J Stat Phys (2009) 137: 1094–1116
DOI 10.1007/s10955-009-9852-8

Nonequilibrium Linear Response for Markov Dynamics, I: Jump Processes and Overdamped Diffusions

Marco Baiesi · Christian Maes · Bram Wynants

Received: 9 July 2009 / Accepted: 20 October 2009 / Published online: 3 November 2009

Abstract Systems out of equilibrium, in stationary as well as in nonstationary regimes, display a linear response to energy impulses simply expressed as the sum of two specific temporal correlation functions. There is a natural interpretation of these quantities. The first term corresponds to the correlation between observable and excess entropy flux yielding a relation with energy dissipation like in equilibrium. The second term comes with a new meaning: it is the correlation between the observable and the excess in dynamical activity or reactivity, playing an important role in dynamical fluctuation theory out-of-equilibrium. It appears as a generalized escape rate in the occupation statistics. The resulting response formula holds for all observables and allows direct numerical or experimental evaluation, for example in the discussion of effective temperatures, as it only involves the statistical averaging of explicit quantities, e.g. without needing an expression for the nonequilibrium distribution. The physical interpretation and the mathematical derivation are independent of many details of the dynamics, but in this first part they are restricted to Markov jump processes and overdamped diffusions.

Keywords Nonequilibrium · Fluctuation-dissipation · Linear response

1 Introduction and Main Result

To know a system is to know its reaction to external stimuli. Linear response theory applies when the effects are comparable to the causes and so far has been systematized only for

M. Baiesi · C. Maes · B. Wynants (✉)
Instituut voor Theoretische Fysica, K. U. Leuven, 3001 Leuven, Belgium
e-mail: bram.wynants@fys.kuleuven.be

M. Baiesi
e-mail: baiesi@pd.infn.it

C. Maes
e-mail: christian.maes@fys.kuleuven.be

M. Baiesi
Dipartimento di Fisica, Università di Padova, Via Marzolo 8, 35131 Padova, Italy

systems in complete thermodynamic equilibrium [5, 26, 44]. The result is summarized in the fluctuation-dissipation theorem [10, 32] stating that the response of an equilibrium system to small external perturbations is proportional to its fluctuations: an impulse changing the potential $U \to U - h_s V$ at time s, will produce a response $R^{\text{eq}}_{QV}(t,s) = \delta\langle Q(t)\rangle/\delta h_s|_{h=0}$ in a quantity Q at time $t \geq s$ given by

$$R^{\text{eq}}_{QV}(t,s) = \beta \frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_{\text{eq}}. \tag{1}$$

Here $\langle V(s)Q(t)\rangle_{\text{eq}}$ is the equilibrium time correlation function between $V(s)$ and $Q(t)$ and $\beta = 1/k_B T$ is the inverse temperature. For example, for an Ising spin system $\sigma_j = \pm 1$ in equilibrium with energy function $U(\sigma)$, where one looks for the response in the magnetization $Q = \sum_j \sigma_j = V$ to a (time-dependent) magnetic field, one finds the (time-dependent) susceptibility in terms of the spin-spin correlations in the equilibrium process:

$$R^{\text{eq}}_{QV}(t,s) = \mathcal{R}(t-s) = \beta \sum_{i,j} \frac{\mathrm{d}}{\mathrm{d}s}\langle \sigma_i(s)\sigma_j(t)\rangle_{\text{eq}}, \quad 0 < s < t.$$

When $h_s = h$ is constant over $s \in [0,\tau]$, then to first order in h and for $t \geq \tau$, we get the thermoremanent magnetization by integrating the previous formula:

$$\sum_i \left[\langle \sigma_i(t)\rangle^h - \langle \sigma_i(0)\rangle_{\text{eq}}\right] = h\beta \sum_{i,j} \langle [\sigma_i(\tau) - \sigma_i(0)]\sigma_j(t)\rangle_{\text{eq}}. \tag{2}$$

The equilibrium process $\langle\cdot\rangle_{\text{eq}}$ is time-reversible and can for example be specified here by a Glauber spin flip dynamics leaving the equilibrium distribution $\sim \exp[-\beta U(\sigma)]$ invariant. The usual proof of (1)–(2) proceeds by applying first-order time-dependent perturbation theory and by inserting the equilibrium condition. We review it in Appendix C and we indicate the problem with this approach for extending it to nonequilibrium conditions.

In the present paper we generalize the equilibrium response formula (1) to nonequilibrium regimes (stationary as well as nonstationary ones). In this first part we treat Markov jump processes and overdamped diffusions as simple nonequilibrium models ignoring inertial degrees of freedom. In this context, formula (1) changes to (5) below. A second part [4] following this paper will give extensions, in particular to underdamped diffusion processes, which are noisy Hamiltonian systems, and to general space-time local processes that are defined from their distribution on path-space.

To get a first understanding of the meaning of our generalization it is instructive to view already (1) not just as the outcome of a perturbative calculation on the level of evolution operators, but rather in its relation with dissipative effects, cf. [22]. The point is that (1) is a correlation with the entropy flux associated to the perturbation. To see it, we apply (1) to write the difference $\langle Q(t)\rangle^h - \langle Q(t)\rangle_{\text{eq}}$ between the perturbed average and the original value as the time-integral

$$\beta \int_0^t \mathrm{d}s h_s \frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_{\text{eq}} = \beta \left\langle \left\{ h_t V(t) - h_0 V(0) - \int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s} h_s V(s) \right\} Q(t) \right\rangle_{\text{eq}} \tag{3}$$

and we recognize the correlation of $Q(t)$ with the entropy flux

$$\frac{1}{k_B T}\left\{ h_t V(t) - h_0 V(0) - \int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s} h_s V(s) \right\}. \tag{4}$$

Indeed $h_t V(t) - h_0 V(0)$ is the extra change of energy in the environment due to the perturbation, and $\int_0^t \mathrm{d}s \dot{h}_s V(s)$ is the work done on the system by the perturbation. To be precise, the entropy flux here is an excess flux of the perturbed process with respect to the unperturbed one ($\propto U(t) - U(0)$). This has an immediate analogue under nonequilibrium conditions when we think of the entropy flux (4) as the excess with respect to whatever steady or transient entropy production the system might already possess.

However, from dynamical fluctuation theory we have learnt that there is more than just dissipation or entropy production that governs the fluctuations around nonequilibrium [37–39]. A novel quantity has appeared, which we call dynamical activity or traffic, a measure for the system's nervosity or internal reactivity. This dynamical activity is time-symmetric, in contrast with the entropy production which leads time's arrow. It was first introduced in [34, 41] and studied in the context of nonequilibrium phase transitions [23], of critical behavior [25], for characterizing the ratchet effect [18], and for large deviation theory, e.g. in [2, 8, 37]. Close to equilibrium, entropy production and dynamical activity merge, as can be understood from observing that when approaching equilibrium the currents loose their direction to become merely elements of (time-symmetric) traffic between the states of the system. Further away from equilibrium, dynamical activity and entropy production are really different and play side by side in characterizing the nonequilibrium. Then, formula (1) splits in two separate terms, one entropic, the other *frenetic*, to be explained below. That is made visible in a generalized response formula, written here for the nonequilibrium situations to be specified below: (1) must be changed into

$$R^{\mu}_{QV}(t,s) = \frac{\beta}{2}\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_{\mu} - \frac{\beta}{2}\langle LV(s)Q(t)\rangle_{\mu}\,. \tag{5}$$

In (5) we have an arbitrary initial distribution μ for a given nonequilibrium process (over which we average on the right-hand side). The first term is exactly like in equilibrium (but with $1/2$); the second term is the correlation with the linearized dynamical activity (the correlation between $Q(t)$ and the function βLV at time s) which is new with respect to the equilibrium (1). For its interpretation we restrict us here first to jump processes. For the dynamical activity we must then look at the escape rates, i.e. at the frequencies by which the Markov jump process leaves a state x. Suppose the process has rates $W(x,y)$ for the transitions $x \to y$. The escape rate at state x gets changed by the perturbation $h_s V$. Its excess is

$$\sum_y W(x,y)\left\{e^{\frac{\beta h_s}{2}[V(y)-V(x)]} - 1\right\} \simeq \frac{\beta h_s}{2}\sum_y W(x,y)[V(y)-V(x)] \equiv \frac{\beta h_s}{2}(LV)(x) \tag{6}$$

to linear order in h, and L is the backward generator of the jump process:

$$LV(x) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}\langle V(s)\rangle_x.$$

In (5) we have abbreviated $LV(s) = LV(x_s)$ for x_s the state at time s. More details will come in the next section, where the precise set-up and meaning will be given. Section 4 gives examples and makes visible the relative contribution of the two terms in (5). Most importantly the new formula is presented in Sects. 5 and 6 with its statistical mechanical interpretation. Section 7 makes the relation with the ambition of effective temperature. Section 8 still connects our work with previous formulations, especially those in [13, 15, 19, 35, 36, 47, 50]. The Appendices give (A) explicit calculations for illustrating the formulæ of Sect. 7, (B) the response-formula for discrete time Markov chains, useful for simulation purposes, and (C) the comparison with formal perturbation theory and with the co-moving frame interpretation of [13].

