Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 340 (2004) 66–75
PHYSICA A
www.elsevier.com/locate/physa

Non-equilibrium thermodynamics and fluctuations

R. van Zon*, E.G.D. Cohen

The Rockefeller University, 1230 York Avenue, NY 10021, USA

Abstract

In the last 10 years, a number of "conventional fluctuation theorems" have been derived for systems with deterministic or stochastic dynamics, in a transient or in a non-equilibrium stationary state. These theorems gave explicit expressions for the ratio of the probability to find the system with a certain value of entropy (or heat) production to that of finding the opposite value. A similar theorem for the fluctuations of the *work* done on a system has recently been demonstrated experimentally for a simple system in a transient state, consisting of a Brownian particle in water, confined by a moving harmonic potential. In this paper, we show that because of the interaction between the stochastic motion of the particle in water and its deterministic motion in the potential, very different new *heat* theorems are found than in the conventional case. One of the consequences of these new heat fluctuation theorems is that the ratio of the probability for the Brownian particle to absorb heat from rather than supply heat to the water is much larger than in the conventional fluctuation theorems. This could be of relevance for micro/nano-technology.

PACS: 05.40.−a; 05.70.−a; 44.05.+e; 02.50.−r

Keywords: Fluctuation theorems; Thermodynamics; Brownian motion

1. Introduction

As is well known the first and second law of thermodynamics only involve averages of the physical quantities of macroscopic systems but say nothing about their fluctuations. In particular, the second law for irreversible processes states that the average[1] heat $\bar{Q}$ internally produced in an irreversible process has to be positive. In the last 10 years a number of fluctuation theorems have been derived for the fluctuations of thermodynamic properties in non-equilibrium stationary [1,2], as well as transient states

* Corresponding author.

[1] Here, as well as in the rest of the paper, a bar means an average.

0378-4371/$ - see front matter
doi:10.1016/j.physa.2004.03.078

[3,4], which constitute refinements of the laws of thermodynamics in so far that they take into account a property of their fluctuations which goes far beyond the statement that $\bar{Q} > 0$. In this respect a generalization of thermodynamics including fluctuations has been in progress.

It should be pointed out that these new fluctuation properties are valid for large fluctuations around a non-equilibrium stationary state possibly far from equilibrium. As such they differ from fluctuations dealt with in the context of irreversible thermodynamics, such as the fluctuation dissipation theorem and the Onsager relations (as treated e.g. in the classical book by De Groot and Mazur [5]), which refer to small fluctuations around equilibrium.

The fluctuation theorems (FT) will be divided into two classes: conventional (CFT) and new (NFT) fluctuation theorems, where in the latter case the word "theorem" is premature. We will not confine ourselves here to heat fluctuations alone but also consider work and energy fluctuations. In the literature until now the overwhelming number of papers has dealt exclusively with heat in the form of entropy production as it occurs in the second law.

1.1. Conventional FT

The first discovery of a FT of the kind we will discuss here was mainly numerical in a computer simulation by Evans et al. (in 1993) [1]. This FT for a non-equilibrium stationary state was inspired by dynamical systems theory notions [2]. A similar FT for a transient state was formulated by Evans and Searles (in 1994) [3]. In this paper, we will restrict ourselves to the non-equilibrium Stationary State Fluctuation Theorem rather than to the transient fluctuation theorem.

The CFT deals with the fluctuations of work as well as heat in a finite dynamical system. The many particle (Hamiltonian) system is subject to an external force, which does work on the system. However, this would heat up the system if it were not for an ingenious internal thermostat, realized by adding a damping term to the equations of motion. The dynamics of this system is purely *deterministic*. The dissipation taking place in the system is manifested in a contraction of the accessible phase space of the system which can be related to a (generalized) physical entropy production in the system.[2] A FT was derived under a number of assumptions for such a system for the heat or entropy production, which will be given below. Some years later, Kurchan (in 1998) [6] and Lebowitz and Spohn (in 1999) [7] derived a similar FT for a system with purely *stochastic* dynamics under a number of assumptions.

