



Work, dissipation, and fluctuations in nonequilibrium physics

Stationary state fluctuation theorems for driven Langevin systems

E.G.D. Cohen^a, Ramses van Zon^{a,b,*}

^a *The Rockefeller University, 1230 York Avenue, New York, NY 10021-6399, USA*

^b *Chemical Physics Theory Group, Department of Chemistry, University of Toronto, 80 St. George St., Toronto, ON M5S 1A1, Canada*

Received 9 March 2007

Available online 27 June 2007

Abstract

Recent results on the stationary state Fluctuation Theorems for work and heat fluctuations of Langevin systems are presented. The relevance of finite time corrections in understanding experimental and simulation results is explained in the context of an exactly solvable model, namely a Brownian particle in a harmonic potential, which is dragged through the surrounding fluid. In this model, work fluctuations obey the conventional form of the fluctuation theorem while heat fluctuations satisfy an extended form. The connection with other work in recent literature is pointed out, and further generalizations are suggested. **To cite this article:** *E.G.D. Cohen, R. van Zon, C. R. Physique 8 (2007).*

© 2007 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Résumé

Théorèmes de fluctuations dans l'état stationnaire pour les systèmes de Langevin entraînés. Des résultats récents sont présentés sur les théorèmes de fluctuations dans l'état stationnaire pour les fluctuations du travail et de la chaleur dans les systèmes de Langevin. L'importance des corrections de temps fini pour comprendre les résultats expérimentaux et de simulations est expliquée dans le contexte d'un modèle exactement soluble, à savoir une particule brownienne dans un potentiel harmonique qui est entraînée dans le fluide environnant. Dans ce modèle, les fluctuations du travail obéissent à la forme conventionnelle du théorème de fluctuations alors que les fluctuations de chaleur satisfont une forme étendue. La connexion à d'autres travaux de la littérature récente est mise en évidence et des généralisations ultérieures sont suggérées. **Pour citer cet article :** *E.G.D. Cohen, R. van Zon, C. R. Physique 8 (2007).*

© 2007 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

Keywords: Fluctuation theorems; Langevin systems

Mots-clés : Théorèmes de fluctuations ; Systèmes de Langevin

1. Introduction

In the last twelve years, the field of non-equilibrium statistical mechanics has been given a boost by the discovery of a number of new, general results regarding the fluctuations of thermodynamic quantities such as heat, work and

* Corresponding author at: Chemical Physics Theory Group, Department of Chemistry, University of Toronto, 80 St. George St., Toronto, ON M5S 1A1, Canada.

E-mail addresses: egdc@mail.rockefeller.edu (E.G.D. Cohen), rzon@chem.utoronto.ca (R. van Zon).

entropy. This paper is not intended as a survey article, rather, we will restrict ourselves to non-equilibrium stationary state Fluctuation Theorems (FTs). A stationary state FT was first discovered in [1]. This FT is a quantitative relation between the probability to observe a positive entropy production rate σ (in units of Boltzmann's constant k_B) versus a negative one $-\sigma$ over a time interval τ . Given the probability distribution π_τ of a fluctuating quantity p , scaled such that its average is equal to one, the Conventional Fluctuation Theorem (CFT) (also called the Gallavotti–Cohen Fluctuation Theorem) takes the form [2]

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \log \frac{\pi_\tau(p)}{\pi_\tau(-p)} = \sigma_+ p \quad \text{for } p < p^* \quad (1)$$

where σ_+ is the average entropy production rate over positive times, $p = \sigma/\sigma_+$, and p^* is an upper bound on the values of p for which the CFT holds.

Whether the CFT holds depends on the kind of system and on what the quantity p represents. Choices for p that have been used in the literature include the phase space contraction rate [2], a “dissipation function” [3], an “action functional” [4], and, more concretely, work, heat and entropy production. The value of p^* can also depend on the quantity under consideration. The stationary state Fluctuation Theorem holds quite generally (i.e. for a large class of deterministic [1–3] and stochastic systems [5,4]), and has been verified in laboratory experiments [6,7].

The behavior for values $p > p^*$ has recently been elucidated for specific systems, and led to an Extended Fluctuation Theorem (EFT) for heat [8–10], which we will discuss in detail below.

In this article, we will give an overview of some interesting results of the stationary state Fluctuation Theorem for an exactly solvable model of a Brownian particle in a harmonic potential which is dragged through a surrounding fluid. While this model is based on a Langevin equation, the conclusions from this model also pertain to certain purely deterministic systems, as we will point out.

We will first introduce the model in Section 2 and discuss its exact solution in Section 3. We will then turn our attention to thermodynamic quantities of the dissipated heat and the supplied work, and their fluctuations, both for finite and infinite time in Section 4. In Section 5 we will show that this results in two kinds of stationary state fluctuation relations: the Conventional Fluctuation Theorem (CFT), for work fluctuations, and an Extended Fluctuation Theorem (EFT), for heat fluctuations. The paper ends with a discussion in Section 6 of the results and generalizations that have appeared in the literature.

2. The model

Consider a spherical Brownian particle of mass m and diameter D suspended in an equilibrium fluid with viscosity η and temperature T . The motion of this Brownian particle can be described by the Langevin equation

$$m\ddot{\mathbf{x}} = -\alpha\mathbf{v} + \boldsymbol{\xi} \quad (2)$$

where \mathbf{x} is the position of the particle, $\mathbf{v} = d\mathbf{x}/dt$ its velocity and $\alpha = 3\pi\eta D$ the Stokes friction coefficient. Furthermore, $\boldsymbol{\xi} = \boldsymbol{\xi}(t)$ is a Gaussian fluctuating force whose average (over realizations) $\langle \boldsymbol{\xi}(t) \rangle = 0$ and whose time correlations are delta-correlated, i.e.

$$\langle \boldsymbol{\xi}(t_1)\boldsymbol{\xi}(t_2) \rangle = 2k_B T \alpha \delta(t_1 - t_2) \mathbf{1} \quad (3)$$

where k_B is Boltzmann's constant and $\mathbf{1}$ is the identity matrix.

