Time-reversal symmetry in nonstationary Markov processes with application to some fluctuation theorems

Carolyne M. Van Vliet*
Department of Physics, University of Miami, P. O. Box 248046, Coral Gables, Florida 33124-0530, USA
(Received 10 August 2012; published 6 November 2012)

Nonequilibrium processes require that the density operator of an interacting system with Hamiltonian $\mathcal{H}(t) = \mathcal{H}^0(t) + \lambda\mathcal{V}$ converges and produces entropy. Employing projection operators in the state space, the density operator is developed to all orders of perturbation and then resummed. In contrast to earlier treatments by Van Hove [Physica **21**, 517 (1955)] and others [U. Fano, Rev. Mod. Phys. **29**, 74 (1959); U. Fano, in *Lectures on the Many-Body Problem*, Vol 2, edited by E. R. Caniello (Academic Press, New York, 1964); R. Zwanzig, in *Lectures in Theoretical Physics*, Vol. III, edited by W. E. Britten, B. W. Downs, and J. Downs (Wiley Interscience, New York, 1961), pp. 116–141; K. M. Van Vliet, J. Math. Phys. **19**, 1345 (1978); K. M. Van Vliet, Can. J. Phys. **56**, 1206 (1978)], closed expressions are obtained. From these we establish the time-reversal symmetry property $P(\gamma,t|\gamma',t') = \tilde{P}(\gamma',t'|\gamma,t)$, where the tilde refers to the time-reversed protocol; also a nonstationary Markovian master equation is derived. Time-reversal symmetry is then applied to thermostatted systems yielding the Crooks-Tasaki fluctuation theorem (FT) and the quantum Jarzynski work-energy theorem, as well as the general entropy FT. The quantum mechanical concepts of work and entropy are discussed in detail. Finally, we present a nonequilibrium extension of Mazo's lemma of linear response theory, obtaining some applications via this alternate route.

DOI: 10.1103/PhysRevE.86.051106 PACS number(s): 05.30.−d, 02.50.Ga, 05.70.Ln

I. INTRODUCTION

The past two decades have seen a strong interest in thermodynamic systems driven far from equilibrium. Generally, these systems are open, having exchange with large reservoirs. Thermodynamic potentials thus exist, as derived from the states of the reservoirs. In particular, as in Clausius's days, isothermal systems coupled to a heat bath have been studied extensively. The driving protocol, in which thermodynamic variables a_i or fields $F(t)$ are changed over a time interval $(t_0 \to t_1)$ is denoted by $\xi(t)$. In order for measurements at the beginning and end point of the interval to be meaningful, it must be assumed that the system equilibrates with the heat bath at those times. However, the canonical equilibrium at time t_1 and possibly also at t_0 is a *constrained* or *driven* equilibrium with respect to the variables varied in the protocol; obviously, the processes envisaged are neither quasistatic nor stationary.

While thermodynamic processes can be studied from a phenomenological or macroscopic point of view, the processes can only be reasonably understood with the methods of nonequilibrium statistical mechanics. Unfortunately, the explosion of papers on driven systems of the last epoch have mainly been based on the classical phase-space description; detailed quantum statistical considerations have been comparatively few. Second, in most quantum papers on the subject, the time evolution is linked to either the unperturbed von Neumann equation or to the behavior of Heisenberg operators, as in Kubo's linear response theory (LRT). However, it is well known that there is no increase of entropy based on the time development of the density operator in the von Neumann equation and any aspects of irreversibility are absent, *thus making the results circumspect as to their applicability to actual nonequilibrium processes*; we note in this regard that nonequilibrium thermodynamics was formerly called "irreversible thermodynamics." As abundantly shown in our previous work (to be cited below), we must work in the interaction picture based on a partitioned Hamiltonian $\mathcal{H}^0 + \lambda\mathcal{V}$ and employ time-dependent perturbation theory to all orders to obtain relevant results. This procedure was started in a seminal paper in *Physica* in 1955 by Leon Van Hove [1]. Whereas his was mostly a "brute-force" procedure in which a doubly infinite series for the exponentiated Liouville operator

$$\exp(-i\mathcal{L}t) = \{\exp[-(i/\hbar)(\mathcal{H}^0 + \lambda\mathcal{V})\,t]^{\rightarrow\ \leftarrow} \exp[(i/\hbar)(\mathcal{H}^0 + \lambda\mathcal{V})\,t]\} \tag{1.1}$$

was evaluated by specifically combining terms of the same order, better techniques employing projection operators were later developed by Fano [2], Zwanzig [3], and the author and collaborators [4–7], largely summarized in the nonequilibrium part of our recent book [8]. The main outcome of Van Hove's paper was that, in the long-time weak-coupling limit, the diagonal part of the so-obtained "reduced density operator" ρ^R satisfies the irreversible Pauli-Van Hove master equation (ME) with a Markovian conditional probability $P(\gamma,t|\gamma',t')$ connecting the quantum states $\{|\gamma_i\rangle\}$ of $\mathcal{H}^0$; the entropy production $-(d/dt)\mathrm{Tr}\rho^R \ln \rho^R$ is now positive semidefinite.

The driven systems in this paper are more complicated in that the processes considered are nonstationary. This is reflected in the evolution operator $U(t,t')$ which is not merely a function of $t - t'$. As will be discussed below, we need to consider both the forward protocol operating over the time interval $(t_0 \to t_1)$ and the backward protocol operating over $(t_1 \to t_0)$. The principal property of the conditional probability for this nonstationary process, now duly randomized by the interactions $\lambda\mathcal{V}(t)$, is the *time-reversal symmetry*,

$$P_t(\gamma_\xi,t|\gamma_{\xi'},t') = P_{-t}(\gamma_{\xi'},t'|\gamma_\xi,t); \tag{1.2}$$

for the subscripts on the states, see below. Both forward and backward processes still satisfy a ME, but the transition

*vanvliet@physics.miami.edu

probabilities must be labeled with the time points t or $-t$, respectively.

The time-reversal symmetry property is the main element that enters into the quantum statistical treatment of the many *nonequilibrium fluctuation theorems* (FT's) and the *work-energy theorems* developed since the mid-1990s. In the former—in the broadest sense—we compare probabilities of "action functionals" for forward processes with those for backward processes and the associated "uncompensated heat" (Clausius); in the latter we are concerned with the relationship of expended work to various thermodynamic potentials.

The quantum mechanical form Eq. (1.2) was first mentioned by the author in a recent paper [9] on the Crooks-Tasaki FT [10,11] and on the work–free energy relationship established by Christopher Jarzynski [12–15]; it was proven there for an unperturbed Hamiltonian, or, what is equivalent, in zero order for the perturbational Hamiltonian. The result (1.2) is also found in a very recent paper of Cohen and Imry, but again, only for a nonpartitioned Hamiltonian [16] and, therefore, not applicable to real irreversible processes. The main charge of this paper is to prove (1.2), or (2.6) below, as well as the associated ME, employing all orders of perturbation for the processes envisaged; this will comprise Secs. I to III. While the computations are still laborious, final closed-form results for the reduced evolution operator $U^R(t,t')$ will be established.

In the later sections of this paper, Secs. IV and V, we shall then apply this property to obtain thermodynamic results.[1] Also, we propose an extension of Mazo's lemma of LRT, giving an "updated" version of the quantum exponentiated work FT by Talkner, Lutz, and Hänggi [17–19]. Truly *microscopic* applications are limited, however, to processes for which the protocol $\xi(t)$ consists of c numbers. In other cases, the variables must be coarse grained since they usually do not commute with the Hamiltonian and with each other. The beautiful microscopic developments must then be diluted and a *mesoscopic* description emerges; results analogous to those of Ref. [9] will be recovered. Also, we obtain the general asymptotic FT for entropy production, cf. Galavotti and Cohen [20], Kurchan [21,22], Lebowitz and Spohn [23], Evans and Searles [24], and Harris and Schütz [25], among many, many others, see the alphabetical list in Ref. [25].

Finally, we must be more specific about the eigenstates employed in Eq. (1.2). Basically, we are dealing with a time-dependent Hamiltonian $\mathcal{H}[\xi(t)] = \mathcal{H}^0[\xi(t)] + \lambda\mathcal{V}(t)$. Usually, the interaction $\lambda\mathcal{V}$ (e.g., electron-phonon interaction or interaction with external agencies such as a heat bath) is independent of time. It will be expedient to consider a set of equivalent time-independent Hamiltonians $\mathcal{H}_\xi$ which over the time interval of the protocol are parametrized by ξ; the so-obtained eigenstates are the sets $\{|\gamma_\xi\rangle\}$. For many-body states both γ and ξ are quasicontinuous.

II. PROOF OF TIME-REVERSAL SYMMETRY

For the unperturbed Hamiltonian the result requires only a few lines. For the diagonal part of the density operator $\rho(t)$ we have with $t > t'$

$$p(\gamma_\xi,t) = \langle\gamma_\xi|U(t,t')\rho(t')U^\dagger(t,t')|\gamma_\xi\rangle = \sum_{\gamma'\gamma''}\langle\gamma_\xi|U(t,t')|\gamma'\rangle\,\langle\gamma'|\rho(t')|\gamma''\rangle\,\langle\gamma''|U^\dagger(t,t')|\gamma_\xi\rangle, \tag{2.1}$$

where we inserted the "decomposition of unity" $\sum|\gamma\rangle\langle\gamma| = \mathbf{1}$; note that U is the evolution operator

$$U(t,t') = \mathcal{T}\exp\left[-i\int_{t'}^{t}\mathcal{H}(\vartheta)d\vartheta/\hbar\right], \tag{2.2}$$

with $\mathcal{T}$ being the time-ordering operator. If we make an initial random phase assumption that $\rho(t')$ is diagonal[2], i.e., $\langle\gamma'|\rho(t')|\gamma''\rangle = p(\gamma',t')\delta_{\gamma'\gamma''}$, we obtain

$$p(\gamma_\xi,t) = \sum_{\gamma_{\xi'}}|\langle\gamma_\xi|U(t,t')|\gamma_{\xi'}\rangle|^2 p(\gamma_{\xi'},t'). \tag{2.3}$$

From Bayes's rule, we see that this implies a conditional probability P,

$$P(\gamma_\xi,t|\gamma_{\xi'},t') = |\langle\gamma_\xi|U(t,t')|\gamma_{\xi'}\rangle|^2. \tag{2.4}$$

Now, since

$$\langle\gamma_\xi|U(t,t')|\gamma_{\xi'}\rangle^* = \langle\gamma_{\xi'}|U^\dagger(t,t')|\gamma_\xi\rangle = \langle\gamma_{\xi'}|U(t',t)|\gamma_\xi\rangle, \tag{2.5}$$

where $U(t',t)$ is the backward evolution with inverse time ordering, it follows that

$$P(\gamma_\xi,t|\gamma_{\xi'},t') = \tilde{P}(\gamma_{\xi'},t'|\gamma_\xi,t), \tag{2.6}$$

where the tilde refers to the backward process (as did our previous subscript, "$-t$").

The above procedure is the quantum equivalent of the various treatments in which a trajectory is followed by the Liouville equation. However, the P here obtained is not Markovian and the full solution of the von Neumann equation involves a convolution integral over all previous times. The correct pathway to irreversibility requires that there are interactions $\lambda\mathcal{V}$ that randomize the motion proper, represented by the principal system Hamiltonian $\mathcal{H}^0$. This is similar as in Boltzmann's 1871 transport equation in which collisions perturbed the ponderomotive motion due to gradients and fields; this was the essence of his "H (capital eta for entropy)" theorem.

A. Projection operators and the state space

The full evolution is still given by the first line of Eq. (2.1), but between interactions the system evolves according to the zero-order unitary operator analogous to Eq. (2.2),

$$U^0(t,t') = \mathcal{T}\exp\left[-i\int_{t'}^{t}\mathcal{H}^0(\vartheta)d\vartheta/\hbar\right]. \tag{2.7}$$

Because of the time dependence, a straightforward resolvent expansion based on Laplace transformation cannot be employed. We will, therefore, stay in the time domain at

[1] These sections can be read independently, if one accepts the statement (2.6) for the irreversible P's without proof.

[2] The von Neumann equation is a first-order differential equation, for which this initial condition is legitimate; note that we do *not* make a repeated random-phase assumption (RRPA) in this paper, except in Sec. IIIA2 with qualification.

first and solve separately for the diagonal and nondiagonal parts. Laplace transforms will be used, however, to evaluate certain operator structures. Whereas the projection-operator procedure is least laborious when applied to the von Neumann equation, the commutators of that formalism need special scrutiny since the bilinear concomitant in Green's theorem is often nonzero, thus rendering the Dirac notation ambiguous unless some rules are established *a priori* ([8], Section 16.10). We shall therefore work directly in the state space $\mathcal{S}$ (dual space denoted by $\bar{\mathcal{S}}$) of $\mathcal{H}^0$, with the initial developments following the pattern of Appendix B of our paper [5].

1. Preliminaries

We, first, briefly consider the case where $\mathcal{H}^0$ and $\mathcal{V}$ are independent of time, $\mathcal{H}^0$ having eigenstates $\{|\gamma\rangle\}$. Any operator can be split into a diagonal part and a nondiagonal part as follows,

$$A = \sum_{\gamma'\gamma''} |\gamma'\rangle\langle\gamma'|A|\gamma''\rangle\langle\gamma''| = \sum_{\gamma'} |\gamma'\rangle\langle\gamma'|A|\gamma'\rangle\langle\gamma'| + \sum_{\gamma'\neq\gamma} |\gamma'\rangle\langle\gamma'|A|\gamma''\rangle\langle\gamma''| = \mathcal{P}A + (1-\mathcal{P})A. \quad (2.8)$$

Here, $\mathcal{P}$ is the projection superoperator on the diagonal Liouville subspace of $\mathcal{S}\otimes\bar{\mathcal{S}}$ and $1-\mathcal{P}$ on its complementary subspace. We have the usual rules $\mathcal{P}^2 = \mathcal{P}$ and $\mathcal{P}(1-\mathcal{P}) \equiv \mathcal{P}\mathcal{Q} = 0$. Explicitly from Eq. (2.8), we have the usual superoperator two-sided action form

$$\mathcal{P} = \sum_{\gamma} \{|\gamma\rangle\langle\gamma|^{\rightarrow}\ {}^{\leftarrow}|\gamma\rangle\langle\gamma|\}, \quad \text{or}$$
$$\mathcal{P}K = \sum_{\gamma} (|\gamma\rangle\langle\gamma|)\langle\gamma|K|\gamma\rangle, \quad (2.9)$$
$$\mathcal{Q} = \sum_{\gamma}\sum_{\gamma'} \{|\gamma\rangle\langle\gamma|^{\rightarrow}\ {}^{\leftarrow}|\gamma'\rangle\langle\gamma'|\}, \quad \gamma\neq\gamma'. \quad (2.9')$$

A product of two operators can be split as follows:

$$AB = \{(A_d + A_{\rm nd})(B_d + B_{\rm nd})\}_d + \{(A_d + A_{\rm nd})(B_d + B_{\rm nd})\}_{\rm nd} = (A_dB_d)_{(d)} + (A_dB_{\rm nd})_{(\rm nd)} + (A_{\rm nd}B_d)_{(\rm nd)} + (A_{\rm nd}B_{\rm nd})_d + (A_{\rm nd}B_{\rm nd})_{\rm nd}.$$

The subscripts in parentheses are automatically fulfilled and can be omitted. As to the Hamiltonian, we note that $\mathcal{H}^0$ only has a diagonal part, whereas $\mathcal{V}$ is nondiagonal (any possible diagonal part must be included in $\mathcal{H}^0$). Thus, from the standard differential equation for U we have

$$dU/dt = (1/\hbar i)\,[\mathcal{H}^0 + \lambda\mathcal{V}]\,U(t,t'), \quad (2.10)$$
$$\frac{dU_d(t,t')}{dt} = (1/\hbar i)\,[\mathcal{H}^0\,U_d(t,t') + \lambda(\mathcal{V}U_{\rm nd}(t,t'))_d] = (1/\hbar i)\,[\mathcal{H}^0\,U_d(t,t') + \lambda\mathcal{P}\mathcal{V}U_{\rm nd}(t,t')], \quad (2.11)$$
$$\frac{dU_{\rm nd}(t,t')}{dt} = (1/\hbar i)\,[\mathcal{H}^0U_{\rm nd}(t,t') + \lambda\mathcal{V}U_d(t,t') + \lambda(\mathcal{V}U_{\rm nd}(t,t'))_{\rm nd}] = (1/\hbar i)\,[\mathcal{H}^0U_{\rm nd}(t,t') + \lambda\mathcal{V}U_d(t,t') + \lambda(1-\mathcal{P})\mathcal{V}U_{\rm nd}(t,t')]; \quad (2.12)$$

we note that the superoperators $\mathcal{P}$ and $1-\mathcal{P}$ work on all operators to their right, unless otherwise indicated by parentheses. The equations (2.11) and (2.12) are fully equivalent to Eq. (2.10); they must be solved conjointly with the weak-coupling, long-time limit being carried out judiciously.

