

A work fluctuation theorem for a Brownian particle in a non confining potential

Christoph Streibnig, Holger Kantz
Max Planck Institute for the Physics of Complex Systems
Nöthnitzer Straße 38 D-01187 Dresden

Using the Feynman-Kac formula, a work fluctuation theorem for a Brownian particle in a non confining potential, e.g., a potential well with finite depth, is derived. The theorem yields an inequality that puts a lower bound on the average work needed to change the potential in time. In comparison to the Jarzynski equality, which holds for confining potentials, an additional term describing a form of energy related to the never ending diffusive expansion appears.

I. INTRODUCTION

Thermal equilibrium is one of the most fundamental concepts in statistical mechanics. Roughly speaking it is a state where time does no longer appear in any of the relevant macroscopic observables. These equilibrium states are very well studied and there exists a set of basic statistical and thermodynamic statements about them. Let us list some of them for the simple special case of an overdamped one dimensional Brownian particle in thermal equilibrium with a heat bath of temperature T and inside a potential $V(x)$. Equilibrium statistical mechanics tells us that the probability density function (PDF) of the particles position is Boltzmann distributed and hence given by

$$P_B(x) = \frac{e^{-V(x)\beta}}{Z_V}. \quad (1)$$

Here $\beta = 1/k_B T$ and Z_V is a normalization factor. From equilibrium thermodynamics we know that an isothermal and quasi-static transition from one equilibrium state to another, which in our case is done by changing the potential from V_1 to V_2 , consumes an average amount of Energy in the form of work W given by

$$\langle W \rangle = \Delta F. \quad (2)$$

Where ΔF is the Helmholtz free energy difference between the initial and the final state. Recall that ΔF is connected to the normalization factor via $\Delta F = -k_B [\log(Z_{V_2}) - \log(Z_{V_1})]$. Also note, due to the stochastic nature of the system W is a random variable and $\langle \dots \rangle$ denotes the expectation value. Relaxing the quasi-static assumption the above equality (2) becomes an inequality

$$\langle W \rangle \geq \Delta F, \quad (3)$$

which can be derived by applying the Clausius inequality, a manifestation of the second law of thermodynamics, to the first law of thermodynamics. Surprisingly the above inequality can also be derived by a more fundamental equality, namely the Jarzynski equality [1]

$$\langle e^{-\beta(W-\Delta F)} \rangle = 1. \quad (4)$$

This equality belongs to a family of so-called integral fluctuation theorems. In the past years a number of integral

and so-called detailed fluctuation theorems for different cases have been discovered, see [2–11] for further reading. Since changing the potential with nonzero speed drives the system away from equilibrium, inequality (3) and the Jarzynski equality (4) are actually out of equilibrium results. Hence it is only required that the system starts in equilibrium and the final equilibrium state exists. The emphasis here is on exists, W does not care if the system relaxes back to equilibrium after the potential has been changed. Now for some systems equilibrium states do not exist. For our simple case, thermal equilibrium can be reached under the condition that the system is enclosed by a potential which diverges faster than logarithmically in space, e.g., a harmonic potential or hard reflective walls. We will call such potentials confining. In principle, since most of the fundamental forces (weak, electromagnetic, gravity) are not diverging it is natural to assume that in reality confining potentials are very exotic. In most cases they are only local approximations of globally non confining potentials, for example a harmonic potential can approximate the Lennard-Jones potential around its minimum.

The general question that this article is trying to tackle is the following: Do thermodynamic equalities and inequalities, structurally similar to the Jarzynski equality(4), and the lower bound (3), also exist in non-confined systems? Or in other words, how important is it to confine the system in order to get these fundamental results? To seek for a general answer is most certainly too ambitious, hence we constrain ourselves to the special case of a Brownian particle inside an asymptotically flat potential which goes to zero at least as fast as $1/x$ and is changed in time via an external protocol. This choice is mainly motivated by the following already existing results. It was shown in [12, 13] that for these kind of systems, assuming that the potential is time independent, to leading-order in the long time limit, the PDF $P(x, t)$ assumes the shape

$$P(x, t) \approx P_{GB}(x, t) = \frac{e^{-\frac{x^2}{4Dt} - \beta V(x)}}{N(t)}, \quad (5)$$

