

PAPER: Classical statistical mechanics, equilibrium and non-equilibrium

Compatibility of linear-response theory with the second law of thermodynamics and the emergence of negative entropy production rates

Pierre Nazé and Marcus V S Bonança

Instituto de Física ‘Gleb Wataghin’, Universidade Estadual de Campinas,
13083-859, Campinas, São Paulo, Brazil
E-mail: p.naze@ifi.unicamp.br and mbonanca@ifi.unicamp.br

Received 26 August 2019

Accepted for publication 28 October 2019

Published 9 January 2020



Online at stacks.iop.org/JSTAT/2020/013206
<https://doi.org/10.1088/1742-5468/ab54ba>

Abstract. The reliability of physical theories depends on whether they agree with well-established physical laws. In this work, we address the compatibility of the Hamiltonian formulation of linear-response theory with the second law of thermodynamics. In order to do so, we verify three complementary aspects often understood as statements of the second law: (1) no dissipation for quasistatic process; (2) dissipation for finite-time processes; and (3) positive entropy production rate. Our analysis focuses on two classes of nonequilibrium isothermal processes: slowly-varying and finite-time but weak ones. For the former, we show that these aspects are easily verified. For the later, we present conditions for the achievement of the first two aspects. We also show that the third one is not always verified, presenting an example based on Brownian motion in which we observe negative values in the entropy production rate. In particular, we compare linear-response and exact results for this example.

Keywords: rigorous results in statistical mechanics

Contents

1. Introduction	2
2. Connecting LR to the second law	3
3. Hamiltonian formulation of LR	6
4. Compatibility for slowly-varying processes	7
4.1. No dissipation for quasistatic processes	8
4.2. Dissipation for finite-time processes.....	8
4.3. Positive entropy production rate	8
5. Compatibility for finite-time but weak processes	9
5.1. No dissipation for quasistatic process	9
5.2. Dissipation for finite-time processes.....	11
5.3. Negative values in the entropy production rate	12
6. Final remarks	17
Acknowledgments	20
Appendix A. Linear-response expression for the generalized force	20
Appendix B. Linear-response expressions for the irreversible work ...	22
Appendix C. Hamiltonian constraints on response functions	23
References	23

1. Introduction

Linear-response theory (LR) is one of the most used frameworks to describe nonequilibrium statistical physics phenomena. Having notorious success in the calculation of kinetic coefficients in transport processes, such as electric conductivity [1, 2], dielectric relaxation [3–6] and nuclear magnetic susceptibility [7–10], the main idea underneath the theory is to provide the response of an observable due to a weak perturbation. The response is encoded by the so-called response function or, equivalently, the relaxation function [11]. Additionally, a great advantage of LR is its twofold aspect of describing a system either by using a microscopical approach or a phenomenological one. In the former case, the relaxation function is deduced directly by solving the Hamiltonian equations (or the Heisenberg ones in the quantum case). In the phenomenological approach, although the underlying Hamiltonian formalism is the same, the relaxation function is obtained from the experimental measurement of the response. In this case, it is quite natural to expect that such relaxation function must be very related to the thermodynamic aspects of the system. However, it is not very clear yet whether the linear-response results derived either from microscopic or phenomenological inputs are entirely compatible with thermodynamics and how this compatibility takes place, although some attempts have been made in that sense [12]. In the present work, we

try to fill this gap, finding mathematical conditions on the relaxation function that make the LR compatible with the second law of thermodynamics. Since we are focusing on the Hamiltonian formulation of LR, our task is related to the old problem of understanding how the macroscopic phenomena are related to the microscopic laws of motion [13–17].

In our analysis, we will consider driven classical systems in the presence of a heat bath and three aspects often taken as statements of the second law will be considered: the first one says that there is no dissipation for systems driven by quasistatic process; the second one says that there is dissipation for systems driven by finite-time process; the third one corresponds to the positivity of the entropy production rate. Considering the two classes of nonequilibrium processes described by LR, namely, slowly-varying and finite-time but weak ones, each one of the aspects just mentioned will be verified, either by finding mathematical conditions on the relaxation function or by presenting a counterexample. This will be the case when we analyze the third aspect for finite-time but weak processes. There we will find that LR predicts negative values of entropy production rates for the paradigmatic example of driven Brownian motion.

The proposition of the positivity of the entropy production rate as an equivalent statement of the second law of thermodynamics dates back to the mid-twentieth century in the works of Prigogine and contemporaries [18, 19]. In their local formulation of the second law, it is stated that the differential of the internal entropy production of the system must be positive, which would imply that the entropy production rate has to be positive as well. In that manner, Prigogine recovers well-established results and since then the positivity of the entropy production rate has been used either as a premise or as a goal to be achieved [20–27]. However, the existence of negative entropy production rates has become a topic of intense research mainly because of its supposed relation with non-Markovian aspects of the dynamics of open quantum systems [28–30]. Although a considerable amount of examples has been presented in the last years trying to establish a connection between negative rates and non-Markovianity [31–35], our understanding about it is still improving [36, 37].

This work is organized in the following form: in section 2 we show how to connect LR to the three aspects of the second law mentioned previously and address the non-equilibrium regions where our analysis will be done; after that, in section 3, we review the main elements of LR that are necessary for the development of this work; in sections 4 and 5, we address the above-mentioned compatibility for slowly-varying and finite-time but weak processes. In particular, section 5.3 presents an example where we observe negative values of the entropy production rate. We make our final remarks in section 6.

2. Connecting LR to the second law

According to the second law of thermodynamics, the entropy variation of a total isolated system, after we have changed a control parameter λ from λ_0 to $\lambda_0 + \delta\lambda$ (we will restrict our analysis to a single control parameter) in a time interval τ , is positive or zero,

$$\Delta S_{\text{tot}} \geq 0. \quad (1)$$

To connect LR with the second law, it is convenient to express (1) in a more suitable way. To this end, we restrict ourselves to the situation in which the total system is composed of a system of interest (or simply system) coupled to a heat bath at temperature T . Considering ΔS and ΔS_B respectively as the entropy variation of the system and the heat bath, Q as the average heat received by the system, W as the average work performed on the system by the external agent and ΔF as the variation of free energy between the final and initial equilibrium state of the system, where $F = U - TS$, we have

$$\begin{aligned} \Delta S_{\text{tot}} = \Delta S + \Delta S_B &\Rightarrow \Delta S_{\text{tot}} = \Delta S - Q/T \\ &\Rightarrow T\Delta S_{\text{tot}} = T\Delta S + W - \Delta U \\ &\Rightarrow T\Delta S_{\text{tot}} = W - \Delta F, \end{aligned}$$

where we used the Clausius theorem and the first law of thermodynamics, $\Delta U = Q + W$, in the first and third implication respectively. Defining the irreversible work W_{irr} as

$$W_{\text{irr}} = W - \Delta F, \quad (2)$$

we have

$$T\Delta S_{\text{tot}} = W_{\text{irr}}. \quad (3)$$

Thus, the irreversible work W_{irr} can be used as a measure of entropy production ΔS_{tot} (see for instance [38]).

In the last decades, expressions for W_{irr} have been derived using LR for the purpose of finding optimal finite-time processes [39–44]. The connection between LR and the second law can be established then through these expressions for W_{irr} which are asked to verify the following statements:

1. No dissipation for quasistatic process:

$$\lim_{\tau \rightarrow \infty} W_{\text{irr}}(\tau) = 0, \quad (4)$$

2. Dissipation for finite-time processes:

$$W_{\text{irr}}(\tau) \geq 0, \quad (5)$$

3. Positive entropy production rate:

$$\dot{W}_{\text{irr}}(t) \geq 0. \quad (6)$$

It is worth clarifying that processes in which the system reaches a stationary state with non-vanishing currents for fixed values of the control parameter λ are ruled out of our considerations.

Our analysis will focus on two nonequilibrium regimes represented in figure 1 by regions 1 and 2. To each of them corresponds a different class of nonequilibrium processes. Slowly-varying processes belong to region 1, where the duration of the process τ is large compared to the system's relaxation time τ_R , but the variation $\delta\lambda$ is arbitrary. In such regime, the expression for W_{irr} deduced by means of LR is given by [39–42]

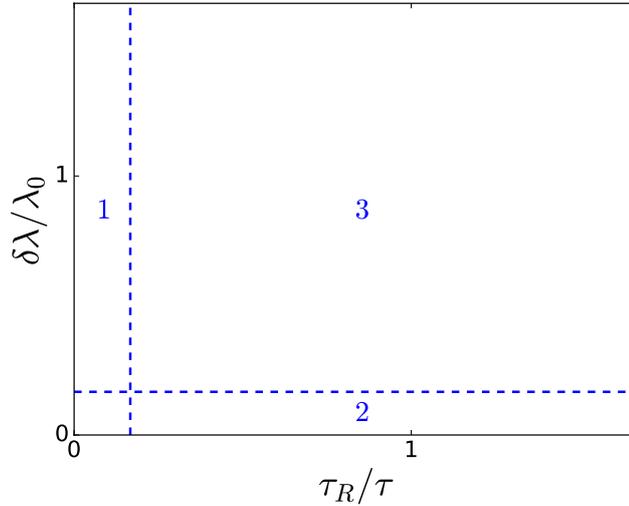


Figure 1. Diagram of nonequilibrium regions. Region 1: slowly-varying processes, Region 2: finite-time but weak processes and Region 3: arbitrarily far from equilibrium processes.

$$W_{\text{irr}} = \beta \int_0^\tau \dot{\lambda}^2(t) \tau_R[\lambda(t)] \chi[\lambda(t)] dt, \quad (7)$$

where $\beta = (k_B T)^{-1}$, k_B is Boltzmann constant and $\dot{\lambda}$ is the time derivative of the protocol $\lambda(t)$. We denote by $\tau_R[\lambda(t)]$ and $\chi[\lambda(t)]$ the parametric variation of the system's relaxation time and of the fluctuations of the generalized force conjugated to λ respectively (in the case of a gas, if the volume is the control parameter then the pressure is the generalized force conjugated to it). We give more precise definitions of these quantities in the following sections (see equations (17) and (18) of section 4).