In both cases a CFT was found, which can be written in the form

$$\frac{P(Q_\tau)}{P(-Q_\tau)} \underset{\tau\to\infty}{\to} \mathrm{e}^{\beta Q_\tau} \,. \tag{1.1}$$

[2] 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.

Here $P(Q_\tau)$ is the probability that the fluctuating heat *produced* in the system during a time τ has a value Q_τ. Therefore, $-Q_\tau$ is a value of the fluctuating heat *absorbed* by the system of the same magnitude (during an equally long time). In taking the limit $\tau \to \infty$ in Eq. (1.1), $Q_\tau \sim \tau$ scales as τ. The Q_τ (and later the work W_τ) in the stationary state can be visualized as fluctuations on segments of duration τ, obtained by cutting a very long stationary state trajectory of the system in phase space into segments.

For systems with deterministic dynamics the *same* CFT holds for the work done in a time τ denoted by W_τ, i.e., Q_τ in Eq. (1.1) can be replaced by W_τ. The reason is that Q_τ and W_τ are represented by the same mathematical expression here, due to the fact that the thermostat converts all external work done on the system into internal heat.

There is a connection between this CFT and theorems discussed in the context of irreversible thermodynamics. Gallavotti proved for deterministic dynamical systems, that near equilibrium, Eq. (1.1) leads to the Onsager relations, the fluctuation dissipation theorem and the Green–Kubo relations for the transport coefficients [8]. This seems to imply that the fluctuations incorporated in Eq. (1.1) go beyond irreversible thermodynamics, i.e., beyond the linear, near equilibrium regime. There is no estimate available for the range of validity of irreversible thermodynamics nor are there any results in this nonlinear regime to date.

1.2. The system

We are interested in this paper in a quasi-many particle system: a Brownian particle suspended in water and restricted in its motion by a laser-induced harmonic potential, which is pulled through the water with a constant velocity $\mathbf{v}^*$ (cf. Fig. 1). This system was introduced earlier in a somewhat different context by Wang et al. (in 2002) [9].

Contrary to the systems in Section 1.1, work and heat are not identical for this system. The fluctuations of the work done on it as well as of the heat produced by it in a time τ were computed based on an overdamped Langevin equation [10,11]. Although this system is in principle a many particle system, the many degrees of freedom of the water have been contracted to those of the (Stokes) friction of the Brownian particle and the strength of its assumed Gaussian white noise. In dimensionless units,[3] the Langevin equation reads

$$0 = -\dot{\mathbf{x}}_t - (\mathbf{x}_t - \mathbf{x}_t^*) + \boldsymbol{\zeta}_t \,. \tag{1.2}$$

Here, the first term on the right-hand side represents the friction of the Brownian particle in the water, the second term represents the harmonic force due to the harmonic potential

$$U_t(\mathbf{x}_t) = \tfrac{1}{2}(\mathbf{x}_t - \mathbf{x}_t^*)^2 \tag{1.3}$$

while the third term represents the fluctuations of the Brownian particle due to the thermal motion of the water molecules. $\mathbf{x}_t$ and $\mathbf{x}_t^*$ are the positions of the Brownian particle and the minimum of the potential at time t, respectively (cf. Fig. 1). Essential is that contrary to the pure dynamics of the two previous systems, satisfying the CFT,

[3] Compared to Ref. [10], this means the force constant k = friction constant, α = temperature, $k_B T = 1$.

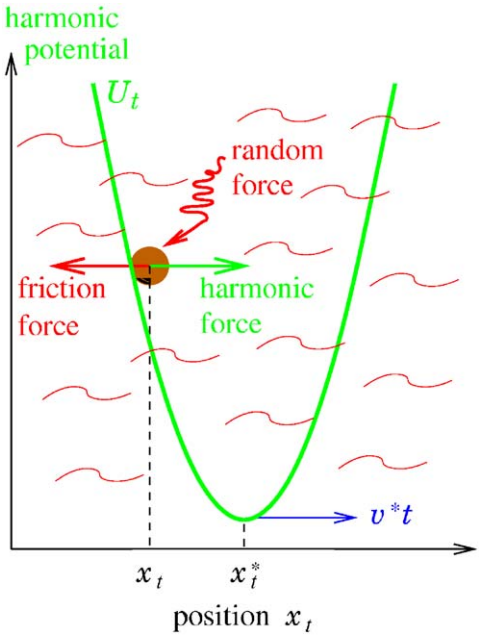


Fig. 1. Brownian particle in a moving harmonic potential (cf. Ref. [9]).