Consider now the situation in which an external force acts on the Brownian particle (but not on the fluid). Under the assumption that the fluid is unperturbed such that α and the properties of $\boldsymbol{\xi}$ are not changed, the Langevin equation gets modified to

$$m\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t) - \alpha\mathbf{v} + \boldsymbol{\xi} \quad (4)$$

where $\mathbf{F}(\mathbf{x}, t)$ is the external force, which may depend on position and time.

Of particular interest here is the case in which the force derives from a moving harmonic potential U , i.e.,

$$\mathbf{F}(\mathbf{x}, t) = -\frac{\partial U(\mathbf{x}, t)}{\partial \mathbf{x}} = -\kappa(\mathbf{x} - \mathbf{x}^*(t)) \quad (5)$$

where the potential $U(\mathbf{x}, t) = \kappa|\mathbf{x} - \mathbf{x}^*(t)|^2/2$ and the center $\mathbf{x}^*(t)$ of the potential changes in time according to

$$\dot{\mathbf{x}}^*(t) = \mathbf{v}^* t \quad (6)$$

This corresponds to the particle being trapped in a harmonic well, the center \mathbf{x}^* of which moves with a constant velocity \mathbf{v}^* , to which we will restrict ourselves here. This closely resembles the experimental system of a Brownian particle in a moving optical trap studied by Wang et al. [11].

One also needs to specify the initial condition of the system. In time, the distribution of \mathbf{x} and \mathbf{v} approach a non-equilibrium stationary state distribution f_{ness} asymptotically. It is not hard to show that

$$f_{\text{ness}}(\mathbf{x}, \mathbf{v}) = \frac{\beta\sqrt{\kappa m}}{2\pi} \exp\left[-\frac{\beta(m|\mathbf{v} - \mathbf{v}^*|^2 + \kappa|\mathbf{x} - \mathbf{v}^*(t - \alpha/\kappa)|^2)}{2}\right] \quad (7)$$

where $\beta = 1/(k_B T)$. Note that if $\mathbf{v}^* = 0$ this reduces to the equilibrium distribution. To focus on the non-equilibrium stationary state fluctuations, one can take as an initial condition that the distribution at time $t = 0$, denoted by f_0 , is equal to f_{ness} .

This system exhibits many features of general non-equilibrium systems, namely the presence of a current, energy dissipation, and entropy and heat production due to the continuous input of work from the outside. Thus it can serve as a prototypical non-equilibrium system. The advantage of looking at this particular system is that it can be solved exactly (as was first exploited for work fluctuations by Mazonka and Jarzynski [12]). In fact, this system can also be solved exactly for a general time dependence of the center $\mathbf{x}^*(t)$ of the harmonic potential, as was shown in [13].

3. The exact solution

The exact solution of the Brownian particle system introduced above proceeds through a time-dependent change of coordinates that brings the equations into a form in which the force does not depend on time, and whose solution is known.

As explained in [13], to eliminate the time dependence of the force, such a coordinate transformation should take the form

$$\mathbf{X}(t) = \mathbf{x}(t) - \tilde{\mathbf{x}}(t) \quad (8)$$

where $\tilde{\mathbf{x}}$ is any solution of the equations of motion in the absence of the fluctuating force. One such solution is

$$\tilde{\mathbf{x}}(t) = \mathbf{v}^*(t - \alpha/\kappa) \quad (9)$$

Note that the new coordinate system corresponds to a frame that is co-moving with the fictitious noiseless trajectory $\tilde{\mathbf{x}}(t)$, and that along this fictitious trajectory, the friction force $-\alpha\mathbf{v}^*$ is exactly balanced by the harmonic force $-\kappa(\tilde{\mathbf{x}} - \mathbf{x}^*) = \alpha\mathbf{v}^*$.

In the co-moving frame, the Langevin equation becomes [cf. Eqs. (4)–(9)]

$$m\ddot{\mathbf{X}} = -\kappa\mathbf{X} - \alpha\mathbf{V} + \boldsymbol{\xi} \quad (10)$$

where $\mathbf{V} = d\mathbf{X}/dt = \mathbf{v} - \mathbf{v}^*$. Eq. (10) is of the same form as Eq. (4) but with $\mathbf{v}^* = 0$, so it describes the system with a stationary harmonic potential. This system is well-studied and is an example of a so-called Ornstein–Uhlenbeck process [14]. Roughly speaking, an Ornstein–Uhlenbeck process is a stationary Markov process that can be described by a Langevin equation with linear forces and a Gaussian random force term. Equation (10) is Markovian because the time-derivatives of \mathbf{X} only depend on quantities at that same time. As a consequence of the Markovian nature of the process, all its properties can be expressed in terms of the Green's function $G_t(\mathbf{X}', \mathbf{V}'; \mathbf{X}, \mathbf{V})$ which is the probability, in the co-moving frame, for the particle to be at position \mathbf{X}' with velocity \mathbf{V}' given that it was at position \mathbf{X} with velocity \mathbf{V} a time t earlier. For Ornstein–Uhlenbeck processes, the Green's function can be solved exactly and is Gaussian in both the initial coordinates (\mathbf{X}, \mathbf{V}) and the final coordinates $(\mathbf{X}', \mathbf{V}')$. It is this Gaussian property that enables one to obtain many explicit results, as we will see.

Although this system can be solved exactly, to avoid unnecessarily complicated expressions, we will consider here only the overdamped case, which means taking the limit $m \ll \alpha^2/k$. In Eq. (10), this limit can be taken by setting m equal to zero, so that it reduces to $0 = -\kappa\mathbf{X} - \alpha\mathbf{V} + \boldsymbol{\xi}$, or

$$\dot{\mathbf{X}} = -\frac{1}{\tau_r}\mathbf{X} + \alpha^{-1}\boldsymbol{\xi} \quad (11)$$

where $\tau_r = \alpha/\kappa$ is the (average) relaxation time of this Langevin equation. Like the general case, this equation also describes an Ornstein–Uhlenbeck process, whose Green’s function is given by [14]

$$G_t(\mathbf{X}'; \mathbf{X}) = \left[\frac{\beta\kappa}{2\pi(1 - e^{-2t/\tau_r})} \right]^{1/2} \exp \left[-\frac{\beta\kappa}{2} \frac{|\mathbf{X}' - e^{-t/\tau_r} \mathbf{X}|^2}{1 - e^{-2t/\tau_r}} \right] \quad (12)$$

Given a set of times $0 < t_1 < t_2 < \dots < t_m$, one can write for the joint probability of the particle to be at position \mathbf{X}_n at time t_n (with $t_0 = 0$):

$$P(\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_m) = f_0(\mathbf{X}_0) \prod_{i=1}^m G_{t_i - t_{i-1}}(\mathbf{X}_i; \mathbf{X}_{i-1}) \quad (13)$$

Because each Green’s function is Gaussian in both the initial and final positions [cf. Eq. (12)], and the initial distribution in Eq. (7) is also Gaussian in the initial position \mathbf{X}_0 , the distribution P is also Gaussian.