2. The state space

In order to parametrize the time dependence of $\mathcal{H}^0[\xi(t)]$, we shall represent the time by a set of discrete time points t_i $(i = 0,1,\ldots,n)$ on the interval with $t_{i=0} = t'$ and $t_{i=n} = t$. The value of ξ will be held constant over an interval, being equal to its value at the lower point. The Hamiltonian now becomes $\mathcal{H}^0[\vartheta] = \mathcal{H}^0_{\xi_i}$, $t_i \leqslant \vartheta < t_{i+1}$. This being a function of bounded variation, its integral still exists, giving for the evolution operator U^0,

$$U^0(t,t') = \mathcal{T}\exp\left[-i\int_{t'}^{t}\mathcal{H}^0(\vartheta)d\vartheta/\hbar\right] = \mathcal{T}\exp\sum_{i=0}^{n-1}\left[-i\mathcal{H}^0_{\xi_i}(t_{i+1}-t_i)/\hbar\right] = \prod_{i=0}^{n-1}\exp\left[-i\mathcal{H}^0_{\xi_i}(t_{i+1}-t_i)/\hbar\right] = \prod_{i=0}^{n-1}U^0(t_{i+1},t_i), \quad (2.13)$$

where we presume that the $\mathcal{H}^0$'s of the various time intervals commute, a likely proposition. The evolution operators on the subintervals are quasistationary, i.e., a function of the time differences only, $U^0(t_{i+1},t_i) = U^0(t_{i+1}-t_i)$, and the limits can be shifted for computational purposes; this is a purely mathematical device, however, and the original time limits must be restored to be useful in the overall result Eq. (2.13), with the subintervals being contiguous.

The full state space is the tensor product of the spaces for the individual Hamiltonians [the situation is not unlike the state space of the grand-canonical ensemble with eigenstates $|\eta_N,N\rangle$]. Accordingly, the full eigenstates are denoted by $|\varphi\rangle$ with

$$|\varphi\rangle = |\gamma,\xi\rangle \otimes |\gamma_{n-1},\xi_{n-1}\rangle \otimes |\gamma_{n-2},\xi_{n-2}\rangle \otimes\cdots\otimes |\gamma_1,\xi_1\rangle \otimes |\gamma',\xi'\rangle \equiv \left(\prod_{\otimes i}|\gamma_{\xi_i}\rangle\right), \quad (2.14)$$

where the far right-hand side is a short-cut notation to avoid overlabelling. The projection operators also involve the full state space with $\mathcal{P} = \Pi\,\mathcal{P}_i$. However, for any operator that acts only over an interval $(t_{\xi_i},t_{\xi_{i+1}})$ only $\mathcal{P}_i$ gives a matrix element that differs from unity. Since the expression for $\mathcal{P}$ [Eq. (2.9) with $|\gamma\rangle$ replaced by $|\varphi\rangle$] requires that we sum over all eigenstates $|\varphi\rangle$, we basically sum all projectors not involving $\mathcal{P}_i$, thus yielding a product of unit operators. Therefore, the diagonal projection for such operators only requires that we apply $\mathcal{P}_i$ and it suffices that we work in the sub-Liouville space $\mathcal{S}_i\otimes\bar{\mathcal{S}}_i$, thereby assuring that the procedure remains manageable.

B. Perturbational form for the evolution operator

1. The diagonal part

Let (τ',τ) be a time interval such that $t_i \leqslant \tau',\tau < t_{i+1}$. For the evolution operator on this interval we already found the equations for its diagonal part, (2.11), and for its nondiagonal part, (2.12); these are solved together. From Eq. (2.12) we readily obtain

$$U_{\rm nd}(\tau,\tau') = (1/\hbar i)\int_{\tau'}^{\tau} d\vartheta\, \mathcal{G}(\tau-\vartheta,\tau')\lambda\mathcal{V}U_d(\vartheta,\tau') = (1/\hbar i)\int_{\tau'}^{\tau} d\vartheta\, \mathcal{G}(\vartheta,\tau')\,\lambda\mathcal{V}U_d(\tau-\vartheta,\tau'), \quad (2.15)$$

where $\mathcal{G}$ is the Green's superoperator

$$\mathcal{G}(\tau,\tau') = \mathcal{I}\big(\exp\{(1/\hbar i)(\tau-\tau')\big[\mathcal{H}^0_{\xi_i}+\lambda(1-\mathcal{P})\mathcal{V}\big]\}\big)\,\Theta(\tau-\tau'), \quad (2.16)$$

with Θ being the Heaviside function and $\mathcal{I}$ the identity superoperator. Equation (2.15) is substituted into Eq. (2.11) to yield the integrodifferential equation

$$\frac{dU_d(\tau,\tau')}{dt} = \frac{1}{\hbar i}\mathcal{H}^0_{\xi_i}U_d(\tau,\tau') - (\lambda^2/\hbar^2)\times\int_{\tau'}^{\tau} d\vartheta\,\mathcal{P}\mathcal{V}\mathcal{G}(\vartheta,\tau')\,\mathcal{V}U_d(\tau-\vartheta,\tau'). \quad (2.17)$$

This result is still exact. We now impose the long-time, weak-coupling (or Van Hove) limit,

$$\lambda\to 0,\quad \tau\to\infty,\quad \lambda^2\tau \quad \text{is} \quad \text{finite;} \quad (2.18)$$

(in reality λ remains small but finite, while τ is just large in comparison with the microscopic transition times, say $\tau_{\rm tr}\sim 10^{-16}$s). The Green's superoperator is then approximated by

$$\mathcal{G}^0(\tau,\tau') = \mathcal{I}\big(\exp\{(1/\hbar i)(\tau-\tau')\mathcal{H}^0_{\xi_i}\}\big)\,\Theta(\tau-\tau'). \quad (2.19)$$

For U_d we now have

$$\frac{dU_d(\tau,\tau')}{d\tau} = \frac{1}{\hbar i}\mathcal{H}^0_{\xi_i}U_d(\tau,\tau') - \frac{\lambda^2}{\hbar^2}\int_{\tau'}^{\tau} d\vartheta\,\mathcal{P}\mathcal{V}e^{-i(\vartheta-\tau')\mathcal{H}^0_{\xi_i}/\hbar}\times\mathcal{V}U_d(\tau-\vartheta,\tau'). \quad (2.20)$$

We note that the integral is of order $\lambda^2\tau$. Further, we set $W = \exp[i\tau\mathcal{H}^0_{\xi_i}/\hbar]U_d$ and we let, at first, $\tau'=0$ (meaning we shift axes, relabeling $\tau-\tau'\to\bar\tau$); then, for W,

$$\frac{dW(\bar\tau,0)}{d\bar\tau} = -\frac{\lambda^2}{\hbar^2}e^{i\bar\tau\mathcal{H}^0_{\xi_i}/\hbar}\int_0^{\bar\tau} d\vartheta\,\mathcal{P}\mathcal{V}e^{-i\vartheta\mathcal{H}^0_{\xi_i}/\hbar}\times\mathcal{V}e^{-i(\bar\tau-\vartheta)\mathcal{H}^0_{\xi_i}/\hbar}W(\bar\tau-\vartheta,0). \quad (2.21)$$

Writing $\mathcal{P}$ in full as in Eq. (2.9) and inserting closure $\sum|\bar\gamma_{\xi_i}\rangle\langle\bar\gamma_{\xi_i}| = \mathbf{1}$ before the second $\mathcal{V}$, we get

$$\begin{aligned}\frac{dW(\bar\tau,0)}{d\bar\tau} &= -\frac{\lambda^2}{\hbar^2}\sum_{\gamma_{\xi_i},\bar\gamma_{\xi_i}}|\gamma_{\xi_i}\rangle\langle\gamma_{\xi_i}|e^{i\varepsilon_{\gamma\xi_i}\bar\tau/\hbar}\int_0^{\bar\tau}d\vartheta\,\langle\gamma_{\xi_i}|\mathcal{V}|\bar\gamma_{\xi_i}\rangle\\ &\quad\times e^{-i\varepsilon_{\bar\gamma\xi_i}\vartheta/\hbar}\langle\bar\gamma_{\xi_i}|\mathcal{V}|\gamma_{\xi_i}\rangle e^{-i\varepsilon_{\gamma\xi_i}(\bar\tau-\vartheta)/\hbar}W(\bar\tau-\vartheta,0)\\ &= -\frac{\lambda^2}{\hbar^2}\sum_{\gamma_{\xi_i},\bar\gamma_{\xi_i}}|\gamma_{\xi_i}\rangle\langle\gamma_{\xi_i}|\int_0^{\bar\tau}d\vartheta\,|\langle\gamma_{\xi_i}|\mathcal{V}|\bar\gamma_{\xi_i}\rangle|^2\\ &\quad\times e^{-i(\varepsilon_{\bar\gamma\xi_i}-\varepsilon_{\gamma\xi_i})\vartheta/\hbar}\,W(\bar\tau-\vartheta,0). \end{aligned}\quad (2.22)$$

The $\varepsilon's$ are the eigenvalues (EV) of $\mathcal{H}^0_{\xi_i}$ on the chosen interval; note that for ease of notation we placed the state symbol γ and the interval label ξ_i of the EV on the same level.

Taking now the Laplace transform of Eq. (2.22) and denoting the transform by $\hat W(s)$, we have

$$s\hat W(s) - W(0) = -\frac{\lambda^2}{\hbar^2}\sum_{\gamma_{\xi_i},\bar\gamma_{\xi_i}}|\gamma_{\xi_i}\rangle\langle\gamma_{\xi_i}|\frac{|\langle\gamma_{\xi_i}|\mathcal{V}|\bar\gamma_{\xi_i}\rangle|^2}{s+i\big(\varepsilon_{\bar\gamma\xi_i}-\varepsilon_{\gamma\xi_i}\big)/\hbar}\,\hat W(s). \quad (2.23)$$

Considering asymptotic times implies $s\to 0+$; thus, we can employ the well-known identity

$$\frac{1}{s+i\big(\varepsilon_{\bar\gamma\xi_i}-\varepsilon_{\gamma\xi_i}\big)/\hbar} = \hbar\left[-i\,\mathsf{P}\frac{1}{\varepsilon_{\bar\gamma\xi_i}-\varepsilon_{\gamma\xi_i}}+\pi\delta\big(\varepsilon_{\bar\gamma\xi_i}-\varepsilon_{\gamma\xi_i}\big)\right], \quad (2.24)$$

where P denotes the principal value. Further, letting $\Sigma_{\gamma_{\xi_i}}\to\int\Delta\gamma_{\xi_i}$, where $\Delta\gamma$ includes the density of states, we set

$$C(\gamma_{\xi_i}) = 2\pi\int\Delta\bar\gamma_{\xi_i}|\langle\gamma_{\xi_i}|\mathcal{V}|\bar\gamma_{\xi_i}\rangle|^2\delta\big(\varepsilon_{\bar\gamma\xi_i}-\varepsilon_{\gamma\xi_i}\big), \quad (2.25)$$

$$D(\gamma_{\xi_i}) = 2\int\Delta\bar\gamma_{\xi_i}|\langle\gamma_{\xi_i}|\mathcal{V}|\bar\gamma_{\xi_i}\rangle|^2\,\mathsf{P}/\big(\varepsilon_{\bar\gamma\xi_i}-\varepsilon_{\gamma\xi_i}\big). \quad (2.26)$$

Solving from Eq. (2.23) for $\hat W(s)$ and noting that $W(0)=1$, the inverse yields, upon restoring τ',

$$W(\tau,\tau') = \exp\left\{-(\lambda^2/2\hbar)(\tau-\tau')\times\sum_{\gamma_{\xi_i}}|\gamma_{\xi_i}\rangle\langle\gamma_{\xi_i}|\,[C(\gamma_{\xi_i})-iD(\gamma_{\xi_i})]\right\}. \quad (2.27)$$

By series expansion, one may verify that, for any $f(\gamma)$,

$$\exp\left[-\sum_{\gamma}|\gamma\rangle\langle\gamma|\,f(\gamma)\right] = \sum_{\gamma}|\gamma\rangle\langle\gamma|\exp[-f(\gamma)]. \quad (2.28)$$

Whence, finally, for U_d,

$$U_d(\tau,\tau') = \sum_{\gamma_{\xi_i}}|\gamma_{\xi_i}\rangle\langle\gamma_{\xi_i}|\{\exp\{-(i/\hbar)\varepsilon_{\gamma\xi_i}(\tau-\tau') - (\lambda^2/2\hbar)(\tau-\tau')\,[C(\gamma_{\xi_i})-iD(\gamma_{\xi_i})]\}. \quad (2.29)$$

We note that the miracle has happened: *the original oscillatory behavior has been altered into a decaying behavior showing irreversibility*; note that $-(\lambda^2/2\hbar)C(\gamma_{\xi_i})$ is negative semidefinite, being an integral over the transition rates $W_{t_{\xi i}}(\gamma_{\xi_i},\gamma_{\xi'})$, as given by Fermi's golden rule. We also give (2.29) in the more compact form, noting that the decomposition theorem allows us to resum,

$$U_d(\tau,\tau') = \exp\big\{-(i/\hbar)(\tau-\tau')\mathcal{H}^0_{\xi_i} - (\lambda^2/2\hbar)(\tau-\tau')\times\big[C\big(\mathcal{H}^0_{\xi_i}\big)-iD\big(\mathcal{H}^0_{\xi_i}\big)\big]\big\}. \quad (2.30)$$

Extending (τ,τ') to the limits (t_{i+1},t_i), we obtain the diagonal part of the evolution operator for the entire interval. From Eq. (2.30), affixing the superscript R for *reduced* operator, we

obtained

$$U_d^R(t,t') = \prod_i \exp\left\{-(i/\hbar)\mathcal{H}^0_{\xi_i}(t_{i+1}-t_i) - (\lambda^2/2\hbar)(t_{i+1}-t_i) \times \left[C\left(\mathcal{H}^0_{\xi_i}\right) - iD\left(\mathcal{H}^0_{\xi_i}\right)\right]\right\}. \quad (2.31)$$

2. *The nondiagonal part*

For the nondiagonal part we may not approximate $\mathcal{G}$ by $\mathcal{G}^0$, since all orders $(\lambda\tau)^n$ occur. Letting $\tau' = 0$, we have for the Laplace transform of Eq. (2.15) with Eq. (2.16)

$$\hat{U}_{\rm nd}(s) = \frac{1}{\hbar i}\,\frac{1}{s + (i/\hbar)\left[\mathcal{H}^0_{\xi_i} + \lambda(1-\mathcal{P})\mathcal{V}\right]}\,\lambda(1-\mathcal{P})\mathcal{V}\hat{U}_d(s), \quad (2.32)$$

where in front of the last $\mathcal{V}$ we can add the $(1-\mathcal{P})$ with impunity. We seek to obtain a convolution series in the diagonal part U_d. To that effect, we rewrite (2.32) as follows,

$$\hat{U}_{\rm nd}(s) = \frac{1}{s + (i/\hbar)\mathcal{H}^0_{\xi_i} + (\lambda^2/\hbar)\mathcal{K} - (1/\hbar i)[\lambda(1-\mathcal{P})\mathcal{V} + i\lambda^2\mathcal{K}]}(1/\hbar i)\,\lambda(1-\mathcal{P})\mathcal{V}\hat{U}_d(s), \quad (2.33)$$

where we introduced a new operator,

$$\mathcal{K} = \Sigma_{\gamma_{\xi_i}} |\gamma_{\xi_i}\rangle\langle\gamma_{\xi_i}|[C(\gamma_{\xi_i}) - iD(\gamma_{\xi_i})] = \left[C\left(\mathcal{H}^0_{\xi_i}\right) - iD\left(\mathcal{H}^0_{\xi_i}\right)\right]. \quad (2.34)$$

Also, from Eq. (2.30) we find

$$\hat{U}_d(s) = \frac{1}{s + (i/\hbar)\mathcal{H}^0_{\xi_i} + (\lambda^2/\hbar)\mathcal{K}}. \quad (2.35)$$

Notice the similarity of the denominators in Eqs. (2.33) and (2.35).