2 What Nonequilibrium?

We consider open system first-order dynamics realized in Markov jump and diffusion processes. These define probabilities on the trajectories $\omega = (\omega_t, t \geq 0)$ of the system by specifying the updating mechanism $\omega_t \to \omega_{t+\mathrm{d}t}$. The state at each time is denoted by $\omega_t = x$ and corresponds to a reduced description of the universe, such as positions or elements of a discrete set of configurations. Steady equilibrium reservoirs are integrated out and replaced by effective external forces and noise. The nonequilibrium condition can be imposed by external driving fields or by installing mechanical displacements or chemical gradients at the boundaries of the system but all external driving is here assumed to be time-independent. Examples will come in Sect. 4; further details will not matter.

The updating is given in terms of a *generator* L in the sense that for all single-time observables $f(x)$ (functions of the configuration x of the system) and for all initial distributions μ:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle f(\omega_t)\rangle_\mu = \langle (Lf)(\omega_t)\rangle_\mu \tag{7}$$

which is equivalent to saying that

$$\langle f(\omega_t)\rangle_\mu = \int \mathrm{d}x\,\mu(x)(e^{tL}f)(x) = \langle e^{tL}f\rangle_\mu.$$

In words, e^{tL} "pulls a function $f(x)$ back to the time of the initial density $\mu(x)$." L is therefore often called the backward generator. More in general, for any two observables f and g, their correlations at times $0 < t < s$ satisfy

$$\frac{\mathrm{d}}{\mathrm{d}s}\langle f(\omega_s)g(\omega_t)\rangle = \langle (Lf)(\omega_s)g(\omega_t)\rangle, \quad 0 < t < s \tag{8}$$

for an arbitrary average at time zero and over all allowed trajectories.

We now specify the two classes of processes we deal with in this first part.

2.1 Jump Processes

We consider a finite space K of states $x, y, \ldots$ on which transition rates $W(x,y)$ give the probability per unit time to jump between the states $x \to y$. No detailed balance is assumed. We write ρ for a stationary distribution, $\sum_{y\in K}[\rho(x)W(x,y) - \rho(y)W(y,x)] = 0$, $x \in K$. In this case the generator L is the matrix with off-diagonal elements $L_{xy} = W(x,y)$ and with diagonal elements equal to minus the escape rates, $L_{xx} = -\sum_y W(x,y)$. Hence

$$Lf(x) = \sum_y W(x,y)[f(y) - f(x)]$$

and by stationarity $\sum_x \rho(x) Lf(x) = 0$.

We think of the states x as configurations of a mesoscopic system undergoing a Markov evolution, such as for chemical kinetics or for models of interacting components on a lattice which are driven away from equilibrium, cf. Example 4.1; the state x is then the total configuration and the transitions are local; response in discrete time is treated in Appendix B.

At time $t = 0$ we draw the initial state from an arbitrary probability distribution $\mu(x)$, $x \in K$. For $t > 0$ we apply the perturbed dynamics with transition rates

$$W_t(x, y) = W(x, y)e^{\frac{\beta h_t}{2}[V(y)-V(x)]} \tag{9}$$

for some potential V with small amplitudes h. We thus switch on another channel of energy exchange with a reservoir at temperature T. This is a standard type of perturbation in Markov jump processes by which we add a potential; we will stick to that. More general types of perturbation are possible, as for example in [19] but for these the specific formula (5) has to be modified, see [40].

2.2 Overdamped Diffusions

Overdamped diffusions are stochastic processes defined in the Itô-sense by

$$\mathrm{d}x_t = \big\{\nu(x_t)\big[F(x_t) - \nabla U(x_t)\big] + \nabla \cdot D(x_t)\big\}\mathrm{d}t + \sqrt{2D(x_t)}\mathrm{d}B_t. \tag{10}$$

Now $x \in \mathbb{R}^d$ and the noise is present in the form of the d-dimensional vector $\mathrm{d}B_t$ having independent standard Gaussian white noise components. We assume that the environment is in thermal equilibrium at inverse temperature $\beta > 0$ imposing the condition $\nu(x) = \beta D(x)$ between the bare mobility ν and the diffusion matrix D; they are strictly positive (symmetric) $d \times d$-matrices. The relation $\nu = \beta D$ is easily confused with the fluctuation-dissipation theorem; they are only equivalent in equilibrium. More in general, $\nu = \beta D$ is an expression of local detailed balance assuring the proper physical identification of the various terms in (10). There is a potential U and the force F represents the nonequilibrium driving. We do not specify here regularity properties or boundary conditions. Equilibrium (1) is obtained when $F = 0$, for stationary $\rho \propto \exp(-\beta U)$. The diffusions (10) are called *overdamped* because they have forces proportional to velocities. These processes are high damping limits (sometimes called Smoluchowski limits) of underdamped or inertial stochastic dynamics that will be considered in Part II [4]. On the other hand, formal extensions to stochastic and driven Ginzburg-Landau models (models A-B-C in [9]) do not seem difficult. Examples come in Sects. 4.2–4.3.

The generator of the dynamics (10) reads

$$L = \nu(F - \nabla U) \cdot \nabla + \nabla D \cdot \nabla. \tag{11}$$

The time-dependent perturbation changes U in (10) into $U - h_t V$ for $t \geq 0$. In other words the Hamiltonian part of the dynamics of the system gets an additional conservative force with a time-dependent amplitude. Here also we assume that ν and D are unchanged under adding the potential $h_t V$.

2.3 Main Question

What to expect at time $t > 0$ for an observable Q? There will be a shift in its value, deviating both from the original expectation and from the stationary value:

$$\begin{aligned}\langle Q(t)\rangle_\mu^h &\neq \langle Q(t)\rangle_\mu \\ &\neq \langle Q(t)\rangle_\rho = \langle Q(0)\rangle_\rho\end{aligned} \tag{12}$$

Note that we abbreviate $Q(t) = Q(x_t)$. The right-hand sides average over the unperturbed dynamics, the upper one starting from μ and, on the next line, when starting from the stationary ρ. Linear response theory for systems out of equilibrium aims at estimating and interpreting the deviations

$$\langle Q(t)\rangle_\mu^h - \langle Q(t)\rangle_\mu$$

to first order in h. In this linear regime, one studies the deviations of an observable $\langle Q(t)\rangle^h$ from its expected value $\langle Q(t)\rangle$ in the unperturbed dynamics,

$$\langle Q(t)\rangle^h = \langle Q(t)\rangle + \int_0^t \mathrm{d}s\, h_s R_{QV}(t,s) \tag{13}$$

where

$$R_{QV}(t,s) \equiv \frac{\delta\langle Q(t)\rangle^h}{\delta h_s}\Big|_{h=0} \tag{14}$$

is the response of a quantity $Q(t)$ to a (small) impulse h_s at a previous time $s < t$.

The initial distribution μ will be remembered via super- and subscripts, where appropriate. A special and interesting case is the stationary response for $\mu = \rho$.

3 Nonequilibrium Response Formula

Under the previous set-up there is a simple general formula for the linear response derived in [3], the same for all observables Q and V:

$$R_{QV}^\mu(t,s) = \frac{\beta}{2}\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_\mu - \frac{\beta}{2}\langle LV(s)Q(t)\rangle_\mu\,. \tag{15}$$

The first example of a formula of this kind is (22) in [36], dealing with discrete spin systems. Since random paths are not smooth, it does not make sense to take the time-derivative inside the expectation in the first term. On the other hand, for $s > t$, (8) can be applied and as a consequence causality is automatically verified in (15): the response vanishes for $s > t$. One can thus also rewrite (15) in a more symmetric way which is however only valid for $t \geq s$:

$$R_{QV}^\mu(t,s) = \frac{\beta}{2}\left[\frac{\mathrm{d}}{\mathrm{d}s}\big\langle V(s)Q(t)\big\rangle_\mu - \frac{\mathrm{d}}{\mathrm{d}t}\big\langle V(t)Q(s)\big\rangle_\mu\right] - \frac{\beta}{2}\big\langle LV(s)Q(t) - LV(t)Q(s)\big\rangle_\mu. \tag{16}$$

Equations of this form can also be found back in the literature [15, 19, 36]. (See more in Sect. 8.) *Our main goal here is to extend these earlier results and others (such as in* [6, 13]) *into the general and usable formula* (15)*, and, most importantly, to accompany it with an interpretation derived from dynamical fluctuation theory.* That will be continued in [4], showing further robustness of the interpretation.

In the introduction Sect. 1 we have already briefly introduced the statistical meaning of the two terms on the right-hand side of (5) = (15). More is to come in Sects. 5 and 6. The first term in (15) (as in (1)) is a correlation with the excess entropy flux. Secondly, there is the correlation between $Q(t)$ and $\beta LV(s)$, a quantity that we call *frenesy*, derived from the adjective frenetic or frantic. In contrast to the entropy production, which has a preferred direction in time, frenesy is a time-symmetric quantity.

Let us also see how formula (15) reconstructs the equilibrium formula (1). In this case $\mu = \rho$ is the equilibrium distribution with the process satisfying time-reversal symmetry, with $s < t$

$$\begin{aligned}\langle LV(s)Q(t)\rangle_{\text{eq}} &= \langle LV(t)Q(s)\rangle_{\text{eq}} = \frac{\mathrm{d}}{\mathrm{d}t}\langle V(t)Q(s)\rangle_{\text{eq}} \\ &= -\frac{\mathrm{d}}{\mathrm{d}s}\langle V(t)Q(s)\rangle_{\text{eq}} = -\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_{\text{eq}}\end{aligned} \tag{17}$$

which is the first term in (15). Since by (17) in equilibrium the frenesy-correlation exactly equals minus the correlation with the entropy flux, the two terms on the right-hand side of (15) add up to give (1).