here a *mixed* dynamics occurs: *deterministic* due to the harmonic force and *stochastic* due to the water fluctuations. For this system, energy conservation reads as the first law of thermodynamics:

$$W_\tau = Q_\tau + \Delta U_\tau \,. \tag{1.4}$$

Here, W_τ is the total work done on the system during time τ, i.e., pulling it with a constant velocity $\mathbf{v}^*$ through the fluid over a time τ; Q_τ is the heat developed in the water due to the friction of the Brownian particle during time τ and $\Delta U_\tau = U_\tau(\mathbf{x}_\tau) - U_0(\mathbf{x}_0)$ is the potential energy difference of the particle in the harmonic potential in time τ. Eq. (1.4) clearly shows the difference between work and heat in this system. Physically, the pulling of the harmonic potential drags the Brownian particle along, but with a delay, because of its friction with the water, while at the same time this delay necessitates a change in its potential energy from its initial position $\mathbf{x}_0$ to its final position $\mathbf{x}_\tau$. A non-equilibrium stationary state will be reached when the friction force cancels the attractive force on the particle due to the harmonic potential (cf. Fig. 1) and the fluctuations around this state will be studied.

In the non-equilibrium stationary state the averages of the thermodynamic quantities $\overline{W_\tau}$ and $\overline{Q_\tau}$ are equal, since $\overline{\Delta U_\tau} = \overline{U_{t+\tau}} - \overline{U_t} = 0$ then. However, unlike in the cases of pure (deterministic or stochastic) dynamics, $P(Q_\tau) \neq P(W_\tau)$, which is the main topic of this paper.

In the following two sections, we first sketch how the distribution functions for W_τ, Q_τ and ΔU_τ for such a system are obtained, after which the fluctuation theorem,

involving the ratios of the probabilities of W_τ and $-W_\tau$, and of Q_τ and $-Q_\tau$ will be stated.

2. Distribution functions

2.1. Work

The probability distribution function of the work W_τ done on the above-described system during a time τ can be derived directly from the Langevin equation [10]. To do that, note first that since the Langevin equation (1.2) is linear in $\mathbf{x}_t$, the distribution function of $\mathbf{x}_t$ is Gaussian. (It is an Ornstein–Uhlenbeck process.) Secondly, since $W_\tau = -\mathbf{v}^* \cdot \int_0^\tau [\mathbf{x}_t - \mathbf{x}_t^*]\,\mathrm{d}t$ depends also linearly on $\mathbf{x}_t$, its distribution is Gaussian too. Therefore, it is completely determined by its first and second moments and can be written in the form

$$P(W_\tau) \sim \mathrm{e}^{-(W_\tau - \overline{W_\tau})^2/2\sigma^2}\,, \tag{2.1}$$

where the variance $\sigma^2 = \overline{(W_\tau - \bar{W}_\tau)^2}$. Here the averages are over all time segments of the trajectory of duration τ. The first and second moment can be computed from the solution of the Langevin equation (1.2) [10].

2.2. Heat

A similar simple direct derivation of $P(Q_\tau)$ as was done for $P(W_\tau)$ cannot be performed, since the Q_τ given by (cf. Eq. (1.4))

$$Q_\tau = W_\tau - \Delta U_\tau \tag{2.2}$$

is quadratic in $\mathbf{x}_t$ via the ΔU_τ (cf. Eq. (1.3)). A way to obtain nevertheless $P(Q_\tau)$ is a much more complicated procedure via its Fourier transform $\hat{P}_\tau(q)$:

$$P(Q_\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}q\, \hat{P}_\tau(q)\mathrm{e}^{-\mathrm{i}qQ_\tau}\,. \tag{2.3}$$

