Journal of Statistical Physics, Vol. 72, Nos. 3/4, 1993

Irreversibility and Randomness in Linear Response Theory

Göran Lindblad[1]

Received February 12, 1993

Recall that the fluctuation-dissipation theorem connects the response function of a passive linear system and the spectral density of the stationary stochastic process which describes the thermal fluctuations in the system. It is shown that the classical limit ($\hbar = 0$) of the fluctuation-dissipation theorem implies a correspondence between systems which are reversible in the sense that the energy used to drive them away from equilibrium is completely recoverable as work and processes which are deterministic in the sense of Wiener's prediction theory, while irreversible systems correspond to nondeterministic processes. This correspondence is expressed by a simple transformation between the operator kernel which determines the optimal choice of the time-dependent force and the linear predictor for the stochastic process. For quantum systems this correspondence does not hold; the fluctuations are always of the deterministic type for any finite temperature, but the system is not necessarily reversible. For irreversible systems a formula is derived for the instantaneous entropy production which is a generalization of the standard one for Markovian dynamics.

KEY WORDS: Fluctuation-dissipation theorem; passive systems; thermal fluctuations; stochastic processes; entropy production.

1. INTRODUCTION

The linear theory of irreversible processes, including linear response theory, Onsager's reciprocal relations, and the fluctuation-dissipation (FD) theorem, is a well-established part of physics. In spite of this it seems that there are certain fundamental problems in the formalism which are almost always overlooked. One such basic aspect concerns the very definition of irreversibility which is relevant for these models of irreversible processes. Of

[1] Department of Theoretical Physics, Royal Institute of Technology, S-100 44 Stockholm, Sweden.

course, this problem is related to that of finding a formula for the entropy production in the system. Long ago J. Meixner introduced an interesting concept of irreversibility in this context which seems to have been largely forgotten.[1,2] In this formalism the irreversibility occurs when the external forces drive the system away from equilibrium and it is not possible to recover completely the work done on the system no matter how we choose the time dependence of the forces. On the other hand, for some systems the response function is such that all the work can be recovered, giving a reversible character to the system even though there is a relaxation to equilibrium in a weaker sense.[3]

An *a priori* unrelated problem concerns the properties of the stationary stochastic process which describes the thermal fluctuations in equilibrium. Such processes can be deterministic in the sense that for any time $t > 0$ there are linear functions of the process in $(-\infty, 0)$ which approximate the random variable at time t with arbitrary accuracy, thus predicting this future outcome from a full knowledge of the past. Others are nondeterministic in the sense that there is an optimal predictor which gives a best approximation for the outcome at time t, but such that the error has a nonzero variance. Obviously the latter situation is seen as intrinsically random (chaotic), and intuitively this is often taken as an important aspect, or even a definition, of irreversibility.

Here it will be shown that for linear systems satisfying the FD theorem in a classical ($\hbar \to 0$) limit there is a simple correspondence between the optimal work cycles in Meixner's theory and the linear predictors for the stochastic processes (Section 4). They satisfy essentially the same Wiener–Hopf integral equation, and there is a simple transformation (4.6) between them. Reversible systems correspond to deterministic processes, irreversible systems to nondeterministic processes.

It is also possible to derive a formula for the entropy production which is a simple quadratic expression in the external forces and which generalizes the familiar form for Markov processes (Section 5). The formula contains the optimal work cycle for the given past history of the system, which means that it is not explicit unless one can solve the Wiener–Hopf equation. In the Markov case (Section 6) there is a solution, in fact the maximal work is given by an infinitely slow, reversible process similar to a Carnot cycle. It corresponds to the familar predictor for Markov processes, which depends on the last observed outcome only.

When the FD theorem has the quantum form the correspondence outlined above does not hold (Section 7). In fact, the quantum correlation functions for thermal fluctuations at a finite temperature always have a deterministic property, while the response functions are not necessarily of the reversible type.

2. LINEAR SYSTEMS

Here the formalism of linear response theory and the FD theorem are introduced in a condensed way to establish the notation. There are numerous standard works[4-6]; here the notation conforms closely to that of ref. 6. Consider a deterministic dynamical system $\mathscr{S}$ where the observational state at time t is defined by n real quantities

$$\mathbf{Y}(t) = \{Y_\mu(t)\}_1^n$$

It is assumed that if no forces act on $\mathscr{S}$, a state of equilibrium will be achieved asymptotically when $t \to \infty$. This state is normalized to be $\mathbf{Y} = 0$. The external forces

$$\mathbf{X}(t) = \{X_\mu(t)\}_1^n$$

drive $\mathscr{S}$ away from equilibrium while performing work on it. The variables are chosen in such a way that the work is given by the following expression, where the summation convention is used for repeated Greek indices:

$$W = \int dt\, \mathbf{X}(t)\, \dot{\mathbf{Y}}(t) \equiv \int dt\, X_\mu(t)\, \dot{Y}_\mu(t) \tag{2.1}$$

for cyclic changes in the forces starting from equilibrium, i.e.,

$$\mathbf{X}(-\infty) = \mathbf{X}(\infty) = 0$$

In the following we will consider only systems where the variables $\mathbf{Y}$ depend in a linear homogeneous way on the forces. Furthermore, it is assumed that this action is time-homogeneous and causal, which means that the linear relation is of the form

$$Y_\mu(t) = \int_{-\infty}^{t} ds\, \Phi_{\mu\nu}(t-s)\, X_\nu(s) \tag{2.2}$$

where it is assumed that

$$\lim_{t \to \infty} \mathbf{\Phi}(t) = 0$$

The real matrix-valued integral kernel $\mathbf{\Phi}(t)$ (the response function) is *a priori* defined only for $t \geqslant 0$, but it is convenient to continue it to a matrix function on $\mathbb{R}$ satisfying (T stands for transpose)

$$\mathbf{\Phi}(-t) = \mathbf{\Phi}(t)^T$$

Also the relaxation function defined by

$$\boldsymbol{\Psi}(t) = \int_t^\infty ds\, \boldsymbol{\Phi}(s)$$

for $t \geqslant 0$ is continued in such a way that $\boldsymbol{\Psi}(-t) = \boldsymbol{\Psi}(t)^T$. Note that at $t = 0$ the derivative $\dot{\boldsymbol{\Psi}}(t)$ will have a step of magnitude

$$\dot{\boldsymbol{\Psi}}(0+) - \dot{\boldsymbol{\Psi}}(0-) = -2\boldsymbol{\Phi}(0) \tag{2.3}$$

The work can be expressed in terms of the response and relaxation functions as follows:

$$W = -\tfrac{1}{2} \iint ds\, dt\, \ddot{\Psi}_{\mu\nu}(s-t)\, X_\mu(s)\, X_\nu(t) \tag{2.4}$$

$$= \tfrac{1}{2} \iint ds\, dt\, \Psi_{\mu\nu}(s-t)\, \dot{X}_\mu(s)\, \dot{X}_\nu(t) \tag{2.5}$$

There are clearly conditions to be fulfilled in order that partial integration will equate the two expressions, but these details will be left out. Again using the notation of ref. 6, we introduce the Fourier transform

$$X_\mu[\omega] = \int_{-\infty}^\infty dt\, e^{i\omega t} X_\mu(t)$$

Hermitian conjugation of the Fourier transform of $\boldsymbol{\Psi}(t)$ gives

$$\boldsymbol{\Psi}[\omega]^\dagger = \boldsymbol{\Psi}[-\omega]$$

Using the Fourier transforms, the work is expressed as

$$W = \frac{1}{4\pi} \int d\omega\, \omega^2 \Psi_{\mu\nu}[\omega]\, X_\mu[\omega]^*\, X_\nu[\omega] \tag{2.6}$$

From (2.6) it is evident that the integrals can develop singularities unless the Fourier transforms of the forces are nice enough. Precise assumptions will not be spelled out here.

The equilibrium state $\mathbf{Y} = 0$ is assumed to be a macroscopic representation of a thermodynamic equilibrium state of given inverse temperature $\beta = 1/k_B T$. Let the thermal fluctuations in the variables $\mathbf{Y}$ be represented by variables $\mathbf{Z}$ which in general must be taken to be noncommuting, self-adjoint operators in the quantum case. Assume for simplicity that the ensemble and time averages of the fluctuations vanish. Introduce the autocorrelation (covariance) function as the ensemble average

$$R_{\mu\nu}(t) = \langle Z_\mu(t)\, Z_\nu(0) \rangle \tag{2.7}$$

Due to the noncommuting nature of the operators, this quantity is complex-valued in general, becoming real-valued in a classical limit. It is a positive semidefinite matrix function, which means that

$$\sum_{k,l,\mu,\nu} \xi^*_{k,\mu} R_{\mu\nu}(t_k - t_l)\, \xi_{l,\nu} \geqslant 0 \tag{2.8}$$

for all $\{t_k \in \mathbb{R},\ \xi_{k,\mu} \in \mathbb{C},\ k = 1,\ldots, N,\ \forall N\}$.