We end the section with the formal proof of formula (15), which is easy when going to path-space. We should write the perturbed expectation value in terms of the unperturbed one. More explicitly we consider paths $\omega = (\omega_s)$ over the time-interval $s \in [0, t]$ to write

$$\langle Q(t)\rangle^h = \int \mathrm{d}P_\mu(\omega)\frac{\mathrm{d}P_\mu^h}{\mathrm{d}P_\mu}(\omega)Q(\omega_t)$$

where the $\mathrm{d}P$'s stand for path-probability densities in the perturbed (superscript h) and unperturbed processes. In particular, for the jump processes of Sect. 2.1,

$$P_\mu(\omega) = \mu(\omega_0)W(\omega_0, \omega_{t_1})e^{-\sum_y \int_0^{t_1}\mathrm{d}s W(\omega_s, y)} \ldots W(\omega_{t_{n-1}}, \omega_{t_n})e^{-\sum_y \int_{t_n}^{t}\mathrm{d}s W(\omega_s, y)}$$

for a path $\omega = (\omega_0 \to \omega_{t_1} \to \cdots \to \omega_{t_n} = \omega_t)$. As a consequence,

$$\begin{aligned}\log\frac{\mathrm{d}P_\mu^h}{\mathrm{d}P_\mu}(\omega) &= \frac{\beta}{2}\sum_{k=1}^{n} h_{t_k}\big[V(\omega_{t_k}) - V(\omega_{t_{k-1}})\big] \\ &\quad - \sum_y \int_0^t \mathrm{d}s W(\omega_s, y)\big[e^{\frac{\beta h_s}{2}[V(y)-V(\omega_s)]} - 1\big]\end{aligned} \tag{18}$$

where the first sum is over all the jump times $(t_1, \ldots, t_n)$ in ω, and we put $t_0 = 0$ and $t_{n+1} = t$. Remember that the path is constant between the jump times so that

$$\begin{aligned}\int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s}h_s V(\omega_s) &= \sum_{k=0}^{n} V(\omega_{t_k})\big[h_{t_{k+1}} - h_{t_k}\big] \\ &= h_t V(\omega_t) - h_0 V(\omega_0) + \sum_{k=1}^{n} h_{t_k}\big[V(\omega_{t_{k-1}}) - V(\omega_{t_k})\big]\end{aligned}$$

by partial summation. We can therefore rewrite the first line of (18) and substitute (4). The rest is expansion to first order in h for a finite number of terms, almost surely under $\mathrm{d}P_\mu(\omega)$. In particular, the frenetic term appears as the first order in the time-symmetric part (second line of (18)), like in (6).

The strategy is unchanged for overdamped diffusions. There, the paths $\omega = (x_s)$ are continuous and the action (18) can be found in the stochastic integral

$$\begin{aligned}\log\frac{\mathrm{d}P_\mu^h}{\mathrm{d}P_\mu}(\omega) = \frac{\beta}{2}\bigg\{&\int_0^t \mathrm{d}x_s \nabla V(x_s)h_s \\ &+ \int_0^t \mathrm{d}s\, \nu(\nabla U - F)(x_s)\cdot h_s \nabla V(x_s) + O(h^2)\bigg\}.\end{aligned} \tag{19}$$

The first Itô-integral can be rewritten in the Stratonovich sense (with the "$\circ$"-notation in the integral)

$$\int_0^t \mathrm{d}x_s \nabla V(x_s) h_s = \int_0^t \mathrm{d}x_s \circ \nabla V(x_s) h_s + \int_0^t \mathrm{d}s h_s \nabla D \cdot \nabla V(x_s) \tag{20}$$

where the last term can be combined with the second line in (19) to make $\beta L V/2$, see (11). Moreover,

$$\left\langle \int_0^t \mathrm{d}x_s \circ \nabla V(x_s) h_s Q(x_t) \right\rangle_\mu = \int_0^t \mathrm{d}s h_s \frac{\mathrm{d}}{\mathrm{d}s} \langle V(s) Q(t) \rangle_\mu .$$

The rest is again trivial expansion.

4 Examples

It is often more convenient to visualize the integrated version of (15) for a small but constant perturbation $h_s = h$, $s \geq 0$. Then, the generalized susceptibility

$$\chi(t) = \frac{1}{h} \left[\langle Q(t) \rangle^h - \langle Q(t) \rangle \right],$$

is given by

$$\chi(t)/\beta = \frac{1}{2}[C(t) + K(t)] \equiv C_{NE}(t) \tag{21}$$

with correlation function (coming from the entropic term in (15))

$$C(t) = \langle V(t) Q(t) \rangle - \langle V(0) Q(t) \rangle$$

and a term (coming from the frenetic term in (15), extra with respect to equilibrium)

$$K(t) = -\int_0^t \mathrm{d}s \langle L V(s) Q(t) \rangle$$

representing an integrated correlation function. The average of C and K, denoted by $C_{NE}(t)$ in (21), thus has to be equal to $\chi(t)/\beta$ in general. If extended to $t \uparrow \infty$, $\chi(t)$ gives the change in nonequilibrium stationary expectation when adding a small potential.

4.1 Driven Kawasaki Dynamics

Consider an exclusion process as a model of ionic transport through a narrow channel. This is described by a collection of n sites, labeled by $i = 1, \ldots, n$, each holding either one particle ($x^i = 1$) or none ($x^i = 0$). In the bulk of this system no particles are created or annihilated, only jumping to neighboring sites is allowed via a Kawasaki dynamics. At the edges $i = 1, n$ particles can move in or out from reservoirs with density d_1 and d_n, respectively. The form of the response formula (15) is unchanged by adding a nearest neighbor interaction with energy $U(x) = -\sum_{i=1}^n x^i x^{i+1}$. Moreover, we can add an "electric" field E promoting particle jumps to the right. Since it is a Markov jump process, there are transition rates for particles hopping to neighboring sites and rates for creation and annihilation at the edges. For example, a particle enters into site $i = 1$ from the reservoir with rate

$$W(x, y) = d_1 \psi(x, y) \exp\left\{ -\frac{\beta}{2}[U(y) - U(x)] \right\}, \quad \psi(x, y) = \psi(y, x)$$

where $y = x$ except that $x^1 = 0$ while $y^1 = 1$. The rate of the reverse transition can be found by imposing local detailed balance. Similarly, in the bulk, when e.g. $x^i = 1, x^{i+1} = 0$ and y is reached by inverting these occupations,

$$W(x, y) = \psi(x, y) \exp\left\{ -\frac{\beta}{2}[U(y) - U(x) - E] \right\}, \quad \psi(x, y) = \psi(y, x).$$

Simple nonequilibrium conditions can thus be introduced either by

(i) setting different reservoir densities $d_1 \neq d_n$, or
(ii) setting a nonzero electric field $E > 0$ in the bulk.

We choose the total number of particles $\mathcal{N}(t) = \sum_{i=1}^{n} x^i$ as observable. We also introduce a perturbation $V(s)$ equal to $\mathcal{N}(s)$, which means that we are changing the chemical potential of both reservoirs with a common shift. Transition rates for the perturbed process are thus multiplied by a factor $e^{\beta h_s/2}$ if a particle enters the system, and by $e^{-\beta h_s/2}$ when a particle leaves; transitions in the bulk are left unchanged.

In this case the frenetic term is with $LV = \mathcal{J}(s)$, the systematic current, and represents the expected change of the potential $\mathcal{N}$ per unit time, i.e., it is the rate of change in the number of particles from the two possible transitions at boundary sites. Thus, for $x \to y$ the transition modifying x^1, and for $x \to z$ the transition modifying x^n, we have

$$\mathcal{J}(x) = [\mathcal{N}(y) - \mathcal{N}(x)]W(x, y) + [\mathcal{N}(z) - \mathcal{N}(x)]W(x, z).$$

We have numerically verified that $\chi = C = K$ under equilibrium conditions. While $C = K$ to excellent precision, the shape of χ depends weakly on h, and is found to converge to C only for h sufficiently small. In fact, one can pretend exact matching only in the limit $h \to 0$, but $h = 0.01$ turns out to be sufficiently small to achieve a good convergence.

A representative example for the nonequilibrium case (i) is shown in Fig. 1(a). One can see that each of the functions $C(t)$ and $K(t)$ is a poor approximation of the response, while the agreement between $C_{NE}(t)$ and $\chi(t)/\beta$ is excellent.

In Fig. 1(b) we show an example of nonequilibrium condition (ii). Again, only $C_{NE}(t)$ matches $\chi(t)/\beta$. Curiously, a comparison of this example with the previous one reveals that C can be either larger or smaller than K, even for two nonequilibrium conditions that look pretty similar, in the sense that they yield a current in the same direction for a relatively simple system.

