

Fluctuations and Response of Nonequilibrium States

Marco Baiesi, Christian Maes, and Bram Wynants*

Instituut voor Theoretische Fysica, Katholieke Universiteit Leuven, 3001, Belgium

(Received 23 February 2009; revised manuscript received 3 June 2009; published 2 July 2009)

A generalized fluctuation-response relation is found for thermal systems driven out of equilibrium. Its derivation is independent of many details of the dynamics, which is only required to be first order. The result gives a correction to the equilibrium fluctuation-dissipation theorem, in terms of the correlation between observable and excess in dynamical activity caused by the perturbation. Previous approaches to this problem are recovered and extended in a unifying scheme.

DOI: 10.1103/PhysRevLett.103.010602

PACS numbers: 05.70.Ln, 05.20.-y

The fluctuation-dissipation theorem is a standard chapter in statistical mechanics [1–3]. A system in thermal equilibrium has statistical fluctuations proportional to its response to external perturbations: a small impulse changing the potential $U \rightarrow U - h_s V$ at time s will produce a response $R_{QV}^{\text{eq}}(t, s) = \delta \langle Q(t) \rangle / \delta h_s$ in a quantity Q at time $t \geq s$ of the form

$$R_{QV}^{\text{eq}}(t, s) = \beta \frac{d}{ds} \langle V(s)Q(t) \rangle_{\text{eq}}, \quad (1)$$

where $\langle V(s)Q(t) \rangle_{\text{eq}}$ is a correlation function quantifying the equilibrium fluctuations in absence of any perturbation, and the proportionality constant $\beta = 1/k_B T$ is the inverse temperature. An early example of this theorem is present in Einstein's treatment of Brownian motion, where the diffusion constant, expressed as a velocity autocorrelation function, is found proportional to the mobility. Other famous examples include the Johnson-Nyquist formula for electronic white noise and the Onsager reciprocity for linear response coefficients.

So far, approaches deriving a fluctuation-dissipation relation (FDR) for nonequilibrium [4–13] have not found a physical unification and do not appear as textbook material. One reason may be that previous work has not been seen to identify a sufficiently general structure with a clear corresponding statistical thermodynamic interpretation. Today, such an interpretation has become available from advances in dynamical fluctuation theory for nonequilibrium systems.

Aiming to provide a simple and general approach to FDRs, in this Letter we put forward a FDR for nonequilibrium regimes in a framework that may represent a unifying scheme for previous formulations. Our main result can be found in a general formula, Eq. (6) below, which can also sometimes be rewritten as (7) or (11).

In order to go beyond equilibrium and beyond formal perturbation theory, it is important to recognize in the right-hand side of (1) the role of entropy production, as usual governing close-to-equilibrium considerations. Equation (1) expresses the correlation between the dissi-

pation represented by entropy production (energy change divided by temperature) and the observable $Q(t)$. Here, for perturbations of a nonequilibrium system (that already has a nonvanishing set of flows and hence a nonzero entropy production), we will rather speak of the excess of entropy produced by the perturbation $h_s V$.

The second ingredient that is essential in nonequilibrium is still a less known quantity, called traffic [14] or dynamical activity [15], first introduced in [16]. Perhaps activity has been somewhat overlooked in the past because it plays a truly significant role only beyond linear order around equilibrium [14,17], and it is not part of the picture in standard irreversible thermodynamics. Our present results will in fact further clarify the relevance of dynamical activity. To get a first idea of its meaning, we note that activity often measures the frequency of transitions in a trajectory: this is a time-symmetric aspect of dynamical fluctuations [16] because it does not depend on the direction chosen to span the trajectory. On the other hand, it is well known that entropy production is time antisymmetric [18], changing sign upon reversal of time (fluxes are inverted if time is run backward). Thus, both quantities arise naturally as two complementary pieces of the space-time action, as reminded below.

