Colloquium: Quantum fluctuation relations: Foundations and applications

Michele Campisi, Peter Hänggi, and Peter Talkner

Institute of Physics, University of Augsburg, Universitätsstrasse 1, D-86135 Augsburg, Germany

(Received 10 December 2010; published 6 July 2011)

Two fundamental ingredients play a decisive role in the foundation of fluctuation relations: the principle of microreversibility and the fact that thermal equilibrium is described by the Gibbs canonical ensemble. Building on these two pillars the reader is guided through a self-contained exposition of the theory and applications of quantum fluctuation relations. These are exact results that constitute the fulcrum of the recent development of nonequilibrium thermodynamics beyond the linear response regime. The material is organized in a way that emphasizes the historical connection between quantum fluctuation relations and (non)linear response theory. A number of fundamental issues are clarified which were not completely settled in the prior literature. The main focus is on (i) work fluctuation relations for transiently driven closed or open quantum systems, and (ii) on fluctuation relations for heat and matter exchange in quantum transport settings. Recently performed and proposed experimental applications are presented and discussed.

DOI: 10.1103/RevModPhys.83.771 PACS numbers: 05.30.−d, 05.40.−a, 05.60.Gg, 05.70.Ln

CONTENTS

I. INTRODUCTION

This Colloquium focuses on fluctuation relations and, in particular, on their quantum versions. These relations constitute a research topic that recently has attracted a great deal of attention. At the microscopic level, matter is in a permanent state of agitation; consequently, many physical quantities of interest continuously undergo random fluctuations. The purpose of statistical mechanics is the characterization of the statistical properties of those fluctuating quantities from the known laws of classical and quantum physics that govern the dynamics of the constituents of matter. A paradigmatic example is the Maxwell distribution of velocities in a rarefied gas at equilibrium, which follows from the sole assumptions that the microdynamics are Hamiltonian, and that the very many system constituents interact via negligible, short range forces (Khinchin, 1949). Besides the fluctuation of velocity (or energy) at equilibrium, one might be interested in the properties of other fluctuating quantities, e.g., heat and work, characterizing nonequilibrium transformations. Imposed by the reversibility of microscopic dynamical laws, the fluctuation relations put severe restrictions on the form that the probability density function (PDF) of the considered nonequilibrium fluctuating quantities may assume. Fluctuation relations are typically expressed in the form

$$p_F(x) = p_B(-x)\exp[a(x-b)], \tag{1}$$

where $p_F(x)$ is the PDF of the fluctuating quantity x during a nonequilibrium thermodynamic transformation, referred to for simplicity as the forward (F) transformation, and $p_B(x)$ is the PDF of x during the reversed (backward B) transformation. The precise meaning of these expressions will be clarified below. The real-valued constants a and b contain information about the *equilibrium* starting points of the B and F transformations. Figure 1 shows a probability distribution satisfying the fluctuation relation, as measured in a recent experiment of electron transport through a nanojunction

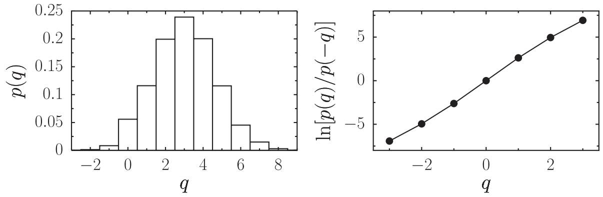

FIG. 1. Example of statistics obeying the fluctuation relation, Eq. (1). Left panel: Probability distribution $p(q)$ of number q of electrons, transported through a nanojunction subject to an electrical potential difference. Right panel: The linear behavior of $\ln[p(q)/p(-q)]$ evidences that $p(q)$ obeys the fluctuation relation, Eq. (1). In this example forward and backward protocols coincide yielding $p_B = p_F \equiv p$, and consequently $b = 0$ in Eq. (1). From Utsumi *et al.*, 2010.

(Utsumi *et al.*, 2010). We analyze this experiment in detail in Sec. VI.

As often happens in science, the historical development of theories is quite tortuous. Fluctuation relations are no exception in this respect. Without any intention of providing a thorough and complete historical account, we mention below a few milestones that, in our view, mark crucial steps in the historical development of quantum fluctuation relations. The beginning of the story might be traced back to the early years of the last century, with the work of Sutherland (1902, 1905) and Einstein (1905, 1906a, 1906b) first, and of Johnson (1928) and Nyquist (1928) later, when it was found that the linear response of a system in thermal equilibrium, as it is driven out of equilibrium by an external force, is determined by the fluctuation properties of the system in the initial equilibrium state. Specifically, Sutherland (1902, 1905) and Einstein (1905, 1906a, 1906b) found a relation between the mobility of a Brownian particle (encoding information about its response to an externally applied force) and its diffusion constant (encoding information about its equilibrium fluctuations). Johnson (1928) and Nyquist (1928)[1] discovered the corresponding relation between the resistance of a circuit and the spontaneous current fluctuations occurring in the absence of an applied electric potential.

The next step was taken by Callen and Welton (1951) who derived the previous results within a general quantum mechanical setting. The starting point of their analysis was a quantum mechanical system described by a Hamiltonian $\mathcal{H}_0$. Initially this system stays in a thermal equilibrium state at the inverse temperature $\beta \equiv (k_B T)^{-1}$, wherein k_B is the Boltzmann constant. This state is described by a density matrix of canonical form; i.e., it is given by a Gibbs state

$$\varrho_0 = e^{-\beta \mathcal{H}_0}/\mathcal{Z}_0, \tag{2}$$

where $\mathcal{Z}_0 = \mathrm{Tr} e^{-\beta \mathcal{H}_0}$ denotes the partition function of the unperturbed system and Tr denotes trace over its Hilbert space. At later times $t > 0$, the system is perturbed by the action of an external, in general, time-dependent force λ_t that couples to an observable $\mathcal{Q}$ of the system. The dynamics of the system then is governed by the modified, time-dependent Hamiltonian

$$\mathcal{H}(\lambda_t) = \mathcal{H}_0 - \lambda_t \mathcal{Q}. \tag{3}$$

The approach of Callen and Welton (1951) was further systematized by Green (1952, 1954) and, in particular, by Kubo (1957) who proved that the linear response is determined by a response function $\phi_{\mathcal{B}\mathcal{Q}}(t)$, which gives the deviation $\langle \Delta B(t) \rangle$ of the expectation value of an observable $\mathcal{B}$ to its unperturbed value as

$$\langle \Delta B(t) \rangle = \int_{-\infty}^{t} \phi_{\mathcal{B}\mathcal{Q}}(t-s)\lambda_s ds. \tag{4}$$

Kubo (1957) showed that the response function can be expressed in terms of the commutator of the observables $\mathcal{Q}$ and $\mathcal{B}^H(t)$ as $\phi_{\mathcal{B}\mathcal{Q}}(s) = \langle [\mathcal{Q}, \mathcal{B}^H(s)] \rangle / i\hbar$ (the superscript H denotes the Heisenberg picture with respect to the unperturbed dynamics.) Moreover, Kubo derived the general relation

$$\langle \mathcal{Q} \mathcal{B}^H(t) \rangle = \langle \mathcal{B}^H(t - i\hbar\beta) \mathcal{Q} \rangle \tag{5}$$

between differently ordered thermal correlation functions and deduced from it the celebrated quantum fluctuation-dissipation theorem (Callen and Welton, 1951), reading

$$\hat{\Psi}_{\mathcal{B}\mathcal{Q}}(\omega) = (\hbar/2i)\coth(\beta\hbar\omega/2)\hat{\phi}_{\mathcal{B}\mathcal{Q}}(\omega), \tag{6}$$

where $\hat{\Psi}_{\mathcal{B}\mathcal{Q}}(\omega) = \int_{-\infty}^{\infty} e^{i\omega s} \Psi_{\mathcal{B}\mathcal{Q}}(s) ds$ denotes the Fourier transform of the symmetrized, stationary equilibrium correlation function $\Psi_{\mathcal{B}\mathcal{Q}}(s) = \langle \mathcal{Q}\mathcal{B}^H(s) + \mathcal{B}^H(s)\mathcal{Q} \rangle / 2$, and $\hat{\phi}_{\mathcal{B}\mathcal{Q}}(\omega) = \int_{-\infty}^{\infty} e^{i\omega s} \phi_{\mathcal{B}\mathcal{Q}}(s) ds$ denotes the Fourier transform of the response function $\phi_{\mathcal{B}\mathcal{Q}}(s)$. Note that the fluctuation-dissipation theorem is valid also for many-particle systems independent of the respective particle statistics. Besides offering a unified and rigorous picture of the fluctuation-dissipation theorem, the theory of Kubo also included other important advancements in the field of nonequilibrium thermodynamics. Specifically, we note the celebrated Onsager-Casimir reciprocity relations (Onsager, 1931a, 1931b; Casimir, 1945). These relations state that, as a consequence of *microreversibility*, the matrix of transport coefficients that connects applied forces to so-called fluxes in a system close to equilibrium consists of a symmetric and an antisymmetric block. The symmetric block couples forces and fluxes that have same parity under time reversal and the antisymmetric block couples forces and fluxes that have different parity.

Most importantly, the analysis of Kubo (1957) opened the possibility for a systematic advancement of response theory, allowing, in particular, one to investigate the existence of higher order fluctuation-dissipation relations beyond linear regime. This task was soon undertaken by Bernard and Callen (1959), who pointed out a hierarchy of irreversible thermodynamic relationships. These higher order fluctuation-dissipation relations were investigated by Stratonovich for the Markovian system, and later by Efremov (1969) for

[1] Nyquist already discussed in his Eq. (8) a precursor of the quantum fluctuation-dissipation theorem as developed later by Callen and Welton (1951). He only missed the correct form by omitting the zero-point energy contribution in his result; see also Hänggi and Ingold (2005).

non-Markovian systems; see Stratonovich (1992), Chap. I, and references therein.

Even for arbitrary systems far from equilibrium the linear response to an applied force can likewise be related to tailored two-point correlation functions of corresponding stationary nonequilibrium fluctuations of the underlying *unperturbed*, stationary nonequilibrium system (Hänggi, 1978; Hänggi and Thomas, 1982). They coined the expression "fluctuation theorems" for these relations. As in the near thermal equilibrium case, in this case higher order nonlinear response can also be linked to corresponding higher order correlation functions of those nonequilibrium fluctuations (Hänggi, 1978; Prost *et al.*, 2009).

At the same time, in the late 1970s Bochkov and Kuzovlev (1977) provided a single compact *classical* expression that contains fluctuation relations of all orders for systems that are at thermal equilibrium when unperturbed. This expression, Eq. (14), can be seen as a fully nonlinear, exact, and universal fluctuation relation. The Bochkov and Kuzovlev formula, Eq. (14), soon turned out useful in addressing the problem of connecting the deterministic and the stochastic descriptions of nonlinear dissipative systems (Bochkov and Kuzovlev, 1978; Hänggi, 1982).

As often happens in physics, the most elegant, compact, and universal relations are consequences of general physical symmetries. In the case of Bochkov and Kuzovlev (1977) the fluctuation relation follows from the time-reversal invariance of the equations of microscopic motion, combined with the assumption that the system initially resides in thermal equilibrium described by the classical analog of the Gibbs state, Eq. (2). Bochkov and Kuzovlev (1977, 1979, 1981a, 1981b) proved Eq. (14) for classical systems. Their derivation will be reviewed in the next section. The quantum version, Eq. (55), was not reported until recently (Andrieux and Gaspard, 2008). In Sec. III.C we discuss the fundamental obstacles that prevented Bochkov and Kuzovlev (1977, 1979, 1981a, 1981b) and Stratonovich (1994), who also studied this very quantum problem, from obtaining Eq. (55).

A new wave of activity in fluctuation relations was initiated by Evans *et al.* (1993) and Gallavotti and Cohen (1995) on the statistics of the entropy produced in nonequilibrium steady states, and by Jarzynski (1997) on the statistics of work performed by a transient, time-dependent perturbation. Since then, the field has generated much interest and flourished. The existing reviews on this topic mostly cover classical fluctuation relations (Jarzynski, 2008, 2011; Marconi *et al.*, 2008; Rondoni and Mejía-Monasterio, 2007; Seifert, 2008), while the comprehensive review by Esposito *et al.* (2009) provided a solid, though in parts technical account of the state of the art of quantum fluctuation theorems. With this work we want to present a widely accessible introduction to quantum fluctuation relations, covering as well the most recent decisive advancements. Particularly, our emphasis will be on (i) their connection to the linear and nonlinear response theories (Sec. II), (ii) the clarification of fundamental issues that relate to the notion of "work" (Sec. III), (iii) the derivation of quantum fluctuation relations for both closed and open quantum systems (Secs. IV and V), and also (iv) their impact for experimental applications and validation (Sec. VI).

II. NONLINEAR RESPONSE THEORY AND CLASSICAL FLUCTUATION RELATIONS

A. Microreversibility of nonautonomous classical systems

Two ingredients are at the heart of fluctuation relations. The first one concerns the initial condition of the system under study. This is supposed to be in thermal equilibrium described by a canonical distribution of the form of Eq. (2). It hence is of *statistical* nature. Its use and properties are discussed in many textbooks on statistical mechanics. The other ingredient, concerning the dynamics of the system, is the principle of microreversibility. This point needs some clarification since microreversibility is customarily understood as a property of autonomous (i.e., nondriven) systems described by a time-independent Hamiltonian (Messiah, 1962, Vol. 2, Ch. XV). On the contrary, here we are concerned with nonautonomous systems, governed by explicitly time-dependent Hamiltonians. In the following we analyze this principle for classical systems in a way that at first glance may appear rather formal but will prove indispensable later on. The analogous discussion of the quantum case is given next in Sec. IV.A.

We deal here with a classical system characterized by a Hamiltonian that consists of an unperturbed part $H_0(\mathbf{z})$ and a perturbation $-\lambda_t Q(\mathbf{z})$ due to an external force λ_t that couples to the conjugate coordinate $Q(\mathbf{z})$. Then the total system Hamiltonian becomes[2]

$$H(\mathbf{z}, \lambda_t) = H_0(\mathbf{z}) - \lambda_t Q(\mathbf{z}), \tag{7}$$

where $\mathbf{z} = (\mathbf{q}, \mathbf{p})$ denotes a point in the phase space of the considered system. In the following we assume that the force acts within a temporal interval set by a starting time 0 and a final time τ. The instantaneous force values λ_t are specified by a function λ, which we refer to as the force protocol. In the sequel, it turned out necessary to clearly distinguish between the function λ and the value λ_t that it takes at a particular instant of time t.

For these systems the principle of microreversibility holds in the following sense. The solution of Hamilton's equations of motion assigns to each initial point in phase space $\mathbf{z}_0 = (\mathbf{q}_0, \mathbf{p}_0)$ a point $\mathbf{z}_t$ at the later time $t \in [0, \tau]$, which is specified by the values of the force in the order of their appearance within the considered time span. Hence, the position

$$\mathbf{z}_t = \varphi_{t,0}[\mathbf{z}_0; \lambda] \tag{8}$$

at time t is determined by the flow $\varphi_{t,0}[\mathbf{z}_0; \lambda]$ which is a function of the initial point $\mathbf{z}_0$ and a functional of the force protocol λ.[3] In a computer simulation one can invert the direction of time and let the trajectory run backward without a problem. Although, as experience evidences, it is impossible to actively revert the direction of time in any experiment, there is yet a way to run a time-reversed trajectory in real time. For simplicity we assume that the Hamiltonian H_0

[2]The generalization to the case of several forces coupling via different conjugate coordinates is straightforward.

[3]Because of causality, $\varphi_{t,0}[\mathbf{z}_0, \lambda]$ may, of course, depend only on the part of the protocol including times from $t = 0$ up to time t.

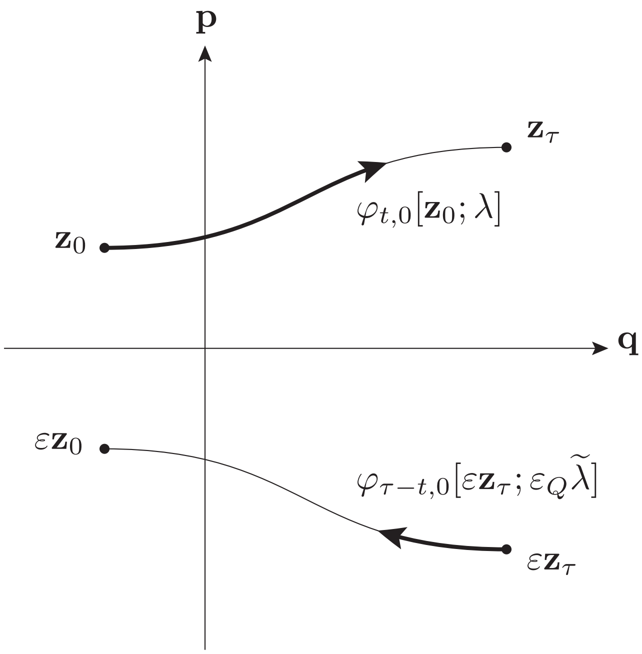

FIG. 2. Microreversibility for nonautonomous classical (Hamiltonian) systems. The initial condition $\mathbf{z}_0$ evolves, under the protocol λ, from time $t = 0$ until time t to $\mathbf{z}_t = \varphi_{t,0}[\mathbf{z}_0; \lambda]$ and until time $t = \tau$ to $\mathbf{z}_\tau$. The time-reversed final condition $\varepsilon\mathbf{z}_\tau$ evolves, under the protocol $\varepsilon_Q\tilde{\lambda}$ from time $t = 0$ until $\tau - t$ to $\varphi_{\tau-t,0}[\varepsilon\mathbf{z}_\tau; \varepsilon_Q\tilde{\lambda}] = \varepsilon\varphi_{t,0}[\mathbf{z}_0; \lambda]$, Eq. (10), and until time $t = \tau$ to the time-reversed initial condition $\varepsilon\mathbf{z}_0$.

is time reversal invariant, i.e., that it remains unchanged if the signs of momenta are reverted. Moreover, we restrict ourselves to conjugate coordinates $Q(\mathbf{z})$ that transform under time reversal with a definite parity $\varepsilon_Q = \pm 1$. Stratonovich (1994, Sec. 1.2.3) showed that the flow under the backward protocol $\tilde{\lambda}$, with

$$\tilde{\lambda}_t = \lambda_{\tau-t}, \tag{9}$$

is related to the flow under the forward protocol λ via

$$\varphi_{t,0}[\mathbf{z}_0; \lambda] = \varepsilon\varphi_{\tau-t,0}[\varepsilon\mathbf{z}_\tau; \varepsilon_Q\tilde{\lambda}], \tag{10}$$

where ε maps any phase-space point $\mathbf{z}$ on its time-reversed image $\varepsilon\mathbf{z} = \varepsilon(\mathbf{q}, \mathbf{p}) = (\mathbf{q}, -\mathbf{p})$. Equation (10) expresses the principle of microreversibility in driven systems. Its meaning is shown in Fig. 2. Particularly, it states that in order to trace back a trajectory, one has to reverse the sign of the velocity, as well as the temporal succession of the force values λ, and the sign of force $\tilde{\lambda}$, if $\varepsilon_Q = -1$.

