Extended heat-fluctuation theorems for a system with deterministic and stochastic forces

R. van Zon and E. G. D. Cohen
The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA
(Received 14 November 2003; published 28 May 2004)

Heat fluctuations over a time τ in a nonequilibrium stationary state and in a transient state are studied for a simple system with deterministic and stochastic components: a Brownian particle dragged through a fluid by a harmonic potential which is moved with constant velocity. Using a Langevin equation, we find the exact Fourier transform of the distribution of these fluctuations for all τ. By a saddle-point method we obtain analytical results for the inverse Fourier transform, which, for not too small τ, agree very well with numerical results from a sampling method as well as from the fast Fourier transform algorithm. Due to the interaction of the deterministic part of the motion of the particle in the mechanical potential with the stochastic part of the motion caused by the fluid, the conventional heat fluctuation theorem is, for *infinite* and for *finite* τ, replaced by an extended fluctuation theorem that differs noticeably and measurably from it. In particular, for large fluctuations, the ratio of the probability for absorption of heat (by the particle from the fluid) to the probability to supply heat (by the particle to the fluid) is much larger here than in the conventional fluctuation theorem.

DOI: 10.1103/PhysRevE.69.056121 PACS number(s): 05.40.−a, 05.70.−a, 44.05.+e, 02.50.−r

I. INTRODUCTION

Knowledge of the behavior of heat fluctuations has an intrinsic value, especially for small systems such as nanosystems and biomolecules, where the fluctuations are relatively large, but the widespread interest in fluctuation theorems (FTs) stems mostly from the fact that although there are few general results in nonequilibrium statistical mechanics, these theorems seem to provide some. The conventional FTs state that the probability distribution function (PDF) π for the average over a time τ of a physical quantity A to have a value a, satisfies [1–6]

$$\frac{\pi(\langle A\rangle_\tau = a;\tau)}{\pi(\langle A\rangle_\tau = -a;\tau)} \approx \exp[a\tau], \tag{1}$$

where $\langle A\rangle_\tau$ denotes the time average of A and the dependence of the PDF on τ has been explicitly indicated. The asymmetry in Eq. (1) is caused by some external field that is present in these models, such that a positive value of $\langle A\rangle_\tau$ can correspond to a behavior "with the field" and a negative value to a behavior "against the field." Note that for large τ, positive values for a become much more probable than negative ones. To be more concrete, in dynamical systems theory, the quantity A pertains to fluctuations in the "phase space contraction rate" [3], while in stochastic systems, A involves an "action functional" [6]. In specific cases, both have been connected with the heat or entropy production. Such a connection is based on the expression of the phase space contraction rate and the action functional, respectively, which typically have the form of a thermodynamic force times a current, divided by the temperature of the system [7], i.e., the same form as the entropy production in irreversible thermodynamics. It is, however, not necessary to make the connection with the entropy production and in this paper, we will simply use the heat.

There are in fact two different kinds of FTs: a transient (TFT) and a stationary state FT (SSFT). While the conventional SSFT only holds for sufficiently large (strictly infinite) times, the conventional TFT holds as an identity for all times [8]. Thus, for the SSFT, the $\approx$ sign in Eq. (1) indicates the large τ behavior, whereas for the TFT, it can be replaced by an equality sign [9].

Apart from an (early) laboratory experiment on the SSFT [10], the FTs were restricted to theoretical and simulation approaches: it was difficult to make laboratory experiments on macroscopic systems, since the large number of particles reduces all fluctuations enormously. Recently, Wang *et al.* [11] measured a TFT in the laboratory, by studying the motion of a *single* Brownian particle dragged through water by a laser-induced moving (confining) potential. But while in Ref. [11] the entropy production (or heat) fluctuations over a time τ were intended to be studied in a transient state, in fact the fluctuations in the work done on the system during that time were studied. These differ from the heat fluctuations due to the joint presence of the confining potential and the water, as we will explain more clearly later in the paper. While for this system the conventional SSFT and TFT do hold for the total *work* done on the system [12], for *heat* fluctuations, very different FTs hold, in which the behavior of large heat deviations is notably and measurably different from the conventional FTs, as shown in a recent paper [13]. Here we present the full theory underlying the conclusions of Ref. [13].

The paper is organized as follows. Section II contains the basic theory. We present the model (Sec. II A), explain the difference between work and heat in the model (Sec. II B), and treat the work-related SSFT and TFT fully (Sec. II C). For the heat fluctuations there is no expression for the PDF in terms of known functions, but its Fourier transform can be calculated exactly (Sec. II D). Since this has no known exact inverse, we first calculate, in Sec. III, the PDF of the fluctuations using two numerical methods: a sampling method as well as the algorithm of the fast Fourier transform to perform the inversion of the Fourier transform. The results of both methods agree well with each other and point to violations of the conventional SSFT and TFT. Next, this is substantiated

by an analytic method developed in Sec. IV based on the saddle-point method, which is shown by a comparison with the numerical methods, to work well when τ is not too small. In Sec. V we show how this method can be used to obtain expressions for the extended heat FTs, both in the limit of $\tau \to \infty$, as well as for finite times. We conclude with a discussion of the results in Sec. VI. At the end, two appendices give some technical details used in the text.

II. BASIC THEORY

A. Model

Theoretical descriptions of the experiment of Wang *et al.* [11] have been given even before the experiment by Mazonka and Jarzynski [14], as well as recently by the authors [12,13]. All were based on an overdamped Langevin equation for the Brownian particle, which reads

$$\dot{\mathbf{x}}_t = -\tau_r^{-1}(\mathbf{x}_t - \mathbf{x}_t^*) + \alpha^{-1}\zeta_t. \tag{2}$$

Here $\mathbf{x}_t$ is the (three-dimensional) position of the Brownian particle at time t, $\alpha = 6\pi\eta R$ is the (Stokes) friction of the particle in water, η is the shear viscosity of water, and R is the radius of the particle. Furthermore, the relaxation time of the position of the particle in Eq. (1), τ_r, equals α/k, where k is the strength of the harmonic force. This force is derived from the potential

$$U(\mathbf{x},t) = \frac{k}{2}|\mathbf{x} - \mathbf{x}_t^*|^2, \tag{3}$$

where $\mathbf{x}_t^*$ is the position of the minimum of the harmonic potential at time t, given in our case by

$$\mathbf{x}_t^* = \mathbf{v}^* t \tag{4}$$

(for $t > 0$) with $\mathbf{v}^*$ the (constant) velocity of the motion of the potential through the water. Finally, the random force ζ_t is taken to be Gaussian and δ function correlated in time: $\langle \zeta_t \rangle = 0$ and $\langle \zeta_t \zeta_s \rangle = 2k_B T \alpha \delta(s-t)\mathbb{1}$. Here, the brackets denote an average over realizations, T is the temperature, k_B is Boltzmann's constant, and $\mathbb{1}$ denotes the 3×3 identity matrix.

In the rest of this paper, we will use a reduced set of units. For our time unit we will use τ_r: thus wherever t appears one should read t/τ_r; for our energy unit, we choose $k_B T$ and for our length unit, we use the width of the equilibrium PDF of $\mathbf{x}_t$ in the potential, i.e., $\sqrt{k_B T/k}$. This change of units has been carried through consistently in the paper by simply setting $\alpha = 1$, $k = 1$, and $k_B T = 1$. Then the Langevin equation above in Eq. (2) takes the simple form

$$\dot{\mathbf{x}}_t = -(\mathbf{x}_t - \mathbf{v}^* t) + \zeta_t, \tag{5}$$

where ζ_t satisfies

$$\langle \zeta_t \rangle = 0, \tag{6}$$

$$\langle \zeta_t \zeta_s \rangle = 2\delta(t-s)\mathbb{1}. \tag{7}$$

In Ref. [12], the distribution of $\mathbf{x}_t$ and its time autocorrelation function were determined, under the assumption that the particle has been in the potential from time $t = -\infty$ up to $t = 0$ and that during that time the potential did not move, so that the initial PDF at time $t = 0$ is the equilibrium one. We found for the PDF of the position $\mathbf{x}$ of the Brownian particle at time t,

$$\rho(\mathbf{x};t) = (2\pi)^{-3/2} e^{-|\mathbf{x} - \mathbf{v}^*(t - 1 + e^{-t})|^2/2}. \tag{8}$$

The resulting average position $\langle \mathbf{x}_t \rangle$ is therefore

$$\langle \mathbf{x}_t \rangle = \mathbf{v}^*(t - 1 + e^{-t}). \tag{9}$$

Equations (8) and (9) show that the system is not in equilibrium, but that, after a transient time period of order 1 (after which one may neglect the term e^{-t}) the system becomes stationary in a comoving frame—the PDF shifts to the right with a constant velocity $\mathbf{v}^*$. The mean position of the particle then becomes $\mathbf{v}^*(t-1)$, which is not equal to the position of the minimum of the potential, $\mathbf{x}_t^* = \mathbf{v}^* t$. The difference $-\mathbf{v}^*$ shows we are in a stationary state where the average harmonic force $-(\langle \mathbf{x}_t \rangle - \mathbf{v}^* t) = \mathbf{v}^*$ balances the friction force $-\mathbf{v}^*$.

Below, we will need the correlations of the position at different times. For the time-autocorrelation function of $\mathbf{x}_t$, Ref. [12] gives in current units

$$\langle \mathbf{x}_{t_2} \mathbf{x}_{t_1} \rangle - \langle \mathbf{x}_{t_2} \rangle \langle \mathbf{x}_{t_1} \rangle = e^{-|t_2 - t_1|}\mathbb{1}. \tag{10}$$

B. Work versus heat fluctuations

We first consider the fluctuations in the total work done on the system. This work is the total amount of energy put into the system. By the system, we mean here the particle in the potential U plus the water. At any time, the harmonic potential exerts a force $\mathbf{F}_t = -(\mathbf{x}_t - \mathbf{v}^* t)$ on the particle. Consequently, the particle exerts a reaction force $-\mathbf{F}_t$ on the potential. The potential has to be kept moving at a fixed velocity $\mathbf{v}^*$, for which an external force on it of magnitude $\mathbf{F}_t$ is required to balance the reaction force. Hence, the fluctuating work W_τ done on the system over a time interval from t to $t+\tau$ is given by

$$W_\tau = \int_t^{t+\tau} dt' \mathbf{v}^* \cdot \mathbf{F}_{t'} = \int_t^{t+\tau} dt' \mathbf{v}^* \cdot [-(\mathbf{x}_{t'} - \mathbf{v}^* t')]. \tag{11}$$

We next consider the heat Q_τ and its fluctuations, which in contrast to the total work W_τ, is only that part of the work which goes into the fluid, while the other part of the total work goes into the potential energy of the particle in the harmonic potential. Therefore,

$$Q_\tau = W_\tau - \Delta U_\tau, \tag{12}$$

where

$$\Delta U_\tau = U(\mathbf{x}_{t+\tau}, t+\tau) - U(\mathbf{x}_t, t). \tag{13}$$

The potential energy U was given in Eq. (3) [15]. Note that in Eqs. (11)–(13) we suppressed the t dependence.