From (2.8) and Bochner's theorem it follows that it is a Fourier transform of a matrix-valued measure (Feller,[7] Chapter 19). Here we are interested in systems with an absolutely continuous spectrum where

$$\lim_{|t| \to \infty} \mathbf{R}(t) = 0$$

and we write it as a Fourier transform of a matrix function

$$\mathbf{R}(t) = \frac{1}{2\pi} \int d\omega\, e^{-i\omega t} \mathbf{R}[\omega]$$

where $\mathbf{R}[\omega]$ is a positive-semidefinite matrix for (almost) every ω, a statement written as

$$\mathbf{R}[\omega] \geqslant 0 \tag{2.9}$$

It is also convenient to introduce the following symmetrized correlation function, which is always real-valued:

$$\mathbf{D}(t) = \tfrac{1}{2}\{\mathbf{R}(t) + \mathbf{R}(-t)^T\}$$

In the commutative case we see that $\mathbf{D}(t) = \mathbf{R}(t)$. One can express the FD theorem in terms of the Fourier transform of $\mathbf{D}$,

$$\begin{aligned} \mathbf{D}[\omega] &= E_\beta(\hbar\omega)\, \mathbf{\Psi}[\omega] \\ E_\beta(\hbar\omega) &= \frac{\hbar\omega}{2} \coth\left(\frac{\beta\hbar\omega}{2}\right) \end{aligned} \tag{2.10}$$

The derivation is based on the microscopic quantum equations of motion and the properties of Gibbs canonical ensemble.[6,8] The problems of justifying the perturbation approach used are discussed in refs. 9 and 10. In the classical limit we find

$$\lim_{\hbar \to 0} E_\beta(\hbar\omega) = \beta^{-1}$$

Thus, in this limit the FD theorem reads

$$\beta \mathbf{R}[\omega] = \mathbf{\Psi}[\omega] \tag{2.11}$$

3. PREDICTION OF STATIONARY PROCESSES

Here we sum up some known facts about the linear prediction of stationary stochastic processes.[11,12] The basic assumption is that only the autocorrelation (2.7) is relevant, which is true for Gaussian processes. Assume that we have observed the outcome of the process $\{\mathbf{Z}(t)\}$ in $(-\infty, 0)$ and that we want to make a prediction of the outcome at some time $t>0$ through a linear combination of the outcomes of the observations

$$PZ_\mu(t) = \int_{-\infty}^{0} ds\, K(t,s)_{\mu\nu}\, Z_\nu(s) \tag{3.1}$$

From the fact that we are dealing only with second moments it is natural to choose as a measure of the error in the prediction the following quadratic expression:

$$\left\langle \sum_\mu |PZ_\mu(t) - Z_\mu(t)|^2 \right\rangle \tag{3.2}$$

A simple variational principle gives that the best prediction in this sense is obtained if the predictor satisfies the equations

$$\langle (PZ_\mu(t) - Z_\mu(t))\, Z_\nu(s) \rangle = 0 \tag{3.3}$$

for all $s \in (-\infty, 0)$. In fact, for any Z_μ which is a linear combination of the type (3.1) it holds that

$$\langle |Z_\mu - Z_\mu(t)|^2 \rangle = \langle |Z_\mu - PZ_\mu(t)|^2 \rangle + \langle |PZ_\mu(t) - Z_\mu(t)|^2 \rangle$$

This shows directly that the minimum is achieved for $Z_\mu = PZ_\mu(t)$. Now (3.3) reads

$$\langle PZ_\mu(t)\, Z_\nu(s) \rangle = R_{\mu\nu}(t-s)$$

or more explicitly, in terms of the kernel of (3.1), for all $s \in (-\infty, 0)$,

$$\int_{-\infty}^{0} du\, K(t,u)_{\mu\sigma}\, R_{\sigma\nu}(u-s) = R_{\mu\nu}(t-s) \tag{3.4}$$

which is a Wiener–Hopf integral equation for $\mathbf{K}(t,s)$. The method of solving this type of equation involves a factorization technique which will not be described here.[12–14]

The predictability properties of stationary processes can be grouped into two exclusive and complementary classes. With the notation used above they can be introduced in the following way.[12]

(1) The process $\{\mathbf{Z}(t)\}$ is *regular* (completely nondeterministic) if

$$P\mathbf{Z}(t) \to \langle \mathbf{Z} \rangle = 0 \qquad \text{as} \quad t \to \infty$$

This means that the knowledge of the past outcome of the process will not allow us to improve the predictions about the distant future. For every t there is a unique linear predictor. The notation linearly regular is used in ref. 12.

(2) The process is *singular* (deterministic) if for all $t > 0$

$$P\mathbf{Z}(t) = \mathbf{Z}(t)$$

This says that the future is determined by a complete knowledge of the past. However, it is clear that there can be no unique linear predictor in this case.

An important theorem says that every process of the type we consider here can be decomposed in a unique way into a regular and a singular component, and the two components are uncorrelated.[12] If there is a regular component the process is called *nondeterministic*. In the case $n = 1$ the characterization can be given in a simple way through the spectral density (ref. 12, §3.2).