4. Thermodynamic fluctuations

4.1. Definitions

While on a macroscopic level, thermodynamic properties such as work, heat and entropy are well-defined, this is not the case on a mesoscopic level. To understand that there is an ambiguity, consider trying to determine the ‘energy of the system’ for the Brownian particle under consideration. Should this include the potential energy of the external harmonic force, or should this only be the (kinetic) energy of the Brownian particle itself? Or consider the ‘work done on the system’. Should this be the work done by the harmonic force on the particle, or should this be the work needed to move the harmonic potential? The first choice is closely related to the mechanical definition of work, but it leads to work being done even in equilibrium (although its average would be zero), a problem that is avoided by using the second choice. There are even more ambiguities when studying dissipation in non-equilibrium systems. When should energy be regarded as dissipated? It would be ambiguous to say that the dissipated energy is that part of the energy input that cannot be used to perform work. For instance, the harmonic potential energy can be used to store energy by displacing the Brownian particle from its origin. Is this dissipated energy? A clever experimentalist may be able to extract useful work from this energy, but should the definition of dissipated energy depend on the ability of an experimentalist?

The ambiguity is thus due to what to call the system, and what forms of energy to consider as inaccessible or dissipated. To avoid this ambiguity, we will now define what we will mean by ‘work’ and ‘heat’ in the rest of the paper. The work will be defined as the work that is required to keep the harmonic potential moving. Because at time t the harmonic potential exerts a force $-\kappa(\mathbf{x}(t) - \mathbf{x}^*(t))$ on the particle, the particle will exert a reaction force $\kappa(\mathbf{x}(t) - \mathbf{x}^*(t))$ on the harmonic potential (or on whatever device generated it). To keep the potential moving therefore requires (at least) a compensating force $-\kappa(\mathbf{x}(t) - \mathbf{x}^*(t))$ to be applied to the potential. Since the velocity of the potential is $\dot{\mathbf{x}}^* = \mathbf{v}^*$ (cf. Eq. (6)), the mechanical work done by this compensating force during a time interval τ is

$$W_\tau = -\kappa \mathbf{v}^* \cdot \int_0^\tau dt [\mathbf{x}(t) - \mathbf{x}^*(t)] \quad (14)$$

Note that this definition of ‘work on the system’ uses the particle plus the harmonic potential as the system.

On the other hand, we will define the heat as the energy which is given off by the Brownian particle to the fluid, taking the point of view that since the energy of the fluid is not explicitly present in the Langevin description, we can suppose it to be inaccessible or dissipated energy. This leads to the following expression for the heat:

$$Q_\tau = W_\tau - \Delta U_\tau \quad (15)$$

where $\Delta U_\tau = U(\mathbf{x}(\tau)) - U(\mathbf{x}(0))$ is the change in harmonic potential from time 0 to time τ . In words, this states that work can be converted both into potential energy of the Brownian particle, or into heat dissipated into the fluid. Note that in the general (non-overdamped) case, the change in kinetic energy of the particle should also be subtracted from W_τ to obtain the heat.

4.2. Fourier transforms

Given the definitions of the work and the heat and the Green’s function of the process, we can study the probability distributions of their fluctuations. Once the probability distributions are obtained, it can be checked whether they obey a stationary state Fluctuation Theorem. In addition (and in contrast to the majority of theoretical works on the subject), *the explicit solution available for this system allows us to consider finite time corrections to these theorems as well. These corrections give information about the time scales on which these stationary state Fluctuation Theorems can be observed.*

An essential ingredient in the calculation will be that any linear function of a Gaussian distributed quantity is again a Gaussian distributed quantity, so that the distribution of such a quantity is fully determined by its mean and its variance. While the work W_τ in Eq. (14) is a linear function of $\mathbf{x}(t)$, making it Gaussian distributed, the heat Q_τ in Eq. (15) is not a linear function of $\mathbf{x}(t)$, but is in fact a quadratic function of $\mathbf{x}(t)$ and consequently not Gaussian distributed. This quadratic nature prevents a straightforward calculation of the heat probability distribution function, but still allows the computation of its Fourier transform, using the following trick.

Consider first the joint probability $P_\tau^*(W_\tau, \mathbf{x}(0), \mathbf{x}(\tau))$ which gives the probability that the particle moves from position $\mathbf{x}(0)$ to position $\mathbf{x}(\tau)$ and an amount W_τ of work is done in the process. Since these three quantities are all linear functions of $\mathbf{x}(t)$, their joint distribution is Gaussian, i.e. of the form

$$P_\tau^*(\mathbf{y}) = \frac{e^{-\frac{1}{2}(\mathbf{y}-\mathbf{M})\cdot\mathbf{V}^{-1}\cdot(\mathbf{y}-\mathbf{M})}}{\sqrt{2\pi \det \mathbf{V}}} \quad (16)$$

where for brevity we denoted $\mathbf{y} = (W_\tau, \mathbf{x}(0), \mathbf{x}(\tau))$, \mathbf{M} as the mean of \mathbf{y} and \mathbf{V} as the variance matrix of \mathbf{y} .

