Available online at www.sciencedirect.com

ScienceDirect

C. R. Physique 8 (2007) 567–578

COMPTES RENDUS
PHYSIQUE

http://france.elsevier.com/direct/COMREN/

Work, dissipation, and fluctuations in nonequilibrium physics

Fluctuation and dissipation

Bart Cleuren [a,*], Christian Van den Broeck [a], Ryoichi Kawai [b]

[a] *Hasselt University, Campus Diepenbeek, Agoralaan–Gebouw D, B-3590 Diepenbeek, Belgium*
[b] *Department of Physics, University of Alabama at Birmingham, Birmingham, AL 35294, USA*

Received 7 March 2007

Abstract

We review the transient version of the work and fluctuation theorem. We verify the theorem explicitly by a microscopically exact analytic calculation of the probability distribution for work or entropy production in three physical examples: the Joule experiment, the compression/expansion of a gas and the process of effusion. ***To cite this article: B. Cleuren et al., C. R. Physique 8 (2007).***

Résumé

Fluctuations et dissipation. Nous passons en revue la version transitoire du théorème du travail et de fluctuations. Nous vérifions explicitement le théorème par un calcul analytique microscopique exact de la distribution de probabilité pour le travail ou la production d'entropie dans trois exemples physiques : l'expérience de Joule, la compression/expansion d'un gaz et le processus d'effusion. ***Pour citer cet article : B. Cleuren et al., C. R. Physique 8 (2007).***

Keywords: Nonequilibrium physics; Work theorem; Fluctuation theorem

Mots-clés : Physique du non-équilibre ; Théorème du travail ; Théorème de fluctuations

1. Introduction

Since the pioneering work of Onsager, we understand that there is a deep relation between fluctuation and dissipation, at least close to equilibrium in the regime of linear response. This relation is formalized in famous results, such as the symmetry of the Onsager coefficients [1], the fluctuation dissipation theorem and the Green–Kubo relations [2]. Over the past decade, new connections have been established that relate fluctuation with dissipation in far from equilibrium situations [3–6]. They may ultimately lead to a formulation of nonequilibrium statistical mechanics akin to the very successful equilibrium theories, and provide a deeper understanding of irreversibility and of the microscopic foundation of the second law. In this paper, we review a few aspects of these new developments and focus more specifically on the transient fluctuation and work theorems. In this case, a system, which is initially at equilibrium, is perturbed arbitrarily far from equilibrium by changing an external control variable, that determines the

---

[*] Corresponding author.
E-mail addresses: bart.cleuren@uhasselt.be (B. Cleuren), christian.vandenbroeck@uhasselt.be (C. Van den Broeck), kawai@uab.edu (R. Kawai).

1631-0705/$ – see front matter © 2007 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.crhy.2007.04.015

Hamiltonian, according to a specific time-dependent protocol. The quantities of interest are the work that is performed and the entropy that is produced. Due to the randomness of the initial condition (which is sampled from an equilibrium ensemble), the quantities are random variables. The work and fluctuation theorem imply that the probability distributions for these quantities obey a specific symmetry relation. The latter expresses micro-reversibility of the underlying dynamics, very much in the same way as the symmetry of the Onsager linear response matrix derives from it.

2. Revisiting the fluctuation and work relation

The work and fluctuation relation have been proven using different approaches and for different types of systems [7–18]. We first review a Hamiltonian derivation based on the microcanonical ensemble [19]. In the next section, we present an abstract argument that highlights the generic nature of these relations.

Consider a Hamiltonian $H(\Gamma, \lambda)$ where Γ represents the set of position and momentum variables of consideration and λ is a control parameter which is varied from an initial value λ_A to a final value λ_B according to a protocol $\lambda(t),\ t \in [t_i, t_f]$, controlled by an external agent. Note that the system is otherwise isolated from the outside world so that no heat is exchanged. We assume that the system is initially in microcanonical equilibrium at energy E. Hence any point in the initial energy surface $\mathcal{A} = \{\Gamma\colon\ H(\Gamma; \lambda_A) = E\}$ is equally likely to be a starting configuration. Each of these configurations evolves according to deterministic Hamiltonian dynamics into a final position in phase space. For a subset $\mathcal{A}'$ of $\mathcal{A}$ of the initial states, the corresponding final position lies in the energy shell $\mathcal{B} = \{\Gamma\colon\ H(\Gamma; \lambda_B) = E + W\}$, and the work performed on the system is equal to W. The probability for this to happen is obviously given by $P_E(W) = \mu(\mathcal{A}')/\mu(\mathcal{A})$ where μ refers to the Liouville measure of the set. We now consider a backward process, in which λ is varied according to the time-reversed protocol from λ_B to λ_A. We suppose that the system is now initially in microcanonical equilibrium at energy $E + W$, i.e., its initial configuration is uniformally sampled from the energy surface $\mathcal{B}$. Consider $\mathcal{B}'$, the subset of initial configurations that end in $\mathcal{A}$, with corresponding work performed on the system equal to $-W$. Obviously one has by a similar argument $\tilde{P}_{E+W}(-W) = \mu(\mathcal{B}')/\mu(\mathcal{B})$ for the work distribution in the backward process. Microscopic time reversibility guarantees that there is a one to one correspondence between the trajectories from $\mathcal{A}'$ and $\mathcal{B}'$, and the sets have the same measure by Liouville's theorem. Hence, the ratio of the probabilities is:

$$\frac{P_E(W)}{\tilde{P}_{E+W}(-W)} = \frac{\mu(\mathcal{B})}{\mu(\mathcal{A})} = \mathrm{e}^{[S(\mathcal{B})-S(\mathcal{A})]/k} \tag{1}$$

The entropy is equal to the Boltzmann constant time the log of the phase volume, e.g., $S(\mathcal{A}) = k \ln \mu(\mathcal{A})$. The above relation simplifies considerably in an appropriate thermodynamic limit, which we clarify in the further discussion. We consider the limit $W/E \to 0$, with the derivative $T = \partial E/\partial S$, i.e., the temperature, a constant ($\beta = 1/kT$). In this limit, P_E and $\tilde{P}_{E+W}$ converge to functions P and $\tilde{P}$ of the temperature T (we drop the subscript T for simplicity of notation). One thus finds to dominant order in W/E:

$$\frac{P(W)}{\tilde{P}(-W)} = \mathrm{e}^{\beta(W-\Delta F)} \tag{2}$$

where we used:

$$S(\mathcal{B}) - S(\mathcal{A}) = S(E + W, \lambda_B) - S(E, \lambda_A) = (W - \Delta F)/T + \mathrm{O}(W/E)$$