B. Bochkov-Kuzovlev approach

We consider a phase-space function $B(\mathbf{z})$ which has a definite parity under time reversal $\varepsilon_B = \pm 1$, i.e., $B(\varepsilon\mathbf{z}) = \varepsilon_B B(\mathbf{z})$. $B_t = B(\varphi_{t,0}[\mathbf{z}_0; \lambda])$ denotes its temporal evolution. Depending on the initial condition $\mathbf{z}_0$ different trajectories B_t are realized. Under the above stated assumption that at time $t = 0$ the system is prepared in a Gibbs equilibrium, the initial conditions are randomly sampled from the distribution

$$\rho_0(\mathbf{z}_0) = e^{-\beta H_0(\mathbf{z}_0)}/Z_0, \tag{11}$$

with $Z_0 = \int d\mathbf{z}_0 e^{-\beta H_0(\mathbf{z}_0)}$. Consequently the trajectory B_t becomes a random quantity. Next we introduce the quantity

$$W_0[\mathbf{z}_0; \lambda] = \int_0^\tau dt \lambda_t \dot{Q}_t, \tag{12}$$

where $\dot{Q}_t$ is the time derivative of $Q_t = Q(\varphi_{t,0}[\mathbf{z}_0; \lambda])$. From Hamilton's equations it follows that (Jarzynski, 2007):

$$W_0[\mathbf{z}_0; \lambda] = H_0(\varphi_{\tau,0}[\mathbf{z}_0; \lambda]) - H_0(\mathbf{z}_0). \tag{13}$$

Therefore, we interpret W_0 as the work injected in the system described by H_0 during the action of the force protocol.[4] The central finding of Bochkov and Kuzovlev (1977) is a formal relation between the generating functional for multitime correlation functions of the phase-space functions B_t and Q_t and the generating functional for the time-reversed multitime autocorrelation functions of B_t, reading

$$\langle e^{\int_0^\tau du_t B_t} e^{-\beta W_0} \rangle_\lambda = \langle e^{\int_0^\tau d\tilde{u}_t \varepsilon_B B_t} \rangle_{\varepsilon_Q \tilde{\lambda}}, \tag{14}$$

where u_τ is an arbitrary test function, $\tilde{u}_t = u_{\tau-t}$ is its temporal reverse, and the average denoted by $\langle\cdot\rangle$ is taken with respect to the Gibbs distribution ρ_0 of Eq. (11). On the left-hand side, the time evolutions of B_t and Q_t are governed by the full Hamiltonian (7) in the presence of the forward protocol as indicated by the subscript λ, while on the right-hand side the dynamics is determined by the time-reversed protocol indicated by the subscript $\varepsilon_Q\tilde{\lambda}$. The derivation of Eq. (14), which is based on the microreversibility principle, Eq. (10), is given in Appendix A. The importance of Eq. (14) lies in the fact that it contains the Onsager reciprocity relations and fluctuation relations of all orders within a single compact formula (Bochkov and Kuzovlev, 1977). These relations may be obtained by means of functional derivatives of both sides of Eq. (14), of various orders, with respect to the force field λ and the test field u at vanishing fields $\lambda = u = 0$. The classical limit of the Callen and Welton (1951) fluctuation-dissipation theorem, Eq. (6), for instance, is obtained by differentiation with respect to u, followed by a differentiation with respect to λ (Bochkov and Kuzovlev, 1977), both at vanishing fields u and λ.

Another remarkable identity is achieved from Eq. (14) by setting $u = 0$, but leaving the force λ finite. This yields the Bochkov-Kuzovlev equality, reading

$$\langle e^{-\beta W_0} \rangle_\lambda = 1. \tag{15}$$

In other words, for any system that initially stays in thermal equilibrium at a temperature $T = 1/k_B\beta$, the work, Eq. (12), done on the system by an external force is a random quantity with an exponential expectation value $\langle e^{-\beta W_0} \rangle_\lambda$ that is independent of any detail of the system and the force acting on it. This, of course, does not hold for the individual nth moments of work. Since the exponential function is convex, a direct consequence of Eq. (15) is

[4]Following Jarzynski (2007) we refer to W_0 as the *exclusive* work, to distinguish from the *inclusive* work $W = H(\mathbf{z}_\tau, \lambda_\tau) - H(\mathbf{z}_0, \lambda_0)$, Eq. (19), which accounts also for the coupling between the external source and the system. We will come back later to these two definitions of work in Sec. III.A.

$$\langle W_0 \rangle_\lambda \geq 0. \tag{16}$$

That is, on average, a driven Hamiltonian system may only absorb energy if it is perturbed out of thermal equilibrium. This does not exclude the existence of energy releasing events which, in fact, must happen with certainty in order that Eq. (15) holds if the average work is larger than zero. Equation (16) may be regarded as a microscopic manifestation of the second law of thermodynamics. For this reason Stratonovich (1994, Sec. 1.2.4) referred to it as the H theorem. We recapitulate that only two ingredients, initial Gibbsian equilibrium and microreversibility of the dynamics, have led to Eq. (14). In conclusion, this relation not only contains linear and nonlinear response theories, but also the second law of thermodynamics.

The complete information about the statistics is contained in the work PDF $p_0[W_0;\lambda]$. The only random element entering the work, Eq. (13), is the initial phase point $\mathbf{z}_0$ which is distributed according to Eq. (11). Therefore, $p_0[W_0;\lambda]$ may formally be expressed as

$$p_0[W_0;\lambda] = \int d\mathbf{z}_0 \rho_0(\mathbf{z}_0)\delta(W_0 - H_0(\mathbf{z}_\tau) + H_0(\mathbf{z}_0)), \tag{17}$$

where δ denotes Dirac's delta function. The functional dependence of $p_0[W_0;\lambda]$ on the force protocol λ is contained in the term $\mathbf{z}_\tau = \varphi_{\tau,0}[\mathbf{z}_0;\lambda]$. Using the microreversibility principle, Eq. (10), one obtains the following fluctuation relation:

$$\frac{p_0[W_0;\lambda]}{p_0[-W_0;\varepsilon_Q\tilde{\lambda}]} = e^{\beta W_0}, \tag{18}$$

in a way analogous to the derivation of Eq. (14). We refer to this relation as the Bochkov-Kuzovlev work fluctuation relation, although it was not explicitly given by Bochkov and Kuzovlev, but was only recently obtained by Horowitz and Jarzynski (2007). This equation has a profound physical meaning. Consider a positive work $W_0 > 0$. Then Eq. (18) says that the probability that this work is injected into the system is larger by the factor $e^{\beta W_0}$ than the probability that the same work is released under the reversed forcing: Energy consuming processes are exponentially more probable than energy releasing processes. Thus, Eq. (18) expresses the second law of thermodynamics at a detailed level which quantifies the relative frequency of energy releasing processes. By multiplying both sides of Eq. (18) by $p_0[-W_0;\varepsilon_Q\tilde{\lambda}]e^{-\beta W_0}$ and integrating over W_0, one recovers the Bochkov-Kuzovlev identity, Eq. (15).

C. Jarzynski approach

An alternative definition of work is based on the comparison of the *total* Hamiltonians at the end and the beginning of a force protocol, leading to the notion of "inclusive" work in contrast to the "exclusive" work defined in Eq. (13). The latter equals the energy difference referring to the unperturbed Hamiltonian H_0. Accordingly, the inclusive work is the difference of the total Hamiltonians at the final time $t = \tau$ and the initial time $t = 0$:

$$W[\mathbf{z}_0;\lambda] = H(\mathbf{z}_\tau, \lambda_\tau) - H(\mathbf{z}_0, \lambda_0). \tag{19}$$

In terms of the force λ_t and the conjugate coordinate Q_t, the inclusive work is expressed as[5]:

$$\begin{aligned} W[\mathbf{z}_0;\lambda] &= \int_0^\tau dt \dot{\lambda}_t \frac{\partial H(\mathbf{z}_t, \lambda_t)}{\partial \lambda_t} = -\int_0^\tau dt \dot{\lambda}_t Q_t \\ &= W_0[\mathbf{z}_0;\lambda] - \lambda_t Q_t|_0^\tau. \end{aligned} \tag{20}$$

For simplicity we confined ourselves to the case of an even conjugate coordinate Q. In a corresponding way, as described in Appendix A, we obtained the following relation between generating functionals of forward and backward processes in analogy to Eq. (14):

$$\langle e^{\int_0^\tau du_t B_t} e^{-\beta W}\rangle_\lambda = \frac{Z(\lambda_\tau)}{Z(\lambda_0)} \langle e^{\int_0^t dt \tilde{u}_t \varepsilon_B B_t}\rangle_{\tilde{\lambda}}. \tag{21}$$

While on the left-hand side the time evolution is controlled by the forward protocol λ and the average is performed with respect to the initial thermal distribution $\rho^\beta(\mathbf{z}, \lambda_0)$, on the right-hand side the time evolution is governed by the reversed protocol $\tilde{\lambda}$ and the average is taken over the reference equilibrium state $\rho^\beta(\mathbf{z}, \tilde{\lambda}_0) = \rho^\beta(\mathbf{z}, \lambda_\tau)$. Here

$$\rho^\beta(\mathbf{z}, \lambda_t) = e^{-\beta H(\mathbf{z}, \lambda_t)}/Z(\lambda_t) \tag{22}$$

formally describes the thermal equilibrium of a system with the Hamiltonian $H(\mathbf{z}, \lambda_t)$ at the inverse temperature β. The partition function $Z(\lambda_t)$ is defined accordingly as $Z(\lambda_t) = \int d\mathbf{z} e^{-\beta H(\mathbf{z}, \lambda_t)}$. Note that in general the reference state $\rho^\beta(\mathbf{z}, \lambda_t)$ is different from the actual phase-space distribution reached under the action of the protocol λ at time t, i.e., $\rho(\mathbf{z}, t) = \rho^\beta(\varphi_{t,0}^{-1}[\mathbf{z};\lambda], \lambda_0)$, where $\varphi_{t,0}^{-1}[\mathbf{z};\lambda]$ denotes the point in phase space that evolves to $\mathbf{z}$ in the time 0 to t under the action of λ.

Setting $u \equiv 0$ we obtain

$$\langle e^{-\beta W}\rangle_\lambda = e^{-\beta \Delta F}, \tag{23}$$

where

$$\Delta F = F(\lambda_\tau) - F(\lambda_0) = -\beta \ln \frac{Z(\lambda_\tau)}{Z(\lambda_0)} \tag{24}$$

is the free energy difference between the reference state $\rho^\beta(\mathbf{z}, \lambda_t)$ and the initial equilibrium state $\rho^\beta(\mathbf{z}\lambda_0)$. As a consequence of Eq. (23) we have

$$\langle W \rangle_\lambda \geq \Delta F, \tag{25}$$

which is yet another expression of the second law of thermodynamics. Equation (23) was first put forward by Jarzynski (1997) and is commonly referred to in the literature as the "Jarzynski equality."

In close analogy to the Bochkov-Kuzovlev approach the PDF of the inclusive work can be formally expressed as

$$p[W;\lambda] = \int d\mathbf{z}_0 \rho(\mathbf{z}_0, \lambda_0)\delta(W - H(\mathbf{z}_\tau, \lambda_\tau) + H(\mathbf{z}_0, \lambda_0)). \tag{26}$$

Its Fourier transform defines the characteristic function of work:

[5]For a further discussion of inclusive and exclusive work we refer to Sec. III.A.

$$
\begin{aligned}
G[u;\lambda] &= \int dW e^{iuW} p[W;\lambda] \\
&= \int d\mathbf{z}_0 e^{iu[H(\mathbf{z}_\tau,\lambda_\tau)-H(\mathbf{z}_0,\lambda_0)]} e^{-\beta H(\mathbf{z}_0,\lambda_0)}/Z(\lambda_0) \\
&= \int d\mathbf{z}_0 \exp\left[iu \int_0^\tau dt \dot{\lambda}_t \frac{\partial H(\mathbf{z}_t,\lambda_t)}{\partial \lambda_t} \right] \frac{e^{-\beta H(\mathbf{z}_0,\lambda_0)}}{Z(\lambda_0)}.
\end{aligned} \tag{27}
$$

Using the microreversibility principle, Eq. (10), we obtain in a way similar to Eq. (18) the (inclusive) work fluctuation relation:

$$
\frac{p[W;\lambda]}{p[-W;\tilde{\lambda}]} = e^{\beta(W-\Delta F)}, \tag{28}
$$

where the probability $p[-W;\tilde{\lambda}]$ refers to the backward process which for the inclusive work has to be determined with reference to the initial thermal state $\rho^\beta(\mathbf{z},\lambda_\tau)$. First put forward by Crooks (1999), Eq. (28) is commonly referred to in the literature as the "Crooks fluctuation theorem." The Jarzynski equality, Eq. (23), is obtained by multiplying both sides of Eq. (28) by $p[-W;\tilde{\lambda}]e^{-\beta W}$ and integrating over W. Equations (21), (23), and (28) continue to hold also when Q is odd under time reversal, with the provision that $\tilde{\lambda}$ is replaced by $-\tilde{\lambda}$.

We here point out the salient fact that, within the inclusive approach, a connection is established between the *nonequilibrium* work W and the difference of free energies ΔF, of the corresponding *equilibrium states* $\rho^\beta(\mathbf{z},\lambda_\tau)$ and $\rho^\beta(\mathbf{z},\lambda_0)$. Most remarkably, Eq. (25) says that the average (inclusive) work is always larger than or equal to the free energy difference, no matter the form of the protocol λ; even more surprising is the content of Eq. (23) saying that the equilibrium free energy difference may be inferred by measurements of nonequilibrium work in many realizations of the forcing experiment (Jarzynski, 1997). This is similar in spirit to the fluctuation-dissipation theorem, also connecting an equilibrium property (the fluctuations) to a nonequilibrium one (the linear response), with the major difference that Eq. (23) is an exact result, whereas the fluctuation-dissipation theorem holds only to first order in the perturbation. Note that as a consequence of Eq. (28) the forward and backward PDFs of exclusive work take on the same value at $W = \Delta F$. This property has been used in experiments (Liphardt *et al.*, 2002; Collin *et al.*, 2005; Douarche *et al.*, 2005) in order to determine free energy differences from nonequilibrium measurements of work. Equations (23) and (28) have further been employed to develop efficient numerical methods for the estimation of free energies (Jarzynski, 2002; Minh and Adib, 2008; Vaikuntanathan and Jarzynski, 2008; Hahn and Then, 2009, 2010).

Both the Crooks fluctuation theorem, Eq. (28), and the Jarzynski equality, Eq. (23), continue to hold for any time-dependent Hamiltonian $H(\mathbf{z},\lambda_t)$ without restriction to Hamiltonians of the form in Eq. (7). Indeed no restriction of the form in Eq. (7) was imposed in the seminal paper by Jarzynski (1997). In the original works of Jarzynski (1997) and Crooks (1999), Eqs. (23) and (28) were obtained directly, without passing through the more general formula in Eq. (21). Notably, neither these seminal papers nor the subsequent literature refer to such general functional identities as Eq. (21). We introduced them here to emphasize the connection between the recent results, Eqs. (23) and (28), with the older results of Bochkov and Kuzovlev (1977), Eqs. (15) and (18). The latter ones were practically ignored, or sometimes misinterpreted as special instances of the former ones for the case of cyclic protocols ($\Delta F = 0$), by those working in the field of nonequilibrium work fluctuations. Only recently Jarzynski (2007) pointed out the differences and analogies between the inclusive and exclusive approaches.

III. FUNDAMENTAL ISSUES

A. Inclusive, exclusive, and dissipated work

As evidenced in the previous section, the studies of Bochkov and Kuzovlev (1977) and Jarzynski (1997) are based on different definitions of work, Eqs. (13) and (19) reflecting two different viewpoints (Jarzynski, 2007). From the exclusive viewpoint of Bochkov and Kuzovlev (1977) the change in the energy H_0 of the unforced system is considered, thus the forcing term ($-\lambda_t Q$) of the total Hamiltonian is not included in the computation of work. From the inclusive point of view the definition of work, Eq. (19), is based on the change of the total energy H including the forcing term ($-\lambda_t Q$). In experiments and practical applications of fluctuation relations, special care must be paid in properly identifying the measured work with either the inclusive (W) or exclusive (W_0) work, bearing in mind that λ represents the prescribed parameter progression and Q is the measured conjugate coordinate.

The experiment of Douarche *et al.* (2005) is well suited to illustrate this point. In that experiment a prescribed torque M_t was applied to a torsion pendulum whose angular displacement θ_t was continuously monitored. The Hamiltonian of the system is

$$
H(\mathbf{y},p_\theta,\theta,M_t) = H_B(\mathbf{y}) + H_{SB}(\mathbf{y},p_\theta,\theta) + \frac{p_\theta^2}{2I} + \frac{I\omega^2\theta^2}{2} - M_t\theta, \tag{29}
$$

where p_θ is the canonical momentum conjugate to θ, $H_B(\mathbf{y})$ is the Hamiltonian of the thermal bath to which the pendulum is coupled via the Hamiltonian H_{SB}, and $\mathbf{y}$ is a point in the bath phase space. Using the definitions of inclusive and exclusive work, Eqs. (12) and (20), and noticing that M plays the role of λ and θ that of Q, we find in this case $W = -\int \theta \dot{M} dt$ and $W_0 = \int M \dot{\theta} dt$.