Using P_τ^* , the distributions of W_τ and Q_τ can be expressed as

$$P_\tau^W(W) = \int d\mathbf{y} P_\tau^*(\mathbf{y})\delta(W - W_\tau) \quad (17)$$

$$P_\tau^Q(Q) = \int d\mathbf{y} P_\tau^*(\mathbf{y})\delta(Q - W_\tau + \kappa|\mathbf{x}(\tau) - \mathbf{x}^*(\tau)|^2/2 - \kappa|\mathbf{x}(0) - \mathbf{x}^*(0)|^2/2) \quad (18)$$

where $\mathbf{x}^*(t) = \mathbf{v}^*t$ as given by Eq. (6). Taking the Fourier transform of these equations yields

$$\hat{P}_\tau^W(q) = \int d\mathbf{y} P_\tau^*(\mathbf{y})e^{iq(W - W_\tau)} \quad (19)$$

$$\hat{P}_\tau^Q(q) = \int d\mathbf{y} P_\tau^*(\mathbf{y})e^{iq(Q - W_\tau + \kappa|\mathbf{x}(\tau) - \mathbf{v}^*\tau|^2/2 - \kappa|\mathbf{x}(0)|^2/2)} \quad (20)$$

where q is the Fourier variable. The integrals on the right hand sides are all just Gaussian integrals, which can be performed explicitly once \mathbf{M} and \mathbf{V} are determined. The results for the stationary state are then [9,13]:

$$\hat{P}_\tau^W(q) = e^{\sigma_+ \tau \tilde{q} [i - \tilde{q} (1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r})]} \quad (21)$$

$$\hat{P}_\tau^Q(q) = \frac{e^{\sigma_+ \tau \tilde{q} [i - \tilde{q} [1 - \frac{2\tilde{q}^2(1 - e^{-\tau/\tau_r})}{\tau/\tau_r(1 + (1 - e^{-2\tau/\tau_r})\tilde{q}^2)}]}}{[1 + (1 - e^{-2\tau/\tau_r})\tilde{q}^2]^{3/2}} \quad (22)$$

where $\tilde{q} = k_B T q$ is the dimensionless Fourier variable and $\sigma_+ = \alpha\beta|\mathbf{v}^*|^2$ is the average heat production rate.

4.3. Saddle point approximations and singularities

To obtain the actual probability distributions of work and heat requires the inversion of the Fourier transforms \hat{P}_τ^W in Eq. (21) and \hat{P}_τ^Q in Eq. (22). Since \hat{P}_τ^W is Gaussian, its inverse is easily obtained:

$$P_\tau^W(W) = \frac{e^{-\frac{(\beta W - \sigma_+ \tau)^2}{4\sigma_+ \tau(1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r})}}}{\sqrt{2\pi \sigma_+ \tau(1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r})}/\beta^2} \quad (23)$$

In fact, it is not necessary to use the Fourier transform to obtain this result for P_τ^W , but it will be useful to have the Fourier transforms of both the work and the heat distribution to understand the origin of the peculiarities of the behavior of the heat fluctuations.

The quantity of interest for the stationary state Fluctuation Theorems to be derived in Section 5 is not the probability of the work fluctuations but that of the *scaled* work fluctuations $p = \beta W / (\sigma_+ \tau)$ (which compares the dimensionless W_τ , βW_τ , with the average dimensionless entropy production over a time τ). Its probability distribution $\pi_\tau^W(p)$ is related to that of W by a Jacobian: $\pi_\tau^W(p) = P_\tau^W(W) dW / dp = P_\tau^W(\beta W / (\sigma_+ \tau)) (\sigma_+ \tau / \beta)$, so that

$$\pi_\tau^W(p) = \frac{e^{-\frac{\sigma_+ \tau (p-1)^2}{4(1 - \frac{1-e^{-\tau/\tau_r}}{\tau/\tau_r})}}}{\sqrt{2\pi(1 - \frac{1-e^{-\tau/\tau_r}}{\tau/\tau_r})} (\sigma_+ \tau)} \quad (24)$$

The Fourier inversion of \hat{P}_τ^Q is more involved because no closed form is known for the Fourier inverse of Eq. (22). We are therefore left with the following formal expression for $\pi_\tau^Q(p)$:

$$\pi_\tau^Q(p) = \frac{\sigma_+ \tau}{2\pi} \int_{-\infty}^{\infty} d\tilde{q} e^{-i\tilde{q}\sigma_+ \tau p} \frac{e^{\sigma_+ \tau \tilde{q} [i - \tilde{q}] [1 - \frac{2\tilde{q}^2(1-e^{-\tau/\tau_r})}{\tau(1+(1-e^{-2\tau/\tau_r})\tilde{q}^2)]}}}{[1 + (1 - e^{-2\tau/\tau_r})\tilde{q}^2]^{3/2}} \quad (25)$$

We are especially interested in the large τ behavior of this function. The presence then of a large parameter in the exponent in the integrand allows us to obtain an approximate result using the saddle point approximation. To exhibit the large parameter τ in Eq. (25) explicitly, we write

$$\pi_\tau^Q(p) = \frac{\sigma_+ \tau}{2\pi} \int_{-\infty}^{\infty} d\tilde{q} e^{\tau h_\tau^Q(\tilde{q})} \quad (26)$$

where

$$h_\tau^Q(\tilde{q}) = \sigma_+ \left\{ -i\tilde{q}p + \tilde{q}[i - \tilde{q}] \left[1 - \frac{2\tilde{q}^2(1 - e^{-\tau/\tau_r})}{(\tau/\tau_r)(1 + (1 - e^{-2\tau/\tau_r})\tilde{q}^2)} \right] \right\} - \frac{3}{2\tau} \log[1 + (1 - e^{-2\tau/\tau_r})\tilde{q}^2] \quad (27)$$

In the saddle point method, the main contribution to the integral is found by considering the maximum \tilde{q}^* of the function $h_\tau^Q(\tilde{q})$. A complicating factor is that h_τ^Q is complex valued. If one takes \tilde{q}^* to be the point on the real line where h_τ^Q has a maximum real part, then the integrand can be highly oscillatory there, causing the contribution of this otherwise dominant point to be washed out. A maximum that does not suffer from this oscillatory behavior can be found, but usually not on the real axis; indeed the solution of

$$\frac{dh_\tau^Q(\tilde{q}^*)}{d\tilde{q}^*} = 0 \quad (28)$$

lies here in the complex plane.