In region 2, we have finite-time but weak processes. In this regime, the relative change $\delta\lambda/\lambda_0$ must be small but the duration τ of the process can be arbitrary. In this case, using LR, the irreversible work expression reads [43]

$$W_{\text{irr}}(\tau) = \int_0^\tau \int_0^t \Psi_0(t-t') \dot{\lambda}(t') \dot{\lambda}(t) dt' dt, \quad (8)$$

where $\Psi_0(t)$ is the so-called relaxation function [11] (see its definition in equation (14) of section 3). Last but not least, region 3 represents processes arbitrarily far from equilibrium, in which both perturbation and duration of the process are arbitrarily chosen outside near-equilibrium regions, and where LR does not hold anymore. Our goal is to find mathematical and physical conditions under which the functionals (7) and (8) and the quantities appearing in them lead to an agreement with the three aspects of the second law mentioned above. Succinct deductions of such irreversible work expressions can be found in appendix B. For more details, see [42, 43].

It is worth emphasizing that W_{irr} is an average over several microscopic realizations of the protocol $\lambda(t)$, each one furnishing a different value of work \mathcal{W} . Hence, $W_{\text{irr}} \equiv \mathcal{W} - \Delta F$ is a fluctuating quantity that obeys very well-known fluctuation theorems, namely, the Jarzynski equality [45]

$$\langle e^{-\beta(W-\Delta F)} \rangle = \langle e^{-\beta\mathcal{W}_{\text{irr}}} \rangle = 1, \quad (9)$$

and Crooks fluctuation theorem [38]

$$\frac{P_F(\mathcal{W})}{P_R(-\mathcal{W})} = \frac{P_F(\mathcal{W}_{\text{irr}})}{P_R(-\mathcal{W}_{\text{irr}})} = e^{\beta(W-\Delta F)} = e^{\beta\mathcal{W}_{\text{irr}}}, \quad (10)$$

where P_F and P_R denote the usual work distributions obtained after several realizations of the forward (described by $\lambda(t)$) and time-reversed protocols.

We show in section 5.3 by means of stochastic thermodynamics (ST) and LR that, along the process induced by $\lambda(t)$, the instantaneous rate \dot{W}_{irr} can assume *negative* values for the paradigmatic example of driven Brownian motion. This seems to be in contrast to previous definitions of entropy production rates in ST [27]. We will postpone to section 6 additional comments about this point. In any case, we would like to stress that the inequality $\langle \mathcal{W}_{\text{irr}} \rangle = W_{\text{irr}} \geq 0$, derived from the integral fluctuation theorem (9), does not imply $\dot{W}_{\text{irr}} \geq 0$.

3. Hamiltonian formulation of LR

In this section, we review the main elements of LR that we find necessary for our purposes. As mentioned in the previous section, we consider a system coupled to a heat bath at temperature T that is driven out of equilibrium by the switch of a certain control parameter λ . We denote this by

$$\lambda(t) = \lambda_0 + g(t)\delta\lambda, \quad (11)$$

with $g(0) = 0$ and $g(\tau) = 1$, where τ is the duration of the process. In a Hamiltonian approach, this means that the time-dependent Hamiltonian $\mathcal{H}(\lambda(t))$ of system plus heat bath is driven from $\mathcal{H}(\lambda_0)$ to $\mathcal{H}(\lambda_0 + \delta\lambda)$ in a time interval τ by some external agent, according to the protocol $\lambda(t)$ (or, equivalently, $g(t)$). We will focus here on the calculation of the work performed on the system since, as explained in section 2, it will be the necessary link to connect LR to the second law of thermodynamics. Our starting point is the following expression for the work performed on the system,

$$W = \int_0^\tau dt \dot{\lambda}(t) \overline{\partial_\lambda \mathcal{H}(t)}, \quad (12)$$

where $\dot{\lambda} := d\lambda/dt$, \overline{A} denotes the nonequilibrium average of the observable A and the quantity $\overline{\partial_\lambda \mathcal{H}} := \overline{\partial \mathcal{H} / \partial \lambda}$ is the generalized force. We are concerned with the linear response of the generalized force due to the variation of the parameter λ . Therefore, we assume that the system is weakly perturbed, that is, $\delta\lambda/\lambda_0 \ll 1$. Using LR (see appendix A for more details), the expression for the generalized force up to first order in $\delta\lambda$ reads

$$\overline{\partial_\lambda \mathcal{H}(t)} = \langle \partial_\lambda \mathcal{H} \rangle_0 - \delta\lambda \Theta_0 g(t) + \delta\lambda \int_0^t du \Psi_0(u) \left. \frac{dg}{dt'} \right|_{t'=t-u}, \quad (13)$$

where $\langle \dots \rangle_0$ is an equilibrium average taken on the initial canonical distribution, $\exp(-\beta\mathcal{H}(\lambda_0))/Z(\beta, \lambda_0)$ ($Z(\beta, \lambda_0)$ being the partition function) and $\Psi_0(t)$ is the relaxation function, given, in our specific case, by

$$\Psi_0(t) = \beta \langle \partial_\lambda \mathcal{H}(q_0, p_0) \partial_\lambda \mathcal{H}(q_t, p_t) \rangle_0 - \mathcal{C}, \quad (14)$$

after using Kubo formula [11] (see equations (A.9)–(A.11) in appendix A). We denote by (q_t, p_t) a phase-space point of the entire system at the instant t and \mathcal{C} is a constant defined by [11]

$$\mathcal{C} = \beta \lim_{s \rightarrow 0} s \int_0^\infty e^{-st} \langle \partial_\lambda \mathcal{H}(q_0, p_0) \partial_\lambda \mathcal{H}(q_t, p_t) \rangle_0 dt, \quad (15)$$

whose purpose is to guarantee that the system attains the correct new equilibrium state (it is assumed then from the very beginning that the auto-correlation function in equation (14) decays due to the interaction with the heat bath). The purpose of the subscript ‘0’ is to emphasize that the initial equilibrium averages were taken with $\lambda = \lambda_0$. Finally, the constant Θ_0 is defined as (see appendix A)

$$\Theta_0 := \Psi_0(0) - \langle \partial_{\lambda\lambda}^2 \mathcal{H} \rangle_0. \quad (16)$$

To calculate the work performed on the system, it is possible to use equation (13) in two different regimes characterized previously by regions 1 and 2 of figure 1 [39–44] (see appendix B). In both cases, the relaxation function $\Psi_0(t)$ is the central object of the theory and its exact expression demands the solutions of Hamilton’s equations, which is not a very easy task to be accomplished. In order to circumvent this problem, the relaxation function can alternatively be modeled using phenomenological information [11, 46]. Although this way of obtaining $\Psi_0(t)$ is never in full agreement with the underlying Hamiltonian dynamics, it can be made approximately consistent with it. This is achieved through the so-called sum rules, which are constraints that the phenomenological relaxation function must satisfy to match its expected microscopic requirements [11, 46] (see appendix C for more details). For instance, Hamiltonian dynamics demand equation (14) to be even under the change $t \rightarrow -t$. Thus, this constraint must be imposed to any phenomenological expression intended to be used as a valid relaxation function.

As it will be shown in section 5, this parity property will be very important to make the linear-response expression of equation (12) compatible with the second law of thermodynamics. Nevertheless, we want to emphasize that $\Psi_0(-t) = \Psi_0(t)$ holds in the particular case of interest here of nonequilibrium work due the variation of a *single* external parameter. Different parities are indeed possible for more general cases [11, 46].

4. Compatibility for slowly-varying processes

The irreversible work for slowly-varying processes is given by equation (7) which depends on two quantities we have not defined properly yet. The first one is the variance χ of the generalized force defined as

$$\chi(\lambda_0) = \langle (\partial_\lambda \mathcal{H})^2 \rangle_0 - \langle \partial_\lambda \mathcal{H} \rangle_0^2, \quad (17)$$

where the dependence in λ_0 emphasizes that the canonical average is taken over the initial equilibrium distribution. The second quantity is the relaxation time defined in LR by

$$\tau_R(\lambda_0) = \int_0^\infty \frac{\Psi_0(t)}{\Psi_0(0)} dt, \quad (18)$$

where the dependence in λ_0 emphasizes again that averages were taken with the initial equilibrium distribution. Observing equation (7), one concludes that W_{irr} is given in terms of equilibrium quantities that vary parametrically in time through the protocol $\lambda(t)$. This suggests that, in this regime, the system slightly deviates from a sequence of equilibrium states as λ changes in time. In other words, the $\chi[\lambda(t)]$ and $\tau_R[\lambda(t)]$ used in equation (7) are still given by equations (17) and (18) evaluated though at the value $\lambda(t)$ instead of λ_0 . Additionally, it only makes sense to consider relaxation functions leading to positive and finite relaxation times.

