) A(\tau) \rangle \quad (15)$$

which is the (classical) Kubo formula. The last equality has used that under detailed balance  $\langle L V(0) A(\tau) \rangle = \langle V(0) L A(\tau) \rangle$  (see further details in Sec. 2.1.3). In a way, the Agarwal formula (13) repeats Kubo's original derivation while stopping short before specifying  $\rho$ .

Also others have re-found the Agarwal formula, such as in Theorem 2, formulæ (2.22) and (2.23), of Chapter 2 in [12]. Weidlich gives a quantum version: his equation (2.17) replaces  $\mathcal{L}_1 \rho \rightarrow \frac{i}{\hbar} [\rho, H_1]$  with the commutator of the perturbing Hamiltonian  $H_1$  [13]. Hänggi and Thomas in [14] find the Agarwal formula in their equation (3.12) for time-dependent processes. In the review [4] formula (2.70)–(2.73) is the Agarwal formula. We also find these formulæ such as (12) involving  $\log \rho$  presented in the book by Risken, formula (7.10)–(7.13) in [8], and as formula (7) in [15]. There also re-started the emphasis on  $\log \rho$  as “generalized” potential.

*2.1.2. Information potential* The formula that Falcioni, Isola and Vulpiani [16] derived for tiny displacements of the initial condition along a unit vector  $\hat{e}$ ,

$$\boxed{R(\tau) = \langle \hat{e} \cdot \nabla \log \rho(0) A(\tau) \rangle} \quad (16)$$

(see it also as formula (3.13) in the review [4]) corresponds to the case of diffusion with constant perturbation  $F_1 = -\hat{e}$  (impulsive constant force in the direction  $-\hat{e}$ ) in (12). Again, as there is no explicit dependence on the noise level the formula can be readily tried for chaotic dynamics with smooth stationary density [17]. The function  $\mathcal{I} \equiv -\log \rho$  is sometimes called the information potential. In fact, that potential gets a prominent place in various works on nonequilibrium linear response that follow the Hatano-Sasa formalism [18], such as in the more recent [19] or [20].

Other formulæ focusing on  $\mathcal{I}$  can be found in the works by Prost, Joanny and Parrondo [20] and by Speck and Seifert [77]. Consider the stationary density  $\rho^h$  for the perturbed dynamics with generator  $L^h$  and the linear response  $\rho^h(x) - \rho(x) = h \rho_1(x) + \mathcal{O}(h^2)$ . From stationarity  $\mathcal{L}^h \rho^h = 0$  we get  $\mathcal{L}_1 \rho + \mathcal{L} \rho_1 = 0$ . For the Agarwal formula (13) we need

$$B = \frac{\mathcal{L}_1 \rho}{\rho} = -\frac{\mathcal{L} \rho_1}{\rho} \quad (17)$$

On the other hand,

$$-\partial_h \mathcal{I}^h = \partial_h \log \rho^h = \lim_{h \rightarrow 0} \frac{\rho^h - \rho}{h \rho^h} = \frac{\rho_1}{\rho} \quad (18)$$

( $\partial_h$  is understood in  $h = 0$ ) so that for inserting in (13),  $B = \mathcal{L}(\rho \partial_h \mathcal{I}^h)/\rho$ . We conclude

$$\begin{aligned} R(\tau) &= \int dx \mathcal{L}(\rho \partial_h \mathcal{I}^h)(x) e^{\tau L} A(x) \\ &= \int dx \rho(x) \partial_h \mathcal{I}^h(x) L e^{\tau L} A(x) \\ &= \frac{d}{d\tau} \int dx \rho(x) \partial_h \mathcal{I}^h(x) e^{\tau L} A(x) \end{aligned} \quad (19)$$

or

$$\boxed{R(\tau) = \frac{d}{d\tau} \langle \partial_h \mathcal{I}^h(0) A(\tau) \rangle} \quad (20)$$

This formula with the special emphasis on the presence of the information potential appears in [77] (but with a different time-derivative), where  $\mathcal{I}$  is called stochastic entropy. The result of [20] starts from expanding the Hatano-Sasa identity, which effectively makes a special choice for the observable  $A$ . There one imagines the process and hence also the  $\mathcal{I}$  to depend on a family of parameters  $\lambda$  and one asks how the expected value of  $A = \partial_\lambda \mathcal{I}$  at  $\lambda = \lambda^*$  changes under a small change in these parameters around a given  $\lambda^*$ . The answer is provided by (20), which is equation (5) in [20].

As already noticed for (13) a disadvantage of incorporating the information potential into the correlation function of (16) and (20) is that it is generally unknown and not directly measurable. Not only do we not know the stationary density, but also no clear physically relevant interpretation in terms of heat, work or thermodynamic

potential has been found for  $\mathcal{I}$ . On the plus side, a clear advantage of (16) and (20) is that no detailed information on the dynamics is needed and that one can use parametrized forms of the stationary density  $\rho$  to put in  $\mathcal{I}$ . Typically, (quasi-)Gaussian approximations are tried for  $\rho$  to make formulæ (12), (16), (20) more explicit and practically useful; see e.g. [21, 12, 22, 23]. Note that there is no reference in formula (12) to the noise strength except through the stationary  $\rho$ . If one replaces the  $\rho$  in the  $B$  of (13) by a Gaussian, its variance will effectively reflect the noise level but needs to be fitted.

Finally, recent works on time-dependent processes with feedback rewrite the positivity of the entropy production in terms of the expected information potential (relative entropies) which appears useful for understanding work relations, see e.g. [24, 52]. At any rate, (20) is the climax of the analytic approach for smooth probability densities, preserving the Kubo form (15) most faithfully.

*2.1.3. State velocity* There is a relation between the information potential  $\mathcal{I}$  and the state-space velocity  $u$ . For diffusion processes with generator (3) this state velocity is

$$u \equiv \frac{j_\rho}{\rho} = F - D\nabla \log \rho = F + D\nabla \mathcal{I} \quad (21)$$

as can be readily checked from the expression for the stationary probability current  $j_\rho$  appearing in the Smoluchowski equation  $\partial_t \mu_t + \nabla j_{\mu_t} = 0$  expressing conservation of probability  $\mu_t$ . Of course this probability velocity needs not be related to physical currents, but an interesting observation writes a nonequilibrium response formula as a co-moving (equilibrium) fluctuation-dissipation theorem.

In a stationary process two quantities  $A$  and  $B$  have time-translationally invariant correlations  $\langle B(0)A(t) \rangle = \langle B(-t)A(0) \rangle$ . As we mentioned above, the Kubo formula (or, the fluctuation-dissipation theorem) holds under detailed balance, i.e., for an unperturbed evolution which gives rise not only to a stationary but also to a time-symmetric distribution of trajectories. Nonequilibrium means breaking time-reversal invariance. We introduce the operator  $L^*$  that generates the time-reversed motion to describe for example

$$\langle B(-t)A(0) \rangle = \int \rho(dx) A(x) (e^{tL^*} B)(x), \quad t \geq 0 \quad (22)$$

One can alternatively write  $\int \rho(dx) (Lf)(x)g(x) = \int \rho(dx) f(x)(L^*g)(x)$ , from which we see that  $L^*g = \mathcal{L}(\rho g)/\rho$ . For the overdamped diffusion processes that we have considered so far, from (3) and (11),  $L^* = L - 2u \cdot \nabla = -F\nabla + D\Delta + 2D\nabla \log \rho \nabla$ .