4.2 Diffusion on the Circle

Consider the Langevin equation for a particle position $x_t \in S^1$ on a circle driven by a constant field $F(x) = f$ and subject to a time-dependent forcing:

$$\mathrm{d}x_t = \nu[f - U'(x_t) - h_t V'(x_t)]\mathrm{d}t + \sqrt{2D}\mathrm{d}B_t. \tag{22}$$

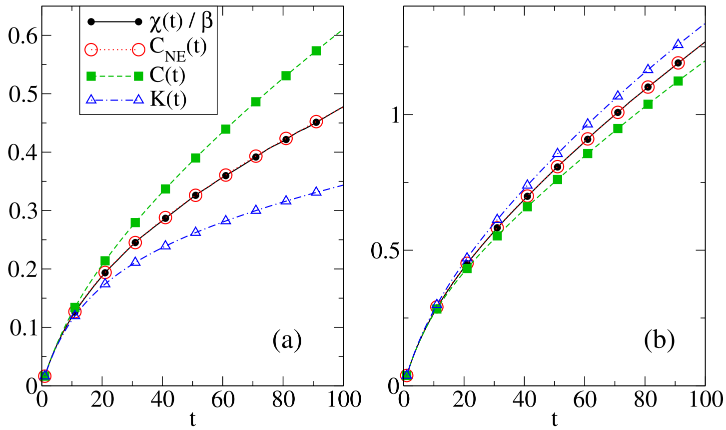

Fig. 1 Plot of the quantities involved in (21), for (**a**) case (i) ($E = 0$) with $L = 10$, $\beta = 1$, $h = -0.01$, and reservoir density unbalance $d_1 = 0.9$, $d_n = 0.1$, and (**b**) for case (ii) ($d_1 = d_n = 0.5$) with $L = 10$, $\beta = 1$, $h = -0.01$, $E = 3$

Here the prime denotes differentiation with respect to space. The term $h_t V'$ is the time-dependent perturbation while $\mathrm{d}B_t$ is a standard Gaussian white noise. The diffusion constant D and the mobility ν are related by the Einstein relation $\nu = \beta D$, with β the inverse temperature.

This example has been recently experimentally realized as reported in [24], for testing the co-moving frame interpretation of [13] also explained in Appendix C. In fact, its nonequilibrium stationary distribution ρ is known analytically, see [38].

Here we show the result of a simple simulation of the overdamped particle studied in the experiment, with potentials and observable $U(x) = Q(x) = V(x) = \sin(x)$, and a constant force f. The frenesy $\beta L V(x)$ of a particle at position x can be computed by applying the generator of the overdamped dynamics

$$L = (f - U')\frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{\beta}\frac{\mathrm{d}^2}{\mathrm{d}x^2}$$

to the potential $V(x)$, which gives $\beta L V(x) = \beta[f - \cos(x)]\cos(x) - \sin(x)$. We have measured the correlations $C(t)$ and $K(t)$ as well as responses $\chi(t)$ for small h, see (21).

Figure 2(a) shows that for strong stationary nonequilibrium $f \gg 0$ there is a large difference between the integrated correlation function of the frenesy and that of the entropy production. However, their average $C_{NE}(t)$ agrees very well with the response $\chi(t)$, as it should.

It is important however to recall that our approach works for nonstationary regimes as well. In Fig. 2(b) we show an example of a particle starting at time $t = 0$ from a nonstationary initial distribution $\mu(x) = \delta_{x,0}$ (i.e. its position is $x_0 = 0$), but for the rest not forced outside equilibrium, $f = 0$, to emphasize the transient character of this situation. Again, we can see that the response is well estimated by $C_{NE}(t)$.

In the second part [4] we will treat the inertial version of this model.

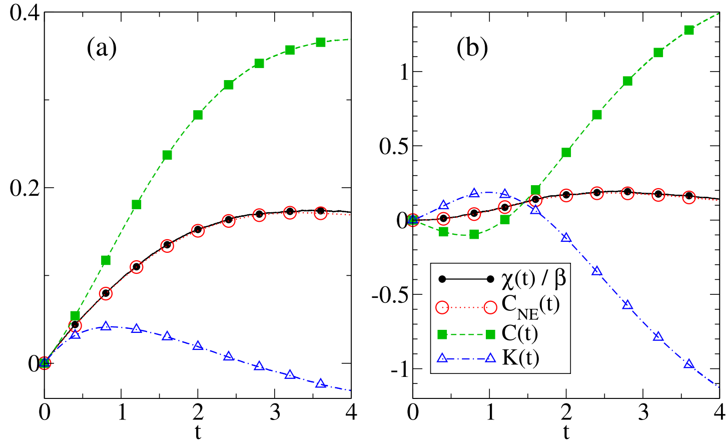

Fig. 2 Response and fluctuations of the overdamped particle in a tilted periodic potential, as discussed in the text, with inverse temperature $\beta = 0.2$, mobility $\nu = 1$ and perturbation $h = -0.02$. (**a**) Steady state regime, with initial distribution equal to the stationary one, and force $f = 0.9$. (**b**) Transient regime, with initial position $x_0 = 0$ and force $f = 0$

4.3 Driven Brownian Motion

Consider again the overdamped Langevin equation (10) but with constant $\nu = \beta D$:

$$(\mathrm{d}x_t^i) = \nu^{ij}[f^j(x_t) + h^j]\mathrm{d}t + \sqrt{2D}^{ij}\,\mathrm{d}B_t^j$$

with repeated indices $j = 1, \dots d$ summed over. The f includes all forces and h is the small constant perturbation. The true mobility M is defined as the response

$$M^{ij} = \lim_{t\to\infty} \frac{d}{dh^j} \left\langle \dot{x}_t^i \right\rangle_\rho^h \bigg|_{h=0} \tag{23}$$

while the real diffusion matrix is given by

$$\mathcal{D}^{ij} = \lim_{t\to\infty} \frac{1}{2t} \int_0^t \mathrm{d}s \int_0^t \mathrm{d}r \left\langle (\dot{x}_s^i - \langle \dot{x}^i \rangle_\rho)(\dot{x}_r^j - \langle \dot{x}^j \rangle_\rho) \right\rangle_\rho .$$

As (23) is a response function, we can easily compute it within our framework. After a straightforward calculation, it follows that

$$M^{ij} = \beta\mathcal{D}^{ij} - \frac{\beta\nu^{jk}}{2} \lim_{t\to\infty} \frac{1}{t} \int_0^t \mathrm{d}s \int_0^t \mathrm{d}r \left\langle f^k(x_s)(\dot{x}_r^i - \langle \dot{x}^i \rangle_\rho) \right\rangle_\rho . \tag{24}$$

Note also that $\langle \dot{x}^i \rangle_\rho = \langle \nu^{ij} f^j(x) \rangle_\rho$. Observe that in equilibrium, the second term in (24) vanishes because the observable $\dot{x}_t$ is anti-symmetric, and $\langle \dot{x} \rangle_{\mathrm{eq}} = 0$. Then, the equilibrium fluctuation-dissipation relation holds with $M = \beta\mathcal{D}$. Its violation for nonequilibrium, i.e., that there is a correction (= second term) in (24), is completely compatible with the condition $\nu = \beta D$ for the reservoir in thermal equilibrium at inverse temperature β.

5 Dissipation: the Entropic Term

Even at equilibrium, the response to an external perturbation is a nonequilibrium process involving a transient with production of entropy. In steady nonequilibrium conditions instead we deal with the *excess* in this entropy production, as the starting point is already a regime with nonvanishing flows.

In equilibrium it is known that the response function is closely related to the energy dissipation of the system into the environment. In systems driven out of equilibrium things get more complicated: even when the system is not perturbed there is already a (time extensive) heat dissipation. This is the housekeeping heat that is needed to maintain the (unperturbed) nonequilibrium stationary state. Perturbing the system then gives additional heat. We see here that the prediction in [46] that the usual equilibrium relation between response and dissipation is preserved when taking into account only the *excess* heating and ignoring the housekeeping heat, is indeed true in the following sense.

During a specific trajectory of the system, the heat dissipation can be split into two parts: $\mathcal{Q}_{hk} + \mathcal{Q}_{ex}$ [30, 46]. The first term is the housekeeping heat while $\mathcal{Q}_{ex}$ is the extra heat generated through the perturbation, and is just the temperature times the excess entropy production, which we already encountered in our calculations (3)–(4).

Let us assume that the amplitude h_s is periodic with period t so that it makes sense to consider the expectation of the excess heat over the interval $[0,t]$:

$$\langle \mathcal{Q}_{ex}\rangle_\rho^h = \int_0^t \mathrm{d}s h_s \frac{\mathrm{d}}{\mathrm{d}s} \langle V(s)\rangle_\rho^h .$$

We expand the expectation value of V using the definition of the response function R_{VV}:

$$\langle \mathcal{Q}_{ex}\rangle_\rho^h = \int_0^t \mathrm{d}s h_s \frac{\mathrm{d}}{\mathrm{d}s} \left[\langle V(s)\rangle_\rho + \int_0^s \mathrm{d}r R_{VV}(s,r) h_r \right].$$

Here $\langle V(s)\rangle_\rho$ is independent of s because of stationarity and $R_{VV}(s,r) = \mathcal{R}(s-r) = 0$ for $s<r$ so that

$$\langle \mathcal{Q}_{ex}\rangle_\rho^h = -\int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s} h_s \int_{-\infty}^{+\infty} \mathrm{d}r \mathcal{R}(r) h_{s-r}$$

where the right-hand side is exactly the same expression as in equilibrium for the dissipated energy over one period. As is standard, see e.g. [26, 44], the last relation can be rewritten in the frequency domain to obtain now the (excess) heat proportional to the imaginary part of the Fourier transform of $R_{VV} = \mathcal{R}$. We conclude that also out-of-equilibrium the imaginary part of the Fourier transform of the generalized susceptibility is related to energy dissipation, but only of the excess heat. Observe finally that the imaginary part of the Fourier transform is the Fourier transform of the time-antisymmetric part of R_{VV} and that we have already written down this time-antisymmetric part: it is the right-hand side of (16), extended to $t \le s$.