Before we discuss the fluctuations in the work and the heat, we want to say a few words on their average behavior. Using Eqs. (8) and (11), we find that the average work in a time interval from t to $t+\tau$ is given by

$$\langle W_\tau \rangle = w[\tau - (1 - e^{-\tau})e^{-t}], \tag{14}$$

where

$$w \equiv |\mathbf{v}^*|^2. \tag{15}$$

From Eq. (14), one sees that after a transient time t of $O(1)$, the second term within the square brackets can be neglected. Then, i.e., in the stationary state, the average work is $w\tau$, which means that the rate of work w done on the system is constant. To calculate the average heat, we use Eq. (12) and consider first the average potential energy

$$\langle U(\mathbf{x}_t,t)\rangle = \int d\mathbf{x}\rho(\mathbf{x},t)\frac{1}{2}|\mathbf{x} - \mathbf{v}^* t|^2 = \frac{3}{2} + \frac{1}{2}w(1 - e^{-t})^2 \tag{16}$$

[by Eqs. (8) and (15)], from which we find, by Eq. (13), that $\langle \Delta U_\tau \rangle = w\left[(1-e^{-\tau})e^{-t} - \frac{1}{2}(1-e^{-2\tau})e^{-2t}\right]$. Combining this with Eqs. (12) and (14), one obtains

$$\langle Q_\tau \rangle = w\left[\tau + \tfrac{1}{2}(1 - e^{-2\tau})e^{-2t}\right]. \tag{17}$$

From this equation, we see that for the heat too, after a transient time of $O(1)$, the rate of heat production becomes a constant. Furthermore, this constant is the same as that for the work in Eq. (14). So in the stationary state, *on average* all work is converted into heat. One might therefore be tempted to identify heat and work, at least in the stationary state. However, it will become clear that such identification is *not possible* for the *fluctuations* in work and heat.

C. Distribution of work fluctuations

Here we will briefly summarize the results for the work fluctuations from Ref. [12] needed in this paper. The work in Eq. (11) is defined as an integral over the path of the particle. Since W_τ is linear in the path $\mathbf{x}_t$, and $\mathbf{x}_t$ is Gaussian distributed, it follows that W_τ is Gaussian distributed as well. Hence, only the first two moments of the PDF need to be considered.

For the work-related TFT, let us first consider the PDF for work done on the system during a *transient state* of time τ, i.e., for an ensemble of transient trajectories of duration τ, starting in equilibrium at $t=0$ [2,8]. The first moment can then be found from Eq. (14), with $t=0$,

$$\langle W_\tau \rangle_{\mathrm{T}} = w(\tau - 1 + e^{-\tau}), \tag{18}$$

where subscript T denotes that the average is over transient trajectories. The second moment can be calculated as in Ref. [12], using the relations in Eqs. (8)–(11), to be

$$\langle [W_\tau - \langle W_\tau \rangle]^2 \rangle_{\mathrm{T}} = 2w(\tau - 1 + e^{-\tau}). \tag{19}$$

This is exactly twice the average in Eq. (18), which turns out to be crucial. Given the mean m and the variance v of a general Gaussian distributed variable s, its PDF is given by $P(s) = \exp[-(s-m)^2/2v]/\sqrt{2\pi v}$, hence

$$\frac{P(s)}{P(-s)} = e^{2ms/v}. \tag{20}$$

If the variance is twice the mean, $v=2m$, the right-hand side (rhs) of this equation becomes e^s, which is of a similar form as the FT in Eq. (1). To apply this equation to the work fluctuations in a convenient way, we define p as the average rate of work in time τ, W_τ/τ, scaled by the average rate w of work in the stationary state:

$$p \equiv \frac{W_\tau}{w\tau}. \tag{21}$$

The transient PDF $P_{\mathrm{T}}^W(W_\tau;\tau)$ of W_τ and the transient PDF $\pi_{\mathrm{T}}^W(p;\tau)$ of p are related by a constant Jacobian:

$$\pi_{\mathrm{T}}^W(p;\tau) = w\tau P_{\mathrm{T}}^W(W_\tau;\tau). \tag{22}$$

Superscript W denotes that these are PDFs related to work. Using Eq. (20) and the fact that for $s=W_\tau$ the variance is twice the mean, one finds the conventional TFT:

$$\frac{\pi_{\mathrm{T}}^W(p;\tau)}{\pi_{\mathrm{T}}^W(-p;\tau)} = \exp[W_\tau] = \exp[w\tau p]. \tag{23}$$

Note that the Jacobian drops out on the left-hand side of this equation. Equation (23) coincides with Eq. (1), with $\langle A \rangle_\tau = wp$, and shows that the TFT holds for work fluctuations. In fact, it holds exactly ("as an identity" [8]) for all times τ.

For the work-related SSFT, one should in principle look at a single (half-infinite) trajectory in the stationary state, and consider the statistics of work done in time intervals of length τ along that trajectory. However, because of its stochastic nature, our system behaves ergodically and the distribution of the initial points of these intervals is simply the stationary one. Thus, the desired statistics can be determined by considering a single interval τ of which the initial point is sampled from the stationary state. In addition, as shown in the previous sections, the system reaches a stationary state as $t\to\infty$, in an exponential fashion. Hence we can generate the sampling of the initial point of the interval by considering the interval $[t, t+\tau]$ for $t\to\infty$. In that limit, Eq. (14) gives

$$\langle W_\tau \rangle_{\mathrm{S}} = w\tau. \tag{24}$$

The subscript S denotes that the average is to be taken over trajectories in the stationary state, in the $t\to\infty$ limit just explained. The variance can be computed similarly, and the result is [12]

$$\langle [W_\tau - \langle W_\tau \rangle]^2 \rangle_{\mathrm{S}} = 2w(\tau - 1 + e^{-\tau}). \tag{25}$$

Note that now the variance in Eq. (25) is not equal to twice the mean in Eq. (24), except when $\tau\to\infty$. We therefore expect that only for $\tau\to\infty$, we have

$$\frac{\pi_{\mathrm{S}}^W(p;\tau)}{\pi_{\mathrm{S}}^W(-p;\tau)} \approx \exp[W_\tau] = \exp[w\tau p], \tag{26}$$

where π_{S}^W is the PDF of the time-averaged scaled work p in the stationary state. Since the rhs goes to infinity in the τ

$\to\infty$ limit, Eq. (26) is not a well-defined result. A more precise form of the SSFT is

$$\lim_{\tau\to\infty} f_{\mathrm{S}}^{W}(p;\tau) = p, \tag{27}$$

where one defines the *work fluctuation function* as

$$f_{\mathrm{S}}^{W}(p;\tau) = \frac{1}{w\tau}\ln\left[\frac{\pi_{\mathrm{S}}^{W}(p;\tau)}{\pi_{\mathrm{S}}^{W}(-p;\tau)}\right]. \tag{28}$$

Indeed, using Eqs. (20), (24), and (25), Eq. (27) follows, i.e., the conventional SSFT for the work fluctuations holds [12].

D. Distribution of heat fluctuations

The difficulty in considering the heat Q_τ instead of the work W_τ lies herein, that it is not linear in the position of the Brownian particle, through the contribution of U in Eqs. (12) and (13), which according to Eq. (3) is quadratic in $\mathbf{x}_t$. As a result, the PDF of Q_τ, denoted by P_{T}^{Q} in the transient state and P_{S}^{Q} in the stationary state (with superscript Q for heat), will not be Gaussian and it does not suffice to calculate their first two moments. Nonetheless, the Fourier transforms of these PDFs,

$$\hat{P}_j(q;\tau) \equiv \int_{-\infty}^{\infty} dQ_\tau e^{iqQ_\tau} P_j^{Q}(Q_\tau;\tau), \tag{29}$$

can be explicitly calculated, as will be shown below. Here, j stands either for T or for S, and this notation will be used throughout the paper.

Consider first the transient case. We start by considering the joint PDF $P_{\mathrm{T}}^{*}(W_\tau,\Delta\mathbf{x}_1,\Delta\mathbf{x}_2;\tau)$ of W_τ, $\Delta\mathbf{x}_1$ and $\Delta\mathbf{x}_2$. Here, $\Delta\mathbf{x}_1$ and $\Delta\mathbf{x}_2$ are the positions of the particle relative to the position of the minimum of the potential U at times 0 and τ, respectively, i.e.,

$$\Delta\mathbf{x}_1 \equiv \mathbf{x}_0, \tag{30}$$

$$\Delta\mathbf{x}_2 \equiv \mathbf{x}_\tau - \mathbf{v}^*\tau. \tag{31}$$

Because W_τ, $\Delta\mathbf{x}_1$, and $\Delta\mathbf{x}_2$ are all linear in the path $\mathbf{x}_t$, P_{T}^{*} is Gaussian, so that only its mean and variance are needed. Using Eqs. (3), (12), and (13), the transient PDF P_{T}^{Q} of Q_τ is related to $P_{\mathrm{T}}^{*}(W_\tau,\Delta\mathbf{x}_1,\Delta\mathbf{x}_2;\tau)$ by

$$P_{\mathrm{T}}^{Q}(Q_\tau;\tau) = \int\int\int dW_\tau d\Delta\mathbf{x}_1 d\Delta\mathbf{x}_2 P_{\mathrm{T}}^{*}(W_\tau,\Delta\mathbf{x}_1,\Delta\mathbf{x}_2;\tau) \times \delta\left(Q_\tau - W_\tau + \frac{1}{2}[|\Delta\mathbf{x}_2|^2 - |\Delta\mathbf{x}_1|^2]\right). \tag{32}$$

The same can be done for the stationary case. The PDF $P_{\mathrm{S}}^{*}(W_\tau,\Delta\mathbf{x}_1,\Delta\mathbf{x}_2;\tau)$ is defined as the infinite time limit of the joint PDF of W_τ, and the relative positions $\Delta\mathbf{x}_1$ and $\Delta\mathbf{x}_2$ defined by

$$\Delta\mathbf{x}_1 \equiv \mathbf{x}_t - \mathbf{v}^* t, \tag{33}$$

$$\Delta\mathbf{x}_2 \equiv \mathbf{x}_{t+\tau} - \mathbf{v}^*(t+\tau). \tag{34}$$

Like P_{T}^{*}, P_{S}^{*} is Gaussian. From P_{S}^{*}, the stationary state PDF of Q_τ can be found from

$$P_{\mathrm{S}}^{Q}(Q_\tau;\tau) = \int\int\int dW_\tau d\Delta\mathbf{x}_1 d\Delta\mathbf{x}_2 P_{\mathrm{S}}^{*}(W_\tau,\Delta\mathbf{x}_1,\Delta\mathbf{x}_2;\tau) \times \delta\left(Q_\tau - W_\tau + \frac{1}{2}[|\Delta\mathbf{x}_2|^2 - |\Delta\mathbf{x}_1|^2]\right). \tag{35}$$

The mean and variance of P_j^{*} in Eqs. (32) and (35) are calculated in Appendix A.