Note that the work $W = -\int \theta \dot{M} dt$, obtained by monitoring the pendulum degree of freedom only, amounts to the energy change of the total pendulum + bath system. This is true in general (Jarzynski, 2004). Writing the total Hamiltonian as

$$
H(\mathbf{x},\mathbf{y},\lambda_t) = H_S(\mathbf{x},\lambda_t) + H_{BS}(\mathbf{x},\mathbf{y}) + H_B(\mathbf{y}), \tag{30}
$$

with $H_S(\mathbf{x},\lambda_t)$ being the Hamiltonian of the system of interest, one obtains

$$
\int_0^\tau dt \frac{\partial H_S}{\partial \lambda_t} \dot{\lambda}_t = \int_0^\tau dt \frac{\partial H}{\partial t} = \int_0^\tau dt \frac{dH}{dt} = W, \tag{31}
$$

because H_{BS} and H_B do not depend on time, and as a consequence of Hamilton's equations of motion $dH/dt = \partial H/\partial t$.

Introducing the notation $W_{\rm diss} = W - \Delta F$, for the dissipated work, one deduces that the Jarzynski equality can be reexpressed in a way that looks exactly like the Bochkov-Kuzovlev identity, namely,

$$\langle e^{-\beta W_{\rm diss}}\rangle_\lambda = 1. \tag{32}$$

This might lead one to believe that the dissipated work coincides with W_0. This, however, would be incorrect. As discussed by Jarzynski (2007) and explicitly demonstrated by Campisi *et al.* (2011a), W_0 and $W_{\rm diss}$ constitute distinct stochastic quantities with different PDFs. The inclusive, exclusive, and dissipated work coincides only in the case of cyclic forcing $\lambda_\tau=\lambda_0=0$ (Campisi *et al.*, 2011a).

B. The problem of gauge freedom

We pointed out that the inclusive work W, and free energy difference ΔF, as defined in Eqs. (19) and (24) are, to use the expression coined by Cohen-Tannoudji *et al.* (1977), not "true physical quantities." That is to say they are not invariant under gauge transformations that lead to a time-dependent shift of the energy reference point. To elucidate this, consider a mechanical system whose dynamics are governed by the Hamiltonian $H(\mathbf{z}, \lambda_t)$. The new Hamiltonian

$$H'(\mathbf{z}, \lambda_t) = H(\mathbf{z}, \lambda_t) + g(\lambda_t), \tag{33}$$

where $g(\lambda_t)$ is an arbitrary function of the time-dependent force, generates the same equations of motion as H. However, the work $W' = H'(\mathbf{z}_\tau, \lambda_\tau) - H'(\mathbf{z}_0, \lambda_0)$ that one obtains from this Hamiltonian differs from the one that follows from H, Eq. (19): $W' = W + g(\lambda_\tau) - g(\lambda_0)$. Likewise we have for the free energy difference $\Delta F' = \Delta F + g(\lambda_\tau) - g(\lambda_0)$. Evidently the Jarzynski equality, Eq. (23), is invariant under such gauge transformations, because the term $g(\lambda_\tau) - g(\lambda_0)$ appearing on both sides of the identity in the primed gauge, would cancel; explicitly this reads

$$\langle e^{-\beta W'}\rangle_\lambda = e^{-\beta\Delta F'} \Leftrightarrow \langle e^{-\beta W}\rangle_\lambda = e^{-\beta\Delta F}. \tag{34}$$

Thus, there is no fundamental problem associated with the gauge freedom.

However, one must be aware that, in each particular experiment, the way by which the work is measured implies a specific gauge. Consider, for example, the torsion pendulum experiment of Douarche *et al.* (2005). The inclusive work was computed as $W = -\int \theta \dot{M} dt$. The condition that this measured work is related to the Hamiltonian of Eq. (29) via $W = H(\mathbf{z}_\tau, \lambda_\tau) - H(\mathbf{z}_0, \lambda_0)$, Eq. (19), is equivalent to $-\int_0^\tau \theta \dot{M} dt = \int_0^\tau (\partial H/\partial M)\dot{M} dt$; see Eq. (31). If this is required for all τ then the stricter condition $\partial H/\partial M = -\theta$ is implied, restricting the remaining gauge freedom to the choice of a constant function g. This residual freedom, however, is not important as it affects neither work nor free energy. We now consider a different experimental setup where the support to which the pendulum is attached is rotated in a prescribed way according to a protocol α_t, specifying the angular position of the support with respect to the laboratory frame. The dynamics of the pendulum are now described by the Hamiltonian

$$H = H_B + H_{SB} + \frac{p_\theta^2}{2I} + \frac{I\omega^2\theta^2}{2} - I\omega^2\alpha_t\theta + g(\alpha_t). \tag{35}$$

If the work $W = \int N\dot{\alpha} dt$ done by the elastic torque $N = I\omega^2(\alpha - \theta)$ on the support is recorded then the requirement $\partial H/\partial\alpha = N$ singles out the gauge $g(\alpha_t) = I\omega^2\alpha_t^2/2 +$ const, leaving only the freedom to chose the unimportant constant. Note that when $M_t = I\omega^2\alpha_t$, the pendulum obeys exactly the same equations of motion in the two examples above, Eqs. (29) and (35). The gauge is irrelevant for the law of motion but is essential for the energy-Hamiltonian connection.[6]

The issue of gauge freedom was first pointed out by Vilar and Rubi (2008c), who questioned whether a connection between work and Hamiltonian may actually exist. Since then this topic has been highly debated,[7] but neither the gauge invariance of fluctuation relations nor the fact that different experimental setups imply different gauges was clearly recognized before.

C. Work is not a quantum observable

Thus far we have reviewed the general approach to work fluctuation relations for classical systems. The question then naturally arises of how to treat the quantum case. Obviously, the Hamilton function $H(\mathbf{z}, \lambda_t)$ is to be replaced by the Hamilton operator $\mathcal{H}(\lambda_t)$, Eq. (3). The probability density $\rho(\mathbf{z}, \lambda_t)$ is then replaced by the density matrix $\varrho(\lambda_t)$, reading

$$\varrho(\lambda_t) = e^{-\beta\mathcal{H}(\lambda_t)}/\mathcal{Z}(\lambda_t), \tag{36}$$

where $\mathcal{Z}(\lambda_t) = {\rm Tr} e^{-\beta\mathcal{H}(\lambda_t)}$ is the partition function and Tr denotes the trace over the system Hilbert space. The free energy is obtained from the partition function in the same way as for classical systems, i.e., $F(\lambda_t) = -\beta^{-1}\ln\mathcal{Z}(\lambda_t)$. Less obvious is the definition of work in quantum mechanics. Originally, Bochkov and Kuzovlev (1977) defined the exclusive quantum work, in analogy with the classical expression, Eqs. (12) and (13), as the operator $\mathcal{W}_0 = \int_0^\tau dt \lambda_t \dot{\mathcal{Q}}_t^H = \mathcal{H}_{0\tau}^H - \mathcal{H}_0$, where the superscript H denotes the Heisenberg picture:

$$\mathcal{B}_t^H = U_{t,0}^\dagger[\lambda]\mathcal{B}U_{t,0}[\lambda]. \tag{37}$$

Here $\mathcal{B}$ is an operator in the Schrödinger picture and $U_{t,0}[\lambda]$ is the unitary time evolution operator governed by the Schrödinger equation

$$i\hbar\frac{\partial U_{t,0}[\lambda]}{\partial t} = \mathcal{H}(\lambda_t)U_{t,0}[\lambda], \qquad U_{0,0}[\lambda] = \mathbb{1}, \tag{38}$$

with $\mathbb{1}$ denoting the identity operator. We use the notation $U_{t,0}[\lambda]$ to emphasize that, similar to the classical evolution $\varphi_{t,0}[\mathbf{z};\lambda]$ of Eq. (8), the quantum evolution operator is a

[6] See also Kobe (1981), in the context of nonrelativistic electrodynamics.

[7] See Horowitz and Jarzynski, 2008, Peliti, 2008a, 2008b, Vilar and Rubi, 2008a, 2008b, 2008c, Adib, 2009, Chen, 2008a, 2008b, Chen, 2009, Crooks, 2009, and Zimanyi and Silbey, 2009.

functional of the protocol λ.[8] The time derivative $\dot{\mathcal{Q}}_t^H$ is determined by the Heisenberg equation. In the case of a time-independent operator $\mathcal{Q}$ it becomes $\dot{\mathcal{Q}}_t^H = i[\mathcal{H}_t^H(\lambda_t), \mathcal{Q}_t^H]/\hbar$.

Bochkov and Kuzovlev (1977) were not able to provide any quantum analog of their fluctuation relations, Eqs. (14) and (15), with the classical work W_0 replaced by the operator $\mathcal{W}_0$.

Yukawa (2000) and Allahverdyan and Nieuwenhuizen (2005) arrived at a similar conclusion when attempting to define an inclusive work operator by $\mathcal{W} = \mathcal{H}_\tau^H(\lambda_\tau) - \mathcal{H}(\lambda_0)$. According to this definition the exponentiated work $\langle e^{-\beta \mathcal{W}}\rangle = \mathrm{Tr}\varrho_0 e^{-\beta\mathcal{W}} = \langle e^{-\beta[\mathcal{H}_\tau^H(\lambda_\tau)-\mathcal{H}(\lambda_0)]}\rangle$ agrees with $e^{-\beta\Delta F}$ only if $\mathcal{H}(\lambda_t)$ commutes at different times $[\mathcal{H}(\lambda_t), \mathcal{H}(\lambda_\tau)] = 0$ for any t, τ. This could lead to the premature conclusion that there exists no direct quantum analog of the Bochkov-Kuzovlev and the Jarzynski identities, Eqs. (15) and (23) (Allahverdyan and Nieuwenhuizen, 2005).

Based on the works by Kurchan (2000) and Tasaki (2000), Talkner *et al.* (2007) demonstrated that this conclusion is based on an erroneous assumption. They pointed out that work characterizes a process, rather than a state of the system; this is also an obvious observation from thermodynamics: unlike internal energy, work is not a state function (its differential is not exact). Consequently, work cannot be represented by a Hermitian operator whose eigenvalues can be determined in a single projective measurement. In contrast, the energy $\mathcal{H}(\lambda_t)$ (or $\mathcal{H}_0$, when the exclusive viewpoint is adopted) must be measured twice, first at the initial time $t = 0$ and again at the final time $t = \tau$.

The difference of the outcomes of these two measurements then yields the work performed on the system in a particular realization (Talkner *et al.*, 2007). That is, if at time $t = 0$ the eigenvalue $E_n^{\lambda_0}$ of $\mathcal{H}(\lambda_0)$ and later, at $t = \tau$, the eigenvalue $E_m^{\lambda_\tau}$ of $\mathcal{H}(\lambda_\tau)$ were obtained,[9] the measured (inclusive) work becomes

$$w = E_m^{\lambda_\tau} - E_n^{\lambda_0}. \tag{39}$$

Equation (39) represents the quantum version of the classical inclusive work, Eq. (19). In contrast to the classical case, this energy difference, which yields the work performed in a single realization of the protocol, cannot be expressed in the form of an integrated power, as in Eq. (20).

The quantum version of the exclusive work, Eq. (13), is $w_0 = e_m - e_n$ (Campisi *et al.*, 2011a), where now e_l are the eigenvalues of $\mathcal{H}_0$. As demonstrated in the next section, with these definitions of work straightforward quantum analogs of the Bochkov-Kuzovlev results, Eqs. (14), (15), and (18) and of their inclusive viewpoint counterparts, Eqs. (21), (23), and (28) can be derived.

IV. QUANTUM WORK FLUCTUATION RELATIONS

Armed with all the proper mathematical definitions of nonequilibrium quantum mechanical work, Eq. (39), we next embark on the study of work fluctuation relations in quantum systems. As in the classical case, in the quantum case one also needs to be careful in properly identifying the exclusive and inclusive work, and must be aware of the gauge freedom issue. In the following we adopt, except when otherwise explicitly stated, the inclusive viewpoint. The two fundamental ingredients for the development of the theory are, in the quantum case as in the classical case, the canonical form of equilibrium and microreversibility.

A. Microreversibility of nonautonomous quantum systems

The principle of microreversibility is introduced and discussed in quantum mechanics textbooks for autonomous (i.e., nondriven) quantum systems (Messiah, 1962). As in the classical case, however, this principle continues to hold in a more general sense also for nonautonomous quantum systems. In this case it can be expressed as

$$U_{t,\tau}[\lambda] = \Theta^\dagger U_{\tau-t,0}[\tilde{\lambda}]\Theta, \tag{40}$$

where Θ is the quantum mechanical time-reversal operator (Messiah, 1962).[10] Note that the presence of the protocol λ and its time-reversed image $\tilde{\lambda}$ distinguishes this generalized version from the standard form of microreversibility for autonomous systems. The principle of microreversibility, Eq. (40), holds under the assumption that at any time t the Hamiltonian is invariant under time reversal,[11] that is,

$$\mathcal{H}(\lambda_t)\Theta = \Theta\mathcal{H}(\lambda_t). \tag{41}$$

A derivation of Eq. (40) is presented in Appendix B. See also Andrieux and Gaspard (2008) for an alternative derivation.

In order to better understand the physics behind Eq. (40) we rewrite it as $U_{t,0}[\lambda] = \Theta^\dagger U_{\tau-t,0}[\tilde{\lambda}]\Theta U_{\tau,0}[\lambda]$, where we used the concatenation rule $U_{t,\tau}[\lambda] = U_{t,0}[\lambda]U_{0,\tau}[\lambda]$, and the inverse $U_{0,\tau}[\lambda] = U_{\tau,0}^{-1}[\lambda]$ of the propagator $U_{t,s}[\lambda]$. Applying it to a pure state $|i\rangle$, and multiplying by Θ from the left, we obtain

$$\Theta|\psi_t\rangle = U_{\tau-t,0}[\tilde{\lambda}]\Theta|f\rangle, \tag{42}$$

where we introduced the notations $|\psi_t\rangle = U_{t,0}[\lambda]|i\rangle$ and $|f\rangle = U_{\tau,0}[\lambda]|i\rangle$. Equation (42) says that, under the evolution generated by the reversed protocol $\tilde{\lambda}$ the time-reversed final state $\Theta|f\rangle$ evolves between time 0 and $\tau - t$, to $\Theta|\psi_t\rangle$. This is shown in Fig. 3. As for the classical case, in order to trace a nonautonomous system back to its initial state, one needs not only to invert the momenta (applying Θ), but also to invert the temporal sequence of Hamiltonian values.

B. The work probability density function

We consider a system described by the Hamiltonian $\mathcal{H}(\lambda_t)$ initially prepared in the canonical state

[8] Because of causality $U_{t,0}[\lambda]$ may, of course, depend only on the part of the protocol including times from 0 up to t.

[9] For a formal definition of these eigenvalues see Eq. (43).

[10] Under the action of Θ coordinates transform evenly, whereas linear and angular momenta, as well as spins change sign. In the coordinate representation, in absence of spin degrees of freedom, the operator Θ is the complex conjugation $\Theta\psi = \psi^*$.

[11] In the presence of external magnetic fields the direction of these fields has also to be inverted in the same way as in the autonomous case.

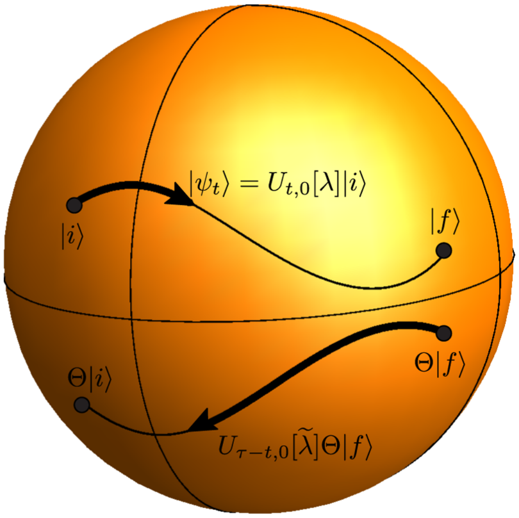

FIG. 3 (color online). Microreversibility for nonautonomous quantum systems. The normalized initial condition $|i\rangle$ evolves, under the unitary time evolution generated by $\mathcal{H}(\lambda_t)$, from time $t=0$ until t to $|\psi_t\rangle = U_{t,0}[\lambda]|i\rangle$ and until time $t=\tau$ to $|f\rangle$. The time-reversed final condition $\Theta|f\rangle$ evolves, under the unitary evolution generated by $\tilde{\lambda}$ from time $t=0$ until $\tau - t$ to $U_{\tau-t,0}[\tilde{\lambda}]\Theta|f\rangle = \Theta|\psi_t\rangle$, Eq. (42), and until time $t=\tau$ to the time-reversed initial condition $\Theta|i\rangle$. The motion occurs on the hypersphere of unitary radius in the Hilbert space.

$\varrho(\lambda_0) = e^{-\beta\mathcal{H}(\lambda_0)}/Z(\lambda_0)$. The instantaneous eigenvalues of $\mathcal{H}(\lambda_t)$ are denoted by $E_n^{\lambda_t}$, and the corresponding instantaneous eigenstates by $|\psi_{n,\gamma}^{\lambda_t}\rangle$:

$$\mathcal{H}(\lambda_t)|\psi_{n,\gamma}^{\lambda_t}\rangle = E_n^{\lambda_t}|\psi_{n,\gamma}^{\lambda_t}\rangle. \tag{43}$$

The symbol n labels the quantum number specifying the energy eigenvalues and γ denotes all further quantum numbers, necessary to specify an energy eigenstate in case of g_n-fold degeneracy. We emphasize that the instantaneous eigenvalue Eq. (43) must not be confused with the Schrödinger equation $i\hbar\partial_t|\psi(t)\rangle = \mathcal{H}(\lambda_t)|\psi(t)\rangle$. The instantaneous eigenfunctions resulting from Eq. (43), in particular, are not solutions of the Schrödinger equation.