The quantity ΔF is given by $-[S(E, \lambda_B) - S(E, \lambda_A)]T$. The latter free energy difference is the work that is performed when the protocol is infinitely slow (quasi-static). Indeed, since no heat is exchanged with the outside world, quasi-static implies reversible, hence $S(\mathcal{B}) - S(\mathcal{A}) = 0$, or $W = W_{\mathrm{rev}} = \Delta F$. An alternative proof of this statement follows from the Jarzynski equality, see below, by observing that this value ΔF is indeed the minimum of the required average work. The physical intuition behind the above limit is as follows: the perturbation acts on a subpart of the very large system. The remainder of the system can be considered as a heat bath, which is during the time of the perturbation not affected by it. On a longer time scale, the overall system relaxes to full equilibrium, which correspond to a microcanonical ensemble for the entire system, but which is canonical for the subsystem. Since $W/E \to 0$, the final temperature of the overall system is the same as the initial temperature. ΔF is then the free energy difference of the subsystem between final and initial canonical equilibrium state, having the familiar meaning of the reversible work for a system in contact with a heat bath. The work that is performed in the transition, and which exceeds this free energy difference $W_{\mathrm{diss}} = W - \Delta F$ is energy going into the heat bath, hence it contributes to entropy production $\Delta S = W_{\mathrm{diss}}/T$.

From the Crooks theorem (2), it is straightforward to derive the Jarzynski equality,

$$\left\langle \mathrm{e}^{-\beta W} \right\rangle = \mathrm{e}^{-\beta \Delta F} \tag{3}$$

which implies:

$$\langle W \rangle \geqslant \Delta F \tag{4}$$

We need to comment on this 'derivation' of the second law. The above calculation is entirely mechanical. We need not make any assumptions about the dynamics, except for reversibility and Liouville theorem (incompressibility in phase space). Ergodicity, mixing or molecular chaos are not invoked. In fact, we have side-stepped the problem of relaxation to equilibrium. We assume equilibrium as a starting configuration (but need not tell how the microcanonical distribution was achieved). The connection with entropy production and the interpretation of ΔF as free energy difference between final and initial state of the subsystem requires the assumption, a posteriori, that the system returns to equilibrium after the perturbation has stopped. In conclusion, one can make exact statements about the 'dissipated' work, without having to address directly the issue of irreversibility. Remarkably, the derivation of the fluctuation and work theorem rely crucially on the conservation of phase volume, a property which is itself often seen as an obstacle for proving irreversibility or for giving a microscopic definition of entropy away from equilibrium.

The above result is sometimes referred to as the transient work theorem. We can, by proceeding along exactly the same lines, derive a transient fluctuation theorem. The latter follows immediately from the Crooks theorem, by realizing that there is a one-to-one correspondence between energy and entropy, so that an energy change from E to $E + W$ implies a corresponding entropy change $\Delta S = S(E + W, \lambda_B) - S(E, \lambda_A)$. The Jacobian of the transformation from the probability distribution for the work, $P(W)$, to the probability distribution for the entropy production, $P(\Delta S)$, drops out so that we can rewrite Eq. (2) as follows:

$$\frac{P(\Delta S)}{\tilde{P}(-\Delta S)} = \mathrm{e}^{\Delta S/k} \tag{5}$$

From this result, it follows that

$$\left\langle \mathrm{e}^{-\Delta S/k} \right\rangle = 1 \tag{6}$$

which implies the second law:

$$\langle \Delta S \rangle \geqslant 0 \tag{7}$$

The above transient theorems have to be compared with the theorems that apply to nonequilibrium steady states. In this case, a symmetry relation can be derived, under certain conditions, for the entropy production per unit time $\sigma = \Delta S/t$, in the limit of asymptotically long times:

$$\frac{P(\sigma)}{P(-\sigma)} \sim \mathrm{e}^{t\sigma/k} \tag{8}$$

Note the absence of a tilde ('time-reversed' scenario) in this formula, in agreement with the observation that there is no obvious way to specify the time-reversed steady state. The restriction to long times is understandable since the initial state is not specified. One therefore needs to make measurements long enough so that the initial configuration is forgotten. The above result is thus in effect a symmetry property for the large deviations of the entropy production. Unfortunately, such deviations are often difficult to observe. Furthermore, the above relation can run into trouble for systems with unbounded energy. The reason is that large deviations of the energy of the initial state can be responsible for the large deviations of the entropy production, even after an asymptotically long time. In such a case, the initial conditions are never forgotten and the fluctuation theorem fails. The similarity with the transient fluctuation theorem is of course striking. In fact, one can imagine perturbations of a subpart of the system that bring it into a steady state, at least as far as work or entropy production is concerned. Examples of this situation will be given below.

3. Generic structure of the work and fluctuation theorem

The form of the work and fluctuation theorem appears to be very specific. As we show below, it is, however, generic for the (information) entropy of a random variable, that is naturally associated to an involution. Let $x \in \Omega$ be a set of random variables with probability distribution $P(x)$, and T an involution on the space Ω, that is $Tx \in \Omega$ and $T^2x = x$. The Jacobian of such a transformation from x to $y = Tx$ has necessarily an absolute value equal to 1. Note that time reversal is an example of an involution.

We now define the following quantity:

$$\sigma(x) = -\log\left[\frac{P(Tx)}{P(x)}\right] \tag{9}$$

The intuitive meaning of this expression is given in information theory: it measures the difference in surprise of observing the outcome Tx as opposed to x. As a result of the involution, one has $\sigma(Tx) = -\sigma(x)$. Consider now the probability distribution P_σ for the quantity σ. We will show that it obeys the fluctuation theorem:

$$\frac{P_\sigma}{P_{-\sigma}} = \mathrm{e}^{\sigma} \tag{10}$$

In words, the probability to observe a positive value σ is exponentially more likely then that of observing $-\sigma$. The proof of Eq. (10) follows directly from the definition of Eq. (9) and the unit absolute value of the Jacobian (analogue of the Liouville theorem):

$$P_\sigma = \int\limits_{\Omega} \mathrm{d}x\, P(x)\delta\big(\sigma(x) - \sigma\big) = \int\limits_{\Omega} \mathrm{d}x\, \mathrm{e}^{\sigma(x)} P(Tx)\delta\big(\sigma(x) - \sigma\big) = \mathrm{e}^{\sigma} \int\limits_{\Omega} \mathrm{d}y\, P(y)\delta\big(\sigma(y) + \sigma\big) = \mathrm{e}^{\sigma} P_{-\sigma} \tag{11}$$