6 Activity: the Frenetic Term

As we have seen, besides the entropy flux there is yet another relevant quantity getting into the picture. It is a younger concept in nonequilibrium studies as witnessed by the fact that its name has not been settled yet: it has been called *traffic* [37–39] or dynamical *activity*

[2, 8, 23, 25], to denote a property related with the volume of transitions or changes performed in time. We like to refer to it as frenesy, a name that sounds more similar to entropy or energy. This quantity is time-symmetric, in the sense that trajectories display the same activity if they are spanned in the normal temporal direction or backward in time. The relevance of time-symmetric quantities has been anticipated in [34, 41]. Activity denotes thus an aspect of the dynamics that complements entropy production (which is time anti-symmetric – flows reverse together with time) in the description of fluctuating quantities in regimes out of equilibrium.

We recall here how the frenesy/dynamical activity appears in dynamical fluctuation theory, choosing the set-up of Sect. 2.1 (finite state space Markov jump processes x_t).

Consider the stationary process P_ρ and measure the fraction of time that the system spends in each state $x \in K$:

$$p_\tau(x) = \frac{1}{\tau}\int_0^\tau \delta_{x_t,x}\,\mathrm{d}t, \qquad \text{with } \delta_{a,b} = 0 \text{ if } a \neq b \text{ and } \delta_{a,b} = 1 \text{ if } a = b.$$

That is the empirical distribution of occupation times over $[0, \tau]$ while drawn at time zero from the stationary ρ. When the stationary process P_ρ is an ergodic Markov process, p_τ is invariant under time-reversal and $p_\tau \to \rho$ for $\tau \uparrow +\infty$. Following the pioneering work in [20] we look here at the fluctuations around that law of large times, for which it is known that

$$\mathrm{Prob}_\rho[p_\tau \simeq \mu] \simeq e^{-\tau I(\mu)}, \qquad \tau \uparrow +\infty.$$

See e.g. [17] for a precise description of this fluctuation formula. The exponent is governed by the functional $I(\mu) \geq 0$, with equality for $I(\rho) = 0$. There is a variational expression for the fluctuation functional $I(\mu)$,

$$I(\mu) = \sup_V \left\{ -\sum_x \mu(x) \left[\sum_y W(x, y) e^{\frac{\beta}{2}[V(y)-V(x)]} - \sum_y W(x, y) \right] \right\} \tag{25}$$

in terms of the excess escape rates when adding a potential $-V$ much as in (9). The actual supremum is then reached for the potential V that makes μ stationary; i.e., by changing $W(x, y) \to W(x, y)\exp\{\beta[V(y) - V(x)]/2\}$. Thus, for that μ-dependent potential V,

$$\begin{aligned} I(\mu) &= \sum_{x,y} \mu(x) W(x, y) - \sum_y \mu(x) W(x, y) e^{\beta[V(y)-V(x)]/2} \\ &\simeq -\frac{\beta}{2} \sum_x \mu(x) L V(x). \end{aligned}$$

In the second line (expected frenesy in μ) we have assumed that μ is close to the stationary measure (small fluctuations), and we have written the first term in an expansion around $\mu - \rho$, where V is supposed to be small as well, again like in (6). In other words, the rate $I(\mu)$ of escape from density μ is given in terms of the expected frenesy under μ.

For further information on the role of activity/frenesy in dynamical fluctuation theory, we refer to [37–39].

7 To Effective Temperature

While the equilibrium fluctuation–dissipation relation (1) is typically violated for nonequilibrium regimes, one may wonder whether it can sometimes be restored by the introduction of an effective temperature T^{eff}, in the sense

$$R^{\mu}_{QV}(t,s) = \frac{1}{k_B T^{\text{eff}}} \frac{\mathrm{d}}{\mathrm{d}s} \langle V(s)Q(t)\rangle_\mu. \tag{26}$$

Many studies have been devoted to the study of the prefactor, in what sense it perhaps resembles a thermodynamic temperature-like quantity for some classes of observables and over some scales of times $(s/t, s)$; we refer to [9, 14, 31, 33] for an entry into the extensive literature. Clearly, whatever the purpose of the discussion, an exact expression of the response should help, especially when entirely in terms of explicit correlation functions. The first calculations in this sense are in [11] and they have been referred to as the "no field-method" [48]. In particular, for purposes of simulation or numerical verification of (26) we do no longer need to perform the perturbation by hand. In fact, now we can write the ratio $T/T^{\text{eff}} = X$ entirely in terms of correlation functions

$$X = X_{QV}(\mu; t, s) = \frac{1}{2}\left[1 - \frac{\langle LV(s)Q(t)\rangle_\mu}{\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_\mu}\right] \tag{27}$$

with numerator and denominator in (27) each having a specific physical meaning as in the previous two sections. An effective temperature is obtained as the ratio between the frenetic and the entropic term: if for some observables (V, Q) and over time-scales $(t/s, t)$,

$$Y \frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_\mu = \langle LV(s)Q(t)\rangle_\mu$$

for some Y, then $X = (1 - Y)/2$. Equilibrium has $X = 1 = -Y$. In the case where $LV \approx 0$ as for a conserved quantity, then $Y = 0$. The simplest example is Brownian motion $L = \Delta$ for V the position, cf. Virasoro's example in [15]. In that last reference, what are called "flat directions" can be associated to perturbations with zero frenesy $LV = 0$. In Appendix A we give some explicit results to illustrate the above ideas and formulæ.

Finally, X and the effective temperature T^{eff} get negative when the frenetic term overwhelms the entropic contribution, which we can expect whenever the perturbation strongly activates the dynamics. It is not clear to us whether that relates with the observations in [47] concerning "activated dynamics," but certainly, those "active states" with negative effective temperature, here interpreted as highly frenetic ones are very different from equilibrium states.

One should understand that (26) represents a rather optimistic scenario. Formula (26) wants to mimic (1) by replacing just one parameter. Why should there be also out-of-equilibrium a single parameter and a useful notion of temperature in its usual thermodynamic understanding, and how would it depend on the observables V and Q? (See [43] for a very recent discussion.) Answers to these questions have been partially given but are often restricted within a context of mean field systems or for small fluctuations, effectively dealing with calculations as in Appendix A, similar to calculations for scalar fields as in [15] and in [9]. In fact, the optimism in (26) is a sort of conservatism as it wants to attach special reference to equilibrium forms and chooses to continue working with equilibrium notions such as the temperature. We take a different attitude: the violation of the equilibrium fluctuation-dissipation relation (FDR) is an opportunity to discover new connections between response and dynamical fluctuations away from equilibrium, and to identify these relations in terms of the relevant newly emerging physical quantities.

Let us then try to see how the notion of effective temperature could be seen as a one-parameter reduction of a general equilibrium-like FDR that is valid also outside equilibrium but with an effective dynamics. The starting point is observing that in the case of equilibrium, formula (1) is equivalent with

$$R^{\text{eq}}_{QV}(t,s) = -\beta\langle (LV)(s)Q(t)\rangle_{\text{eq}}. \tag{28}$$

That follows from the calculation (17). In other words, the fluctuation-dissipation theorem in equilibrium can also be called a fluctuation-frenesy theorem; the two terms on the right-hand side of (15) are simply the same. Therefore, for purposes of getting closer to equilibrium response formulæ one really has the choice to mimic either (1) or rather (28). The first leads to the ambition of effective temperature (26), the latter to the new notion of effective frenesy. But the latter is also much richer. In fact a simple calculation, that we postpone to the end of Appendix C, shows that the exact nonequilibrium response formula (15) can indeed be written in the equilibrium form (28):

$$R^{\mu}_{QV}(t,s) = -\langle G_{\mu_s} V(s)Q(t)\rangle_{\mu} \tag{29}$$

with a new effective frenesy

$$\begin{aligned} G_\mu V &= \frac{\beta}{2\mu}[L^\dagger(\mu V) - VL^\dagger\mu + \mu LV] \\ &= \frac{\beta\rho}{2\mu}\left[L^*\left(\frac{\mu}{\rho}V\right) - VL^*\left(\frac{\mu}{\rho}\right) + \frac{\mu}{\rho}LV\right]. \end{aligned} \tag{30}$$

Here $L^\dagger$ is the forward generator[1] and for jump processes $L^\dagger g(x) = \sum_y [W(y,x)g(y) - W(x,y)g(x)]$, and L^* is the adjoint generator, see (35) in Appendix C. The first line in (30) does not need to assume a stationary distribution ρ, while it is explicitly and implicitly (in L^*) present in the second line of (30). At any rate the operator G_μ acting on V in (30) has the following exact property: it is itself a generator (just like the original L in (7) or (8)) but of a new dynamics for which μ is an equilibrium distribution (i.e. $\langle f(G_\mu g)\rangle_\mu = \langle g(G_\mu f)\rangle_\mu$ for all f, g). Thus, in (29) the generator G_{μ_s} is the instantaneous equilibrium generator with respect to the time-evolved distribution μ_s. This underlies the conclusions of [12].