An interesting property of complex functions is that they have no true maximum. According to the Cauchy–Riemann identities for complex functions, the point \tilde{q}^* defined by (28) has the property that there is a path through it called the path of steepest descent, on which h_τ^Q has a constant imaginary part and its real part has a maximum, while along another (orthogonal) path the function has a minimum (also with constant imaginary part); thus \tilde{q}^* is a saddle point of h_τ^Q . The saddle point character of \tilde{q}^* does not interfere with it being the dominant contribution to the integral, as long as we can deform the original line of integration (the real axis) into a path which overlaps with the path of steepest descent through \tilde{q}^* . Furthermore, it can be shown that along this line the function does not oscillate, so that the point \tilde{q}^* truly gives the dominant contribution. Note that during the deformation of the integration path, the end-points ($\pm\infty$ on the real axis) must be fixed and no singularities of the function must be crossed. The value of integral along this deformed path is then exactly equal to that of the original integral. To get the saddle point approximation, one expands the function $h_\tau^Q(q)$ around \tilde{q}^* up to second order in $\delta\tilde{q} = \tilde{q} - \tilde{q}^*$, and integrates along the deformed path. As τ increases, the function becomes more peaked around the saddle point, and the saddle point expression becomes an increasingly good approximation to the integral.

To illustrate this method, let us first apply it to the work distribution, which is simpler than the heat distribution and for which the result can be compared with the exact result in Eq. (24). From Eq. (21) one then obtains

$$\pi_\tau^W(p) = \frac{\sigma_+\tau}{2\pi} \int_{-\infty}^{\infty} d\tilde{q} e^{\tau h_\tau^W(\tilde{q})} \quad (29)$$

where

$$h_\tau^W(\tilde{q}) = \sigma_+ \left\{ -i\tilde{q}p + \tilde{q} \left[i - \tilde{q} \left(1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r} \right) \right] \right\} \quad (30)$$

Since this is a quadratic function in \tilde{q} , the saddle point \tilde{q}^* can be obtained analytically from $dh_\tau^W/d\tilde{q}^* = 0$, with the result

$$\tilde{q}^* = i \frac{1 - p}{2 \left(1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r} \right)} \quad (31)$$

Note that the saddle point lies on the imaginary axis. Noting also that $h_\tau^W(\tilde{q})$ has no singularities as a function of \tilde{q} , one can deform the integration path (i.e. the real axis) to go through \tilde{q}^* . Around \tilde{q}^* the function h_τ^W has the following behavior:¹

$$h_\tau^W(\tilde{q}^* + \delta\tilde{q}) = -\sigma_+ \left\{ \frac{(1 - p)^2}{4 \left(1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r} \right)} + \delta\tilde{q}^2 \left(1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r} \right) \right\} \quad (32)$$

Note that for real $\delta\tilde{q}$, this expression is real (i.e., non-oscillatory) and has a maximum at $\delta\tilde{q} = 0$: the path of steepest descent therefore crosses \tilde{q}^* parallel to the real axis. Thus we can indeed deform the original line of integration (the real axis) by shifting the integration path up or down (depending on p) until it crosses \tilde{q}^* ; this changes the end points $\pm\infty$, but the integrand vanishes there, so that the integral over the shifted line is still the same as for the original integration line. This is sketched in the left panel of Fig. 1. The final integral now becomes

$$\pi_\tau^W(p) = \frac{\sigma_+\tau}{2\pi} e^{-\frac{\sigma_+\tau(1-p)^2}{4 \left(1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r} \right)}} \int_{-\infty}^{\infty} d\delta\tilde{q} e^{-\sigma_+\tau \left(1 - \frac{1 - e^{-\tau/\tau_r}}{\tau/\tau_r} \right) \delta\tilde{q}^2} \quad (33)$$

After performing the Gaussian integral, one recovers Eq. (24).

We now turn to the distribution of heat fluctuations. To find the saddle point in this case, the derivative of $h_\tau^Q(\tilde{q})$ has to be equal to zero. This derivative is rather complicated, in contrast to the simple linear function that it was for the work functions. To make the expressions in the following a bit more manageable, let us assume that τ is so large that $e^{-\tau/\tau_r}$ may be neglected in Eq. (28) (by comparison with numerical methods, we found that this typically leads to accurate results for $\tau > 3\tau_r$ [9]). In that case, h_τ^Q becomes

$$h_\tau^Q(\tilde{q}) = \sigma_+ \left\{ -i\tilde{q}p + \tilde{q}(i - \tilde{q}) + \frac{2\tilde{q}^3}{(\tau/\tau_r)(i + \tilde{q})} \right\} - \frac{3}{2\tau} \log(1 + \tilde{q}^2) \quad (34)$$

while

$$\frac{d}{d\tilde{q}} h_\tau^Q(\tilde{q}) = \sigma_+ \left\{ -ip + i - 2\tilde{q} + \frac{6\tilde{q}^2}{(\tau/\tau_r)(i + \tilde{q})} - \frac{2\tilde{q}^3}{(\tau/\tau_r)(i + \tilde{q})^2} \right\} - \frac{3\tilde{q}}{\tau(1 + \tilde{q}^2)} \quad (35)$$

Note that the naive limits $\tau \rightarrow \infty$ of h_τ^Q in Eq. (34) and h_τ^W in Eq. (30) coincide. However, the terms in Eq. (34) that are then neglected have singularities, which have to be taken into account, because even though the factor $1/\tau$ may tend to zero as $\tau \rightarrow \infty$, it multiplies an expression that can, because of the singularities, go to infinity, making the limit $\tau \rightarrow \infty$ ill-defined.

The saddle point equation $dh_\tau^Q/d\tilde{q} = 0$ leads, using Eq. (35), to a fourth order polynomial equation. Such an equation has four solutions; hence there are four saddle points, in contrast to the Fourier transform of the work

¹ For more general forms of h , there would be $O(\delta\tilde{q}^3)$ correction terms.