At $t=0$ the first measurement of $\mathcal{H}(\lambda_0)$ is performed, with outcome $E_n^{\lambda_0}$. This occurs with probability

$$p_n^0 = g_n e^{-\beta E_n^{\lambda_0}}/Z(\lambda_0). \tag{44}$$

According to the postulates of quantum mechanics, immediately after the measurement of the energy $E_n^{\lambda_0}$ the system is found in the state

$$\varrho_n = \Pi_n^{\lambda_0}\varrho(\lambda_0)\Pi_n^{\lambda_0}/p_n^0, \tag{45}$$

where $\Pi_n^{\lambda_0} = \sum_\gamma|\psi_{n,\gamma}^{\lambda_0}\rangle\langle\psi_{n,\gamma}^{\lambda_0}|$ is the projector onto the eigenspace spanned by the eigenvectors belonging to the eigenvalue $E_n^{\lambda_0}$. The system is assumed to be thermally isolated at any time $t \geq 0$, so that its evolution is determined by the unitary operator $U_{t,0}[\lambda]$, Eq. (38); hence it evolves according to

$$\varrho_n(t) = U_{t,0}[\lambda]\varrho_n U_{t,0}^\dagger[\lambda]. \tag{46}$$

At time τ a second measurement of $\mathcal{H}(\lambda_\tau)$ yielding the eigenvalue $E_m^{\lambda_\tau}$ with probability

$$p_{m|n}[\lambda] = \mathrm{Tr}\Pi_m^{\lambda_\tau}\varrho_n(\tau) \tag{47}$$

is performed. The PDF to observe the work w is thus given by

$$p[w;\lambda] = \sum_{m,n}\delta(w - [E_m^{\lambda_\tau} - E_n^{\lambda_0}])p_{m|n}[\lambda]p_n^0. \tag{48}$$

The work PDF has been calculated explicitly for a forced harmonic oscillator (Talkner, Burada, and Hänggi, 2008, 2009) and for a parametric oscillator with varying frequency (Deffner and Lutz, 2008; Deffner *et al.*, 2010).

C. The characteristic function of work

The characteristic function of work $G[u;\lambda]$ is defined as in the classical case as the Fourier transform of the work probability density function

$$G[u;\lambda] = \int dw e^{iuw}p[w;\lambda]. \tag{49}$$

Similar to $p[w;\lambda]$, $G[u;\lambda]$ contains full information regarding the statistics of the random variable w. Talkner *et al.* (2007) showed that the work PDF has the form of a two-time nonstationary quantum correlation function, i.e.,

$$\begin{aligned} G[u;\lambda] &= \langle e^{iu\mathcal{H}_\tau^H(\lambda_\tau)}e^{-iu\mathcal{H}(\lambda_0)}\rangle \\ &= \mathrm{Tr}e^{iu\mathcal{H}_\tau^H(\lambda_\tau)}e^{-(iu+\beta)\mathcal{H}(\lambda_0)}/Z(\lambda_0), \end{aligned} \tag{50}$$

where the average symbol stands for quantum expectation over the initial state density matrix $\varrho(\lambda_0)$, Eq. (36), i.e., $\langle\mathcal{B}\rangle = \mathrm{Tr}\varrho(\lambda_0)\mathcal{B}$, and the superscript H denotes the Heisenberg picture, i.e., $\mathcal{H}_\tau^H(\lambda_\tau) = U_{\tau,0}^\dagger[\lambda]\mathcal{H}(\lambda_\tau)U_{\tau,0}[\lambda]$.

Equation (50) was derived first by Talkner *et al.* (2007) in the case of nondegenerate $\mathcal{H}(\lambda_t)$ and later generalized by Talkner, Hänggi, and Morillo (2008) to the case of possibly degenerate $\mathcal{H}(\lambda_t)$.[12]

The product of the two exponential operators $e^{iu\mathcal{H}_\tau^H(\lambda_\tau)}e^{-iu\mathcal{H}(\lambda_0)}$ can be combined into a single exponent under the protection of the time ordering operator $\mathcal{T}$ to yield $e^{iu\mathcal{H}_\tau^H(\lambda_\tau)}e^{-iu\mathcal{H}(\lambda_0)} = \mathcal{T}e^{iu[\mathcal{H}_\tau^H(\lambda_\tau)-\mathcal{H}(\lambda_0)]}$. In this way one can convert the characteristic function of work to a form that is analogous to the corresponding classical expression, Eq. (27),

$$\begin{aligned} G[u;\lambda] &= \mathrm{Tr}\mathcal{T}e^{iu[\mathcal{H}_\tau^H(\lambda_\tau)-\mathcal{H}(\lambda_0)]}e^{-\beta\mathcal{H}(\lambda_0)}/Z(\lambda_0) \\ &= \cancel{\mathrm{Tr}\mathcal{T}\exp\left[iu\int_0^\tau dt\dot{\lambda}_t\frac{\partial\mathcal{H}_t^H(\lambda_t)}{\partial\lambda_t}\right]e^{-\beta\mathcal{H}(\lambda_0)}/Z_0}. \end{aligned} \tag{51}$$

[12]The derivation in Talkner, Hänggi, and Morillo (2008) is more general in that it does not assume any special form of the initial state, thus allowing the study, e.g., of microcanonical fluctuation relations. The formal expression, Eq. (50), remains valid for any initial state ϱ, with the provision that the average is taken with respect to $\bar{\varrho} = \sum_n\Pi_n^{\lambda_0}\varrho\Pi_n^{\lambda_0}$ representing the diagonal part of ρ in the eigenbasis of $\mathcal{H}(\lambda_0)$.

~~The second equality follows from the fact that the total time derivative of the Hamiltonian in the Heisenberg picture coincides with its partial derivative.~~

As a consequence of quantum microreversibility, Eq. (40), the characteristic function of work obeys the following important symmetry relation (see Appendix C):

$$\mathcal{Z}(\lambda_0)G[u;\lambda] = \mathcal{Z}(\lambda_\tau)G[-u+i\beta;\tilde{\lambda}]. \tag{52}$$

By applying the inverse Fourier transform and using $\mathcal{Z}(\lambda_t) = \mathrm{Tr}e^{-\beta\mathcal{H}(\lambda_t)} = e^{-\beta F(\lambda_t)}$ one ends up with the quantum version of the Crooks fluctuation theorem in Eq. (28):

$$\frac{p[w;\lambda]}{p[-w;\tilde{\lambda}]} = e^{\beta(w-\Delta F)}. \tag{53}$$

This result was first accomplished by Kurchan (2000) and Tasaki (2000). Later Talkner and Hänggi (2007) gave a systematic derivation based on the characteristic function of work. The quantum Jarzynski equality,

$$\langle e^{-\beta w}\rangle_\lambda = e^{-\beta\Delta F}, \tag{54}$$

follows by multiplying both sides by $p[-w;\tilde{\lambda}]e^{-\beta w}$ and integrating over w. Given the fact that the characteristic function is determined by a two-time quantum correlation function rather than by a single time expectation value is another clear indication that work is not an observable but instead characterizes a process.

As discussed by Campisi *et al.* (2010a) the Tasaki-Crooks relation, Eq. (53), and the quantum version of the Jarzynski equality, Eq. (54), continue to hold even if further projective measurements of any observable $\mathcal{A}$ are performed within the protocol duration $(0,\tau)$. These measurements, however, do alter the work PDF (Campisi *et al.*, 2011b).

D. Quantum generating functional

The Jarzynski equality can also immediately be obtained from the characteristic function by setting $u = i\beta$, in Eq. (50) (Talkner *et al.*, 2007). In order to obtain this result it is important that the Hamiltonian operators at initial and final times enter into the characteristic function, Eq. (50), as arguments of two factorizing exponential functions, i.e., in the form $e^{-\beta\mathcal{H}^H(\lambda_\tau)}e^{\beta\mathcal{H}(\lambda_0)}$. In general, this, of course, is different from a single exponential $e^{-\beta[\mathcal{H}^H(\lambda_\tau)-\mathcal{H}(\lambda_0)]}$. In the definitions of generating functionals, Bochkov and Kuzovlev (1977, 1981a) and Stratonovich (1994) employed yet different ordering prescriptions which do not lead to the Jarzynski equality. In order to maintain the structure of the classical generating functional, Eq. (21), for quantum systems the classical exponentiated work $e^{-\beta W}$ also has to be replaced by the product of exponentials as it appears in the characteristic function of work, Eq. (50). This then leads to a desired generating functional relation

$$\left\langle \exp\left[\int_0^\tau dt\, u_t \mathcal{B}_t^H\right] e^{-\beta\mathcal{H}^H(\lambda_\tau)} e^{\beta\mathcal{H}(\lambda_0)}\right\rangle_\lambda = \left\langle \exp\left[\int_0^\tau dt\,\tilde{u}_t \varepsilon_{\mathcal{B}} \mathcal{B}_t^H\right]\right\rangle_{\tilde{\lambda}} e^{-\beta\Delta F}, \tag{55}$$

where $\mathcal{B}$ is an observable with definite parity $\varepsilon_{\mathcal{B}}$ (i.e., $\Theta\mathcal{B}\Theta^\dagger = \varepsilon_{\mathcal{B}}\mathcal{B}$), $\mathcal{B}_t^H$ denotes the observable $\mathcal{B}$ in the Heisenberg representation, Eq. (37), u_t is a real function, and $\tilde{u}_t = u_{\tau-t}$. This can be proved by using the quantum microreversibility principle, Eq. (40), in a similar way as in the classical derivation, Eq. (A1).

The derivation of Eq. (55) was provided by Andrieux and Gaspard (2008), who also recovered the formula of Kubo, Eq. (4), and the Onsager-Casimir reciprocity relations.[13]

These relations are obtained by means of functional derivatives of Eq. (55) with respect to the force fields λ_t and test fields u_t at $\lambda_t = u_t = 0$ (Andrieux and Gaspard, 2008). Relations and symmetries for higher order response functions follow in an analogous way as in the classical case, Eq. (14), by means of higher order functional derivatives with respect to the force field λ_t. Such relations were investigated experimentally by Nakamura *et al.* (2010, 2011); see Sec. V, Eq. (89).

Within the exclusive viewpoint approach, the counterparts of Eqs. (50) and (52)–(55) are obtained by replacing $\mathcal{H}^H$ with $\mathcal{H}_0^H$, $\mathcal{H}(\lambda_0)$ with $\mathcal{H}_0$, $\mathcal{Z}(\lambda_t)$ with $\mathcal{Z}_0 = \mathrm{Tr}e^{-\beta\mathcal{H}_0}$, and setting accordingly ΔF to 0 (Campisi *et al.*, 2011a).

E. Microreversibility, conditional probabilities, and entropy

For a Hamiltonian $\mathcal{H}(\lambda_t)$ with nondegenerate instantaneous spectrum for all times t and instantaneous eigenvectors $|\psi_n^{\lambda_t}\rangle$, the conditional probability $p_{m|n}[\lambda]$, Eq. (47), is given by the simple expression $p_{m|n}[\lambda] = |\langle\psi_m^{\lambda_\tau}|U_{\tau,0}[\lambda]|\psi_n^{\lambda_0}\rangle|^2$. As a consequence of the assumed invariance of the Hamiltonian, Eq. (41), the eigenstates $|\psi_n^{\lambda_t}\rangle$ are invariant under the action of the time-reversal operator up to a phase $\Theta|\psi_n^{\lambda_t}\rangle = e^{i\varphi_n^t}|\psi_n^{\lambda_t}\rangle$. Then, expressing microreversibility as $U_{\tau,0}[\lambda] = \Theta^\dagger U_{0,\tau}[\tilde{\lambda}]\Theta$, see Eq. (40), one obtains the following symmetry relation for the conditional probabilities:

$$p_{m|n}[\lambda] = p_{n|m}[\tilde{\lambda}]. \tag{56}$$

Note the exchanged position of m and n in the two sides of this equation. From Eq. (56) the Tasaki-Crooks fluctuation theorem is readily obtained for a canonical initial state, using Eq. (48).

Using instead an initial microcanonical state at energy E, described by the density matrix[14]

$$\rho_0(E) = \delta(E-\mathcal{H}(\lambda_0))/\omega(E,\lambda_0), \tag{57}$$

where $\omega(E,\lambda_t) = \mathrm{Tr}\delta(E-\mathcal{H}(\lambda_t))$, we obtain (Talkner, Hänggi, and Morillo, 2008)

$$\frac{p[E,W;\lambda]}{p[E+W,-W;\tilde{\lambda}]} = e^{[S(E+W,\lambda_\tau)-S(E,\lambda_0)]/k_B}, \tag{58}$$

[13] Andrieux and Gaspard (2008) also allowed for a possible dependence of the Hamiltonian on a magnetic field $\mathcal{H} = \mathcal{H}(\lambda_t,\mathbf{B})$. Then it is meant that the dynamical evolution of $\mathcal{B}$ on the right-hand side is governed by the Hamiltonian $\mathcal{H}(\tilde{\lambda}_t,-\mathbf{B})$, i.e., besides inverting the protocol, the magnetic field needs to be inverted as well.

[14] Strictly speaking, in order to obtain well-defined expressions the δ function has to be understood as a sharply peaked function with infinite support.

where $S(E, \lambda_t) = k_B \ln\omega(E, \lambda_t)$ denotes Boltzmann's thermodynamic equilibrium entropy. The corresponding classical derivation was provided by Cleuren *et al.* (2006). A classical microcanonical version of the Jarzynski equality was put forward by Adib (2005) for non-Hamiltonian isoenergetic dynamics. It was recently generalized to energy controlled systems by Katsuda and Ohzeki (2011).

F. Weak-coupling case

In Secs. IV.B–IV.D we studied a quantum mechanical system at canonical equilibrium at time $t = 0$. During the subsequent action of the protocol it is assumed to be completely isolated from its surrounding apart from the influence of the external work source and hence to undergo a unitary time evolution. The quality of this approximation depends on the relative strength of the interaction between the system and its environment, compared to typical energies of the isolated system as well as on the duration of the protocol. In general, though, a treatment that takes into account possible environmental interactions is necessary. As will be shown, the interaction with a thermal bath does not lead to a modification of the Jarzynski equality, Eq. (54), nor of the quantum work fluctuation relation, Eq. (53), in both the cases of weak and strong coupling (Campisi *et al.*, 2009a; Talkner, Campisi, and Hänggi, 2009), a main finding which holds true as well for classical systems (Jarzynski, 2004). In this section we address the weak-coupling case, while the more intricate case of strong coupling is discussed in the next section.

We consider a driven system described by the time-dependent Hamiltonian $\mathcal{H}_S(\lambda_t)$, in contact with a thermal bath with time-independent Hamiltonian $\mathcal{H}_B$; see Fig. 4. The Hamiltonian of the compound system is

$$\mathcal{H}(\lambda_t) = \mathcal{H}_S(\lambda_t) + \mathcal{H}_B + \mathcal{H}_{SB}, \tag{59}$$

where the energy contribution stemming from $\mathcal{H}_{SB}$ is assumed to be much smaller than the energies of the system and bath resulting from $\mathcal{H}_S(\lambda_t)$ and $\mathcal{H}_B$. The parameter λ that is manipulated according to a protocol solely enters in the system Hamiltonian $\mathcal{H}_S(\lambda_t)$.

The compound system is assumed to be initially ($t = 0$) in the canonical state

$$\varrho(\lambda_0) = e^{-\beta\mathcal{H}(\lambda_0)}/Y(\lambda_0), \tag{60}$$

where $Y(\lambda_t) = \mathrm{Tr}e^{-\beta\mathcal{H}(\lambda_t)}$ is the corresponding partition function. This initial state may be provided by contact with a superbath at inverse temperature β; see Fig. 4. It is then assumed either that the contact to the superbath is removed for $t \geq 0$ or that the superbath is so weakly coupled to the compound system that it bears no influence on its dynamics over the time span 0 to τ.

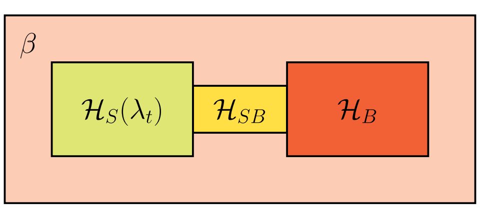

FIG. 4 (color online). Driven open system. A driven system [represented by its Hamiltonian $\mathcal{H}_S(\lambda_t)$] is coupled to a bath (represented by $\mathcal{H}_B$) via the interaction Hamiltonian $\mathcal{H}_{SB}$. The compound system is in vanishingly weak contact with a superbath that provides the initial canonical state at inverse temperature β, Eq. (60).

Because the system and the environmental Hamiltonians commute with each other, their energies can be simultaneously measured. We denote the eigenvalues of $\mathcal{H}_S(\lambda_t)$ as $E_i^{\lambda_t}$, and those of $\mathcal{H}_B$ as E_α^B. In analogy with the isolated case we assume that at time $t = 0$ a joint measurement of $\mathcal{H}_S(\lambda_0)$ and $\mathcal{H}_B$ is performed, with outcomes $E_n^{\lambda_0}$ and E_ν^B. A second joint measurement of $\mathcal{H}_S(\lambda_\tau)$ and $\mathcal{H}_B$ at $t = \tau$ yields the outcomes $E_m^{\lambda_\tau}$ and E_μ^B.

In analogy to the energy change of an isolated system, the differences of the eigenvalues of system and bath Hamiltonians yield the energy changes of system and bath, ΔE and ΔE^B, respectively, in a single realization of the protocol, i.e.,

$$\Delta E = E_m^{\lambda_\tau} - E_n^{\lambda_0}, \tag{61}$$

$$\Delta E^B = E_\mu^B - E_\nu^B. \tag{62}$$

In the weak-coupling limit, the change of the energy content of the total system is given by the sum of the energy changes of the system and bath energies apart from a negligibly small contribution due to the interaction Hamiltonian $\mathcal{H}_{SB}$. The work w performed on the system coincides with the change of the total energy because the force is assumed to act only directly on the system. For the same reason, the change of the bath energy is due solely to an energy exchange with the system and hence can be interpreted as negative heat $\Delta E^B = -Q$. Accordingly we have[15]

$$\Delta E = w + Q. \tag{63}$$

Following the analogy with the isolated case, we considered the joint probability distribution function $p[\Delta E, Q; \lambda]$ that the system energy changes by ΔE and the heat Q is exchanged, under the protocol λ:

$$p[\Delta E, Q; \lambda] = \sum_{m,n,\mu,\nu} \delta(\Delta E - E_m^{\lambda_\tau} + E_n^{\lambda_0}) \times \delta(Q + E_\mu^B - E_\nu^B) p_{m\mu|n\nu}[\lambda] p_{n\nu}^0, \tag{64}$$

where $p_{m\mu|n\nu}[\lambda]$ is the conditional probability to obtain the outcome $E_m^{\lambda_\tau}$, E_μ^B at τ, provided that the outcome $E_n^{\lambda_0}$, E_ν^B was obtained at time $t = 0$, whereas $p_{n\nu}^0$ is the probability to find the outcome $E_n^{\lambda_0}$, E_ν^B in the first measurement. The conditional probability $p_{m\mu|n\nu}[\lambda]$ can be expressed in terms of the projectors on the common eigenstates of $\mathcal{H}_S(\lambda_t)$, $\mathcal{H}_B$, and the unitary evolution generated by the total Hamiltonian $\mathcal{H}(\lambda_t)$ (Talkner, Campisi, and Hänggi, 2009).