We compare (29) with (28) and we recognize the equilibrium form with L for the equilibrium ρ replaced by G_{μ_s} for the transient μ_s. In the stationary nonequilibrium case, we have $\mu_s = \rho$ and (30) is

$$G_\rho = \frac{\beta}{2}(L + L^*) \tag{31}$$

replacing L in (28). Hence, if the perturbation V is time-direction independent in the precise sense that $LV = L^*V$, then the nonequilibrium response ((37) = (15) for $\mu = \rho$) reduces to the equilibrium formula (1) and $X = 1$. See [45] for very related conjectures and observations.

[1] It is called the forward generator because of the definition of the time-evolved measure μ_t:

$$\langle f(x_t)\rangle_\mu = \int \mathrm{d}x\,\mu_t(x) f(x).$$

From this we see that $\mu_t = e^{tL^\dagger}\mu$, so that $L^\dagger$ pushes the measure μ forward in time.

8 Some Other Previous Formulations

The literature on (extensions of) the fluctuation-dissipation theorem is vast. We apologize for not being able to list all possible and even essential contributions. The equilibrium formulation spans all of the previous century, while nonequilibrium versions started to appear since the 1970's and very much continue up to now. Early works include [1, 51] and also [16, 27], where we see a discussion within the theory of stochastic dynamics. In contrast to [27, 28] our unperturbed process is time-homogeneous. Coming to more recent times, violations of equilibrium FDR have been most often discussed in transient regimes. For example, in the context of aging phenomena [14, 31], much thought has been given to making sense of an effective temperature as briefly discussed in Sect. 7.

However, recently much new work has been also directed to find generic extensions of FDR in nonequilibrium steady states and to discussions of the dissipative elements in relaxations to nonequilibrium. There is for example the line starting from [29, 30, 45, 46] which treats nonequilibrium heat effects. In particular [29] might also be useful in real experiments because the FDR violation is there directly connected with the energy dissipation. We do not know yet how to relate that to the new ideas surrounding the frenetic term in our work or to the physics of "active states" as also in [47]. The latter follows the line of [22], where the connection between the so called fluctuation theorem and the fluctuation-dissipation theorem was first explained.

For other recent extensions of the FDR, we refer to [6, 7, 21, 49, 50]. We also mentioned before (at the end of Sect. 7 and more will come in Appendix C) how our approach is related to the co-moving frame interpretation of [12, 13, 24]. There, as in the previous references, one disadvantage is that one keeps the stationary density ρ (or its logarithm) as observable in the fluctuation formula. In our approach the largely unknown distribution only enters in the statistical averaging. Being optimistic one could say that the generalized FDR of [21, 42] can be used to verify hypothesis on the phase space distribution function. Or, one can imagine approximation schemes for making the numerics possible at all, cf. [50]. In the same sense the extensions in [6, 13] are also not explicit as they need the information on the adjoint hydrodynamics or again on the local probability current, but at least they come with a new and still interesting interpretation (reverse response in [6] and co-moving frame in [12, 13]).

A first generalized fluctuation-dissipation relation giving a response formula *in our sense* appears in Sect. 2 of [15]. It treats a Langevin dynamics $y(t)$ for soft spin models and the resulting equation (2.10) in [15] is exactly our formula (16) for $Q = V$ equal to y. A very similar treatment is repeated in [36] for systems of Ising spins (i.e., within the class of Markov jump processes). Some more general treatment again for jump processes is offered in [19], in particular its (16)–(17). However as before and as mentioned already in the abstract of [19], "the asymmetry... is not related to any physical observable." In contrast, we have emphasized the interpretation via dynamical activity/frenesy in nonequilibrium fluctuation theory. In fact, that interpretation is exactly what makes a systematic generalization possible at all, as will become even more clear in [4]. The only study in which we recognize some of the ideas related to the frenetic term is in [47], in the context of dynamical systems.

9 Conclusions

We have studied linear response relations under general nonequilibrium conditions (stationary and not). This first part has dealt with jump processes and with overdamped diffusions, which is the usual set-up for discussions on the violation of the fluctuation-dissipation theorem and for the possible emergence of an effective temperature. Here we have stressed the emergence of the dynamical activity, or frenesy in linear order, as complementary to the entropy flux. Out of equilibrium, dissipation and activity detach and the response needs to be evaluated in terms of both entropic and frenetic correlation functions. As both correlations are expressed in terms of explicit averages, they constitute a formula ready to use in a general context. For example, estimates of these correlations can be obtained with usual averaging in simulations, without any need to know further details or approximations of the stationary density of states. On the theoretical side, several previous approaches are recovered or have been extended within the same scheme. In many cases they have been discussed in specific model dynamics or for specific observables. It is interesting that there is a unifying approach with statistical interpretation behind this very broad variety of previous results.

Within the models we considered, more general types of perturbation are possible and interesting. For example, what will happen if the perturbation is not of the potential type but actually changes the nonequilibrium part or driving of the dynamics? In all these cases our results must be modified, of course, but we think that the general method and framework we use is applicable to this wider set of questions. Specifically, the concepts of entropy and frenesy will remain the major actors in nonequilibrium fluctuation-response theory.

Acknowledgements The authors are grateful to Wojciech De Roeck, Andrea Gambassi, Krzysztof Gawędzki, Dragi Karevski and Karel Netočný for very useful discussions. B.W. receives support from FWO, Flanders. M.B. benefits from K.U.Leuven grant OT/07/034A.

Appendix A: Explicit Calculations for Linear Diffusions

We give here the results of some explicit calculations to illustrate formula (27) in Sect. 7. Similar calculations can be found in [9, 15].

Fix parameters $\alpha, B \in \mathbb{R}$, $D > 0$ and look at the linear Langevin dynamics for a global order parameter $M \in \mathbb{R}$,

$$\dot{M}(t) = -\alpha M(t) + h_t B + \sqrt{2D}\xi(t)$$

for standard white noise $\xi(t)$. The $h_t, t > 0$, is a small time-dependent field. The generator of the unperturbed dynamics (on observables f) is $Lf(M) = -\alpha M f'(M) + Df''(M)$. We can think of a Gaussian approximation to a relaxational dynamics of the scalar magnetization M (no conservation laws and no spatial structure) valid in high enough dimensions (above $d = 4$ for the standard Ising model). Then, in a way, $\alpha = 0$ corresponds to the critical (massless) dynamics and $\alpha > 0$ is a paramagnetic dynamics (high temperature). By taking $D \downarrow 0$ we exclude the diffusive aspects and we can think then of gradient relaxation in the low temperature regime.

The equilibrium (reversible stationary density on $\mathbb{R}$ for perturbation $B = 0$) is

$$\rho(M) = \frac{1}{Z} \exp\left\{-\alpha \frac{M^2}{2D}\right\}$$

with zero mean and variance $\langle M^2\rangle = D/\alpha$. (Here, $\alpha > 0$ is needed but not for the finite-time existence of the dynamics.)

We now start from an initial fixed $M(t = 0) = M_0$. The response function is obtained from

$$\langle M(t)\rangle_{M_0}^h = M_0 e^{-\alpha t} + B \int_0^t \mathrm{d}s\, h_s e^{-\alpha(t-s)}$$

or

$$\frac{\delta}{\delta h_s}\langle M(t)\rangle_{M_0}^{h}(h=0) = B e^{-\alpha(t-s)} \tag{32}$$

which does in fact not depend on M_0 (and thus also equals the equilibrium result).

The correlation function for $0 < s < t$ is

$$\langle M(s)M(t)\rangle_{M_0} = M_0^2 e^{-\alpha(t+s)} + \frac{D}{\alpha}\big[e^{-\alpha(t-s)} - e^{-\alpha(t+s)}\big]$$

and hence

$$\frac{\mathrm{d}}{\mathrm{d}s}\langle M(s)M(t)\rangle_{M_0} = -\alpha M_0^2 e^{-\alpha(t+s)} + D\big[e^{-\alpha(t-s)} + e^{-\alpha(t+s)}\big].$$

When we average that last expression over the equilibrium density (thus replacing M_0^2 by D/α) we find

$$\frac{\mathrm{d}}{\mathrm{d}s}\langle M(s)M(t)\rangle_{\rho} = D e^{-\alpha(t-s)}$$

which, in comparison with (32) specifies the equilibrium temperature to be equal to $T = D/k_B$.

The frenetic term is obtained from $LM = -\alpha M$, and thus

$$\langle LM(s)M(t)\rangle_{M_0} = -\alpha M_0^2 e^{-\alpha(t+s)} - D\big[e^{-\alpha(t-s)} - e^{-\alpha(t+s)}\big].$$

Clearly,

$$\frac{\delta}{\delta h_s}\langle M(t)\rangle_{M_0}^{h}(h=0) = \frac{B}{2D}\left\{\frac{\mathrm{d}}{\mathrm{d}s}\langle M(s)M(t)\rangle_{M_0} - \langle LM(s)M(t)\rangle_{M_0}\right\}$$

as it should.