We consider a system in an environment at inverse temperature β (with $k_B = 1$), where a time-independent nonequilibrium condition can be imposed by external driving fields or by installing mechanical displacements or chemical gradients at the boundaries of the system. Moreover, the system evolves according to a Markovian/first-order dynamics: for any two observables f and g , their correlations at times $s < t$ satisfy $\frac{d}{dt} \langle f(x_t)g(x_s) \rangle = \langle (Lf)(x_t)g(x_s) \rangle$ where L is a linear operator, called a generator, acting on observables, and x_t is the state at time t at some reduced, e.g., mesoscopic level of description: $e^{tL} f(x) = \langle f(x_t) | x_0 = x \rangle$, $t \geq 0$ [in the sequel we abbreviate $f(x_t) = f(t)$]. The generator L is a standard tool: for jump processes it is the matrix with transition rates as off-diagonal elements, while $L = (\text{force}) \cdot \nabla + \Delta/\beta$ for overdamped diffusions. However, it is our aim to leave the generator L as unspecified as possible, to emphasize the

generality of the results, referring, e.g., to [5] for further background.

Our formulation is based on the point of view that the system can be described by a distribution $\mathcal{P}(\omega)$ over its possible trajectories ω (space-time paths), giving the state ω_s of the system at each time $s \leq t$. In general $\mathcal{P}(\omega) = \mathcal{P}(\omega|\omega_0)\mu(\omega_0)$, namely, $\mathcal{P}(\omega)$ is the probability of the initial state ω_0 , i.e., $\mu(\omega_0)$, times the probability of the trajectory $\mathcal{P}(\omega|\omega_0)$ given the starting point. The ensemble averages of a quantity $O(\omega)$ are then $\langle O(\omega) \rangle \equiv \sum_{\omega} \mathcal{P}(\omega)O(\omega)$ where \sum_{ω} is a simple notation including also a continuum of trajectories. Here we are interested in comparing probabilities of trajectories for $h_s \neq 0$ (for $s \geq 0$), $\mathcal{P}^h(\omega) = \mathcal{P}^h(\omega|\omega_0)\mu(\omega_0)$, with those of the unperturbed system. In analogy with the $h = 0$ case, we use the notation $\langle O(\omega) \rangle^h \equiv \sum_{\omega} \mathcal{P}^h(\omega)O(\omega)$ for averages in the perturbed system. We start by writing

$$\mathcal{P}^h(\omega) = e^{-A(\omega)}\mathcal{P}(\omega) \quad (2)$$

and by focusing on the action $A(\omega) = -\ln[\mathcal{P}^h(\omega)/\mathcal{P}(\omega)]$. In this setup we put ourselves in line with the Onsager-Machlup approach but outside equilibrium, possibly beyond quadratic or diffusive approximations and also dealing with jump processes.

We ask what determines the path-space distribution and the space-time local action governing it. To find an answer, it is useful to decompose $A(\omega) = [T(\omega) - S(\omega)]/2$ in terms of its time-antisymmetric component $S(\omega) \equiv A(\theta\omega) - A(\omega)$ and the time-symmetric $T(\omega) \equiv A(\theta\omega) + A(\omega)$, where a time-reversed state $(\theta\omega)_s$ is equal to ω_{t-s} (with reversed momenta when applicable). In the time interval $[0, t]$, in standard physical situations one finds that

$$S(\omega) = \beta \left[h_t V(t) - h_0 V(0) - \int_0^t ds \frac{dh_s}{ds} V(s) \right] \quad (3)$$

is the excess entropy flux from the system to its environment. Excess is always meant in the sense of the perturbed process with respect to the original one. From basic thermodynamics we know that this entropy flux is βQ , where Q is the heat flowing to the thermal environment, that is, minus the change in potential energy of the system minus the work performed on the system. The reason for (3) is the physical condition of local detailed balance, which ensures that the ratio of probabilities of the forward with respect to the backward trajectory is given by the exponential of the entropy flux in the forward trajectory [18].

The time-symmetric part $T(\omega)$ of the action A is the excess in activity, having an essential role in dynamical fluctuation theory for the large deviations of the occupations, and in nonequilibrium studies of phase transitions [14,15]. Being in linear response theory (small h_s), we are actually more concerned with the first order in excess activity $\tau(\omega, s) = \frac{\delta}{\delta h_s} T(\omega)|_{h_s=0}$. The excess in activity quantifies how “frenetic” is the motion in the perturbed process, compared with the unperturbed one. For

example, for a Markov jump process with transition rates $W(x \rightarrow y)$ between states $x \rightarrow y$, the traffic is equal to twice the time-integrated escape rates over a trajectory, $2 \int_0^t ds \sum_y W(\omega_s \rightarrow y)$, and its excess $\tau(s)$ [here $\tau(\omega, s) = \tau(s)$ depends only on the state ω_s] to linear order in V is

$$\begin{aligned} \tau(s) &= 2 \sum_y W(\omega_s \rightarrow y) \{ e^{\beta/2[V(y) - V(\omega_s)]} - 1 \} \\ &\simeq \beta \sum_y W(\omega_s \rightarrow y) [V(y) - V(\omega_s)] \\ &= \beta \frac{dV}{ds}. \end{aligned} \quad (4)$$

The last expression is of course formal, the mathematical meaning being $dV/ds = LV(s)$.