$\hat{P}_\tau(q)$ can be computed exactly for all τ (see Refs. [11,13]):

$$\hat{P}_\tau(q) = \int_{-\infty}^{\infty} \mathrm{d}Q_\tau\, P(Q_\tau)\mathrm{e}^{\mathrm{i}qQ_\tau} = \frac{\exp\left[q(\mathrm{i}-q)v^{*2}\{\tau - \frac{2q^2(1-\mathrm{e}^{-\tau})^2}{1+(1-\mathrm{e}^{-2\tau})q^2}\}\right]}{[1+(1-\mathrm{e}^{-2\tau})q^2]^{3/2}}\,. \tag{2.4}$$

Note that $\hat{P}_\tau(q)$ is a function in the complex plane with branch cuts (due to the square root in its denominator) and two singularities at $q_1 = \mathrm{i}(1-\mathrm{e}^{-2\tau})^{-1/2}$ and $q_2 = -q_1$, which are also the end points of the branch cuts from q_1 to $+\mathrm{i}\infty$ and q_2 to $-\mathrm{i}\infty$ (cf. Fig. 2). All singularities occur for imaginary q-values and introduce exponential (rather than Gaussian) tails in $P(Q_\tau)$ for large Q_τ.

The approximate evaluation of $P(Q_\tau)$ can be performed for large τ using the saddle point method (SPM) [12], capitalizing on the fact that Q_τ in the exponent on the right-hand side of Eq. (2.3) is proportional to τ. As seen in Fig. 2, the integral $\int_{-\infty}^{\infty}$ along the real axis R in Eq. (2.3) can, for every Q_τ, be deformed to a path of steepest

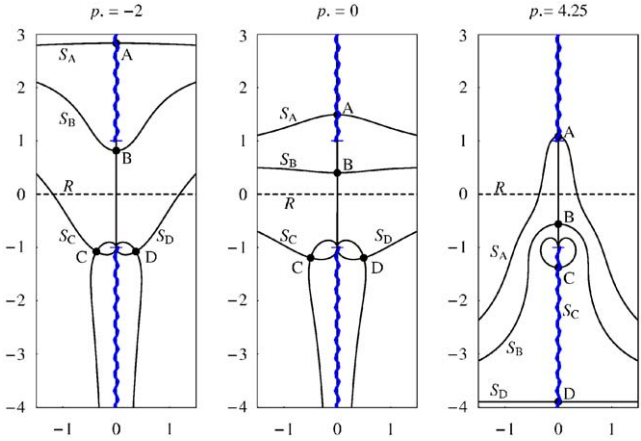


Fig. 2. Structure of the complex function $\hat{P}_\tau(q)e^{-iqQ_\tau}$ in the complex plane, with $p = Q_\tau/\overline{Q_\tau} = -2$, 0 and 4.25, respectively. The saddle points A, B, C and D are shown as dots. Wiggly curves along the imaginary axis are branch cuts, ending in the branch points q_1 (top cut) and q_2 (bottom cut) indicated with a thin horizontal line. The dashed line R is the real axis and S_A, S_B, S_C and S_D are paths of steepest descent though the saddle points, at which the function attains a maximum.

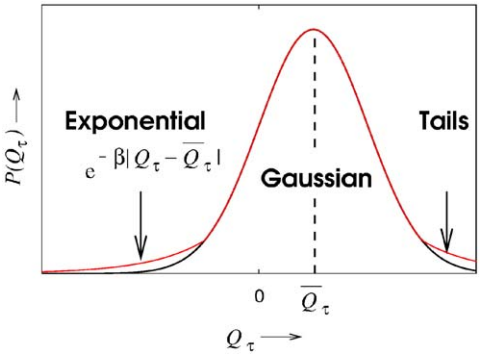


Fig. 3. Mixed Gaussian and exponential behavior of $P(Q_\tau)$ (sketch).

descent S_B in the complex plane that goes through the saddle point B, without passing through a singularity—which is not possible for the saddle points A, C and D. All that is needed to evaluate the integral along the real axis are then the properties of the function at the point B. For details we refer to a forthcoming paper [13].