Now the following operator identity is easily verified, $B^{-1} = A^{-1} + A^{-1}(A-B)\,B^{-1}$. This can be iterated *ad infinitum*. We, thus, obtain the perturbation series,

$$\hat{U}_{\rm nd}(s) = \hat{U}_d(s)\sum_{n=0}^{\infty}\left\{\frac{1}{\hbar i}[\lambda(1-\mathcal{P})\mathcal{V} + i\lambda^2\mathcal{K}]\,\hat{U}_d(s)\right\}^n \times \frac{1}{\hbar i}(1-\mathcal{P})\mathcal{V}\hat{U}_d(s). \quad (2.36)$$

In the square brackets we neglect the term with λ^2 versus the term in λ, to obtain

$$\hat{U}_{\rm nd}(s) = \hat{U}_d(s)\sum_{n=1}^{\infty}\left\{\frac{\lambda}{\hbar i}[(1-\mathcal{P})\mathcal{V}\hat{U}_d(s)\right\}^n. \quad (2.37)$$

Since the product operator $\mathcal{V}U_d(s)$ is clearly nondiagonal, it is destroyed by the $\mathcal{P}$, so we can *a posteriori* simplify to

$$\hat{U}_{\rm nd}(s) = \hat{U}_d(s)\sum_{n=1}^{\infty}\{(\lambda/\hbar i)\mathcal{V}\hat{U}_d(s)\}^n. \quad (2.38)$$

Its inversion yields a time-ordered $(n+1)-$fold convolution; hence,

$$U_{\rm nd}(\bar{\tau}) = \sum_{n=1}^{\infty}\left(\frac{\lambda}{\hbar i}\right)^n\int_0^{\bar{\tau}} d\vartheta_n\int_0^{\vartheta_n} d\vartheta_{n-1}\int_0^{\vartheta_{n-1}} d\vartheta_{n-2}\cdots \times\int_0^{\vartheta_2} d\vartheta_1 U_d(\bar{\tau}-\vartheta_n)\mathcal{V}U_d(\vartheta_n-\vartheta_{n-1})\mathcal{V} \times U_d(\vartheta_{n-1}-\vartheta_{n-2})\cdots\mathcal{V}U_d(\vartheta_2-\vartheta_1)\mathcal{V}U_d(\vartheta_1). \quad (2.39)$$

In contrast to Van Hove's series for the *matrix elements*, this series of *operators* can be summed!

To this purpose, we introduce the interaction operator,

$$U^I_{\rm nd}(\bar{\tau}) = U_d^{-1}(\bar{\tau})\,U_{\rm nd}(\bar{\tau})\,, \quad (2.40)$$

with superscript "-1" denoting the inverse; also let $\mathcal{V}^I(\bar{\tau}) = U_d^{-1}(\bar{\tau})\mathcal{V}\,U_d(\bar{\tau})$. Noting the composition property

$$U_d^{-1}(\bar{\tau}) = U_d^{-1}(\bar{\tau}-\vartheta_n)\,U_d^{-1}(\vartheta_n)\,, \quad (2.41)$$

and the time-inversion property[3]

$$U_d^{-1}(\bar{\tau}-\vartheta) = U_d(\vartheta-\bar{\tau})\,, \quad (2.42)$$

one finds

$$U^I_{\rm nd}(\bar{\tau}) = \sum_{n=1}^{\infty}\left(\frac{\lambda}{\hbar i}\right)^n\int_0^{\bar{\tau}} d\vartheta_n\int_0^{\vartheta_n} d\vartheta_{n-1}\int_0^{\vartheta_{n-1}} d\vartheta_{n-2}\cdots \times\int_0^{\vartheta_2} d\vartheta_1\;\mathcal{V}^I(\vartheta_n)\,\mathcal{V}^I(\vartheta_{n-1})\,\mathcal{V}^I(\vartheta_{n-2})\cdots \times\mathcal{V}^I(\vartheta_1). \quad (2.43)$$

This can be changed into the equivalent form, resetting the time axis,

$$U^I_{\rm nd}(\tau,\tau') = \frac{1}{n!}\sum_{n=1}^{\infty}\left(\frac{\lambda}{\hbar i}\right)^n\mathcal{T}\int_{\tau'}^{\tau} d\vartheta_n\int_{\tau'}^{\tau} d\vartheta_{n-1}\int_{\tau'}^{\tau} d\vartheta_{n-2} \times\cdots\int_{\tau'}^{\tau} d\vartheta_1\,\mathcal{V}^I(\vartheta_n,\tau')\,\mathcal{V}^I(\vartheta_{n-1},\tau') \times\cdots\mathcal{V}^I(\vartheta_1,\tau'). \quad (2.44)$$

Going back to $U_{\rm nd}$, we obtain

$$U_{\rm nd}(\tau,\tau') = U_d(\tau,\tau')\left\{\mathcal{T}\exp\left[(\lambda/\hbar i)\int_{\tau'}^{\tau} d\vartheta\;U_d^{-1}(\vartheta,\tau') \times\mathcal{V}U_d(\vartheta,\tau')\right] - 1\right\}. \quad (2.45)$$

[3]Time inversion shall also involve reversal of the magnetic field $\vec{H}$, when present.

Adding now the diagonal part, *the surprising final result is*

$$U(\tau,\tau') = U_d(\tau,\tau')\left\{\mathcal{T}\exp\left[(\lambda/\hbar i)\int_{\tau'}^{\tau} d\vartheta\; U_d^{-1}(\vartheta,\tau')\mathcal{V}U_d(\vartheta,\tau')\right]\right\}, \tag{2.46}$$

with U_d given by Eq. (2.30). We can extend the result to the entire subinterval, $\tau = t_{i+1}$ and $\tau' = t_i$.

3. *Reduced forms*

For the full reduced (i.e., postperturbation procedure) evolution operator we obtained

$$U^R(t,t') = \left(\prod_i \exp\left\{-(i/\hbar)(t_{i+1}-t_i)\left[\mathcal{H}^0_{\xi_i} - (\lambda^2/2)D\left(\mathcal{H}^0_{\xi_i}\right)\right]\right\}\right)\left(\prod_i \exp\left[-(\lambda^2/2\hbar)(t_{i+1}-t_i)\,C\left(\mathcal{H}^0_{\xi_i}\right)\right]\right)$$
$$\times\left\{\mathcal{T}\exp\left[-i\,(\lambda/\hbar)\int_{t'}^{t} d\vartheta\; U_d^{-1}(\vartheta,t')\mathcal{V}U_d(\vartheta,t')\right]\right\}. \tag{2.47}$$

We notice that the last factor has a structure quite similar to the point of departure, Eq. (2.2). However, the second factor shows that the reduced evolution operator is *no longer unitary*.

We must now again consider the matrix elements, the analog of Eq. (2.4) being

$$P(\gamma_\xi,t|\gamma_{\xi'},t') = |\langle\gamma_\xi|U^R(t,t')|\gamma_{\xi'}\rangle|^2. \tag{2.48}$$

From Eq. (2.47) we have, splitting off the first and the last factors in the product,

$$\langle\gamma_\xi|U^R(t,t')|\gamma_{\xi'}\rangle = \langle\gamma_\xi|\exp\{-(i/\hbar)(t-t_{n-1})\left[\varepsilon_{\gamma\xi} - (\lambda^2/2)D(\gamma_\xi)\right]\}\mathbf{exp}[-(\lambda^2/2\hbar)\,(t-t_{n-1})C(\gamma_\xi)]$$
$$\times\left(\prod_{i=1}^{n-2}\exp\left\{-(i/\hbar)(t_{i+1}-t_i)\left[\mathcal{H}^0_{\xi_i} - (\lambda^2/2)D\left(\mathcal{H}^0_{\xi_i}\right)\right]\right\}\right)\left(\prod_{i=1}^{n-2}\mathbf{exp}\left[-(\lambda^2/2\hbar)(t_{i+1}-t_i)\,C\left(\mathcal{H}^0_{\xi_i}\right)\right]\right)$$
$$\times\left\{\mathcal{T}\exp\left[-i\,(\lambda/\hbar)\int_{t'}^{t} d\vartheta\; U_d^{-1}(\vartheta)\mathcal{V}U_d(\vartheta)\right]\right\}\exp\{-(i/\hbar)(t_1-t')\left[\varepsilon_{\gamma\xi'} - (\lambda^2/2)D(\gamma_{\xi'})\right]\}$$
$$\times\,\mathbf{exp}[-(\lambda^2/2\hbar)\,(t_1-t')C(\gamma_{\xi'})]|\gamma_{\xi'}\rangle. \tag{2.49}$$

Next, we consider the complex conjugate matrix element,

$$\langle\gamma_\xi|U^R(t,t')|\gamma_{\xi'}\rangle^* = \langle\gamma_{\xi'}|U^{R\dagger}(t,t')|\gamma_\xi\rangle. \tag{2.50}$$

For $U^{R\dagger}(t,t')$ all unitary parts in Eq. (2.49) simply require $i \to -i$, which is tantamount to time reversal (or transposition, denoted by the superscript tr). Also note that the time-ordered integral, apart from being complex conjugated, retains its form with the integrand being the same cf. (2.42)

$$\left[U_d^{-1}(\vartheta,t_i)\mathcal{V}U_d(\vartheta,t_i)\right]^{\mathrm{tr}} = U_d^{-1}(\vartheta,t_i)\mathcal{V}U_d(\vartheta,t_i). \tag{2.51}$$

The state of the exponentials with *real* arguments bears more scrutiny; for clarity, we set the exp symbol in bold. Under Hermitian conjugation, these remain invariant! To see this, we must go back to Eqs. (2.20), (2.22), and (2.24). Hermitian conjugation of Eqs. (2.20) and (2.22) requires complex conjugation of Eq. (2.24), which leaves the δ part unchanged, since

$$\frac{1}{s \pm i\left(\varepsilon_{\bar\gamma\xi_i} - \varepsilon_{\gamma\xi_i}\right)/\hbar} = \mp i\hbar\,\mathsf{P}\frac{1}{\varepsilon_{\bar\gamma\xi_i} - \varepsilon_{\gamma\xi_i}} + \pi\hbar\delta\left(\varepsilon_{\bar\gamma\xi_i} - \varepsilon_{\gamma\xi_i}\right). \tag{2.52}$$

Therefore, the arguments $(-\lambda^2/2\hbar)\Delta t_{\xi i}C(\gamma_{\xi i})$, with the $C's$ being the integrals (2.25), are unchanged. However, the entire form Eq. (2.49) only holds for the forward process since we solved for forward time, cf. the Heaviside function for each subinterval in Eqs. (2.16) and (2.19). The real argument exponentials ensure forward irreversibility; from the ME (next section) we can prove that the entropy increases during the exercise of the protocol.

We now turn to the reverse process. The differential equations of departure for $U_d(\tau,\tau')$ and $U_{\mathrm{nd}}(\tau,\tau')$ still hold but they must be solved for backward time $d/d(-\tau)$, using the *adjunct* Green's operator, $\tilde{\mathcal{G}}(\tau',\tau)$ [26], the tilde indicating that τ precedes τ'. We can then find the time-reversed result $\langle\gamma_{\xi'}|\tilde{U}^R(t',t)|\gamma_\xi\rangle$; in the **exp** factors the times are now interchanged, the argument still being negative semi-definite (viz. $+(\lambda^2/2\hbar)(t_i - t_{i+1})C \leqslant 0$) . Whence we established

$$\langle\gamma_{\xi'}|U^{R\dagger}(t,t')|\gamma_\xi\rangle \to \langle\gamma_{\xi'}|\tilde{U}^R(t',t)|\gamma_\xi\rangle, \tag{2.53}$$

where the tilde on the U^R is *essential*. Altogether, we ascertained,

$$\langle\gamma_\xi|U^R(t,t')|\gamma_{\xi'}\rangle^* = \langle\gamma_{\xi'}|\tilde{U}^R(t',t)|\gamma_\xi\rangle. \tag{2.54}$$

Multiplying both sides with their complex conjugates, we definitively confirmed Eq. (2.6),

$$P(\gamma_\xi,t|\gamma_{\xi'},t') = |\langle\gamma_\xi|U^R(t,t')|\gamma_{\xi'}\rangle|^2 = |\langle\gamma_{\xi'}|\tilde{U}^R(t',t)|\gamma_\xi\rangle|^2$$
$$= \tilde{P}(\gamma_{\xi'},t'|\gamma_\xi,t). \quad \text{QED}$$

III. THE MASTER EQUATION

The purpose of this section is to show that the obtained postperturbation conditional probability $P(\gamma_\xi,t|\gamma_{\xi'},t')$ is Markovian and as such satisfies the master equation, postulated in Ref. [9]. We will only do this for the forward time process; the backward time equation can be obtained along the same lines but is seldom needed. The superscripts R will henceforth be omitted.

A. The quasistationary case

We, first, discuss the ME for a quasistationary process $\tau' \to \tau$, with τ and τ' being, again, times in an interval $t_i \leqslant \tau',\tau < t_{i+1}$, the zero-order Hamiltonian being $\mathcal{H}^0_{\xi_i}$. Again, let us consider the perturbational solution of the von Neumann equation for $\tau > \tau'$,

$$\rho(\tau) = U(\tau,\tau')\rho(\tau')\,U^\dagger(\tau,\tau'). \tag{3.1}$$

Assuming the initial operator to be diagonal, we have

$$\rho(\tau) = \sum_{\bar\gamma_{\xi'}} U(\tau,\tau')|\bar\gamma_{\xi'}\rangle\; p(\bar\gamma_{\xi'},\tau')\,\langle\bar\gamma_{\xi'}|\,U^\dagger(\tau,\tau'). \tag{3.2}$$

The states $|\bar\gamma_{\xi'}\rangle$ are eigenstates of the Hamiltonian near the initial point τ' ; the subscript ξ' will henceforth be suppressed. For $U(\tau,\tau')$ we shall use Eq. (2.46); for its Hermitian conjugate $U^\dagger(\tau,\tau')$ we keep forward times. Hence, from Eqs. (2.46), (2.51), and (3.2),

$$\begin{aligned}\rho(\tau) = \sum_{\bar\gamma} U_d(\tau,\tau')\,\mathcal{T}(\vartheta)\left\{\exp\left[-i\,(\lambda/\hbar)\int_{\tau'}^{\tau} d\vartheta\; U_d^{-1}(\vartheta,\tau')\mathcal{V}U_d(\vartheta,\tau')\right]\right\}|\bar\gamma\rangle\,\langle\bar\gamma| \\ \times\,\mathcal{T}(\bar\vartheta)\left\{\exp\left[i\,(\lambda/\hbar)\int_{\tau'}^{\tau} d\bar\vartheta\; U_d^{-1}(\bar\vartheta,\tau')\mathcal{V}U_d(\bar\vartheta,\tau')\right]\right\}U_d^\dagger(\tau,\tau')\cdot p(\bar\gamma,\tau')\,.\end{aligned} \tag{3.3}$$

We proceed to obtain the derivatives of the time-ordered operators, which act on the initial time projectors $|\bar\gamma\rangle\,\langle\bar\gamma|$. For the first time-ordered exponential operator we have

$$\mathcal{T}(\vartheta)\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{-i\lambda}{\hbar}\right)^n\int_{\tau'}^{\tau}d\vartheta_n\int_{\tau'}^{\tau}d\vartheta_{n-1}\cdots\int_{\tau'}^{\tau}d\vartheta_1\,U_d^{-1}(\vartheta_n,\tau')\mathcal{V}U_d(\vartheta_n,\tau')\cdots U_d^{-1}(\vartheta_1,\tau')\mathcal{V}U_d(\vartheta_1,\tau'). \tag{3.4}$$