Since $S(\omega)$ is already linear in h_s , we have

$$-A(\omega) = \frac{1}{2}S(\omega) - \frac{1}{2} \int_0^t ds h_s \tau(\omega, s) + O(h_s^2). \quad (5)$$

We are now ready to derive the FDR for a single-time observable $O(\omega) = Q(t)$ at time $t > 0$. Its expectation in the linear response regime is $\langle Q(t) \rangle^h = \langle Q(t) \rangle + \int_0^t ds h_s R_{QV}(t, s)$, which we can rewrite with (2) as $\langle Q(t) e^{-A(\omega)} \rangle - \langle Q(t) \rangle \simeq -\langle Q(t) A(\omega) \rangle = \int_0^t ds h_s R_{QV}(t, s)$. With (3) and (5), it is thus straightforward to see that the response function is equal to

$$R_{QV}(t, s) = \frac{\beta}{2} \frac{d}{ds} \langle V(s) Q(t) \rangle - \frac{1}{2} \langle \tau(\omega, s) Q(t) \rangle, \quad (6)$$

which, as we will see, is a nonequilibrium FDR that generalizes and unifies previous formulations. We stress that the derivation of (6) is independent of many details of the dynamics and hence of its generator L : we went beyond formalities of a perturbative first-order calculation by specifying the origin of the two parts on the right-hand side of (6) in terms of (i) the entropy flux and (ii) the dynamical activity [14–16], in excess due to the time-dependent perturbation.

When $\tau(\omega, s) = \tau(s)$, as in (4) or for overdamped diffusions [19], formula (6) simplifies to

$$R_{QV}(t, s) = \frac{\beta}{2} \frac{d}{ds} \langle V(s) Q(t) \rangle - \frac{\beta}{2} \left\langle \frac{d}{ds} V(s) Q(t) \right\rangle. \quad (7)$$

The right-hand side of (7) is nonzero because of the correlation of $V(s)$ with $Q(t)$ for $t > s$ [20].

When the unperturbed dynamics is in detailed balanced, we can recover (1) as a special case of (7) by using time-reversal symmetry: for $s < t$, $\langle (dV/ds) Q(t) \rangle_{\text{eq}} = \langle (dV/dt) Q(s) \rangle_{\text{eq}} = (d/dt) \langle V(t) Q(s) \rangle_{\text{eq}}$, where the last equality is a consequence of the Markov property of the dynamics. Because of stationarity this is then also equal to $-(d/ds) \langle V(s) Q(t) \rangle_{\text{eq}}$, so that in equilibrium the second term of (7) gives exactly the same contribution as its first term, and (1) is recovered.

As an illustration of the general formula, consider a system with a flow of interacting particles due to the presence of reservoirs at different chemical potentials. We choose $Q = V = \mathcal{N}$, where \mathcal{N} is the particle number. Hence, h_s induces a global shift in chemical potentials, and the linearized excess activity $\tau(s) = \beta d\mathcal{N}/ds = \beta \mathcal{J}(s)$ is proportional to the systematic particle current $\mathcal{J}(s)$ into the system via its boundaries, i.e., to the expected current given the configuration at time s . Numerically it is convenient to test the time-integrated version of (7) for constant perturbation $h_s = h$ for $s \geq 0$,

$$\chi(t)/\beta = [C(t) + D(t)]/2 \quad (8)$$

with susceptibility $\chi(t) = [\langle \mathcal{N}(t) \rangle^h - \langle \mathcal{N} \rangle]/h$, correlation function $C(t) = \langle \mathcal{N}^2 \rangle - \langle \mathcal{N}(0)\mathcal{N}(t) \rangle$, and nonequilibrium term $D(t) = -\int_0^t ds \langle \mathcal{J}(s)\mathcal{N}(t) \rangle$. Again, (8) holds irrespective of the details of interaction or driving.