4. Three case studies

We proceed with an illustration of the work and fluctuation theorem in three case studies, each one corresponding to an important and well studied physical situation. Explicit analytic results are obtained by considering the simplest possible constituent, namely an ideal gas. It will be supposed that the gas is initially at equilibrium and that the container is infinitely large. The use of such an ideal heat bath is of course a familiar trick in statistical mechanics, cf. the derivation of a Langevin equation for a particle connected to a bath of harmonic oscillators. Our main new contribution is that we focus on the probability distribution of work and entropy production in situations that are arbitrarily far from equilibrium. The explicit results all vindicate the work and fluctuation theorem. One essential ingredient that allows for a full and detailed solution in all three examples is the fact that the molecular chaos ansatz is exact. More precisely, the physical processes (collision or effusion) that require work or cause entropy production have no memory. There is in this sense no distinction between steady state or transient work and fluctuation theorems: a steady state is reached right from the start of the perturbation. These results are compared with molecular dynamics for a dilute hard disk gas (for more details: see [19–21]).

4.1. Joule experiment

In 1845, James P. Joule demonstrated the equivalence between mechanical work and heat in his historical experiment. A paddle is rotated inside a thermally isolated container. Due to friction, work input, e.g., by a falling weight, is required to keep the paddle moving. This work is converted into heat, as demonstrated by the increasing temperature inside the container. As is well known, the work that has to be injected is generally speaking not a state function. For example, the work to rotate the paddle a fixed number of times will be larger if the stirring is more violent (e.g., using a heavier weight). Thermodynamics teaches us only the lower limit of the work. When the procedure is performed reversibly, the work is minimal. If the container is sufficiently large, it can be considered to function as a heat bath at constant temperature and the reversible work is then $W_{\mathrm{rev}} = \Delta F$. In the present case ΔF is the change in free energy of the paddle, which is zero if we disregard any possible change in its internal structure.

In view of the historic and fundamental importance of the Joule experiment, it is somewhat surprising that an exact analysis has never been attempted, that is, until recently [19]. To formulate a model that can be analyzed microscopically, we need to make a number of simplifications. We consider a isosceles triangle moving at *constant* speed V

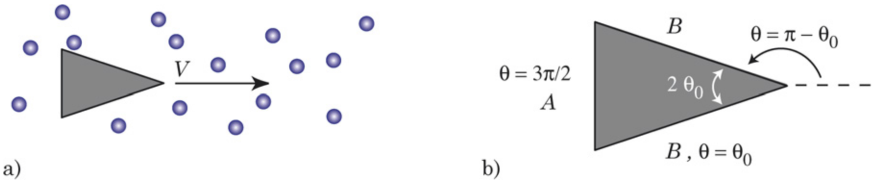

Fig. 1. (a) A triangle moves at constant speed through an ideal gas for a fixed time interval t. (b) Sketch of an isosceles triangle in two dimensions.

through a gas (Fig. 1(a)). For simplicity, the focus here is on a two dimensional system and pure translational motion of the triangle. We also assume that the container is infinitely large and that the gas is initially at equilibrium and consists of point particles (ideal gas assumption) with mean free path much longer than the linear dimension of the object (large Knudsen number). Since there are no pre-collisional correlations between the gas particle and the triangle, the molecular chaos assumption is exact and a Master equation provides a rigorous starting point to study the dissipation in this set-up.

To do so we need some further preparatory work. The shape of the isosceles triangle is determined by the length of its sides, taken here to be A for the base and B for the two other sides, cf. Fig. 1(b). The apex angle $2\theta_0$ follows from $\sin(\theta_0) = A/(2B)$. Upon collision of a gas particle (mass m) with the object, the amount of work ΔW transferred from the outside world to the gas is given by the change of kinetic energy of the colliding particle. Given a gas particle with pre-collisional velocity (v_x, v_y), the resulting work contribution can be deduced from the conservation of total energy and total momentum in the x-direction, and is found to be:

$$\Delta W = -2mV\sin^2\theta(v_x - V - v_y\cot\theta) \tag{12}$$

Here, the angle θ is the inclination angle of the triangle at impact. It is measured counterclockwise from the horizontal x-direction, cf. Fig. 1(b). For the triangle considered here, oriented symmetrically with respect to the horizontal direction, the three possible values are $\pi - \theta_0$, $3\pi/2$ and θ_0 respectively.

Clearly, ΔW is a random variable. Because of the aforementioned molecular chaos property (note that the speed of the object is fixed) the subsequent collisions are independent random events. Hence the accumulated work W is a stochastic process with independent increments. The time evolution of the work distribution $P_t(W)$ is described by the following Master equation:

$$\partial_t P_t(W) = \int_{-\infty}^{+\infty} T(\Delta W)\big(\mathrm{e}^{-\Delta W\partial_W} - 1\big)P_t(W)\,\mathrm{d}\Delta W \tag{13}$$

The probability per unit time $T(\Delta W)$ for a change in W by an amount ΔW, obtained following the basic methods of kinetic theory, is given by:

$$\begin{aligned} T(\Delta W) = \frac{\rho}{\sqrt{32\pi m^3 kT}}\Bigg[& A\frac{|\Delta W|}{V^2}H\left(\frac{\Delta W}{-V}\right)\mathrm{e}^{\frac{-m}{2kT}(V-\frac{\Delta W}{2mV})^2} \\ &+ \frac{2B|\Delta W|}{V^2\sin^2\theta_0}H\left(\frac{\Delta W}{V\sin\theta_0}\right)\mathrm{e}^{\frac{-m}{2kT}(V\sin\theta_0-\frac{\Delta W}{2mV\sin\theta_0})^2}\Bigg] \end{aligned} \tag{14}$$

where H denotes the Heaviside function. One distinguishes two contributions in $T(\Delta W)$, one from the backside and another one from the front sides of the triangle. In the calculation of $T(\Delta W)$, the gas is assumed to have a uniform density ρ and a Maxwellian velocity distribution $\phi(v_x, v_y)$ at temperature T:

$$\phi(v_x, v_y) = \frac{m}{2\pi kT}\mathrm{e}^{-\frac{m}{2kT}(v_x^2+v_y^2)} \tag{15}$$