[15]By use of the probability distribution in Eq. (64), the averaged quantity $\langle\Delta E\rangle_\lambda = \int d(\Delta E) dQ p[\Delta E, Q; \lambda]\Delta E = \mathrm{Tr}\varrho_\tau\mathcal{H}_S(\lambda_\tau) - \mathrm{Tr}\varrho(\lambda_0)\mathcal{H}_S(\lambda_0)$ cannot, in general, be interpreted as a change in thermodynamic internal energy. The reason is that the final state, ϱ_τ, reached at the end of the protocol is typically not a state of thermodynamic equilibrium; hence its thermodynamic internal energy is not defined.

By taking the Fourier transform of $p[\Delta E, Q; \lambda]$ with respect to both ΔE and Q, one obtains the characteristic function of system energy change and heat exchange, reading

$$G[u, v; \lambda] = \int d(\Delta E) dQ e^{i(u\Delta E + vQ)} p[\Delta E, Q; \lambda], \tag{65}$$

which can be further simplified and cast, as in the isolated case, in the form of a two-time quantum correlation function (Talkner *et al.*, 2009):

$$G[u, v; \lambda] = \langle e^{i[u\mathcal{H}^H_{S\tau}(\lambda_\tau) - v\mathcal{H}^H_{B\tau}]} e^{-i[u\mathcal{H}_S(\lambda_0) - v\mathcal{H}_B]} \rangle, \tag{66}$$

where the average is over the state $\bar{\varrho}(\lambda_0)$ that is the diagonal part of $\varrho(\lambda_0)$, Eq. (60), with respect to $\{\mathcal{H}_S(\lambda_0), \mathcal{H}_B\}$. Notably, in the limit of weak coupling this state $\bar{\varrho}(\lambda_0)$ approximately factorizes into the product of the equilibrium states of system and bath with the deviations being of second order in the system-bath interaction (Talkner, Campisi, and Hänggi, 2009).

Using Eq. (66) in combination with microreversibility, Eq. (40), leads, in analogy with Eq. (52), to

$$\mathcal{Z}_S(\lambda_0) G[u, v; \lambda] = \mathcal{Z}_S(\lambda_\tau) G[-u + i\beta, -v - i\beta; \tilde{\lambda}], \tag{67}$$

where

$$\mathcal{Z}_S(\lambda_t) = \mathrm{Tr}_S e^{-\beta\mathcal{H}_S(\lambda_t)}, \tag{68}$$

with Tr_S denoting the trace over the system Hilbert space. Upon applying an inverse Fourier transform of Eq. (67) one arrives at the following relation:

$$\frac{p[\Delta E, Q; \lambda]}{p[-\Delta E, -Q; \tilde{\lambda}]} = e^{\beta(\Delta E - Q - \Delta F_S)}, \tag{69}$$

where

$$\Delta F_S = -\beta^{-1} \ln[\mathcal{Z}_S(\lambda_\tau)/\mathcal{Z}_S(\lambda_0)] \tag{70}$$

denotes the system free energy difference. Equation (69) generalizes the Tasaki-Crooks fluctuation theorem, Eq. (53), to the case where the system can exchange heat with a thermal bath.

Performing the change of the variable $\Delta E \to w$, Eq. (63), in Eq. (69) leads to the following fluctuation relation for the joint probability density function of work and heat:

$$\frac{p[w, Q; \lambda]}{p[-w, -Q; \tilde{\lambda}]} = e^{\beta(w - \Delta F_S)}. \tag{71}$$

Notably, the right-hand side does not depend on the heat Q but depends on the work w only. This fact implies that the marginal probability density of work $p[w; \lambda] = \int dQ p[w, Q; \lambda]$ obeys the Tasaki-Crooks relation:

$$\frac{p[w; \lambda]}{p[-w; \tilde{\lambda}]} = e^{\beta(w - \Delta F_S)}. \tag{72}$$

Subsequently the Jarzynski equality $\langle e^{-\beta w} \rangle = e^{-\beta \Delta F_S}$ is also satisfied. Thus, the fluctuation relation, Eq. (53), and the Jarzynski equality, Eq. (54), keep holding, unaltered, also in the case of weak coupling. This result was originally found upon assuming a Markovian quantum dynamics for the reduced system dynamics S.[16] With the above derivation we followed Talkner, Campisi, and Hänggi (2009) in which one does not rely on a Markovian quantum evolution and consequently the results hold true as well for a general non-Markovian reduced quantum dynamics of the system S.

G. Strong-coupling case

In the case of strong coupling, the system-bath interaction energy is non-negligible, and therefore it is no longer possible to identify the heat as the energy change of the bath. How to define heat in a strongly coupled driven system and whether it is possible to define it at all currently present open problems. This, however, does not hinder the possibility to prove that the work fluctuation relation, Eq. (72), remains valid also in the case of strong coupling. For this purpose it suffices to properly identify the work w done on and the free energy F_S of an open system, without entering the issue of what heat means in a strong-coupling situation. As for the classical case (see Sec. III.A), the system Hamiltonian $\mathcal{H}_S(\lambda_t)$ is the only time-dependent part of the total Hamiltonian $\mathcal{H}(\lambda_t)$. Therefore, the work done on the open quantum system coincides with the work done on the total system, as in the weak-coupling case treated in Sec. IV.F. Consequently, the work done on an open quantum system in a single realization is

$$w = \mathcal{E}^{\lambda_\tau}_m - \mathcal{E}^{\lambda_0}_n, \tag{73}$$

where $\mathcal{E}^{\lambda_t}_l$ are the eigenvalues of the total Hamiltonian $\mathcal{H}(\lambda_t)$.

Regarding the proper identification of the free energy of an open quantum system, the situation is more involved because the bare partition function $\mathcal{Z}^0_S(\lambda_t) = \mathrm{Tr}_S e^{-\beta\mathcal{H}_S(\lambda_t)}$ cannot take into account the full effect of the environment in any case other than the limiting situation of weak coupling. For strong coupling the equilibrium statistical mechanical description has to be based on a partition function of the open quantum system that is given as the ratio of the partition functions of the total system and the isolated environment,[17] i.e.,

$$\mathcal{Z}_S(\lambda_t) = Y(\lambda_t)/\mathcal{Z}_B, \tag{74}$$

where $\mathcal{Z}_B = \mathrm{Tr}_B e^{-\beta\mathcal{H}_B}$ and $Y(\lambda_t) = \mathrm{Tr} e^{-\beta\mathcal{H}(\lambda_t)}$ with Tr_B, and Tr denoting the traces over the bath Hilbert space and the total Hilbert space, respectively. It must be stressed that, in general, the partition function $\mathcal{Z}_S(\lambda_t)$ of an open quantum system differs from its partition function in absence of a bath:

$$\mathcal{Z}_S(\lambda_t) \neq \mathrm{Tr}_S e^{-\beta\mathcal{H}_S(\lambda_t)}. \tag{75}$$

[16]See Mukamel (2003), de Roeck and Maes (2004), Esposito and Mukamel (2006), and Crooks (2008).

[17]See Feynman and Vernon (1963), Caldeira and Leggett (1983), Grabert *et al.* (1984), Ford *et al.* (1985), Grabert *et al.* (1988), Dittrich *et al.* (1998), Ingold (2002), Nieuwenhuizen and Allahverdyan (2002), Hänggi and Ingold (2005, 2006), Hörhammer and Büttner (2008), Campisi *et al.* (2009a, 2009b), and Campisi, Zueco, and Talkner (2010).

The equality is restored, though, in the limit of a weak coupling.

The free energy of an open quantum system follows according to the standard rule of equilibrium statistical mechanics as

$$F_S(\lambda_t) = F(\lambda_t) - F_B = -\frac{1}{\beta}\ln \mathcal{Z}_S(\lambda_t). \tag{76}$$

In this way the influences of the bath on the thermodynamic properties of the system are properly taken into account. Besides, Eq. (76) complies with all the grand laws of thermodynamics (Campisi *et al.*, 2009a).

For a total system initially prepared in the Gibbs state, Eq. (60), the Tasaki-Crooks fluctuation theorem, Eq. (53), applies yielding

$$\frac{p[w;\lambda]}{p[-w;\tilde{\lambda}]} = \frac{Y(\lambda_\tau)}{Y(\lambda_0)}e^{\beta w}. \tag{77}$$

Since $\mathcal{Z}_B$ does not depend on time, the salient relation

$$Y(\lambda_\tau)/Y(\lambda_0) = \mathcal{Z}_S(\lambda_\tau)/\mathcal{Z}_S(\lambda_0) \tag{78}$$

holds, leading to

$$\frac{p[w;\lambda]}{p[-w;\tilde{\lambda}]} = \frac{\mathcal{Z}_S(\lambda_\tau)}{\mathcal{Z}_S(\lambda_0)}e^{\beta w} = e^{\beta(w-\Delta F_S)}, \tag{79}$$

where $\Delta F_S = F_S(\lambda_\tau) - F_S(\lambda_0)$ is the proper free energy difference of an open quantum system. Because w coincides with the work performed on the open system, both the Tasaki-Crooks relation, Eq. (72), and the Jarzynski equality, Eq. (54), are recovered also in the case of strong coupling (Campisi *et al.*, 2009a).

V. QUANTUM EXCHANGE FLUCTUATION RELATIONS

The transport of energy and matter between two reservoirs that stay at different temperatures and chemical potentials represents an important experimental setup (see also Sec. VI), as well as a central problem of nonequilibrium thermodynamics (de Groot and Mazur, 1984). Here the two-measurement scheme described above in conjunction with the principle of microreversibility leads to fluctuation relations similar to the Tasaki-Crooks relation, Eq. (53), for the probabilities of energy and matter exchanges. The resulting fluctuation relations have been referred to as "exchange fluctuation theorems" (Jarzynski and Wójcik, 2004), to distinguish them from the "work fluctuation theorems."

The first quantum exchange fluctuation theorem was put forward by Jarzynski and Wójcik (2004). It applies to two systems initially at different temperatures that are allowed to interact over the lapse of time $(0, \tau)$, via a possibly time-dependent interaction. This situation was later generalized by Saito and Utsumi (2008) and Andrieux *et al.* (2009), to allow for the exchange of energy and particles between several interacting systems initially at different temperatures and chemical potentials; see Fig. 5.

The total Hamiltonian $\mathcal{H}(\mathcal{V}_t)$ consisting of s subsystems is

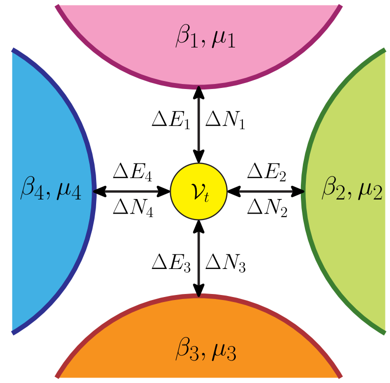

FIG. 5 (color online). Exchange fluctuation relation setup. Several reservoirs (large semicircles) interact via the coupling $\mathcal{V}_t$ (symbolized by the small circle), which is switched on at time $t = 0$ and switched off at $t = \tau$. During the on period $(0, \tau)$ the reservoirs exchange energy and matter with each other. The resulting net energy change of the ith reservoir is ΔE_i and its particle content changes by ΔN_i. Initially the reservoirs prescribed temperatures $T_i = (k_B\beta_i)^{-1}$, and chemical potentials μ_i of the particle species that are exchanged, respectively.

$$\mathcal{H}(\mathcal{V}_t) = \sum_{i=1}^{s}\mathcal{H}_i + \mathcal{V}_t, \tag{80}$$

where $\mathcal{H}_i$ is the Hamiltonian of the ith system, and $\mathcal{V}_t$ describes the interaction between the subsystems, which sets in at time $t = 0$ and ends at time $t = \tau$. Consequently, $\mathcal{V}_t = 0$ for $t \notin (0, \tau)$, and, in particular, $\mathcal{V}_0 = \mathcal{V}_\tau = 0$. As before, it is important to distinguish between the values $\mathcal{V}_t$ at a specific time and the whole protocol $\mathcal{V}$.

Initially, the subsystems are supposed to be isolated from each other and to stay in a factorized grand-canonical state

$$\varrho_0 = \prod_i \varrho_i = \prod_i e^{-\beta_i[\mathcal{H}_i - \mu_i\mathcal{N}_i]}/\Xi_i, \tag{81}$$

with μ_i, β_i, and $\Xi_i = \mathrm{Tr}_i e^{-\beta_i(\mathcal{H}_i - \mu_i\mathcal{N}_i)}$ the chemical potential, inverse temperature, and grand potential, respectively, of subsystem i. Here $\mathcal{N}_i$ and Tr_i denote the particle number operator and the trace of the ith subsystem, respectively.

We also assume that in the absence of interaction the particle numbers in each subsystem are conserved, i.e., $[\mathcal{H}_i, \mathcal{N}_i] = 0$. Since operators acting on Hilbert spaces of different subsystems commute, we find $[\mathcal{H}_i, \mathcal{N}_j] = 0$, $[\mathcal{N}_i, \mathcal{N}_j] = 0$, and $[\mathcal{H}_i, \mathcal{H}_j] = 0$ for any i, j. Accordingly, one may measure all the $\mathcal{H}_i$'s and all the $\mathcal{N}_i$'s simultaneously. Adopting the two-measurement scheme discussed above in the context of the work fluctuation relation, we make a first measurement of all the $\mathcal{H}_i$'s and all the $\mathcal{N}_i$'s at $t = 0$. Accordingly, the wave function collapses onto a common eigenstate $|\psi_n\rangle$ of all these observables with eigenvalues E_n^i and N_n^i. Subsequently, this wave function evolves according to the evolution $U_{t,0}[\mathcal{V}]$ generated by the total Hamiltonian, until time τ when a second measurement of all $\mathcal{H}_i$'s and $\mathcal{N}_i$'s is performed leading to a wave function collapse onto an eigenstate $|\psi_m\rangle$, with eigenvalues E_m^i and N_m^i. As in the case studied in Sec. IV.F, the joint probability density of energy and particle exchanges $p[\Delta\mathbf{E}, \Delta\mathbf{N}; \mathcal{V}]$ completely describes the effect of the interaction protocol $\mathcal{V}$:

$$p[\Delta\mathbf{E},\Delta\mathbf{N};\mathcal{V}]=\sum_{m,n}\prod_i\delta(\Delta E_i-E_m^i+E_n^i)\times\delta(\Delta N_i-N_m^i+N_n^i)p_{m|n}[\mathcal{V}]p_n^0, \tag{82}$$

where $p_{m|n}[\mathcal{V}]$ is the transition probability from state $|\psi_n\rangle$ to $|\psi_m\rangle$ and $p_n^0=\prod_i e^{-\beta_i[E_n^i-\mu_i N_n^i]}/\Xi_i$ is the initial distribution of energies and particles. Here the symbols $\Delta\mathbf{E}$ and $\Delta\mathbf{N}$ are shorthand notations for the individual energy and particle number changes of all subsystems $\Delta E_1,\Delta E_2,\ldots,\Delta E_s$ and $\Delta N_1,\Delta N_2,\ldots,\Delta N_s$, respectively.

Assuming that the total Hamiltonian commutes with the time-reversal operator at any instant of time and using the time-reversal property of the transition probabilities, Eq. (56), one obtains

$$\frac{p[\Delta\mathbf{E},\Delta\mathbf{N};\mathcal{V}]}{p[-\Delta\mathbf{E},-\Delta\mathbf{N};\tilde{\mathcal{V}}]}=\prod_i e^{\beta_i[\Delta E_i-\mu_i\Delta N_i]}. \tag{83}$$

This equation was derived by Andrieux *et al.* (2009) and expresses the exchange fluctuation relation for the case of transport of energy and matter.

In the case of a single isolated system ($s=1$), it reduces to the Tasaki-Crooks work fluctuation theorem, Eq. (53), upon rewriting $\Delta E_1=w$ and assuming that the total number of particles is conserved also when the interaction is switched on, i.e., $[\mathcal{H}(\mathcal{V}_t),\mathcal{N}]=0$, to obtain $\Delta N=0$. The free energy difference does not appear in Eq. (83) because we have assumed the protocol $\mathcal{V}$ to be cyclic.

In the case of two weakly interacting systems ($s=2$, $\mathcal{V}$ small) that do not exchange particles, it reduces to the fluctuation theorem of Jarzynski and Wójcik (2004) for heat exchange:

$$\frac{p[Q;\mathcal{V}]}{p[-Q;\tilde{\mathcal{V}}]}=e^{(\beta_1-\beta_2)Q}, \tag{84}$$

where $Q=\Delta E_1=-\Delta E_2$, with the second equality following from the assumed weak interaction.

In case of two weakly interacting systems ($s=2$, $\mathcal{V}$ small) that do exchange particles and are initially at the same temperature, the fluctuation relation takes on the form

$$\frac{p[q;\mathcal{V}]}{p[-q;\tilde{\mathcal{V}}]}=e^{\beta(\mu_1-\mu_2)q}, \tag{85}$$

where $q\equiv\Delta N_2=-\Delta N_1$.

One basic assumption leading to the exchange fluctuation relation, Eq. (83), is that the initial state is a factorized state, in which the various subsystems are uncorrelated from each other. In most experimental situations, however, unavoidable interactions between the systems would lead to some correlations and a consequent deviation from the assumed factorized state, Eq. (81). The resulting deviation from the exchange fluctuation relation, Eq. (83), is expected to vanish for observation times τ larger than some characteristic time scale τ_c, determined by the specific physical properties of the experimental setup (Andrieux *et al.*, 2009; Esposito *et al.*, 2009; Campisi *et al.*, 2010a):

$$\frac{p[\Delta\mathbf{E},\Delta\mathbf{N};\mathcal{V}]}{p[-\Delta\mathbf{E},-\Delta\mathbf{N};\tilde{\mathcal{V}}]}\xrightarrow{\tau\gg\tau_c}\prod_i e^{\beta_i(\Delta E_i-\mu_i\Delta N_i)}. \tag{86}$$

For those large times $t\gg\tau_c$ a nonequilibrium steady state sets in under the condition that the reservoirs are chosen macroscopic. For this reason Eq. (86) is referred to as a steady state fluctuation relation. This is in contrast to the other fluctuation relations discussed above, which instead are valid for any observation time τ and are accordingly referred to as transient fluctuation relations. Saito and Dhar (2007) provided an explicit demonstration of Eq. (86) for the quantum heat transfer across a harmonic chain connecting two thermal reservoirs at different temperatures. Ren *et al.* (2010) reported on the breakdown of Eq. (86) induced by a nonvanishing Berry-phase heat pumping. The latter occurs when the temperatures of the two baths are adiabatically modulated in time.