where $N(t)$ is the normalization constant which is $\sim \sqrt{t}$ for sufficiently large t . Eq. (5), has a simple intuitive explanation: The Gaussian factor in the asymptotic shape of the PDF is dominant in the tails of the system, at $x > \sqrt{\pi Dt}$ where the potential is effectively zero whereas at small x and $t \gg 1$, the Gaus-

sian factor is = 1 and the Boltzmann factor is dominant. When $t \rightarrow \infty$, according to Eq. (5), the PDF approaches a non-normalizable Boltzmann infinite invariant density [12, 13] (see also the related works [14, 15]) $\lim_{t \rightarrow \infty} N(t)P_t(x) \rightarrow \exp(-V(x)/k_B T)$, which replaces the standard Boltzmann distribution in its role in determining integrable physical observables such as energy and occupation times, and leads to infinite ergodic theory, see e.g., [14, 16–19].

II. SETTING THE STAGE

We begin with the overdamped Langevin dynamics of a Brownian particle in an external potential field

$$\dot{x}_t = -\frac{V'(x_t, \lambda_t)}{\gamma} + \sqrt{2D} \xi_t, \quad (6)$$

where $V(x_t, \lambda_t)$ is a potential depending on an externally controlled protocol λ_t , and D , γ , ξ_t are respectively the diffusion constant, the friction and Gaussian white noise with zero mean and

$$\langle \xi_t \xi_{t'} \rangle = \delta(t - t'). \quad (7)$$

Furthermore $V(x_t, \lambda_t)$ is assumed to be an asymptotically flat potential well which falls off at least as rapidly as $1/x$ hence

$$\lim_{x \rightarrow \pm\infty} V(x, \lambda_t) = 0. \quad (8)$$

The evolution of the PDF $P(x, t)$ is given by

$$\partial_t P(x, t) = LP(x, t), \quad (9)$$

where L is the Fokker-Planck operator

$$L = \left(D\partial_x^2 + \frac{1}{\gamma}\partial_x V' \right). \quad (10)$$

For a fixed λ_t and sufficiently long times $P(x, t)$ converges to [13]

$$P_{\text{GB}}(x, t, \lambda_t) = \frac{e^{-\frac{x^2}{4Dt} - \beta V(x, \lambda_t)}}{N(t, \lambda_t)}, \quad (11)$$

here $\beta = \frac{1}{k_B T}$, k_B is the Boltzmann constant and $N(t, \lambda_t)$ is the normalization constant

$$N(t, \lambda_t) = \int_{-\infty}^{\infty} e^{-\frac{x^2}{4Dt} - \beta V(x, \lambda_t)} dx. \quad (12)$$

Although we mentioned in the introduction that for large enough t , $N(t, \lambda_t) \sim \sqrt{t}$, we choose to keep the full normalization constant since it leads to faster convergence.

The particular scenario that we consider throughout this article is the following. At $t = 0$ the particle is placed inside the potential well. From $t = 0$ to $t = t_0$ the system relaxes such that at $t = t_0$ the density is

approximately given by $P_{\text{GB}}(x, t_0)$, Eq. (11). From $t = t_0$ to $t = t_1$ the potential is changing according to an externally controlled protocol λ_t . At $t = t_1$ the potential stops changing and in principle the system relaxes back to a state described by (11). The relaxation in the end however will not play a role in the results. In this scenario the work done by the protocol along a trajectory up to time t is given by

$$W_t = \int_{t_0}^t \dot{\lambda}_\tau \frac{\partial V(x_\tau, \lambda_\tau)}{\partial \lambda_\tau} d\tau = \int_{t_0}^t \frac{\partial V(x_\tau, \lambda_\tau)}{\partial \lambda_\tau} d\lambda_\tau. \quad (13)$$