Note now that the Kubo formula (15) would follow from the Agarwal formula when in (13) we take  $B = -\beta L^*V$ . Indeed,

$$\langle L^*V(0)A(\tau) \rangle = \langle V(0)LA(\tau) \rangle = \frac{d}{d\tau} \langle V(0)A(\tau) \rangle \quad (23)$$

Detailed balance corresponds to  $L = L^*$  (i.e.,  $\langle B(-t)A(0) \rangle = \langle B(t)A(0) \rangle$ ) and corrections to the Kubo formula will thus arise from choosing in (13)

$$B(x) = -\frac{L + L^*}{2}V(x) = -L^*V(x) - \frac{L - L^*}{2}V(x) \quad (24)$$

for some function  $V$ . In other words the antisymmetric part  $\frac{L-L^*}{2}V$  in the formula (24) for  $B$  will be responsible for violating the Kubo formula and will show the nonequilibrium aspect. Moreover,  $L - L^* = 2u \cdot \nabla$  relates to the state velocity (21). Substituting these relations via (24) in (13), we find that for diffusions (3) (with  $D = 1/\beta$ )

$$R(t, s) = \beta \frac{d}{ds} \int dx \rho(x) V(x) e^{(t-s)L} A(x) - \beta \int dx \rho(x) u(x) \cdot \nabla V(x) e^{(t-s)L} A(x) \quad (25)$$

abbreviated as

$$\boxed{R(t, s) = \beta \frac{d}{ds} \langle V(s) A(t) \rangle - \beta \langle [u(s) \cdot \nabla V(s)] A(t) \rangle} \quad (26)$$

The equation (26) shows that the equilibrium Kubo form gets “restored” when describing the system in the Lagrangian frame moving with drift velocity  $u$ . The passage to the Lagrangian frame of local velocity  $d/ds \rightarrow d/ds - u \cdot \nabla$  “removes” the non-conservative forcing from the formula, as explained in [15, 26, 27]. Still, if we do not know the stationary  $\rho$ , the formula (26) implies a statistical average over the unknown vector  $u$ .

## 2.2. Acting on the observable

We go back to the original (9) and we move to an algorithm focusing on the evolution of the observable  $A$  rather than that of the density  $\rho$ . Various contributions to the theory of dynamical systems start exactly from that formula for the formulation of linear response as there still no assumption is needed on the smoothness of the probability  $\rho$ .

The main player now is  $L_1 e^{tL} A$  in (7), and  $L$  need not to commute with  $L_1$ . Typically the generator of the perturbation is  $L_1 = F_1 \cdot \nabla$  for states  $x \in \mathbb{R}^d$ . We are thus interested in obtaining a useful relation for  $\nabla e^{tL} A$ . In other words, we need to start from

$$\boxed{R(\tau) = \int \rho(dx) F_1(x) \cdot \nabla e^{\tau L} A(x)} \quad (27)$$

Ruelle treated a formula with this form for deterministic dynamical systems [28, 11], recently applied especially in studies of simple models for climate response [22, 29, 30]. Indeed (27) is especially suited in case the stationary probability law is *strange* in the sense that its lack of smoothness forbids further partial integration to go back to the Agarwal formula (13). In that same context however, it is useful to further split (27) into two parts [11], one describing fluctuations along the stable directions of motion and one parallel to unstable directions (those with positive Lyapunov exponents). Informally you would expect that the vector  $\nabla e^{\tau L} A(x)$  can be given, at least for large  $\tau$ , in its natural components  $\exp[\tau \lambda_{\hat{e}}(x)] \hat{e} \cdot \nabla A(x)$  for local Lyapunov exponents  $\lambda_{\hat{e}}$  corresponding to the various (stable and unstable) directions  $\hat{e}$ . Such a decomposition is natural for stationary distributions  $\rho$  which belong to the class of Sinai-Ruelle-Bowen (SRB) measures; these



have a density in the unstable (expanding) direction (positive Lyapunov exponents) so that there a further *partial* integration remains possible towards an Agarwal formula (13). That is exactly what is suggested in the hybrid form of Section 2.3 of [22]. See [31, 32] for an introduction on SRB-measures with their physical interpretation.

*2.2.1. Algorithm for deterministic and stochastic systems* The numerical evaluation of linear response for chaotic dynamics has recently been studied by various groups. In particular, Abramov and Majda have developed new computational approaches based on the theory of SRB-measures [29]. We simplify here the presentation to introduce an algorithm that works also for stochastic systems. To exploit (27), no knowledge of  $\rho$  is required, but we need to get a hold on  $\nabla_x e^{\tau L} A(x) = \nabla_x \langle A(x_\tau) \rangle_{x_0=x}$  where we have emphasized that the differentiation is on an evolved quantity with respect to the initial condition  $x$ . To solve the problem that  $\nabla$  (state derivative) and  $L$  (generating the time-evolution  $e^{tL} A(x)$ ) need not to commute, we unfold the formula a little further. A practical numerical tool for estimating  $\nabla_x A(x_\tau)$  in deterministic dynamics (where  $x_\tau$  is uniquely determined from  $x_0 = x$ ) was already presented in [33], but the following numerical method is more efficient for steady states.

Consider first an evolution in discrete time  $n = 0, 1, \dots$  (but with a parameter  $\epsilon$  that will allow a continuous time limit  $\epsilon \rightarrow 0$ ),

$$\begin{aligned} x_{n+1} &= g_n(x_n) = x_n + v_n(x_n), & x_0 &= x \\ \text{with} & & v_n(x) &= \epsilon[F(x) + \xi_n] \end{aligned} \quad (28)$$

for  $\xi_n$  anything stationary in time, including possible “noise” depending on the time  $n$ , but that does not depend on the state  $x$ . This ( $\xi_n$ ) is considered frozen so that for its given realization we take (28) as a deterministic dynamics. The main point is that  $\nabla_x \langle A(x_n) \rangle_x = \langle \nabla_x A(x_n) \rangle_x$  where the  $\langle \cdot \rangle$  averages over the “noise” ( $\xi_n$ ). We then need to deal with  $\nabla_x A(x_n)$  where  $x_n$  depends on the initial state  $x$  through (28). By the chain rule, applied recursively,

$$\begin{aligned} \nabla_x A(x_n) &= (\nabla A)(x_n) \cdot \nabla_x (g_{n-1} \circ \dots \circ g_0)(x) \\ &= (\nabla A)(x_n) \cdot \mathbb{G}_{n-1}(x) \\ &= (\nabla A)(x_n) (\nabla g_{n-1})(x_{n-1}) \cdot \nabla_x (g_{n-2} \circ \dots \circ g_0)(x) \\ &= (\nabla A)(x_n) (\nabla g_{n-1})(x_{n-1}) \cdot \mathbb{G}_{n-2}(x) = \dots \end{aligned} \quad (29)$$

where  $\nabla A$  is a  $1 \times d$  row vector (easily computed at all times) and  $\mathbb{G}_k(x) \equiv \nabla(g_k \circ g_{k-1} \circ \dots \circ g_0)(x)$  is a  $d \times d$  matrix obeying the recursive relations

$$\mathbb{G}_n(x) = \nabla g_n(x_n) \cdot \mathbb{G}_{n-1}(x) \quad (30a)$$

$$\mathbb{G}_0(x) = \nabla g_0(x) \quad (30b)$$

Note now that for each time  $k$ , the analytical form of the matrix  $\nabla g_k = \mathbb{I} + \epsilon \nabla F$  does not depend on  $k$  or on the “noise” ( $\xi_n$ ). Hence,

$$\begin{aligned} \nabla_x A(x_n) &= (\nabla A)(x_n) \cdot [\mathbb{I} + \epsilon(\nabla F)(x_{n-1})] \\ &\quad \cdot [\mathbb{I} + \epsilon(\nabla F)(x_{n-2})] \cdots [\mathbb{I} + \epsilon(\nabla F)(x)] \end{aligned} \quad (31)$$

The continuous time version  $\mathbb{G}_t(x)$  can be obtained in a suitable limit  $n \rightarrow \infty$  for  $\mathbb{G}_n(x)$  with time  $t = n\epsilon$  and time step  $\epsilon = t/n$ . In this limit (28) returns the evolution  $\dot{x}_t = F(x_t)$  for the deterministic case, while for stochastic equations an additional suitable rescaling to reproduce e.g. Gaussian noise with  $\xi_t$  is needed. At any rate, from  $\mathbb{I} + \epsilon(\nabla F)(x) \simeq e^{\epsilon(\nabla F)(x)}$  we get formally

$$\nabla_x A(x_\tau) = (\nabla A)(x_\tau) \cdot \mathbf{T} \exp \left[ \int_0^\tau ds (\nabla F)(x_s) \right] \quad (32)$$

where each  $x_s$  depends deterministically on  $x$  for frozen  $(\xi_r, 0 \leq r < s)$ , and  $\mathbf{T}$  indicates a time-ordered integral. The formula (32) or its discretization (31) is ready to be inserted into (27) where the  $\langle \cdot \rangle$  will first average there over the possible “noise” and then, with  $\rho$ , over the initial condition  $x$ . That provides the main algorithm for the analytic approach working on the observable instead of on the probability distribution.