VI. EXPERIMENTS

A. Work fluctuation relations

Regarding the experimental validation of the work fluctuation relation, a fundamental difference exists between the classical and the quantum regime. In classical experiments work is accessible by measuring the trajectory $\mathbf{x}_t$ of the possibly open system and integrating the instantaneous power according to $W=\int dt\partial H_S/\partial t$, Eq. (31). In clear contrast, in quantum mechanics the work is obtained as the difference of two measurements of the energy, and an "integrated power" expression does not exist for the work; see Sec. III.C.

Closely following the prescriptions of the theory one should perform the following steps in order to experimentally verify the work fluctuation relation, Eq. (53): (i) Prepare a quantum system in the canonical state, Eq. (2), at time $t=0$. (ii) Measure the energy at $t=0$. (iii) Drive the system by means of some forcing protocol λ_t for times t between 0 and τ, and make sure that during this time the system is well insulated from its environment. (iv) Measure the energy again at τ and record the work w, according to Eq. (39). (v) Repeat this procedure many times and construct the histogram of the statistics of work as an estimate of the work PDF $p[w;\lambda]$. In order to determine the backward probability the same type of experiment has to be repeated with the inverted protocol, starting from an equilibrium state at inverse temperature β and at those parameter values that are reached at the end of the forward protocol.

1. Proposal for an experiment employing trapped cold ions

Huber *et al.* (2008) suggested an experiment that exactly follows the procedure described above. They proposed to implement a quantum harmonic oscillator by optically trapping an ion in the quadratic potential generated by a laser trap, using the setup developed by Schulz *et al.* (2008). In principle, the setup of Schulz *et al.* (2008) allows, on the one hand, to drive the system by changing in time the stiffness of the trap, and, on the other hand, to probe whether the ion is in a certain Fock state $|n\rangle$, i.e., in an energy eigenstate of the harmonic oscillator. The measurement apparatus may be understood as a single Fock state "filter," whose outcome is "yes" or "no," depending on whether the ion is or is not in the probed state. Thus the experimentalist probes each possible outcome (n, m), where (n, m) denotes the Fock states at time $t = 0$ and $t = \tau$, respectively. Then the relative frequency of the outcome (n, m) occurrence is recorded by repeating the driving protocol many times always preparing the system in the same canonical initial state. In this way the joint probabilities $p_{m|n}[\lambda]p_n^0$ are measured.

The work histogram is then constructed as an estimate of the work PDF, Eq. (48), thus providing experimental access to the fluctuation relation, Eq. (53). Likewise the relative frequency of the outcomes having n as the initial state gives the experimental initial population p_n^0. Thus, with this experiment one can actually check the symmetry relation of the conditional probabilities $p_{m|n}[\lambda] = p_{n|m}[\tilde{\lambda}]$, Eq. (56), and compare their experimental values with the known theoretical values (Husimi, 1953; Deffner and Lutz, 2008; Talkner, Burada, and Hänggi, 2008, 2009).

Another suitable quantum system to test quantum fluctuation relations are quantum versions of nanomechanical oscillator setups that with present day nanotechnology are at the verge of entering the quantum regime.[18] In these systems work protocols can be imposed by optomechanical means.

2. Proposal for an experiment employing circuit quantum electrodynamics

Currently, the experiment proposed by Huber *et al.* (2008) has not yet been carried out. An analogous experiment could, in principle, be performed in a circuit quantum electrodynamics (QED) setup as the one described by Hofheinz *et al.* (2008, 2009). The setup consists of a Cooper pair box qubit (a two state quantum system) that can be coupled to and decoupled from a superconducting 1D transmission line, where the latter mimics a quantum harmonic oscillator. With this architecture it is possible to implement various functions with a very high degree of accuracy. Among them the following tasks are of special interest in the present context: (i) creation of pure Fock states $|n\rangle$, i.e., the energy eigenstates of the quantum harmonic oscillator in the resonator; (ii) measurement of photon statistics p_m, i.e., measurements of the population of each quantum state $|m\rangle$ of the oscillator; and (iii) driving the resonator by means of an external field.

Hofheinz *et al.* (2008) reported, for example, on the creation of the ground Fock state $|0\rangle$, followed by a driving protocol λ (a properly engineered microwave pulse applied to the resonator) that "displaces" the oscillator and creates a coherent state $|\alpha\rangle$, whose photon statistics $p_{m|0}[\lambda]$ was measured with good accuracy up to $n_{\max} \sim 10$. In more recent experiments (Hofheinz *et al.*, 2009) the accuracy was improved and $n_{\max}$ was raised to ~ 15. The quantity $p_{m|0}[\lambda]$ is actually the conditional probability to find the state $|m\rangle$ at time $t = \tau$, given that the system was in the state $|0\rangle$ at time $t = 0$. Thus, by preparing the oscillator in the Fock state $|n\rangle$ instead of the ground state $|0\rangle$, and repeating the same driving and readout as before, the matrix $p_{m|n}[\lambda]$ can be determined experimentally. Accordingly one can test the validity of the symmetry relation $p_{m|n}[\lambda] = p_{n|m}[\tilde{\lambda}]$, Eq. (56), which in turn implies the work fluctuation relation; see Sec. IV.E. At variance with the proposal of Huber *et al.* (2008), in this case the initial state would not be randomly sampled from a canonical state, but would be rather created deterministically by the experimenter.

The theoretical values of transition probabilities for this case corresponding to a displacement of the oscillator were first reported by Husimi (1953); see also Campisi (2008). Talkner, Burada, and Hänggi (2008) provided an analytical expression for the characteristic function of work and investigated in detail the work probability distribution function and its dependence on the initial state, such as, for example, canonical, microcanonical, and coherent states.

So far we addressed possible experimental tests of the Tasaki-Crooks work fluctuation theorem, Eq. (53), for isolated systems. The case of open systems, interacting with a thermal bath, poses extra difficulties related to the fact that in order to measure the work in this case one should make two measurements of the energy of the total macroscopic system, made up of the system of interest and its environment. This presents an extra obstacle that at the moment seems difficult to surmount except for a situation at (i) weak coupling and (ii) $Q \sim 0$, then yielding, together with Eq. (63) $w \sim \Delta E$.

B. Exchange fluctuation relations

Similar to the quantum work fluctuation relations, the quantum exchange fluctuation relations are understood in terms of two-point quantum measurements. In an experimental test, the net amount of energy and/or number of particles [depending on which of the three exchange fluctuation relations, Eqs. (83)–(85), is studied] has to be measured in each subsystem twice, at the beginning and at the end of the protocol. However, typically these are macroscopic reservoirs, whose energy and particle number measurement are practically impossible.[19] Thus, seemingly, the verification of the exchange fluctuation relations would be even more problematic than the validation of the quantum work fluctuation relations. Indeed, while experimental tests of the work fluctuation relations have not yet been reported, experiments concerning quantum exchange fluctuation relations have already been performed. In the following we discuss two of them, one by Utsumi *et al.* (2010) and the other by Nakamura *et al.* (2010). In doing so we demonstrated how the obstacle of energy or particle content measurement of macroscopic reservoirs was circumvented.

1. An electron counting statistics experiment

Utsumi *et al.* (2010) recently performed an experimental verification of the particle exchange fluctuation relation, Eq. (85), using bidirectional electron counting statistics (Fujisawa *et al.*, 2006). The experimental setup consists of

---

[18] Note the exciting recent advancements obtained with the works LaHaye *et al.* (2004), Kippenberg and Vahala (2008), Anetsberger *et al.* (2009), and O'Connell *et al.* (2010).

[19] See Andrieux *et al.* (2009), Esposito *et al.* (2009), and Campisi *et al.* (2010a), and also Jarzynski (2000) regarding the classical case.

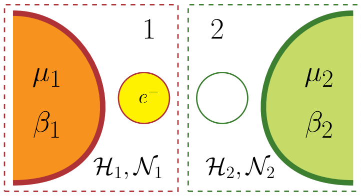

FIG. 6 (color online). Scheme of a bidirectional counting statistics experiment. Two leads (large semicircles) with different electronic chemical potentials ($\mu_1 \neq \mu_2$) and same temperature ($\beta_1 = \beta_2$) are connected through a double quantum dot (small circles), whose quantum state is continuously monitored. The state (1,0), i.e., one electron in the left dot, no electrons in the right dot, is depicted. The transition from this state to the state (0,1) signals the exchange of one electron from subsystem 1 to subsystem 2. $\mathcal{H}_{1,2}$ and $\mathcal{N}_{1,2}$ denote the Hamiltonian and electron number operators of the subsystems, respectively.

two electron reservoirs (leads) at the same temperature. The two leads are connected via a double quantum dot; see Fig. 6.

When an electric potential difference $V = \mu_1 - \mu_2$ is applied to the leads, a net flow of electrons starts transporting charges from one lead to the other, via lead-dot and dot-dot quantum tunnelings. The measurement apparatus consists of a secondary circuit in which a current flows due to an applied voltage. Thanks to a properly engineered coupling between the secondary circuit and the double quantum dot, the current in the circuit depends on the quantum state of the double dot. The latter has four relevant states, which we denote as $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ corresponding, respectively, to no electrons in the left dot and no electrons in the right dot, no electrons in the left dot and one electron in the right dot, etc. Each of these states leads to a different value of the current in the secondary circuit. In the experiment an electric potential difference is applied to the two leads for a time τ. During this time the state of the double quantum dot is monitored by registering the current in the secondary circuit. This current was found to switch between the four values corresponding to the four quantum states mentioned above. The outcome of the experiment is a sequence I_k of current values, with I_k taking only four possible values. In other words, the outcome of the experiment consists of a sequence $\{(l, r)\}_k$ (with $l, r = 0, 1$) of joint eigenvalues of two commuting observables $\mathcal{L}$, $\mathcal{R}$ specifying the occupation of the left (l) and right (r) dots by single electrons at the time of the kth measurement. The presence of an exchange of entries within one time step of the form $(1, 0)_n$, $(0, 1)_{n+1}$ signals the transfer of one electron from left to right, and vice versa $(0, 1)_n$, $(1, 0)_{n+1}$ the transfer from right to left. Thus, given a sequence $\{(l, r)\}_k$, the total number $q[\{(l, r)\}_k]$ of electrons transferred from left to right is obtained by subtracting the total number of right-to-left transfers from the total number of left-to-right transfers. It was found that, for observation times larger than a characteristic time τ_c, the fluctuation relation $p[q] = p[-q]e^{\beta V q}$, Eq. (85), was satisfied with the actual temperature of the leads replaced by an effective temperature; see Fig. 1. The renormalization of temperature was explained as an effect due to an exchange of electrons occurring between the dots and the secondary circuit (Utsumi *et al.*, 2010).

The question, however, remains of how to connect this experiment in which the flux of electrons through an interface is monitored and the theory, leading to Eq. (85), which instead prescribes only two measurements of total particle numbers in the reservoirs. The answer was given by Campisi *et al.* (2010a), who showed that the exchange fluctuation relation, Eq. (83), remains valid, if in addition to the two measurements of total energy and particle numbers occurring at 0 and τ, the evolution of a quantum system is interrupted by means of projective quantum measurements of any observable $\mathcal{A}$ that commutes with the quantum time-reversal operator Θ. In other words, while the forward and backward probabilities are affected by the occurrence of intermediate measurement processes, their ratio remains unaltered.

In the experiment by Utsumi *et al.* (2010) one does not need to measure the initial and final content of particles in the reservoirs because the number of exchanged particles is inferred from the sequence of intermediate measurements outcomes $\{(l, r)\}_k$. Thus, thanks to the fact that quantum measurements do not alter the fluctuation relation, one may overcome the problem of measuring the energy and number of particles of the macroscopic reservoirs, by monitoring instead the flux through a microscopic junction.

2. Nonlinear response relations in a quantum coherent conductor

As discussed in the Introduction, the original motivation for the study of fluctuation relations was to overcome the limitations of linear response theory and to obtain relations connecting higher order response functions to fluctuation properties of the unperturbed system. As an indirect and partial confirmation of the fluctuation relations higher order static fluctuation-response relations can be tested experimentally.

Such a validation was recently accomplished in coherent quantum transport experiments by Nakamura *et al.* (2010, 2011), where the average current I and the zero-frequency current noise power S generated in an Aharonov-Bohm ring were investigated as a function of an applied dc voltage V and magnetic field $\mathbf{B}$. In the nonlinear response regime, the current and noise power may be expressed as power series of the applied voltage:

$$I(V, \mathbf{B}) = G_1(\mathbf{B})V + \frac{G_2(\mathbf{B})}{2}V^2 + \frac{G_3(\mathbf{B})}{3!}V^3 + \cdots, \tag{87}$$

$$S(V, \mathbf{B}) = S_0(\mathbf{B}) + S_1(\mathbf{B})V + \frac{S_2(\mathbf{B})}{2}V^2 + \cdots, \tag{88}$$

where the coefficients depend on the applied magnetic field $\mathbf{B}$. The steady state fluctuation theorem, Eq. (86), then predicts the following fluctuation relations (Saito and Utsumi, 2008):

$$S_0 = 4k_B T G_1, \qquad S_1^S = 2k_B T G_2^S, \qquad S_1^A = 6k_B T G_2^A, \tag{89}$$

where $S_i^S = S_i(\mathbf{B}) + S_i(-\mathbf{B})$, $S_i^A = S_i(\mathbf{B}) - S_i(-\mathbf{B})$, and analogous definitions for G_i^S and G_i^A. The first equation in (89) is the Johnson-Nyquist relation (Johnson, 1928; Nyquist, 1928). In the experiment by Nakamura *et al.* (2010) good quantitative agreement with the first and third expressions in Eq. (89) was established, whereas, for the time being, only qualitative agreement was found with the second relation.

The higher order static fluctuation-dissipation relations (89) were obtained from a steady state fluctuation theorem for particle exchange under the simplifying assumption that no heat exchange occurs (Nakamura *et al.*, 2010). Then the probability of transferring q particles is related to the probability of the reverse transfer by $p(q) = p(-q)e^{Aq}$, where $A = \beta V = \beta(\mu_1 - \mu_2)$ is the so-called affinity. If both sides are multiplied by q and integrated over q a comparison of equal powers of applied voltage V yields Eq. (89) (Nakamura *et al.*, 2011). An alternative approach, that also allows one to include the effect of heat conduction, is offered by the fluctuation theorems for currents in open quantum systems. This objective has been put forward by Saito and Utsumi (2008) and also by Andrieux *et al.* (2009), based on a generating function approach in the spirit of Eq. (55).

VII. OUTLOOK

In closing this Colloquium we stress that the known fluctuation relations are based on two facts: (a) microreversibility for nonautonomous Hamiltonian systems Eq. (40), and (b) the special nature of the initial equilibrium states which is expressible in either microcanonical, canonical, or grand-canonical form, or products thereof. The final state reached at the end of a protocol though is in no way restricted. It evolves from the initial state according to the governing dynamical laws under a prescribed protocol. In general, this final state may markedly differ from any kind of equilibrium state.

For quantum mechanical systems it also is of utmost importance to correctly identify the work performed on a system as the difference between the energy of the system at the end and the beginning of the protocol. In case of open systems the difference of the energies of the total system at the end and beginning of the protocol coincides with the work done on the open system as long as the forces exclusively act on this open system. With the free energy of an open system determined as the difference of free energies of the total system and that of the isolated environment the quantum and classical Jarzynski equality and the Tasaki-Crooks theorem continue to hold true even for systems strongly interacting with their environment. Deviations from the fluctuation relations, however, must be expected if protocol forces not only act on the system alone but as well directly on the environmental degrees of freedom, for example, if a time-dependent system-bath interaction protocol is applied.

The most general and compact formulation of quantum work fluctuation relations also containing the Onsager-Casimir reciprocity relations and nonlinear response to all orders is the Andrieux-Gaspard relation, Eq. (55), which represents the proper quantum version of the classical Bochkov-Kuzovlev formula (Bochkov *et al.*, 1977), Eq. (14). These relations provide a complete theoretical understanding of those nonequilibrium situations that emerge from arbitrary time-dependent perturbations of equilibrium initial states.

Less understood are exchange fluctuation relations with their important applications to counting statistics (Esposito *et al.*, 2009). The theory there so far is restricted to situations where the initial state factorizes into grand-canonical states of reservoirs at different temperatures or chemical potentials. The interaction between these reservoirs is turned on and it is assumed that it will lead to a steady state within the duration of the protocol. Experimentally, it is in general difficult to exactly follow this prescription and therefore a comparison of theory and experiment is only meaningful for the steady state. Alternative derivations of exchange relations for more realistic, nonfactorizing initial states would certainly be of interest. In this context, the issue of deriving quantum fluctuation relations for open systems that initially are in nonequilibrium steady quantum transport states constitutes an interesting challenge. Likewise, from the theoretical point of view little is known thus far about quantum effects for transport in presence of time-dependent reservoirs, for example, using a varying temperature and/or chemical potentials (Ren *et al.*, 2010).

The experimental applications and validation schemes involving nonlinear quantum fluctuation relations still are in a state of infancy, as detailed in Sec. VI, so that there is plenty of room for advancements. The major obstacle for the experimental verification of the work fluctuation relation is posed by the necessity of performing quantum projective measurements of energy. Besides the proposal of Huber *et al.* (2008) employing trapped ions, we suggested here the scheme of a possible experiment employing circuit-QED architectures. In regard to exchange fluctuation relations instead, the main problem is related to the difficulty of measuring microscopic changes of macroscopic quantities pertaining to heat and matter reservoirs. Continuous measurements of fluxes seemingly provide a practical and efficient loophole for this dilemma (Campisi *et al.*, 2010a).