4.1. No dissipation for quasistatic processes

Defining $u = t/\tau$, we can rewrite equation (7) as

$$W_{\text{irr}} = \frac{\beta}{\tau} \int_0^1 \dot{\lambda}^2(u) \tau_R[\lambda(u)] \chi[\lambda(u)] du. \quad (19)$$

We note that the dependence on the switching time occurs only in the factor $1/\tau$ [42]. This is so because $\lambda(t) = \lambda_0 + \delta\lambda g(t)$ satisfies the boundary conditions $\lambda(0) = \lambda_0$ and $\lambda(\tau) = \lambda_0 + \delta\lambda$, implying that $g(t)$ is indeed a function of t/τ . In this manner, we have immediately

$$\lim_{\tau \rightarrow \infty} W_{\text{irr}} = \lim_{\tau \rightarrow \infty} \frac{\beta}{\tau} \int_0^1 \dot{\lambda}^2(u) \tau_R[\lambda(u)] \chi[\lambda(u)] du = 0. \quad (20)$$

4.2. Dissipation for finite-time processes

As the integrand of equation (19) is composed of positive functions, the integral leads to a positive result as well. Therefore, for a finite switching time τ , we have

$$W_{\text{irr}} = \frac{\beta}{\tau} \int_0^1 \dot{\lambda}^2(u) \tau_R[\lambda(u)] \chi[\lambda(u)] du > 0. \quad (21)$$

4.3. Positive entropy production rate

Considering again equation (7), the time derivative of W_{irr} is composed of positive functions only. Thus,

$$\dot{W}_{\text{irr}} = \beta \dot{\lambda}^2(t) \tau_R[\lambda(t)] \chi[\lambda(t)] > 0. \quad (22)$$

In addition, by analogy with the thermodynamics of linear irreversible processes [18, 19], we will define some nomenclature that will be useful for our discussion in what follows. Consider the factorization

$$\dot{W}_{\text{irr}} = \dot{\lambda}(t) \left(\beta \tau_R[\lambda(t)] \chi[\lambda(t)] \dot{\lambda}(t) \right) = \dot{\lambda}(t) \mathcal{F}[\lambda(t)]. \quad (23)$$

We call $\dot{\lambda}$ and \mathcal{F} respectively the *affinity* and the associated *flux*. We observe that, in the situation where the process is driven externally, the affinity is related to how the external agent performs the process, namely, $\dot{\lambda}(t)$. In particular, for slowly-varying processes, the flux responds instantaneously to the affinity, which characterizes a *memoryless* process [47].

5. Compatibility for finite-time but weak processes

In the previous section, we have shown that the compatibility of the linear-response expression (7) with the second law is straightforward and the only necessary requirement on the relaxation function is the convergence of the integral (18) defining the relaxation time for all values of λ between λ_0 and $\lambda_0 + \delta\lambda$. Our goal hereafter is to provide further constraints that any relaxation function must fulfill in order to achieve compatibility with the second law in region 2. As a first step, we will consider that the relaxation function $\Psi_0(t)$ has the following parity due to time-reversal symmetry,

$$\Psi_0(t) = \Psi_0(-t). \quad (24)$$

As we will see, such property will be essential to demonstrate the aspects that follow below. This is indeed a property inherited from the Hamiltonian definition of the relaxation function, which is an auto-correlation function calculated in the canonical equilibrium (see equation (14)), and already discussed in section 3.

5.1. No dissipation for quasistatic process

We demonstrate now that equation (8) satisfies the first aspect of compatibility with the second law (see equation (4)) assuming that the system thermalises with the heat bath. In other words, we assume as before that the relaxation time, defined by equation (18), is finite,

$$\int_0^\infty \frac{\Psi_0(t)}{\Psi_0(0)} dt < \infty. \quad (25)$$

We observe first that we can rewrite the relaxation time as

$$\tau_R = \frac{\tilde{\Psi}_0(0)}{\Psi_0(0)}, \quad (26)$$

where $\tilde{\Psi}_0$ is the Laplace transform of Ψ_0 . Thus, equation (25) leads to

$$\begin{aligned} \tilde{\Psi}_0(0) < \infty &\Rightarrow \lim_{s \rightarrow 0} \tilde{\Psi}_0(s) < \infty \\ &\Rightarrow \lim_{s \rightarrow 0} s \tilde{\Psi}_0(s) = 0 \\ &\Rightarrow \lim_{t \rightarrow \infty} \Psi_0(t) = 0, \end{aligned}$$

where in the last line we used the final value theorem. We remark that the reverse implication is not true as shown by the counterexample $\Psi_0(t) = \Psi_0(0)/(1 + |t|)$.

Considering equation (24) and the definitions $u = t/\tau$ and $v = t'/\tau$, we rewrite equation (8) as

$$W_{\text{irr}}(\tau) = \frac{1}{2} \int_0^1 \int_0^1 \Psi_0(\tau(u - v)) \dot{\lambda}(u) \dot{\lambda}(v) du dv, \tag{27}$$

(for more details, see [43]). Therefore, we have

$$\begin{aligned} \lim_{\tau \rightarrow \infty} W_{\text{irr}}(\tau) &\propto \lim_{\tau \rightarrow \infty} \int_0^1 \int_0^1 \Psi_0(\tau(u - v)) \dot{\lambda}(u) \dot{\lambda}(v) du dv \\ &= \int_0^1 \int_0^1 \left(\lim_{\tau \rightarrow \infty} \Psi_0(\tau(u - v)) \right) \dot{\lambda}(u) \dot{\lambda}(v) du dv \\ &= 0, \end{aligned}$$

in which we moved the limit inside the integral assuming that the integrand is well-behaved and used equation (24) again. Figure 2 corroborates our result presenting the irreversible work given by equation (8) for different protocols and relaxation functions (the subscript ‘0’ was dropped for the sake of simplicity of notation)

$$\Psi_1(t) = \Psi_1(0)e^{-a_1|t|}, \tag{28a}$$

$$\Psi_2(t) = \Psi_2(0)e^{-a_2|t|}(\cos(a_2t) + \sin(a_2|t|)), \tag{28b}$$

$$\Psi_3(t) = \Psi_3(0)J_0(a_3t), \tag{28c}$$

which satisfy equations (24) and (25). We remark that J_0 is the Bessel function of the first kind for index $\alpha = 0$ and a_1 , a_2 and a_3 are positive free parameters. The relaxation functions Ψ_1 and Ψ_2 are commonly used to model respectively overdamped and underdamped Brownian motions [11, 42, 44] and Ψ_3 is motivated by correlation functions of spin systems [48].

It is worth emphasizing that there are no free parameters when the relaxation function is exactly obtained from the solutions of the equations of motion. Hence the expressions above represent possible phenomenological models for the behavior of the equilibrium correlation function given in equation (14). The free parameters of such models can be expressed in terms of thermodynamic quantities and parameters of the

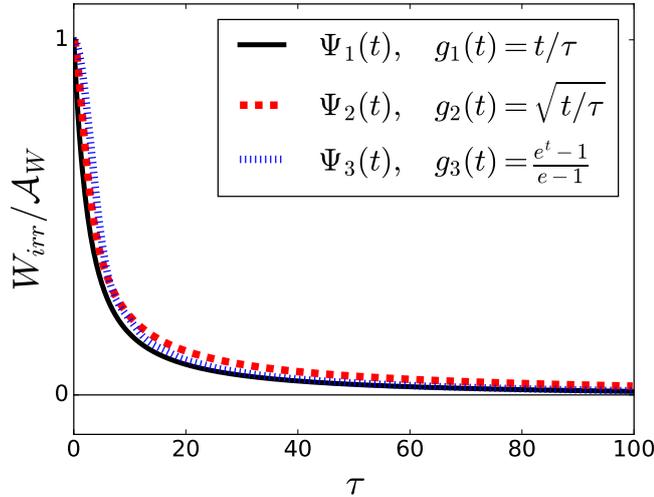


Figure 2. Irreversible work given by equation (8) as a function of switching time τ for the relaxation functions defined in equation (28) and protocols shown in the inset. In all cases we observe that dissipation decreases monotonically as the quasistatic limit is approached, $\tau \rightarrow \infty$. We chose the irreversible work unit $\mathcal{A}_W = \Psi_0(0)(\delta\lambda)^2/2$.

Hamiltonian through the expressions in equation (C.3a) (for examples of how to do this, see [11, 42]).

5.2. Dissipation for finite-time processes

Dissipation for finite-time processes occurs if the following theorem is satisfied.

Theorem 1. Consider that the relaxation function is even with respect to the change $t \rightarrow -t$ due to time-reversal symmetry. The irreversible work given by equation (8) is nonnegative if, and only if, the Fourier transform of the relaxation function is a non-negative function

$$\hat{\Psi}_0(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\omega t} \Psi_0(t) dt \geq 0. \quad (29)$$

We will restrict ourselves to prove the first implication. The other one can be seen in detail in [49] under the name *Bochner's theorem*.

Proof. Firstly, $\hat{\Psi}_0$ is a real function since $\Psi_0(t) = \Psi_0(-t)$. Hence, if $\hat{\Psi}_0$ is a positive function, we have

$$\begin{aligned} W_{\text{irr}}[\dot{g}] &\propto \int_0^1 \int_0^1 \Psi_0(\tau(u-v)) \dot{\lambda}(u) \dot{\lambda}(v) du dv \\ &= \frac{1}{\sqrt{2\pi}} \int_0^1 \int_0^1 \int_{\mathbb{R}} e^{i\omega\tau(u-v)} \hat{\Psi}_0(\omega) \dot{\lambda}(u) \dot{\lambda}(v) d\omega du dv \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left| \int_0^1 e^{i\omega\tau u} \dot{\lambda}(u) du \right|^2 \hat{\Psi}_0(\omega) d\omega \geq 0. \end{aligned}$$

□

We present below the Fourier transform of the relaxation functions defined in equation (28), which satisfy the hypotheses of theorem 1,

$$\hat{\Psi}_1(\omega) = \Psi_1(0) \sqrt{\frac{2}{\pi}} \frac{a_1}{a_1^2 + \omega^2}, \tag{30a}$$

$$\hat{\Psi}_2(\omega) = \Psi_2(0) \sqrt{\frac{2}{\pi}} \frac{4a_2^3}{4a_2^4 + \omega^4}, \tag{30b}$$

$$\hat{\Psi}_3(\omega) = \Psi_3(0) \sqrt{\frac{2}{\pi}} \frac{\theta(\omega/a_3 + 1) - \theta(\omega/a_4 - 1)}{\sqrt{a_3^2 - \omega^2}}, \tag{30c}$$

where θ is the Heaviside step-function. We conclude that these relaxation functions are compatible with aspect (5) the second law of thermodynamics. In summary, phenomenological models of $\Psi_0(t)$ whose Fourier transform does not follow the conditions of theorem 1 can violate aspect (5) and therefore cannot be accepted.