A similar expression can be found in the Appendix B of Ref. [33] for the case of deterministic dynamics. In that work, however, the estimate of  $(\nabla A)(x_n)\mathbb{G}_t(x)$  was performed with an adjoint scheme, by integrating numerically backward in time the final value  $(\nabla A)(x_n)$  with the equation

$$\partial_t(\nabla A)(x_t) + (\nabla A)(x_t) \cdot (\nabla g)(x_t) = 0 \quad (33)$$

(In [33] they wrote the equation for the column vector rather than for the row vector  $\nabla A$ , and a transpose of  $(\nabla g)$  was used). Although that scheme is equivalent to our (31), it requires a CPU time  $\mathcal{O}(n^2)$  for a perturbation active during all  $n$  iterations, as opposed to the  $\mathcal{O}(n)$  matrix multiplications (2.2.1) for estimating matrices  $\mathbb{G}_k$  for  $k \leq n$ . Our scheme is exploiting stationarity: the propagator from time  $n - k$  to time  $n$  is the same as  $\mathbb{G}_k$  from time 0 to  $k$ . In transient regimes we would lose this property and we would also need  $\mathcal{O}(n^2)$  operations.

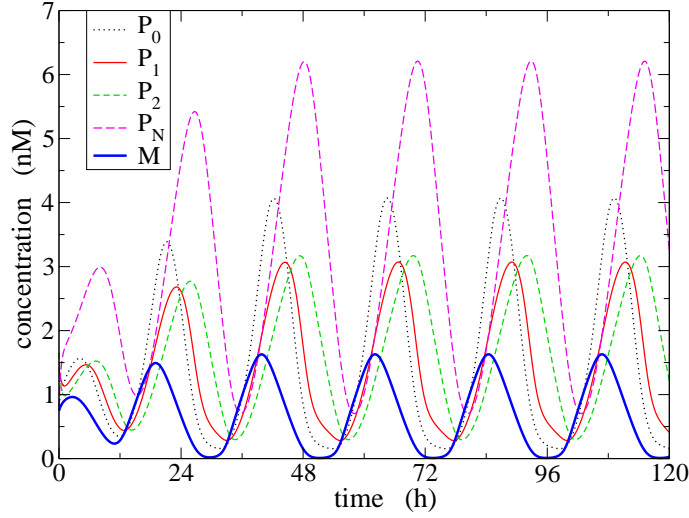
A further study of such a numerical algorithm is contained in [22, 29]; in particular Appendix A of [29] explains the derivation above (without the generalization to stochastic evolutions). For questions of dealing in a similar context with the impact of stochastic perturbations, we refer to [34].

*2.2.2. Numerical illustration* To illustrate the numerical scheme (29) for estimating Ruelle’s linear response formula (27) in a simple context, we consider a set of equations introduced in biology to describe circadian cycles, that is the periodicity of biorhythms, for *Drosophila* [35, 36]. The state space has  $d = 5$  dimensions, with states  $x = (P_0, P_1, P_2, P_N, M)$ . The dynamics couples the concentration  $M$  of mRNA with those of four types of proteins, written as  $P_0, P_1, P_2$ , and  $P_N$  in Ref. [36], where one can find the details of these equations. Denoting

$$\left[ \frac{P}{K} \right] \equiv \frac{P}{P + K} \quad (34)$$

we have equations of motion  $\dot{x} = F(x)$  of the form

$$\frac{dP_0}{dt} = k_S M - \nu_1 \left[ \frac{P_0}{K_1} \right] + \nu_2 \left[ \frac{P_1}{K_2} \right] \quad (35a)$$



**Figure 1.** Evolution of the concentration of mRNA and of the four proteins of the model from random initial conditions, according to (2.2.2).

$$\frac{dP_1}{dt} = \nu_1 \left[ \frac{P_0}{K_1} \right] - \nu_2 \left[ \frac{P_1}{K_2} \right] - \nu_3 \left[ \frac{P_1}{K_3} \right] + \nu_4 \left[ \frac{P_2}{K_4} \right] \quad (35b)$$

$$\frac{dP_2}{dt} = \nu_3 \left[ \frac{P_1}{K_3} \right] - \nu_4 \left[ \frac{P_2}{K_4} \right] + \nu_d \left[ \frac{P_2}{K_d} \right] - k_1 P_2 + k_2 P_N \quad (35c)$$

$$\frac{dP_N}{dt} = k_1 P_2 - k_2 P_N \quad (35d)$$

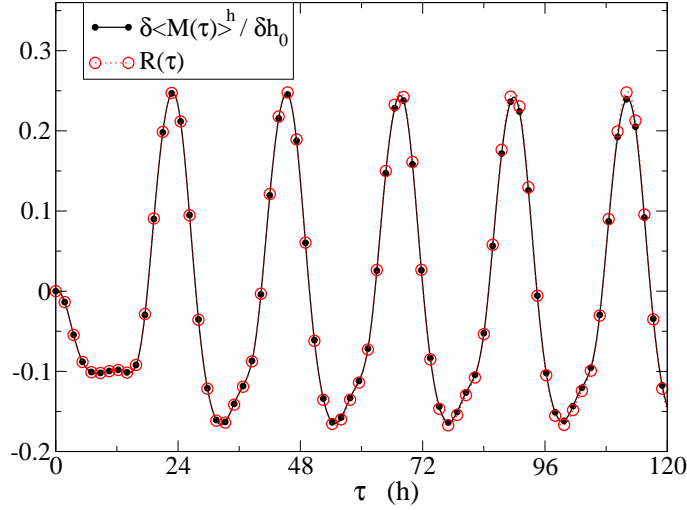
$$\frac{dM}{dt} = \nu_S \left[ \frac{K_I^n}{P_N^n} \right] - \nu_m \left[ \frac{M}{K_m} \right] \quad (35e)$$

with parameters as in previous papers,

$$\begin{aligned} (\nu_S = 0.5, \nu_m = 0.3, \nu_1 = 6, \nu_2 = 3, \nu_3 = 6, \nu_4 = 3) \text{ nM h}^{-1} \\ (K_m = 0.2, K_I = 2, K_1 = 1.5, K_2 = 2, K_3 = 1.5, K_4 = 2) \text{ nM} \\ (k_S = 2, k_1 = 2, k_2 = 1) \text{ h}^{-1}, \quad n = 4 \end{aligned}$$

Integration of (2.2.2) was performed with a simple Verlet scheme with a discrete time step  $\epsilon = 0.0025$ . Starting from random concentrations, the model reaches quickly a cyclic regime, as shown in Fig. 1.