The exact solution to Eq. (13) can be readily obtained by Fourier transform, as a result of the convolution structure of the integral kernel, which is typical for processes with independent increments. It is, however, convenient, before doing so, to switch to the dimensionless variables:

$$w = \beta W, \qquad v = V\left(\frac{\beta m}{2}\right)^{1/2}, \qquad \tau = \frac{(A+2B)\rho t}{(2\beta m)^{1/2}} \tag{16}$$

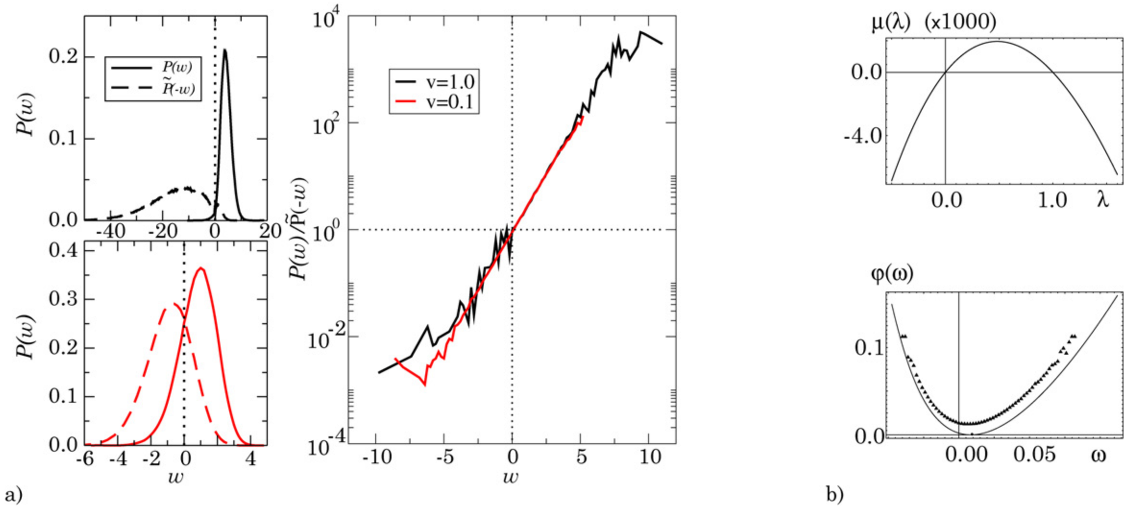

Fig. 2. (a) Histograms of $P(w)$, $\tilde{P}(-w)$ and their ratio, obtained from molecular dynamics. Parameter values are $v = 1$ and $v = 0.1$ with $|v|\tau = 10$ fixed. (b) Plot of the cumulant generating function $\mu(\lambda)$ and its Legendre transform, the large deviation function $\varphi(\omega)$ for $v = -0.1$. The solid lines correspond to the theoretical results; the dots come from hard sphere molecular dynamics simulations with $\tau = 100$.

being the work and speed measured in terms of the thermal energy and speed of the gas particles, and the time in terms of the average time between collisions. Furthermore, the Fourier transform is presented here under the form of the cumulant generating function $\mu(\lambda)$, defined by:

$$e^{-\tau\mu(\lambda)} = \langle e^{-\lambda w}\rangle = \int_{-\infty}^{+\infty} e^{-\lambda w} P_\tau(w)\,dw \tag{17}$$

The expression of $\mu(\lambda)$ follows directly by multiplying the master equation (13) with $e^{-\lambda w}$ and integrating over w (with obviously $w = 0$ at the initial time), leading to:

$$\begin{aligned}\mu(\lambda) ={}& \frac{A}{A+2B} v\big[\mathrm{erfc}\big[(1-2\lambda)v\big](1-2\lambda)e^{4\lambda(\lambda-1)v^2} - \mathrm{erfc}[v]\big] \\ &+ \frac{2B}{A+2B} v\sin\theta_0\big[1+\mathrm{erf}[v\sin\theta_0] - \big(1+\mathrm{erf}\big[(1-2\lambda)v\sin\theta_0\big]\big)(1-2\lambda)e^{4\lambda(\lambda-1)v^2\sin^2\theta_0}\big]\end{aligned} \tag{18}$$

Expanding $\mu(\lambda)$ around $\lambda = 0$ gives the cumulants of the random variable w. In particular, the first moment is given by:

$$\begin{aligned}\langle w\rangle ={}& \tau v \frac{A}{A+2B}\left[\frac{4ve^{-v^2}}{\sqrt{\pi}} - 2\big(1+2v^2\big)\big(\mathrm{erfc}[v]\big)\right] \\ &+ \tau v \frac{2B}{A+2B}\left[\frac{4v\sin^2\theta_0 e^{-v^2\sin^2\theta_0}}{\sqrt{\pi}} + 2\sin\theta_0\big(1+2v^2\sin^2\theta_0\big)\big(1+\mathrm{erf}[v\sin\theta_0]\big)\right]\end{aligned} \tag{19}$$

The dissipation goes to zero for $V \to 0$ (with Vt fixed), in agreement with the quasi-static prediction. For non-zero but small V, the dissipation has a familiar Joule expression, with dissipation rate $\langle\dot{W}\rangle = \gamma V^2$, where

$$\gamma = 4\rho\sqrt{\frac{k_B T m}{2\pi}}\big(A + 2B\sin^2\theta_0\big) \tag{20}$$

is the friction coefficient.

We now turn to the verification of the Crooks relation, which is in fact immediate. For the present problem, with $\Delta F = 0$, the Crooks relation reads $P(w) = \exp(w)\tilde{P}(-w)$. This implies the following symmetry relation for $\mu(\lambda)$:

$$\mu(\lambda) - \tilde{\mu}(1-\lambda) = 0 \tag{21}$$

Taking into account that the inverse process corresponds to the inversion of the velocity, inspection of Eq. (18) allows to verify that this symmetry condition is indeed satisfied. We finally point out that the asymptotic behaviour of $P_\tau(w)$ for large time $\tau \to \infty$, is described in terms of the large deviation function $\varphi(\omega)$ ($\omega = w/\tau$):

$$\varphi(\omega) = -\lim_{t\to\infty} \frac{1}{\tau} \ln P_\tau(\omega\tau) \tag{22}$$

This positive convex function can be obtained as the Legendre transform of the generating function $\mu(\lambda)$:

$$\varphi(\omega) = \sup_\lambda \bigl\{\mu(\lambda) - \lambda\omega\bigr\} \tag{23}$$

The symmetry relation (21) for $\mu(\lambda)$ implies the following corresponding property of the large deviation function:

$$\varphi(\omega) - \tilde{\varphi}(-\omega) = \omega \tag{24}$$

$\varphi(\omega)$ has a unique zero at the most probable value $\omega = \langle w\rangle/\tau$. In Fig. 2(b), we have reproduced both $\mu(\lambda)$ and $\varphi(\omega)$.