The idea that useful work may be obtained by using information (Maruyama *et al.*, 2009) has established a connection between the topical fields of quantum information theory (Vedral, 2002) and quantum fluctuation relations. Piechocinska (2000) and Kawai *et al.* (2007) used fluctuation relations and information theoretic measures to derive Landauer's principle. A generalization of the Jarzynski equality to the case of feedback controlled systems was provided in the classical case by Sagawa and Ueda (2010), and in the quantum case by Morikuni and Tasaki (2010). Recently Deffner *et al.* (2010) gave bounds on the entropy production in terms of quantum information concepts. In a similar spirit, Hide and Vedral (2010) presented a method by relating relative quantum entropy to the quantum Jarzynski fluctuation identity in order to quantify multipartite entanglement within different thermal quantum states. A practical application of the Jarzynski equality in quantum computation was shown by Ohzeki (2010).

In conclusion, we are confident in our belief that this topic of quantum fluctuation relations will exhibit an ever growing activity within nanosciences and further may invigorate readers to pursue their own research and experiments as this theme certainly offers many more surprises and unforeseen applications.

ACKNOWLEDGMENTS

We thank Y. Utsumi, D. S. Golubev, M. Marthaler, K. Saito, T. Fujisawa, and G. Schön for providing the data for Fig. 1. This work was supported by the cluster of excellence Nanosystems Initiative Munich (NIM) and the Volkswagen Foundation (Project No. I/83902).

APPENDIX A: DERIVATION OF THE BOCHKOV-KUZVLEV RELATION

We report the steps leading to Eq. (14):

$$\begin{aligned}\left\langle \exp\left[\int_0^\tau ds u_s B_s\right] e^{-\beta W_0}\right\rangle_\lambda &= \int d\mathbf{z}_0 \frac{e^{-\beta[H_0(\mathbf{z}_0)+W_0]}}{Z(t_0)} \exp\left[\int_0^\tau ds u_s B(\varphi_{s,0}[\mathbf{z}_0;\lambda])\right] \\ &= \int d\mathbf{z}_\tau \rho_0(\mathbf{z}_\tau) \exp\left[\int_0^\tau ds u_s B(\varepsilon\varphi_{\tau-s,0}[\varepsilon\mathbf{z}_\tau;\varepsilon_Q\tilde{\lambda}])\right] \\ &= \int d\mathbf{z}'_\tau \rho_0(\mathbf{z}'_\tau) \exp\left[\int_0^\tau dr u_{\tau-r}\varepsilon_B B(\varphi_{r,0}[\mathbf{z}'_\tau;\varepsilon_Q\tilde{\lambda}])\right],\end{aligned} \tag{A1}$$

where the first equality provides an explicit expression for the left-hand side of Eq. (14). In going from the second to the third line we employed the expression of work in Eq. (13), the microreversibility principle (10) and made the change of variable $\mathbf{z}_0 \to \mathbf{z}_\tau$. The Jacobian of this transformation is unity, because the time evolution in classical mechanics is a canonical transformation. A further change of variables $\mathbf{z}_\tau \to \mathbf{z}'_\tau = \varepsilon\mathbf{z}_\tau$, whose Jacobian is unity as well, and the change $s \to r = \tau - s$, yields the expression in the last line that coincides with the right-hand side of Eq. (14). In the last line we used the property $\rho_0(\mathbf{z}) = \rho_0(\varepsilon\mathbf{z})$, inherited by $\rho_0 = e^{-\beta H_0}/Z_0$ from the assumed time-reversal invariance of the Hamiltonian $H(\mathbf{z}) = H(\varepsilon\mathbf{z})$.

APPENDIX B: QUANTUM MICROREVERSIBILITY

In order to prove the quantum principle of microreversibility, we first discretize time and express the time evolution operator $U_{t,0}[\tilde{\lambda}]$ as a time ordered product (Schleich, 2001):

$$\begin{aligned}U_{\tau-t,0}[\tilde{\lambda}] = \lim_{N\to\infty} & e^{-(i/\hbar)\mathcal{H}(\tilde{\lambda}_{\tau-N\varepsilon})\varepsilon} \cdots e^{-(i/\hbar)\mathcal{H}(\tilde{\lambda}_\varepsilon)\varepsilon} \\ & \times e^{-(i/\hbar)\mathcal{H}(\tilde{\lambda}_0)\varepsilon},\end{aligned} \tag{B1}$$

where $\varepsilon = t/N$ denotes the time step. Using Eq. (9), we obtain

$$\begin{aligned}U_{\tau-t,0}[\tilde{\lambda}] = \lim_{N\to\infty} & e^{-(i/\hbar)\mathcal{H}(\lambda_{N\varepsilon})\varepsilon} \cdots e^{-(i/\hbar)\mathcal{H}(\lambda_{\tau-\varepsilon})\varepsilon} \\ & \times e^{-(i/\hbar)\mathcal{H}(\lambda_\tau)\varepsilon}.\end{aligned} \tag{B2}$$

Thus,

$$\begin{aligned}\Theta^\dagger U_{\tau-t,0}[\tilde{\lambda}]\Theta = \lim_{N\to\infty} & \Theta^\dagger e^{-(i/\hbar)\mathcal{H}(\lambda_{N\varepsilon})\varepsilon}\Theta\Theta^\dagger \\ & \times e^{-(i/\hbar)\mathcal{H}(\lambda_{N\varepsilon+\varepsilon})\varepsilon}\Theta \cdots \Theta^\dagger \\ & \times e^{-(i/\hbar)\mathcal{H}(\lambda_\tau)\varepsilon}\Theta,\end{aligned} \tag{B3}$$

where we inserted $\Theta\Theta^\dagger = \mathbb{1}$, $N-1$ times. Assuming that $\mathcal{H}(\lambda_t)$ commutes at all times with the time-reversal operator Θ, Eq. (41), we find

$$\Theta^\dagger e^{-(i/\hbar)\mathcal{H}(\lambda_t)u}\Theta = e^{(i/\hbar)\mathcal{H}(\lambda_t)u^*}, \tag{B4}$$

for any complex number u. Using this equation and the fact that ε is real valued, $\varepsilon^* = \varepsilon$, we obtain Eq. (40):

$$\begin{aligned}\Theta^\dagger U_{\tau-t,0}[\tilde{\lambda}]\Theta &= \lim_{N\to\infty} e^{(i/\hbar)\mathcal{H}(\lambda_{N\varepsilon})\varepsilon} e^{(i/\hbar)\mathcal{H}(\lambda_{N\varepsilon+\varepsilon})\varepsilon} \cdots \\ &\quad \times e^{(i/\hbar)\mathcal{H}(\lambda_\tau)\varepsilon} \\ &= \lim_{N\to\infty} [e^{-(i/\hbar)\mathcal{H}(\lambda_\tau)\varepsilon} \cdots e^{-(i/\hbar)\mathcal{H}(\lambda_{N\varepsilon+\varepsilon})\varepsilon} \\ &\quad \times e^{-(i/\hbar)\mathcal{H}(\lambda_{N\varepsilon})\varepsilon}]^\dagger \\ &= U^\dagger_{\tau,t}[\lambda] = U_{t,\tau}[\lambda].\end{aligned} \tag{B5}$$

APPENDIX C: TASAKI-CROOKS RELATION FOR THE CHARACTERISTIC FUNCTION

From Eq. (50) we have

$$\begin{aligned}\mathcal{Z}(\lambda_0)G[u;\lambda] = \mathrm{Tr} U^\dagger_{\tau,0}[\lambda] e^{iu\mathcal{H}(\lambda_\tau)} U_{\tau,0}[\lambda] e^{-iu\mathcal{H}(\lambda_0)} \\ \times e^{-\beta\mathcal{H}(\lambda_0)}.\end{aligned} \tag{C1}$$

For $t = 0$, the microreversibility principle, Eq. (40), becomes $U_{0,\tau}[\lambda] = U^\dagger_{\tau,0}[\lambda] = \Theta^\dagger U_{\tau,0}[\tilde{\lambda}]\Theta$. Therefore,

$$\begin{aligned}\mathcal{Z}(\lambda_0)G[u;\lambda] = \mathrm{Tr}\Theta^\dagger U_{\tau,0}[\tilde{\lambda}]\Theta e^{iu\mathcal{H}(\lambda_\tau)}\Theta^\dagger U^\dagger_{\tau,0}[\tilde{\lambda}] \\ \times \Theta e^{-iu\mathcal{H}(\lambda_0)} e^{-\beta\mathcal{H}(\lambda_0)}\Theta^\dagger\Theta,\end{aligned} \tag{C2}$$

where we inserted $\Theta^\dagger\Theta = \mathbb{1}$ under the trace. Using Eq. (B4), we obtain

$$\begin{aligned}\mathcal{Z}(\lambda_0)G[u;\lambda] = \mathrm{Tr}\Theta^\dagger U_{\tau,0}[\tilde{\lambda}] e^{-iu^*\mathcal{H}(\lambda_\tau)} U^\dagger_{\tau,0}[\tilde{\lambda}] e^{iu^*\mathcal{H}(\lambda_0)} \\ \times e^{-\beta\mathcal{H}(\lambda_0)}\Theta.\end{aligned} \tag{C3}$$

The antilinearity of Θ implies, for any trace class operator $\mathcal{A}$, $\mathrm{Tr}\Theta^\dagger\mathcal{A}\Theta = \mathrm{Tr}\mathcal{A}^\dagger$. Using this we can write

$$\begin{aligned}\mathcal{Z}(\lambda_0)G[u;\lambda] = \mathrm{Tr} e^{-\beta\mathcal{H}(\lambda_0)} e^{-iu\mathcal{H}(\lambda_0)} U_{\tau,0}[\tilde{\lambda}] e^{iu\mathcal{H}(\lambda_\tau)} \\ \times U^\dagger_{\tau,0}[\tilde{\lambda}].\end{aligned} \tag{C4}$$

Using the cyclic property of the trace one then obtains the important result

$$\begin{aligned} \mathcal{Z}(\lambda_0)G[u;\lambda] &= \mathrm{Tr} U^\dagger_{\tau,0}[\tilde{\lambda}] e^{i(-u+i\beta)\mathcal{H}(\lambda_0)} \\ &\quad \times U_{\tau,0}[\tilde{\lambda}] e^{-i(-u+i\beta)\mathcal{H}(\lambda_0)} e^{-\beta\mathcal{H}(\lambda_0)} \\ &= \mathcal{Z}(\lambda_\tau)G[-u+i\beta;\tilde{\lambda}]. \end{aligned} \tag{C5}$$

REFERENCES

Adib, A. B., 2005, ‘‘Entropy and density of states from isoenergetic nonequilibrium processes,’’ Phys. Rev. E **71**, 056128.

Adib, A. B., 2009, ‘‘Comment on ’On the Crooks fluctuation theorem and the Jarzynski equality,’ ’’ [J. Chem. Phys. 129, 091101 (2008)], J. Chem. Phys. **130**, 247101.

Allahverdyan, A. E., and T. M. Nieuwenhuizen, 2005, ‘‘Fluctuations of work from quantum subensembles: The case against quantum work-fluctuation theorems,’’ Phys. Rev. E **71**, 066102.

Andrieux, D., and P. Gaspard, 2008, ‘‘Quantum Work Relations and Response Theory,’’ Phys. Rev. Lett. **100**, 230404.

Andrieux, D., P. Gaspard, T. Monnai, and S. Tasaki, 2009, ‘‘The fluctuation theorem for currents in open quantum systems,’’ New J. Phys. **11**, 043014.

Anetsberger, G., O. Arcizet, Q. P. Unterreithmeier, R. Rivière, A. Schliesser, E. M. Weig, J. P. Kotthaus, and T. J. Kippenberg, 2009, ‘‘Near-field cavity optomechanics with nanomechanical oscillators,’’ Nature Phys. **5**, 909.

Bernard, W., and H. B. Callen, 1959, ‘‘Irreversible Thermodynamics of Nonlinear Processes and Noise in Driven Systems,’’ Rev. Mod. Phys. **31**, 1017.

Bochkov, G. N., and Y. E. Kuzovlev, 1977, ‘‘General theory of thermal fluctuations in nonlinear systems,’’ Zh. Eksp. Teor. Fiz. **72**, 238 [Sov. Phys. JETP **45**, 125 (1977)].

Bochkov, G. N., and Y. E. Kuzovlev, 1978, ‘‘Nonlinear stochastic models of oscillator systems,’’ Radiophys. Quantum Electron. **21**, 1019.

Bochkov, G. N., and Y. E. Kuzovlev, 1979, ‘‘Fluctuation-dissipation relations for nonequilibrium processes in open systems,’’ Zh. Eksp. Teor. Fiz. **76**, 1071 [Sov. Phys. JETP **49**, 543 (1979)].

Bochkov, G. N., and Y. E. Kuzovlev, 1981a, ‘‘Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics I. Generalized fluctuation-dissipation theorem,’’ Physica (Amsterdam) **106A**, 443.

Bochkov, G. N., and Y. E. Kuzovlev, 1981b, ‘‘Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics II. Kinetic potential and variational principles for nonlinear irreversible processes,’’ Physica (Amsterdam) **106A**, 480.

Caldeira, A. O., and A. J. Leggett, 1983, ‘‘Quantum tunnelling in a dissipative system,’’ Ann. Phys. (N.Y.) **149**, 374.

Callen, H. B., and T. A. Welton, 1951, ‘‘Irreversibility and Generalized Noise,’’ Phys. Rev. **83**, 34.

Campisi, M., 2008, ‘‘Increase of Boltzmann entropy in a quantum forced harmonic oscillator,’’ Phys. Rev. E **78**, 051123.

Campisi, M., P. Talkner, and P. Hänggi, 2009a, ‘‘Fluctuation Theorem for Arbitrary Open Quantum Systems,’’ Phys. Rev. Lett. **102**, 210401.

Campisi, M., P. Talkner, and P. Hänggi, 2009b, ‘‘Thermodynamics and fluctuation theorems for a strongly coupled open quantum system: an exactly solvable case,’’ J. Phys. A **42**, 392002.

Campisi, M., P. Talkner, and P. Hänggi, 2010a, ‘‘Fluctuation theorems for continuously monitored quantum fluxes,’’ Phys. Rev. Lett. **105**, 140601.

Campisi, M., P. Talkner, and P. Hänggi, 2011a, ‘‘Quantum Bochkov-Kuzovlev Work Fluctuation Theorems,’’ Phil. Trans. R. Soc. A **369**, 291.

Campisi, M., P. Talkner, and P. Hänggi, 2011b, ‘‘Influence of measurements on the statistics of work performed on a quantum system,’’ Phys. Rev. E **83**, 041114.

Campisi, M., D. Zueco, and P. Talkner, 2010, ‘‘Thermodynamic anomalies in open quantum systems: Strong coupling effects in the isotropic *XY* model,’’ Chem. Phys. **375**, 187.

Casimir, H. B. G., 1945, ‘‘On Onsager’s Principle of Microscopic Reversibility,’’ Rev. Mod. Phys. **17**, 343.

Chen, L. Y., 2008a, ‘‘Nonequilibrium fluctuation-dissipation theorem of Brownian dynamics,’’ J. Chem. Phys. **129**, 144113.

Chen, L. Y., 2008b, ‘‘On the Crooks fluctuation theorem and the Jarzynski equality,’’ J. Chem. Phys. **129**, 091101.

Chen, L. Y., 2009, Response to ‘‘Comment on ‘On the Crooks fluctuation theorem and the Jaraynski equality’ and ’Nonequilibrium fluctuation dissipation theorem of Brownian dynamics’,’’ [J. Chem. Phys. 130, 107101 (2009)] J. Chem. Phys. **130**, 107102.

Cleuren, B, C. Van den Broeck, and R. Kawai, 2006, ‘‘Fluctuation and Dissipation of Work in a Joule Experiment,’’ Phys. Rev. Lett. **96**, 050601.

Cohen-Tannoudji, C., B. Diu, and F. Laloë, 1977, *Quantum Mechanics* (Wiley, New York), Vol. 1, p. 318.

Collin, D., F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante, 2005, ‘‘Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies,’’ Nature (London) **437**, 231.

Crooks, G. E., 1999, ‘‘Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences,’’ Phys. Rev. E **60**, 2721.

Crooks, G. E., 2008, ‘‘On the Jarzynski relation for dissipative quantum dynamics,’’ J. Stat. Mech. P10023.

Crooks, G. E., 2009, Comment regarding ‘‘On the Crooks fluctuation theorem and the Jarzynski equality’’ [J. Chem. Phys. 129, 091101 (2008)] and ‘‘Nonequilibrium fluctuation-dissipation theorem of Brownian dynamics’’ [J. Chem. Phys. 129, 144113 (2008)]; J. Chem. Phys. **130**, 107101.

Deffner, S., O. Abah, and E. Lutz, 2010, ‘‘Quantum work statistics of linear and nonlinear parametric oscillators,’’ Chem. Phys. **375**, 200.

Deffner, S., and E. Lutz, 2008, ‘‘Nonequilibrium work distribution of a quantum harmonic oscillator,’’ Phys. Rev. E **77**, 021128.

Deffner, S., and E. Lutz, 2010, ‘‘Generalized Clausius Inequality for Nonequilibrium Quantum Processes,’’ Phys. Rev. Lett. **105**, 170402.

de Groot, S. R., and P. Mazur, 1984, *Non-equilibrium Thermodynamics* (Dover, New York).

de Roeck, W., and C. Maes, 2004, ‘‘Quantum version of free-energy–irreversible-work relations,’’ Phys. Rev. E **69**, 026115.

Dittrich, T., P. Hänggi, G.-L. Ingold, B. Kramer, G. Schön, and W. Zwerger, 1998, *Quantum Transport and Dissipation* (Wiley-VCH, Weinheim).

Douarche, F., S. Ciliberto, A. Petrosyan, and I. Rabbiosi, 2005, ‘‘An experimental test of the Jarzynski equality in a mechanical experiment,’’ Europhys. Lett. **70**, 593.

Efremov, G. F., 1969, ‘‘A fluctuation dissipation theorem for nonlinear media,’’ Zh. Eksp. Teor. Fiz. **55**, 2322 [Sov. Phys. JETP **28**, 1232 (1969)].

Einstein, A., 1905, ‘‘Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen,’’ Ann. Phys. (Leipzig) **322**, 549, translated into English in Einstein, 1926.

Einstein, A., 1906a, ‘‘Eine neue Bestimmung der Moleküldimensionen,’’ Ann. Phys. (Leipzig) **324**, 289, translated into English in Einstein, 1926.