We check the response in the mRNA concentration  $M$  to a change of the rate  $k_S \rightarrow k_S(1+h)$ , starting from the steady state, i.e., a random phase of the cycle. The dynamical perturbation introduced by that change is  $h F_1 = (h M k_S, 0, 0, 0, 0)$ , since a change in  $k_S$  affects only the evolution equation of variable  $P_0$ , see (35a). The observable  $A(x) = M$  has a gradient  $(\nabla A)(x) = (0, 0, 0, 0, 1)$ , which is coupled to the perturbation via the “propagator” matrix  $\mathbb{G}_n$ . With the setup of Sec. 2.2.1, we can estimate  $\mathbb{G}_n$  by a sequence of matrix multiplications. Given an evolution in discrete time  $g(x) = x + F(x)\epsilon$ , we are interested in matrix  $\nabla g(x) = \mathbb{I} + \nabla F(x)\epsilon$ , or in coordinates  $(\nabla g)_{ij}(x) = \delta_{ij} + dF_i/dx_j \epsilon$ . The rows of the matrix  $(\nabla F)$  are



**Figure 2.** Linear response function  $R(\tau)$  of the mRNA density  $M$  to an impulsive change in the parameter  $k_S \rightarrow k_S(1 + h_0)$  as a function of time and response  $\delta\langle M(\tau)\rangle/\delta h_0$  computed with  $h_0 = 10^{-3}$ . Units are nM vs. hours.

$$\left( -\nu_1 \left[ \frac{P_0}{K_1} \right]_{P_0}, \nu_2 \left[ \frac{P_1}{K_2} \right]_{P_1}, 0, 0, k_S \right) \quad (36a)$$

$$\left( \nu_1 \left[ \frac{P_0}{K_1} \right]_{P_0}, -\nu_2 \left[ \frac{P_1}{K_2} \right]_{P_1} - \nu_3 \left[ \frac{P_1}{K_3} \right]_{P_1}, \nu_4 \left[ \frac{P_2}{K_4} \right]_{P_2}, 0, 0 \right) \quad (36b)$$

$$\left( 0, \nu_3 \left[ \frac{P_1}{K_3} \right]_{P_1}, -\nu_4 \left[ \frac{P_2}{K_4} \right]_{P_2} - \nu_d \left[ \frac{P_2}{K_d} \right]_{P_2}, -k_1, k_2, 0 \right) \quad (36c)$$

$$(0, 0, k_1, -k_2, 0) \quad (36d)$$

$$\left( 0, 0, 0, \nu_S \left[ \frac{K_I^n}{P_N^n} \right]_{P_N}, -\nu_m \left[ \frac{M}{K_m} \right]_M \right) \quad (36e)$$

where derivatives with respect to  $P$  are denoted as  $\left[ \frac{P}{K} \right]_P = \frac{1}{P+K} - \frac{P}{(P+K)^2}$  and similarly for  $\left[ \frac{K_I^n}{P_N^n} \right]_{P_N}$ . This matrix with  $x_0 = (P_0, P_1, P_2, P_N, M)_0$  yields  $(\nabla g)(x_0) = \mathbb{I} + (\nabla F)(x_0)\epsilon$ , coinciding with  $\mathbb{G}_0$ . Iteratively,  $\mathbb{G}_1 = (\nabla g)(x_1) \cdot \mathbb{G}_0$ , and so on.

By sampling many trajectories, in parallel for perturbed and unperturbed dynamics starting from an initial condition from the steady state, the estimate of  $\delta M(\tau)/\delta h_0$  has been obtained by setting a small constant  $h$  for  $\tau > 0$  and by evaluating the time derivative of  $(\langle M(\tau) \rangle^h - \langle M(\tau) \rangle)/h$  (we used  $h = 10^{-3}$ ). The response function instead has been calculated with (29) and (27), and well overlaps with the response, as shown in Fig. 2. Asymptotic stability in this model [36] probably favors a good performance of formula (27). We computed both the perturbed dynamics and the unperturbed fluctuation formula just to show graphically that the results are the same, but of course there is no additional understanding of the process. Obviously, computing the response

only with unperturbed simulations would be convenient as long as the convergence of the method is good. We postpone a more detailed study of the efficiency of this and other numerical schemes to a future work.

### 3. Probabilistic approach

Linear response makes sense in general only within a statistical theory. That is to say, sensitive dependence on initial and boundary conditions can create strong and relevant effects beyond linear order on the microscopic scale while macroscopic linearity remains valid. For estimating the mobility we do not investigate the microscopic particle's individual motion at specific times and how it changes under an external field, but we ask for a spatio-temporal averaged current and that includes noise. In fact, for stronger microscopic chaoticity we expect the statistical approach to be more relevant. While various physical variables can show chaotic behavior in their time-evolution, their spatial or temporal averages will typically have a much smoother behavior, see also [9] and chapter 6.2 in [37] answering the so called van Kampen objection.

#### 3.1. As noise gets important

In coarsening dynamics of low temperature spin systems, or in spin glasses, a very long transient regime may exist towards equilibrium, i.e., the nonequilibrium is not imposed by external gradients and the dynamic equations are actually equilibrium ones. It is in this context that an extensive response-literature has been produced during the last decades. We mention briefly some results from 2003, whose focus was mostly to develop zero-field algorithms, in other words, exploiting fluctuation–response relations to estimate numerically responses from fluctuations in unperturbed dynamics. Works of Chatelain [38], Ricci-Tersenghi [39], and Crisanti and Ritort [40] introduced new schemes for computing the response of the system from correlations in the spins. They were followed by Diezeman [41] and by Lippiello, Corberi, and Zannetti [42, 43, 44]. The derivations of these results cannot be reduced in our brief scheme and various forms of response relations have been obtained. However, the one by Lippiello et al., a specific case of (41) below, emerged as having clear physical significance beyond numerical usefulness, because it matched the form of an earlier study by Cugliandolo, Kurchan and Parisi [45], who proposed a response relation for autocorrelations in Langevin equations. With the present notation (1), perturbation  $h_s x$  and observable  $A(x) = x$ , that would read

$$R(t, s) = \frac{1}{2D} \left\{ \left( \frac{d}{ds} - \frac{d}{dt} \right) \langle x_s x_t \rangle - (\langle F(x_t) x_s \rangle - \langle F(x_s) x_t \rangle) \right\} \quad (37)$$

showing an “asymmetry” term in addition to the form of the Kubo formula. The derivation of (37) is based on the equality  $\langle x_t \xi_t \rangle = 2D\beta R(t, s)$ , hence it is centered on the presence of noise  $\xi_t$ . This was new, certainly with respect to derivations presented in previous sections where eventually noise was just an aspect of the dynamics, not

fundamental for the derivations. An equation similar to (37) appears in a more recent study for nonequilibrium steady states [46], based on a path-space formulation by Harada and Sasa [47], see their Appendix B. The Harada-Sasa approach has really pioneered the path-space approach of the next subsection. The consistency between the approaches of Refs. [45, 42, 47] indeed shows that these belong to a general framework with new significant physical content. In the following section we discuss the path-space formulation embracing these results and we mention the physical interpretation that goes beyond merely dissipative aspects. Noise becomes visible then as dynamical activity, ruling the time-symmetric fluctuations.

### 3.2. Path space approach

The origin of dynamical ensembles is the projection of a microscopic Hamiltonian dynamics on the dynamics of reduced variables. That Mori-Zwanzig projection [48, 49] originates from making a physical partition of the phase space, depending on the physical situation at hand, in which each microstate  $X$  is mapped (many-to-one) to a reduced state  $x(X)$ . That induces noise in the reduced dynamics, for example on the mesoscopic level of description. It is then natural to consider dynamical ensembles, i.e., probability distributions on path space where a path refers to a trajectory on the mesolevel. Such was already the approach of Onsager and Machlup starting dynamical fluctuation theory in [50]. These trajectories, under certain limiting conditions (e.g. via a weak coupling limit or via adiabatic elimination), can be described via first-order equations, in which case we meet the Markov processes of the previous Section 2. Dynamical ensembles can however also describe non-Markovian processes describing important memory effects, see e.g. [51] for the application of a response relation in a visco-elastic medium.

