On a response-formula and its interpretation

Christian Maes

Instituut voor Theoretische Fysica, K.U.Leuven

Abstract: We present a systematic and physically inspired generalization of the equilibrium response formula, also called fluctuation-dissipation theorem, to Markov processes possibly describing interacting particle systems out-of-equilibrium.

The subject covered here originates from discussions with Marco Baiesi, Wojciech De Roeck, Karel Netočný and Bram Wynants. This text is the written-out version of a talk given at the ETH Zürich on 10 June 2009, at the conference on *Open systems: Non-Equilibrium Phenomena–Dissipation, Decoherence, Transport*. I am grateful to Profs. J. Fröhlich and G.M. Graf for their kind invitation and hospitality. The first announcement of the results was made in the paper [1] by M.Baiesi, C.Maes and B.Wynants, soon to appear. A more elaborate exposition with many more references will also be finished soon.

1. Response in equilibrium

Relations between fluctuations, response behavior and dissipation in equilibrium systems have been obtained and applied throughout the development of statistical mechanics in the 20th century, [7]. Let us consider a simple situation, which falls in the context of the present discussion. Take an Ising spin system on a finite graph $(\Lambda, \sim)$; at each vertex $i \in \Lambda$ there is a spin $\sigma_i = \pm 1$. We consider a spin flip Markov dynamics on the configurations $\sigma \in \{+1, -1\}^\Lambda \equiv K$ with possible transitions $\sigma \rightarrow \sigma^j$ where $\sigma^j_i = \sigma_i$ for $j \neq i$ and $\sigma^j_j = -\sigma_j$ is the new configuration spin flipped at vertex j. Physically, we imagine that there is a thermal reservoir perhaps in the form of lattice vibrations or of electronic degrees of freedom attached to the system so that for each transition $\sigma \rightarrow \sigma^j$ there is an energy exchange $U(\sigma^j) - U(\sigma)$ with and an entropy flux $[U(\sigma) - U(\sigma^j)]/T$ in the reservoir at equilibrium temperature T. There is no need here to specify that energy function $U(\sigma)$. The transition rates for $\sigma \rightarrow \sigma^j$ are chosen à la Glauber

$$W(\sigma, \sigma^j) = \psi(\sigma, j)\, \exp -\frac{\beta}{2}[U(\sigma^j) - U(\sigma)] \tag{1}$$

where the prefactor $\psi(\sigma, j) = \psi(\sigma^j, j)$ does not depend on σ_j.
That defines a purely dissipative relaxational dynamics. There is a reversible stationary distribution ρ giving probabilities

$$\rho(\sigma) = \frac{1}{Z} e^{-\beta U(\sigma)} \tag{2}$$

to the spin configurations. We call it the equilibrium distribution. The reversibility is expressed by the condition of detailed balance

$$W(\sigma, \sigma^j)\, \rho(\sigma) = W(\sigma^j, \sigma)\, \rho(\sigma^j), \quad \text{for all } j \in \Lambda, \sigma \in K$$

but the really important property is the time-reversal symmetry for the stationary process: denoting by P_ρ the stationary Markov (equilibrium) process with fixed time marginals equal to ρ, we have equilibrium correlations

$$\begin{aligned} \langle f(\sigma(s))\, g(\sigma(t))\rangle_\rho &\equiv \int \mathrm{d}P_\rho(\omega)\, f(\omega(s))\, g(\omega(t)) \\ &= \langle f(\sigma(t))\, g(\sigma(s)) \equiv C_{f,g}(s,t) \end{aligned} \tag{3}$$

function of $|t-s|$ and time-reversal invariant, $C_{f,g}(s,t) = C_{f,g}(t,s)$.