Equations (32) and (35) cannot be integrated explicitly, due to the quadratic nature of ΔU_τ, but their Fourier transforms can be obtained, as is seen when one combines Eq. (32) or (35) with Eq. (29) to yield

$$\hat{P}_j(q;\tau) = \int\int\int dW_\tau d\Delta\mathbf{x}_1 d\Delta\mathbf{x}_2 P_j^{*}(W_\tau,\Delta\mathbf{x}_1,\Delta\mathbf{x}_2;\tau) \times e^{iq[W_\tau - (|\Delta\mathbf{x}_2|^2 - |\Delta\mathbf{x}_1|^2)/2]}. \tag{36}$$

Because P_j^{*} and $e^{iq[W_\tau-(|\Delta\mathbf{x}_2|^2-|\Delta\mathbf{x}_1|^2)/2]}$ are both Gaussian, the integrals in Eq. (36) can be explicitly performed. The details of the calculation are given in Appendix B, with the result

$$\hat{P}_j(q;\tau) = \frac{\exp\left\{wq(i-q)\left[\tau - \dfrac{[1-e^{-\tau}][\Delta_j + (\Delta_j/2 + 2q^2)(1-e^{-\tau})]}{1+(1-e^{-2\tau})q^2}\right]\right\}}{[1+(1-e^{-2\tau})q^2]^{3/2}}, \tag{37}$$

where for j=T, Δ_{T}=1, while for j=S, Δ_{S}=0. The PDFs P_j^{Q} are related to the $\hat{P}_j(q;\tau)$ by the inverse Fourier transform

$$P_j^{Q}(Q_\tau;\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dq\, e^{-iqQ_\tau}\hat{P}_j(q;\tau). \tag{38}$$

Before we proceed with the evaluation of $P_j^{Q}(Q_\tau;\tau)$ from Eqs. (37) and (38), we note that if the conventional FT were to hold exactly, i.e., [cf. Eq. (1) with $\langle A\rangle_\tau = Q_\tau/\tau$]

$$\frac{P_j^{Q}(Q_\tau;\tau)}{P_j^{Q}(-Q_\tau;\tau)} = e^{Q_\tau}, \tag{39}$$

then, by the definition (29), we would have

$$\hat{P}_j(q;\tau) = \hat{P}_j(i-q;\tau). \tag{40}$$

If Eq. (40) does not hold, the fluctuation theorem in Eq. (39) cannot hold exactly for all τ. Neither $\hat{P}_{\mathrm{T}}(q;\tau)$ nor $\hat{P}_{\mathrm{S}}(q;\tau)$ in Eq. (37) satisfy Eq. (40). That means we can rule out the possibility of a TFT or a SSFT for heat fluctuations that holds as an identity for all τ.

The first term in the exponent in Eq. (37) does have the symmetry in Eq. (40), and this term is the only one that grows with the time τ. However, one cannot conclude from this that the FTs hold for large τ, due to the singularities in Eq. (37). In the rest of the paper, we will be concerned with whether Eq. (39) holds, both for transient and stationary states, for $\tau\to\infty$, or, if it does not, what the deviations from this behavior are. To be precise, similar to the treatment in Sec. II C, we define the scaled heat fluctuations—denoted by the same symbol p as scaled work fluctuations—as

$$p \equiv \frac{Q_\tau}{w\tau}, \tag{41}$$

[cf. Eq. (21) and the remark below Eq. (17)]. Its PDF is

$$\pi_j^Q(p;\tau) = w\tau P_j^Q(w\tau p;\tau) \tag{42}$$

$$= \frac{w\tau}{2\pi}\int_{-\infty}^{\infty} dq e^{-iqw\tau p}\hat{P}_j(q;\tau) \tag{43}$$

[cf. Eq. (22)], where Eq. (38) was used. Furthermore, we define the *heat fluctuation function* [cf. Eq. (26)]

$$f_j^Q(p;\tau) \equiv \frac{1}{w\tau}\ln\left[\frac{\pi_j^Q(p;\tau)}{\pi_j^Q(-p;\tau)}\right], \tag{44}$$

and we investigate whether [cf. Eq. (28)]

$$\lim_{\tau\to\infty} f_j^Q(p;\tau) = p \tag{45}$$

as is required for the conventional FT.

III. NUMERICAL APPROACH

Since no exact Fourier inverse of Eq. (37) is known, we first treat the problem of finding the PDF of heat fluctuations numerically.

We used two numerical methods. The first is a sampling method which starts from the expressions in Eqs. (32) and (35) that give the P_j^Q in terms of the P_j^*. From the Gaussian PDF P_j^* with the mean and variance calculated in Appendix A, one draws many sets of values $(W_\tau, \Delta\mathbf{x}_1, \Delta\mathbf{x}_2)$, using a random number generator [16]. From each set, using Eqs. (12), (13), and (41), $p=(W_\tau-\Delta U_\tau)/w\tau$ is calculated. From these p values we build a histogram by constructing bins of width δp in an interval $[-p_{max}, p_{max}]$ and counting how many of the values fall into each bin. If the number of sample points N_s is large and δp is small, the histogram gives a good approximation for the PDF π_j^Q. As a simple error estimate, one performs this procedure a few times with varying random seeds, and determines the spread in the results.

The second numerical method is the standard *fast Fourier transform* algorithm [16] applied to the inverse Fourier transform. This algorithm takes a discrete set of values of the function $\hat{P}_j(q)$ for q in an interval $[-q_{max}, q_{max}]$ with equal spacing δq between the q-values in the interval, and returns values of the (inverse) transform, i.e., of the function P_j^Q, on a reciprocal grid of values for Q_τ. If δq is small enough, and q_{max} is large enough, this can be used to obtain a good approximation for the (inverse) Fourier transform. An error estimate can be found by varying δq and q_{max} and observing to what extent the results have converged. In all plots in this paper, these errors would be unobservable and are therefore not plotted. Note that once P_j^Q is known, by Eq. (42) we can obtain π_j^Q as well.

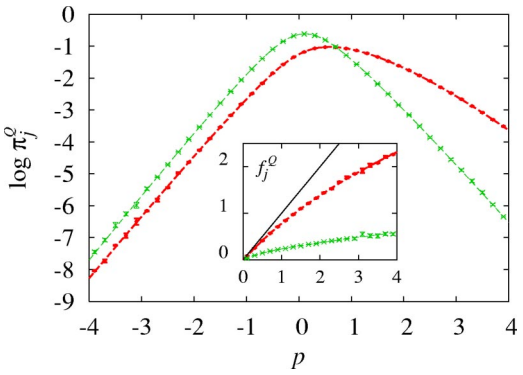

FIG. 1. The PDFs π_j^Q obtained by two numerical methods, for $w=2.0$ and $\tau=1.0$. Sampling method results ($N_s=2\,000\,000$, $\delta p=0.2$) are shown as crosses for π_{T}^Q and as dots for π_{S}^Q, respectively. Fast Fourier transform results ($\delta q=4\times10^{-4}$, $q_{max}=200$) are shown as a thin dashed line for π_{T}^Q and as a bold dashed line for π_{S}^Q. The inset shows the corresponding fluctuation functions f_j^Q [Eq. (44)] and a solid straight line indicating the prediction of the conventional FT.

The results of these two methods for the PDFs π_{T}^Q and π_{S}^Q for the parameter values $w=2.0$ and $\tau=1.0$ have been plotted in Fig. 1 as a function of p, and the corresponding fluctuation functions f_{T}^Q and f_{S}^Q are plotted in the inset in that figure. The two numerical methods agree very well, but do not agree at all with the straight line with slope 1 of the conventional FT [i.e., p cf. Eq. (45)], which is also drawn in the inset in Fig. 1, neither in the stationary nor in the transient case.

To investigate the discrepancies with the conventional FTs further, we have explored a range of values for the parameters w and τ, using the two numerical methods. Some results are shown in Fig. 2. One sees that for large values for p, deviations of $f_j^Q(p;\tau)$ from the straight line p of the conventional FT are generic.

Furthermore, Fig. 3 shows (for $w=2.0$, but the same holds for other w values) that for large τ, the straight line is approached for small p, both in the transient and the stationary case, although the approach is considerably slower for the former.

IV. ANALYTIC APPROACH—SADDLE POINT METHOD

The numerical methods are not suitable to give reliable results for the tails of the PDFs for large τ. To get around

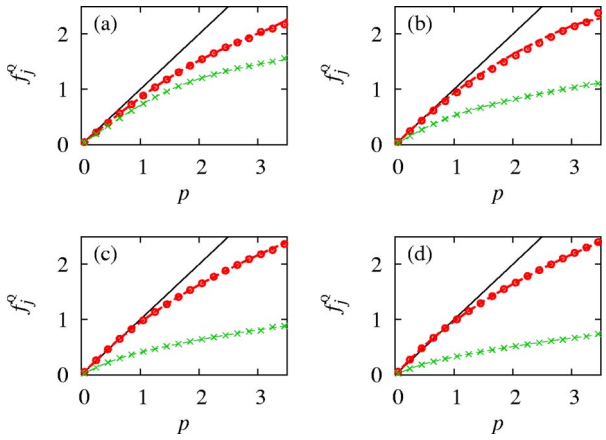

FIG. 2. Numerical results for the fluctuation functions f_j^Q for the stationary ($j=$S) and the transient case ($j=$T), for (a) $w=1, \tau=4$, (b) $w=2, \tau=2.5$, (c) $w=3, \tau=1.3333$, and (d) $w=4, \tau=1$. Sampling method results ($N_s=1.2\times10^9, \delta p=0.2$) are shown as thin crosses for f_T^Q and as bold dots for f_S^Q. Fast fourier transform results ($\delta q=4\times10^{-4}, q_{max}=200$) are shown as the dashed thin curves for f_T^Q and as dashed bold curves for f_S^Q. The solid straight line indicates the prediction of the conventional FT.

this, we developed an asymptotic analytic approach.

The starting point of the approach is to notice that the exponent in Eq. (37) grows linearly with τ. Thus, π_j^Q, given by Eq. (43), can be written as

$$\pi_j^Q(p;\tau) = \frac{w\tau}{2\pi}\int_{-\infty}^{\infty} dq e^{-w\tau[e_j(q;\tau)+iqp]}, \tag{46}$$

where

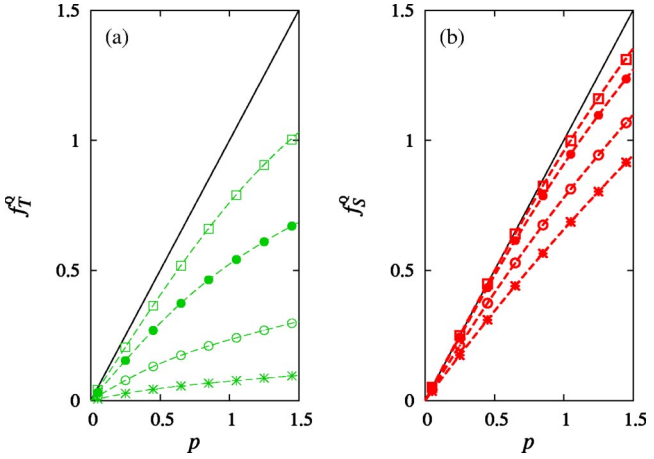

FIG. 3. Numerical results for the τ dependence of the fluctuation functions f_j^Q for $w=2.0$ for (a) the transient ($j=T$) and (b) the stationary case ($j=S$). Sampling results ($N_s=1.2\times10^9, \delta p=0.2$) are shown as symbols: (*) $\tau=0.5$, (○) $\tau=1$, (•) $\tau=2$, (□) $\tau=4$. Fast Fourier transform results ($\delta q=4\times10^{-4}, q_{max}=200$) are shown as the dashed curves (the τ values of these curves are the same as those of the coinciding sampling points). The solid straight line is the conventional FT result.

$$e_j(q;\tau) \equiv -\frac{1}{w\tau}\ln[\hat{P}_j(q;\tau)] \tag{47}$$

is of order τ^0 [see Eq. (37)]. The presence of a large parameter τ in Eq. (46) makes it suitable for the saddle-point method [17], which we will briefly describe now.