Paths are trajectories  $\omega = (x_s, 0 \leq s \leq t)$  in state space, say looking at the states  $x_s$  in the time-interval  $[0, t]$ . For the sake of simplicity we characterize the unperturbed (perturbed) process by the probability weight  $P(\omega)$  [ $P^h(\omega)$ ] with respect to some reference  $d\omega$ . The mathematical idea is to turn perturbed expectations  $\langle A(t) \rangle^h$  into unperturbed ones  $\langle A(t) P^h(\omega) / P(\omega) \rangle$ . Defining the relative action  $\mathcal{U}(\omega) = \log P(\omega) / P^h(\omega)$  and splitting  $\mathcal{U}(\omega) = [\mathcal{T}(\omega) - \mathcal{S}(\omega)] / 2$  into a time-symmetric  $\mathcal{T}(\omega)$  and a time-antisymmetric  $\mathcal{S}(\omega)$ , we get

$$\langle A(t) \rangle^h - \langle A(t) \rangle = \int d\omega P(\omega) A(x_t) (e^{-\mathcal{U}(\omega)} - 1) \quad (38a)$$

$$= \langle A(t) (e^{-\mathcal{U}(\omega)} - 1) \rangle \quad (38b)$$

$$= \frac{1}{2} \langle A(t) \mathcal{S}_1(\omega) \rangle - \frac{1}{2} \langle A(t) \mathcal{T}_1(\omega) \rangle + \mathcal{O}(h^2) \quad (38c)$$

where  $\mathcal{T}_1$  and  $\mathcal{S}_1$  are the linear contributions in  $h$  of  $\mathcal{T}$  and  $\mathcal{S}$ , respectively, around  $h = 0$  (where  $\mathcal{U} = 0$ ). This formula is general but the point is that in physical systems with local detailed balance [47] we have a good physical understanding of the path functions  $\mathcal{T}$  and  $\mathcal{S}$  [53, 54, 55, 56]. Let us first look under global detailed balance. We already see here that since the observable  $A(t) - A(0)$  is time-antisymmetric (ignoring momenta),

under detailed balance

$$\begin{aligned}\langle A(t) \rangle^h - \langle A(t) \rangle &= \langle A(t) - A(0) \rangle^h \\ &= \frac{1}{2} \langle [A(t) - A(0)] \mathcal{S}_1(\omega) \rangle \\ &= \langle A(t) \mathcal{S}_1(\omega) \rangle\end{aligned}\tag{39}$$

because  $\langle [A(t) - A(0)] \mathcal{T}_1(\omega) \rangle = 0$  and  $\langle A(t) \mathcal{S}_1(\omega) \rangle = -\langle A(0) \mathcal{S}_1(\omega) \rangle$  by time-reversal symmetry of the reference equilibrium process. The result (39) should equal the Kubo formula (15), and indeed it does as will become clear. When the perturbation is from a potential  $-h_s V(x_s)$ , it is well established that the time-antisymmetric  $\mathcal{S}(\omega) = \mathcal{S}_1(\omega)$  must be the path-dependent entropy flux into the environment as caused by the perturbation potential  $V$  [57, 58]. If the environment is at uniform temperature, for small constant  $h$  for times  $t > 0$  the change in entropy is  $\mathcal{S}(\omega) = \beta h [V(t) - V(0)]$ , namely dissipated energy  $h[V(t) - V(0)]$  divided by temperature. The first term on the right-hand side of (38c) has thus the same form one finds in equilibrium (39), but in general the unperturbed process is in steady nonequilibrium and  $\mathcal{S}(\omega) = \beta h [V(t) - V(0)]$  is an excess of entropy production with respect to that already generated by the nonequilibrium dynamics. Hence, the derivative  $\delta \langle A(t) \mathcal{S}(\omega) \rangle / \delta h_s = \beta \frac{d}{ds} \langle V(s) A(t) \rangle$  also mimics the equilibrium Kubo formula.

We now turn to the second term on the right-hand side of (38c). Eq. (38c) must be true also for a constant  $A$ , for which the response is zero. We thus get

$$\beta \frac{d}{ds} \langle V(s) \rangle = \frac{\delta}{\delta h_s} \langle \mathcal{T}_1(\omega) \rangle = \left\langle \frac{\delta}{\delta h_s} \mathcal{T}_1(\omega) \right\rangle\tag{40}$$

If  $\frac{\delta}{\delta h_s} \mathcal{T}_1(\omega)$  equals a state function  $B(s)$ , then, since  $\frac{d}{ds} \langle V(s) \rangle = \langle LV(s) \rangle$ , one deduces that  $B(s) = \beta LV(s)$ , arriving at

$$\boxed{R(t, s) = \frac{\beta}{2} \frac{d}{ds} \langle V(s) A(t) \rangle - \frac{\beta}{2} \langle LV(s) A(t) \rangle}\tag{41}$$

For diffusion processes (1) with generator (3) and temperature  $D = 1/\beta$  we see that indeed  $LV(x) = (F \cdot \nabla V)(x) + D \Delta V(x)$  is a state function. The same is true for Markov jump processes (for specific derivations of (41) using stochastic calculus, see [55]). The function  $LV$  quantifies the time-symmetric volatility of  $V$  under the unperturbed dynamics, also named “frenesy” [53]. It is related to the “dynamical activity” in discrete systems, where it quantifies the change in escape rates;  $\mathcal{T}(\omega)$  is the excess in dynamical activity over  $[0, t]$ . We refer to [54, 55, 56] for more details and examples. Thus, there is a new contribution next to the first term in (41) (taking one half of the Kubo formula (15) and referring as usual to dissipation). In particular, its form now clearly deviates from response formulæ such as (20). One then wonders in what sense the decomposition in formulæ like (41) is intrinsically natural. We think it is, as that splitting makes the first term time-antisymmetric and the second term symmetric in time  $s$ , which corresponds to the dissipative and the reactive part of the susceptibility, respectively.

It remains to be discovered what is the operational meaning of the notion of dynamical activity, and whether the frenetic contribution is on an equal footing with the

entropic part in the response. What is certain however is that the dynamical activity and the entropy flux merge for systems close to equilibrium, as it should be indeed for recovering the linear response around equilibrium. So the frenetic part of the response only starts to play its independent role from second order around equilibrium. We then expect that thermodynamic forces in that regime pick up non-gradient contributions that are directly related to the frenetic response.

Note that in (41) one computes averages without needing an explicit knowledge on the stationary probability. On the other hand some knowledge is required on the dynamics sitting in the generator  $L$ . It is still not clear how much kinetic information is truly needed and how practical that gets. We summarize the general physical idea in the formula

$$\langle A(t) \rangle^h - \langle A(t) \rangle = \frac{1}{2} \langle \text{Entr}^{[0,t]}(\omega) A(t) \rangle - \frac{1}{2} \langle \text{Esc}^{[0,t]}(\omega) A(t) \rangle \quad (42)$$

where  $\text{Entr}^{[0,t]}(\omega)$  is the excess in entropy flux over the time period  $[0, t]$  and  $\text{Esc}^{[0,t]}(\omega)$  is the excess in dynamical activity due to the perturbation. The physical challenge is to learn to guess or to find that  $\text{Esc}^{[0,\tau]}(\omega)$  from partial information on the dynamics. For the overdamped dynamics of e.g. (41) it means to consider the expected rate of change of the perturbing potential  $V$  under the original dynamics. The response relation out of equilibrium is no longer a fluctuation-dissipation relation but a fluctuation-dissipation-activation relation. In fact, the formula can now be turned around and from measuring “violations” of the fluctuation-dissipation relation one obtains informations about the active forces [51]. Harada and Sasa [47, 46] already expressed the mean rate of dissipation as the deviation from the equilibrium fluctuation-dissipation relation: such approach was also used to study active forces for a single F<sub>1</sub>-ATPase molecule [52].