Suppose now that we start at equilibrium ρ at time $t = 0$ but thereafter we slightly modify the dynamics in a time-dependent way. For example, for times $t \in [0, \tau]$ during some interval of length τ we switch on a magnetic field of small amplitude h. The transition rates have now become

$$W_t(\sigma, \sigma^j) = W(\sigma, \sigma^j)\, e^{-\beta\, h(t)\, \sigma_j}, \quad t > 0$$

where we allow for a general time-dependence $h(t)$. That is the external stimulus by which we change the energy function U in (1) into $U - h(t)\,V$ for $V(\sigma) = \sum_{i\in\Lambda} \sigma_i$. How will the equilibrium system respond? The easiest case is when the $h(t)$ is small and we look at the linear response

$$\langle Q(t)\rangle_\rho^h = \langle Q(t)\rangle_\rho^{\text{eq}} + \int_0^t \mathrm{d}s\, h(s) R_{QV}^{\text{eq}}(t,s) + o(h)$$

Here, $Q(t) = Q(\sigma(t))$ is a function of the random spin configuration evaluated at time $t > 0$. The left-hand side averages over the perturbed dynamics, depending on h, and over the initial equilibrium ρ; the right-hand side averages over the unperturbed dynamics always starting in ρ: $\langle Q(t)\rangle_\rho^{\text{eq}} = \sum_\sigma \rho(\sigma)\, Q(\sigma)$ as the equilibrium is time-invariant. The linear correction contains the response function or generalized susceptibility $R_{QV}^{\text{eq}}(t,s)$ which is our object of study. Formally and leaving away further decorations,

$$R_{QV}(t,s) = \frac{\delta}{\delta h(s)} \langle Q(t)\rangle^h (h=0)$$

An interesting case would be to look at the response in the magnetization itself, taking $Q = V = \sum \sigma(i)$; let us write $R_{QV}^{\text{eq}}(t,s) = \chi^{\text{eq}}(t-s)$ for the response function then. It turns out, as we will see under more general circumstances below that then

$$\chi^{\text{eq}}(t-s) = \beta \sum_{i,j\in\Lambda} \frac{\mathrm{d}}{\mathrm{d}s} \langle \sigma_i(s)\, \sigma_j(t)\rangle_\rho^{\text{eq}}, \quad 0 < s < t$$

is expressible as a time-correlation function in the equilibrium process. That formula is valid for all times $0 < s < t$ and we can see what it implies for the shift in magnetization when we would take $h(t) = h, t \in [0,\tau]$: to first order in h and for $t \geq \tau$

$$\sum_{i\in\Lambda} \left[\langle \sigma_i(t)\rangle^h - \langle \sigma_i(0)\rangle_\rho^{\text{eq}}\right] = h\,\beta \sum_{i,j\in\Lambda} \langle [\sigma_i(\tau) - \sigma_i(0)]\, \sigma_j(t)\rangle_\rho^{\text{eq}} \qquad (4)$$

All that is an example of the fluctuation-dissipation theorem for finite-time perturbations. The more general equilibrium formula of which (4) is a special case reads

$$R_{QV}^{\text{eq}}(t,s) = \beta \frac{\mathrm{d}}{\mathrm{d}s} \langle V(s) Q(t)\rangle_\rho^{\text{eq}}, \quad 0 < s < t \qquad (5)$$

A proof of this is easy by applying first-order time-dependent perturbation theory and by inserting the equilibrium condition (2).

2. What in nonequilibrium ?

Take a Markov stochastic dynamics for a finite system. Denote the state space by K. We have transition rates $W(x,y), x,y \in K$. We do no longer assume that there is a potential, i.e. a function $E(x), x \in K$ for which $W(x,y)\exp -E(x) = W(y,x)\exp -E(y)$. In particular, for a stationary distribution $\rho(x), x \in K$, while

$$\sum_{y\in K}[\rho(x)\,W(x,y) - \rho(y)\,W(y,x)] = 0, \quad x \in K$$

still, there are nonzero currents of the form $\rho(x)\,W(x,y) - \rho(y)\,W(y,x) \neq 0$ for some pairs $x \neq y \in K$. The stationary process (Markov dynamics in ρ) is then no longer time-reversible. We have in mind systems of stochastically interacting particles which are driven away from equilibrium; the state x is then the total configuration of particles and the transitions are local.
Secondly we also do not assume that we start at time $t = 0$ from a stationary distribution. Rather, we have an arbitrary probability distribution $\mu(x), x \in K$, from which the initial data are drawn and then for $t > 0$ we apply the perturbed dynamics. That perturbed dynamics uses transition rates