To get a good approximation of the integral in Eq. (46), one considers

$$h(q) \equiv -w[e_j(q;\tau) + iqp] \tag{48}$$

as a function of a complex-valued q (where the dependences of h on j, p, and τ are suppressed). Then, one determines its *saddle point*, i.e., the complex number $q=q^*$ for which $\partial h/\partial q=0$. Through this point q^* in the complex plane lies a *path of steepest descent* S along which Im h is constant and Re h attains a maximum at q^* [17]. Next, one continuously deforms the original path of integration R (here the real axis) to this path of steepest descent S without crossing singularities. The integral over S is for large τ dominated by the (small) segment on S around the saddle point q^* and can be expanded in inverse powers of τ by Taylor expanding the function h on S around the saddle point. For the leading term one uses a second-order Taylor expansion and finds [17]

$$\int_{-\infty}^{\infty} dq\, e^{\tau h(q)} = \sqrt{\frac{2\pi}{\tau|h''(q^*)|}} e^{\tau h(q^*)+i\alpha}[1+O(\tau^{-1})]. \tag{49}$$

Here, α is the angle between the direction of the path S when it traverses the saddle point, and the real axis. Furthermore, $h''(q^*)$ is the second derivative of h with respect to q at the saddle point q^*, which is assumed to be nonzero. The correction term in Eq. (49) is $O(\tau^{-1})$, though that strictly applies only when the function h does not depend on τ. In our case it does, and we will see that consequently the correction terms can then be $O(\tau^{-1/2})$, although the leading behavior in Eq. (49) is not affected.

The form of $h(q)$ in Eq. (48) has the following simplifying consequences for the expression in Eq. (49). First, the equation for the saddle point, $\partial h/\partial q|_{q=q^*}=0$, becomes

$$e_j'(q^*;\tau) = -ip, \tag{50}$$

where e_j' is the derivative of e_j with respect to q. Second, $h''(q^*)$ in Eq. (49) is equal to $-we_j''(q^*)$. This second derivative can be further simplified by noticing that Eq. (50) holds for all p, so we can take the derivative with respect to p on both sides to find $e_j''(q^*;\tau)(\partial q^*/\partial p)=-i$. Hence

$$|h''(q^*)| = w|e_j''(q^*;\tau)| = w\left|\frac{\partial q^*}{\partial p}\right|^{-1}. \tag{51}$$

Third, the approximation is for a PDF, which should be real and positive. Thus, if the saddle-point method is to be a consistent calculation of a PDF, we must have

$$\alpha = 0. \tag{52}$$

We will prove later that this is indeed the case. Using Eqs. (48), (51), and (52), we apply the saddle point approximation Eq. (49) to the integral in Eq. (46) to find

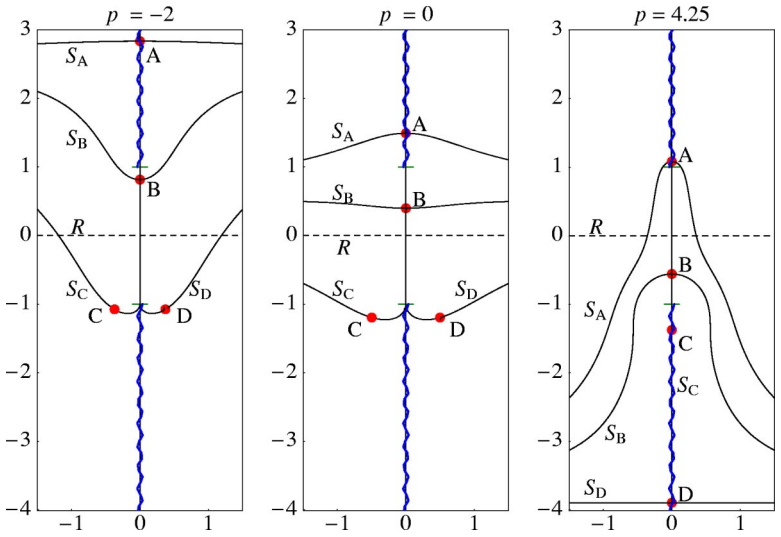

FIG. 4. Geometry in the complex q plane of the saddle points, singularities, and paths of steepest descent of $h(q)$ [Eqs. (48) and (54)] for the stationary state [$\Delta_j=0$], for $w=1$, $\tau=4$, and $p=-2$, 0, and 4.25, respectively. In all graphs, solid dots depict the position of the saddle points A, B, C, and D. Wiggly curves give the branch cuts ending in the branch points at $q=\pm i$ which are indicated with a thin horizontal line. The branch point at $-i$ is also a pole. The dashed line R gives the original line of integration, i.e., the real axis. Solid lines labeled S_A, S_B, S_C, and S_D are paths of steepest descent though the saddle point, on which Re h attains a maximum.

$$\pi_j^Q(p;\tau) \sim \sqrt{\frac{w\tau}{2\pi}\left|\frac{\partial q^*}{\partial p}\right|}\, e^{-w\tau[e_j(q^*;\tau)+ipq^*]}. \qquad (53)$$

Here, the symbol $\sim$ is for convenience used to denote the asymptotic behavior, instead of explicitly denoting the correction terms as $O(\tau^{-1})$ as we did in Eq. (49).

As mentioned above Eq. (49), this method gives an asymptotic expansion in inverse powers of τ. Now if τ is large enough so that terms of relative order τ^{-1} can be neglected, then terms which are exponentially small in τ can also be neglected [18]. Therefore, we use in the definition of the exponents e_j in Eq. (47) the expression of the Fourier transform $\hat{P}_j(q;\tau)$ in Eq. (37) without exponential terms, so that [19]

$$e_j(q;\tau) = -q(i-q) - \frac{4q^3+3\Delta_j q}{2\tau(i+q)} + \frac{3\,\ln(1+q^2)}{2w\tau}. \qquad (54)$$

When substituted in Eq. (50), we obtain a fourth-order polynomial equation for the saddle points, so that there are four saddle points. To know which of the four saddle points to use, one needs to determine their location, the paths of steepest descent that traverse them, and to see if our initial integration line (the real axis) can be deformed to lie on one of these paths without crossing any singularities.

For three typical values for p we have depicted the four saddle points of h for the stationary state in Fig. 4, calculated analytically by MATHEMATICA [20] which uses Ferrari's method for solving the quartic equation [21]. We see in Fig. 4 that for all p there is one saddle point (A) on the imaginary axis above i, one in between $-i$ and i (B), and two (C and D) who are either both purely imaginary but below $-i$, or are complex with the same imaginary part (also below $-i$), and with opposite real parts. In Fig. 4 there are a pole at $q=-i$, and branch cuts which we have taken along the imaginary axis, one from i upward and one from $-i$ downward. These cuts come from the logarithmic term in Eq. (54). Finally and most importantly, Fig. 4 shows the paths of steepest descent S_A, S_B, S_C, and S_D through the saddle points A, B, C, and D. These paths were found by having MATHEMATICA solve for lines of constant Im h, where the values of the constant for Im h are chosen to match the values of Im h at the four saddle points.

Knowing now the geometry of h in the complex plane, we can determine which saddle point to use. One sees that for all p the path S_B through saddle point B is a possible deformation of the real axis: each point on the real axis is simply shifted up or down until it lies on this path and no poles or other singularities are crossed in doing so. Strictly speaking, the deformation changes the end points because they get shifted by an imaginary amount ia [where a can be shown to be $(1-p)/2$], but one can add portions ia to the integration path from $-\infty$ to $-\infty+ia$ and from $\infty+ia$ to ∞, and as the integrand vanishes on these portions, the integral taken over R and over S_B have the same value. This is not true for S_A, S_C, and S_D.

The graphs in Fig. 4 suggest that the path S_B goes through B horizontally, i.e., that $\alpha=0$ as Eq. (52) conjectured. To prove this, notice that for purely imaginary values for q, i.e., $q=i\lambda$, $h(i\lambda)=-w[e(i\lambda;\tau)-\lambda p]$ is real if $-1<\lambda<1$ [cf. Eq. (54)]. Hence, the imaginary axis from $-i$ to i is a line of constant imaginary part of h. In the saddle point, lines of constant imaginary part cross perpendicularly [17]. Thus, in saddle point B, the curve of steepest descent is seen to be perpendicular to the imaginary axis, i.e., in the direction of the real axis, so that indeed $\alpha=0$.

We remark that the above analysis for the stationary case can be carried through also for the transient case, with only a slight change in the saddle points.

Our saddle-point method now consists of the following. We use MATHEMATICA to determine analytically the purely imaginary solution q^* of Eq. (50) using Eq. (54) with $|q^*|<1$, as a function of p. Given that solution, the PDF π_j^Q is explicitly calculated using Eqs. (53) and (54). Once π_j^Q is known, we determine the fluctuation functions f_j^Q from Eq. (44). We can then investigate the FTs, even for finite τ, something which is usually not possible, but is doable in this case because the Fourier transform is known exactly for all time.

To see how good the saddle-point expression for π_j^Q in Eq. (53) is, we compare it with the numerical results from

Sec. III, specifically, the numerically determined Fourier inverse. This is shown in Fig. 5, both for the transient and for the stationary case. In Fig. 5(a), we chose $\tau=2.5$ and $w=0.5$, i.e., not too large, and we are therefore still able to see the discrepancies between the saddle-point result and the inverse Fourier transform for π_j^Q, which are about 15 % near the peak, and just a few percents in the tails. In Fig. 5(b), we chose $\tau=5.0$, which is still not too large, but we now see that the saddle-point method has become quite accurate, up to about 7 % in the peak of the distribution, while in the tails one can hardly distinguish the results from the two methods. As τ gets larger, the results from these two methods approach each other even more and plotting them together as in Fig. 5 would show basically two curves on top of the other.

While the numerical methods in Sec. III could give us the large τ behavior of $f_j^Q(p;\tau)$ only for small p values, using the saddle-point approximation, we can now determine the behavior of $f_j^Q(p;\tau)$ with increasing τ for the full range of p values. The results are plotted in Fig. 6 and show that both for the transient and for the stationary case, the fluctuation function f_j^Q approaches the conventional FT only for small p values. In contrast, for large p values ($p>3$), a completely different limit emerges as $\tau\to\infty$, one where $f_j^Q(p;\tau)$ appears to approach 2, both for the transient and for the stationary case. The exact form of this $\tau\to\infty$ limit of $f_j^Q(p;\tau)$ is given by Eq. (68) in Sec. V.