(1′) The process $\{Z(t)\}$ is nondeterministic if the spectral density $R[\omega]$ defining the part of the spectrum which is absolutely continuous relative to Lebesgue measure satisfies

$$\int_{-\infty}^{\infty} d\omega \, \frac{\ln R[\omega]}{1+\omega^2} > -\infty \tag{3.5}$$

Then the absolutely continuous part of the spectrum defines the regular component of the process, the singular part defines the singular component.

(2′) The process is singular if there is no absolutely continuous spectrum or if

$$\int_{-\infty}^{\infty} d\omega \, \frac{\ln R[\omega]}{1+\omega^2} = -\infty$$

For $n > 1$ the description is a bit more complex, and more about the subject can be found in Rozanov.[12]

4. PASSIVE SYSTEMS AND IRREVERSIBILITY

Meixner[1,2] developed a formalism for linear passive systems and a notion of irreversibility which is the basis for this work. It was also employed in ref. 15 in a more general nonlinear setting. The thermodynamic equilibrium state has the property of passivity: the work performed on the system by the forces starting from equilibrium is never negative, i.e., $W \geqslant 0$ in (2.1). In other words, a passive system cannot produce work in cyclic processes starting from the equilibrium state $\mathbf{Y} = 0$. This is Kelvin's form of the second law of thermodynamics. With some additional assumptions the KMS property of the autocorrelation functions characterizing the thermal equilibrium states in quantum statistical mechanics can be recovered from the passivity.[16,17] In the linear regime the passivity condition is inevitably weaker in its consequences.

If the form (2.6) of the work is used, it is clear that the passivity takes the form: for almost all real ω

$$\mathbf{\Psi}[\omega] \geqslant 0 \tag{4.1}$$

The validity of this property follows directly from the form (2.10) of the FD theorem, as $\mathbf{D}[\omega] \geqslant 0$. Consequently $\mathbf{\Psi}(t)$ and $-\ddot{\mathbf{\Psi}}(t)$ are positive semidefinite in the sense (2.8). From this and (2.3) it follows that $\mathbf{\Phi}(0) \geqslant 0$. Of course, the derivation of the FD theorem from the microscopic dynamics uses the properties of the Gibbs canonical state which is the origin of the KMS condition as well.

For non-Markovian systems this initial condition of being in equilibrium must include the assumption that it was not acted upon by external forces in the entire past history of the system. In fact, such systems can retain a memory of the past and the instantaneous state does not determine the future evolution for a given time-dependent force. For some systems this memory effect does not decay with time. In general, if work is performed on the system by a cyclic force in $(-\infty, 0]$, then it will be possible to recover a part of this energy during $(0, \infty)$. Meixner saw the possibility of basing the concepts of reversibility and irreversibility on the property of recoverability of work, and he used the following definition.

Definition 1. A passive system is called *reversible* if for every time-dependent force $\{\mathbf{X}(t),\ t \in (-\infty, 0]\}$ there is a continuation $\{\mathbf{X}(t),\ t \in (0, \infty)\}$ such that the total work on the system over $(-\infty, \infty)$ is zero, otherwise it is called *irreversible*.

König and Tobergte[3] derived a necessary and sufficient condition for the irreversibility of the system in the case $n = 1$. With the present notation it may be written

$$\int_{-\infty}^{\infty} d\omega \frac{\ln(\omega^2 \Psi[\omega])}{1 + \omega^2} > -\infty \tag{4.2}$$

From the FD theorem in its classical form (2.7) and Eq. (3.5) one can immediately recognize the similarity between the condition for irreversibility in Meixner's sense and that for regularity of the associated stationary process in the sense of Wiener. It is evident that the factor ω^2 makes no difference to the convergence or not of the integral in (4.2), so the two conditions are actually identical.

The reason for this relation between irreversibility and unpredictability becomes evident from the variational equation for the total work performed on the system. When the $\mathbf{X}(t)$ is given up to time 0, a variation

$$\mathbf{X}(t) \to \mathbf{X}(t) + \delta\mathbf{X}(t)$$
$$\delta\mathbf{X}(t) = 0 \qquad \text{for} \quad t \in (-\infty, 0]$$

gives a variation in W which is set to zero to solve for $\mathbf{X}(t)$, $t > 0$. In order to underline the similarity with the variational equation (3.4), we use the expression (2.5). We find for all $t > 0$

$$\int_{-\infty}^{\infty} ds\, \Psi_{\mu\nu}(t-s)\, \dot{X}_\nu(s) = 0$$

Hence we want to solve for $\dot{\mathbf{X}}(t)$, $t > 0$, in

$$\int_0^{\infty} ds\, \Psi_{\mu\nu}(t-s)\, \dot{X}_\nu(s) = -\int_{-\infty}^{0} ds\, \Psi_{\mu\nu}(t-s)\, \dot{X}_\nu(s) \tag{4.3}$$