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

for some potential V with small amplitudes h. The inverse temperature β signals that the perturbation concerns an additional energy exchange with a reservoir at temperature β^{-1}.
At time $t > 0$ the expected value of an observable Q will most probably deviate both from the expectation under the unperturbed dynamics and from the stationary expectation:

$$\begin{aligned} \langle Q(t)\rangle_\mu^h &\neq \langle Q(t)\rangle_\mu \\ &\neq \langle Q(t)\rangle_\rho = \sum_{x\in K}\rho(x)Q(x) \end{aligned} \qquad (6)$$

Note that we abbreviate $Q(t) = Q(x(t))$. The right-hand sides concern the unperturbed dynamics, the upper one starting from μ and, on the next line, when starting from the stationary ρ. Linear response theory out-of-equilibrium is interested in estimating and interpreting the deviations

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

to first order in h.

Here we start by letting h small and then we could take other limits like $t \uparrow \infty$. That is not the most interesting or unique physical set-up. Moreover, also other types of perturbations than via a potential become very interesting when away from equilibrium. Yet, this paper deals with the simplest non-technical case as described above. There are various extensions, such as treating diffusion processes including inertial dynamics or even non-Markovian dynamics or driven systems without a uniquely defined reservoir temperature which can be treated via the very same set-up as the present one, but they will not be discussed except for Section 8.

3. Response formula

Keeping in mind the previous set-up there is a simple formula for the response. Let L be the backward generator, acting on observables: for $s > t$,

$$\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)\,Q(t)\rangle_\mu = \langle LV(s)\,Q(t)\rangle_\mu\,, \qquad s > t > 0 \tag{7}$$

For our case of Markov jump processes, we have simply

$$Lf(x) = \sum_{y\in K} W(x,y)[f(y) - f(x)]$$

In case of equilibrium, formula (5) is equivalent with

$$R^{\mathrm{eq}}_{QV}(t,s) = -\beta\langle (LV)(s)\,Q(t)\rangle^{\mathrm{eq}}_\rho \tag{8}$$

Then, our generalized response formula [1] also valid in nonequilibrium is

$$R^{\mu}_{QV}(t,s) = \frac{\beta}{2}\frac{d}{ds}\langle V(s)Q(t)\rangle_\mu - \frac{\beta}{2}\langle LV(s)\,Q(t)\rangle_\mu \tag{9}$$

Remark that in the first term it does not make sense to take the time-derivative inside the expectation; the random paths are piecewise constant. On the other hand, for $s > t$ (7) applies and therefore, causality in the sense that the response should vanish for $s > t$ is automatically verified in (9).
Observe that (9) shares its first term with (5) except for the factor $1/2$ and we start from an arbitrary μ in (9); in general we have no explicit formula for the stationary ρ.
Remark also that the formula (9) as such keeps making sense for a much broader class of dynamics, including diffusion processes or infinite volume dynamics when V and Q are local observables. We will briefly comment on these generalizations in Section 8.

We explain the meaning of the two terms on the right-hand side of (9).
The first term in (9) (as in (5)) is a correlation with the excess entropy production. To see it, we write its contribution to the deviation

$\langle Q(t)\rangle_\mu^h - \langle Q(t)\rangle_\mu$ in the time-integral

$$\beta \int_0^t \mathrm{d}s\, h(s)\, \frac{\mathrm{d}}{\mathrm{d}s}\, \langle V(s) Q(t)\rangle_\mu$$
$$= \beta\langle\{h(t)V(t) - h(0)V(0) - \int_0^t \mathrm{d}s\, \dot{h}(s)\, V(s)\}\, Q(t)\rangle_\mu \qquad (10)$$

which is the correlation of $Q(t)$ in the unperturbed process with the entropy flux

$$\beta\{h(t)V(t) - h(0)V(0) - \int_0^t \mathrm{d}s\, \frac{\mathrm{d}}{\mathrm{d}s} h(s)\; V(s)\} \qquad (11)$$

Indeed we recognize the change of energy $h(t)V(t) - h(0)V(0)$ in the environment minus the work done on the system $\int_0^t \mathrm{d}s\, \dot{h}(s)\, V(s)$. Of course, the system may already show a steady or transient entropy production; in the response formula only the excess appears caused by the time-dependent perturbation.