Although U_d is in the integrands, we will see below that the parts to be retained are unitary. We will adopt a time-reversed order for the integrals, writing instead

$$\mathcal{T}^{-1}(\vartheta)\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{-i\lambda}{\hbar}\right)^n\int_{\tau'}^{\tau}d\vartheta_1\int_{\tau'}^{\tau}d\vartheta_2\;\ldots\;\int_{\tau'}^{\tau}d\vartheta_n\,U_d^{-1}(\vartheta_1,\tau')\mathcal{V}U_d(\vartheta_1,\tau')\cdots U_d^{-1}(\vartheta_n,\tau')\mathcal{V}U_d(\vartheta_n,\tau')\,, \tag{3.5}$$

where $\mathcal{T}^{-1}$ means that the equivalent convolution form still starts with $\int_{\tau'}^{\tau}d\vartheta_n\int_{\tau'}^{\vartheta_n}d\vartheta_{n-1}\;\ldots$. For the product operators we then obtain

$$\begin{aligned}\mathcal{T}^{-1}(\ldots)|\bar\gamma\rangle\,\langle\bar\gamma|\mathcal{T}(\ldots) = \sum_{n=0}^{\infty}\left(\frac{1}{n!}\right)^2\left(\frac{\lambda^2}{\hbar^2}\right)^n\mathcal{T}^{-1}(\vartheta)\mathcal{T}(\bar\vartheta)\int_{\tau'}^{\tau}d\vartheta_1\;\ldots\;\int_{\tau'}^{\tau}d\vartheta_{n-1}\prod_{i=1}^{n-1}U_d^{-1}(\vartheta_i,\tau')\mathcal{V}U_d(\vartheta_i,\tau') \\ \times\left(\int_{\tau'}^{\tau}d\vartheta_n U_d^{-1}(\vartheta_n,\tau')\mathcal{V}U_d(\vartheta_n,\tau')|\bar\gamma\rangle\,\langle\bar\gamma|\int_{\tau'}^{\tau}d\bar\vartheta_n\;U_d^{-1}(\bar\vartheta_n,\tau')\mathcal{V}U_d(\bar\vartheta_n,\tau')\right) \\ \times\int_{\tau'}^{\tau}d\bar\vartheta_{n-1}\int_{\tau'}^{\tau}d\bar\vartheta_{n-2}\;\ldots\int_{\tau'}^{\tau}d\bar\vartheta_1\prod_{i=n-1}^{1}U_d^{-1}(\bar\vartheta_i,\tau')\mathcal{V}U_d(\bar\vartheta_i,\tau').\end{aligned} \tag{3.6}$$

The double integral in the large round brackets is readily evaluated. First, from Eq. (2.30), we note that

$$U_d(\vartheta_n,\tau') = \exp\left\{-(i/\hbar)(\vartheta_n-\tau')\left[\mathcal{H}^0_{\xi i}+O(\lambda^2)\right]\right\}, \tag{3.7}$$

$$U_d^{-1}(\vartheta_n,\tau') = \exp\left\{(i/\hbar)(\vartheta_n-\tau')\left[\mathcal{H}^0_{\xi i}+O(\lambda^2)\right]\right\}. \tag{3.8}$$

The terms with $O(\lambda^2)$ will be omitted (*only* in the time-ordered integrals being dealt with here) since further computation would show that they give $\lambda^4\tau$ terms that should not be retained; whence the modified U_d and U_d^{-1} are unitary. Next we insert a closure sum $\sum_{\gamma'}|\gamma'\rangle\,\langle\gamma'| = \mathbf{1}$ before the first $\mathcal{V}$ and another one $\sum_{\gamma''}|\gamma''\rangle\langle\gamma''| = \mathbf{1}$ behind the second $\mathcal{V}$. The double integral now yields

$$\begin{aligned}(\;) &= \sum_{\gamma'}\sum_{\gamma''}|\gamma'\rangle\int_{\tau'}^{\tau}d\vartheta_n e^{(i/\hbar)(\vartheta_n-\tau')(\varepsilon_{\gamma'}-\varepsilon_{\bar\gamma})}\langle\gamma'|\mathcal{V}|\bar\gamma\rangle \\ &\quad\times\int_{\tau'}^{\tau}d\bar\vartheta_n e^{(i/\hbar)(\bar\vartheta_n-\tau')(\varepsilon_{\bar\gamma}-\varepsilon_{\gamma''})}\langle\bar\gamma|\mathcal{V}|\gamma''\rangle\,\langle\gamma''| \\ &\doteq \sum_{\gamma'}|\gamma'\rangle\langle\gamma'|\int_{\tau'}^{\tau}d\vartheta_n\int_{\tau'}^{\tau}d\bar\vartheta_n\,e^{(i/\hbar)(\vartheta_n-\bar\vartheta_n)(\varepsilon_{\gamma'}-\varepsilon_{\bar\gamma})} \\ &\quad\times|\langle\gamma'|\mathcal{V}|\bar\gamma\rangle|^2;\end{aligned} \tag{3.9}$$

in the conditional equality $\doteq$ we omitted the zero-trace terms $\Sigma_{\gamma'}\Sigma_{\gamma''\neq\gamma'}|\gamma'\rangle\cdots\langle\gamma''|$ which do not contribute to the density operator. Thus, we only retain terms with $\gamma'=\gamma''$. For large τ this double integral yields $2\pi\hbar(\tau-\tau')\,\delta(\varepsilon_{\gamma'}-\varepsilon_{\bar\gamma})$; note the result is linear in τ. With the usual transition rate

$$W(\gamma'|\bar\gamma)=(2\pi\lambda^2/\hbar)|\langle\gamma'|\mathcal{V}|\bar\gamma_{\xi'}\rangle|^2\,\delta(\varepsilon_{\gamma'}-\varepsilon_{\bar\gamma}), \quad (3.10)$$

(where entries behind the vertical bar are the data of departure), we see that the double integral has given us

$$\frac{\partial}{\partial\tau}\Big(\ \Big)=(\lambda^2/\hbar^2)^{-1}\sum_{\gamma'}|\gamma'\rangle\,\langle\gamma'|\ \ W(\gamma'|\bar\gamma)\,. \quad (3.11)$$

Obviously, we have only differentiated one pair of integrals. We now repeat this procedure for all other pairs of integrals, involving $d\vartheta_{n-i}$ and $d\bar\vartheta_{n-j}$; this requires that we make i transpositions to the right in the reverse ordered set and j transpositions to the left in the normally ordered set. These integrals are now adjacent and are evaluated as before, giving for the derivative the same result as in Eq. (3.11). Notice also that the time ordering for the remaining integrals on both sides has not been affected, since the pair, after evaluation, is removed from the two sets. Altogether, there are n^2 contributions (3.11). In the first two factors on the right-hand side of Eq. (3.6) we end up with $(1/m!)^2\,(\lambda^2/\hbar^2)^m$, where $m=n-1$. But, obviously, to commence the pair differentiation we need to have at least one pair, i.e., n has to start at 1; consequently, m starts again at zero. Clearly, the differentiation recreates the original series of time-ordered integrals, hence, also the time-ordered exponentials which, moreover, yield an algebraic sum involving the projectors of $\mathcal{H}^0$. Thus, the entire expression for ρ as given in Eq. (3.3) is recovered, except for the original projector $|\bar\gamma\rangle\,\langle\bar\gamma|$, which has been "used up." For the complete differentiation of the two series of integrals we, thus, obtained

$$\begin{aligned}\left.\frac{\partial\rho_d}{\partial\tau}\right|_{\text{part(a)}}&=\sum_{\gamma'}(|\gamma'\rangle\langle\gamma'|)\,(|\bar\gamma\rangle\langle\bar\gamma|)^{-1}\,W(\gamma'|\bar\gamma)\,\mathcal{P}\rho(\tau)\\&=\sum_{\bar\gamma}\sum_{\gamma'}|\gamma'\rangle\langle\gamma'|W(\gamma'|\bar\gamma)\,\langle\bar\gamma|\rho(\tau)|\bar\gamma\rangle. \quad (3.12)\end{aligned}$$

We added the subscript d (for diagonal part) on the left-hand side, noting from Eq. (3.11) that this differential is diagonal at all times; since the right-hand side must then also be diagonal, we added the $\mathcal{P}$ in front of ρ and evaluated the result in the second line, based on the explicit form Eq. (2.9).

Next, we must differentiate the terms U_d and $U_d^\dagger$ in Eq. (3.3). It is noticed that they commute with the projector expression inside, so $U_d^\dagger$ near the far end of Eq. (3.3) can be brought to the front next to U_d. The reduced operators in their full form are *not* unitary anymore, as shown by the detailed expressions of Eqs. (2.29) and (2.30). Using the latter, the imaginary exponentials yield unity, so that remains

$$U_d(\tau,\tau')\,U_d^\dagger(\tau,\tau')=\exp\left[-(\lambda^2/\hbar)(\tau-\tau')\,C\left(\mathcal{H}^0_{\xi_i}\right)\right]. \quad (3.13)$$

For the derivative we now have $(\partial\rho/\partial\tau)_{\text{part(b)}}=-(\lambda^2/\hbar)\,C(\mathcal{H}^0_{\xi_i})\rho(\tau)$. This part is also diagonal, as expected. Writing $C(\mathcal{H}^0_{\xi_i})$ in spectral decomposition and employing also (2.28), we find

$$\begin{aligned}\left.\frac{\partial\rho_d}{\partial\tau}\right|_{\text{part(b)}}&=-(\lambda^2/\hbar)\left[C\left(\mathcal{H}^0_{\xi_i}\right)\rho(\tau)\right]\\&=-(\lambda^2/\hbar)\sum_{\gamma'}|\gamma'\rangle\,\langle\gamma'|C(\gamma'_{\xi_i})\,\langle\gamma'|\rho(\tau)|\gamma'\rangle\\&=-(2\pi\lambda^2/\hbar)\sum_{\gamma'}\sum_{\bar\gamma}|\gamma'\rangle\,\langle\gamma'||\langle\gamma'|\mathcal{V}|\bar\gamma\rangle|^2\\&\quad\times\delta(\varepsilon_{\bar\gamma}-\varepsilon_{\gamma'})\langle\gamma'|\rho(\tau)|\gamma'\rangle\\&=-\sum_{\gamma'}\sum_{\bar\gamma}|\gamma'\rangle\,\langle\gamma'|W(\bar\gamma|\gamma')\,\langle\gamma'|\rho(\tau)|\gamma'\rangle. \quad (3.14)\end{aligned}$$

Combining now both results, Eqs. (3.12) and (3.14), the final result reads

$$\begin{aligned}\frac{\partial\rho_d(\tau)}{\partial\tau}&=\sum_{\gamma'}\sum_{\bar\gamma}|\gamma'\rangle\,\langle\gamma'|\{W(\gamma'|\bar\gamma)\,\langle\bar\gamma|\rho(\tau)|\bar\gamma\rangle\\&\quad-W(\bar\gamma|\gamma')\,\langle\gamma'|\rho(\tau)|\gamma'\rangle\}\ \equiv-\Lambda_d\rho_d(\tau). \quad (3.15)\end{aligned}$$

This is the ME form in *Liouville space*; we shall comment on it in part B of this section. Taking the matrix element $\langle\gamma|\rho_d|\gamma\rangle$ on both sides, we obtain the more familiar form in the *function space*,

$$\begin{aligned}\frac{\partial p(\gamma,\tau)}{\partial\tau}&=\sum_{\bar\gamma}\{W(\gamma|\bar\gamma)\,p(\bar\gamma,\tau)-W(\bar\gamma|\gamma)\,p(\gamma,\tau)\}\\&\equiv-M_\gamma\,[p(\gamma,\tau)]. \quad (3.16)\end{aligned}$$

We still notice the property of *microscopic reversibility*, $W(\gamma|\bar\gamma)=W(\bar\gamma|\gamma)$. However, we wrote Eq. (3.16) with distinct W's, of equal value, so that it reads like a conservation principle: going from $\bar\gamma$ to γ we get a gain-term, while going from γ to $\bar\gamma$ we have a loss term.

Next, we write

$$p(\gamma,\tau)=\sum_{\gamma'}P(\gamma,\tau|\gamma',\tau')p(\gamma',\tau'), \quad (3.17)$$

where P is the conditional Markov probability. Substituting this into Eq. (3.16), we obtain

$$\begin{aligned}&\sum_{\gamma'}p(\gamma',\tau')\,\frac{\partial P(\gamma,\tau|\gamma',\tau')}{\partial\tau}\\&=-\sum_{\gamma'}p(\gamma',\tau')\,M_\gamma[\,P(\gamma,\tau|\gamma',\tau')]\,,\quad\tau\geqslant\tau'. \quad (3.18)\end{aligned}$$

Since this holds for any $p(\gamma',\tau')$, we also have, noting the initial condition for P,

$$\begin{aligned}&\frac{\partial P(\gamma,\tau|\gamma',\tau')}{\partial\tau}+M_\gamma\,[P(\gamma,\tau|\gamma',\tau')]\\&\quad=\delta(\tau-\tau')\,\delta(\gamma-\gamma')/\chi(\gamma), \quad (3.19)\end{aligned}$$

where $\chi(\gamma)$ is the density of states. This is the usual Markovian master equation (MME). Or, writing M_γ in full,

$$\begin{aligned}\frac{\partial P(\gamma,\tau|\gamma',\tau')}{\partial\tau}&=\sum_{\bar\gamma}\{W(\gamma|\bar\gamma)\,P(\bar\gamma,\tau|\gamma',\tau')\\&\quad-W(\bar\gamma|\gamma)\,P(\gamma,\tau|\gamma',\tau')\},\quad\tau>\tau'. \quad (3.20)\end{aligned}$$

B. The nonstationary process

The far right-hand side of Eq. (3.15) defines the *master superoperator in Liouville space*; it was introduced by the author in 1976 during a series of seminars at the University of Utrecht. Its advantage is that it allows a direct solution for the density operator, which mirrors the solution of the von Neumann equation, $\rho(\tau) = \{\exp[-i\mathcal{L}(\tau-\tau')]\}\,\rho(\tau')$, cf. (1.1), for the weak-coupling, long-time limit:

$$\rho_d(\tau) = (\exp[-\Lambda_d(\tau-\tau')])\rho_d(\tau'). \tag{3.21}$$

We can now extend the τ and τ' to the entire interval (t_i,t_{i+1}). The full solution over the interval (t',t) then is

$$\rho_d(t) = \left(\exp\left[-\Lambda_d^{t_{n-1}}(t-t_{n-1})\right]\right)\left(\exp\left[-\Lambda_d^{t_{n-2}}(t_{n-1}-t_{n-2})\right]\right) \times \ldots \left(\exp\left[-\Lambda_d^{t'}(t_1-t')\right]\right)\rho(t'). \tag{3.22}$$

The superscripts on the master operator Λ_d indicate the time points (with corresponding set of eigenstates $|\gamma_{\xi_i}\rangle$) that are applicable to that interval. Differentiating to t we find $\partial\rho_d/\partial t = -\Lambda_d^{t_{n-1}}\rho_d(t)$. Since the eigenstates $|\gamma_{\xi_{n-1}}\rangle \to |\gamma_\xi\rangle$ cover the entire last interval (t_{n-1},t), we will use the superscript t instead. The final explicit form of the ME in Liouville space then reads,

$$\frac{\partial\rho_d(t)}{\partial t} = \sum_{\gamma_\xi}\sum_{\bar\gamma} |\gamma_\xi\rangle\,\langle\gamma_\xi|\{W_t(\gamma_\xi|\bar\gamma)\,\langle\bar\gamma|\rho(t)|\bar\gamma\rangle - W_t(\bar\gamma|\gamma_\xi)\,\langle\gamma_\xi|\rho(t)|\gamma_\xi\rangle\}. \tag{3.23}$$

Here $|\gamma_\xi\rangle$ is an eigenstate at t while $|\bar\gamma\rangle$ can be any state of the entire state space. The equivalent ME in function space has the form, cf. Ref. [9] and Gaspard [27],

$$\frac{\partial p(\gamma_\xi,t)}{\partial t} = \sum_{\bar\gamma}\{W_t(\gamma_\xi|\bar\gamma)\,\langle\bar\gamma|\rho(t)|\bar\gamma\rangle - W_t(\bar\gamma|\gamma_\xi)\,\langle\gamma_\xi|\rho(t)|\gamma_\xi\rangle\}. \tag{3.24}$$

Lastly, it may not have escaped the reader that we used an initial random-phase assumption on each subinterval (t_i,t_{i+1}). Basically, this means that we have "reset" the conditions when we switch to a new set of eigenstates. However, another interpretation is that the density operator needs not to be diagonal at any time [6,8]; we just seek the response of the diagonal part. [There also is a nondiagonal part, featuring the reduced Liouville operator $\mathcal{L}^0$ [8], not considered here.]

a. Discussion. While the master equation itself is not necessary for the applications to the fluctuation theorems dealt with in Sec. IV A and can be dispensed with in the first part of Sec. IV B, we included it here to show that the perturbational form of the time-reversal property is associated with a Markovian ME that has a positive entropy production, as expected for a driven non-equilibrium process; the proof is provided in Sec. IV B2. In any quantum mechanical theory for irreversible processes, we need either internal dissipative processes or an explicit randomizing coupling with the reservoir(s); previous papers have mostly failed to satisfy these criteria.

The ME for nonstationary processes differs only from that for stationary processes in that the transition probabilities must be subscripted with the time t for which $\partial\rho/\partial t$ is evaluated. If one does not want to connect with the physical processes via the von Neumann equation as done here, that result can also be obtained from the Chapman-Kolmogoroff equation of the stochastic literature, which serves as a "consistency criterion" for Markov processes, cf. van Kampen [28].