For the issue of effective temperature we compute the ratio "frenesy versus entropy" as

$$Y = Y(M_0; s,t) = \frac{-\alpha\langle M(s)M(t)\rangle_{M_0}}{\frac{\mathrm{d}}{\mathrm{d}s}\langle M(s)M(t)\rangle_{M_0}} = \frac{-\alpha M_0^2 e^{-\alpha(t+s)} - D[e^{-\alpha(t-s)} - e^{-\alpha(t+s)}]}{-\alpha M_0^2 e^{-\alpha(t+s)} + D[e^{-\alpha(t-s)} + e^{-\alpha(t+s)}]}.$$

In that notation, the effective inverse temperature is $T^{\mathrm{eff}} = 2T/(1-Y)$. In equilibrium $Y = -1$ while $Y = 1$ for $D = 0$ and $M_0 \neq 0$.

For the limit of the "frenetic ratio"

$$\lim_{s\uparrow+\infty}\lim_{t\uparrow+\infty} Y(M_0; s,t) = \begin{cases} -1, & \text{if } \alpha > 0,\\ 0, & \text{if } \alpha = 0\end{cases}$$

with respectively, $T^{\mathrm{eff}} = T$ (paramagnetic) and $T^{\mathrm{eff}} = 2T$ (critical quench), cf. [9].

Appendix B: Discrete Time

So far we have been dealing with continuous time Markov processes. Nevertheless simulations often use an updating in discrete time. It is therefore useful to give the response formula also for discrete time, here in analogy with the continuous time jump processes of Sect. 2.1. Similar efforts have appeared before in [11, 36].

Consider a discrete time Markov process with configurations $x, y, \ldots$ and transition probabilities $p(x, y)$. We perturb the system as follows: the new transition probability at time i reads

$$p_i^h(x, y) = \frac{p(x, y)}{z_i(x)} e^{\frac{h_i \beta}{2}[V(y)-V(x)]}$$

with normalization

$$z_i(x) = \sum_y p(x, y) e^{\frac{h_i \beta}{2}[V(y)-V(x)]}.$$

Now consider a path $\omega = (x_i)$, with $i = 1, \ldots, n$, where at time i the system jumps to configuration x_{i+1}. Then the relative probability of the path in the perturbed versus the unperturbed dynamics is

$$\frac{P^h}{P}(\omega) = \exp\left\{\frac{\beta}{2}\sum_{i=1}^{n-1} h_i[V(x_{i+1}) - V(x_i)] - \sum_{i=1}^{n-1} \log z(x_i)\right\}.$$

The response function becomes

$$\begin{aligned} R_{Q,V}(n, m) &= \frac{\partial}{\partial h_m} \langle Q(x_n)\rangle_\mu^h \Big|_{h=0} \\ &= \frac{\beta}{2} \langle V(x_{m+1}) Q(x_n)\rangle_\mu - \frac{\beta}{2}\left\langle \sum_y p(x_m, y) V(y) Q(x_n)\right\rangle_\mu, \quad m = 1, \ldots, n-1. \end{aligned}$$

Again we recognize this as the sum of an entropic term

$$\frac{\beta}{2} \langle [V(x_{m+1}) - V(x_m)] Q(x_n)\rangle_\mu$$

and a frenetic term

$$\frac{\beta}{2}\left\langle \sum_y p(x_m, y)[V(y) - V(x_m)] Q(x_n)\right\rangle_\mu.$$

In equilibrium ρ, when the process is time-reversal symmetric, that simplifies to $(m < n)$

$$R_{Q,V}^{\text{eq}}(n, m) = \frac{\beta}{2} \langle [V(x_{m+1}) - V(x_{m-1})] Q(x_n)\rangle_{\text{eq}}.$$

Note that in this discrete version, the entropic and the frenetic terms do not give exactly the same contribution to the equilibrium formula, although this discrepancy disappears in the limit to continuous time.

Appendix C: First Order Perturbation

The usual and natural approach to response theory is that of time-dependent perturbation theory, see for example [44]. We show here what that gives for the overdamped diffusion processes of Sect. 2.2.

We prepare the system at time $t = 0$ according to its stationary distribution ρ; the perturbation $-hV$ is added for positive times. Therefore, for times $t \geq 0$ the dynamics has (backward) generator (working on observables)

$$L^h = L + h\nu \nabla V \cdot \nabla, \quad L = \nu(F - \nabla U + \nabla D) \cdot \nabla$$

with $\nu = \beta D$. For the change in expectations at times t with respect to what we had at time zero

$$\langle Q(t)\rangle^h - \langle Q(0)\rangle = \int \mathrm{d}x \rho(x) \big(e^{tL^h} - e^{tL}\big) Q(x)$$

we get the linear order

$$e^{tL^h} - e^{tL} = \int_0^t e^{(t-s)L}(L^h - L)e^{sL}\mathrm{d}s + O(h^2).$$

Or, always to leading order in $h \downarrow 0$,

$$\frac{1}{h}\big[\langle Q(t)\rangle^h - \langle Q(0)\rangle\big] = \int_0^t \mathrm{d}s R_{QV}(t,s)$$

with response function

$$R_{QV}(t,s) = \int \mathrm{d}x \rho(x) \nu \nabla V(x) \cdot \nabla e^{(t-s)L} Q(x). \tag{33}$$

(We are still writing the dependence separately on time s and on time t, for greater generality in case the perturbation is time-dependent (through h_s).) Equation (33) is not useful as such because the derivatives do not commute with the time-evolution. Another way would be trying the partial integration

$$\nabla \cdot \big(\rho \nu \nabla V\big) = \nabla \rho \cdot \nu \nabla V + \rho \nabla \cdot \nu \nabla V = \beta \rho \left\{ LV - \frac{j_\rho}{\rho} \cdot \nabla V \right\}$$

where we have inserted the local stationary current j_ρ (assuming that ρ nowhere vanishes). As a result

$$R_{QV}(t,s) = -\beta \int \mathrm{d}x \rho(x) \mathcal{L} V(x) e^{(t-s)L} Q(x) \quad \text{for } \mathcal{L} = L - u(x) \cdot \nabla \tag{34}$$

with local velocity

$$u(x) = \frac{j_\rho}{\rho}(x)$$

function of the stationary density ρ.

We can further rewrite that using the adjoint generator L^* with respect to ρ: for any pair of functions f, g

$$\int \mathrm{d}x \rho(x) g(x) L^* f(x) = \int \mathrm{d}x \rho(x) f(x) L g(x). \tag{35}$$

Its physical interpretation is that it generates the time-reversed stationary process. It is important to realize though that this adjoint generator L^* depends on the stationary distribution

ρ, and is therefore most often not explicitly known in nonequilibrium systems. For equilibrium where $\rho \propto \exp(-\beta U)$, $L = L^*$.

We can verify that $L - L^* = 2u \cdot \nabla$ so that $\mathcal{L} = L^* + u \cdot \nabla$ in (34):

$$R_{QV}(t,s) = \beta \frac{\mathrm{d}}{\mathrm{d}s} \int \mathrm{d}x \rho(x) V(x) e^{(t-s)L} Q(x) - \beta \int \mathrm{d}x \rho(x) u(x) \cdot \nabla V(x) e^{(t-s)L} Q(x). \tag{36}$$

For jump processes, under Sect. 2.1,

$$((L - L^*)V)(x) = 2 \sum_y \frac{j(x,y)}{\rho(x)} [V(y) - V(x)],$$
$$j(x,y) = W(x,y)\rho(x) - W(y,x)\rho(y).$$

Equations (34)–(36) contain the interpretation that the equilibrium form gets "restored" when describing the system in the Lagrangian frame moving with drift velocity u, as in [13]. Note that $\mathcal{L} = L - L_A$, which subtracts the antisymmetric part $L_A = (L - L^*)/2$ from the original generator $L = L_S + L_A$. What remains is the symmetric part $\mathcal{L} = L_S$, of course defining an evolution which is now detailed balance with respect to ρ. That aspect can also be realized path-wise: when $\Phi_t(x)$ is the one-parameter group corresponding to the flow with velocity $u(x)$,

$$\partial_t \Phi_t(x) = u(\Phi_t(x))$$

then, $y_t = \Phi_{-t}(x_t)$ satisfies an equilibrium Langevin equation with potential $-\beta^{-1} \ln \rho$ but with time-dependent coefficients (in mobility). In other words, the passage to the Lagrangian frame of local velocity removes the non-conservative forcing, as explained in [12, 13]. Still, if we do not know ρ, the formulæ (34)–(36) contain unknown observables for statistical averaging. Moreover, thinking of spatial processes, the probability current has little relation with the real physical currents. For our formula (15), one has an explicit expression in terms of known observables and ρ only enters the statistical averaging.

Let us now see how the above expressions are mathematically related to our formula (15). For stationary nonequilibrium, taking $\mu = \rho$, it is immediate to rewrite (15) as

$$R_{QV}(t,s) = \beta \frac{\mathrm{d}}{\mathrm{d}s} \langle V(s) Q(t) \rangle_\rho - \frac{\beta}{2} \big\langle \big(L - L^*\big) V(s) Q(t) \big\rangle_\rho . \tag{37}$$

In fact,

$$\frac{\mathrm{d}}{\mathrm{d}s} \langle V(s) Q(t) \rangle_\rho = -\frac{\mathrm{d}}{\mathrm{d}t} \langle V(s) Q(t) \rangle_\rho = -\langle (L^* V)(s) Q(t) \rangle_\rho \tag{38}$$

so that also

$$R_{QV}(t,s) = -\beta \left\langle \left(\frac{L + L^*}{2} V \right)(s) Q(t) \right\rangle_\rho \tag{39}$$

as we have also seen in (31), or in (34) because $\mathcal{L} = (L + L^*)/2$.