Secondly, in (9) there is also the correlation between $Q(t)$ and the function $\beta\, LV$ at time s. For its interpretation we must look at the escape rates, i.e. the frequencies at which the Markov jump process leaves a state x. The escape rate at $x \in K$ gets changed by the perturbation. Its excess is

$$\sum_y W(x,y)\left\{e^{\frac{\beta\, h}{2}[V(y)-V(x)]} - 1\right\}$$
$$\simeq \frac{\beta\, h}{2} \sum_y W(x,y)[V(y) - V(x)] \;\; = \;\; \frac{\beta\, h}{2}(LV)(x) \qquad (12)$$

to linear order in h. We call this the ***frenesy***, derived from the adjective frenetic or frantic. The latter refers to the nervosity or dynamical activity in the system. In contrast to the entropy production which has a preferred direction in time, frenesy is a time-symmetric quantity much like traffic is, while current is not.

The mathematical proof of formula (9) is elementary for finite state space Markov processes. The expectation value in the perturbed process can be related to the unperturbed one via a Girsanov formula over the time-interval $[0, t]$:

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

for logarithmic density

$$\log \frac{\mathrm{d}P_\mu^h}{\mathrm{d}P_\mu}(\omega) = \frac{\beta}{2}\sum_s h(s)\,\big[V(\omega(s)) - V(\omega(s^-))\big] - \sum_y \int_0^t \mathrm{d}s\, W(\omega(s),y)\big[e^{\frac{\beta h(s)}{2}(V(y)-V(\omega(s)))} - 1\big] \quad (13)$$

where the first sum is over the jump times $s \in [0,t]$ in ω with $\omega(s^-)$ the state before the jump. On the other hand, since the path is constant between the jump times $(t_0 = 0, t_1, \ldots, t_n, t_{n+1} = t)$ in ω

$$\int_0^t \mathrm{d}s\,\frac{\mathrm{d}}{\mathrm{d}s}h(s)\;V(\omega(s)) = \sum_{k=0}^n V(\omega(t_k))\,\big[h(t_{k+1}) - h(t_k)\big]$$
$$= h(t)V(\omega(t)) - h(0)V(\omega(0)) + \sum_s h(s)\,\big[V(\omega(s^-)) - V(\omega(s))\big]$$

by partial summation. We can therefore substitute (11) in the first line of (13). The rest is expansion to first order in h for a finite number of terms, with probability one.

Let us finally repeat (8) and see how formula (9) reconstructs the equilibrium formula (5). For that we must take $\mu = \rho$ the equilibrium distribution and apply time-reversal symmetry so that with $s < t$

$$\langle LV(s)\,Q(t)\rangle_\rho^{\text{eq}} = \langle LV(t)\,Q(s)\rangle_\rho^{\text{eq}}$$
$$= \frac{\mathrm{d}}{\mathrm{d}t}\langle V(t)\,Q(s)\rangle_\rho^{\text{eq}} = -\frac{\mathrm{d}}{\mathrm{d}s}\langle V(t)\,Q(s)\rangle_\rho^{\text{eq}}$$
$$= -\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)\,Q(t)\rangle_\rho^{\text{eq}}$$

reconstructing the first term in (9). That is how (9) leads to (5) when there is a stationary time-reversal symmetry (= equilibrium). In other words, in equilibrium, the two terms on the right-hand side of (9) add up to give (5); the frenesy-correlation then equals minus the correlation with the entropy flux.

4. Example

We come back to the example (1) in the beginning of a Glauber spin-flip dynamics but we add a mixing dynamics. More specifically, we not only have transitions $\sigma \to \sigma^j$ with corresponding rates $W(\sigma,\sigma^j)$, but we also allow now transitions $\sigma \to \sigma^{ij}$ where the spins at neighboring vertices $i \sim j \in \Lambda$ get exchanged: $\sigma_k^{ij} = \sigma_k$, if $i \neq k \neq j$ while $\sigma_i^{ij} = \sigma_j, \sigma_j^{ij} = \sigma_i$. The rate for these exchanges is $\lambda > 0$. We now have a reaction-diffusion process on $K = \{+1,-1\}^\Lambda$ with generator L

acting on $f : K \to \mathbb{R}$,

$$Lf(\sigma) = \sum_{j\in\Lambda} W(\sigma,\sigma^j)[f(\sigma^j) - f(\sigma)] + \lambda \sum_{i\sim j}[f(\sigma^{ij}) - f(\sigma)]$$