Here we simulate a boundary driven one-dimensional exclusion process, a simple paradigmatic model for transport. A state x is an array with empty sites $x^i = 0$ and particles $x^i = 1$. Nonequilibrium is imposed at boundary sites $i = 1$ and $i = n$, which are in contact with reservoirs with particle densities d_1 and $d_n = 1 - d_1$, respectively. There is a nearest neighbor attraction with energy $H(x) = -\sum_{i=1}^{n-1} x^i x^{i+1}$ determining the Kawasaki dynamics in the bulk. In this context the systematic current arises from two possible transitions: $x \rightarrow y$, with $y^1 = 1 - x^1$, and $x \rightarrow z$, with $z^n = 1 - x^n$. Hence

$$\begin{aligned} \mathcal{J}(x) = & [\mathcal{N}(y) - \mathcal{N}(x)]W(x \rightarrow y) \\ & + [\mathcal{N}(z) - \mathcal{N}(x)]W(x \rightarrow z), \end{aligned}$$

where, e.g., $W(x \rightarrow y) \propto d_1 e^{-\beta[H(y)-H(x)]/2}$ if a particle enters at $i = 1$. Examples in Fig. 1(a) illustrate Eq. (8), while the magnitude of $C(t = 50)$ and $D(t = 50)$ is shown in Fig. 1(b) as a function of the driving $d_1 - d_n$. Far from equilibrium $C(t) \neq D(t)$, and thus the correlation function

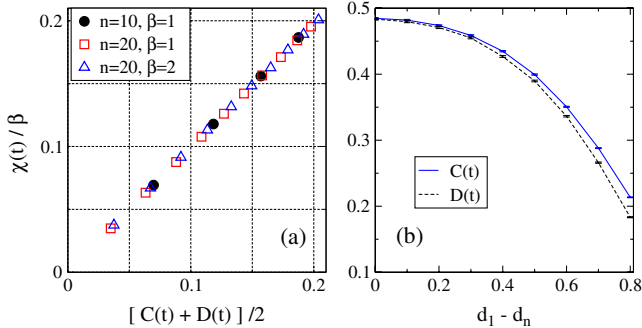


FIG. 1 (color online). (a) Plots of $\chi(t)/\beta$ vs $[C(t) + D(t)]/2$ for times $t = 5, 10, \dots, 50$ (other parameters: see legend and $h = -0.1$, $d_1 - d_n = 0.8$). (b) $C(t = 50)$ and $D(t = 50)$ as functions of the unbalance between reservoir densities, for $\beta = 1$, $n = 20$. At equilibrium ($d_1 = d_n$) one has $C(50) = D(50)$ as expected, while the difference between $C(50)$ and $D(50)$ is not negligible for strong nonequilibrium.

$C(t)$ alone cannot provide a good estimate of the susceptibility $\chi(t)$.

In order to connect Eq. (6) [or (7)] with some previous formulations, it is easy enough to rewrite it in another form. First remember that causality implies $R_{QV}(s, t) = 0$ for $s \leq t$. Hence, for (6) to be valid for all s, t , one needs

$$\beta \frac{d}{dt} \langle V(t)Q(s) \rangle = \langle \tau(\omega, t)Q(s) \rangle \quad \text{for } s \leq t, \quad (9)$$

which is automatically satisfied in our framework: it follows from $\langle e^{-A(\omega)} \rangle = 1$, equivalent to $\langle A \rangle = 0$ to first order in h_s , via partial integration of (3). By subtracting (9) from (6) one finds

$$\begin{aligned} R_{QV}(t, s) = & \frac{\beta}{2} \left(\frac{d}{ds} \langle V(s)Q(t) \rangle - \frac{d}{dt} \langle V(t)Q(s) \rangle \right) \\ & - \frac{1}{2} \langle \tau(\omega, s)Q(t) - \tau(\omega, t)Q(s) \rangle. \end{aligned} \quad (10)$$

In stationarity the second line in (10) is an extra term compared to the equilibrium version (1), in which the time-antisymmetric correlation between activity τ and the observable Q is evaluated. This term has been called *asymmetry* in studies of overdamped Langevin equations [8] and of discrete stochastic systems [9,10].