Einstein, A., 1906b, ‘‘Zur Theorie der Brownschen Bewegung,’’ Ann. Phys. (Leipzig) **324**, 371, translated into English in Einstein, 1926.

Einstein, A., 1926, *Investigations on the Theory of Brownian Movement* (Methuen, London), reprinted by Dover, New York, 1956.

Esposito, M., U. Harbola, and S. Mukamel, 2009, ‘‘Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems,’’ Rev. Mod. Phys. **81**, 1665.

Esposito, M., and S. Mukamel, 2006, ‘‘Fluctuation theorems for quantum master equations,’’ Phys. Rev. E **73**, 046129.

Evans, D. J., E. G. D. Cohen, and G. P. Morriss, 1993, ‘‘Probability of second law violations in shearing steady states,’’ Phys. Rev. Lett. **71**, 2401.

Feynman, R. P., and J. F. L. Vernon, 1963, ‘‘The theory of a general quantum system interacting with a linear dissipative system,’’ Ann. Phys. (N.Y.) **24**, 118.

Ford, G. W., J. T. Lewis, and R. F. O'Connell, 1985, ‘‘Quantum oscillator in a blackbody radiation field,’’ Phys. Rev. Lett. **55**, 2273.

Fujisawa, T., T. Hayashi, R. Tomita, and Y. Hirayama, 2006, ‘‘Bidirectional counting of single electrons,’’ Science **312**, 1634.

Gallavotti, G., and E. G. D. Cohen, 1995, ‘‘Dynamical Ensembles in Nonequilibrium Statistical Mechanics,’’ Phys. Rev. Lett. **74**, 2694.

Grabert, H., P. Schramm, and G.-L. Ingold, 1988, ‘‘Quantum Brownian motion: The functional integral approach,’’ Phys. Rep. **168**, 115.

Grabert, H., U. Weiss, and P. Talkner, 1984, ‘‘Quantum theory of the damped harmonic oscillator,’’ Z. Phys. B **55**, 87.

Green, M. S., 1952, ‘‘Markoff Random Processes and the Statistical Mechanics of Time-Dependent Phenomena,’’ J. Chem. Phys. **20**, 1281.

Green, M. S., 1954, ‘‘Markov Random Processes and the Statistical Mechanics of Time-Dependent Phenomena. II. Irreversible Processes in Fluids,’’ J. Chem. Phys. **22**, 398.

Hahn, A. M., and H. Then, 2009, ‘‘Using bijective maps to improve free-energy estimates,’’ Phys. Rev. E **79**, 011113.

Hahn, A. M., and H. Then, 2010, ‘‘Measuring the convergence of Monte Carlo free-energy calculations,’’ Phys. Rev. E **81**, 041117.

Hänggi, P., 1978, ‘‘Stochastic Processes II: Response theory and fluctuation theorems,’’ Helv. Phys. Acta **51**, 202.

Hänggi, P., 1982, ‘‘Nonlinear fluctuations: The problem of deterministic limit and reconstruction of stochastic dynamics,’’ Phys. Rev. A **25**, 1130.

Hänggi, P., and G.-L. Ingold, 2006, ‘‘Quantum Brownian motion and the third law of thermodynamics,’’ Acta Phys. Pol. B **37**, 1537 [http://www.actaphys.uj.edu.pl/vol37/pdf/v37p1537.pdf].

Hänggi, P., and G. Ingold, 2005, ‘‘Fundamental aspects of quantum Brownian motion,’’ Chaos **15**, 026105.

Hänggi, P., and H. Thomas, 1982, ‘‘Stochastic processes: Time evolution, symmetries and linear response,’’ Phys. Rep. **88**, 207.

Hide, J., and V. Vedral, 2010, ‘‘Detecting entanglement with Jarzynski's equality,’’ Phys. Rev. A **81**, 062303.

Hofheinz, M., *et al.*, 2009, ‘‘Synthesizing arbitrary quantum states in a superconducting resonator,’’ Nature (London) **459**, 546.

Hofheinz, M., E. M. Weig, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, H. Wang, J. M. Martinis, and A. N. Cleland, 2008, ‘‘Generation of Fock states in a superconducting quantum circuit,’’ Nature (London) **454**, 310.

Hörhammer, C., and H. Büttner, 2008, ‘‘Information and Entropy in Quantum Brownian Motion Thermodynamic Entropy versus von Neumann Entropy,’’ J. Stat. Phys. **133**, 1161.

Horowitz, J., and C. Jarzynski, 2008, Comment on ‘‘Failure of the work-Hamiltonian connection for free-energy calculations,’’ Phys. Rev. Lett. **101**, 098901.

Horowitz, J., and C. Jarzynski, 2007, ‘‘Comparison of work fluctuation relations,’’ J. Stat. Mech. P11002.

Huber, G., F. Schmidt-Kaler, S. Deffner, and E. Lutz, 2008, ‘‘Employing trapped cold ions to verify the quantum Jarzynski equality,’’ Phys. Rev. Lett. **101**, 070403.

Husimi, K., 1953, ‘‘Miscellanea in Elementary Quantum Mechanics, I,’’ Prog. Theor. Phys. **9**, 238.

Ingold, G.-L., 2002 Lect. Notes Phys. **611**, 1.

Jarzynski, C., 1997, ‘‘Nonequilibrium equality for free energy differences,’’ Phys. Rev. Lett. **78**, 2690.

Jarzynski, C., 2000, ‘‘Hamiltonian derivation of a detailed fluctuation theorem,’’ J. Stat. Phys. **98**, 77.

Jarzynski, C., 2002, ‘‘Targeted free energy perturbation,’’ Phys. Rev. E **65**, 046122.

Jarzynski, C., 2007, ‘‘Comparison of far-from-equilibrium work relations,’’ C.R. Physique **8**, 495.

Jarzynski, C., 2008, ‘‘Nonequilibrium work relations: foundations and applications,’’ Eur. Phys. J. B **64**, 331.

Jarzynski, C., 2011, ‘‘Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale,’’ Annu. Rev. Condens. Matter Phys. **2**, 329.

Jarzynski, C., 2004, ‘‘Nonequilibrium work theorem for a system strongly coupled to a thermal environment,’’ J. Stat. Mech. P09005.

Jarzynski, C., and D. K. Wójcik, 2004, ‘‘Classical and quantum fluctuation theorems for heat exchange,’’ Phys. Rev. Lett. **92**, 230602.

Johnson, J. B., 1928, ‘‘Thermal agitation of electricity in conductors,’’ Phys. Rev. **32**, 97.

Katsuda, H., and M. Ohzeki, 2011, ‘‘Jarzynski Equality for an Energy-Controlled System,’’ J. Phys. Soc. Jpn. **80**, 045003.

Kawai, R., J. M. R. Parrondo, and C. V. den Broeck, 2007, ‘‘Dissipation: The Phase-Space Perspective,’’ Phys. Rev. Lett. **98**, 080602.

Khinchin, A., 1949, *Mathematical foundations of statistical mechanics* (Dover, New York).

Kippenberg, T. J., and K. J. Vahala, 2008, ‘‘Cavity optomechanics: Back-action at the mesoscale,’’ Science **321**, 1172.

Kobe, D. H., 1981, ‘‘Gauge-invariant classical Hamiltonian formulation of the electrodynamics of nonrelativistic particles,’’ Am. J. Phys. **49**, 581.

Kubo, R., 1957, ‘‘Statistical-mechanical theory of irreversible processes. I,’’ J. Phys. Soc. Jpn. **12**, 570.

Kurchan, J., 2000, ‘‘A quantum fluctuation theorem,’’ arXiv:cond-mat/0007360.

LaHaye, M. D., O. Buu, B. Camarota, and K. C. Schwab, 2004, ‘‘Approaching the Quantum Limit of a Nanomechanical Resonator,’’ Science **304**, 74.

Liphardt, J., S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante, 2002, ‘‘Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality,’’ Science **296**, 1832.

Marconi, U. M. B., A. Puglisi, L. Rondoni, and A. Vulpiani, 2008, ‘‘Fluctuation-dissipation: Response theory in statistical physics,’’ Phys. Rep. **461**, 111.

Maruyama, K., F. Nori, and V. Vedral, 2009, "Colloquium: The physics of Maxwell's demon and information," Rev. Mod. Phys. **81**, 1.

Messiah, A., 1962, *Quantum Mechanics* (North-Holland, Amsterdam).

Minh, D. D. L., and A. B. Adib, 2008, "Optimized free energies from bidirectional single-molecule force spectroscopy," Phys. Rev. Lett. **100**, 180602.

Morikuni, Y., and H. Tasaki, 2011, "Quantum Jarzynski-Sagawa-Ueda relations," J. Stat. Phys. **143**, 1.

Mukamel, S., 2003, "Quantum extension of the Jarzynski relation: Analogy with stochastic dephasing," Phys. Rev. Lett. **90**, 170604.

Nakamura, S., *et al.*, 2010, "Nonequilibrium fluctuation relations in a quantum coherent conductor," Phys. Rev. Lett. **104**, 080602.

Nakamura, S., *et al.*, 2011, "Fluctuation Theorem and Microreversibility in a Quantum Coherent Conductor," Phys. Rev. B **83**, 155431.

Nieuwenhuizen, T. M., and A. E. Allahverdyan, 2002, "Statistical thermodynamics of quantum Brownian motion: Construction of perpetuum mobile of the second kind," Phys. Rev. E **66**, 036102.

Nyquist, H., 1928, "Thermal agitation of electric charge in conductors," Phys. Rev. **32**, 110.

O'Connell, A. D., *et al.*, 2010, "Quantum ground state and single-phonon control of a mechanical resonator," Nature (London) **464**, 697.

Ohzeki, M., 2010, "Quantum Annealing with the Jarzynski Equality," Phys. Rev. Lett. **105**, 050401.

Onsager, L., 1931a, "Reciprocal relations in irreversible processes. I.," Phys. Rev. **37**, 405.

Onsager, L., 1931b, "Reciprocal relations in irreversible processes. II.," Phys. Rev. **38**, 2265.

Peliti, L., 2008a, Comment on "Failure of the Work-Hamiltonian Connection for Free-Energy Calculations," Phys. Rev. Lett. **101**, 098903.

Peliti, L., 2008b, "On the work-Hamiltonian connection in manipulated systems," J. Stat. Mech. P05002.

Piechocinska, B., 2000, "Information erasure," Phys. Rev. A **61**, 062314.

Prost, J., J.-F. Joanny, and J. M. R. Parrondo, 2009, "Generalized Fluctuation-Dissipation Theorem for Steady-State Systems," Phys. Rev. Lett. **103**, 090601.

Ren, J., P. Hänggi, and B. Li, 2010, "Berry-Phase-Induced Heat Pumping and Its Impact on the Fluctuation Theorem," Phys. Rev. Lett. **104**, 170601.

Rondoni, L., and C. Mejía-Monasterio, 2007, "Fluctuations in nonequilibrium statistical mechanics: models, mathematical theory, physical mechanisms," Nonlinearity **20**, R1.

Sagawa, T., and M. Ueda, 2010, "Generalized Jarzynski Equality under Nonequilibrium Feedback Control," Phys. Rev. Lett. **104**, 090602.

Saito, K., and A. Dhar, 2007, "Fluctuation Theorem in Quantum Heat Conduction," Phys. Rev. Lett. **99**, 180601.

Saito, K., and Y. Utsumi, 2008, "Symmetry in full counting statistics, fluctuation theorem, and relations among nonlinear transport coefficients in the presence of a magnetic field," Phys. Rev. B **78**, 115429.

Schleich, W. P., 2001, *Quantum Optics in Phase Space* (Wiley-VCH, Berlin).

Schulz, S. A., U. Poschinger, F. Ziesel, and F. Schmidt-Kaler, 2008, "Sideband cooling and coherent dynamics in a microchip multi-segmented ion trap," New J. Phys. **10**, 045007.

Seifert, U., 2008, "Stochastic thermodynamics: Principles and perspectives," Eur. Phys. J. B **64**, 423.

Stratonovich, R. L., 1992, *Nonlinear Nonequilibrium Thermodynamics I: linear and Nonlinear Fluctuation-dissipation Theorems*, Springer Series in Synergetics Vol. 57 (Springer-Verlag, Berlin).

Stratonovich, R. L., 1994, *Nonlinear Nonequilibrium Thermodynamics II: Advanced Theory*, Springer Series in Synergetics Vol. 59 (Springer-Verlag, Berlin).

Sutherland, W., 1902, "Ionization, ionic velocities, and atomic sizes," Philos. Mag. **3**, 161.

Sutherland, W., 1905, "Dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin," Philos. Mag. **9**, 781.

Talkner, P., P. S. Burada, and P. Hänggi, 2008, "Statistics of work performed on a forced quantum oscillator," Phys. Rev. E **78**, 011115.

Talkner, P., P. S. Burada, and P. Hänggi, 2009, "Erratum: Statistics of work performed on a forced quantum oscillator" [Phys. Rev. E 78, 011115 (2008)], Phys. Rev. E **79**, 039902(E).

Talkner, P., M. Campisi, and P. Hänggi, 2009, "Fluctuation theorems in driven open quantum systems," J. Stat. Mech. P02025.

Talkner, P., and P. Hänggi, 2007, "The Tasaki-Crooks quantum fluctuation theorem," J. Phys. A **40**, F569.

Talkner, P., P. Hänggi, and M. Morillo, 2008, "Microcanonical quantum fluctuation theorems," Phys. Rev. E **77**, 051131.

Talkner, P., E. Lutz, and P. Hänggi, 2007, "Fluctuation theorems: Work is not an observable," Phys. Rev. E **75**, 050102.

Tasaki, H., 2000, "Jarzynski relations for quantum systems and some applications," arXiv:cond-mat/0009244.

Utsumi, Y., D. S. Golubev, M. Marthaler, K. Saito, T. Fujisawa, and G. Schön, 2010, "Bidirectional single-electron counting and the fluctuation theorem," Phys. Rev. B **81**, 125331.

Vaikuntanathan, S., and C. Jarzynski, 2008, "Escorted free energy simulations: Improving convergence by reducing dissipation," Phys. Rev. Lett. **100**, 190601.

Vedral, V., 2002, "The role of relative entropy in quantum information theory," Rev. Mod. Phys. **74**, 197.

Vilar, J. M. G., and J. M. Rubi, 2008a, Comment on "Failure of the work-Hamiltonian connection for free-energy calculations–Reply," Phys. Rev. Lett. **101**, 098902.

Vilar, J. M. G., and J. M. Rubi, 2008b, Comment on "Failure of the work-Hamiltonian connection for free-energy calculations"–Reply," Phys. Rev. Lett. **101**, 098904.

Vilar, J. M. G., and J. M. Rubi, 2008c, "Failure of the work-Hamiltonian connection for free-energy calculations," Phys. Rev. Lett. **100**, 020601.

Yukawa, S., 2000, "A quantum analogue of the Jarzynski equality," J. Phys. Soc. Jpn. **69**, 2367.

Zimanyi, E. N., and R. J. Silbey, 2009, "The work-Hamiltonian connection and the usefulness of the Jarzynski equality for free energy calculations," J. Chem. Phys. **130**, 171102.

Erratum: *Colloquium*: Quantum fluctuation relations: Foundations and applications [Rev. Mod. Phys. 83, 771 (2011)]

Michele Campisi, Peter Hänggi, and Peter Talkner

(published 19 December 2011)

DOI: 10.1103/RevModPhys.83.1653 PACS numbers: 05.30.−d, 05.40.−a, 05.60.Gg, 05.70.Ln, 99.10.Cd

The first line of Eq. (51) contains some typos: it correctly reads

$$G[u;\lambda] = \mathrm{Tr}\mathcal{T} e^{iu[\mathcal{H}_\tau^H(\lambda_\tau) - \mathcal{H}(\lambda_0)]} e^{-\beta\mathcal{H}(\lambda_0)}/Z(\lambda_0). \tag{51}$$

This compares with its classical analog, i.e., the second line of Eq. (27).

Quite surprisingly, notwithstanding the identity

$$\mathcal{H}_\tau^H(\lambda_\tau) - \mathcal{H}(\lambda_0) = \int_0^\tau dt\dot{\lambda}_t \frac{\partial \mathcal{H}_t^H(\lambda_t)}{\partial \lambda_t}, \tag{1}$$

one finds that generally

$$\mathcal{T} e^{iu[\mathcal{H}_\tau^H(\lambda_\tau) - \mathcal{H}(\lambda_0)]} \neq \mathcal{T} \exp\left[iu \int_0^\tau dt\dot{\lambda}_t \frac{\partial \mathcal{H}_t^H(\lambda_t)}{\partial \lambda_t}\right]. \tag{2}$$

As a consequence, it is not allowed to replace $\mathcal{H}_\tau^H(\lambda_\tau) - \mathcal{H}(\lambda_0)$, with $\int_0^\tau dt\dot{\lambda}_t \partial\mathcal{H}_t^H(\lambda_t)/\partial\lambda_t$ in Eq. (51). Thus, there is no quantum analog of the classical expression in the third line of Eq. (27). This is yet another indication that "work is not an observable" (Talkner, Lutz, and Hänggi, 2007)). This observation also corrects the second line of Eq. (4) of the original reference (Talkner, Lutz, and Hänggi, 2007).

The correct expression is obtained from the general formula

$$\mathcal{T}\exp[A(\tau) - A(0)] = \mathcal{T}\exp\left[\int_0^\tau dt \left(\frac{d}{dt} e^{A(t)}\right) e^{-A(t)}\right], \tag{3}$$

where $A(t)$ is any time dependent operator [in our case $A(t) = iu\mathcal{H}_t^H(\lambda_t)$]. Equation (3) can be proved by demonstrating that the operator expressions on either side of Eq. (3) obey the same differential equation with the identity operator as the initial condition. This can be accomplished by using the operator identity $de^{A(t)}/dt = \int_0^1 ds e^{sA(t)}\dot{A}(t)e^{(1-s)A(t)}$.

There are also a few minor misprints: (i) The symbol ds in the integral appearing in the first line of Eq. (55) should read dt. (ii) The correct year of the reference (Morikuni and Tasaki, 2010) is 2011 (not 2010).

The authors are grateful to Professor Yu. E. Kuzovlev for providing them with this insight, and for pointing out the error in the second line of Eq. (51).

REFERENCES

Talkner P., E. Lutz, and P. Hänggi, 2007, Phys. Rev. E **75**, 050102(R).