That unperturbed dynamics does not satisfy the condition of detailed balance when $\beta \neq 0$ for a nontrivial energy function $U(\sigma)$. There is a stationary distribution ρ of which very little is known.
We still consider the magnetization $V(\sigma) = Q(\sigma) = \sum_i \sigma_i$ for organizing and evaluating the perturbation of amplitude $h(t), t > 0$. Note that $V(\sigma^{ij}) - V(\sigma) = 0$ and the transition $\sigma \to \sigma^{ij}$ leaves the total magnetization unchanged. Hence, for the frenesy, $LV(\sigma) = -2\sum_i \sigma_i W(\sigma,\sigma^i)$. Let us abbreviate $W(\sigma,\sigma^i) = c(i,\sigma)$. We thus get the response around steady nonequilibrium

$$\frac{\delta}{\delta h_j(s)}\langle\sigma_j(t)\rangle_\rho^h \qquad (h=0) = \frac{\beta}{2}\frac{\mathrm{d}}{\mathrm{d}s}\langle\sigma_i(s)\,\sigma_j(t)\rangle_\rho + \beta\langle\sigma_i(s)c(i,\sigma(s))\,\sigma_j(t)\rangle_\rho$$

For a constant perturbation $h(s) = h, s \in [0,t]$ as in (4), we can integrate over $s \in [0,t]$ to get the leading order of the response:

$$\begin{aligned}\frac{1}{h}\sum_i\langle\sigma_i(t)\rangle^h - \langle\sigma_i(0)\rangle_\rho \quad &= \quad \frac{\beta}{2}\sum_{i,j\in\Lambda}\langle[\sigma_i(t) - \sigma_i(0)]\,\sigma_j(t)\rangle_\rho \\ &+ \quad \frac{\beta}{2}\sum_{i,j\in\Lambda}\int_0^t \mathrm{d}s\langle\sigma_i(0)c(i,\sigma(0))\,\sigma_j(s)\rangle_\rho\end{aligned} \tag{14}$$

Note that the rate λ is hiding in the correlation functions but the form (14) is unchanged no matter what is λ.

5. Relation with the notion of effective temperature

After having observed a violation of the equilibrium fluctuation–dissipation relation for nonequilibrium regimes, people have asked whether and have sometimes verified that there is an effective temperature in the sense that still

$$R^\mu_{QV}(t,s) = \frac{1}{T^{\text{eff}}}\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_\mu \tag{15}$$

to resemble the equilibrium formula (5) but with a prefactor in terms of what perhaps resembles a thermodynamic temperature-like quantity for some classes of observables and over some scales of times $(s/t, s)$, [4]. A simple scenario takes our magnetic system of (1) but starts with a highly disordered distribution μ. That μ could be the high-temperature distribution of the Ising model on Λ. The unperturbed

dynamics defined by (1) is however taken at a different bath temperature $(k_B\,\beta)^{-1} = T$, mostly at lower temperatures than initially. For example, μ could be a product measure in which each spin is independently ± 1 with probability $1/2$, while the environment temperature T to which the system is exposed from $t > 0$ would be much lower. Depending on the graph and on the energy function U, the system can exhibit metastable or long lived transient behavior which shows as (15). 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 such calculations are in [2].
In fact, now we can write the ratio $T/T^{\text{eff}} = X$ in terms of correlation functions

$$X = X_{QV}(\mu;t,s) = \frac{1}{2}\big[1 - \frac{\langle LV(s)\,Q(t)\rangle_\mu}{\partial_s\langle V(s)Q(t)\rangle_\mu}\big] \tag{16}$$

In turn, the verification of the existence of an effective temperature can now proceed by investigating 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 $Y = 0$. X and the effective temperature T^{eff} get negative when the frenetic term overwhelms the entropic contribution.