Again, when $\tau(\omega, s) = \tau(s)$, as in (4) and (7), Eq. (9) implies that $\tau(s) = \beta LV(s)$. Hence (6) turns into

$$R_{QV}(t, s) = \frac{\beta}{2} \frac{d}{ds} \langle V(s)Q(t) \rangle - \frac{\beta}{2} \langle LV(s)Q(t) \rangle, \quad (11)$$

which is the rigorous version of (7). Let us now make contact with some previous formulations for stationary processes, where $\frac{d}{ds} \langle V(s)Q(t) \rangle = \frac{d}{ds} \langle V(0)Q(t-s) \rangle = -\frac{d}{dt} \langle V(s)Q(t) \rangle$. From our definitions, the last term equals $-\langle V(s)LQ(t) \rangle$. Alternatively we can think of $Q(t)$ correlated with $V(s)$ evolved backward in time by the adjoint L^* (which generates the time-reversed process): $\langle V(s)LQ(t) \rangle = \langle L^*V(s)Q(t) \rangle$. Equation (11) can then be rewritten as

$$R_{QV}(t, s) = -\frac{\beta}{2} \langle L^*V(s)Q(t) + LV(s)Q(t) \rangle, \quad (12)$$

which is a generalization of Eq. 2.15 in [7]. Indeed, in the context of fluctuation theory around diffusive scaling limits, the L^* is referred to as the adjoint hydrodynamics in the infinite-dimensional treatment in Sec. 2.3 of [7]. It is emphasized there that the response for the adjoint process is typically along the reversed trajectory from a spontaneous fluctuation in the original dynamics. This is also explicit in (12), as the response in the time-reversed process amounts to the same expression upon exchanging L with L^* and t with s . Note however that one needs the stationary distribution to know L^* . The knowledge of this distribution enables also the derivation of other FDRs [2,4–6,21]. On the other hand, Eq. (11) does not involve the stationary law, except for the statistical averaging.

We also recover the interpretation of [11]: by using (12), we rewrite the stationary version of (6) as

$$R_{QV}(t, s) = \beta \frac{d}{ds} \langle V(s)Q(t) \rangle - \frac{\beta}{2} \langle [(L - L^*)V](s)Q(t) \rangle. \quad (13)$$

Note that the response regains its equilibrium form [only the first term in the right-hand side of (13)] whenever V is a time-direction neutral observable, in the precise sense that $LV = L^*V$. Furthermore one can check (say, for overdamped diffusions and for jump processes) that $L - L^* = 2 \frac{j}{\rho} \cdot \nabla$, where j is the probability current and ρ is the stationary distribution on the states. Their ratio is a drift velocity $v = j/\rho$. Hence, for a time-direction neutral potential V , the probability current is orthogonal to its gradient, $v \cdot \nabla V = 0$, and the equilibrium form of the response is obtained, in agreement with the observations in [22]. The same effect is achieved by describing the system in the Lagrangian frame moving with drift velocity v . The second term in (13) vanishes also in this case, and (13) yields exactly the interpretation (and Eq. 4.5) of [11], which was recently experimentally verified [23].

To finish, we revisit the relation with dissipation to confirm the prediction in [24] that the usual equilibrium relation between response and dissipation is preserved when taking into account only the excess heating and ignoring the housekeeping heat. During a specific trajectory of the system, the heat dissipation can be split into two parts: $\mathcal{Q}_{\text{hk}} + \mathcal{Q}_{\text{ex}}$ [24]. The first term is the housekeeping heat, which is the heat produced in the unperturbed dynamics. This is the heat that drives (or is the result of) the system out of equilibrium. The second part \mathcal{Q}_{ex} is the extra heat generated through the perturbation, and is just $1/\beta$ times the excess entropy production, which we already encountered in our calculations:

$$\mathcal{Q}_{\text{ex}} = \frac{1}{\beta} \langle S(\omega) \rangle^h = \int_0^t ds h_s \frac{d}{ds} \langle V(s) \rangle^h. \quad (14)$$

This expression is of the same form as in equilibrium, so we will not repeat the calculations here, only the conclusion: the (excess) heat is given by the imaginary part of the Fourier transform of R_{VV} , or, equivalently, the Fourier transform of the time-antisymmetric part of R_{VV} . Moreover, we have already written down this time-antisymmetric part: it is the right-hand side of (10), extended to $t \leq s$.