To summarize the results of the distribution functions for the fluctuations of all three quantities occurring in Eq. (1.4):

1. $P(W_\tau) \sim e^{-(W_\tau-\bar{W}_\tau)^2/2\sigma^2}$ is Gaussian, from the Langevin Eq. (2.1),
2. $P(\Delta U_\tau) \sim e^{-\beta|\Delta U_\tau|}$ is exponential. This can be physically understood by observing that for large $\tau, \Delta U_\tau$ is the difference of two independent quantities, distributed as the

potential energy of a (Brownian) particle in a potential in contact with a heat bath, i.e., with a Boltzmann weight $e^{-\beta U}$.

3. Since $Q_\tau = W_\tau - \Delta U_\tau$, $P(Q_\tau)$ results from an interplay of the Gaussian $P(W_\tau)$ and the exponential $e^{-\beta \Delta U(\tau)}$. Thus the SPM leads to a mixed curve for $P(Q_\tau)$ which is Gaussian-like in the center, but has exponential tails (Fig. 3).

3. Fluctuation theorems

3.1. Work

We start by introducing a new formulation of the FT for $P(W_\tau)$ which is more precise than that given (for $P(Q_\tau)$) in Eq. (1.1). Taking the logarithm of both sides of Eq. (1.1), with W_τ instead of Q_τ, and dividing both sides then by $\bar{W}_\tau$, we obtain a reformulation of Eq. (1.1) of the form

$$\lim_{\tau\to\infty} F_\tau(p_W) = p_W \,, \tag{3.1}$$

where F_τ is the fluctuation function

$$F_\tau(p_W) = \frac{1}{\bar{W}_\tau} \ln \frac{P(W_\tau)}{P(-W_\tau)} \tag{3.2}$$

and $p_W = W_\tau/\bar{W}_\tau$ is a scaled work fluctuation. Eq. (3.1) expresses then the CFT for W_τ, proven in Ref. [10]. A new FT for *finite* τ for the work can be obtained from the Langevin equation (1.2) and Eq. (2.1), of the form [10]

$$F_\tau(p_W) = \frac{p_W}{1-\varepsilon(\tau)} \approx p_W + O\left(\frac{1}{\tau}\right) \,, \tag{3.3}$$

where $\varepsilon(\tau) = (1 - e^{-\tau})/\tau$. We note that all CFTs used so far in the literature restrict themselves to $\tau \to \infty$. In this model one can also discuss the finite τ behavior and its correction to the infinite time behavior, which shows that the slope of $F_\tau(p_W)$ is bigger than one for all finite τ (cf. Fig. 4).

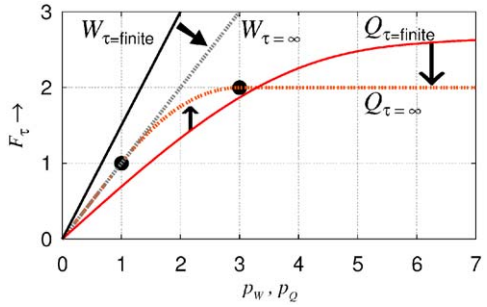


Fig. 4. CFT and NFTs for work and heat fluctuations.

3.2. Heat

Proceeding similarly with Eq. (1.1) for $P(Q_\tau)$ as in 3.1 for $P(W_\tau)$ one finds the CFT in a more satisfactory form

$$\lim_{\tau\to\infty} F_\tau(p_Q) = p_Q \,, \tag{3.4}$$

where $p_Q = Q_\tau/\overline{Q_\tau}$ is a scaled heat fluctuation, and

$$F_\tau(p_Q) = \frac{1}{\bar{Q}_\tau} \ln \frac{P(Q_\tau)}{P(-Q_\tau)} \,. \tag{3.5}$$

Contrary to Eq. (3.1) for W_τ, the relation in Eq. (3.4) is in fact incorrect, due to the interaction of $P(Q_\tau)$ with the (exponential) $P(\Delta U_\tau)$.