6. Relation with co-moving frame interpretation

It is easy to find a relation with the interpretation in [3].
We take our formula (9) for stationary nonequilibrium, taking $\mu = \rho$, and it is immediate to rewrite it as

$$R_{QV}(t,s) = \beta\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_\rho - \frac{\beta}{2}\,\big\langle\big(L - L^*\big)V(s)\,Q(t)\big\rangle_\rho \tag{17}$$

in terms of the adjoint generator L^* for which $\sum_x \rho(x)(Lf)(x)\,g(x) = \sum_x \rho(x)\,f(x)\,(L^*g)(x)$. In fact,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\langle V(s)Q(t)\rangle_\rho &= -\frac{\mathrm{d}}{\mathrm{d}t}\langle V(s)Q(t)\rangle_\rho \\ &= -\langle (L^*V)(s)\,Q(t)\rangle_\rho \end{aligned} \tag{18}$$

so that also

$$R_{QV}(t,s) = -\beta\langle(\frac{L+L^*}{2}V)(s)\,Q(t)\rangle_\rho \tag{19}$$

From (17) we see that when the perturbation V is time-direction independent in the precise sense that $LV = L^*V$, then the nonequilibrium response (17) (= (9) for $\mu = \rho$) reduces to the equilibrium formula (5).

Let us look more into (17) and rewrite

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

For overdamped diffusion processes, which is the context of [3], we would get $L - L^* = 2u\cdot\nabla$ for the local velocity $u(x) = \frac{j_\rho}{\rho}(x)$ with respect to the probability current j_ρ. In other words, the linear response formula for stationary nonequilibrium (17) can be interpreted as modifying the time-derivative $\mathrm{d}/\mathrm{d}s$ of the equilibrium formula (5) into $\mathrm{d}/\mathrm{d}s - u\cdot\nabla$ which is a co-moving derivative. The more precise version of the above is formula (19) when compared with (8). We see that the equilibrium formula (8) is simply reproduced when, in nonequilibrium, we use the symmetric part of the generator $(L + L^*)/2$. The change to a co-moving frame is like subtracting the antisymmetric part of the generator: the passage to the Lagrangian frame of local velocity removes the non-conservative forcing.
While that offers an interesting interpretation it is not clear how useful that rewriting can be for spatial processes where the probability current has little relation with the real physical currents.
For example, knowing that there is detailed balance in a co-moving frame for energy function $-\beta^{-1}\ln\rho$ and with time-dependent coefficients offers little information. The point is that we do not know L^* and we do not know ρ for general nonequilibrium systems; yet they appear explicitly in (17) or in (19). In contrast, for our formula (9), one has an explicit expression in terms of known observables and ρ only enters the statistical averaging. For the purpose of numerical simulation or experiment, no explicit knowledge of ρ is needed for verifying (9) of for finding (16).

7. More on frenesy

We have been interpreting the second term in (9) in terms of the nervosity, or what we have called the frenesy or the activity of the dynamics. There would be little reason for it if that quantity had not a larger role. Frenesy/Dynamical activity also appears in dynamical fluctuation theory, [8, 10]. To start and always in the context of finite state space Markov processes $x(t)$ we look at

$$p_\tau = \frac{1}{\tau}\int_0^\tau \delta_{x(t),\cdot}\,\mathrm{d}t, \quad \delta_{a,b} = 0 \text{ if } a \neq b, = 1 \text{ if } a = b$$

in the stationary process P_ρ. That p_τ defines the empirical distribution of occupation times: $p_\tau(x)$ is the fraction of time the system spends in state x in the time $[0,\tau]$ while initially drawn from the stationary ρ. Clearly, p_τ is invariant under time-reversal and $p_\tau \rightarrow \rho, \tau \uparrow +\infty$ when P_ρ is an ergodic Markov process. The fluctuations around that law of large times has been the origin of much pleasure in the theory of large deviations. Our context is again the simplest and it is well known since the pioneering work in [6], that the steady-state probability that the occupation statistics is given by the law μ obeys a fluctuation formula

$$\boldsymbol{P}_\rho[p_\tau \simeq \mu] \simeq e^{-\tau I(\mu)}, \quad \tau \uparrow +\infty$$

in the usual logarithmic sense, see e.g. [5, 6]. In the exponent sits the functional $I(\mu) \geq 0$, with equality only for $I(\rho) = 0$. There is a well-known variational expression for the fluctuation functional $I(\mu)$,

$$I(\mu) = \sup_{g>0} - \sum_x \mu(x) \, \frac{Lg}{g}(x) \tag{20}$$

where the supremum is over all positive functions $g : K \rightarrow \mathbb{R}$. (When the dynamics is reversible and ρ is equilibrium, then we know more: $I^{\text{eq}}(\mu) = -\sum_x \rho(x) \sqrt{f(x)} \, (L\sqrt{f})(x)$ for $f(x) = \mu(x)/\rho(x)$ when the latter makes sense.)