A question that naturally arises is which are the properties on the relaxation function itself (i.e. not on its Fourier transform) that guarantee the non-negativity of the irreversible work. In fact, there is mathematical research currently investigating related issues [50–54].

5.3. Negative values in the entropy production rate

Our last task is to verify the non-negativity of the entropy production rate as described by LR in region 2. Surprisingly, we will show that equation (8) predicts negative values of entropy production rate for a rather simple example. Our results also reveal how such negative values disappear as the process becomes ‘memoryless’, a notion that we clarify later on.

In what follows, we will compare our linear-response results with exact ones provided by ST [55]. This comparison will show a nice agreement between both sets of results. Additionally, it will become clear how the linear-response expression for the entropy production rate clarifies the origin of negative values in contrast to the unappealing numerical solution furnished by ST.

Our example consists on a particle of mass m , immersed in a heat bath of temperature T and subjected to a time-dependent harmonic potential with stiffness $\lambda(t)$. We model the dynamics of this particle through the following Langevin equation,

$$m\ddot{x} + \gamma\dot{x} + \lambda(t)x = f(t), \tag{31}$$

where $x(t)$ is the particle position at time t , $\gamma > 0$ is the dissipation constant and $f(t)$ is a white noise which has the properties

$$\overline{f(t)} = 0, \tag{32a}$$

$$\overline{f(t)f(t')} = 2\gamma k_B T \delta(t - t'), \tag{32b}$$

where $\overline{(\dots)}$ is the average taken over the noise history. We will consider the underdamped regime in which $\gamma/\omega_0 < 2$, with $\omega_0^2 = \lambda_0/m$.

Due to equations (2) and (12), the instantaneous power given by ST reads

$$\dot{W}_{\text{irr}}^{\text{ST}}(t) = \frac{\dot{\lambda}(t)}{2} \left(\overline{x^2}(t) - \frac{dF}{d\lambda}(\lambda(t)) \right), \quad (33)$$

where the term $dF/d\lambda$ is a generalized force that measures the rate of change of the equilibrium free energy along the process. In the underdamped regime, the mean squared displacement $\overline{x^2}$ can be found solving the following system of equations [56]

$$\partial_t \overline{x^2} = 2\overline{x\dot{p}}/m, \quad (34a)$$

$$\partial_t \overline{p^2} = -2\lambda(t)\overline{x\dot{p}} - 2\gamma\overline{p^2}/m + 2\gamma T, \quad (34b)$$

$$\partial_t \overline{x\dot{p}} = \overline{p^2}/m - \lambda(t)\overline{x^2} - \gamma\overline{x\dot{p}}/m. \quad (34c)$$

According to equation (8), the linear-response expression for the instantaneous power in region 2 reads

$$\dot{W}_{\text{irr}}^{\text{LR}}(t) = \dot{\lambda}(t) \left(\int_0^t \Psi_0(t-t') \dot{\lambda}(t') dt' \right), \quad (35)$$

in which, again by analogy with the thermodynamics of linear irreversible processes, the first factor, $\dot{\lambda}(t)$, is the affinity and the second one the associated flux. In contrast to equation (23), the delayed response of the flux with respect to the affinity characterizes a process *with memory*.

To obtain $\dot{W}_{\text{irr}}^{\text{LR}}(t)$, we need an expression for the relaxation function $\Psi_0(t)$. According to its definition, equation (14), an exact expression demands the knowledge of the solutions of Hamilton's equations of particle plus heat bath. As explained in section 3, to circumvent this problem we can use a phenomenological model that is minimally compatible with the Hamiltonian dynamics. In the present case, this can be done as follows: we first obtain $x(t)$ by solving equation (31) for $\lambda(t) = \lambda_0$; then, since $\partial_\lambda \mathcal{H} = x^2/2$ (we are considering a Brownian particle in a harmonic trap), we plug into $\Psi_0(t)$ the following expression

$$\Psi_0(t) = \frac{\beta}{4} \left\langle \overline{x^2}(t) x^2(0) \right\rangle_0 - \mathcal{C}, \quad (36)$$

according to equations (14) and (15). It is clear that the phenomenological correlation function in equation (36) is obtained after taking two averages, namely, one over the noise history and another, a canonical one, over the initial conditions. The final analytical expression for $\Psi_0(t)$ reads

$$\Psi_0(t) = \frac{e^{-\gamma|t|}}{2\beta m^2 \omega_0^4 \omega^2} \left[(2\omega_0^2 - \omega^2) \cos(\omega t) - \gamma\omega \sin(\omega|t|) - 2\omega_0^2 \right], \quad (37)$$

where $\omega^2 = 4\omega_0^2 - \gamma^2$. The absolute value of t was added by hand as a minimal Hamiltonian requirement due to time-reversal symmetry (see the discussion in section 3). We observe that equation (37) leads to a finite relaxation time and a positive

Fourier transform. Therefore, according to what was discussed in section 4.2, it is an acceptable relaxation function to model our system.

The relaxation time obtained from equations (18) and (37) is

$$\tau_R = \frac{1}{2\gamma} + \frac{\gamma}{2\omega_0^2}. \tag{38}$$

It is convenient to express γ and ω_0 in terms of the coupling between the system and heat bath, given by $\eta := \gamma/\omega_0$, and the relaxation time (38)

$$\gamma = \frac{1 + \eta^2}{2\tau_R}, \quad \omega_0 = \frac{1 + \eta^2}{2\eta\tau_R}. \tag{39}$$

In particular, we have set $\tau_R = 1$ in all the following results. The entropy production rates were computed using the protocol (11), with

$$g(t) = \frac{t}{\tau} + \sin\left(\frac{2\pi t}{\tau}\right). \tag{40}$$

Figure 3 shows the comparison between the results obtained using ST and LR, i.e. equations (33) and (35). The system of equation (34) were solved numerically using equation (40). As the quasistatic limit is approached, $\tau \gg \tau_R$, the entropy production rate becomes non-negative. Although $\gamma/\omega_0 = 1$ for this set of results, we have not observed any noteworthy changes in the results for other ratios (see, for example, figure 4). The relative change of the control parameter λ was chosen to be $\delta\lambda/\lambda_0 = 0.1$ and $m = 1$.

The appearance of negative values in the entropy production rate is closely related to the system’s memory in responding to the perturbation and the arbitrary affinity or protocol $\lambda(t)$ that the external agent can choose. We remind that the system’s memory manifests itself as a delayed response of the flux in respect to the affinity of the system. Roughly speaking, the convolution between Ψ_0 and $\dot{\lambda}(t)$ in (35) gives rise to a flux that is out-of-phase with $\dot{\lambda}(t)$. In addition, the affinity can be positive or negative as long as the protocol $\lambda(t)$ is non-monotonic. The combination of these aspects entails a product of terms with different signs in different regions possibly yielding negative rates. To illustrate this, equation (35) can be written as the product of a delayed propagation of the affinity,

$$\text{Prop}_{\Psi_0}[\dot{\lambda}] = \int_0^t \Psi_0(t - t')\dot{\lambda}(t')dt', \tag{41}$$

and an instantaneous one,

$$\text{Prop}_{\delta}[\dot{\lambda}] = \int_0^t \delta(t - t')\dot{\lambda}(t')dt', \tag{42}$$

yealding

$$\dot{W}_{\text{irr}}^{\text{LR}}(t) = \text{Prop}_{\delta}[\dot{\lambda}] \times \text{Prop}_{\Psi_0}[\dot{\lambda}]. \tag{43}$$

Figure 5 shows the functions obtained from the propagations in equations (41) and (42) for a fast and slow realization of protocol (40). The positivity of the entropy