It will follow from the calculations of Section 5 that the solution is really a minimum. It is clear from the linearity of the system and the quadratic nature of the expression to be minimized that the solution is a linear function of the given data, that is, there is a relation

$$\dot{X}_\mu(t) = \int_{-\infty}^{0} du\, L_{\mu\nu}(t, u)\, \dot{X}_\nu(u) \tag{4.4}$$

which should hold for any choice of $\dot{\mathbf{X}}(t)$, $t < 0$. From this requirement it follows immediately that $\mathbf{L}$ is the solution to the integral equation

$$\int_0^{\infty} du\, \Psi_{\mu\sigma}(t-u)\, L_{\sigma\nu}(u, s) = -\Psi_{\mu\nu}(t-s) \tag{4.5}$$

This is again a Wiener–Hopf equation. In the commutative case where (2.11) holds we can rewrite this equation in the same form as (3.4),

$$\int_{-\infty}^{0} du\, L_{\sigma\mu}(-u, -t)\, R_{\sigma\nu}(u-s) = -R_{\mu\nu}(t-s)$$

and consequently we can identify the solutions for the two problems

$$\mathbf{K}(t, u) = -\mathbf{L}(-u, -t)^T \tag{4.6}$$

In fact the two problems are isomorphic, as they have the same Hilbert space structure. Consequently the minima in the quadratic forms (3.2) and W are zero and nonzero simultaneously. The same goes for the unicity of the solutions. Furthermore, the regular processes will correspond to systems where the available work at time t goes to zero as $t \to \infty$ when $\mathbf{X}$ is held constant in $(0, t)$. This, by definition, represents a relaxation to equilibrium (Section 5). We can conclude that in linear response models satisfying the *classical* form of the FD theorem there is the correspondence:

deterministic processes ⇔ *reversible* systems

nondeterministic processes ⇔ *irreversible* systems

regular processes ⇔ systems relaxing to equilibrium

5. ENTROPY PRODUCTION

With the formalism adopted in Section 4 it is natural to use the available work at any instant as a measure of the distance to equilibrium of the state at that instant. In a finite system we could interpret this as follows. The system has a higher energy than the equilibrium state of the same entropy, and an optimal work process allows us to reach the equilibrium state of the same entropy but a lower energy, the rest being delivered as work to the exterior. In general the entropy of nonequilibrium states is not uniquely defined, so for a given set of variables and forces this allows us to define a nonequilibrium entropy starting from a set of work processes.[15] With this definition processes which are not optimal do increase the entropy of the system, as a part of the work available at one moment becomes unavailable.

In the present context the system is infinite (there is a temperature which is constant, independent of the dissipation of the work put into the system). Still we can define an entropy increase when a part ΔW of the available work goes to waste

$$\Delta S = \beta \, \Delta W$$

Note that the entropy has been chosen to be dimensionless. This approach allows us to define an instantaneous entropy production in the following way. Assume for simplicity that we have a situation where the variational equations have unique solutions (which corresponds to a regular process).

Furthermore, let $\ddot{\Psi}(t)$ in (2.4) have the delta-function singularity defined by (2.3) where $\Phi(0)$ is positive definite (hence, nonsingular), while the remainder is a continuous bounded function on $\mathbb{R}$.

Now start from the expression (2.3) for the work and write it in a short-hand way as a quadratic form

$$W = \mathbf{X} * \mathbf{W} * \mathbf{X}$$

where $*$ denotes the convolution and matrix multiplication and where the transpose on the left-hand $\mathbf{X}$ factor is left out for simplicity. Consider the variational equation (4.3) for the optimal work process in $(0, \infty)$, or rather its analog for $\mathbf{X}$. Index the time interval $(-\infty, 0]$ by 0 and $(0, \infty)$ by 1 and write, for example,

$$\Pi_0 \mathbf{X} = \mathbf{X}_0 = \begin{cases} \mathbf{X}(t) & \text{for} \quad t \in (-\infty, 0] \\ 0 & \text{for} \quad t \in (0, \infty) \end{cases}$$

The variational equation for $\mathbf{X}_1$, given $\mathbf{X}_0$, then reads

$$\Pi_1 \mathbf{W} * \mathbf{X}_1 = -\Pi_1 \mathbf{W} * \mathbf{X}_0$$

If $\mathbf{X}_1$ is a solution, then it holds that

$$\mathbf{X} * \mathbf{W} * \mathbf{X} = \mathbf{X}_0 * \mathbf{W} * \mathbf{X}_0 - \mathbf{X}_1 * \mathbf{W} * \mathbf{X}_1$$

Furthermore, let there be a force which is not an optimal solution,

$$\mathbf{X}' = \mathbf{X}_0 + \mathbf{X}'_1 \equiv \mathbf{X}_0 + \mathbf{X}_1 + \Delta \mathbf{X}_1$$

Then

$$\mathbf{X}' * \mathbf{W} * \mathbf{X}' = \mathbf{X} * \mathbf{W} * \mathbf{X} + \Delta \mathbf{X}_1 * \mathbf{W} * \Delta \mathbf{X}_1$$

We see that the variational equation really gives a minimum for the total work performed on the system.