4.2. *Compression and expansion of an ideal gas*

The compression and expansion of an ideal gas is one of the first illustrations in classes on thermodynamics and statistical mechanics. The quasi-static work $W_{\mathrm{rev}} = -\int P\,\mathrm{d}V$ is probably one of the best know expressions from these theories. It is therefore surprising, even more so than for the Joule experiment, that the exact probability distribution for the work has never been calculated, that is, until recently [22,20]. This calculation provides another interesting test of the work theorem. In the discussion presented below, we will focus on the limiting case of an ideal gas in an infinitely large container, sealed at one end by a movable piston, cf. Fig. 3. The case of a finite container has been discussed in detail in [22]. In the latter situation, a gas particle will in the course of time perform a sequence of correlated recollisions with the piston. While it is obviously of theoretical interest to see that the work theorem remains valid is this extreme situation, one does not expect that the result has much to do with what happens in a real gas (except in the case of an extremely small container with just a few gas particles). For an infinitely large container, a particle that has collided with the piston in our scenario (see below) will never collide again, hence there is a complete absence of correlated recollisions. It is expected that this opposite limit will give results which compare well with the behavior of a real gas. This is corroborated in the comparison with hard disk molecular dynamics simulations. The position of the piston plays the role of the parameter λ in the derivation of the work theorem. For simplicity, we will focus here on a piston moving at constant speed V for a time t. In this case ΔF is given by:

$$\Delta F = -NkT \ln\left[\frac{(L + Vt)}{L}\right] \tag{25}$$

N is the total number of gas particles, T the temperature of the gas, and L the length of the container. In order to ensure the molecular chaos assumption, one considers the thermodynamic limit, with both $N, L \to \infty$ such that $N/(LA) = \rho$ is kept fixed. A is the length of the piston in contact with the gas, and ρ the density of the gas. In this limit, the free energy change reduces to:

$$\Delta F = -\rho AkTVt \tag{26}$$

In complete analogy with the Joule experiment, the work done by the outside world during the motion of the piston is determined by the uncorrelated collisions of the gas particles with the piston. In fact, the work distribution satisfies the same master equation as that shown in Eq. (13). The only difference with the Joule experiment lies in the calculation

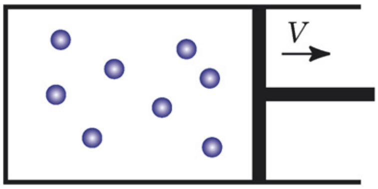

Fig. 3. An ideal gas is expanded (compressed) by moving the piston at constant speed V for a fixed time interval t. Due to the collisions with the piston, the gas loses (expansion) or gains (compression) kinetic energy.

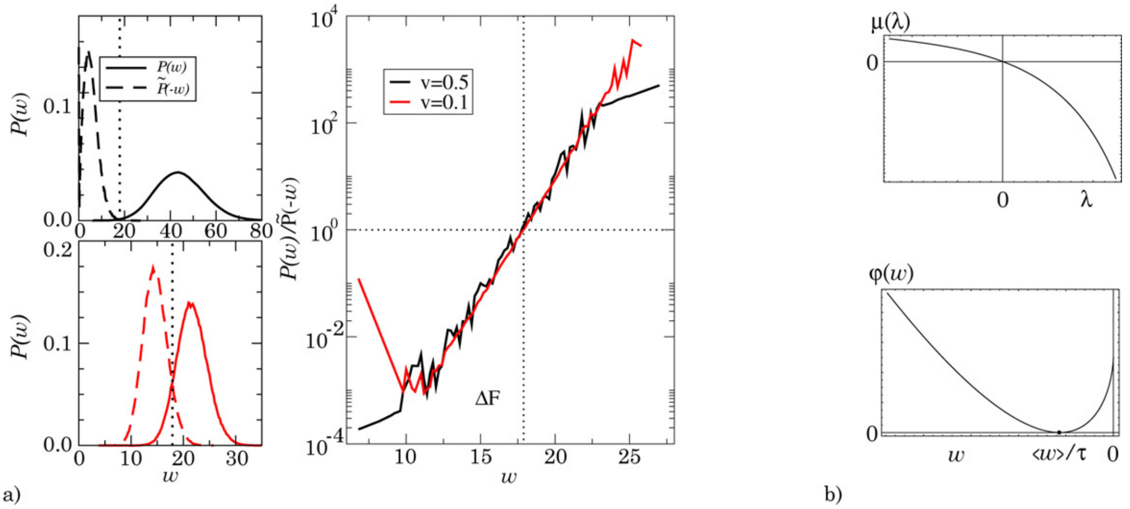

Fig. 4. (a) Histograms of $P(w)$, $\tilde{P}(-w)$ and their ratio, obtained from molecular dynamics. Parameter values are $v = 0.5$ and $v = 0.1$ with $|\Delta F| = 2|v|\tau = 20$ fixed. (b) Plot of the cumulant generating function $\mu(\lambda)$ and its Legendre transform, the large deviation function $\varphi(\omega)$, for the expansion process of an ideal gas ($v > 0$).

of the transition rate. As is obvious from Fig. 3, only the left side of the piston is in contact with the gas. Moreover, the contribution of the left side is exactly the same as the one for the base of the triangle. Hence, the expression for $T(\Delta W)$ is obtained by putting $B \to 0$ in Eq. (14). In fact, all the results for the piston problem can be obtained from the corresponding results in the Joule experiment by taking the limit $B \to 0$. Hence, introducing the dimensionless variables defined in Eq. (16), the explicit analytical expression of $\mu(\lambda)$ is found to be:

$$\mu(\lambda) = v(1-2\lambda)\,\mathrm{erfc}\big[v(1-2\lambda)\big]\mathrm{e}^{-4\lambda(1-\lambda)v^2} - v\,\mathrm{erfc}[v] \tag{27}$$

In the new variables, the free energy difference (26) becomes $\Delta F \to -2v\tau$. The Crooks relation then implies the following symmetry relation for $\mu(\lambda)$:

$$\mu(\lambda) - \tilde{\mu}(1-\lambda) = -2v \tag{28}$$

which holds, as can be easily verified.