### 3.3. Inertial case

We open a separate subsection on the inertial case of Langevin dynamics because its linear response is much less discussed in the literature despite the obvious interest for example for models of heat conduction. The main point however is that the ideas summarized under (3.2) and (42) remain unchanged.

For states  $(q, p) = (q^1, q^2, \dots, q^n; p^1, p^2, \dots, p^n) \in \mathbb{R}^{2n}$  of positions and momenta we attach standard white noise  $\xi_t^i$  to each  $1 \leq i \leq n$ , with constant strength  $D^i$  and a friction coefficient  $\gamma^i$  to model heat baths at temperature  $D^i/\gamma^i = T^i$ :

$$\begin{aligned} \dot{q}^i &= p^i \\ \dot{p}^i &= F^i(q) - \gamma^i p^i + h_t \frac{\partial V}{\partial q^i} + \sqrt{2D^i} \xi_t^i \end{aligned} \quad (43)$$

The forces  $F^i$  can contain a nonconservative part but are confining when we want a stationary regime where the particles typically reside in a bounded region. We already inserted the perturbation  $V(q)$  with small time-dependent amplitude  $h_s$  for  $s \geq 0$ . The



linear response is given by formula (17) in [53]:

$$\begin{aligned}
R(t, s) = & \sum_i \frac{1}{2T^i} \left\langle \frac{\partial V}{\partial q^i}(q_s) p_s^i A(t) \right\rangle \\
& - \sum_i \frac{1}{2D^i} \left\{ \left\langle \frac{\partial V}{\partial q^i}(q_s) F^i(q_s) A(t) \right\rangle \right. \\
& \quad \left. - \frac{d}{ds} \left\langle \frac{\partial V}{\partial q^i}(q_s) p_s^i A(t) \right\rangle \right. \\
& \quad \left. + \sum_j \left\langle \frac{\partial^2 V}{\partial q^j \partial q^i}(q_s) p_s^j p_s^i A(t) \right\rangle \right\} \quad (44)
\end{aligned}$$

The first sum again corresponds to the dissipative part from the entropy fluxes in the reservoirs at temperatures  $T^i$ . The remaining sums give the frenetic contribution. As not recognized yet, formula (44) can still be rewritten in a similar way as done in (41). Supposing  $D^i = \gamma/\beta$  we must replace there  $V \rightarrow \dot{V}/\gamma = p \cdot \nabla_q V/\gamma$ ,

$$\boxed{R(t, s) = \frac{\beta}{2\gamma} \frac{d}{ds} \langle p_s \cdot \nabla_q V(s) A(t) \rangle - \frac{\beta}{2\gamma} \langle L(p_s \cdot \nabla_q V)(s) A(t) \rangle} \quad (45)$$

in which the generator  $L$  for (43) now reads

$$L f(q, p) = p \cdot \nabla_q f + (F - \gamma p) \cdot \nabla_p f + D \Delta_p f \quad (46)$$

and is the underdamped version of (3). Note however that the dissipative last term  $D \Delta_p f$  does not contribute in (45) and that these formulæ do not work in case  $D^i = 0$  for an  $i$  with  $\partial V/\partial q^i \neq 0$ . In that case the best alternative is probably to apply the algorithm of Section 2.2.

An important application of (44) is to the modification of the Sutherland-Einstein formula relating transport coefficients such as mobility with fluctuation quantities such as the diffusion constant [59].

## 4. Further remarks and conclusions

### 4.1. Nonlinear responses

There is also a growing number of works on higher-order terms around equilibrium. In fact part of the book by Evans and Morriss is devoted to that [60]. Other references include [61], Section 10 in [58] or the more recent [4, 62, 63, 64]. One typical start is the fluctuation symmetry in the distribution of the entropy flux, transient from the reference equilibrium system as also explained in [65] for thermostated dynamics.

As was emphasized in [66] the main point is probably not to be able to write formal expansions and formulæ, but to find useful structures and unifying interpretations. We will not deal with that here, except for mentioning one particular general relation between second order and first order terms that has been largely unnoticed [67], and an instance of which has appeared as identity (25) in [66].

$$\left. \frac{\partial^2}{\partial h^2} \langle A(t) \rangle^h \right|_{h=0} = \beta \left. \frac{\partial}{\partial h} \langle [A(t) - \bar{A}(0)][V(t) - V(0)] \rangle^h \right|_{h=0} \quad (47)$$

The  $\bar{A}$  equals  $A$  up to flipping the sign of the momenta. All terms are explicitly expressed as correlation functions in the equilibrium reference process (indicated with the superscript  $\langle \cdot \rangle^0$ ). The derivation of (47) uses linear response around nonequilibrium for non-state functions (i.e., the correlation function in the right-hand side).

#### 4.2. Non-state functions

In case of observables that depend explicitly on time or are functions of the trajectory over several times (such as products  $A_1(t_1)A_2(t_2)\dots A_n(t_n)$ ), some of the formulæ must change. The basic techniques remain however in place. A typical application is how the heat depends on a change in parameters, e.g., for estimating nonequilibrium heat capacities [68]. Heat is not a state function but varies with the trajectory and the applied protocol.

#### 4.3. Note on effective temperature

A traditional approach to violation of the equilibrium fluctuation-dissipation theorem is to imagine an effective temperature in the otherwise unchanged Kubo formula [69]. That idea has had most success with mean field type systems, but it has remained more unclear how the effective temperature can provide a consistent and general tool for realistic systems. Still today it serves as a paradigm for interpreting experimental results, see e.g. [70]. One possible approach for the future would be to associate an effective kinetic temperature to the ratio between the frenetic and the entropic contribution in (42), see Appendix A in [55].

#### 4.4. Transient case

Lots of attention have been devoted to the linear response behavior for relaxational processes. We mentioned some of that in Section 3.1, but there is no way to be complete. For example, the interest in ageing and glassy dynamics has much stimulated the search of modified fluctuation-dissipation relations [71]. Here we emphasize that the analytic approach and in particular the methods of Section 2.1 lose their simple structure when the unperturbed reference is time-dependent (and not stationary as was assumed). On the other hand, formulæ like (41) are unchanged when one is not starting at time zero from the stationary distribution, but one needs to take the average in the unperturbed transient regime. Also causality is automatically verified there as the second term in (41) equals the first term when  $s > t$ . This unification in expressions for the transient and the steady regime is only natural as even the stationary regime is physically and ultimately but a very long transient.

#### 4.5. General properties of linear response

Some formal properties of linear response go basically unchanged from equilibrium to nonequilibrium contexts. For example, sum rules and Kramers-Kronig relations only

depend on the presence of an underlying Hamiltonian dynamics for the total system plus environment or on causality. Of course, specific expressions can differ but there is physically nothing new, see also [72, 73]. Nevertheless, there do exist essential differences. We have already alluded to the fact that the name “fluctuation–dissipation theorem” is no longer so appropriate because of the importance and complimentary character of changes in the dynamical activity (and not only in the dissipation). In fact, the word “fluctuation” also becomes less correct as the time-correlations in the expressions for  $R(t, s)$  no longer express a symmetrized time-correlation. To make that point, let us evaluate the equilibrium Kubo formula (15)) when the perturbing potential  $V$  equals the observable  $A$ :

$$\langle V(t) \rangle^h - \langle V \rangle = -\langle V(0)[V(t) - V(0)] \rangle = \frac{1}{2} \langle [V(t) - V(0)]^2 \rangle \geq 0 \quad (48)$$

which is the variance (or the “diffusion”) of the displacement  $V(x_t) - V(x_0)$ . That is sometimes called a generalized Einstein relation. In fact all linear transport coefficients for equilibrium can be expressed like that, as was understood already in 1960 by Helfand [74]. For nonequilibrium linear response, that relation including its positivity gets violated, see also [59, 7]. In the words of the previous subsection, negative effective temperatures become possible in nonequilibrium. The origin lies in the frenetic contribution, e.g. the second term on the right-hand side of (41), which can overrun the first dissipative term. Examples of negative effective temperature in active matter are already known [75, 76, 77].