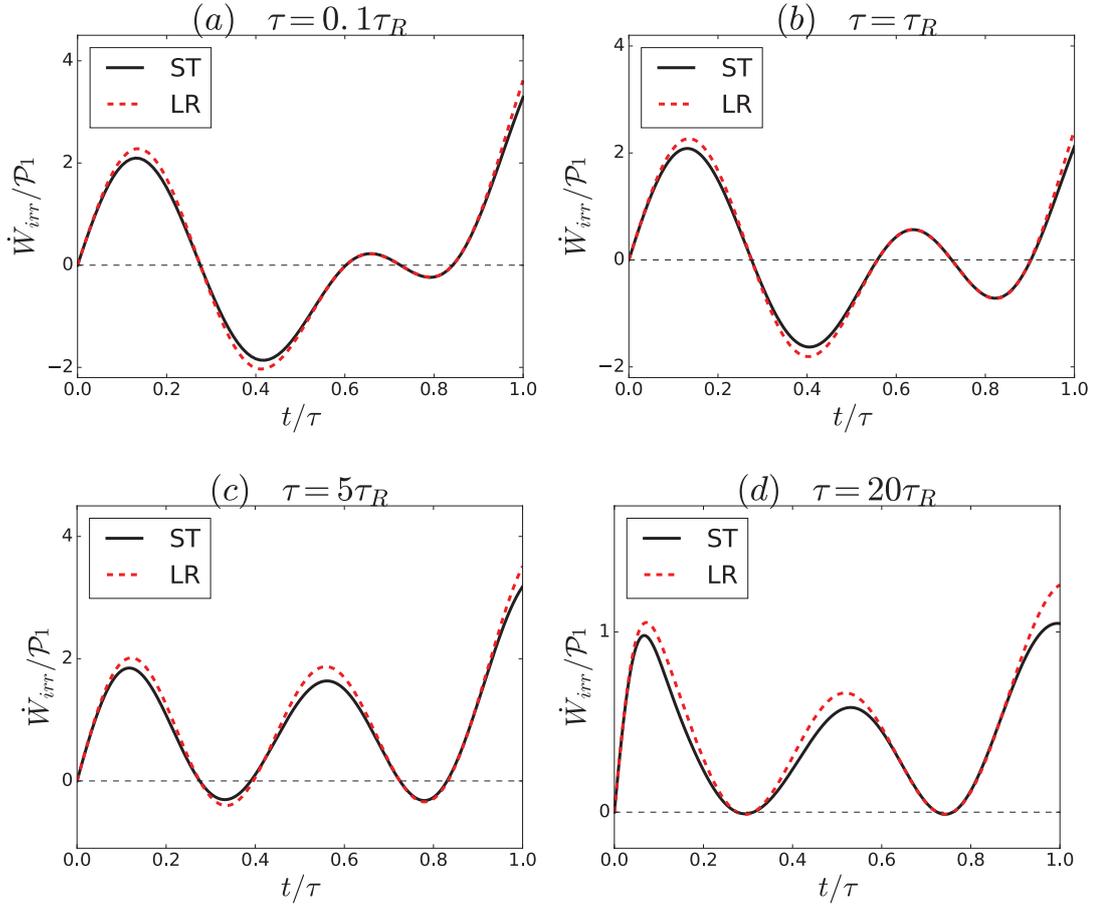


Figure 3. Comparison between the entropy production rate calculated by ST and LR, using respectively equations (33) and (35) with the protocol (40). As illustrated in panels (a)–(d), negative values of the entropy production rate vanish as the process approaches the quasistatic regime, i.e. the ratio τ/τ_R increases. We chose $\tau_R = 1$, $\gamma/\omega_0 = 1$, $\delta\lambda/\lambda_0 = 0.1$ and $m = 1$. We chose also the instantaneous power unit as $\mathcal{P}_1 = k_B T/(100\tau)$.

production rate for the slow process is quite understandable if we remind ourselves of the discussion in section 4. In fact, when $\tau/\tau_R \gg 1$, the system will almost relax completely at each small piece of the process. In other words, the process will be entering the slowly-varying regime (see in figure 1 the intersection of regions 1 and 2). Therefore, the entropy production rate will become a positive function. From a mathematical point of view, such attainment of the quasistatic regime can be understood as the limit in which the relaxation function becomes a Dirac delta function and the response of the associated flux is basically instantaneous (see figure 5(b)).

Figure 6 shows a comparison between the entropy production rates calculated by equations (33) and (35) with $\delta\lambda/\lambda_0 = 0.5$. Although this is a regime in which LR already deviates from the exact result, it still predicts correctly the time intervals in which negative rates exist. Additionally, the order of magnitude and the outline of the entropy production rate also follow reasonably well the exact result. This shows how useful LR predictions can be even in the fully nonequilibrium regime.

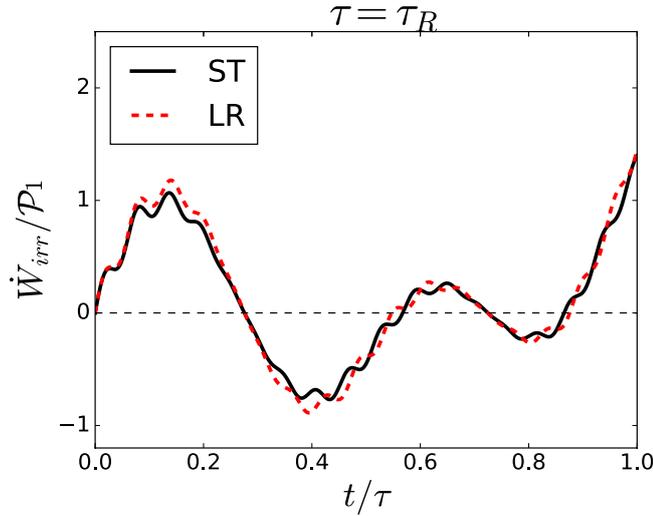


Figure 4. Comparison between the entropy production rate calculated by ST and LR, using respectively equations (33) and (35) with the protocol (40). The weak coupling between the system and heat bath does not affect the emergence of negative entropy production rate, although the outlines of the curves are more oscillatory than those of figure 3. We chose $\tau_R = 1$, $\gamma/\omega_0 = 0.01$, $\delta\lambda/\lambda_0 = 0.1$ and $m = 1$. We chose also the instantaneous power unit as $\mathcal{P}_1 = k_B T/(100\tau)$.

We emphasize that the existence of the negative values observed in the entropy production rate is not an exclusive consequence of the memory of the system. For instance, if we consider a monotonic protocol such as $\lambda(t) = \lambda_0 + \delta\lambda(t/\tau)^2$, whose affinity is $\dot{\lambda} = 2t\delta\lambda/\tau^2$, although the process can be fast the entropy production rate is always positive (see figure 7). As pointed out previously, the form of $\lambda(t)$ is essential in obtaining negative rates. On the other hand, if we consider different systems described by relaxation functions such as $\Psi_1(t)$, $\Psi_3(t)$ and $\Psi_4(t)$ and maintain the protocol (40), we still observe negative values in the entropy production rate (see figure 8). For instance, if we consider Brownian motion in the overdamped regime, expression (36) yields a simple exponential as the phenomenological model for the relaxation function. The instantaneous power (33) is obtained from the solution of the equation below [57]

$$\overline{\partial_t x^2} = -2\lambda(t)\overline{x^2}/\gamma + 2k_B T/\gamma, \quad (44)$$

and the instantaneous power provided by LR in equation (35) can be calculated using $\Psi_1(t)$ [42, 44]. Figure 8 shows a comparison of both results with great agreement. Surprisingly, we still observe negative values in the entropy production rate for $\tau_R = 1/\gamma = 1$ and $\delta\lambda/\lambda_0 = 0.1$.

We would like to stress that the results of figures 3, 4, 6 and 8 should not be misunderstood. The regions with negative values do not lead to violations of the inequality (5) since an integration of \dot{W}_{irr} over an interval of t/τ with negative values do not give the W_{irr} for a process that drives the system between the corresponding values of λ in that interval.

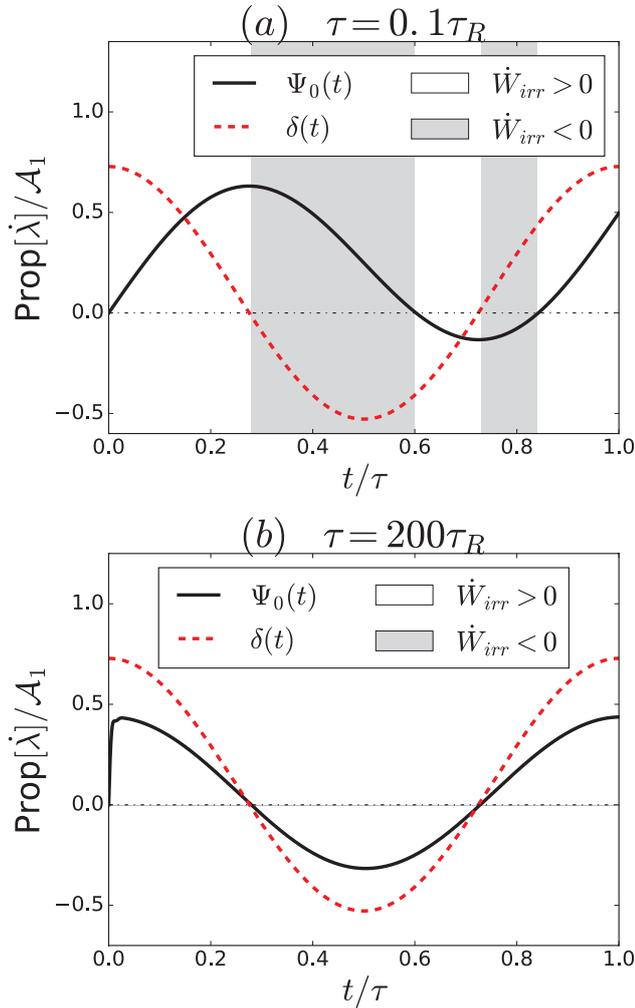


Figure 5. Comparisons between the delayed and instantaneous propagations of $\dot{\lambda}$ via the relaxation function and Dirac delta function (see equations (41) and (42) respectively) for the protocol (40). Panels (a) and (b) illustrate such propagations for $\tau = 0.1\tau_R$ and $\tau = 200\tau_R$. In (a), the negative values in the entropy production rate are a consequence of the system’s memory and the non-monotonicity of the protocol. In (b), the positivity is acquired as long as the response of the flux in respect to the affinity becomes instantaneous. We chose $\tau_R = 1$, $\gamma/\omega_0 = 1$, $\delta\lambda/\lambda_0 = 0.1$ and $m = 1$. We chose also the affinity unit as $\mathcal{A}_1 = \delta\lambda/(10\tau)$. We remark that the delayed propagations $\text{Prop}_{\Psi_0}[\dot{\lambda}]$ were rescaled by factors (a) 100 and (b) 1.2 for a better presentation.