The above theoretical results compare very well with molecular dynamics simulations of a dilute two-dimensional gas of hard disks. As an additional check of the Crooks relation, we have generated, by an average over half a million runs, the histogram of the probability distributions $P(W)$ and $\tilde{P}(W)$. Results are shown in Fig. 4(a). Note that the Crooks relation allows one to determine the value of ΔF as the (unique) W-value for which the two distributions intersect, $P(\Delta F) = \tilde{P}(-\Delta F)$. The value obtained for ΔF in the simulations yields $\Delta F \approx 18$, whereas the theoretical value equals 20. This small deviation can be attributed to finite size effects, including correlated collisions, and to the non-ideality corrections for a hard disk gas.

The average work performed on the gas is:

$$\langle w\rangle = \tau v\left[\frac{4v\mathrm{e}^{-v^2}}{\sqrt{\pi}} - 2\big(1+2v^2\big)\big(\mathrm{erfc}[v]\big)\right] \tag{29}$$

In the quasistatic limit, with $v \to 0$, it coincides with ΔF.

By Legendre transform of $\mu(\lambda)$, see Eq. (23), we obtain the large deviation function $\varphi(\omega)$, cf. Fig. 4(b). Note that, unlike the Joule experiment, the function $\varphi(\omega)$ is only defined for $\omega < 0$ when $v > 0$ (expansion: the gas performs work) and for $\omega > 0$ when $v < 0$ (compression: the gas receives work).

4.3. *Effusion of an ideal gas*

Effusion is the escape of a gas out of a large container through a small pore or opening. This phenomenon has received a lot of attention since the early days of kinetic theory of gases. It is also not without practical applications.

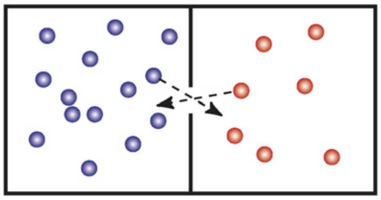

Fig. 5. The system under consideration: two ideal gases, in equilibrium at their respective temperatures and densities, can exchange energy and particles through a small hole in the common adiabatic wall.

Before discussing effusion from the point of view of the fluctuation theorem, we briefly review some of its salient features. It is intuitively clear that the flux of particles, leaving the hole, will primarily be determined by the speed of particles. From the equipartition relation $m\langle v^2\rangle/2 = 3/2kT$, it follows that this speed is $\sim\sqrt{kT/m}$. The $1/\sqrt{m}$ dependence is referred to as Graham's law [23]. It implies that in a gas mixture, the lighter species will leak out faster. This technique was used in the production of the atomic bomb during the Manhattan project for the enrichment of uranium. It is also intuitively clear that faster particles have more chance to escape. In fact, one verifies that the average kinetic energy of escaping particles is equal to $2kT$. In other words, effusion can and is used as a mechanism for cooling [24]. In the early 1900s, Knudsen investigated other aspects of effusion [25]. He discovered the cosine law, corresponding to the fact that the particles exit with an isotropic angular distribution, a property which was subsequently using for coating spherical bulbs. This property is also related to the Kirchoff's laws of radiation and to proper boundary conditions for particles that are thermally reemitted from an absorbing wall [26].

When effusion is taking place between two reservoirs, cf. Fig. 5, it forms an interesting case study for irreversible thermodynamics. Indeed, as the linear dimension of the opening is assumed to be smaller than the mean free path of the gas particles, the leak does not disturb the state of the gas. If the latter is at equilibrium, it will remain so and many properties of the effusing gas particles can be easily calculated, using basic arguments from kinetic theory of gases. While local equilibrium (inside each container) is preserved, the existing density and temperature gradients, which can be arbitrarily large, will lead to both particle and energy fluxes. A somewhat surprising observation, already noted by Knudsen, is that for $\rho_A\sqrt{T_A} = \rho_B\sqrt{T_B}$ an energy flux is present in the absence of a corresponding particle flux, cf. Eq. (42) (this peculiar property also appears in the so-called Jepsen gas [27]). Similarly, a particle flux is present with no corresponding energy flux when $\rho_A {T_A}^{3/2} = \rho_B {T_B}^{3/2}$, cf. Eq. (43).

The presence of fluctuations in the effusion process is obvious, as the transfer of energy and particles is a result of individual particles crossing the opening. In order to study these fluctuations, we focus on the following generic setup. Two infinitely large reservoirs (A and B) are separated by a common (adiabatic) wall. Each reservoir contains an ideal gas at equilibrium, with uniform density ρ_i and Maxwellian velocity distribution $\phi_i(\vec{v})$ at temperature T_i, $i \in \{A, B\}$ (cf. Fig. 5):

$$\phi_i(\vec{v}) = \left(\frac{m}{2\pi kT_i}\right)^{3/2} \mathrm{e}^{-\frac{mv^2}{2kT_i}} \tag{30}$$

Starting from this initial equilibrium configuration, we open a small opening (surface area σ) in the common wall, and measure the net transfer of energy ΔU and of particles ΔN from A to B during a time interval of length t. Assuming that this small transfer does not perturb the equilibrium states of both reservoirs, the entropy change in the total system ΔS can be written as:

$$\Delta S = \Delta S_A + \Delta S_B = -\frac{1}{T_A}\Delta U + \frac{\mu_A}{T_A}\Delta N + \frac{1}{T_B}\Delta U - \frac{\mu_B}{T_B}\Delta N = \mathcal{A}_U\Delta U + \mathcal{A}_N\Delta N \tag{31}$$

We introduced, in accordance with the definitions from irreversible thermodynamics, the following thermodynamic forces (affinities) for energy and particle flow, respectively:

$$\mathcal{A}_U = \frac{1}{T_B} - \frac{1}{T_A}; \qquad \mathcal{A}_N = \frac{\mu_A}{T_A} - \frac{\mu_B}{T_B} = k\log\left[\frac{\rho_A}{\rho_B}\left(\frac{T_B}{T_A}\right)^{3/2}\right] \tag{32}$$