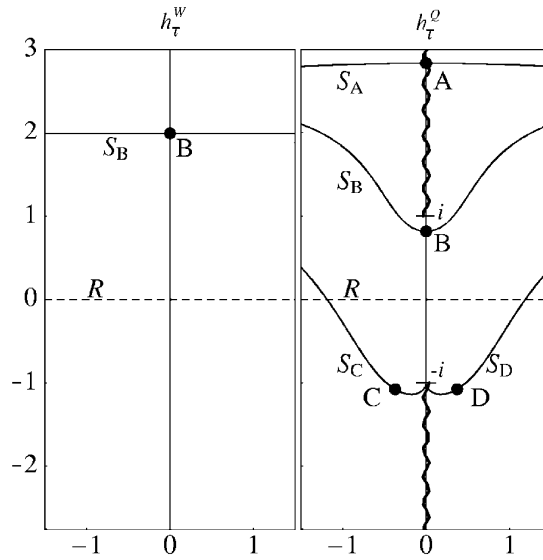


Fig. 1. Saddle points and paths of steepest descent of the functions h_τ^W (left) and h_τ^Q (right) for $\sigma_+ = 1$, $\tau = 4\tau_r$ and $p = -2$. Solid circles represent saddle points and solid curves S_A, \dots, S_D drawn through these are the paths of steepest descent. For h_τ^W , shown on the left, only one saddle point B exists and the (dashed) real axis R can be deformed to the path of steepest descent S_B going horizontally through B, without crossing a singularity. For h_τ^Q shown on the right, wiggly lines represent branch cuts, two singularities appear at $\pm i$, and there are four saddle points A, B, C and D. The path of steepest descent S_B through B is the only way to deform the real axis to a path of steepest descent without crossing a singularity.

distribution, which had only one. Which saddle point(s) to use depends on how we can deform the real axis integration path to a new path which includes saddle points, without crossing any singularities. As Eq. (34) shows, there are singularities at $\tilde{q} = \pm i$. The point $\tilde{q} = i$ is a branch point associated with the logarithm, while $\tilde{q} = -i$ is in addition a simple pole. The structure of the saddle points is shown in the right panel of Fig. 1. As a function of p , there is only one saddle point that we can reach by deforming the real axis without crossing singularities. This saddle point lies on the imaginary axis and is in a sense trapped between the singularities at $\pm i$; this is quite different from the saddle point for the work fluctuations [i.e. $i(1 - p)/2$ for large τ , cf. Eq. (31)] which is unbounded as a function of p .

The full calculation of the Fourier inverse of \hat{P}_τ^Q would be too detailed for this short paper and can be found in [9]. In an asymptotic expansion for large τ , the result is

$$\pi_\tau^Q(p) = \begin{cases} \sqrt{\frac{\sigma_+^3 \tau^3 |p + 1|}{36\pi}} e^{-\sigma_+[\tau_r - \tau p] + 3/2} & \text{if } p < -1 \\ \sqrt{\frac{16\sigma_+ \tau}{\pi} \frac{e^{-\sigma_+(1-p)^2[\tau + 2(1-p)/(3-p)\tau_r]/4}}{[(3-p)(1+p)]^{3/2}}} & \text{if } -1 < p < 3 \\ \sqrt{\frac{\sigma_+ \tau^2 / \tau_r}{32\pi}} e^{-\sigma_+[(p-2)\tau - \sqrt{8(p-3)\tau\tau_r} + \frac{20-6p}{p-3}\tau_r]} & \text{if } p > 3 \end{cases} \quad (36)$$

Note that this function has a predominantly Gaussian center, for large τ , for $-1 < p < 3$ and exponential tails for $p < -1$ and $p > 3$, i.e., for large τ , one has $\pi_\tau \sim \exp[\sigma_+ \tau p]$, $\pi_\tau \sim \exp[-\sigma_+ \tau (p - 1)^2 / 4]$ and $\pi_\tau \sim \exp[-\sigma_+ \tau (p - 2)]$ for $p < -1$, $-1 < p < 3$ and $p > 3$, respectively. Qualitatively, this result can be understood as follows: if p is such that the saddle point \tilde{q}^* is relatively far away from the singularities, the function and its relevant saddle point resemble those of the work fluctuations, i.e., \tilde{q}^* scales linearly with p and the Fourier inverse is consequently Gaussian in p . Since for the work fluctuations $\tilde{q}^* = i(1 - p)/2$, the values of p for which \tilde{q}^* is not close to the singularities $\pm i$ lie within the interval $p \in [-1, 3]$. For values $p > 3$, the work saddle point lies below $-i$, but the heat saddle point has to stay within $[-i, i]$. As a result it is basically stuck just above $-i$, i.e., $\tilde{q}^* \approx -i$, yielding, cf. Eq. (25), an exponential dependence of $\pi_\tau^Q(p)$ on p . Likewise, for $p < -1$, the heat saddle point is stuck just below i and again an exponential tail of the distribution is present.

Eq. (36) also includes finite τ effects, obtained using the saddle point approximation. Note that the approach to the large τ behavior has a different character for different values of p . This corresponds to the different character of the singularities at $\pm i$.

5. Stationary state fluctuation theorems

Given the distribution functions, one can now investigate whether the stationary state Fluctuation Theorem holds. The quantity considered in the Fluctuation Theorems is the *fluctuation function* $f_\tau(p) = \frac{1}{\sigma_+\tau} \log[\pi_\tau(p)/\pi_\tau(-p)]$. This function measures the asymmetry between positive and negative fluctuations in the quantity p . When p represents work, $\pi_\tau(p) = \pi_\tau^W(p)$ and Eq. (24) gives:

$$f_\tau^W(p) = \frac{p}{1 - \frac{1-e^{-\tau/\tau_r}}{\tau/\tau_r}} \tag{37}$$

As $\tau \rightarrow \infty$, this tends to p from above (cf. Fig. 2), hence the Conventional Fluctuation Theorem (CFT) is satisfied by the work fluctuations. For these fluctuations, the characteristic time is τ_r .