6. Final remarks

Three complementary aspects, often taken as statements of the second law of thermodynamics, were considered in this work in the context of isothermal processes: no dissipation for quasistatic process, dissipation for finite-time processes and positive entropy production rate. We have shown that the linear-response formulation of slowly-varying processes satisfies almost automatically all of them, while for finite-time

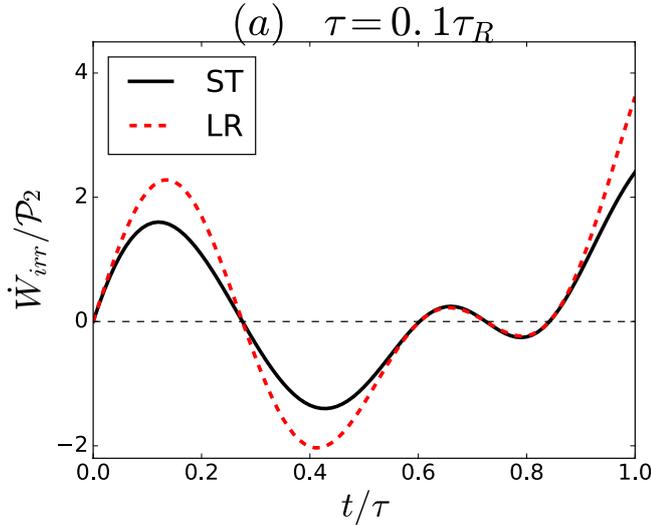


Figure 6. Comparison between the entropy production rate calculated by ST and LR, using equations (33) and (35), respectively, with the protocol (40). We observe that LR result has the same outline and order of magnitude of ST result. We chose $\tau_R = 1$, $\gamma/\omega_0 = 1$, $\delta\lambda/\lambda_0 = 0.5$ and $m = 1$. We chose also the instantaneous power unit as $\mathcal{P}_2 = k_B T/(4\tau)$.

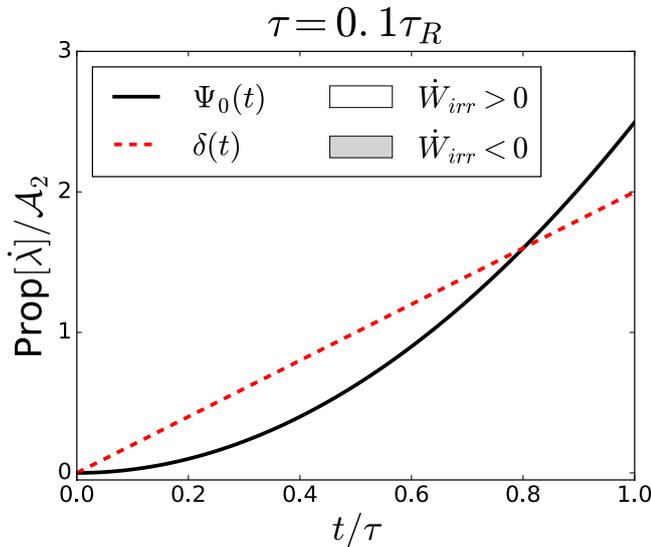


Figure 7. Comparison between the delayed and instantaneous propagations of $\dot{\lambda}$ via the relaxation function, given by equation (37), and Dirac delta function (see equations (41) and (42) respectively). We used the quadratic protocol $g(t) = (t/\tau)^2$. The entropy production rate is positive although the protocol is fast, $\tau = 0.1\tau_R$. We chose $\tau_R = 1$, $\gamma/\omega_0 = 1$, $\delta\lambda/\lambda_0 = 0.1$ and $m = 1$. We chose also the affinity unit as $\mathcal{A}_2 = \delta\lambda/\tau$. We remark that the delayed propagation $\text{Prop}_{\Psi_0}[\dot{\lambda}]$ was rescaled by a factor 50 for a better presentation.

but weak processes more detailed properties of the relaxation function are demanded, namely, finite relaxation time, time-reversal symmetry and positive Fourier transform. Concerning the entropy production rate, we have shown a simple example in which

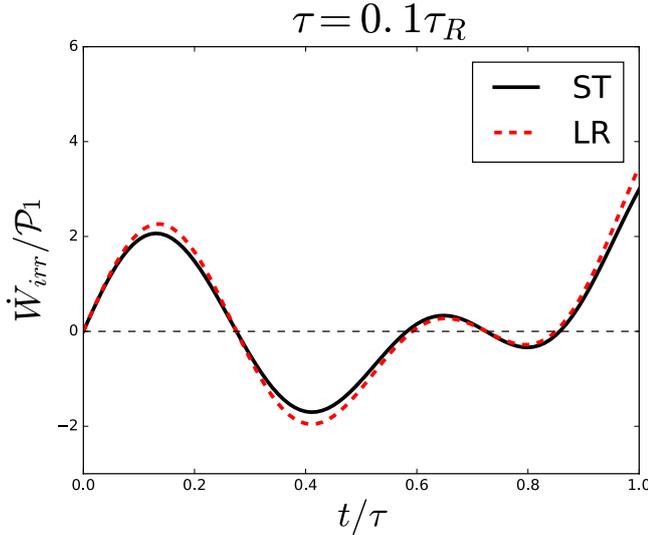


Figure 8. Entropy production rates calculated via ST and LR using protocol (40) for overdamped Brownian motion. For the ST calculation, we used the solution of equation (44) and for LR the relaxation function $\Psi_0(t) = \Psi_0(0)e^{-t/\tau_R}$, where $\Psi_0(0) = \beta\langle x_0^4 \rangle_0/4$. We chose $\tau_R = 1/\gamma = 1$ and $\delta\lambda/\lambda_0 = 0.1$. The instantaneous power unit was chosen as $\mathcal{P}_1 = k_B T/(100\tau)$.

negative values are present. Surprisingly, these results were corroborated by exact calculations. The linear-response formulation has also proven useful in understanding the origin of such phenomenon. The absence of time-scale separation between the performed process and the relaxation of the system and the non-monotonicity of the protocol chosen by the external agent are essential ingredients for the emergence of negative rates according to this formulation. Negative entropy production rates have been observed for the same model of a Brownian particle considered here in [37] although using a different measure of entropy production in terms of a nonequilibrium free energy.

The clear contrast between our results and previous analysis of the entropy production rates for driven Brownian motion deserves further investigation. In [27], the non-negativity of the entropy production rate is obtained after considering the ensemble average over microscopic, or trajectory dependent, entropy production rates. Here, we have verified that the rate of a macroscopically-motivated definition of entropy production namely, the irreversible work, can be negative. A careful analysis of the differences between this two quantities will be done in a future work. Nevertheless, it is interesting to observe that in either case the time integral of the corresponding rates leads to non-negative results.

In addition, we remark that it is quite interesting how Hamiltonian constraints are important to make the connection with the second law of thermodynamics. If, on one hand, the ‘breakdown’ of time-reversal symmetry of microscopic laws in macroscopic irreversible phenomena is often stated as a conundrum, our formulation shows that time-reversal symmetry is necessary in order to achieve compatibility with the second law. Along the same line of thought, we wonder whether the positivity of the Fourier transform of the relaxation function, required by thermodynamic constraints, might be also related to Hamiltonian dynamics. We also wonder how this is related to the fact

that, although there are regions of negative entropy production rates, the integral over the curves in figures 3 and 8 must be positive and hence the maximal amount of negative regions must be already encoded in the relaxation function no matter the protocol we choose.

Finally, the introduction of positive entropy production rate as a statement of the second law was made under the assumption of local equilibrium [18, 19]. From the point of view adopted in this work, this would correspond to the regime of slowly-varying processes whose duration is longer than the relaxation time. Once the ratio between relaxation and switching time increases, we enter the region where local equilibrium does not hold anymore and hence we should not expect only positive rates. We consider this result as a significant step towards the understanding of optimal processes in finite time. As in [57], our results suggest that, if we are constrained to switching times of the order of the relaxation time, a non-monotonic protocol does a better job in minimizing dissipation and entropy production. We leave for future work the extensions of our results to more than one control parameter and to the quantum regime.

Acknowledgments

The authors thank Sebastian Deffner and Philipp Strasberg for very useful discussions. The authors acknowledge support from FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) (Brazil) (Grant No. 2018/06365-4 and No. 2018/21285-7) and support from CNPq (Conselho Nacional de Desenvolvimento Científico e Pesquisa) (Brazil) (Grant No. 141018/2017-8).