Since we approached the topic by considering a set of quasistationary subintervals, our present proof of the Pauli-Van Hove master equation deserves particular attention. Basically, we followed van Hove's procedure in extracting the relevant terms from the double series expansion for $U\rho_0U^\dagger$, rather than Zwanzig's method in considering the single series expansion for the perturbed Liouville operator, followed also in our book [8], Sections 16.11 and 12. However, in contrast to the very lengthy expressions in Refs. [1] and [5], we obtained a closed-form result for $U(\tau,\tau')$ and the subsequent differentiation of $\rho(\tau)$ was straightforward and without "gimmicks." In particular, there was no need for "Van Hove's *functional rule*." Rather than inserting closure sums at particular places in the computations, Van Hove relied on a postulated linear relationship pertaining to a product of $\mathcal{V}$ operators, separated by any operator expression Ω:

$$\langle\gamma|\mathcal{V}\,\Omega\mathcal{V}|\gamma'\rangle = \delta(\gamma-\gamma')\int d\gamma''\langle\gamma|\,\Omega|\gamma'\rangle X(\gamma'',\gamma) + \int d\gamma''\langle\gamma''|\,\Omega|\gamma''\rangle Y(\gamma'';\gamma,\gamma'), \tag{3.25}$$

where X and Y are at first unknown kernels in the indicated variables. [As for X, one easily shows that $X(\gamma,\bar\gamma) = |\langle\gamma|\mathcal{V}|\bar\gamma\rangle|^2$.] In subsequent computations, the terms in Y do not yield terms of order $\lambda^2\tau$, in contrast to the terms with X. Thus, Van Hove ascribes the obtained irreversible behavior to the *diagonal singularity* of the functional rule (3.25). The present calculation shows that the diagonal singularity plays no preponderant role in the emergence of irreversible results; rather, we believe that the partitioning of the Hamiltonian as $\mathcal{H}^0+\lambda\mathcal{V}$, with small but finite λ, is *of essence* in the quest for irreversibility.

IV. THERMODYNAMIC APPLICATIONS

A. Crooks-Tasaki and Jarzynski revisited

1. Microscopic considerations

As noted by a number of authors, among others [16,18], the deduction of the Crooks-Tasaki FT, as well as the Jarzynski work–energy theorem from standard quantum mechanical considerations involving the time evolution, such as set forth here, is basically straightforward. However, there is a catch: What is work? We shall discuss this in detail below.

The pivotal point is the time-reversal symmetry. Thus, starting from Eq. (2.6), applied for the time interval (t_0,t_1), we multiply the left-hand side with $p_{\rm can}(\gamma_0,t_0)$ and the right-hand side with $p_{\rm can}(\gamma_1,t_1)$ to obtain

$$P(\gamma_1,t_1|\gamma_0,t_0)\,p_{\rm can}(\gamma_0,t_0) = \tilde P(\gamma_0,t_0|\gamma_1,t_1)\,p_{\rm can}(\gamma_1,t_1)[p_{\rm can}(\gamma_0,t_0)/p_{\rm can}(\gamma_1,t_1)], \tag{4.1}$$

whereby it is understood that the eigenstates $|\gamma_1\rangle$ and $|\gamma_0\rangle$ belong to different Hamiltonians, as parametrized by $\mathcal{H}[\xi(t_1)]$ and $\mathcal{H}[\xi(t_0)]$, respectively. With $p_{\rm can}$ given by the Gibbs

distribution $(1/Z)e^{-\beta\varepsilon_\gamma}$, we have, by Bayes's rule [29],

$$W_2(\gamma_1,t_1;\gamma_0,t_0)\,e^{-\beta(\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0})}=\tilde{W}_2(\gamma_0,t_0;\gamma_1,t_1)\,e^{-\beta\,(F_1^0-F_0^0)}; \tag{4.2}$$

here W_2 is the two-point probability and F^0 denotes the equilibrium free energy of the system. Or also, with $\Delta F^0 = F_1^0 - F_0^0$,

$$W_2(\gamma_1,t_1;\gamma_0,t_0)/\tilde{W}_2(\gamma_0,t_0;\gamma_1,t_1)=\ e^{\beta\,[(\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0})-\,\Delta F^0]}. \tag{4.3}$$

Let us assume a one-to-one correspondence between energies and eigenstates; any extraneous degeneracy, such as due to spin, will be the same for both W_2's. Then, alternately for Eq. (4.3),

$$p(\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0})/\tilde{p}(\varepsilon_{\gamma 0}-\varepsilon_{\gamma 1})=e^{\beta\,[(\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0})-\,\Delta F^0]}. \tag{4.4}$$

Note that this implies that the measurement of the difference $(\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0})$ is a two-time exercise.[4] Clearly, we obtain the Crooks-Tasaki FT if we can establish that this energy difference is an eigenvalue w of the quantum mechanical work W performed on the system. We, first, give some heuristic—and, at face value, incorrect—arguments as found in the literature [16].

We shall restrict ourselves to the scenario that the protocol $\xi(t)$ generates a set of c numbers. For that case, the imposed changes can be modeled by an external field Hamiltonian, as is customary in LRT, $\mathcal{H}_{\rm ext}=-\mathcal{A}\,F(t)$, where $\mathcal{A}$ is a system operator and F a c number; for a system of charged particles, $\mathcal{A}$ is the shift of all charges, $\Sigma_i\,q_i\,(\vec{r}_i-\vec{r}_{i,\,\rm eq})$, and $F(t)$ an applied electric field. We can make a Fourier analysis of $F(t)$ over the applicable interval, with Fourier amplitudes F_ω. This will leave a linewidth $\delta\varepsilon=\hbar/\Delta\tau$ with $\Delta\tau\sim(t_{i+1}-t_i)$, the width of a subinterval over which the system Hamiltonian can be approximated as being constant. As shown elsewhere [30,31], the truncated phasors contribute to zero-point energy fluctuations. To the extent that these can be neglected, there will be conservation of system energy. It is now argued that the electric field component F_ω will couple to the system, causing corresponding harmonic displacements of the charges, which in turn emit virtual photons of the same frequency. The reabsorption by the system gives excitations of its quantum states with the selection rule $\hbar\omega=(\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0})$. The probability for this to occur is proportional to $|F_\omega|^2$. Now let, at first, the system plus external agent be fully isolated. Though the external agent has been described here as having a Hamiltonian *operator*, we may assume that the agent has so many degrees of freedom that the correspondence rule applies; then, $\Delta\|\mathcal{A}\,F_\omega\|=\mathfrak{W}$, where $\mathfrak{W}$ is the *classical work* performed. For the described sequence of events, in the correspondence limit we have $\mathfrak{W}\sim w=\hbar\omega$.

When, next, we reconnect with the heat bath, the work must be broken up into two parts. Prior to the heat transfer to the reservoir we have $w_{\rm system}=\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0}+\Delta Q$. After transfer, the heat gained by the reservoir is ΔQ. For simplicity, let the reservoir consist of an isothermal ideal gas; then $\Delta\mathcal{E}_{\rm reserv}=0$, thus $w_{\rm reserv}=-\Delta Q$. Adding, we find again

$$w=w_{\rm system}+w_{\rm reserv}=\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0}\,. \tag{4.5}$$

We note hereby that Tasaki's FT concerns closed quantum systems, whereas Crooks FT deals with thermostatted classical systems; this confirms that the bath is often not essential. In practice, however, truly closed systems are academic in that they are not amenable to measurements. From Eqs. (4.4) and (4.5) we now obtain

$$p(w)/\tilde{p}(-w)=e^{\beta[w-\Delta F^0]}, \tag{4.6}$$

which is the Crooks-Tasaki FT. Despite its simplicity, the "proof" and the model presented here have some heuristic as well as extremely unrealistic elements in it.[5]

To improve on the argument, one could consider the quantized field and its forward autocorrelation function $\mathrm{Tr}\{\rho_{\rm can}[\mathcal{F}(\tau)\mathcal{F}(0)]\}$ in a canonical ensemble; its Fourier transform (times a factor of 2) yields the spectral density $S(\omega)$. Evaluating the trace for the Fourier transform in the representation $\{|\gamma_\xi\rangle\}$, one obtains

$$S(\omega)=4\pi\hbar\sum_{\gamma_1\,\gamma_0}p_{\rm can}(\gamma_0)\,|\langle\gamma_0|\mathcal{F}|\gamma_1\rangle|^2\,\delta\big[\hbar\omega-\big(\varepsilon_{\gamma_1}-\varepsilon_{\gamma_0}\big)\big]. \tag{4.7}$$

The δ function of the spectral analysis confirms Eq. (4.5); see also Sec. V B.

a. Definition of work. The concept of work must be clarified before any results can be meaningful. In the many-body phase space, work is directly related to the difference of the classical Hamiltonians. When this is carried over to quantum statistics, we must resort to the Heisenberg picture noting that, by the Heisenberg equation of motion, the total derivative of the Heisenberg Hamiltonian is equal to the local derivative. Hence, $d\mathcal{H}_H(t)=[\partial\mathcal{H}_H(t)/\partial t]dt$ or also

$$\mathcal{H}_H(t_1)-\mathcal{H}_H(t_0)=\int_{t_0}^{t_1}(\partial\mathcal{H}_H/\partial\xi)\,d\xi. \tag{4.8}$$

With the notion that $\partial\mathcal{H}_H/\partial\xi$ relates to the local work, it is now tempting to equate the left-hand side of Eq. (4.8) with the operator for work $\mathcal{W}$; cf. the extensive article in this journal by Allahverdyan and Nieuwenhuizen [32]. They stress, in particular, that in a quantum statistical mixture of states, there will be sizeable fluctuations given by the Bochkov-Kusov (BK) equality [33]. For this reason, it seems to us that work should not be defined by Eq. (4.8) except in the classical limit when $\mathcal{H}_H\sim\mathcal{H}$.

Here we posit that *work is not an observable in the Hilbert space* $\mathcal{S}$, as was also noted by Talkner *et al.* [17]. However, while $\mathcal{W}$ is not a bounded operator in $\mathcal{S}$, the unitary operator $e^{iu\,\mathcal{W}}$ exists and, by Stone's theorem [34,35], has a spectral decomposition

$$e^{iu\mathcal{W}}=\int e^{iuw}d\,\hat{p}(w), \tag{4.9}$$

[4]The problem of collapse of the state and decoherence will not be touched on, since the time-measurement interval is very large compared to any quantum transition times.

[5]It may have been grossly overlooked that, for normal coupling frequencies, this "Gedanken experiment" pertains to femto- or atto-Joules of work delivered, unless one supposes the process to be repeated a great many times.

where w are the *measurable eigenvalues* and $d\hat{p}(w)$ the incremental projectors. Upon averaging, this implies a characteristic function,[6]

$$\langle e^{iu\mathcal{W}}\rangle = \int e^{iuw} dP(w), \tag{4.10}$$

with dP being the incremental cumulative distribution function (cdf). Generally the right-hand side is a Lebesgue-Stieltjes integral, cf. [8] Section 1.7. Formally, we can write $dP(w) = p(w)\,dw$, but the probability density (pdf) may be singular. It is now expedient to define work by the relationship

$$\exp[iu\mathcal{W}] \equiv \exp\{iu\,[\mathcal{H}_H(t_1) - \mathcal{H}_H(t_0)]\}. \tag{4.11}$$

As to the characteristic function (4.10), it follows that the averaged exponentiated work is rigorously equal to the forward time-ordered correlation function,

$$\begin{aligned}\langle \exp[iu\mathcal{W}]\rangle &= \langle \exp\{iu\,[\mathcal{H}_H(t_1) - \mathcal{H}_H(t_0)]\}\rangle \\ &= \langle \mathcal{T}\exp[iu\mathcal{H}_H(t_1)]\exp[-iu\mathcal{H}_H(t_0)]\rangle.\end{aligned} \tag{4.12}$$

To obtain the connection with the Schrödinger forms, we will write the operator average on the left-hand side in the form of a trace, i.e., $\langle \exp[iu\mathcal{W}]\rangle = \mathrm{Tr}\{\rho \exp[iu\mathcal{W}]\}$ [where, as in footnote 6, $\langle w|\rho|w\rangle = p(w)$]. We then have, with U being the original evolution operator of Eq. (2.2),

$$\begin{aligned}&\mathrm{Tr}\{\rho e^{iu\mathcal{H}_H(t_1)} e^{-iu\mathcal{H}_H(t_0)}\} \\ &\quad= \mathrm{Tr}\{\rho U^\dagger(t_1,t_0) e^{iu\mathcal{H}(t_1)} U(t_1,t_0) e^{-iu\mathcal{H}(t_0)}\} \\ &\quad\overset{\text{c.i.}}{=} \mathrm{Tr}\{e^{iu\mathcal{H}(t_1)} U(t_1,t_0) e^{-iu\mathcal{H}(t_0)} \rho U^\dagger(t_1,t_0)\} \\ &\quad\overset{\text{com}}{=} \mathrm{Tr}\{U(t_1,t_0) e^{iu\mathcal{H}(t_1)} e^{-iu\mathcal{H}(t_0)} \rho U^\dagger(t_1,t_0)\} \\ &\quad\overset{\text{c.i.}}{=} \mathrm{Tr}\{e^{iu\mathcal{H}(t_1)} e^{-iu\mathcal{H}(t_0)} \rho U^\dagger(t_1,t_0) U(t_1,t_0)\} \\ &\quad= \mathrm{Tr}\{e^{iu[\mathcal{H}(t_1)-\mathcal{H}(t_0)]}\rho\} = \mathrm{Tr}\big\{e^{iu[\mathcal{H}^0(t_1)-\mathcal{H}^0(t_0)]}\rho\big\}.\end{aligned} \tag{4.13}$$

In the various transitions we employed the property of cyclic invariance (c.i.) of the trace and that two particular adjacent operators commute (com); also we observed that the perturbations $\lambda\mathcal{V}$ are independent of time as presumed throughout this article, thus yielding the last right-hand side. Evaluating the trace in the representation $\{|\gamma_\xi\rangle\}$ now gives

$$\langle e^{iu\mathcal{W}}\rangle = \langle e^{iu(\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0})}\rangle. \tag{4.14}$$

Taking the inverse Fourier transform, $(1/2\pi)\int_{-\infty}^{\infty} du\, e^{-iuw}\cdots$, on both sides, we obtain

$$p(w) = \delta\,[w - (\varepsilon_{\gamma 1} - \varepsilon_{\gamma 0})]. \tag{4.15}$$

Finally, let us go back to Eq. (4.4). Substituting (4.15) into Eq. (4.4), we established the Crooks-Tasaki FT in a fully quantum statistical manner,

$$p(w)/\tilde{p}(-w) = e^{\beta[w-\Delta F^0]}. \tag{4.16}$$

Next, we notice that the characteristic function with $u = z + iv$ is analytic for $0 \leqslant v \leqslant \beta$; for a proof based on Schwartz's inequality, cf. Ref. [22]. Hence, letting $u \to i\beta$, Eq. (4.14) yields

$$\langle e^{-\beta\mathcal{W}}\rangle = \langle e^{-\beta(\varepsilon_{\gamma 1}-\varepsilon_{\gamma 0})}\rangle. \tag{4.17}$$

Now, the average at the right-hand side of Eq. (4.17) can be obtained by a double integration $[\iint d\gamma_1 d\gamma_0]$ over W_2. Thus, going back to Eq. (4.2) and integrating both sides, noticing that the distribution $\tilde{W}_2$ is normalized, Eqs. (4.2) and (4.17) yield *the quantum Jarzynski work-energy theorem*,

$$\langle e^{-\beta\mathcal{W}}\rangle = e^{-\beta\,\Delta F^0}. \tag{4.18}$$

b. Discussion. With the work being given as a functional of the Heisenberg operators, we, first, used the standard time-dependent evolution operator of Eq. (2.2) to connect with the Schrödinger Hamiltonian. We were then able to connect with the eigenstates of the system, obtaining Eqs. (4.12)–(4.17). But, as in LRT, which manipulates the properties of Heisenberg operators, the theory must gain *physical content* by an *a posteriori* randomization (called by Kubo "stochasticization"), cf. Ref. [36], p. 196. Thus, the results must be applied in conjunction with the fully stochastic results of Eqs. (4.1)–(4.4), which form the basis of our treatment. Also, we notice that the connection between the work and the eigenstates of the system emerges explicitly.