#### 4.6. Outlook

Response theory is primarily about predicting the reaction of a system in terms of its unperturbed behavior. As we mostly have in mind response in time, that involves temporal correlations. Therefore, *dynamical* fluctuation theory can be expected to be most prominent. (See [78] for the distinction between static and dynamic fluctuation theory: statics looks at deviations around the law of large numbers at single times, for example for the average over many copies of the system, while dynamical fluctuations are around the law of large times, for deviations around time-averages.) Around equilibrium, dynamical fluctuations are governed, just as static fluctuations, by the entropy and dissipation functions. That further enables connecting different response coefficients, such as in Onsager reciprocity or via Maxwell relations. Nonequilibrium makes a more drastic difference between static and dynamical fluctuations. In particular, no useful connections between different types of responses have been discovered for nonequilibrium processes. We think of the analogue of relations between compressibilities, heat capacities and conductivities. In our opinion, that challenge will require finding experimental access to quantities like the dynamical activity which is complementing entropic characterizations in the description of nonequilibrium processes.

#### 4.7. Conclusions

We have presented a concise guide (with plenty of references) to the multi-faceted world of linear response for systems out of equilibrium. From there we have discovered similarities but also some missing pieces that otherwise would not be evident. One such piece was the extension to stochastic systems of Ruelle's formulation, which we have introduced, together with an efficient algorithm. That also allows for numerical calculations where an explicit knowledge of the density of states is not required, as opposed to other (analytic) formulations that we have described. However, the resulting linear response formula contains a correlation function whose physical meaning is not very clear.

In contrast, a rich physical picture emerges from a probabilistic approach based on path space weights, where the stationary distribution is also not needed. Besides the quite different mathematical apparatus compared to the more standard analytical approach, the probabilistic way indeed emerges as the one that currently offers more relations with dynamical fluctuation theory: one has to study also how the system correlates with the activity of the perturbing potential, a time-symmetric quantity complementary to the time-antisymmetric fluxes of entropy. The combination fluctuation-dissipation therefore does not suffice to characterize the linear response of nonequilibrium systems.

*Acknowledgments* M.B. thanks M. Colangeli, A. Vulpiani and J. Wouters for useful discussions.

## References

- [1] Callen H B and Welton T A 1951 Irreversibility and generalized noise *Phys. Rev.* **83** 34
- [2] Kubo R 1966 The fluctuation-dissipation theorem *Rep. Prog. Phys.* **29** 255–284
- [3] Kubo R, Toda M and Hashitsume N 1992 *Statistical Physics: Nonequilibrium statistical mechanics* 2nd ed vol 2 (Springer)
- [4] Marini Bettolo Marconi U, Puglisi A, Rondoni L and Vulpiani A 2008 Fluctuation-dissipation: response theory in statistical physics *Phys. Rep.* **461** 111–195
- [5] Seifert U and Speck T 2010 Fluctuation-dissipation theorem in nonequilibrium steady states *Europhys. Lett.* **89** 10007
- [6] Chetrite R and Gupta S 2011 Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale *J. Stat. Phys.* **143** 543
- [7] Maes C, Netočný K and Wynants B 2011 Monotonic return to steady nonequilibrium *Phys. Rev. Lett.* **107** 010601
- [8] Risken H 1989 *The Fokker-Planck Equation* 2nd ed (Berlin: Springer-Verlag)
- [9] Colangeli M, Rondoni L and Vulpiani A 2012 Fluctuation-dissipation relation for chaotic non-Hamiltonian systems *J. Stat. Mech.* L04002
- [10] Agarwal G S 1972 Fluctuation-dissipation theorems for systems in non-thermal equilibrium and applications *Z. Physik* **252** 25–38
- [11] Ruelle D 2009 A review of linear response theory for general differentiable dynamical systems *Nonlin.* **22** 855–870

- [12] Majda A J, Abramov R V and Grote M J 2005 Information theory and stochastics for multiscale nonlinear systems *CRM Monograph Series* (AMS)
- [13] Weidlich W 1971 Fluctuation-dissipation theorem for a class of stationary open systems *Z. Physik* **248** 234–243
- [14] Hänggi P and Thomas H 1982 Stochastic processes: time-evolution, symmetries and linear response *Phys. Rep.* **88** 207–319
- [15] Speck T and Seifert U 2006 Restoring a fluctuation-dissipation theorem in a nonequilibrium steady state *Europhys. Lett.* **74** 391–396
- [16] Falcioni M, Isola S and Vulpiani A 1990 Correlation functions and relaxation properties in chaotic dynamics and statistical mechanics *Phys. Lett. A* **144** 341
- [17] Falcioni M and Vulpiani A 1995 The relevance of chaos for the linear response theory *Physica A* **215** 481
- [18] Hatano T and Sasa S-i 2001 Steady-state thermodynamics of langevin systems *Phys. Rev. Lett.* **86** 3463–3466
- [19] Verley G, Chétrite R and Lacoste D 2011 Modified fluctuation-dissipation theorem near non-equilibrium states and applications to the Glauber-Ising chain *J. Stat. Mech.* P10025
- [20] Prost J, Joanny J F and Parrondo J M 2009 Generalized fluctuation-dissipation theorem for steady-state systems *Phys. Rev. Lett.* **103** 090601
- [21] Leith C 1975 Climate response and fluctuation-dissipation *J. Atmos. Sci.* **32** 2022
- [22] Abramov R V and Majda A J 2007 Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems *Nonlin.* **20** 2793–2821
- [23] Lacorata G, Puglisi A and Vulpiani A 2009 On the fluctuation-response relation in geophysical systems *Int. J. Mod. Phys. B* **23** 5515–5529
- [24] Sagawa T and Ueda M 2009 Minimum energy cost for thermodynamic information processing: Measurement and information erasure *Phys. Rev. Lett.* **102** 250602
- [25] Toyabe S, Sagawa T, Ueda M, Muneyuki E and Sano M 2010 Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality *Nature Phys.* **6** 988–992
- [26] Chétrite R, Falkovich G and Gawędzki K 2008 Fluctuation relations in simple examples of non-equilibrium steady states *J. Stat. Mech.* P08005
- [27] Chétrite R and Gawędzki K 2009 Eulerian and Lagrangian pictures of non-equilibrium diffusions *J. Stat. Phys.* **137** 890–91
- [28] Ruelle D 1998 General linear response formula in statistical mechanics and the fluctuation-dissipation theorem far from equilibrium *Phys. Lett. A* **245** 220–224
- [29] Abramov R V and Majda A J 2008 New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems *J. Nonlin. Sci.* **18** 303–341
- [30] Lucarini V and Sarno S 2011 A statistical mechanical approach for the computation of the climatic response to general forcings *Nonlin. Proc. Geophys.* **18** 7–28
- [31] Young L S 2002 What are SRB measures, and which dynamical systems have them? *Dyn. Sys.* **108** 733–754
- [32] Tasaki S, Gilbert T and Dorfman J R 1998 An analytical construction of the SRB measures for Baker-type maps *Chaos* **8** 424–443
- [33] Eyink G L, Haine T W N and Lea D J 2004 Ruelle’s linear response formula, ensemble adjoint schemes and Lévy flights *Nonlin.* **17** 1867–1889
- [34] Lucarini V 2012 Stochastic perturbations to dynamical systems: A response theory approach *J. Stat. Phys.* **146** 774–786
- [35] Goldbeter A 1995 A model for circadian oscillations in the *Drosophila* period protein (PER) *Proc. Royal Soc. Lond. B* **261** 319–324
- [36] Ogawa K, Takekawa N, Uchida K and Shibata S 2006 On robust stability and sensitivity of circadian rhythms *Asian J. Control* **8** 281–289
- [37] Dorfman J R 1999 An introduction to chaos in nonequilibrium statistical mechanics *Cambridge*