Let us parameterize $g(x) = \exp \beta V(x)/2$ in (20) so that

$$\sum_x \mu(x) \, \frac{Lg}{g}(x) = \sum_x \mu(x) \, \sum_y W(x,y)\{e^{\frac{\beta}{2}[V(y)-V(x)]} - 1\}$$

and we recognize the expected excess in escape rates when we add a potential to the rates, see (12):

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

The potential $V = v$ that reaches (or approximates) the supremum has an interpretation. It is that potential v which makes μ (approximately) stationary; i.e., by changing $W(x,y) \rightarrow W(x,y) \, \exp \beta(v(y) - v(x)/2$. Thus, with $\tilde{L}$ the generator for that modified process making μ stationary:

$$\begin{aligned} I(\mu) &= \sum_y \mu(x) \, W(x,y) - \sum_y \mu(x) \, W(x,y) \, e^{\beta(v(y)-v(x)/2} \\ &\simeq -\beta \sum_x \mu(x) \, Lv(x) \end{aligned}$$

where the second line (expected frenesy in μ) is the first term in an expansion around $\mu - \rho$; second order in the distance from ρ as v is then small as well.

For further information on the frenesy functional, we refer to [10]. Let us just add that we can of course also consider the change in escape rates by adding a general antisymmetric function $F(x,y)$ to the rates

$$W(x,y) \longrightarrow W(x,y)\, e^{\frac{1}{2}F(x,y)}, \qquad F(x,y) = -F(y,x)$$

Then, the expected new escape rate

$$H(\mu,F) = \sum_{x,y} \mu(x)\, W(x,y)\, e^{\frac{1}{2}F(x,y)}$$

is a potential function for the (expected transient) currents in the sense

$$\delta H(\mu,F) = \frac{1}{4}\sum_{x,y} j_{\mu,F}(x,y)\delta F(x,y)$$

for $j_{\mu,F}(x,y) = \sum_{x,y}\big[\mu(x)W(x,y)\, e^{\frac{1}{2}F(x,y)} - \mu(y)W(y,x)\, e^{-\frac{1}{2}F(x,y)}\big]$.

8. Extensions

The above analysis applies unchanged to overdamped diffusion. By this we mean $d-$dimensional processes defined in the Itô-sense by

$$\mathrm{d}x_t = \big\{\chi(x_t)\big[F(x_t) - \nabla U(x_t)\big] + \nabla\cdot D(x_t)\big\}\,\mathrm{d}t + \sqrt{2D(x_t)}\,\mathrm{d}B_t \quad (21)$$

where the $d-$dimensional vector $\mathrm{d}B_t$ has independent standard Gaussian white noise components. As the thermal reservoir is assumed to be in equilibrium the mobility $\chi(x) = \beta D(x)$ equals the diffusion matrix up to the inverse temperature $\beta > 0$; they are strictly positive (symmetric) $d \times d-$matrices. There is a potential U and the force F represents the nonequilibrium driving. We do not specify here regularity properties or boundary conditions. The unperturbed generator is here

$$L = \chi(F - \nabla U)\cdot\nabla + \nabla D\cdot\nabla$$

Equilibrium (5) is regained when $F = 0$, for $\rho \propto \exp -\beta U$. At $t \geq 0$, there is a time-dependent perturbation changing U in (21) into $U - h(t)\,V$. The resulting response relation is exactly the same as in (9). Its dynamical fluctuation theory is treated in [9].