A NFT can be derived using the SPM [12], both for infinite and for finite τ. The behavior is determined by the above mentioned singularities in the complex plane in carrying out the SPM. The result for $F_\tau(p_Q)$ versus p_Q is given in Fig. 4, together with that of $F_\tau(p_W)$ versus p_W for comparison.

While $F_\tau(p_W)$ versus p_W is linear for all τ, the behavior of $F_\tau(p_Q)$ versus p_Q is much more complicated. In fact, for $\tau \to \infty$ there are three regimes [11]:

$$F_\tau(p_Q) = \begin{cases} p_Q & \text{for } 0 < p_Q < 1 \,, \\ p_Q - \dfrac{(1-p_Q)^2}{4} & \text{for } 1 < p_Q < 3 \,, \\ 2 & \text{for } p_Q > 3 \,. \end{cases} \tag{3.6}$$

Thus, for infinite τ, the NFT coincides with the CFT for small fluctuations $0 < p_Q < 1$, then exhibits a parabolic behavior between $1 < p_Q < 3$ and finally reaches a plateau, where $F_\tau(p_Q)=2$ for all $p_Q > 3$. The behavior for $p_Q < 0$ follows from the asymmetry of $F_\tau(p_Q)$ (cf. Eq. (3.5)).

The SPM also allows to study analytically the approach of the finite τ behavior to the infinite τ behavior, giving for large but finite τ:

$$F_\tau(p_Q) = \begin{cases} p_Q + \dfrac{h(p_Q)}{\tau} + O\left(\dfrac{1}{\tau^2}\right) & \text{for } |p_Q| < 1 \,, \\ 2 + \dfrac{g(p_Q)}{\sqrt{\tau}} + O\left(\dfrac{1}{\tau}\right) & \text{for } p_Q > 3 \,, \end{cases} \tag{3.7}$$

where

$$h(p) = \frac{8p}{9-p^2} - \frac{3}{2v^{*2}} \ln\left[\frac{(3-p)(1+p)}{(3+p)(1-p)}\right] \,. \tag{3.8}$$

$$g(p) = \sqrt{8(p-3)} \,. \tag{3.9}$$

Eqs. (3.7)–(3.9) show that the asymptotic behavior of $F_\tau(p_Q)$ for $p_Q > 3$, is a slowly increasing function $\sim \sqrt{p_Q - 3}$, while the asymptotic $\tau \to \infty$ curve is approached as $\tau^{-1/2}$ for $p_Q > 3$ and as τ^{-1} for $p_Q < 1$.

4. Discussion

In Table 1, the results for the CFT and the NFT for work and heat are summarized. A number of questions and remarks present themselves:

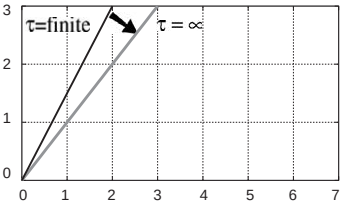
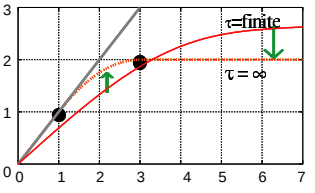
1. How general is the NFT for systems with mixed deterministic and stochastic dynamics? Is the Boltzmann factor and the ensuing exponential decay for large fluctuations more general than in this model? One would be inclined to think so, in view of the physical argument of point 2 in 2.2.

2. The relative probability for the Brownian particle to gain rather than supply heat to the water is much larger in the NFT than in the CFT (cf. Fig. 4). This might be of relevance in designing micro or nano-machines sensitive in their functioning to the heat absorbed during large fluctuations. The CFT would not be a good basis to judge this effect. We remark that since $F_\tau(p_Q) > 0$ for $p_Q > 0$ in the NFT, $\bar{Q}_\tau > 0$, in accordance with the second law.