V. ANALYTIC EXPRESSIONS FOR EXTENDED FLUCTUATION THEOREMS

Because we do not find it satisfactory to have the analytical result of the saddle-point method only in the memory of MATHEMATICA, we will determine in this section explicitly the leading asymptotic behavior of π_j^Q and corrections as τ gets large, using Eqs. (50), (53), and (54).

To calculate the saddle-point approximation for π_j^Q given in Eq. (53), we need to find the saddle-point q^* first. With e_j from Eq. (54), Eq. (50) for q^* takes the form

$$1+2iq^*+\frac{4q^{*2}(3-2iq^*)+3\Delta_j}{2\tau(i+q^*)^2}-\frac{3iq^*}{w\tau(1+q^{*2})}=p. \tag{55}$$

This equation always has four (complex) solutions as explained below Eq. (54). Among these four solutions, one is always purely imaginary and lies between $-i$ and i, which is the one we need, as it corresponds to saddle-point B (see preceeding section). In order to find this solution, we try the expansion

$$q^*=q_0^*+\frac{q_1^*}{\tau}+O(\tau^{-2}), \tag{56}$$

where q_0^* is given by the solution of Eq. (55) with $\tau\to\infty$, i.e., $1+2iq^*=p$. Hence

$$q_0^*=i(1-p)/2. \tag{57}$$

Since q^* lies between $-i$ and i, the above zeroth-order solution is only appropriate if it lies in that interval, i.e., if $-1<p<3$. We will refer to this as case (a), which will be dealt with first. The cases (b) $p<-1$ and (c) $p>3$ will be treated later.

(a) For $-1<p<3$, Eq. (57), when substituted into Eq. (53) using e_j from Eq. (54), yields a zeroth-order solution for $\pi_j^Q(p;\tau)$. Because q_0^* is the maximum of h in Eq. (49), there is no need to calculate q_1^*, as far as the exponent h is concerned [22]. Furthermore, in the prefactor in Eq. (53), $\partial q_0^*/\partial p$ is of order 1, so adding $\tau^{-1}\partial q_1^*/\partial p$ gives a correction of relative order $O(\tau^{-1})$, i.e., of the same order as the correction terms in Eq. (49). So there is no point to work out q^* to higher orders than q_0^*. Substituting the expression for q_0^*, given in Eq. (57), for q^* in Eq. (53) and using Eq. (54), yields

$$\pi_j^Q(p;\tau)\sim\sqrt{\frac{16w\tau}{\pi}}\times\frac{e^{-w[\tau(1-p)^2+2(1-p)/(3-p)\{(1-p)^2-3\Delta_j\}]/4}}{[(3-p)(1+p)]^{3/2}}\quad\text{for}-1<p<3, \tag{58}$$

with correction terms that are relatively $O(\tau^{-1})$.

Outside of the range $-1<p<3$, the zeroth-order Ansatz in Eq. (57) fails to produce a solution for q^* between $-i$ and $+i$. The explanation is that the expression for q_0^* in Eq. (57) was based on neglecting the last two terms on the left-hand side (lhs) of Eq. (55), which clearly does not work for $p>3$ or $p<-1$. However, the only way that these terms can become comparable to the first term in that equation, is if their denominators become $O(1)$ instead of $O(\tau)$ as they appear to be. For this another ansatz than Eq. (56) is necessary. As there are two different denominators in Eq. (55), there are two cases. In the first case, the denominator of the last term on the lhs of Eq. (55) is taken to be $O(1)$, i.e., $\tau(1+q^{*2})=O(1)$, or $q^*=\pm i+O(\tau^{-1})$. The sign ambiguity is lifted by noting that for $q^*=-i+O(\tau^{-1})$, the second term in Eq. (55) would diverge $\propto\tau$, so only

$$q^*=+i+O(\tau^{-1}) \tag{59}$$

is correct. In the second case, the denominator of the second term on the lhs of Eq. (55) is taken to be $O(1)$, i.e., $\tau(i+q^*)^2=O(1)$, or

$$q^*=-i+O(\tau^{-1/2}). \tag{60}$$

(b) We first consider the case $q^*=i+O(\tau^{-1})$ of Eq. (59), which we write for convenience as [23]

$$q^*=i+\frac{ia}{w\tau}+O(\tau^{-2}). \tag{61}$$

Substituting Eq. (61) into Eq. (55), and keeping only $O(1)$ terms, one gets $-1+(3/2a)=p$, so $a=3/2(p+1)$, and

$$q^*=i+\frac{3i}{2w\tau(p+1)}+O(\tau^{-2}). \tag{62}$$

Therefore, for $p<-1$, this is an approximation of the solution q^* below i. Using Eqs. (53), (54), and (62), one finds

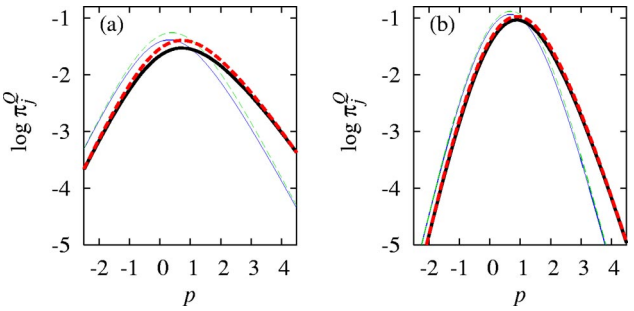

FIG. 5. Comparison of the saddle-point approximation with the numerically obtained solution, for $w=0.5$ and (a) $\tau=2.5$ and (b) $\tau=5$. The solid thin line is the transient PDF $\pi_T^Q(p;\tau)$ as calculated using the saddle-point approximation of Sec. IV and the dashed thin curve is the same PDF calculated using the numerical inverse Fourier transform of Sec. III ($\delta q=2\cdot 10^{-4}, q_{max}=2^{18}\delta q=52.4288$). The solid bold line is the stationary PDF $\pi_S^Q(p;\tau)$ calculated using the saddle-point approximation, and the bold dashed curve is that same PDF obtained using the numerical inverse Fourier transform.

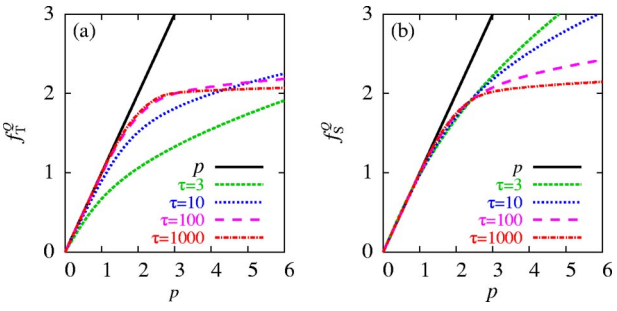

FIG. 6. Results for the large τ behavior of (a) $f_T^Q(p;\tau)$ and (b) $f_S^Q(p;\tau)$ for $w=1$, as found from the saddle-point approximation. The figures show four lines corresponding to the result for $\tau=3$, $\tau=10$, $\tau=100$, and $\tau=1000$, as well as a solid black straight line (p) that gives the behavior expected from the conventional FT.

$$\pi_j^Q(p;\tau) \sim \sqrt{\frac{w^3\tau^3|p+1|}{36\pi}} e^{-w[-\tau p+1-(3/4)\Delta_j]+3/2} \quad \text{for } p<-1. \tag{63}$$

Again, correction terms to this are of relative $O(\tau^{-1})$.

(c) Finally, we consider the case $q^*=-i+O(\tau^{-1/2})$ of Eq. (60) and write $q^*=-i+(ib/\sqrt{\tau})+(ic/\tau)$. Note that in principle, we needed to include two correction terms here to get a similar level of approximation as in Eqs. (56) and (61). However, as the sum of the first two terms (i.e., $-i+ib/\sqrt{\tau}$) is to leading order the maximum of the exponent, we need not calculate c [22]. Substituting $q^*=-i+ib\tau^{-1/2}$ into Eq. (55), gathering the $O(1)$ terms and using that $\Delta_j^2=\Delta_j$, we find $b=\pm(2-\Delta_j)/\sqrt{2(p-3)}$. We need the positive solution, so that q^* is above $-i$, i.e.,

$$q^* = -i + i\frac{2-\Delta_j}{\sqrt{2(p-3)\tau}} + O(\tau^{-1}). \tag{64}$$

Note that we need $p>3$ for this to be purely imaginary, as it should be. Combining Eqs. (53), (54), and (64) gives

$$\pi_j^Q(p;\tau) \sim \sqrt{\frac{w\tau^2}{16\pi(2-\Delta_j)^2}} \times \exp\left\{-w\left[(p-2)\tau-(2-\Delta_j)\sqrt{2(p-3)\tau} + \frac{40-12p+3\Delta_j(p-4)}{2(p-3)}\right]\right\}, \quad \text{for } p>3, \tag{65}$$

where $\Delta_j^2=\Delta_j$ has been used to simplify the expression. Due to the different form of the expansion of q^* here, compared to the cases (a) and (b), this result is valid up to corrections of relative order $O(\tau^{-1/2})$.

Note that we now have the asymptotic forms for π_j^Q in the three separate regions, $p<-1$, $-1<p<3$ and $p>3$ in the Eqs. (63), (58), and (65), respectively. This shows that the behavior in the center of the PDF is Gaussian-like, while in the tails it is exponential.

To investigate the FTs, we will now consider the fluctuation functions f_j^Q. For that, we have to take the above asymptotic expressions for large but finite τ to calculate f_j^Q of Eq. (44) (after which one can take the limit $\tau\to\infty$). In this way, one finds for the large but finite τ behavior of f_j^Q [24]:

$$f_j^Q(p;\tau) = \begin{cases} p + \dfrac{(8-6\Delta_j)p}{\pi(9-p^2)} - \dfrac{3\ln\dfrac{3-p^2+2p}{3-p^2-2p}}{2w\tau} + O(\tau^{-2}) & \text{for } 0\leqslant p<1 \\ p-(1-p)^2/4 - \dfrac{\ln\tau}{w\tau} + \dfrac{r_j(p)}{w\tau} + O(\tau^{-2}) & \text{for } 1<p<3 \\ 2+(2-\Delta_j)\sqrt{\dfrac{2(p-3)}{\tau}} - \dfrac{\ln\tau}{2w\tau} + O(\tau^{-1}) & \text{for } p>3, \end{cases} \tag{66}$$

where

$$r_j(p) = -\frac{1}{2}\ln[w^2(3-p)^3(1+p)^3(p-1)/576] - \frac{3}{2} + \frac{w(p+1)(3\Delta_j - 2p^2 + 8p - 10)}{4(3-p)}. \quad (67)$$

To obtain Eq. (66), we have used Eq. (58) for $0 \leqslant p < 1$, and we have combined Eqs. (58) and (63) for $1 < p < 3$, and Eqs. (63) and (65) for $p > 3$. Only the case $p \geqslant 0$ is given by Eq. (66). For $p < 0$, one can use that by the definition Eq. (44), $f_j^Q(p;\tau)$ is an odd function of p.