On the other hand, considering the heat fluctuations and using Eq. (36), one finds an Extended Fluctuation Theorem (EFT):

$$f_\tau^Q(p) = \begin{cases} p + \frac{8p}{(9-p^2)\tau/\tau_r} - \frac{3 \ln \frac{3-p^2+2p}{3-p^2-2p}}{2\sigma_+\tau} & \text{if } 0 < p < 1 \\ p - \frac{(1-p)^2}{4} + \frac{(p+1)(-2p^2+8p-10)}{4(3-p)\tau/\tau_r} & \\ - \frac{1}{\sigma_+\tau} \left\{ \ln(\sigma_+\tau) - \frac{1}{2} \ln \frac{(3-p)^3(1+p)^3(p-1)}{576} - \frac{3}{2} \right\} & \text{if } 1 < p < 3 \\ 2 + \sqrt{\frac{8(p-3)}{\tau/\tau_r}} - \frac{\ln(\tau\sigma^2\tau_r)}{2\sigma_+\tau} & \text{if } p > 3 \end{cases} \tag{38}$$

Note that the behavior for negative p can simply be found from the antisymmetry of the function $f_\tau^Q(p)$, i.e. $f_\tau^Q(-p) = -f_\tau^Q(p)$.

To illustrate the difference between the work and heat fluctuation functions f_τ^W and f_τ^Q , we sketched these functions as a function of p in Fig. 2. One sees that in the case $\tau \rightarrow \infty$, for $p < 1$, the fluctuation functions tend to a straight line with slope one. This behavior persists beyond $p > 1$ for f_τ^W , while f_τ^Q bends down away from the straight line and settles to a plateau value of 2 for $p > 3$.

Thus, even in the limit $\tau \rightarrow \infty$, Eq. (38) for heat fluctuations differs from the CFT for work fluctuations when $p > 1$, since the fluctuation function has for heat a plateau at a value of 2 rather than that it increases without bound

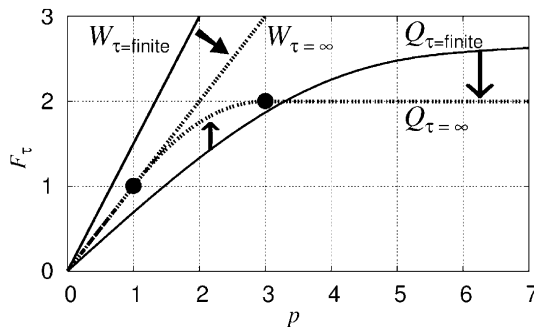


Fig. 2. Conventional (CFT) and Extended Fluctuation Theorem (EFT). $W_{\tau=\text{finite}}$ and $W_{\tau=\infty}$ represent the CFT and $Q_{\tau=\text{finite}}$ and $Q_{\tau=\infty}$ the EFT for finite and infinite times, respectively. The arrows indicate how the infinite time limit is approached, i.e., for the work fluctuations from above, and for the heat fluctuations from below for small p and from above for large p .

as for work. Alternatively, one could say that f_τ^Q does satisfy Eq. (1), but with $p^* = 1$, whereas f_τ^W has $p^* = \infty$. We note that for the heat fluctuations there is in addition to τ_r a relaxation time σ_+^{-1} , which can exceed τ_r .

The plateau value 2 in the EFT for $p > 3$ (cf. Eq. (38)) is a consequence of the exponential tails of the distribution P_τ^Q . These tails are related to the presence of singularities at $\pm i$, which limit the position of the saddle point \tilde{q}^* , such that for $p > 3$ one has $\tilde{q}^* \approx -i$, while for $p < -1$, $\tilde{q}^* \approx +i$. As the saddle point gives the dominant contribution to the Fourier inverse, this leads to the exponential tails

$$\pi_\tau^Q(p) \sim \begin{cases} e^{-\sigma_+\tau(p-2)} & \text{for } p > 3 \\ e^{\sigma_+\tau p} & \text{for } p < -1 \end{cases} \quad (39)$$

The exponential tails directly give $f_\tau^Q(p) = (\sigma_+\tau)^{-1} \ln \pi_\tau^Q(p)/\pi_\tau^Q(-p) \sim 2$. Note that conversely, exponential tails give rise to singularities in the Fourier transform since $\int_{-\infty}^{\infty} dp e^{iqp} e^{-|p|} = i/(q+i) - i/(q-i)$.

6. Discussion

We have shown that for a dragged Brownian particle described by a Langevin equation, the work fluctuations satisfy the Conventional Fluctuation Theorem, and the heat fluctuations satisfy an Extended Fluctuation Theorem. These results may seem to only be peculiarities of this model. However, these results largely carry over to other more general systems too:

1. The model of a Brownian particle in a harmonic potential dealt with here can be mapped onto a system of an electric circuit with a current source [10]. Furthermore, the model can be extended to the Rouse model for a polymer, which consists of a harmonically bound system of Brownian particles. It has been shown that there too the work distribution is Gaussian and satisfies the CFT [15], while the heat fluctuations appear to have exponential tails [16].
2. Another extension is to consider the non-overdamped case. This has been done in the context of work fluctuations by Douarche et al. [17]. While the details of the work fluctuations for finite times can be quite different, the general conclusion, namely, that the work fluctuations satisfy a CFT, still holds. It is hard to imagine that the heat would not satisfy an EFT, since this model is just a more detailed description of the same Brownian system.
3. Extensions to non-harmonic potentials have been explored as well. Blicke et al. have studied the work fluctuations in [18], while Baiesi et al. [19] have argued for an EFT for heat under some general assumptions, one of which is the existence of exponential tails for the distribution of Q_τ .
4. The limitation $p < p^*$ to the CFT has been found analytically by other authors as well [20,21]. The limit is in all these cases also due to singularities in the complex plane of the Fourier transform of the distribution, which correspond to exponential tails of the distribution itself.
5. Turning to deterministic systems, Evans [22] argued that there should be a relation between the FT in such systems and the EFT. Indeed, Gilbert has shown numerically that the EFT holds in a Nose–Hoover thermostated Lorentz Gas [23] and [20] explains in a general way why the EFT should hold using a large deviation formalism, but still assuming that there are exponential tails and that the work fluctuations have faster than exponential tails. They also provide a strategy on how to apply the chaotic hypothesis to singular systems such as Lennard-Jones systems, by using a Poincaré section that circumvents the singularities in the potential.
6. A derivation of a general form of the EFT has been given for stochastic systems by Baiesi et al. [19] and for deterministic systems by Bonetto et al. [20], both based (among others) on the assumption of the presence of exponential tails. They found that universal features in the EFT of the $\lim_{\tau \rightarrow \infty} f_\tau^Q(p)$ are the slope one for $p < 1$ and the presence of a plateau for large p . On the other hand, the shape and extent of the intermediate- p region is not universal, nor is the height of the plateau, unless the heat distribution is symmetric around its average, in which case the plateau value is always 2. This can be understood by realizing that symmetry around the average (i.e. $\langle p \rangle = 1$) can be expressed as $\pi_\tau^Q(p) = \pi_\tau^Q(2-p)$. Hence an exponential tail for positive heat fluctuations $\pi_\tau^Q(p) \sim e^{-\beta Q_\tau} = e^{-\sigma_+\tau p}$ entails for negative heat fluctuations that $\pi_\tau^Q(-p) = \pi_\tau^Q(2+p) \sim e^{-\sigma_+\tau(p+2)}$, giving $f_\tau^Q(p) = (\sigma_+\tau)^{-1} \ln \pi_\tau^Q(p)/\pi_\tau^Q(-p) \sim 2$.
7. Deviations from the CFT are often found to be a consequence of so-called boundary terms [8,9,19,20,24], such as the ΔU_τ term in Eq. (15). A boundary term is the difference of a function (here the potential energy) at the