*Lecture Notes in Physics* vol 14 (Cambridge)

- [38] Chatelain C 2003 A far-from-equilibrium fluctuation-dissipation relation for an Ising-Glauber-like model *J. Phys. A: Math. Gen* **36** 10739
- [39] Ricci-Tersenghi F 2003 Measuring the fluctuation-dissipation ratio in glassy systems with no perturbing field *Phys. Rev. E* **68** 065104(R)
- [40] Crisanti A and Ritort J 2003 Violation of the fluctuation-dissipation theorem in glassy systems: basic notions and the numerical evidence *J. Phys. A: Math. Gen* **36** R181–R290
- [41] Diezemann G 2005 Fluctuation-dissipation relations for Markov processes *Phys. Rev. E* **72** 011104
- [42] Lippiello E, Corberi F and Zannetti M 2005 Off-equilibrium generalization of the fluctuation dissipation theorem for Ising spins and measurement of the linear response function *Phys. Rev. E* **71** 036104
- [43] Andrenacci N, Lippiello E and Corberi F 2006 Fluctuation-dissipation relation in an Ising model without detailed balance *Phys. Rev. E* **73** 046124
- [44] Lippiello E, Corberi F and Zannetti M 2007 Fluctuation dissipation relations far from equilibrium *J. Stat. Mech.* P07002
- [45] Cugliandolo L, Kurchan J and Parisi G 1994 Off equilibrium dynamics and aging in unfrustrated systems *J. Phys. I* **4** 1641
- [46] Harada T and Sasa S-i 2006 Energy dissipation and violation of the fluctuation-response relation in nonequilibrium langevin systems *Phys. Rev. E* **73** 026131
- [47] Harada T and Sasa S-i 2005 Equality connecting energy dissipation with violation of fluctuation-response relation *Phys. Rev. Lett.* **95** 130602
- [48] Zwanzig R 1961 Memory effects in irreversible thermodynamics *Phys. Rev.* **124** 983
- [49] Mori H 1965 Transport, collective motion, and Brownian motion *Prog. Theor. Phys.* **33** 423
- [50] Onsager L and Machlup S 1953 Fluctuations and irreversible processes *Phys. Rev.* **91** 1505–1512
- [51] Bohec P, Gallet F, Maes C, Safaverdi S, Visco P and Van Wijland F 2012 Probing active forces via a fluctuation-dissipation relation *arXiv:1203.3571*
- [52] Toyabe S, Okamoto T, Watanabe-Nakayama T, Taketani H, Kudo S, and Muneyuki E 2010 Nonequilibrium Energetics of a Single F<sub>1</sub>-ATPase Molecule, *Phys. Rev. Lett.* **104** 198103
- [53] Baiesi M, Boksenbojm E, Maes C and Wynants B 2010 Linear response of nonequilibrium states, II: inertial dynamics *J. Stat. Phys.* **139** 492–505
- [54] Baiesi M, Maes C and Wynants B 2009 Fluctuations and response of nonequilibrium states *Phys. Rev. Lett.* **103** 010602
- [55] Baiesi M, Maes C and Wynants B 2009 Nonequilibrium linear response for Markov dynamics, I: Jump processes and overdamped diffusions *J. Stat. Phys.* **137** 1094–1116
- [56] Maes C and Wynants B 2010 On a response function and its interpretation *Markov Proc. Rel. Fields* **16** 45–58
- [57] Maes C and Netočný K 2003 Time-reversal and entropy *J. Stat. Phys.* **110** 269–310
- [58] Maes C 2003 On the origin and the use of fluctuation relations for the entropy *Séminaire Poincaré* **2** 29–62
- [59] Baiesi M, Maes C and Wynants B 2011 The modified Sutherland-Einstein relation for diffusive non-equilibria *Proc. Royal Soc. A* **467** 2792–2809
- [60] Evans D J and Morriss G P 1990 *Statistical Mechanics of NonEquilibrium Liquids* Theoretical Chemistry Monograph Series (London: Academic Press)
- [61] Andrieux D and Gaspard P 2007 A fluctuation theorem for currents and non-linear response coefficients *J. Stat. Mech.* P02006
- [62] Mallick K, Moshe M and Orland H 2011 A field-theoretic approach to nonequilibrium work identities *J. Phys. A: Math. Gen* **44** 095002
- [63] Lippiello E, Corberi F, Sarracino A and Zannetti M 2008 Nonlinear response and fluctuation dissipation relations *Phys. Rev. E* **78** 041120
- [64] Villamaina D, Baldassarri A, Puglisi A and Vulpiani A 2009 The fluctuation-dissipation relation: how does one compare correlation functions and responses? *J. Stat. Mech.* P07024

- [65] Evans D J, Searles D J and Williams S R 2008 On the fluctuation theorem for the dissipation function and its connection with response theory *J. Chem. Phys.* **128** 014504
- [66] Colangeli M, Maes C and Wynants B 2011 A meaningful expansion around detailed balance *J. Phys. A: Math. Gen* **44** 095001
- [67] Saito K Private communication (Leuven), January 2012
- [68] Boksenbojm E, Maes C, Netočný K and Pesek J 2011 Heat capacity in nonequilibrium steady states *Europhys. Lett.* **96** 40001
- [69] Cugliandolo L F 2011 The effective temperature *J. Phys. A: Math. Gen* **44** 483001
- [70] Ben-Isaac E, Park Y K, Popescu G, Brown F L H, Gov N S and Shokef Y 2011 Effective temperature of red-blood-cell membrane fluctuations *Phys. Rev. Lett.* **106** 238103
- [71] Calabrese P and Gambassi A 2005 Ageing properties of critical systems *J. Phys. A: Math. Gen* **38** R133–R193
- [72] Shimizu A 2010 Universal properties of response functions of nonequilibrium states *J. Phys. Soc. Jpn.* **79** 113001
- [73] Shimizu A and Yuge T 2010 General properties of response functions of nonequilibrium steady states *J. Phys. Soc. Jpn.* **79** 013002
- [74] Helfand E 1960 Transport coefficients from dissipation in a canonical ensemble *Phys. Rev.* **119** 1–9
- [75] Martin P, Hudspeth A J, and Jülicher F 2001 Comparison of a hair bundle’s spontaneous oscillations with its response to mechanical stimulation reveals the underlying active process *Proc. Natl. Acad. Sci. U.S.A.* **98** 14380
- [76] Kikuchi N, Ehrlicher A, Koch D, Käs J, Ramaswamy S, and Rao M 2009 Buckling, stiffening, and negative dissipation in the dynamics of a biopolymer in an active medium *Proc. Natl. Acad. Sci. U.S.A.* **106** 19776
- [77] Seifert U 2010 Generalized Einstein or Green-Kubo relations for active biomolecular transport *Phys. Rev. Lett.* **104** 138101
- [78] Maes C and Netočný K 2007 Static and dynamical nonequilibrium fluctuations *Comptes Rendus Physique* **8** 591–597