Also for transient regimes we can rewrite the response formula (15) as what one would get in equilibrium plus a correction term:

$$R^\mu_{QV}(t,s) = \beta \frac{\mathrm{d}}{\mathrm{d}s} \langle V(s) Q(t) \rangle_\mu + \beta \langle \tilde{G}_{\mu_s} V(s) Q(t) \rangle_\mu \tag{40}$$

where now

$$\tilde{G}_\mu V = \frac{1}{2\mu}[-VL^\dagger\mu + L^\dagger(\mu V) - \mu LV].$$

Formula (40) reduces to (37) for $\mu = \rho$.

References

1. Agarwal, G.S.: Fluctuation-dissipation theorems for systems in non-thermal equilibrium and applications. Z. Phys. **252**, 25–38 (1972)
2. Appert-Rolland, C., Derrida, B., Lecomte, V., Van Wijland, F.: Universal cumulants of the current in diffusive systems on a ring. Phys. Rev. E **78**, 021122 (2008)
3. Baiesi, M., Maes, C., Wynants, B.: Fluctuations and response of nonequilibrium states. Phys. Rev. Lett. **103**, 010602 (2009)
4. Baiesi, M., Boksenbojm, E., Maes, C., Wynants, B.: Linear response of nonequilibrium states, II: inertial dynamics. In preparation
5. Balakrishnan, V.: Elements of Nonequilibrium Statistical Mechanics. Ane Books India (2008)
6. Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Macroscopic fluctuation theory for stationary non equilibrium, states. J. Stat. Phys. **107**, 635–675 (2002)
7. Blickle, W., Speck, T., Lutz, C., Seifert, U., Bechinger, C.: Einstein relation generalized to nonequilibrium. Phys. Rev. Lett. **98**, 210601 (2007)
8. Bodineau, T., Lefevere, R.: Large deviations of lattice Hamiltonian dynamics coupled to stochastic thermostats. J. Stat. Phys. **133**, 1 (2008)
9. Calabrese, P., Gambassi, A.: Ageing properties of critical systems. J. Phys. A, Math. Gen. **38**, R133–R193 (2005)
10. Callen, H.B., Welton, T.A.: Irreversibility and generalized noise. Phys. Rev. **83**, 34 (1951)
11. Chatelain, C.: A far-from-equilibrium fluctuation-dissipation relation for an Ising-Glauber-like model. J. Phys. A **36**, 10739 (2003)
12. Chetrite, R., Gawędzki, K.: Eulerian and Lagrangian pictures of non-equilibrium diffusions. arXiv:0905.4667
13. Chetrite, R., Falkovich, G., Gawędzki, K.: Fluctuation relations in simple examples of non-equilibrium steady states. J. Stat. Mech. P08005 (2008)
14. Crisanti, A., Ritort, J.: Violation of the fluctuation-dissipation theorem in glassy systems: basic notions and the numerical evidence. J. Phys. A, Math. Gen. **36**, R181–R290 (2003)
15. Cugliandolo, L., Kurchan, J., Parisi, G.: Off equilibrium dynamics and aging in unfrustrated systems. J. Phys. I **4**, 1641 (1994)
16. Deker, U., Haake, F.: Fluctuation-dissipation theorems for classical processes. Phys. Rev. A **11**, 2043 (1975)
17. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. Springer, New York (1998)
18. De Roeck, W., Maes, C.: Symmetries of the ratchet current. Phys. Rev. E **76**, 051117 (2007)
19. Diezemann, G.: Fluctuation-dissipation relations for Markov processes. Phys. Rev. E **72**, 011104 (2005)
20. Donsker, M.D., Varadhan, S.R.: Asymptotic evaluation of certain Markov process expectations for large time, I. Commun. Pure Appl. Math. **28**, 1–47 (1975)
21. Falcioni, M., Isola, S., Vulpiani, A.: Correlation functions and relaxation properties in chaotic dynamics and statistical mechanics. Phys. Lett. A **144**, 341 (1990)
22. Gallavotti, G.: Chaotic hypothesis: Onsager reciprocity and fluctuation-dissipation theorem. J. Stat. Phys. **84**, 899–926 (1996)
23. Garrahan, J.P., Jack, R.L., Lecomte, V., Pitard, E., van Duijvendijk, K., van Wijland, F.: First-order dynamical phase transition in models of glasses: an approach based on ensembles of histories. J. Phys. A, Math. Gen. **42**, 075007 (2009)
24. Gomez-Solano, J.R., Petrosyan, A., Ciliberto, S., Chetrite, R., Gawędzki, K.: Experimental verification of a modified fluctuation-dissipation relation for a micron-sized particle in a non-equilibrium steady state. Phys. Rev. Lett. **103**, 040601 (2009)
25. Gorissen, M., Hooyberghs, J., Vanderzande, C.: DMRG-study of current and activity fluctuations near non-equilibrium phase transitions. Phys. Rev. E. **79**, 020101 (2009)
26. de Groot, S.R., Mazur, P.: Non-Equilibrium Thermodynamics. North-Holland, Amsterdam (1962), Chap. VIII
27. Hänggi, P., Thomas, H.: Linear response and fluctuation theorems for nonstationary stochastic processes. Z. Phys. B **22**, 295–300 (1975)

28. Hänggi, P., Thomas, H.: Stochastic processes: time-evolution, symmetries and linear response. Phys. Rep. **88**, 207–319 (1982)
29. Harada, T., Sasa, S.-Y.: Equality connecting energy dissipation with violation of fluctuation-response relation. Phys. Rev. Lett. **95**, 130602 (2005)
30. Hatano, T., Sasa, S.: Steady-state thermodynamics of Langevin systems. Phys. Rev. Lett. **86**, 3463 (2001)
31. Henkel, M., Pleimling, M.: Non-Equilibrium Phase Transitions, vol. 2. Dynamical Scaling far from Equilibrium. Canopus Academic (2009)
32. Kubo, R.: The fluctuation-dissipation theorem. Rep. Prog. Phys. **29**, 255–284 (1966)
33. Kurchan, J.: In and out of equilibrium. Nature **433**, 222 (2005)
34. Lecomte, V., Appert-Rolland, C., van Wijland, F.: Thermodynamic formalism for systems with Markov dynamics. J. Stat. Phys. **127**, 51–106 (2007)
35. Lippiello, E., Corberi, F., Sarracino, A., Zannetti, M.: Nonlinear response and fluctuation dissipation relations. Phys. Rev. E **78**, 041120 (2008)
36. Lippiello, E., Corberi, F., Zannetti, M.: Off-equilibrium generalization of the fluctuation dissipation theorem for Ising spins and measurement of the linear response function. Phys. Rev. E **71**, 036104 (2005)
37. Maes, C., Netočný, K.: The canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. Europhys. Lett. **82**, 30003 (2008)
38. Maes, C., Netočný, K., Wynants, B.: Steady state statistics of driven diffusions. Physica A **387**, 2675–2689 (2008)
39. Maes, C., Netočný, K., Wynants, B.: On and beyond entropy production; the case of Markov jump processes. Markov Proc. Relat. Fields **14**, 445–464 (2008)
40. Maes, C., Netočný, K., Wynants, B.: On a response function and its interpretation. Markov Proc. Relat. Fields (2009)
41. Maes, C., van Wieren, M.H.: Time-symmetric fluctuations in nonequilibrium systems. Phys. Rev. Lett. **96**, 240601 (2006)
42. Marini Bettolo Marconi, U., Puglisi, A., Rondoni, L., Vulpiani, A.: Fluctuation-dissipation: response theory in statistical physics. Phys. Rep. **461**, 111–195 (2008)
43. Martens, K., Bertin, E., Droz, M.: Observable dependence of the fluctuation-dissipation temperature. arXiv:0906.3482
44. Martin, P.A.: Physique statistique des processus irreversibles. In: Lecture Notes of the DEA de Physique Théorique, Notes by F. Coppex, ENS Lyon, Fall 2001–Spring 2004
45. Nakamura, T., Sasa, S.: A fluctuation-response relation of many Brownian particles under non-equilibrium conditions. Phys. Rev. E **77**, 021108 (2008)
46. Oono, Y., Paniconi, M.: Steady state thermodynamics. Prog. Theor. Phys. Suppl. **130**, 29 (1998)
47. Ruelle, D.: A review of linear response theory for general differentiable dynamical systems. arXiv:0901.0484
48. Ricci-Tersenghi, F.: Measuring the fluctuation-dissipation ratio in glassy systems with no perturbing field. Phys. Rev. E **68**, 065104(R) (2003)
49. Speck, T., Seifert, U.: Restoring a fluctuation-dissipation theorem in a nonequilibrium steady state. Europhys. Lett. **74**, 391–396 (2006)
50. Speck, T., Seifert, U.: Extended fluctuation-dissipation theorem for soft matter in stationary flow. Phys. Rev. E **79**, 040102 (2009)
51. Weidlich, W.: Fluctuation-dissipation theorem for a class of stationary open systems. Z. Phys. **248**, 234–243 (1971)