Now consider three time intervals: that indexed 0 is as before, 1 stands for $(0, \tau]$, where τ is going to approach zero, and 2 stands for (τ, ∞). Let

$$\mathbf{X} = \mathbf{X}_0 + \mathbf{X}_1 + \mathbf{X}_2$$

where $\mathbf{X}_1 + \mathbf{X}_2$ is an optimal solution for given $\mathbf{X}_0$, while

$$\mathbf{X}' = \mathbf{X}_0 + \mathbf{X}'_1 + \mathbf{X}'_2$$

where $\mathbf{X}_2'$ is an optimal process in (τ, ∞) given $\mathbf{X}_0 + \mathbf{X}_1'$. With $\Delta\mathbf{X} = \mathbf{X}' - \mathbf{X}$ it follows that the extra work performed on the system by having a nonoptimal process in $(0, \tau]$ equals

$$\begin{aligned} \Delta\mathbf{X} * \mathbf{W} * \Delta\mathbf{X} = \Delta\mathbf{X}_1 * \mathbf{W} * \Delta\mathbf{X}_1 + \Delta\mathbf{X}_1 * \mathbf{W} * \Delta\mathbf{X}_2 \\ + \Delta\mathbf{X}_2 * \mathbf{W} * \Delta\mathbf{X}_1 + \Delta\mathbf{X}_2 * \mathbf{W} * \Delta\mathbf{X}_2 \end{aligned} \tag{5.1}$$

The singularity at $t = 0$ of $\ddot{\mathbf{\Psi}}(t)$ gives a nontrivial contribution from the diagonal in (2.4). It appears in the first term in the rhs of (5.1) and it is $O(\tau)$. From the equation

$$\Pi_2 \mathbf{W} * \Delta\mathbf{X}_2 = -\Pi_2 \mathbf{W} * \Delta\mathbf{X}_1$$

which follows from the variational equations for $\mathbf{X}_2$ and $\mathbf{X}_2'$, we conclude that $\Delta\mathbf{X}_2$ is $O(\tau)$. Then we find that the three remaining terms in the rhs of (5.1) are $O(\tau^2)$. Hence it follows that

$$\frac{dW}{d\tau} = \Delta\mathbf{X}(0+)\, \mathbf{\Phi}(0)\, \Delta\mathbf{X}(0+) \tag{5.2}$$

where the notation $(0+)$ signifies a limit where $\tau \to 0$ from above and $\Delta\mathbf{X}$ is the difference between the value of the external force and the optimal solution of the variational problem given the past up to $t = 0$, that is, $\mathbf{X}_0$.

When $\mathbf{\Phi}(0) = 0$ or a singular matrix the optimal solution may have a singularity at $t = 0$, which will make the limit (5.2) nonzero for irreversible systems. This case corresponds to Wiener's prediction formula when the optimal predictor contains derivatives.[11]

Using the temperature defined through the FD theorem one can now define the instantaneous entropy production

$$\sigma(t) = \beta\, \Delta\mathbf{X}(t+)\, \mathbf{\Phi}(0)\, \Delta\mathbf{X}(t+) \tag{5.3}$$

which involves the optimal solution in (t, ∞) given the past in $(-\infty, t]$. The total entropy production over all time is proportional to the total amount of work performed on the system

$$\Delta S = \int_{-\infty}^{\infty} dt\, \sigma(t) = \beta W \tag{5.4}$$

as one expects. In the standard texts this is the only entropy production which is discussed in the general non-Markovian case. Without taking into account the dependence of the available work on the past history of the system the instantaneous entropy production can be defined only for

Markovian systems (Ref. 4, §8.5). In next section it is shown that (5.3) gives the expected answer in that case.

We note that the expression for the entropy production depends on the dimension n of the system. If the forces are restricted to a subspace of dimension $m < n$, then the value of (5.3) will be different from that obtained if the system is described by an m-component quantity from the start. This is due to the fact that the solutions of the variational equation (6.1) are taken to be n-dimensional instead of m-dimensional, and more of the work can potentially be recovered. Thus the entropy production depends on the level of description. Only the total integrated entropy increase (5.4) will be independent of this arbitrariness.