In this respect, we mention another intriguing brief communication of Talkner *et al.* [18]. They consider the characteristic function of $\mathcal{W}$ for $u = z + i\beta$, which is in the domain of analyticity as indicated above. Let $Z(t_0)$ be the partition function at the beginning of the protocol and $Z(t_1)$ be the partition function at the end of the protocol and let $\Phi(u)$ be the characteristic function of the work operator $\mathcal{W}$ as defined above. Then, by manipulation of the trace,

$$\begin{aligned}&Z(t_0)\Phi(u) \\ &\quad= \mathrm{Tr}\{U^\dagger(t_1,t_0)\, e^{i(z+i\beta)\mathcal{H}(t_1)} U(t_1,t_0)\, e^{-i(z+i\beta)\mathcal{H}(t_0)} e^{-\beta\mathcal{H}(t_0)}\} \\ &\quad= \mathrm{Tr}\{e^{iz\mathcal{H}(t_1)} e^{-\beta\mathcal{H}(t_1)} U(t_1,t_0)\, e^{-iz\mathcal{H}(t_0)} U^\dagger(t_1,t_0)\} \\ &\quad= \mathrm{Tr}\{e^{iz\mathcal{H}(t_1)} e^{-\beta\mathcal{H}(t_1)} U^\dagger(t_0,t_1)\, e^{-iz\mathcal{H}(t_0)} U(t_0,t_1)\} \\ &\quad= \mathrm{Tr}\{U^\dagger(t_0,t_1)\, e^{-iz\mathcal{H}(t_0)} U(t_0,t_1)\, e^{iz\mathcal{H}(t_1)} e^{-\beta\mathcal{H}(t_1)}\} \\ &\quad= \mathrm{Tr}\{e^{-iz[\mathcal{H}_H(t_0)-\mathcal{H}(t_1)]}\, e^{-\beta\mathcal{H}(t_1)}\} = Z(t_1)\,\tilde{\Phi}(-u + i\beta),\end{aligned} \tag{4.19}$$

or

$$\Phi(u) = e^{-\beta\Delta F^0}\tilde{\Phi}[-(u - i\beta)]. \tag{4.20}$$

The Fourier inversion of the left-hand side gives $p(w)$. For the right-hand side we obtain (proof is ours),

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\infty}^{\infty} du\, e^{-iuw}\int_{-\infty}^{\infty} dw' e^{-iw'(u-i\beta)}\tilde{p}(w') \\ &\quad= \frac{1}{2\pi}\int_{-\infty}^{\infty} dw'\,\tilde{p}(w')\int_{-\infty}^{\infty} du\, e^{-iu(w+w')} e^{-\beta w'} \\ &\quad= \int_{-\infty}^{\infty} dw'\,\tilde{p}(w')\,\delta(w+w')\, e^{-\beta w'} = e^{\beta w}\tilde{p}(-w).\end{aligned} \tag{4.21}$$

Equations (4.20) and (4.21) directly yield Tasaki's quantum FT (4.16). In Sec. V we will show that (4.20) is a special form of Mazo's lemma of LRT [37], extended to nonequilibrium states. While this is a beautiful derivation, it is *incomplete* in two

[6] In Dirac notation, $\langle \exp[iu\mathcal{W}]\rangle = \mathrm{Tr}\{\rho\int \exp[iuw]\,d\,(|w\rangle\langle w|)\} = \int \exp[iuw]\,d\,(\langle w|\rho|w\rangle) = \int \exp[iuw]\,dP(w)$. However, a correct application of Stone's theorem requires the mathematical notation for the scalar product and the projectors.

respects: it offers no connection between the work performed and the excited states of the system and the randomization is not carried out, without which Kubo correlation functions do not converge. As indicated by van Kampen [38] and by the present author [5,8,39], linear response theory is a "hollow shell" *until* the physics for dissipative behavior is put in place.

Our microscopic developments in this section yielded the quantum results with full reference to randomizing interactions, either stemming from internal dissipative causes like electron-phonon collisions or due to interactions with external reservoir(s). However, the condition that the protocol $\xi(t)$ does not involve a variation of system operators is stringent.

2. *Mesoscopic considerations*

When intensive observables other than fields are varied in the processes that drive the system far from equilibrium such as shearing in an isothermal fluid or volume expansion in a thermostatted system (see below), a mesoscopic treatment becomes necessary since the operators corresponding to these observables, $\{a_i\}$, $i = 1, \ldots, s$, do not commute with each other or the Hamiltonian. The operators must then be given a certain leeway δa_i. Coarse-graining will erase cross-matrix elements for different cells; the process is described in some detail in Ref. [9]. A mesoscopic ME can now be derived, cf. Ref. [8], Sec. 18.3; however, it will not be needed. Therefore, suffice it to consider the connection

$$P(a_1,t_1|a_0,t_0) = P(\gamma_1,t_1|\gamma_0,t_0)\,\chi(a), \tag{4.22}$$

where $\chi(a)$ is the density of states; the symbol a represents all variables of the process. We noted hereby that a given initial state $|\gamma_0\rangle$ engenders a given set of variables a_{i0}, but many states $|\gamma_1\rangle$ correspond with the leeway $(a_i, a_i + \delta a_i)$. The mesoscopic conditional probability densities for forward and backward driving processes are, therefore, related by combining (2.6) and (4.22),

$$P(a_1,t_1|a_0,t_0) = \tilde{P}(a_0,t_0|a_1,t_1)\,[\chi(a_1)/\chi(a_0)]. \tag{4.23}$$

We now need the initial distributions for both sides. They are not simply the canonical distributions of the previous section, since the a's are subject to additional fluctuations. In Ref. [9] we proved the canonical Boltzmann-Einstein probability density function (pdf),

$$W(a_0,t_0) = \hat{c}_0^{-1} e^{-\beta[F(a_0)-F_0^0]}, \quad W(a_1,t_1) = \hat{c}_1^{-1} e^{-\beta[F(a_1)-F_1^0]}, \tag{4.24}$$

where the nonsuperscripted F's are nonequilibrium Helmholtz free-energy functions, while the $\hat{c}$'s are normalization constants that vanish logarithmically. Multiplying both sides of Eq. (4.23) with the initial pdf's, some algebra yields

$$W_2(a_1,t_1;a_0,t_0)\,[\hat{c}_0\chi(a_0)/\hat{c}_1\chi(a_1)]\,e^{\beta[F(a_0)-F(a_1)]} = \tilde{W}_2(a_0,t_0;a_1,t_1)\,e^{-\beta\,\Delta F^0}; \tag{4.25}$$

here $\hat{c}\chi(a) = \Delta\Gamma(a)$ is the accessible number of states in the coarse-grained intervals. Using further the well-known result $\Delta\Gamma(a) = \exp(S/k_B)$, we established

$$W_2(a_1,t_1;a_0,t_0)\,e^{-\beta\,[\mathcal{E}(a_1)-\mathcal{E}(a_0)]} = \tilde{W}_2(a_0,t_0;a_1,t_1)\,e^{-\beta\,\Delta F^0}. \tag{4.26}$$

For the coarse-grained energy difference, the connection with the (classical) work is immediate,

$$\mathcal{E}(a_1) - \mathcal{E}(a_0) = \mathfrak{W}(a_0 \to a_1). \tag{4.27}$$

Also, we have $W_2(a_1,t_1;a_0,t_0) = p\,[\mathfrak{W}(a_0 \to a_1)]$ and similarly for $\tilde{W}_2$. This, then, gives the mesoscopic form of the Crooks-Tasaki FT,

$$p(\mathfrak{W})/\tilde{p}(-\mathfrak{W}) = \exp[\beta(\mathfrak{W} - \Delta F^0)]. \tag{4.28}$$

Integrating both sides of Eq. (4.26), $\iint da_1 da_0 \cdots$, the mesoscopic Jarzynski equality appears,

$$\langle \exp(-\beta\mathfrak{W}(a)) \rangle = \exp(-\beta\,\Delta F^0). \tag{4.29}$$

Further results along these lines are found in Ref. [9], Sec. III.

A final note concerns isothermal compression or expansion of a gas in a cylinder of cross section A. The average pressure is 2/3 of the energy density for all ideal gases, BE, FD, and Boltzmann. This suggests that a microscopic pressure operator can be defined as a constant times the Hamiltonian volume derivative $A^{-1}(\partial\mathcal{H}/\partial\xi)$, where ξ monitors the position x of the piston. Hence, it might be tempting to consider a work-related operator $\int(\partial\mathcal{H}_H/\partial\xi)\dot{\xi}\,dt$. However, ξ itself is now an operator, noncommuting with the Hamiltonian and subject to substantial fluctuations $\Delta\xi$; a mesoscopic approach is, therefore, imperative, contrary to the treatment in Ref. [16].

B. Entropy fluctuation theorems

The entropy FT's have been obtained in a number of articles, cf. Refs. [20–25] and references therein; however, except for the small contribution in Ref. [22], the derivations are primarily based on phase-space or classical trajectory considerations. Yet a number of papers deal with stochastic dynamics [22,23,25,40] which, although not linked to the von Neumann equation as point of departure (as in this paper) can easily be adapted to a quantum mechanical viewpoint. Even so, we gathered from a study of these papers that a fully microscopic description, while allowing for a Gibbs entropy, is not compatible with the concepts of entropy current I_η and entropy production η. In fact, the entropy current which serves for the exchange with a thermostat, will be zero if based on microscopically reversible transition rates $W(\gamma'|\gamma) = W(\gamma|\gamma')$; see the form for I_η [Eq. (4.44)] below. Thus, we must assume that a thermostatted process is described by a mesoscopic ME, for which the "states" will be labelled by σ rather than by a, as in our previous subsection, in order to be in line with the basic papers on the subject [23,25]. We shall omit these complications at first and give a brief microscopic discussion for isolated systems.

1. *The entropy FT for isolated systems*

From the time-reversal symmetry (2.6), we find again the microscopic two-point probability with Bayes' rule,

$$W_2(\gamma_1,t_1;\gamma_0,t_0) = \tilde{W}_2(\gamma_0,t_0;\gamma_1,t_1)\,[p(\gamma_0,t_0)/p(\gamma_1,t_1)], \tag{4.30}$$

or also

$$W_2(\gamma_1,t_1;\gamma_0,t_0) = \tilde{W}_2(\gamma_0,t_0;\gamma_1,t_1)\, e^{-[\ln p(\gamma_1,t_1) - \ln p(\gamma_0,t_0)]}. \tag{4.31}$$

Now we remember that the nonequilibrium Gibbs entropy is given by

$$\begin{aligned} S_G(t) &= -k_B \mathrm{Tr}\rho(t) \ln \rho(t) \\ &= -k_B \sum_{\gamma} p(\gamma,t) \ln p(\gamma,t) \\ &= -k_B \langle \ln p(\gamma,t) \rangle_{\text{nonequil}}. \end{aligned} \tag{4.32}$$

However, the Gibbs entropy does not fluctuate—as does the Boltzmann entropy—but is an average over $p(\gamma,t)$. We therefore introduce a *more* microscopic nonequilibrium entropy by $\mathcal{S} = -k_B \ln p(\gamma,t)$. Substituting this into Eq. (4.31), we find

$$W_2(\gamma_1,t_1;\gamma_0,t_0) = \tilde{W}_2(\gamma_0,t_0;\gamma_1,t_1)\, e^{\Delta\mathcal{S}/k_B}, \tag{4.33}$$

where $\Delta\mathcal{S} = \mathcal{S}(\gamma_1,t_1) - \mathcal{S}(\gamma_0,t_0)$. Obviously, the distribution function for $\Delta\mathcal{S}$ is just the two-point probability function W_2; hence, we obtain the "transient entropy FT,"

$$p(\Delta\mathcal{S})/\tilde{p}(-\Delta\mathcal{S}) = e^{\Delta\mathcal{S}/k_B}. \tag{4.34}$$

Next, let us assume that the driven system has reached a nonequilibrium steady state at time t_0 from where on the protocol will be time independent, so that $p = \tilde{p}$. We then find the "steady-state entropy FT,"

$$p(-\Delta\mathcal{S})/p(\Delta\mathcal{S}) = e^{-\Delta\mathcal{S}/k_B}. \tag{4.35}$$

This relationship provides a quantitative answer to Loschmidt's objections to Boltzmann's irreversible H theorem: decreasing entropy can be observed, but with an exponentially low relative probability. This far-reaching quantitative statement was first obtained in a limited way by Evans, Cohen, and Morris in 1993 [41]. An asymptotic form will be presented later.

2. *Thermostatted systems*

The general ME for a nonstationary process, randomized by the nondeterministic interactions $\lambda\mathcal{V}$ with the heat reservoir, reads[7]

$$\begin{aligned} \frac{\partial p(\sigma,t)}{\partial t} &= \sum_{\sigma' \neq \sigma} [w_{\sigma',\sigma}(t)p(\sigma',t) - w_{\sigma,\sigma'}(t)p(\sigma,t)] \\ &\equiv -M_\sigma[p(\sigma,t)], \end{aligned} \tag{4.36}$$

where M_σ is the function-space master operator and $w_{\sigma,\sigma'}(t)$ is the time-dependent transition rate *from* σ *to* σ'; these rates are connected to the microscopic rates of Sec. III B by

$$w_{\sigma,\sigma'}(t) = W_t(\gamma'|\gamma)\,\chi(\sigma',t), \quad w_{\sigma',\sigma}(t) = W_t(\gamma|\gamma')\,\chi(\sigma,t), \tag{4.37}$$

where χ is the density of states. Because of microscopic reversibility for the W's, we have

$$w_{\sigma,\sigma'}(t)/w_{\sigma',\sigma}(t) = \chi(\sigma',t)/\chi(\sigma,t). \tag{4.38}$$

Next, we introduce the conditional probability, $P(\sigma_1,t_1|\sigma_0,t_0)$, related to the microscopic conditional probability by $P(\gamma_1,t_1|\gamma_0,t_0)\,\chi(\sigma_1,t_1)$. Notice that a given state σ_0 is engendered by a given state γ_0. We also write $\Delta\Gamma(\sigma) = \chi(\sigma)\,\delta\sigma$, being the accessible number of quantum states. From the time-reversal symmetry property (2.6) we find

$$P(\sigma_1,t_1|\sigma_0,t_0) = \tilde{P}(\sigma_0,t_0|\sigma_1,t_1)\,[\Delta\Gamma(\sigma_1,t_1)/\Delta\Gamma\ (\sigma_0,t_0)]. \tag{4.39}$$

While the classical papers abundantly speak of the stochastic "trajectory," such notions are foreign to a quantum description; only the initial and final states have meaning. If in Eq. (4.36) the σ's referred to occupancies of sites in a lattice gas, we could use the tenets of generation-recombination noise theory or employ quantum field concepts. However, in a mesoscopic description, a simpler approach will be justified. So, we content ourselves here with a heuristic picture. The two-point probabilities for the forward and backward protocol are related by

$$\begin{aligned} W_2(\sigma_1,t_1;\sigma_0,t_0) = \tilde{W}_2(\sigma_0,t_0;\sigma_1,t_1)\big(&\exp\{-[\ln p(\sigma_1,t_1) \\ &- \ln p(\sigma_0,t_0)]\}\big)\,\big(\Delta\Gamma(\sigma_1,t_1)/\Delta\Gamma(\sigma_0,t_0)\big). \end{aligned} \tag{4.40}$$

As we noted before, the first factor is $\exp(\Delta\mathcal{S}_{\text{syst}}/k_B)$; the second factor represents the nonequilibrium entropy flow to the thermostat, $\exp(\Delta\mathcal{S}_{\text{flow}}/k_B)$. Multiplying the two factors, the arguments of the exponentials relate to the total entropy difference, giving $\exp(\Delta\mathcal{S}_{\text{total}}/k_B)$. We thus obtain the more complete transient entropy FT, analogous to Eq. (4.34),

$$p(\Delta\mathcal{S}_{\text{total}})/\tilde{p}(-\Delta\mathcal{S}_{\text{total}}) = e^{\Delta\mathcal{S}_{\text{total}}/k_B}. \tag{4.41}$$

The word *transient* stems from the classical description, in which the stochasticity is attributed to the succession of jumps $\{w_t(\sigma,\sigma')\}$ of the classical path.

a. Nonequilibrium entropy. For a thermostatted system the Gibbs entropy mentioned in subsection *1* is generalized to $S_G[\sigma,t] = -k_B \sum_\sigma p(\sigma,t) \ln p(\sigma,t)$. Employing the ME (4.36), this yields, for the time derivative in a straightforward fashion,

$$\begin{aligned} \frac{\partial S_G[p(\sigma,t)]}{\partial t} &= \frac{1}{2} k_B \sum_{\sigma,\sigma'} [p(\sigma,t)w_{\sigma,\sigma'}(t) \\ &\quad - p(\sigma',t)w_{\sigma',\sigma(t)}] \ln\left[\frac{p(\sigma,t)}{p(\sigma',t)}\right] \\ &= k_B \sum_{\sigma,\sigma'} p(\sigma,t)w_{\sigma,\sigma'}(t) \ln\left[\frac{p(\sigma,t)}{p(\sigma',t)}\right]. \end{aligned} \tag{4.42}$$

Following Schnakenberg [42], this is now split into an entropy production η and an entropy current I_η

[7] We use the partial derivative, $\partial/\partial t$, since $p(\sigma)$ stems from $p(\gamma) = \langle\gamma|\rho|\gamma\rangle$ in which ρ is a Schrödinger operator (not a Heisenberg operator), cf. (3.24).

as follows:

$$\eta\,[p(\sigma,t)] = \frac{1}{2}k_B \sum_{\sigma,\sigma'} [p(\sigma,t)w_{\sigma,\sigma'}(t) - p(\sigma',t)w_{\sigma',\sigma}(t)] \times \ln\left(\frac{p(\sigma,t)w_{\sigma,\sigma'}(t)}{p(\sigma',t)w_{\sigma',\sigma}(t)}\right) = k_B \sum_{\sigma,\sigma'} p(\sigma,t)w_{\sigma,\sigma'}(t) \ln\left(\frac{p(\sigma,t)w_{\sigma,\sigma'}(t)}{p(\sigma',t)w_{\sigma',\sigma}(t)}\right), \tag{4.43}$$

$$I_\eta\,[p(\sigma,t)] = k_B \sum_{\sigma} p(\sigma,t) \sum_{\sigma'} w_{\sigma,\sigma'}(t) \ln\left(\frac{w_{\sigma,\sigma'}(t)}{w_{\sigma',\sigma}(t)}\right). \tag{4.44}$$

Clearly, for the entropy production we have

$$\partial S_G/\partial t + I_\eta = \eta \geqslant 0, \tag{4.45}$$

by Klein's lemma; in a driven system the equals sign will rarely occur since the "states" σ,σ' mostly stem from different Hamiltonians $\mathcal{H}\,[\xi(t)]$ of the protocol. On the contrary, entropy current can have either sign.[8] We now follow Ref. [25] in associating $\partial S_G/\partial t$ with the system's entropy rate, contrary to Ref. [23] wherein I_η plays that role.