The diffusions (21) are *overdamped* because they have forces proportional to velocities. These processes are high damping limits (sometimes called, Smoluchowski limits) of underdamped or inertial stochastic dynamics. We will not give many details here as their treatment requires the introduction of more notation and of more technicalities. The important thing though is that the interpretation of the response relation gets unchanged. We just give one example to highlight some points:

Example: **Brownian motion under nonequilibrium driving**:
We consider a particle with momentum and position $(q, p) \in \mathbb{R}^{2d}$ undergoing

$$\mathrm{d}q = p\,\mathrm{d}t$$

$$\mathrm{d}p = -\nabla U(q)\,\mathrm{d}t + F(q)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + h(t)E\,\mathrm{d}t + \sqrt{2\gamma/\beta}\,\mathrm{d}B(t)$$

where γ is the friction. The potential U is assumed sufficiently confining the particle's position and allowing a stationary distribution ρ. The dynamics is nonequilibrium because of the presence of a non-gradient F. The perturbation is in the form of a constant external field E. We look at the excesses, first in entropy flux: for a path ω during $[0, \tau]$

$$S(\omega) = \beta \int_0^\tau \mathrm{d}t\, E \cdot p(t)\, h(t) \tag{22}$$

is the excess entropy flux by the field E. Secondly, the path-dependent frenesy (in linear order in h) gives

$$\mathcal{T}(\omega) = -\frac{\beta}{\gamma} \int_0^\tau h(t)\, \left[\mathrm{d}t\big(\nabla U - F\big) + \mathrm{d}p(t)\right] \cdot E \tag{23}$$

which is time-symmetric under reversing the paths (together with flipping the momentum).
The response relation has the same form as (9) but where $\beta L V$ is replaced by the frenesy (23), yielding the nonequilibrium mobility

$$\frac{\delta}{\delta h(s)} \langle p(t)\rangle^h \,(h = 0) = \frac{\beta}{2} \langle E \cdot p(s)\, p(t)\rangle_\rho$$
$$+\frac{\beta}{2\gamma} \langle E \cdot [\nabla U(q(s)) - F(q(s))]\; p(t)\rangle_\rho +$$
$$\frac{\beta}{2\gamma}\frac{\mathrm{d}}{\mathrm{d}s} \langle E \cdot p(s)\, p(t)\rangle_\rho$$

The last term vanishes in the overdamped case.
Further extensions could perhaps ask to reduce the noise level or to demand less regularity of the forces. These are mostly technical issues, but we should remember that driven nonequilibrium systems are open systems and the reduced description of their dynamics, after integrating out the bath degrees of freedom will always show some amount of stochasticity. Whether the reduced dynamics is Markovian, is more subtle and has to do much more with the *right choice* of variables. However, the notions of entropy flux and of frenesy should remain available in principle and hence, also the generalized response formula.

References

[1] M. Baiesi, C. Maes and B. Wynants: Fluctuations and response of nonequilibrium states, arXiv:0902.3955v3, to appear in *Phys. Rev. Lett.* 2009.

[2] C. Chatelain: A far-from-equilibrium fluctuation-dissipation relation for an Ising-Glauber-like model, *J. Phys. A* **36**, 10739 (2003).

[3] R.Chetrite, G.Falkovich and K.Gawedzki: Fluctuation relations in simple examples of non-equilibrium steady states: *J. Stat. Mech.* P08005 (2008).

[4] A. Crisanti and J.Ritort: Violation of the fluctuation-dissipation theorem in glassy systems : basic notions and the numerical evidence, *J. Phys. A : Math. Gen.* **36**, R181–R290 (2003)

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

[6] M. D. Donsker and S. R. Varadhan: Asymptotic evaluation of certain Markov process expectations for large time, I., *Comm. Pure Appl. Math.*, **28**:1–47 (1975).

[7] R.Kubo: The fluctuation-dissipation theorem, *Rep. Prog. Phys.* **29**, 255–284 (1966).

[8] C.Maes and K. Netočný: The canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states, *Europhys. Lett.* **82**, 30003 (2008).

[9] C. Maes, K. Netočný, and B. Wynants, Steady state statistics of driven diffusions, *Physica A* **387**, 2675–2689 (2008).

[10] C. Maes, K. Netočný and B. Wynants: On and beyond entropy production; the case of Markov jump processes, *Markov Proc. Rel. Fields* **14**, 445–464 (2008).

http://itf.fys.kuleuven.be/~christ/