In conclusion, from physical constraints on the probability of trajectories, we have obtained a general FDR for the response of a driven system to the addition of a potential. It lifts the nonequilibrium FDR beyond formal first-order perturbation theory applied to a specific dynamics, by identifying in general physical terms the statistical quantities that determine the response. Previous formulations are recovered as specific cases. Finally, we have given

observational significance to the notion of dynamical activity, which so far has mostly appeared as a theoretical concept in fluctuation theories.

We are grateful to W. De Roeck, K. Gawędzki, and K. Netočný for enlightening discussions. B.W. receives support from FWO, Flanders. M.B. benefits from K.U. Leuven Grant No. OT/07/034A.

*Bram.Wynants@fys.kuleuven.be

- [1] R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).
- [2] U. Marini Bettolo Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani, Phys. Rep. **461**, 111 (2008).
- [3] J.P. Sethna, *Entropy, Order Parameters, and Complexity* (Oxford University Press, Oxford, 2006).
- [4] U. Decker and F. Haake, Phys. Rev. A **11**, 2043 (1975).
- [5] P. Hänggi and H. Thomas, Phys. Rep. **88**, 207 (1982).
- [6] M. Falcioni, S. Isola, and A. Vulpiani, Phys. Lett. A **144**, 341 (1990).
- [7] L. Bertini *et al.*, J. Stat. Phys. **107**, 635 (2002).
- [8] L. Cugliandolo, J. Kurchan, and G. Parisi, J. Phys. I **4**, 1641 (1994).
- [9] E. Lippiello, F. Corberi, and M. Zannetti, Phys. Rev. E **71**, 036104 (2005).
- [10] G. Diezemann, Phys. Rev. E **72**, 011104 (2005).
- [11] R. Chetrite, G. Falkovich, and K. Gawędzki, J. Stat. Mech. (2008) P08005.
- [12] T. Speck and U. Seifert, Europhys. Lett. **74**, 391 (2006).
- [13] V. Blickle *et al.*, Phys. Rev. Lett. **98**, 210601 (2007).
- [14] C. Maes and K. Netočný, Europhys. Lett. **82**, 30003 (2008); C. Maes, K. Netočný, and B. Wynants, Markov Proc. Rel. Fields **14**, 445 (2008); Physica (Amsterdam) **387A**, 2675 (2008).
- [15] C. Appert-Rolland *et al.*, Phys. Rev. E **78**, 021122 (2008); T. Bodineau and R. Lefevre, J. Stat. Phys. **133**, 1 (2008); M. Gorissen, J. Hooyberghs, and C. Vanderzande, Phys. Rev. E **79**, 020101(R) (2009); J.P. Garrahan *et al.*, J. Phys. A: Math. Theor. **42**, 075007 (2009).
- [16] C. Maes and M.H. van Wieren, Phys. Rev. Lett. **96**, 240601 (2006); V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).
- [17] W. De Roeck and C. Maes, Phys. Rev. E **76**, 051117 (2007).
- [18] C. Maes and K. Netočný, J. Stat. Phys. **110**, 269 (2003); C. Maes, Séminaire. Poincaré **2**, 29 (2003).
- [19] We do not discuss here some extra difficulties for underdamped diffusions, but it is valid for Markov jump processes and for overdamped diffusions.
- [20] $(d/ds)\langle V(s)Q(t) \rangle = (d/ds)\sum_x \mu(x) [e^{sL}(Ve^{(t-s)L}Q)](x) \neq \sum_x \mu(x) \{e^{sL}[(LV)e^{(t-s)L}Q]\}(x) = \langle (dV/ds)Q(t) \rangle$.
- [21] T. Speck and U. Seifert, Phys. Rev. E **79**, 040102(R) (2009).
- [22] T. Nakamura and S. Sasa, Phys. Rev. E **77**, 021108 (2008).
- [23] J.R. Gomez-Solano *et al.*, arXiv:0903.1075.
- [24] Y. Oono and M. Paniconi, Prog. Theor. Phys. Suppl. **130**, 29 (1998); T. Hatano and S. Sasa, Phys. Rev. Lett. **86**, 3463 (2001).