As noted above, all these "entropy derivates" are still averages over $p(\sigma,t)$. True nonequilibrium concepts are obtained by omitting $\sum_\sigma p(\sigma,t)\cdots$; the corresponding quantities will be denoted by $\partial\mathcal{S}/\partial t$, $\mathcal{I}_\eta$, and $\hat{\eta}$. Basically, this makes sense only for $\mathcal{I}_\eta \equiv \partial\mathcal{S}_{\rm flow}/\partial t$ since the others still contain $\ln[p(\sigma,t)/p(\sigma',t')]$; however, these can always be added on as "boundary values."

The integration of the system entropy, $\int dt(\partial\mathcal{S}_{\rm syst}/\partial t)$, will lead directly to our previous result, $\Delta\mathcal{S}_{\rm syst} = -k_B \ln[p(\sigma_1,t_1)/p(\sigma_0,t_0)]$. The integration of the entropy current $\mathcal{I}_\eta$, leading to $\Delta\mathcal{S}_{\rm flow}$, is a far more complex problem. The flow is caused by the interactions $\lambda\mathcal{V}$ with the reservoir but the mesoscopic flow is not quantifiable as an attribute in a "quantum phase space." In fact, it renders the process to be non-Markovian in a statistical sense. One might think of a lattice gas, the σ's being site occupancies, which are triggered to change (flip) by the entropy current flowing across a bond. The primary stochastic variable is $\mathcal{I}_\eta(t)$, with the σ's being modulated random variables. The compound process is best described in an abstract vector space $\mathcal{R}$ with the σ's being represented by elements (kets) $|\sigma\rangle$ and their dual space elements (bras) $\langle\sigma|$. The space $\mathcal{R}$ will be the tensor product of the space $\mathcal{F}$ for the flipping and the space $\mathcal{C}$ for the entropy current $\mathcal{I}_\eta(t)$, i.e., $\mathcal{R} = \mathcal{C}\otimes\mathcal{F}$. The generic operator of $\mathcal{C}$ is the master operator M and the corresponding operator in $\mathcal{R}$ will be denoted by $\mathcal{M}$; we note that both operators are linear, an advantage that is always there when we operate on the many-body level.

The formal solution of the ME (4.36) and of its concomitant equation of the same form for the conditional probability is given by

$$P(\sigma,t|\sigma',0) = \langle\sigma|\mathcal{T}\exp\left[-\int_0^t d\vartheta\, M(\vartheta)\right]|\sigma'\rangle, \tag{4.46}$$

with initial value $\delta_{\sigma,\sigma'}$. Here the matrix element denotes the following operation:

$$\langle\sigma|M|\sigma'\rangle = \left[\left[\sum_\sigma\right]\right]\left(w_{\sigma,\sigma'} - w_{\sigma',\sigma}\sum_\sigma \delta_{\sigma,\sigma'}\right), \tag{4.47}$$

where the sum is omitted if it acts on a vector but included when acting on a scalar; action on $p(\sigma,t)$ yields the ME. The operator form in projectors $|\sigma\rangle\langle\sigma|$ and pseudoprojectors $|\sigma\rangle\langle\sigma'|$ is found in Ref. [25], Eq. (2.4) [replace H by M and omit the parentheses (σ',σ)]. These results will be used below.

3. *Asymptotic entropy fluctuation theorem*

We need more detail to find the generating function of the entropy due to flow $\Delta\mathcal{S}_{\rm flow} = \int^t \mathcal{I}_\eta d\vartheta$, whereby we assume that the entropy flow is zero at $t = 0$. Also, henceforth we set $k_B = 1$, i.e., entropy is expressed in dimensionless units $\mathcal{S}/k_B$. In the space $\mathcal{R}$ this now yields

$$\langle e^{-\lambda\,\Delta\mathcal{S}_{\rm flow}}\rangle_> = \sum_{\sigma,\sigma'}\langle\sigma|\mathcal{T}\exp\left[-\int_0^t d\vartheta\,\mathcal{M}_{\mathcal{I}_\eta}(\lambda,\vartheta)\right]|\sigma'\rangle\,p(\sigma',0)\rangle; \tag{4.48}$$

the subscript $>$ refers, as before, to the forward protocol. The matrix elements of the exponential in Eq. (4.48) should not be misconstrued: one must expand the exponential and take the matrix elements in each term, using matrix multiplication for higher orders. Therefore, suffice it to consider the matrix element in the first order. The operator $\mathcal{M}_{\mathcal{I}_\eta}(\lambda,\vartheta)$ can be found from the compounding theorem [29] or from the entropy flow (4.44). Including the variable $\exp(-\lambda)$ in the master expression (4.48), we have, using only the first term of Eq. (4.47) with the second term implied,

$$\langle\sigma|\mathcal{M}_{\mathcal{I}_\eta}(\lambda,\vartheta)|\sigma'\rangle = -w_{\sigma,\sigma'}(\vartheta)\exp\{-\lambda\ln[w_{\sigma,\sigma'}(\vartheta)/w_{\sigma',\sigma}(\vartheta)]\}. \tag{4.49}$$

The following symmetry property is the key to the FT to be obtained,

$$\langle\sigma|\mathcal{M}_{\mathcal{I}_\eta}(\lambda,\vartheta)|\sigma'\rangle = \langle\sigma'|\tilde{\mathcal{M}}_{\mathcal{I}_\eta}(1-\lambda,\vartheta)|\sigma\rangle, \tag{4.50}$$

where the tilde—as always—refers to the backward process. The proof is simple; we note,

$$\begin{aligned}\langle\sigma|\mathcal{M}_{\mathcal{I}_\eta}(\lambda,\vartheta)|\sigma'\rangle &= -w_{\sigma,\sigma'}(\vartheta)\exp\{-\lambda\ln[w_{\sigma,\sigma'}(\vartheta)/w_{\sigma',\sigma}(\vartheta)]\}\\ &= -w_{\sigma',\sigma}(\vartheta)\exp\{(1-\lambda)\ln[w_{\sigma,\sigma'}(\vartheta)/w_{\sigma',\sigma}(\vartheta)]\}\\ &= -w_{\sigma',\sigma}(\vartheta)\exp\{-(1-\lambda)\ln[w_{\sigma',\sigma}(\vartheta)/w_{\sigma,\sigma'}(\vartheta)]\}\\ &= \langle\sigma'|\tilde{\mathcal{M}}_{\mathcal{I}_\eta}(1-\lambda,\vartheta)|\sigma\rangle.\end{aligned} \tag{4.51}$$

We shall now also include the system entropy, $\exp(-\Delta\mathcal{S}_{\rm syst})$, in the generating function. One then obtains

[8] We still note that (4.45) is the "quantum phase-space" equivalent of Boltzmann's H-function statement in μ space, $\partial{\rm H}/\partial t + {\rm div}\mathbf{J}_h = \eta_h(\mathbf{r},t) \leqslant 0$.

the more complete expression,

$$\langle e^{-\lambda\,\Delta\mathcal{S}}\rangle_{>} = \sum_{\sigma,\sigma'} (p(\sigma,t))^{\lambda}\langle\sigma|\mathcal{T}\exp\left[-\int_0^t d\vartheta\,\mathcal{M}_{\mathcal{I}_\eta}(\lambda,\vartheta)\right]|\sigma'\rangle(p(\sigma',0))^{1-\lambda}\rangle, \tag{4.52}$$

where $\Delta\mathcal{S} = \Delta\mathcal{S}_{\rm total}$. Using the symmetry property (4.50), as well as $\int_0^t f(\vartheta)d\vartheta = \int_0^t f(t-\vartheta)\,d\vartheta$, needed in the reverse process, it is a minor task to obtain our main result; hence,

$$\begin{aligned}\langle e^{-\lambda\,\Delta\mathcal{S}}\rangle_{>} &= \sum_{\sigma,\sigma'} (p(\sigma,t))^{\lambda}\langle\sigma|\mathcal{T}\exp\left[-\int_0^t d\vartheta\,\mathcal{M}_{\mathcal{I}_\eta}(\lambda,\vartheta)\right]|\sigma'\rangle(p(\sigma',0))^{-\lambda}p(\sigma',0)\rangle\\ &= \sum_{\sigma,\sigma'} (p(\sigma',0))^{1-\lambda}\langle\sigma'|\mathcal{T}\exp\left[-\int_0^t d\vartheta\,\tilde{\mathcal{M}}_{\mathcal{I}_\eta}(1-\lambda,t-\vartheta)\right]|\sigma\rangle(p(\sigma,t))^{\lambda}\rangle\\ &= \sum_{\sigma,\sigma'} (p(\sigma',0))^{1-\lambda}\langle\sigma'|\mathcal{T}\exp\left[-\int_0^t d\vartheta\,\tilde{\mathcal{M}}_{\mathcal{I}_\eta}(1-\lambda,t-\vartheta)\right]|\sigma\rangle(p(\sigma,t))^{-(1-\lambda)}p(\sigma,t)\rangle.\end{aligned} \tag{4.53}$$

We thus established

$$\langle e^{-\lambda\,\Delta\mathcal{S}}\rangle_{>} = \langle e^{-(1-\lambda)\,\Delta\mathcal{S}}\rangle_{<}. \tag{4.54}$$

It remains to invert the generating function. The left-hand side yields $p(\Delta\mathcal{S})$. For the right-hand side we have the inverse, setting $R \equiv \Delta\,\mathcal{S}$,

$$\begin{aligned}&\int_{c-i\infty}^{c+i\infty} d\lambda\, e^{\lambda R}\int_0^\infty dR'\,\tilde{p}(R')\,e^{-(1-\lambda)R'}\\ &= \int_0^\infty dR' e^{-R'}\,\tilde{p}(R')\int_{-i\infty}^{i\infty} d\lambda\, e^{\lambda(R+R')}\\ &= \int_0^\infty dR' e^{-R'}\,\tilde{p}(R')\,\delta(R+R')\\ &= e^{R}\tilde{p}(-R).\end{aligned} \tag{4.55}$$

Or, for the FT,

$$p(\Delta\mathcal{S}) = e^{\Delta\mathcal{S}}\tilde{p}(-\Delta\mathcal{S}). \tag{4.56}$$

Let us assume that the system reaches a steady state at some point in time, so that $\tilde{p} = p$, from where on there is a steady entropy production $\hat{\eta}$; then $\Delta\mathcal{S} \sim \hat{\eta}\,t$. The asymptotic FT now reads

$$p(\hat{\eta}) \sim e^{\hat{\eta}t}p(-\hat{\eta}) \tag{4.57}$$

or, more succinctly,

$$\lim_{t\to\infty}\left(-\frac{1}{t}\ln\left[\frac{p(-\hat{\eta})}{p(\hat{\eta})}\right]\right) = \hat{\eta}, \tag{4.58}$$

which is the form implied in most treatments, see, e.g., Kurchan [21] or Lebowitz and Spohn [23].

Finally, we should emphasize that the treatment in this Sec. V B for the entropy FT fully takes account of the irreversible character of the time evolution based on the master equation; this in contrast to what superficially looks like an equivalent derivation by Talkner *et al.* [18] for the Crooks-Tasaki work FT discussed at the end of the Sec. V A, which employs the evolution operator of the unperturbed von Neumann equation.

V. OTHER ASPECTS

A. Nonequilibrium extension of Mazo's lemma of linear response theory

While we believe that the time-reversal symmetry property of Eq. (2.6) is at the core of nonequilibrium fluctuation theorems, there are some developments in standard linear response theory that are worth extending to driven nonequilibrium systems. As such, we will briefly consider what has become known as *Mazo*'s *lemma* published in 1969 [37], although the lemma may have been around in earlier years. The essence of the lemma is that it relates forward-time correlation functions to reverse-time correlation functions; it is necessary for the derivation of the quantum version of the fluctuation-dissipation theorem, cf. Kubo [36], which ties the response function to the Fourier transform of equilibrium correlation functions.