We remark that as we approach the limits of the range of p values for which each of the expressions for the PDF π_j^Q, [i.e., Eqs. (63), (58), and (65)] holds, their rhs's diverge. Similarly, the rhs's of Eq. (66) diverge there. These divergences are an artifact of the expansion in inverse powers of τ, as can be seen because such divergences are not observed in Sec. IV (i.e., in Figs. 5 and 6. In fact, for any τ, MATHEMATICA shows us that q^* is a continuous and differentiable function of p [21] (with a discontinuity in the derivative $\partial q^*/\partial p$ only as $\tau \to \infty$), leading to a continuously varying $\pi_j^Q(p;\tau)$ with p. Furthermore, considering the regions of p values in which Eq. (66) differs (due to these divergences) noticeably from the full saddle-point approximation, one finds from MATHEMATICA that these regions shrink to zero as $\tau \to \infty$.

For $\tau \to \infty$, the results in Eq. (66) become

$$\lim_{\tau\to\infty} f_j^Q(p;\tau) = \begin{cases} p & \text{for } 0 \leqslant p < 1 \\ p - (p-1)^2/4 & \text{for } 1 < p < 3 \\ 2 & \text{for } p > 3. \end{cases} \quad (68)$$

This is the extension of the conventional FT as it holds for infinite time. Comparing with Eq. (45), it is clear that the conventional FT only holds for not too large fluctuations, i.e., $|p| < 1$. For larger p, there is first a quadratic deviation after which the ratio between the probability for positive and for negative fluctuations becomes a constant. In contrast, for the conventional FT, f_j^Q keeps increasing linearly with the size of the fluctuation p.

Equation (66) gives the extension of the fluctuation theorem for *finite* τ. The deviations from the infinite τ results are of a different character for small ($|p| < 3$) and large ($|p| > 3$) fluctuations. The $f_j^Q(p;\tau)$ for large positive fluctuations ($p > 3$) approaches 2 with corrections that scale as $1/\sqrt{\tau}$. But whereas in the infinite τ limit in Eq. (68), the function no longer grows above 2 beyond $p = 3$, in the Eq. (66) for finite τ this function keeps growing to leading order as a square root, i.e., $f_j^Q(p;\tau) \sim \sqrt{(p-3)/\tau}$. The prefactor of $\sqrt{(p-3)/\tau}$ depends on whether we consider the transient or the stationary case—it is multiplied by $\sqrt{2}$ for the former and by $2\sqrt{2}$ for the latter—but it does not depend on w. For small fluctuations, on the other hand, which might be more relevant for many applications, we see corrections to the infinite τ limit of $O(\tau^{-1})$, which do depend on w. The value of w determines whether the curve starts around $p = 0$ above or below the line with slope 1, and thus also whether the slope of 1 is approached from below or from above. In fact, expanding Eq. (66) around $p = 0$, we get

$$f_j^Q(p;\tau) = \left[1 + \frac{2}{\tau}\left(\frac{4 - 3\Delta_j}{9} - \frac{1}{w}\right)\right]p + O(p^2). \quad (69)$$

This shows that the critical value for w is $w_c = 9/(4 - 3\Delta_j)$, i.e., $w_c = 9$ for the transient and $w_c = 9/4$ for the stationary case. For $w = w_c$, the slope (near $p = 0$) is 1 for finite times (up to corrections of order τ^{-2}), while the slope of 1 is approached for large τ from above for $w > w_c$, and from below for $w < w_c$. In contrast, for the conventional TFT, the slope is 1 for all τ, while for the conventional SSFT, at least for the work done on the system, the slope always approaches 1 from above, irrespective of w, with increasing τ.

VI. DISCUSSION

(1) This paper treats fluctuations in the heat developed in a system of a Brownian particle in water, confined by a harmonic potential, which moves at constant velocity through the fluid, dragging the Brownian particle with it. The theory of heat fluctuations developed in this system was based on an overdamped Langevin equation for the position of the particle. This theory required a far more sophisticated analysis than was used in the previous paper [12] for work fluctuations. It should be mentioned that some of this same sophistication is also found in the work of Farago for a quantity different from both the work and the heat [25].

(2) In essence, our theory deals with the fluctuations of the quantities occurring in the first law of thermodynamics, i.e., work, heat and internal energy. The energy balance for the system is

$$Q_\tau = W_\tau - \Delta U_\tau. \quad (70)$$

Here W_τ is the total work done on the system during a time τ, Q_τ is the heat produced by the Brownian particle in water in that time, and ΔU_τ is the difference in potential energy of the particle in the harmonic potential in the same time interval. Equation (70), which is basically the first law, can be applied both to averages as well as to fluctuations, because it expresses energy conservation, which holds both macroscopically and microscopically.

(3) The theory gives extensions of the conventional SSFT and TFT in Eq. (66). In the limit $\tau \to \infty$, leading to Eq. (68), this new theorem coincides with the conventional TFT and SSFT only for small fluctuations $p < 1$ (as Rey-Bellet and Thomas also found for a different system [26]), while for larger fluctuations the behavior is completely different from the conventional ones. We will now explain the qualitative behavior of work and heat fluctuations. For that, it is useful first to consider, in point (4) below, the system in equilibrium, i.e., in the situation in which the harmonic confining potential does not move. In point (5), we then discuss the nonequilibrium case.

(4) In equilibrium, work, potential energy, and heat behave as follows. (i) Work: There is no displacement and hence no work is done. Therefore, the PDF of W_τ is a δ function: $P(W_\tau) = \delta(W_\tau)$. (ii) Potential: The PDF of the po-

tential energy of the particle in the harmonic potential U is given by a Boltzmann factor $\sim\exp[-U]$ ($k_BT=1$), since the particle in the harmonic potential can be seen as a subsystem of a larger one. From this, one sees that the PDF of ΔU_τ behaves for large τ similarly in its tails, i.e., $P(\Delta U_\tau) \sim\exp(-|\Delta U_\tau|)$. (iii) Heat: For the PDF of the heat fluctuation, using Eq. (70) with $W_\tau=0$, we see that Q_τ also has exponential tails: $P(Q_\tau)\sim\exp(-|Q_\tau|)$.

(5) We now turn to the nonequilibrium, stationary state case. (i) Work: the PDF of W_τ is now a Gaussian [12]. (ii) Potential: The PDF of ΔU_τ is expected to have still exponential tails, at least near equilibrium, i.e., $P(\Delta U_\tau) \sim\exp(-|\Delta U_\tau|)$. (iii) Heat: The effect of the interplay between W_τ and ΔU_τ on the behavior of heat fluctuations can be deduced using Eq. (70). We have to separately consider small and large fluctuations. For large τ, W_τ and Q_τ grow on average linearly in time, while ΔU_τ stays of $O(1)$. Hence, for small fluctuations of these quantities (near their averages) one can neglect ΔU_τ in Eq. (70), so that $Q_\tau\approx W_\tau$, and the behavior of Q_τ is very similar behavior to that of W_τ, i.e., Gaussianlike. On the other hand, when a large fluctuation of Q_τ occurs, it is less likely to be due to a large fluctuation of W_τ than to a large fluctuation of ΔU_τ. This is so because the tails of the Gaussian PDF for W_τ are much smaller than the exponential tails of the PDF for ΔU_τ. As a result, W_τ will be near its average while ΔU_τ is large. For the sake of the argument, we put W_τ equal to its average, $\langle W_\tau\rangle$, which coincides with $\langle Q_\tau\rangle$ for large τ. Hence by Eq. (70), $Q_\tau\approx\langle W_\tau\rangle -\Delta U_\tau=\langle Q_\tau\rangle-\Delta U_\tau$. Using that the PDF for ΔU_τ behaves as $\exp(-|\Delta U_\tau|)$ in its tails, we see that the PDF for Q_τ has tails of the form $\exp(-|Q_\tau-\langle Q_\tau\rangle|)$, in agreement with Eqs. (63) and (65).

(6) We saw in Ref. [12] that the work obeys the conventional FT in the limit $\tau\to\infty$. Heat and work fluctuations are expected to behave similarly for small fluctuations (cf. point 5), and therefore the conventional FT is obeyed also by the heat fluctuations for $\tau\to\infty$ for *small* enough fluctuations. However, for *large* fluctuations, we get a different behavior. Using $P(Q_\tau)\sim\exp(-|Q_\tau-\langle Q_\tau\rangle|)$, the fluctuation function f_j^Q which is defined by Eq. (44) and can be written as $\ln[P(Q_\tau)/P(-Q_\tau)]/\langle Q_\tau\rangle$ becomes $(-|Q_\tau-\langle Q_\tau\rangle|+|Q_\tau +\langle Q_\tau\rangle|)/\langle Q_\tau\rangle=2$. This explains qualitatively the behavior expressed in Eq. (68).

(7) The symmetry relation in Eq. (40) is very reminiscent of the one used by Lebowitz and Spohn [6] in their work on the fluctuation theorem [i.e., $e(\lambda)=e(1-\lambda)$], although their method of large deviation theory allows only a treatment of the behavior of P_j^Q and f_j^Q for $\tau\to\infty$. The precise connection between their models (and the conventional FT they find), and our model (and the extended FTs) is not clear.

(8) In dynamical systems, microscopic reversibility and chaoticity are the requirements for the conventional FTs [1,2,3]. According to Lebowitz and Spohn [6], for stochastic systems, one only needs that if a state A can be reached from a state B, the reverse process from B to A can also occur. The SSFT then holds for an abstract "action functional," which resembles the entropy production if detailed balance holds. Our model is stochastic and does obey detailed balance, but it does not fall precisely into one of the classes discussed in Ref. [6] (mainly due to the infinite available volume in our case rather than the torus in Ref. [6]). It is presently unclear to what extent our model is representative of a larger class of systems in which heat fluctuations obey the extended FTs.

(9) We derived the extended infinite-τ FT of Eq. (68) already in Ref. [13] using large deviation theory. In fact, the saddle-point method reduces to the large deviation theory of that paper in the limit $\tau\to\infty$. We will not prove this here, but remark that if one takes $\tau\to\infty$ in Eq. (54), then one gets $-q(i-q)$. Setting $q=i\lambda$, this becomes $\lambda(1-\lambda)$, which is the form of the quantity $e(\lambda)$ used in the large deviation theory in Ref. [13] for $|\lambda|<1$. The expression breaks down when the correction terms to $-q(i-q)$ in Eq. (54) become infinite, i.e., at $\lambda=\pm1$. This restriction of $|\lambda|<1$ was also crucial in obtaining the extended FT in Ref. [13]. We remark that likewise, in Fig. 4, the singularities at $q=\pm i$ restrict the saddle point to the region $|q|<1$, and that this in turn leads to exponential tails of $P(Q_\tau)$, which finally give rise to the extensions of the heat FT.

(10) One of the important and striking results of the extended FTs is that the probability ratio for negative to positive fluctuations in the heat production by the Brownian particle is much larger than that given by the conventional FTs.

ACKNOWLEDGMENT

This work has been supported by the Office of Basic Engineering of the U.S. Department of Energy, under Grant No. DE-FG-02-88-ER13847.