6. MARKOVIAN SYSTEMS

The stationary process $\mathbf{Z}(t)$ is Markovian if the function (2.5) has the following form for $t > 0$ (ref. 7, §3.8), using matrix notation:

$$\mathbf{R}(t) = \{\mathbf{R}(0)\}^{1/2}\,\mathbf{T}(t)\{\mathbf{R}(0)\}^{1/2} \tag{6.1}$$

$$\mathbf{T}(t) = \begin{cases} \exp(t\mathbf{L}) & \text{for} \quad t > 0 \\ \exp(-t\mathbf{L}^{\dagger}) & \text{for} \quad t < 0 \end{cases} \tag{6.2}$$

Here the square root is defined due to the fact that $\mathbf{R}(0) \geqslant 0$, which follows from (2.8) (for all t_k equal). In order to satisfy (2.8) it is necessary and sufficient that

$$-\mathbf{L} - \mathbf{L}^{\dagger} \geqslant 0$$

a result derived using Bochner's theorem on the Fourier transform of (6.2) or from the general theory of contractive semigroups ($\mathbf{L}$ is a dissipative operator).[18] In the classical limit $\mathbf{L}$ is real. From (2.3) and (2.11) it then follows that

$$2\mathbf{\Phi}(0) = -\beta(\dot{\mathbf{R}}(0+) - \dot{\mathbf{R}}(0-)) = -\beta\{\mathbf{R}(0)\}^{1/2}\,(\mathbf{L} + \mathbf{L}^{T})\{\mathbf{R}(0)\}^{1/2} \geqslant 0 \tag{6.3}$$

In this case the solution of (3.4) by factorization techniques is rather straightforward, but it will be left out, as the solution is well known for Markov processes, where the last (complete) observation of the system is sufficient for the prediction. We find that the kernel in (3.1) is

$$\mathbf{K}(t, u) = \delta(u)\,\mathbf{R}(t)\,\mathbf{R}(0)^{-1}$$

which, together with (5.2), implies that (3.4) is satisfied.

On the other hand, we know what the optimal work processes look like in this case. For Markovian systems the dynamics is described by the instantaneous state $\mathbf{Y}(t)$ through the following linear equation, which is obtained from (2.2), (2.11), and (6.2):

$$\dot{\mathbf{Y}}(t) = \mathbf{R}(0)^{1/2}\, \mathbf{L}\mathbf{R}(0)^{-1/2}[-\boldsymbol{\Psi}(0)\, \mathbf{X}(t) + \mathbf{Y}(t)] \tag{6.4}$$

The maximal work starting from the state $\mathbf{Y}$ is performed by a reversible Carnot-like process where the starting value in $t = 0+$ of the force is

$$\mathbf{X} = \boldsymbol{\Psi}(0)^{-1}\, \mathbf{Y}$$

and which has an infinitely slow return to the equilibrium value $\mathbf{Y} = 0$. The solution to the variational problem should then be a force constant in $(0, \infty)$,

$$\begin{aligned}\mathbf{X}(0+) = \boldsymbol{\Psi}(0)^{-1}\, \mathbf{Y}(0) &= \boldsymbol{\Psi}(0)^{-1} \int_{-\infty}^{0} du\, \boldsymbol{\Phi}(-u)\, \mathbf{X}(u)\\ &= \mathbf{X}(0-) - \boldsymbol{\Psi}(0)^{-1} \int_{-\infty}^{0} du\, \boldsymbol{\Psi}(-u)\, \dot{\mathbf{X}}(u)\end{aligned}$$

With the classical limit of the FD theorem (2.11) we find that the derivative of the optimal solution reads

$$\dot{\mathbf{X}}(t) = -\delta(t)\, \mathbf{R}(0)^{-1} \int_{-\infty}^{0} du\, \mathbf{R}(-u)\, \dot{\mathbf{X}}(u)$$

and consequently the kernel in (4.4) is

$$\mathbf{L}(t, u) = -\delta(t)\, \mathbf{R}(0)^{-1}\, \mathbf{R}(-u)$$

which means that (4.6) is satisfied. The form of the entropy production is obtained by writing the deviation of the entropy in the state $\mathbf{Y}$ from the equilibrium (maximum) value as the negative quadratic form

$$S(t) = -\frac{\beta}{2}\, \mathbf{Y}(t)\, \boldsymbol{\Psi}(0)^{-1}\, \mathbf{Y}(t) \tag{6.5}$$

This formula is valid only in the Markovian case. The entropy production

$$\sigma(t) = \frac{dS(t)}{dt}$$

is then shown, using (2.11), (6.3), and (6.4), to equal (5.3).