3. The plateau value of 2 for $F(p_Q)$ for large p_Q (> 3) can be understood physically. For the probability for large Q_τ (i.e., p_Q), the exponential distribution $P(\Delta U_\tau) \sim \mathrm{e}^{-\beta|\Delta U_\tau|}$ dominates over the Gaussian distribution $P(W_\tau)$ (cf. Fig. 3). This implies, with $\bar{W}_\tau = \bar{Q}_\tau$ (see the end of Section 1.2), that $P(Q_\tau) \sim \mathrm{e}^{-\beta|Q_\tau - \bar{Q}_\tau|}$ so that $F_\tau(p_Q) = 2$, with $\beta = 1$.

4. Similar results as discussed in this paper are obtained for the Transient Fluctuation Theorem [3,4]. It is certainly, at least in this model, not the identity for all τ which obtains in the CFT [10,13].

Table 1
Results for the work (Section 3.1) and heat (Section 3.2) fluctuations

	Work	Heat
Symbol X_τ	W_τ	Q_τ
Distribution	Gaussian	Gauss.+Exp. tails
Fluctuation function	$F_\tau = \frac{1}{\bar{X}_\tau} \ln \frac{P(X_\tau)}{P(-X_\tau)}$	idem
$\tau \to \infty$	*Conventional* F_τ straight line with slope 1 for all W_τ	*New* F_τ has slope 1 for small Q_τ $F_\tau = 2$ for large Q_τ
τ finite	*New* F_τ straight line with slope > 1 for all W_τ	*New* F_τ no slope 1 for small Q_τ F_τ increasing for large Q_τ
Plots $F_\tau(p)$ versus p		

5. All our analytic results have been verified by comparison with the results of two numerical methods: a sampling method and a fast inverse Fourier transform of Eq. (2.3) [13]. In particular, the SPM turns out to give good results already for $\tau > 3$, whereas curves for $\tau \leqslant 3$ need to be obtained numerically.

6. The connection between the NFT and theorems of irreversible thermodynamics is unclear, although the linear (CFT-like) behavior of the NFT for small fluctuations with $0 < p_Q < 1$ (cf. Fig. 4) suggests that the same relations hold as for the CFT for small deviations near equilibrium.

7. So far, in all cases dealt with here, only one property of the fluctuations of the thermodynamic quantities work and heat in a non-equilibrium stationary state—has been discussed, viz. the fluctuation function F_τ. Whether something can be said about other properties of fluctuations of thermodynamic quantities, also beyond the linear regime, remains an interesting open question.

Acknowledgements

This research was supported by the Office of Basic Engineering of the US Department of Energy, under Grant No. DE-FG-02-88-ER13847.

References

[1] D.J. Evans, E.G.D. Cohen, G.P. Morriss, Phys. Rev. Lett. 71 (1993) 2401.
[2] G. Gallavotti, E.G.D. Cohen, Phys. Rev. Lett. 74 (1995) 2694;
G. Gallavotti, E.G.D. Cohen, J. Stat. Phys. 80 (1995) 931;
E.G.D. Cohen, Physica 240 (1997) 43.
[3] D.J. Evans, D.J. Searles, Phys. Rev. E 50 (1994) 1645.
[4] E.G.D. Cohen, G. Gallavotti, J. Stat. Phys. 96 (1999) 1343.
[5] S.R. de Groot, P. Mazur, Non-Equilibrium Thermodynamics, Dover Publications, Inc., New York, 1984.
[6] J. Kurchan, J. Phys. A, Math. Gen. 31 (1998) 3719.
[7] J.L. Lebowitz, H. Spohn, J. Stat. Phys. 95 (1999) 333.
[8] G. Gallavotti, Phys. Rev. Lett. 77 (1996) 4334.
[9] G.M. Wang, et al., Phys. Rev. Lett. 89 (2002) 050601.
[10] R. van Zon, E.G.D. Cohen, Phys. Rev. E 67 (2003) 046102.
[11] R. van Zon, E.G.D. Cohen, Phys. Rev. Lett. 91 (2003) 110601.
[12] H. Jeffreys, B.S. Jeffreys, Methods of Mathematical Physics, 3rd Edition, Cambridge University Press, Cambridge, 1956.
[13] R. van Zon, E.G.D. Cohen, arXiv: cond-mat/0311357, to appear in Phys. Rev. E.