Thus, let C and D be two observables represented by Schrödinger operators $\mathcal{C}$ and $\mathcal{D}$. Their Heisenberg forms will have a bracketed time dependence, whereby at $t = 0$ they coincide with their Schrödinger equivalents. Mazo's lemma now reads

$$\begin{aligned}&\int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathrm{Tr}\{\rho_{\rm eq}\mathcal{C}(t)\mathcal{D}\}\\ &\quad = e^{-\beta\hbar\omega}\int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathrm{Tr}\{\rho_{\rm eq}\mathcal{D}\,\mathcal{C}(t)\}.\end{aligned} \tag{5.1}$$

The proof, reproduced nearly verbatim in our book (Ref. [8], subsection 16.3.3) is based on analytical extension in the complex plane. Thus, starting with the right-hand side and multiplying with the partition function, we have

$$\begin{aligned}&\int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathrm{Tr}\{e^{-\beta\mathcal{H}}\mathcal{D}\,\mathcal{C}(t)\}\\ &= \int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathrm{Tr}\{e^{-\beta\mathcal{H}}\mathcal{D}\,e^{i\mathcal{H}t/\hbar}\mathcal{C}e^{-i\mathcal{H}t/\hbar}\}\\ &= \int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathrm{Tr}\{e^{i\mathcal{H}t/\hbar}\mathcal{C}e^{-i\mathcal{H}t/\hbar}e^{-\beta\mathcal{H}}\mathcal{D}\}\end{aligned}$$

$$= e^{\beta\hbar\omega}\int_{-\infty}^{\infty} dt\, e^{-i\omega(t-i\hbar\beta)}\mathrm{Tr}\,\{e^{-\beta\mathcal{H}}e^{i\mathcal{H}(t-i\hbar\beta)/\hbar} \times \mathcal{C}e^{-i\mathcal{H}(t-i\hbar\beta)/\hbar}\mathcal{D}\}$$
$$= e^{\beta\hbar\omega}\int_{-\infty-i\hbar\beta}^{\infty-i\hbar\beta} dz\, e^{-i\omega z}\mathrm{Tr}\{\, e^{-\beta\mathcal{H}}\mathcal{C}(z)\,\mathcal{D}\}. \quad (5.2)$$

There is analyticity on $0 \geqslant \mathrm{Im}z \geqslant -\hbar\beta$, cf. Ref. [22], and we make a contour integration along the real axis from $-R$ to $+R$, the line $z = x - i\hbar\beta$, and the lines $z = \pm R - iy$. Noting that the contributions along the latter vanish for $R \to \infty$ due to the mixing property, the final integral of Eq. (5.2) can be replaced by the real axis integral $e^{\beta\hbar\omega}\int_{-\infty}^{\infty} dt \cdots$ and (5.1) follows. Alternately, we can avoid analytical extension and evaluate the trace in the representation $|\eta\rangle$ of $\mathcal{H}$. For the left-hand side of Eq. (5.1) one gets

lhs (5.1)
$$= \sum_{\eta,\eta'}\int_{-\infty}^{\infty} dt\, e^{-i\omega t}\, p_{\mathrm{eq}}(\eta)\,\langle\eta|\mathcal{C}|\eta'\rangle e^{i(\varepsilon_\eta-\varepsilon_{\eta'})t/\hbar}\langle\eta'|\mathcal{D}|\eta\rangle$$
$$= 2\pi\hbar\sum_{\eta,\eta'} Z^{-1}e^{-\beta\varepsilon_\eta}\langle\eta|\mathcal{C}|\eta'\rangle\,\langle\eta'|\mathcal{D}|\eta\rangle\,\delta[\hbar\omega-(\varepsilon_\eta-\varepsilon_{\eta'})]. \quad (5.3)$$

Observing the δ function, we substitute $\exp(-\beta\varepsilon_\eta) = \exp(-\beta\varepsilon_{\eta'})\exp(-\beta\hbar\omega)$; Eq. (5.3) then yields the right-hand side of Eq. (5.1). In the notation of the present article, Mazo's lemma has given us the connection

$$\int_{-\infty}^{\infty} dt\, e^{-i\,\omega t}\mathcal{T}[\langle\mathcal{C}(t)\mathcal{D}(0)\rangle_{\mathrm{eq}}] = e^{-\beta\hbar\omega}\int_{-\infty}^{\infty} dt e^{-i\omega t}\mathcal{T}^{-1}[\langle\mathcal{D}(0)\,\mathcal{C}(t)\rangle_{\mathrm{eq}}]. \quad (5.4)$$

In a driven system subject to the protocol $\xi(t)$, the evolution operator (2.2) can be approximated by a stepwise defined function; the integral in Eq. (2.2) then yields a form as in the middle member of Eq. (2.13), which we repeat here,

$$U(t,0) = \mathcal{T}\exp\sum_{i=0}^{n-1}\big[-i\mathcal{H}_{\xi_i}(t_{i+1}-t_i)/\hbar\big]; \quad (5.5)$$

note that the superscript 0 on $\mathcal{H}$ has been left off and that we begin at $t' = 0$. For the forward protocol it is convenient to start from a canonical equilibrium state after which we leave the system isolated until time t; after finishing the process, we reconnect with the bath without doing work. Likewise, for the reverse protocol, we start from the equilibrated state at t and run the reverse process to time 0, after which we reconnect with the reservoir. Mazo's lemma is still valid on each subinterval. However, when we reach a point of discontinuity, the Hamiltonian and its eigenstates change. On each subinterval, the Hamiltonian will, as before, be associated with the preceding breakpoint, i.e., on $t_i \leqslant t < t_{i+1}$ it will be denoted by $\mathcal{H}_i$, with eigenstates $|\eta_i\rangle$. We can now use the same algebra as before; additionally, we must multiply the new result by the ratio of the partition functions, $Z(\eta_1)/Z(\eta_0)$.

We thus obtain the extension of Mazo's lemma,

$$\int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathcal{T}[\langle\mathcal{C}(t)\mathcal{D}(0)\rangle_{\mathrm{can}}] = e^{\beta(\varepsilon_{\eta1}-\varepsilon_{\eta0})}\, e^{-\beta\,[F^0(t_1)-F^0(0)]} \times \int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathcal{T}^{-1}[\langle\mathcal{D}(0)\,\mathcal{C}(t)\rangle_{\mathrm{can}}]. \quad (5.6)$$

We note that this result involved the δ function $\delta(|\hbar\omega| - \varepsilon_{\eta1} - \varepsilon_{\eta0})$; for the implications, see footnote 5. Consequently, only miniexcitations of the system are accomplished; the many required repeats mean that, in practice, the system must be "mostly" connected with the reservoir, as was also foreseen in Crooks' original paper [10].

The next question is in regard to what will happen to these results after we carry out the perturbation procedure to the evolution operator and, consequently, to all Heisenberg operators. "Convergent LRT" has been considered by the author and coworkers in the cited literature and is summarized in the second part of Chapter XVI of Ref. [8]. Modified response functions are found in subsection 16.10.4; correlation functions can be done similarly. Since diagonal parts and nondiagonal parts do not mix in the evaluation of the trace, there will be separate diagonal and nondiagonal correlation functions, which must be evaluated; we note hereby that the diagonal parts involve the resolvent of the master superoperator Λ_d while the nondiagonal parts are based on the resolvent of the interaction Liouville operator $\mathcal{L}^0$. The results obtained from the diagonal correlation functions via this lemma should, therefore, be identical to the results foreseen in the previous sections of this article; the nondiagonal contributions, which do not produce entropy, appear additionally and at times can be significant.

B. Application to the characteristic function of quantum mechanical work

In the present context, we consider once more "work" as given by Eq. (4.12). Now in Eq. (5.6) let $\mathcal{C}(t) = \exp[it\mathcal{H}_H(t_1)/\hbar]$ and $\mathcal{D}(0) = \exp[-it\mathcal{H}_H(0)/\hbar] = \exp[-it\mathcal{H}/\hbar]$; we then have

$$\int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathcal{T}\langle e^{it\,\mathcal{H}_H(t_1)/\hbar}e^{-it\mathcal{H}/\hbar}\rangle = e^{\beta(\varepsilon_{\eta1}-\varepsilon_{\eta0})}e^{-\beta[\,F^0(t_1)-F^0(0)]} \times \int_{-\infty}^{\infty} dt\, e^{-i\omega t}\mathcal{T}^{-1}\langle\, e^{-it\mathcal{H}/\hbar}e^{it\mathcal{H}_H(t_1)/\hbar}\rangle. \quad (5.7)$$

Substituting $t = \hbar u$ and letting $\varepsilon_{\eta_1} - \varepsilon_{\eta_0} = w$, this yields

$$\int_{-\infty}^{\infty} du\, e^{-iwu}\langle e^{iu\mathcal{W}}\rangle_{>} = e^{\beta w}e^{-\beta[F^0(t_1)-F^0(0)]}\int_{-\infty}^{\infty} du e^{-iwu}\langle e^{-iu\mathcal{W}}\rangle_{<}, \quad (5.8)$$

These integrals ($\div 2\pi$) are the inverse Fourier transforms of the characteristic functions, giving $p(w)$ and $\tilde{p}(-w)$, respectively. Whence the Crooks-Tasaki FT naturally reemerges,

$$p(w)/\tilde{p}(-w) = \exp[\beta(w-\Delta F^0)]. \quad (5.9)$$

Next, let us write (5.8) as follows:

$$(1/2\pi)\, e^{-\beta w} \int_{-\infty}^{\infty} du\, e^{-iwu} \langle e^{iu\mathcal{W}} \rangle_{>} = (1/2\pi)\, e^{-\beta \Delta F^0} \int_{-\infty}^{\infty} du\, e^{-iwu} \langle e^{-iu\mathcal{W}} \rangle_{<}. \tag{5.10}$$

With the Fourier transform of $e^{-\beta w}$ being $2\pi\, \delta(u + i\beta)$, the inverse transform of (5.10) is by the Faltung theorem,

$$\int_{-\infty}^{\infty} du' \delta\,[i\beta + (u - u')]\, \langle e^{iu'\mathcal{W}} \rangle_{>} = \; e^{-\beta \Delta F^0} \langle e^{-iu\mathcal{W}} \rangle_{<} \tag{5.11}$$

or,

$$\langle e^{(iu-\beta)\mathcal{W}} \rangle_{>} \; = e^{-\beta \Delta F^0} \langle e^{-iu\mathcal{W}} \rangle_{<}. \tag{5.12}$$

This is actually *a new fluctuation theorem* involving the characteristic functions for work in the forward and reverse protocol. Setting $u = 0$ gives the quantum Jarzynski work-energy theorem.

Altogether, Mazo's lemma, extended to non-equilibrium driven thermostatted systems, is very useful for quantum-mechanical considerations involving time reversal.

VI. CONCLUSIONS

This article was motivated by two facts: First, there are relatively few papers on fluctuation theorems (FT's), as well as work–energy relationships (W–E's), that are strictly quantum mechanical in origin. Second, the papers that do emanate from quantum principles are based on the von Neumann equation or on the "pure" Heisenberg picture, neither of which predicts an increase in entropy, a fact already known to Gibbs in the equivalent Liouville treatment of his time. All FT's and W–E's of this nature cannot really be applicable to driven nonequilibrium processes. We strongly believe that such theorems are either *wrong* or, at best, *incomplete*. The situation is similar as with LRT; LRT gets physical content only after a "stochasticization" – (Kubo's terminology) has been applied. In the present context, it is often possible to do the operations in the normal order, i.e., we incorporate the necessary asymptotic requirements for the system's evolution by working in the interaction picture, thus dealing *a priori* with internal interactions, as well as with the random exchange with the reservoir(s). The need for such a modified treatment was announced in our previous paper [9]. The "conversion" of the von Neumann equation to the entropy-producing master equation for the stationary case had been carried out by Van Hove as early as 1955 [1] and later by Zwanzig, Fano, and the author and collaborators [2–7]. The quest for convergent LRT expressions had also been considered in some of these papers and has been summarized in our recent book [8] in the second part, dealing with nonequilibrium statistical mechanics.

The first part of this article reconsiders the derivation of the ME in the interaction picture, but now for stationary as well as nonstationary processes. In Sec. II we showed that the time-reversal symmetry of the conditional probability for the forward and backward protocol remains valid after an infinite-order perturbational treatment of the evolution operator. In Sec. III the master equation is newly derived, whereby we obtained an explicit closed-form expression for the evolution operator, not established hitherto, both for stationary and nonstationary protocols. These derivations are rather tedious but yield the basic tools necessary for a realistic convergent treatment of FT's and W–E's.

The second part of this article, which, if needs be, can be read without the details of the preceding sections, is then set forth in Secs. IV A, IV B, and V. In Sec. IV A we redo the quantum version of the Crooks-Tasaki FT but based on our convergent entropy-producing evolution probabilities with special emphasis on the meaning of "work," which, in our opinion, cannot be represented by a bounded quantum-mechanical operator in the Hilbert space $\mathcal{S}$ of the Hamiltonian but does have a unitary operator representation in $\mathcal{S}$, in agreement with a similar earlier view expressed in a recent letter by Talkner, Lutz, and Hänggi [17].

In Sec. IV B we then dwell on the various treatments of the entropy FT's. We did not find a satisfactory derivation based on true quantum concepts; most papers, in particular, those of the earliest discoverers, Galavotti-Cohen, Kurchan, Evans-Searles, Lebowitz-Spohn, and Harris-Schütz [20–25], work in classical phase space or deal with "stochastic trajectories," concepts foreign to quantum statistics. Our treatment is exclusively based on the convergent form of time-reversal symmetry derived in the first part of this paper and on probability aspects for the initial and final quantum states of the protocols, with no reference whatsoever to the "path" connecting these states. With the Gibbsian definition of nonequilibrium fluctuating quantum entropy as $-k_B \ln p(\sigma)$, where σ denotes a "mesoscopic state," the various entropy FT's are easily obtained. The symmetry relations between the generating functions of the forward and backward protocol reappear, as well as the asymptotic entropy FT for the stationary state, first formulated by Evans, Cohen, and Morris in 1993 [41].

Altogether, we do not claim fundamentally new results, except for the derivation of the full, nonstationary ME and time-reversal symmetry based on the interaction picture, with entropy-producing probabilities $p(\sigma,t)$. We have looked on what has transpired in the past two decades in the proliferation of papers on FT's and W–E's with critical intent; we may have been iconoclastic for some researchers, whose tenets on work and heat we simply do not share. We have provided new quantum-mechanical meaning for these FT's, it being our view that thermodynamics, no less than the standard model in particle physics, is based on the quantum nature of the universe.

ACKNOWLEDGMENTS

We have appreciated discussions with H. Capel of the University of Amsterdam (ITFA) regarding the derivation of the Pauli master equation by the late Léon Van Hove. Also, we are indebted to J. Ashkenazi and N. Johnson of the physics department of the University of Miami for their critical reading and discussions of this manuscript. Finally, we thank the physics department for financial support.

[1] L. Van Hove, Physica **21**, 517 (1955).
[2] U. Fano, Rev. Mod. Phys. **29**, 74 (1959); U. Fano, in *Lectures on the Many-Body Problem*, Vol 2, edited by E. R. Caniello (Academic Press, New York, 1964).
[3] R. Zwanzig, in *Lectures in Theoretical Physics*, Vol. III, edited by W. E. Britten, B. W. Downs, and J. Downs (Wiley Interscience, New York, 1961), pp. 116–141.
[4] K. M. Van Vliet, J. Math. Phys. **19**, 1345 (1978).
[5] K. M. Van Vliet, J. Math. Phys. **20**, 2573 (1979).
[6] M. Charbonneau, K. M. Van Vliet, and P. Vasilopoulos, J. Math. Phys. **23**, 318 (1982).
[7] K. M. Van Vliet, Can. J. Phys. **56**, 1206 (1978).
[8] C. M. Van Vliet, *Equilibrium and Non-equilibrium Statistical Mechanics* (World Scientific, Singapore, 2010).
[9] C. M. Van Vliet, Physica A **390**, 1917 (2011).
[10] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[11] H. Tasaki, arXiv:cond-mat/0009244.
[12] B. L. Holian, R. Blumenfeld, and P. Gumbsch, Phys. Rev. Lett. **78**, 78 (1997).
[13] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).
[14] C. Jarzynski, C. R. Phys. **8**, 495 (2007).
[15] C. Jarzynski, Prog. Theor. Phys. Suppl. **165**, 1 (2006).
[16] D. Cohen and Y. Imry, Phys. Rev. E **86**, 011111 (2012).
[17] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102(R) (2007).
[18] P. Talkner and P. Hänggi, J. Phys. A: Math. Theor. **40**, F569 (2007).
[19] P. Talkner, M. Campisi, and P. Hänggi, J. Stat. Mech. (2009) P02025.
[20] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995); J. Stat. Phys. **80**, 931 (1995).
[21] J. Kurchan, J. Phys. A: Math Gen. **31**, 3719 (1998).
[22] J. Kurchan, arXiv:cond-mat/0007360.
[23] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
[24] D. J. Evans and D. Searles, Adv. In Phys. **51**, 1528 (2002).
[25] R. J. Harris and G. M. Schütz, J. Stat. Mech. (2007) P07020.
[26] P. Morse and H. Feshbach, *Methods of Theoretical Physics*, Vol. I (McGraw-Hill, New York, 1953), chap. 9.
[27] P. Gaspard, Physica A **369**, 201 (2006).
[28] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007), p. 78ff.
[29] W. Feller, *An Introduction to Probability Theory and its Applications*, 2nd ed. (Wiley, New York, 1957), p. 114.
[30] W. J. Kleen, in *Proceedings of the Eighth International Conference on Noise in Physical Systems*, edited by A. D'Amico and P. Mazetti (North Holland, Amsterdam, 1985), p. 331.
[31] C. M. Van Vliet, in *Proceedings of the Ninth International Conference on Noise in Physical Systems*, edited by C. M. Van Vliet (World Scientific, Singapore, 1987), p. 191.
[32] A. E. Allahverdyan and Th. M. Nieuwenhuizen, Phys. Rev. E **71**, 066102 (2005).
[33] G. N. Bochkov and Yu. E. Kuzovlev, Sov. Phys. JETP **45**, 125 (1977).
[34] J. von Neumann, Ann. Math. **33**, 567 (1932).
[35] F. Riesz and B. S. Z. Nagy, *Leçons d'analyse fonction-nelle*, 4th ed. (Gauthier-Villars, Paris, 1965).
[36] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II (Non-Equilibrium Statistical Mechanics)* (Springer-Verlag, Berlin, 1985).
[37] R. M. Mazo, *Statistical Mechanical Theories of Transport Processes* (Pergamon Press, Oxford, 1967), chap. 10.
[38] N. G. van Kampen, Phys. Norv. **5**, 279 (1971).
[39] C. M. Van Vliet, J. Stat. Phys. **53**, 49 (1988).
[40] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[41] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[42] J. Schnakenberg, Rev. Mod. Phys. **48**, 571 (1976).