APPENDIX A: PDF OF WORK OVER A TIME INTERVAL AND ENDPOINT POSITIONS

Here, we will determine the Gaussian joint distributions P_j^* of the work W_τ over a time interval of length τ from time t to $t+\tau$ and the positions of the Brownian particle at the beginning, $\Delta\mathbf{x}_t$, and the end, $\Delta\mathbf{x}_{t+\tau}$, where

$$\Delta\mathbf{x}_t=\mathbf{x}_t-\mathbf{v}^*t. \tag{A1}$$

These are needed for the numerical sampling method in Sec. III as well as in the calculating of the Fourier transform of P_j^Q in Eq. (36) which is evaluated in Appendix B. We are interested both in the transient case, for which $j=\mathrm{T}$ and $t=0$, and in the stationary state, for which $j=\mathrm{S}$ and $t\to\infty$.

For notational convenience, we introduce a seven dimensional vector $\mathbf{a}=(W_\tau,\Delta\mathbf{x}_1,\Delta\mathbf{x}_2)$. The PDF P_j^* is characterized by the moments

$$\overline{\mathbf{a}_j}=\int d\mathbf{a}\ P_j^*(\mathbf{a};\tau)\mathbf{a}, \tag{A2}$$

which is a seven-dimensional vector, and

$$\mathsf{A}=\int d\mathbf{a}\ P_j^*(\mathbf{a};\tau)(\mathbf{a}-\overline{\mathbf{a}_j})(\mathbf{a}-\overline{\mathbf{a}_j})^\dagger. \tag{A3}$$

Here the superscript dagger ($\dagger$) denotes the transpose. It is clear from this definition that A is a real symmetric 7×7 matrix. Also, it has $\det\mathsf{A}\geqslant0$.

Once these moments are known, P^* is given by

$$P_j^*(\mathbf{a};\tau) = \frac{e^{-(1/2)(\mathbf{a}-\overline{\mathbf{a}_j})^\dagger \cdot \mathsf{A}^{-1}\cdot(\mathbf{a}-\overline{\mathbf{a}_j})}}{\sqrt{\det(2\pi\mathsf{A})}}. \tag{A4}$$

Note that if $\det \mathsf{A}=0$, then the PDF P^* is a δ function (in one or more directions).

To give the specific form for $\mathbf{a}$ and A, we first write

$$\overline{\mathbf{a}_j} = \begin{pmatrix} \overline{\mathbf{a}_j^{(1)}} \\ \overline{\mathbf{a}_j^{(2)}} \\ \overline{\mathbf{a}_j^{(3)}} \end{pmatrix} \tag{A5}$$

(where $\overline{\mathbf{a}_j^{(1)}}$ is a scalar and $\overline{\mathbf{a}_j^{(2)}}$ and $\overline{\mathbf{a}_j^{(3)}}$ are three vectors) and

$$\mathsf{A} = \begin{pmatrix} A_{11} & \mathbf{A}_{21}^\dagger & \mathbf{A}_{31}^\dagger \\ \mathbf{A}_{21} & \mathsf{A}_{22} & \mathsf{A}_{32}^\dagger \\ \mathbf{A}_{31} & \mathsf{A}_{32} & \mathsf{A}_{33} \end{pmatrix} \tag{A6}$$

(where A_{11} is a scalar, $\mathbf{A}_{21}$ and $\mathbf{A}_{31}$ are three vectors, A_{22}, A_{23}, and A_{33} are 3×3 matrices).

The specific form for $\overline{\mathbf{a}_j}$ for the transient case, i.e., $\overline{\mathbf{a}_\mathrm{T}}$ is obtained using Eqs. (8), (18), (30), (31), (A1), (A2), and (A5), which yield

$$\overline{\mathrm{a}_\mathrm{T}^{(1)}} = \langle W_\tau\rangle_\mathrm{T} = w(\tau - 1 + e^{-\tau}), \tag{A7}$$

$$\overline{\mathbf{a}_\mathrm{T}^{(2)}} = \langle \Delta\mathbf{x}_0\rangle = 0, \tag{A8}$$

$$\overline{\mathbf{a}_\mathrm{T}^{(3)}} = \langle \Delta\mathbf{x}_\tau\rangle = (e^{-\tau}-1)\mathbf{v}^*. \tag{A9}$$

The subelements of A for the transient case are explicitly determined from Eqs. (9)–(11), (19), (30), (31), (A1), (A2), and (A5)–(A9), giving

$$A_{11} = \langle [W_\tau - \langle W_\tau\rangle]^2\rangle_\mathrm{T} = 2w(\tau - 1 + e^{-\tau}), \tag{A10}$$

$$\mathbf{A}_{21} = \langle [\Delta\mathbf{x}_0 - \langle\Delta\mathbf{x}_0\rangle][W_\tau - \langle W_\tau\rangle]\rangle = (e^{-\tau}-1)\mathbf{v}^*, \tag{A11}$$

$$\mathbf{A}_{31} = \langle [\Delta\mathbf{x}_\tau - \langle\Delta\mathbf{x}_\tau\rangle][W_\tau - \langle W_\tau\rangle]\rangle = (e^{-\tau}-1)\mathbf{v}^*, \tag{A12}$$

$$\mathsf{A}_{22} = \langle [\Delta\mathbf{x}_0 - \langle\Delta\mathbf{x}_0\rangle][\Delta\mathbf{x}_0 - \langle\Delta\mathbf{x}_0\rangle]^\dagger\rangle = \mathbb{1}, \tag{A13}$$

$$\mathsf{A}_{33} = \langle [\Delta\mathbf{x}_\tau - \langle\Delta\mathbf{x}_\tau\rangle][\Delta\mathbf{x}_\tau - \langle\Delta\mathbf{x}_\tau\rangle]^\dagger\rangle = \mathbb{1}, \tag{A14}$$

$$\mathsf{A}_{32} = \langle [\Delta\mathbf{x}_\tau - \langle\Delta\mathbf{x}_\tau\rangle][\Delta\mathbf{x}_0 - \langle\Delta\mathbf{x}_0\rangle]^\dagger\rangle = e^{-\tau}\mathbb{1}. \tag{A15}$$

For the stationary case, the specific forms for the components of $\overline{\mathbf{a}_\mathrm{S}}$ are found from Eqs. (8), (24), (33), (34), (A1), (A2), and (A5):

$$\overline{\mathrm{a}_\mathrm{S}^{(1)}} = \langle W_\tau\rangle_\mathrm{S} = w\tau, \tag{A16}$$

$$\overline{\mathbf{a}_\mathrm{S}^{(2)}} = \lim_{t\to\infty}\langle \Delta\mathbf{x}_t\rangle = -\mathbf{v}^*, \tag{A17}$$

$$\overline{\mathbf{a}_\mathrm{S}^{(3)}} = \lim_{t\to\infty}\langle \Delta\mathbf{x}_{t+\tau}\rangle = -\mathbf{v}^*, \tag{A18}$$

while the subelements of A are in that case, by Eqs. (33), (34), (A2), (A5), and (A6),

$$A_{11} = \lim_{t\to\infty}\langle [W_\tau - \langle W_\tau\rangle]^2\rangle = \langle [W_\tau - \langle W_\tau\rangle]^2\rangle_\mathrm{S}, \tag{A19}$$

$$\mathbf{A}_{21} = \lim_{t\to\infty}\langle [\Delta\mathbf{x}_t - \langle\Delta\mathbf{x}_t\rangle][W_\tau - \langle W_\tau\rangle]\rangle, \tag{A20}$$

$$\mathbf{A}_{31} = \lim_{t\to\infty}\langle [\Delta\mathbf{x}_{t+\tau} - \langle\Delta\mathbf{x}_{t+\tau}\rangle][W_\tau - \langle W_\tau\rangle]\rangle, \tag{A21}$$

$$\mathsf{A}_{22} = \lim_{t\to\infty}\langle [\Delta\mathbf{x}_t - \langle\Delta\mathbf{x}_t\rangle][\Delta\mathbf{x}_t - \langle\Delta\mathbf{x}_t\rangle]^\dagger\rangle, \tag{A22}$$

$$\mathsf{A}_{33} = \lim_{t\to\infty}\langle [\Delta\mathbf{x}_{t+\tau} - \langle\Delta\mathbf{x}_{t+\tau}\rangle][\Delta\mathbf{x}_{t+\tau} - \langle\Delta\mathbf{x}_{t+\tau}\rangle]^\dagger\rangle, \tag{A23}$$

$$\mathsf{A}_{32} = \lim_{t\to\infty}\langle [\Delta\mathbf{x}_{t+\tau} - \langle\Delta\mathbf{x}_{t+\tau}\rangle][\Delta\mathbf{x}_t - \langle\Delta\mathbf{x}_t\rangle]^\dagger\rangle, \tag{A24}$$

which turn out to be identical to those of the transient case in Eqs. (A10)–(A15) upon direct evaluation using Eqs. (9)–(11), (24), (A1), and (A16)–(A18). This is why we did not denote a j dependence of A.

APPENDIX B: FOURIER TRANSFORM OF π_j^Q

The Fourier transform of the PDF of heat will be calculated here, starting from Eq. (36). To calculate the $\hat{P}_j(q;\tau)$ from that equation, we define the quantities

$$\mathbf{c} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{B1}$$

$$\mathsf{B} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{B2}$$

so that one can write the exponent in Eq. (36) as $W_\tau - \frac{1}{2}\{|\Delta\mathbf{x}_2|^2 - |\Delta\mathbf{x}_1|^2\} = \mathbf{c}\cdot\mathbf{a} + \frac{1}{2}\mathbf{a}^\dagger\cdot\mathsf{B}\cdot\mathbf{a}$. Using Eq. (A4), one obtains then

$$\hat{P}_j(q;\tau) = \int d\mathbf{a}\frac{e^{-(1/2)(\mathbf{a}-\overline{\mathbf{a}_j})^\dagger\cdot\mathsf{A}^{-1}\cdot(\mathbf{a}-\overline{\mathbf{a}_j})+(1/2)iq\mathbf{a}^\dagger\cdot\mathsf{B}\cdot\mathbf{a}+iq\mathbf{c}\cdot\mathbf{a}}}{\sqrt{\det(2\pi A)}}. \tag{B3}$$

To evaluate this, the exponent is first rewritten as

$$\begin{aligned} &-\tfrac{1}{2}(\mathbf{a}-\overline{\mathbf{a}_j})^\dagger\cdot\mathsf{A}^{-1}\cdot(\mathbf{a}-\overline{\mathbf{a}_j}) + \tfrac{1}{2}iq\mathbf{a}^\dagger\cdot\mathsf{B}\cdot\mathbf{a} + iq\mathbf{c}\cdot\mathbf{a} \\ &\quad = -\tfrac{1}{2}(\mathbf{a}-\mathbf{a}_j')^\dagger\cdot[\mathsf{A}^{-1} - iq\mathsf{B}]\cdot(\mathbf{a}-\mathbf{a}_j') + d_j, \end{aligned} \tag{B4}$$

where $\mathbf{a}_j' = [\mathsf{I} - iq\,\mathsf{A}\cdot\mathsf{B}]^{-1}\cdot(\overline{\mathbf{a}_j} + iq\,\mathsf{A}\cdot\mathbf{c})$ and

$$d_j = \frac{iq}{2}[(\mathsf{B}\cdot\overline{\mathbf{a}_j} + \mathbf{c})^\dagger \cdot (\mathsf{I} - iq\ \mathsf{A}\cdot\mathsf{B})^{-1}\cdot(\overline{\mathbf{a}_j} + iq\ \mathsf{A}\cdot\mathbf{c}) + \overline{\mathbf{a}_j}\cdot\mathbf{c}]. \tag{B5}$$