Appendix A. Linear-response expression for the generalized force

In this appendix we provide a quick derivation of equation (13) for the generalized force. The general idea of the method is to expand the quantities involved in the calculation in terms of $\delta\lambda$ which is assumed to be small. Consider then the non-equilibrium average of $\partial_\lambda \mathcal{H}$, given by

$$\overline{\partial_\lambda \mathcal{H}(t)} = \int_\Gamma d\Gamma \rho(\Gamma, t) \partial_\lambda \mathcal{H}(\Gamma, \lambda(t)), \quad (\text{A.1})$$

where ρ is the nonequilibrium ensemble of the total system and Γ is a point in the phase space. Firstly, the expansion of the Hamiltonian $\mathcal{H}(\Gamma, \lambda(t))$ in the perturbation $\delta\lambda$ reads

$$\mathcal{H}(\Gamma, \lambda(t)) = \mathcal{H}(\Gamma, \lambda_0) + \partial_\lambda \mathcal{H}(\Gamma, \lambda_0) g(t) \delta\lambda + \mathcal{O}(\delta\lambda^2), \quad (\text{A.2})$$

and, consequently,

$$\partial_\lambda \mathcal{H}(\Gamma, \lambda(t)) = \partial_\lambda \mathcal{H}(\Gamma, \lambda_0) + \partial_{\lambda\lambda}^2 \mathcal{H}(\Gamma, \lambda_0) g(t) \delta\lambda + \mathcal{O}(\delta\lambda^2). \quad (\text{A.3})$$

On the other hand, the non-equilibrium ensemble ρ must satisfy the Liouville equation

$$\partial_t \rho = -\{\rho, \mathcal{H}\} := \mathcal{L}\rho, \quad (\text{A.4})$$

Compatibility of linear-response theory with the second law of thermodynamics and the emergence

where \mathcal{L} is called the Liouville operator. Equation (A.4) can be written in the following integral form,

$$\rho(\Gamma, t) = \rho_{\text{eq}}(\Gamma) + \int_0^t e^{-\mathcal{L}(t-t')} \mathcal{L}\rho(\Gamma, t') dt', \quad (\text{A.5})$$

where $e^{-\mathcal{L}t}$ is the dynamical evolution operator of the ensemble $\rho(\Gamma, 0)$. Using equations (A.2) and (A.5), the non-equilibrium ensemble ρ expanded until its first order in $\delta\lambda$ is given by

$$\rho(\Gamma, t) = \rho_{\text{eq}}(\Gamma) + \int_0^t e^{-\mathcal{L}_0(t-t')} \mathcal{L}_1 \rho_{\text{eq}}(\Gamma) dt' + \mathcal{O}(\delta\lambda^2), \quad (\text{A.6})$$

where we defined the following Liouville operators

$$\mathcal{L}_0(\cdot) := -\{\cdot, \mathcal{H}_0\}, \quad \mathcal{L}_1(\cdot) := -\{\cdot, \partial_\lambda \mathcal{H}(\Gamma, \lambda_0)\} g(t) \delta\lambda. \quad (\text{A.7})$$

Using equations (A.1), (A.3) and (A.6), the anti-Hermiticity property of the Liouville operators and restraining us to the first order in $\delta\lambda$, we obtain

$$\begin{aligned} \overline{\partial_\lambda \mathcal{H}(t)} &= \langle \partial_\lambda \mathcal{H}(\Gamma, \lambda_0) \rangle_0 + \langle \partial_{\lambda\lambda}^2 \mathcal{H}(\Gamma, \lambda_0) \rangle_0 g(t) \delta\lambda \\ &+ \delta\lambda \int_0^t dt' \Phi_0(t-t') g(t'), \end{aligned} \quad (\text{A.8})$$

where $\langle A \rangle_0$ denotes the average of the observable A taken with ρ_{eq} ,

$$\Phi_0(t) = \langle \{\partial_\lambda \mathcal{H}(\Gamma, \lambda_0), \partial_\lambda \mathcal{H}(\Gamma_t, \lambda_0)\} \rangle_0 \quad (\text{A.9})$$

is the so-called response function and Γ_t is the phase-space point evolved up to time t . Defining the relaxation function $\Psi_0(t)$ as

$$\Psi_0(t) := - \int \Phi_0(t) dt - \mathcal{C}, \quad (\text{A.10})$$

and using Kubo formula [11] for the canonical distribution $\rho_{\text{eq}}(\Gamma) = \exp(-\beta\mathcal{H}(\Gamma))/Z(\beta)$,

$$\Phi_0(t) = -\beta \frac{d}{dt} \langle \partial_\lambda \mathcal{H}(0) \partial_\lambda \mathcal{H}(t) \rangle_0, \quad (\text{A.11})$$

one obtains equation (14) in section 3. Besides, after an integration by parts in equation (A.8), we obtain

$$\overline{\partial_\lambda \mathcal{H}(t)} = \langle \partial_\lambda \mathcal{H} \rangle_0 - \delta\lambda \Theta_0 g(t) + \delta\lambda \int_0^t dt' \Psi_0(t') \dot{g}(t-t') + \mathcal{O}(\delta\lambda^2), \quad (\text{A.12})$$

where

$$\Theta_0 := \Psi_0(0) - \langle \partial_{\lambda\lambda}^2 \mathcal{H}(\Gamma, \lambda_0) \rangle_0. \quad (\text{A.13})$$

We refer to [11, 42, 43] for more details.

Appendix B. Linear-response expressions for the irreversible work

We shall start with the irreversible work for slowly varying processes, given by

$$W_{\text{irr}} = \beta \int_0^\tau \dot{\lambda}^2(t) \tau_R[\lambda(t)] \chi[\lambda(t)] dt. \quad (\text{B.1})$$

We split now the processes of duration τ in N time steps of length $\delta t := \tau/N$. Along the n th time step, the control parameter λ , whose change is described by the protocol $\lambda_n(t) = \lambda_n + \delta\lambda_n g_n(t)$, varies only $\delta\lambda_n$, which is considered to be small enough so that we can use LR to describe the generalized force. In particular, the generalized force $\overline{\partial_\lambda \mathcal{H}_n(t)}$ for the n th time step is

$$\overline{\partial_\lambda \mathcal{H}_n(t)} = \langle \partial_\lambda \mathcal{H} \rangle_n - \delta\lambda_n \Theta_n g_n(t) + \delta\lambda_n \dot{g}_n \Psi_n(0) \tau_R(\lambda_n), \quad (\text{B.2})$$

where the index n indicates that all quantities involved are calculated in the n th time step and the equilibrium averages are taken with $\lambda = \lambda_n$ (we are using here the same notation introduced in appendix A for the averages \overline{A} and $\langle A \rangle$ of a given observable A). Expression (B.2) is obtained from equation (A.12) considering that, in the interval δt , the protocol g_n can be taken as linear (which implies that \dot{g}_n is constant but dependent on the n th step), and that the system relaxes faster than δt , so that we can use the definition of the relaxation time, given by equation (18) of section 4, in equation (A.12). We denote by $\Psi_n(0)$ the amplitude of the relaxation function at the beginning of the n th time step. In other words, the argument ‘0’ in $\Psi_n(0)$ refers to the instant of time in which the n th step begins. The constant Θ_n is given by equation (A.13) with λ_n replacing λ_0 .

Finally, the work δW_n performed on the system in the interval δt during the n th steps reads

$$\begin{aligned} \delta W_n = & \langle \partial_\lambda \mathcal{H} \rangle_n \dot{g}_n \delta\lambda_n \delta t - \Theta_n \dot{g}_n g_n(t) (\delta\lambda_n)^2 \delta t \\ & + (\delta\lambda_n)^2 \dot{g}_n^2 \Psi_n(0) \tau_R(\lambda_n) \delta t. \end{aligned} \quad (\text{B.3})$$

It can be shown [42] that the first two terms of the previous expression lead to the free-energy change. Therefore, the irreversible contribution comes from the third term, which has already the form of the integrand in equation (B.1). Considering that $\Psi_n(0) = \beta \chi(\lambda_n)$, where χ is the variance of the $\partial_\lambda \mathcal{H}$ evaluated at $\lambda = \lambda_n$, and summing up all the irreversible contributions of each n th step, we obtain in the continuum limit,

$$W_{\text{irr}} = \beta \int_0^\tau \dot{\lambda}^2(t) \tau_R[\lambda(t)] \chi[\lambda(t)] dt, \quad (\text{B.4})$$

where we consider the n th dependence of the quantities as a instantaneous dependence on the parameter $\lambda(t)$.

The irreversible work expression for finite-time but weak processes comes up from the very definition of work,

$$W(\tau) = \int_0^\tau dt \dot{\lambda}(t) \overline{\partial_\lambda \mathcal{H}(t)}, \quad (\text{B.5})$$

Compatibility of linear-response theory with the second law of thermodynamics and the emergence and the linear-response expression we have already derived in appendix A for the generalized force,

$$\overline{\partial_\lambda \mathcal{H}}(t) = \langle \partial_\lambda \mathcal{H} \rangle_0 - \delta\lambda \Theta_0 g(t) + \delta\lambda \int_0^t dt' \Psi_0(t') \dot{g}(t-t') + \mathcal{O}(\delta\lambda^2). \quad (\text{B.6})$$

Plugging the third term of the previous expression into (B.5), we obtain the irreversible contribution given by equation (8). For more details, see [43].