In the last equality, we used the expression for the chemical potential of an ideal gas. We will show below that the joint probability density $P_t(\Delta U, \Delta N)$ satisfies the following fluctuation theorem:

$$\frac{P_t(\Delta U, \Delta N)}{P_t(-\Delta U, -\Delta N)} = \mathrm{e}^{\Delta S/k} \tag{33}$$

Since the increments of both ΔU and ΔN are again independent, entailing a cumulant generating function of the form:

$$\left\langle \mathrm{e}^{-(\lambda_U \Delta U + \lambda_N \Delta N)} \right\rangle = \mathrm{e}^{-t\mu(\lambda_U, \lambda_N)} \tag{34}$$

this fluctuation theorem can be expressed as the following symmetry relation:

$$\mu(\lambda_U, \lambda_N) = \mu(\mathcal{A}_U/k - \lambda_U, \mathcal{A}_N/k - \lambda_N) \tag{35}$$

Eq. (33) implies the usual fluctuation theorem:

$$\frac{P_t(\Delta S)}{P_t(-\Delta S)} = \mathrm{e}^{\Delta S/k} \tag{36}$$

Note that the steady state which is reached immediately after opening the hole, has no preferred direction of time, so that $\tilde{P} = P$.

Eq. (33) also implies fluctuation theorems for energy and particle transport separately, when the corresponding thermodynamic force for the other process vanishes. Indeed, in this case only one of the variables, ΔU or ΔN, appears in ΔS. Bringing the denominator from lhs to rhs in Eq. (33), subsequent integration over the other variable implies:

$$\frac{\mathcal{P}_t(\Delta U)}{\mathcal{P}_t(-\Delta U)} = \mathrm{e}^{\Delta S/k} \quad \text{when } \mathcal{A}_N = 0 \tag{37}$$

and

$$\frac{\mathcal{P}_t(\Delta N)}{\mathcal{P}_t(-\Delta N)} = \mathrm{e}^{\Delta S/k} \quad \text{when } \mathcal{A}_U = 0 \tag{38}$$

We now proceed to a direct verification of the detailed fluctuation theorem (35) and its implications Eqs. (36)–(38) by an explicit evaluation of the generating function $\mu(\lambda_U, \lambda_N)$.

The probability density $P_t(\Delta U, \Delta N)$ obeys the following Master equation:

$$\begin{aligned}\frac{\partial}{\partial t} P_t(\Delta U, \Delta N) = {} & \int_0^\infty T_{A\to B}(E) P_t(\Delta U - E, \Delta N - 1)\,\mathrm{d}E + \int_0^\infty T_{B\to A}(E) P_t(\Delta U + E, \Delta N + 1)\,\mathrm{d}E \\ & - P_t(\Delta U, \Delta N) \int_0^\infty \left[T_{A\to B}(E) + T_{B\to A}(E)\right] \mathrm{d}E\end{aligned} \tag{39}$$

where the transition probabilities $T_{A\to B}(E)$ and $T_{B\to A}(E)$, to observe a particle with kinetic energy $\frac{1}{2}mv^2 = E$ crossing the hole from $A \to B$ and $B \to A$ respectively, are given by:

$$T_{A\to B}(E) = \frac{\sigma\rho_A}{\sqrt{2\pi m k T_A}} \frac{E}{kT_A} \mathrm{e}^{-\frac{E}{kT_A}}; \qquad T_{B\to A}(E) = \frac{\sigma\rho_B}{\sqrt{2\pi m k T_B}} \frac{E}{kT_B} \mathrm{e}^{-\frac{E}{kT_B}} \tag{40}$$

Notice the $1/\sqrt{m}$ dependence, in accordance with Graham's law. Again, the integral operators are of the convolution type, in agreement with the fact that the processes ΔU and ΔN have independent increments. Calculation of the cumulant generating function $\mu(\lambda_U, \lambda_N)$ leads to the following exact result [21]:

$$\mu(\lambda_U, \lambda_N) = \frac{\sigma\sqrt{k}}{\sqrt{2\pi m}} \left(\rho_A \sqrt{T_A} \left[1 - \frac{\mathrm{e}^{-\lambda_N}}{(1 + kT_A\lambda_U)^2} \right] + \rho_B \sqrt{T_B} \left[1 - \frac{\mathrm{e}^{\lambda_N}}{(1 - kT_B\lambda_U)^2} \right] \right) \tag{41}$$

One easily verifies, using the explicit expressions for the thermodynamic forces given in Eq. (32), that this expression indeed verifies the detailed fluctuation theorem, Eq. (35). The average energy and particle transfer, being the first order cumulants, are given as:

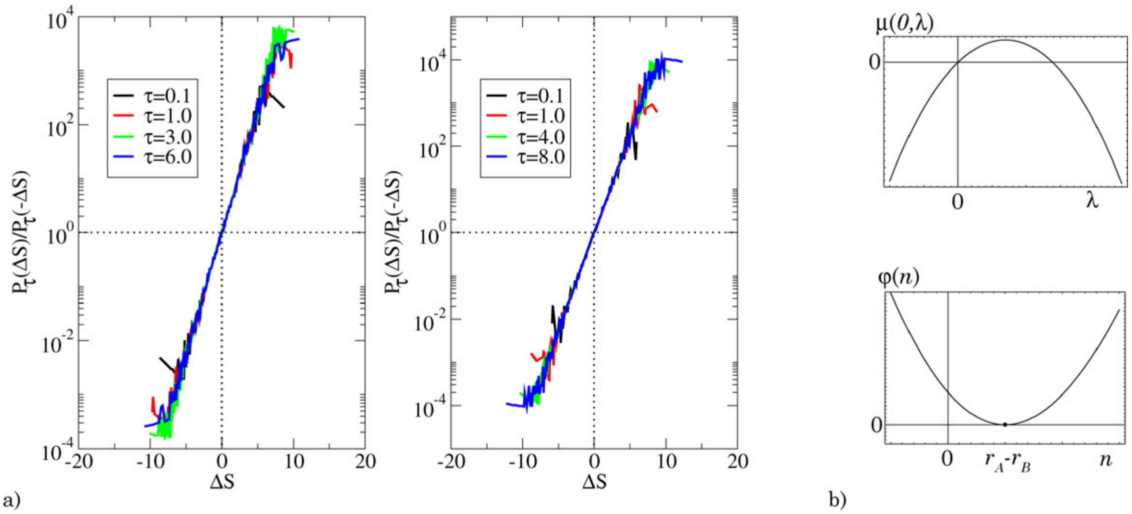

Fig. 6. (a) Plot of the ratio $P_\tau(\Delta S)/P_\tau(-\Delta S)$ for different values of τ, obtained from molecular dynamics. Left panel: $2N_A = N_B = 2000$ and $T_A = 4T_B = 1$. Right panel: $N_A = 2N_B = 2000$ and $T_A = 2T_B = 1$. (b) Plot of the cumulant generating function $\mu(0,\lambda)$ and its Legendre transform, the large deviation function $\varphi(n)$.