Here, I is the 7×7 identity matrix. Then substituting Eq. (B4) in Eq. (B3) and changing the integration variable to $\mathbf{x} = \mathbf{a} - \mathbf{a}_j'$ yields

$$\hat{P}_j(q;\tau) = \int d\mathbf{x}\ e^{-(1/2)\mathbf{x}^\dagger\cdot(\mathsf{A}^{-1} - iq\mathsf{B})\cdot\mathbf{x}}\frac{e^{d_j}}{\sqrt{\det(2\pi\mathsf{A})}} \tag{B6}$$

$$= \frac{e^{d_j}}{\sqrt{\det(\mathsf{I} - iq\ \mathsf{A}\cdot\mathsf{B})}}, \tag{B7}$$

where the identity $\det(\mathsf{A})\det(\mathsf{B}) = \det(\mathsf{AB})$ has been used. To make Eq. (B7) into an explicit expression for $\hat{P}_j$, the inverse of the matrix $(\mathsf{I} - iq\mathsf{A}\cdot\mathsf{B})$ is required in the expression for d_j in Eq. (B5), and in Eq. (B7), its determinant. These are obtained as follows. Using Eqs. (A6), (A10)–(A15), and (B2), it follows that

$$\mathsf{I} - iq\ \mathsf{A}\cdot\mathsf{B} = \begin{pmatrix} 1 & iq(e^{-\tau}-1)\mathbf{v}^{*\dagger} & iq(e^{-\tau}-1)\mathbf{v}^{*\dagger} \\ \mathbf{0} & (\mathbb{1} - iq)\mathbb{1} & iqe^{-\tau}\mathbb{1} \\ \mathbf{0} & -iqe^{-\tau}\mathbb{1} & (1+iq)\mathbb{1} \end{pmatrix}. \tag{B8}$$

The determinant of this matrix is

$$\det(\mathsf{I} - iq\ \mathsf{A}\cdot\mathsf{B}) = [1 + (1-e^{-2\tau})q^2]^3. \tag{B9}$$

For the inverse of Eq. (B8), we get

$$(\mathsf{I} - iq\ \mathsf{A}\cdot\mathsf{B})^{-1} = \begin{pmatrix} 1 & \dfrac{iq(e^{-\tau}-1)[1+iq(1-e^{-\tau})]\mathbf{v}^{*\dagger}}{1+(1-e^{-2\tau})q^2} & \dfrac{iq(1-e^{-\tau})[1-iq(1-e^{-\tau})]\mathbf{v}^{*\dagger}}{1+(1-e^{-2\tau})q^2} \\ \mathbf{0} & \dfrac{1+iq}{1+(1-e^{-2\tau})q^2}\mathbb{1} & \dfrac{-iqe^{-\tau}}{1+(1-e^{-2\tau})q^2}\mathbb{1} \\ \mathbf{0} & \dfrac{iqe^{-\tau}}{1+(1-e^{-2\tau})q^2}\mathbb{1} & \dfrac{1-iq}{1+(1-e^{-2\tau})q^2}\mathbb{1} \end{pmatrix}. \tag{B10}$$

We now have the material needed to calculate d_j from Eq. (B5) explicitly. To calculate Eq. (B5), we use Eqs. (A5), (A6), (A10)–(A12), (B1), (B1), (B2), and (B10), to find, after some rearrangements,

$$\begin{aligned} d_j = iq\Bigg\{ \frac{1}{1+(1-e^{-2\tau})q^2}\Bigg[& -iq^3 w(1-e^{-\tau})^3 - iqe^{-\tau}\overline{\mathbf{a}}_j^{(2)}\cdot\overline{\mathbf{a}}_j^{(3)} \\ & + \frac{1}{2}|\overline{\mathbf{a}}_j^{(2)}|^2(1+iq) - \frac{1}{2}|\overline{\mathbf{a}}_j^{(3)}|^2(1-iq) - iq(1-e^{-\tau})[1+iq \\ & \times(1-e^{-\tau})]\mathbf{v}^*\cdot\overline{\mathbf{a}}_j^{(2)} + iq(1-e^{-\tau})[1-iq(1 \\ & - e^{-\tau})]\mathbf{v}^*\cdot\overline{\mathbf{a}}_j^{(3)}\Bigg] + \overline{\mathbf{a}}_j^{(1)} + iqw(\tau - 1 + e^{-\tau})\Bigg\}. \end{aligned} \tag{B11}$$

Furthermore, from Eqs. (A7)–(A9) and (B11), d_{T} for the transient case follows as

$$d_{\mathrm{T}} = wq(i-q)\left\{\tau - \frac{[1-e^{-\tau}]\left[1+\left(\frac{1}{2}+2q^2\right)(1-e^{-\tau})\right]}{1+(1-e^{-2\tau})q^2}\right\} \tag{B12}$$

while using Eqs. (A16)–(A18) and (B11), d_{S} for the stationary case, is, after some rewriting, found to be

$$d_{\mathrm{S}} = wq(i-q)\left\{\tau - \frac{2q^2(1-e^{-\tau})^2}{1+q^2(1-e^{-2\tau})}\right\}. \tag{B13}$$

Finally, by Eqs. (B7) and (B9), these expressions for d_j yield for the Fourier transforms explicitly

$$\hat{P}_{\mathrm{T}}(q;\tau)\ .2 = \frac{\exp wq(i-q)\left(\tau - \frac{[1-e^{-\tau}][1+(1/2+2q^2)(1-e^{-\tau})]}{1+(1-e-2\tau)q^2}\right)}{[1+(1-e^{-2\tau})q^2]^{3/2}}, \tag{B14}$$

$$\hat{P}_{\mathrm{S}}(q;\tau) = \frac{\exp wq(i-q)\left(\tau - \frac{2q^2(1-e^{-\tau})^2}{1+(1-e-2\tau)q^2}\right)}{[1+(1-e^{-2\tau})q^2]^{3/2}}. \tag{B15}$$

[1] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).

[2] D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994).

[3] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995); J. Stat. Phys. **80**, 931 (1995).

[4] E. G. D. Cohen, Physica A **240**, 43 (1997).

[5] J. Kurchan, J. Phys. A **31**, 3719 (1998).

[6] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).

[7] The dynamical systems are required to have an isoenergetic Gaussian thermostat for this to be strictly true, otherwise correction terms appear, though these might possibly vanish in the large time or system size limit.

[8] E. G. D. Cohen and G. Gallavotti, J. Stat. Phys. **96**, 1343 (1999).

[9] Also, the integrated TFT and SSFT exist [11,12], which we do not discuss here since they are constructed only to improve the statistics in experiments or simulations.

[10] S. Ciliberto and C. Laroche, J. Phys. IV **8**(6), 215 (1998).

[11] G. M. Wang *et al.*, Phys. Rev. Lett. **89**, 050601 (2002).

[12] R. van Zon and E. G. D. Cohen, Phys. Rev. E **67**, 046102 (2003).

[13] R. van Zon and E. G. D. Cohen, Phys. Rev. Lett. **91**, 110601 (2003).

[14] O. Mazonka and C. Jarzynski, e-print cond-mat/9912121.

[15] In Eq. (12) the possibility of energy being stored as kinetic energy of the Brownian particle has been neglected. While this in principle should be included, fluctuations in velocity decay on a time scale $\ll 1$ in the overdamped case, so that we may neglect them on time scales of $O(1)$.

[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran, The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).

[17] H. Jeffreys and B. S. Jeffreys, *Methods of Mathematical Physics*, 3rd ed. (Cambrige University Press, Cambridge, 1956); P. Dennery and A. Krzywicki, *Mathematics for Physicists* (Harper and Row, New York, 1967). L. Sirovich, *Techniques of Asymptotic Analysis* (Springer-Verlag, New York, 1971).

[18] We have verified using the numerical fast Fourier transform inverse of Sec. III, that neglecting the exponential terms in the Fourier transform $\hat{P}_j$ in Eq. (37) has negligible consequences for π_j^Q in Eq. (38) if $\tau \gtrsim 3$.

[19] With exponential terms neglected in Eq. (37), $P_j(q;\tau)$ behaves as $\exp[-(\tau-2)q^2]$ for large q, so that on the rhs of Eq. (38) the integral for the Fourier inverse diverges for $\tau<2$. This is an artifact of neglecting the exponentials.

[20] S. Wolfram, *The Mathematica Book* (Cambridge University Press, Cambridge, 1999).

[21] It is useful to have an analytic result for the saddle points in MATHEMATICA's memory because this shows that q^* is a smooth function of p and allows for the derivative $\partial q^*/\partial p$ needed in Eq. (53) to be obtained as well.

[22] Corrections to the position of the saddle point are here of higher order. To see this, let us split up h in Eq. (49) as $h_0 + \delta h$, and the position of the maximum of h as $q^* = x + \delta x$. Assuming δx and δh to scale as $\tau^{-\alpha}$ and $\tau^{-\beta}$, respectively, with $\alpha, \beta > 0$, we expand $h(q^*) = h(x) + \delta h'(x)\delta x + h_0''(x)\delta x^2/2 + \delta h''(x)\delta x^2/2 + \cdots$, using $h_0'(x) = 0$. The terms in this expansion are $O(1)$, $O(\tau^{-\alpha-\beta})$, $O(\tau^{-2\alpha})$, and $O(\tau^{-2\alpha-\beta})$, respectively, so that $h(q^*) = h(x) + O(\tau^{-\alpha-\min(\alpha,\beta)})$.

[23] In case (b), the zeroth order i for q^* is to zeroth order a maximum of h at the *boundary* of the interval from $-i$ to i, implying that its derivative need not be zero there. Hence, the argument in [22] does not go through and we need to keep the $O(\tau^{-1})$ correction in Eq. (61).

[24] In general, in d dimensions, the corresponding expressions are found by changing $3/2$ to $d/2$ in the denominator on the rhs of Eq. (37) and carrying through the calculation. For the fluctuation function, this amounts to replacing the prefactor "3" of the logarithm by d in the first line of Eq. (66) and to adding prefactors $(d-1)/2$ to the logarithmic terms $\ln \tau/(w\tau)$ and $\ln \tau/(2w\tau)$ in its second and third lines, respectively. In Eq. (67), the last term remains the same, but the first two become $-\frac{1}{2}\ln[w^2 d^{1-d}(3-p)^d(1+p)^d(p-1)^{d-2}/64]$ and $-d/2$, resp. Finally, in Eq. (69) $1/w$ becomes $d/(3w)$.

[25] J. Farago, J. Stat. Phys. **107**, 781 (2002).

[26] L. Rey-Bellet and L. E. Thomas, Ann. Henri Poincare **3**, 483 (2002); L. Rey-Bellet, e-print math-ph/0303021.