boundaries of the interval τ . Boundary terms *typically* do not grow with τ , in contrast to bulk terms like W_τ in Eq. (15), which *typically* grow linearly with τ . It is therefore tempting to neglect boundary terms, but the Fluctuation Theorem concerns *atypical* fluctuations, for which boundary terms possibly play a role. Indeed, for the exactly solved model, neglecting ΔU_τ in Q_τ would eliminate the exponential tails and would thus give the CFT instead of the correct EFT for heat fluctuations.

8. Evidence for the EFT has been seen in experiments as well, notably by Garnier and Ciliberto in an electric circuit [7].
9. It has been suggested that the Gallavotti–Cohen Fluctuation Theorem (Eq. (1)), meaning the stationary state Fluctuation Theorem for entropy production, or heat, would not hold near equilibrium, but the FT for the ‘dissipation function’ of Evans and Searles (which often coincides with work) would [25]. However, it seems that in the appropriate limit $\tau \rightarrow \infty$, with a restriction $p < p^*$, the Gallavotti–Cohen Fluctuation Theorem does hold. Nonetheless, there is a problem in observing the EFT for heat near equilibrium that the CFT for work does not have. As we can see in the exactly solved model of the Brownian particle, the FT for work decays to the correct FT form on a time scale of $O(\tau_r)$ [cf. Eq. (37)]. On the other hand, the heat FT for $p < 1$ decays to the correct FT on a time scale of $O(\sigma_+^{-1})$ [cf. Eq. (38)]. The latter time scale diverges as σ_+ gets smaller, i.e., as one gets closer to equilibrium, whereas the time scale for the work FT (i.e., τ_r) does not. This explains why it is very hard to see the FT for heat in simulations [26].

Acknowledgements

This work was supported by grants from the Department of Energy’s Office of Basic Engineering under grant DE-FG-02-88-ER13847 and the National Science Foundations’ Mathematics–Physics Division under grant PHY-0501315.

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;
G. Gallavotti, Math. Phys. Electron. J. 1 (1995) 1;
G. Gallavotti, cond-mat/0404699.
- [3] D.J. Evans, D.J. Searles, Adv. Phys. 51 (2002) 1529.
- [4] J.L. Lebowitz, H. Spohn, J. Stat. Phys. 95 (1999) 333.
- [5] J. Kurchan, J. Phys. A, Math. Gen. 31 (1998) 3719.
- [6] S. Ciliberto, N. Garnier, S. Hernandez, C. Lacpatia, J.-F. Pinton, G. Ruiz Chavarria, Physica A 340 (2004) 240.
- [7] N. Garnier, S. Ciliberto, Phys. Rev. E 71 (2005) 060101(R).
- [8] R. van Zon, E.G.D. Cohen, Phys. Rev. Lett. 91 (2003) 110601.
- [9] R. van Zon, E.G.D. Cohen, Phys. Rev. E 69 (2004) 056121.
- [10] R. van Zon, S. Ciliberto, E.G.D. Cohen, Phys. Rev. Lett. 92 (2004) 130601.
- [11] G.M. Wang, E.M. Sevick, E. Mittag, D.J. Searles, D.J. Evans, Phys. Rev. Lett. 89 (2002) 050601.
- [12] O. Mazonka, C. Jarzynski, cond-mat/9912121.
- [13] R. van Zon, E.G.D. Cohen, Phys. Rev. E 67 (2003) 046102.
- [14] G.E. Uhlenbeck, L.S. Ornstein, Phys. Rev. 36 (1930) 823.
- [15] T. Speck, U. Seifert, Eur. Phys. J. B 43 (2005) 521.
- [16] A. Dhar, Phys. Rev. E 71 (2005) 036126.
- [17] F. Douarache, S. Joubaud, N.B. Garnier, A. Pertosyan, S. Ciliberto, Phys. Rev. Lett. 97 (2006) 140603.
- [18] V. Blickle, T. Speck, L. Heiden, U. Seifert, C. Bechinger, Phys. Rev. Lett. 96 (2006) 070603.
- [19] M. Baiesi, T. Jacobs, C. Maes, N.S. Skantzos, Phys. Rev. E 74 (2006) 021111.
- [20] F. Bonetto, G. Gallavotti, A. Giuliani, F. Zamponi, J. Stat. Phys. 123 (2006) 39.
- [21] L. Rey-Bellet, L.E. Thomas, Ann. H. Poincaré 3 (2002) 483;
P. Visco, J. Stat. Mech. (2006) P06006.
- [22] D.J. Evans, Mol. Simulation 31 (2005) 389.
- [23] T. Gilbert, Phys. Rev. E 73 (2006) 035102.
- [24] A. Puglisi, L. Rondoni, A. Vulpiani, J. Stat. Mech. (2006) P08010.
- [25] D.J. Evans, D.J. Searles, L. Rondoni, Phys. Rev. E 71 (2005) 056120.
- [26] D.J. Searles, D.J. Evans, J. Chem. Phys. 113 (2000) 3503;
M. Dolowschiak, Z. Kovacs, Phys. Rev. E 71 (2005) 025202(R).