$$\langle \Delta U \rangle = \frac{t\sigma k^{3/2}}{\sqrt{2\pi m}} 2\left(\rho_A T_A^{3/2} - \rho_B T_B^{3/2}\right) \tag{42}$$

$$\langle \Delta N \rangle = \frac{t\sigma k^{1/2}}{\sqrt{2\pi m}} \left(\rho_A T_A^{1/2} - \rho_B T_B^{1/2}\right) \tag{43}$$

nicely reproducing the findings of Knudsen mentioned before.

The effusion process considered here has the additional property that explicit results can also be obtained in the original variables. For example, a close inspection of the master equation (39) reveals that the time evolution of $\mathcal{P}_t(\Delta N)$ is identical to that of a biased random walk in 1D taking steps to the right and left with probabilities per unit time:

$$r_A = \sigma\rho_A\sqrt{kT_A/2\pi m}; \qquad r_B = \sigma\rho_B\sqrt{kT_B/2\pi m} \tag{44}$$

respectively. Its exact time dependent solution is therefore given by the well known expression in terms of the modified Bessel functions. This relation is intuitively clear since the passages of particles from reservoir A to B and vice-versa induce a random walk on the variable ΔN. The asymptotic behaviour of $\mathcal{P}_t(\Delta N)$ for large time $t \to \infty$ is described in terms of the variable $n \equiv \Delta N/t$, and characterised by its large deviation function $\varphi(n)$:

$$\mathcal{P}_t(\Delta N) \sim \mathrm{e}^{-t\varphi(n)} \tag{45}$$

which can either be obtained as the Legendre transform of the generating function $\mu(0,\lambda)$:

$$\varphi(n) = \sup_\lambda \left\{\mu(0,\lambda) - \lambda n\right\} \tag{46}$$

or by making an asymptotic expansion of the Bessel functions for large orders. The final result reads:

$$\varphi(n) = r_A + r_B - \sqrt{4r_A r_B + n^2} - n \ln\left(\frac{\sqrt{4r_A r_B + n^2} - n}{2r_B}\right) \tag{47}$$

A sketch of both $\mu(0,\lambda)$ and $\varphi(n)$ is given in Fig. 6(b). The function $\varphi(n)$ is positive everywhere, and has a single zero at the most probable value $n = \bar{n} = r_A - r_B$.
Similar explicit calculations can be performed for the energy transport, cf. [21].

5. Discussion

The fluctuation and work theorem are interesting but also intriguing results whose status and range of applications has not yet been fully understood. It is also not clear whether these results are of any practical use, since they express a connection between probable and exponentially less probable events. However, it is possible, using similar ideas and techniques, to make new statements about average quantities such as dissipated work and entropy production [28]. This and other developments are certainly deepening our understanding of nonequilibrium statistical mechanics.

References

[1] L. Onsager, Phys. Rev. 37 (1931) 405.
[2] S. de Groot, P. Mazur, Nonequilibrium Thermodynamics, Holland Publishing Co., Amsterdam, 1969.
[3] D. Evans, E.G.D. Cohen, G.P. Morris, Phys. Rev. Lett. 71 (1993) 2401.
[4] G. Gallavotti, E.G.D. Cohen, Phys. Rev. Lett. 74 (1995) 2694.
[5] C. Jarzynski, Phys. Rev. Lett. 78 (1997) 2690.
[6] G.E. Crooks, Phys. Rev. E 60 (1999) 2721.
[7] J. Kurchan, J. Phys. A 31 (1998) 3719.
[8] J.L. Lebowitz, H. Spohn, J. Stat. Phys. 95 (1999) 333.
[9] C. Maes, J. Stat. Phys. 95 (1999) 367.
[10] T. Hatano, S. Sasa, Phys. Rev. Lett. 86 (2001) 3463.
[11] D.J. Evans, D.J. Searles, Adv. Phys. 51 (2002) 1529.
[12] U. Seifert, Phys. Rev. Lett. 95 (2005) 040602.
[13] D. Andrieux, P. Gaspard, J. Chem. Phys. 121 (2004) 6167.
[14] D. Andrieux, P. Gaspard, J. Stat. Mech. (2006) P01011.
[15] F. Ritort, Semin. Poincaré 2 (2003) 195.
[16] B. Derrida, B. Doucot, P.-E. Roche, J. Stat. Phys. 115 (2004) 717.
[17] B. Duplantier (Ed.), Einstein aujourd'hui, EDP Sciences, 2005.
[18] G.E. Crooks, C. Jarzynski, Phys. Rev. E. 75 (2007) 021116.
[19] B. Cleuren, C. Van den Broeck, R. Kawai, Phys. Rev. Lett. 96 (2006) 050601.
[20] I. Bena, C. Van den Broeck, R. Kawai, Europhys. Lett. 71 (2005) 879.
[21] B. Cleuren, C. Van den Broeck, R. Kawai, Phys. Rev. E. 74 (2006) 021117.
[22] R.C. Lua, A.Y. Grosberg, J. Phys. Chem. B 109 (2005) 6805.
[23] T. Graham, Philos. Trans. R. Soc. London 136 (1846) 573.
[24] A.K. Shukla, S. Banik, R.S. Dhaka, C. Biswas, S.R. Barman, H. Haak, Rev. Sci. Instrum. 75 (2004) 4467.
[25] M. Knudsen, Ann. Phys. (Leipzig) 28 (1909) 75.
[26] A.L. Garcia, Numerical Methods for Physics, Prentice Hall, Englewood Cliffs, NJ, 2000.
[27] V. Balakrishnan, C. Van den Broeck, Phys. Rev. E 72 (2005) 046141.
[28] R. Kawai, J.M.R. Parrondo, C. Van den Broeck, Phys. Rev. Lett. 98 (2007) 080602.