7. DISCUSSION

The relation between determinism and reversibility set down in Section 4 is not valid when $\beta\hbar \neq 0$. The KMS condition for the correlation functions (ref. 17, Chapter 5.3) is equivalent to

$$\mathbf{R}[\omega] = e^{\beta\hbar\omega}\mathbf{R}[-\omega]^T$$

which implies that

$$\int_{-\infty}^{\infty} d\omega\, \mathbf{R}[\omega] = \int_0^{\infty} d\omega\, \{\mathbf{R}[-\omega] + e^{-\beta\hbar\omega}\mathbf{R}[-\omega]^T\} = 2\pi\mathbf{R}(0)$$

But $\mathbf{R}(0)$ is finite, so from (2.9) it is evident that $\mathbf{R}[\omega] \to 0$ exponentially when $\omega \to -\infty$. This means that the spectral density decays too fast in this direction to be of the regular type (this argument was also used in ref. 15, Appendix A.5). It is possible to come to this conclusion directly from the KMS condition using analytic continuation arguments.[19] In ref. 19 it is used to show that in thermal equilibrium with $\beta > 0$ the full set of quantum correlation functions of all orders have a deterministic property. The reversibility properties are determined by $\mathbf{\Psi}$, which is no longer proportional to $\mathbf{R}$. It holds that

$$\mathbf{D}[\omega] = \tfrac{1}{2}(1 + e^{-\beta\hbar\omega})\,\mathbf{R}[\omega]$$

which is not necessarily of fast decay in any direction, and the same goes for $\mathbf{\Psi}[\omega]$.

For a dissipative Markovian quantum dynamics, given by a semigroup of completely positive maps, the entropy production was defined in ref. 20. However, this quantity corresponds to the present formalism only for a semigroup of quasifree maps on the CCR algebra of fluctuation variables, where the relative entropy is a quadratic expression which is essentially (6.5).[21] Note that neither the KMS condition nor the quantum form of the FD theorem is satisfied by the correlation functions in this case.[15,22] On the other hand, a derivation of the dynamics of the macroscopic fluctuation observables shows that it inherits the KMS property of the microscopic dynamics.[10]

It is shown in ref. 15 that the entropy as defined through the available work will be nondecreasing when the description (here the number n of variables) is reduced, a simple consequence of the variational principle. The same conclusion applies to the quantity

$$\Delta S(t) = \int_{-\infty}^{t} ds\, \sigma(s)$$

for a given function $\mathbf{X}(t)$ (belonging to the reduced set of variables). As the total entropy increase is given by (5.4), it is evident that there can be no monotonic behavior of $\sigma(t)$ for all times under a reduction in the description. Any other definition of the entropy increase as a function of the time (for the given external forces) must not be larger than $\Delta S(t)$ if there is not going to be an apparent contradiction of the second law of thermodynamics when formulated in terms of this nonequilibrium entropy.

ACKNOWLEDGMENT

This work was partially supported by the Swedish Natural Science Research Council.

REFERENCES

1. J. Meixner, *Z. Naturforsch.* **46a**:721–726 (1961).
2. J. Meixner, In *Statistical Mechanics of Equilibrium and Non-Equilibrium*, J. Meixner, ed. (North-Holland, Amsterdam, 1965), pp. 52–68.
3. H. König and J. Tobergte, *J. Reine Angew. Math.* **212**:104–108 (1963).
4. S. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland, Amsterdam, 1962).
5. L. E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, Texas, 1980).
6. R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II* (Springer, Berlin, 1985).
7. W. Feller, *An Introduction to Probability Theory and Applications*, Vol. 2, 2nd ed. (Wiley, New York, 1971).
8. R. Kubo, *J. Phys. Soc. Japan* **12**:570–586 (1957).
9. J. Naudts, A. Verbeure, and R. Weder, *Commun. Math. Phys.* **44**:87–99 (1975).
10. D. Goderis, A. Verbeure, and P. Vets, *Commun. Math. Phys.* **136**:265–283 (1991).
11. N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series* (MIT Press, Cambridge, Massachusetts, 1950).
12. Yu. A. Rozanov, *Stationary Random Processes* (Holden-Day, San Francisco, 1967).
13. I. C. Gohberg and M. G. Krein, *Usp. Mat. Nauk* **13**(2):3–72 (1958) [*AMS Transl.* **14**:217–287 (1960)].
14. I. Gohberg and M. A. Kaashoek, in *Modern Mathematical Methods in Transport Theory*, W. Greenberg and J. Polewczak, eds. (Birkhäuser, Basel, 1991), pp. 20–33.
15. G. Lindblad, *Non-Equilibrium Entropy and Irreversibility* (Reidel, Dordrecht, 1983).
16. W. Pusz and S. L. Woronowicz, *Commun. Math. Phys.* **58**:273–290 (1978).
17. O. Bratteli and D. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vol. II (Springer, New York, 1981).
18. E. B. Davies, *One-Parameter Semigroups* (Academic Press, London, 1980).
19. G. Lindblad, Determinism and randomness in quantum dynamics, Preprint, Stockholm (1992).
20. H. Spohn and J. L. Lebowitz, *Adv. Chem. Phys.* **38**:109–142 (1978).
21. D. Goderis, A. Verbeure, and P. Vets, *J. Stat. Phys.* **56**:721–746 (1989).
22. P. Talkner, *Ann. Phys.* (N.Y.) **167**:390–436 (1986).