Appendix C. Hamiltonian constraints on response functions

As mentioned in section 3, Hamiltonian dynamics imposes certain constraints on phenomenological expressions for the response functions. To exemplify what we mean by this, consider the response function $\Phi_0(t)$ (see appendix A for more details and definitions),

$$\Phi_0(t) = \langle \{ \partial_\lambda \mathcal{H}(0), \partial_\lambda \mathcal{H}(t) \} \rangle_0 = - \frac{d\Psi_0(t)}{dt}, \quad (\text{C.1})$$

where $\{.,.\}$ denotes the Poisson bracket and $\langle A \rangle_0$ denotes an equilibrium average of observable A . The short-time expansion of $\Phi_0(t)$ reads

$$\Phi_0(t) = \Phi_0^{(0)}(0) + \Phi_0^{(1)}(0)t + \Phi_0^{(2)}(0)\frac{t^2}{2} + \dots, \quad (\text{C.2})$$

where the coefficients are given by [11, 46]

$$\Phi_0^{(0)}(0) = \langle \{ \partial_\lambda \mathcal{H}(0), \partial_\lambda \mathcal{H}(0) \} \rangle_0 = 0, \quad (\text{C.3a})$$

$$\Phi_0^{(1)}(0) = \langle \{ \partial_\lambda \mathcal{H}(0), \{ \partial_\lambda \mathcal{H}(0), \mathcal{H} \} \} \rangle_0, \quad (\text{C.3b})$$

$$\Phi_0^{(2)}(0) = \langle \{ \partial_\lambda \mathcal{H}(0), \{ \{ \partial_\lambda \mathcal{H}(0), \mathcal{H} \}, \mathcal{H} \} \} \rangle_0 = 0. \quad (\text{C.3c})$$

Indeed, the Hamiltonian dynamics demand $\Phi_0(t)$ to be *odd* with respect to the change $t \rightarrow -t$ and therefore all $\Phi_0^{(k)}(0)$ with k even are zero. Consequently, due to equation (C.1), any phenomenological model of the relaxation function $\Psi_0(t)$ must be *even* under the change $t \rightarrow -t$. Each of the equation (C.3a) are examples of Hamiltonian constraints leading to different sum-rules. These constraints must be imposed to any expression intended to be used as a valid relaxation or response function. Since there is a infinite hierarchy of them, a phenomenological expression with a finite number of free parameters can only fulfill a finite amount of them (for more details of this method, see [46]).

References

- [1] Mori H 1956 A quantum-statistical theory of transport processes *J. Phys. Soc. Japan* **11** 1029–44
- [2] Nakano H 1956 A method of calculation of electrical conductivity *Prog. Theor. Phys.* **15** 77–9
- [3] Debye P J W 1929 *Polare Molekeln* (Leipzig)
- [4] Fröhlich H 1949 *Theory of Dielectrics: Dielectric Constant and Dielectric Loss* (Oxford: Oxford University Press)

- [5] Böttcher C J F 1973 *Theory of Electric Polarization* (Amsterdam: Elsevier)
- [6] Böttcher C J F and Bordewijk P 1978 *Theory of Electric Polarization* vol 2 (Amsterdam: Elsevier)
- [7] Bloch F 1946 Nuclear induction *Phys. Rev.* **70** 460
- [8] Abragam A 1961 *The Principles of Nuclear Magnetism* (Oxford: Clarendon) p 32
- [9] Slichter C P 1978 *Principles of Magnetic Resonance* (Berlin: Springer) pp 1–10
- [10] Kubo R and Hashitsume N 1970 Brownian motion of spins *Prog. Theor. Phys. Suppl.* **46** 210–20
- [11] Kubo R, Toda M and Hashitsume N 1985 *Statistical Physics II: Nonequilibrium Statistical Mechanics* vol 31 (Berlin: Springer)
- [12] Mannella R, West B J, Bianucci M and Grigolini P 1995 From dynamics to thermodynamics: linear response and statistical mechanics *Phys. Rev. E* **51** 3002
- [13] Kemble E C 1939 Fluctuations, thermodynamic equilibrium and entropy *Phys. Rev.* **56** 1013
- [14] Ruelle D 1993 *Chance and Chaos* vol 11 (Princeton, NJ: Princeton University Press)
- [15] Lebowitz J L 1993 Macroscopic laws, microscopic dynamics, time's arrow and boltzmann's entropy *Physica A* **194** 1–27
- [16] Penrose O 2005 *Foundations of Statistical Mechanics: a Deductive Treatment* (New York: Courier Corporation)
- [17] Lebowitz J L 2007 From time-symmetric microscopic dynamics to time-asymmetric macroscopic behavior: an overview *Boltzmann's Legacy* ed W L Reiter *et al* (Zürich: European Mathematical Society) pp 63–88
- [18] De Groot S R and Mazur P 2013 *Non-Equilibrium Thermodynamics* (New York: Courier Corporation)
- [19] Prigogine I 1967 *Introduction to Thermodynamics of Irreversible Processes* (New York: Interscience)
- [20] Schnakenberg J 1976 Network theory of microscopic and macroscopic behavior of master equation systems *Rev. Mod. Phys.* **48** 571
- [21] McAdory R T Jr and Schieve W C 1977 On entropy production in a stochastic model of open systems *J. Chem. Phys.* **67** 1899–903
- [22] Spohn H 1978 Entropy production for quantum dynamical semigroups *J. Math. Phys.* **19** 1227–30
- [23] Spohn H and Lebowitz J L 1978 Irreversible thermodynamics for quantum systems weakly coupled to thermal reservoirs *Adv. Chem. Phys.* **38** 109–42
- [24] Alicki R 1979 The quantum open system as a model of the heat engine *J. Phys. A: Math. Gen.* **12** L103
- [25] Gallavotti G and Cohen E G D 1995 Dynamical ensembles in nonequilibrium statistical mechanics *Phys. Rev. Lett.* **74** 2694
- [26] Ruelle D 1996 Positivity of entropy production in nonequilibrium statistical mechanics *J. Stat. Phys.* **85** 1–23
- [27] Seifert U 2005 Entropy production along a stochastic trajectory and an integral fluctuation theorem *Phys. Rev. Lett.* **95** 040602
- [28] Rivas A, Huelga S F and Plenio M B 2014 Quantum non-Markovianity: characterization, quantification and detection *Rep. Prog. Phys.* **77** 094001
- [29] Laine E-M, Piilo J, Breuer H-P and Vacchini B 2016 Colloquium: non-Markovian dynamics in open quantum systems *Rev. Mod. Phys.* **88** 021002
- [30] De Vega I and Alonso D 2017 Dynamics of non-Markovian open quantum systems *Rev. Mod. Phys.* **89** 015001
- [31] Tukiainen M, Chruscinski D, Piilo J, Bylicka B and Maniscalco S 2016 Thermodynamic power of non-Markovianity *Sci. Rep.* **6** 27989
- [32] Bhattacharya S, Misra A, Mukhopadhyay C and Pati A K 2017 Exact master equation for a spin interacting with a spin bath: non-markovianity and negative entropy production rate *Phys. Rev. A* **95** 012122
- [33] Alipour S, Benatti F, Marcantoni S and Rezakhani A T 2017 Entropy production and non-Markovian dynamical maps *Sci. Rep.* **7** 12447
- [34] Vacchini B, Campbell S and Popovic M 2018 Entropy production and correlations in a controlled non-Markovian setting *Phys. Rev. A* **98** 012130
- [35] Xu Y Y, Liu J and Feng M 2018 Positive entropy production rate induced by non-markovianity *Phys. Rev. E* **98** 032102
- [36] Bhattacharya B, Bhattacharya S and Majumdar A S 2019 Thermodynamic advantages of non-Markovianity from the perspective of resource interconversion (arXiv:1902.05864)
- [37] Strasberg P and Esposito M 2019 Non-markovianity and negative entropy production rates *Phys. Rev. E* **99** 012120
- [38] Crooks G E 1999 Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences *Phys. Rev. E* **60** 2721
- [39] Tsao L-W, Sheu S-Y and Mou C-Y 1994 Absolute entropy of simple point charge model water by adiabatic switching processes *J. Chem. Phys.* **101** 2302–8

- [40] de Koning M and Antonelli A 1997 Adiabatic switching applied to realistic crystalline solids: vacancy-formation free energy in copper *Phys. Rev. B* **55** 735
- [41] Sivak D A and Crooks G E 2012 Thermodynamic metrics and optimal paths *Phys. Rev. Lett.* **108** 190602
- [42] Bonança M V S and Deffner S 2014 Optimal driving of isothermal processes close to equilibrium *J. Chem. Phys.* **140** 244119
- [43] Acconcia T V and Bonança M V S 2015 Degenerate optimal paths in thermally isolated systems *Phys. Rev. E* **91** 042141
- [44] Bonança M V S and Deffner S 2018 Minimal dissipation in processes far from equilibrium *Phys. Rev. E* **98** 042103
- [45] Jarzynski C 1997 Nonequilibrium equality for free energy differences *Phys. Rev. Lett.* **78** 2690
- [46] Kubo R and Ichimura M 1972 Kramers–Kronig relations and sum rules *J. Math. Phys.* **13** 1454–61
- [47] Callen H B 1998 *Thermodynamics and an Introduction to Thermostatistics* (New York: Wiley)
- [48] Glauber R J 1963 Time-dependent statistics of the ising model *J. Math. Phys.* **4** 294–307
- [49] Feller W 2008 *An Introduction to Probability Theory and Its Applications* vol 2 (New York: Wiley)
- [50] Tuck E O 2006 On positivity of fourier transforms *Bull. Aust. Math. Soc.* **74** 133–8
- [51] Giraud B G and Peschanski R B 2006 On Positive functions with positive Fourier transforms *Acta Phys. Polon. B* **37** 331–46
- [52] Giraud B G and Peschanski R B 2014 On the positivity of fourier transforms (arXiv:1405.3155)
- [53] Giraud B G and Peschanski R B From ‘dirac combs’ to fourier-positivity *Acta Phys. Polon. B* **47**
- [54] Giraud B G and Peschanski R B 2016 Fourier-positivity constraints on QCD dipole models *Phys. Lett. B* **760** 26–30
- [55] Seifert U 2012 Stochastic thermodynamics, fluctuation theorems and molecular machines *Rep. Prog. Phys.* **75** 126001
- [56] Schmiedl T, Gomez-Marin A and Seifert U 2008 Optimal protocols for minimal work processes in under-damped stochastic thermodynamics *J. Chem. Phys.* **129** 024114
- [57] Schmiedl T and Seifert U 2007 Optimal finite-time processes in stochastic thermodynamics *Phys. Rev. Lett.* **98** 108301