III. A MOTIVATING SPECIAL CASE: THE INFINITELY FAST PROTOCOL

Let us start by considering a simple special case where the potential changes instantaneously. This can be expressed mathematically by stating that the change of the potential $V(x, \Theta(t - t_0))$ in time is only through a heaviside/theta function $\Theta(t - t_0)$. The natural choice for the protocol here is

$$\lambda_t = \Theta(t - t_0). \quad (14)$$

Introducing the abbreviate notation

$$\Delta V(x) := V(x, 1) - V(x, 0) \quad (15)$$

we write the potential as

$$V(x, \lambda_t) = V(x, 0) + \lambda_t \Delta V(x). \quad (16)$$

According to (13) the trajectory dependent work is then given by the difference between the potential after and before the change evaluated at x_{t_0} ,

$$W_t = \Delta V(x_{t_0}). \quad (17)$$

As mentioned in the introduction we are interested in a Jarzynski like equality. Due to the simple expression for the work we can straight forwardly calculate

$$\langle e^{-\beta W_t} \rangle = \int_{-\infty}^{\infty} e^{-\Delta V(x)} P_{\text{GB}}(x, t_0, 0) dx \quad (18)$$

$$= \frac{\int_{-\infty}^{\infty} e^{-\frac{x^2}{4Dt_0} - \beta V(x, 1)} dx}{N(t_0, 0)}. \quad (19)$$

Introducing a quantity ΔG analogue to the Helmholtz free energy difference

$$\Delta G = -\beta \ln \left(\frac{N(t_0, 0)}{N(t_0, 1)} \right) \quad (20)$$

$$= -\beta \ln \left(\frac{\int_{-\infty}^{\infty} e^{-\frac{x^2}{4Dt_0} - \beta V(x, 1)} dx}{\int_{-\infty}^{\infty} e^{-\frac{x^2}{4Dt_0} - \beta V(x, 0)} dx} \right) \quad (21)$$

we arrive at

$$\langle e^{-\beta W_t} \rangle = e^{-\beta \Delta G}. \quad (22)$$

Eq. (22) is analogous to Eq. (4), but in contrast to the standard Jarzynski equality, it is now valid even though the system has no equilibrium state. By the so called Jensen's inequality this relation yields

$$\langle W_t \rangle \geq \Delta G. \quad (23)$$

In the next section we will derive a version of (22) valid for arbitrary protocol speed.

IV. DERIVATION OF THE WORK FLUCTUATION THEOREM

Our derivation is essentially an adjusted version of an elegant derivation of the Jarzynski equality using the Feynman-Kac formula, first presented in [20]. Let us briefly state a version of the Feynman-Kac formula which is sufficient for our purpose, for a proof see [21]. Assume a Langevin process x_t whose phase space density $P(x, t) = \langle \delta(x_t - x) \rangle$ obeys

$$\partial_t P(x, t) = LP(x, t). \quad (24)$$

Here $\langle \dots \rangle$ denotes an average over all trajectories ending at time t and $\delta(x_t - x)$ being the delta-distribution picks out the ones that end at position x . The Feynman Kac formula then says that

$$g(x, t) = \langle \delta(x - x_t) e^{-\Omega_t} \rangle, \quad (25)$$

with

$$\Omega_t = \int_{t_0}^t f(x_\tau, t) d\tau, \quad (26)$$

being a stochastic functional obeys

$$\partial_t g(x, t) = Lg(x, t) - f(x, t)g(x, t). \quad (27)$$

Now we apply this statement to our case by making the initially arbitrary seeming choice

$$\Omega_t := \beta \left[W_t - \int_{t_0}^t \left(\frac{k_B T}{2\tau} + \frac{x_\tau F(x_\tau, \lambda_\tau)}{2\tau} \right) d\tau \right], \quad (28)$$

or equivalently

$$f(x, \tau) := \beta \left[\frac{\partial V(x, \tau, \lambda_\tau)}{\partial \tau} - \frac{k_B T}{2\tau} - \frac{x_\tau F(x_\tau, \lambda_\tau)}{2\tau} \right], \quad (29)$$

with $F = -V'$ being the force acting on the particle. Equation (27) then becomes

$$\begin{aligned} \partial_t g(x, t) &= Lg(x, t) \\ &+ \beta \left[\frac{k_B T}{2t} + \frac{x F(x, t)}{2t} - \dot{\lambda} \frac{\partial V(x, \lambda_t)}{\partial \lambda_t} \right] g(x, t). \end{aligned} \quad (30)$$

It can be verified by direct substitution that

$$g(x, t) = \frac{e^{-\frac{x^2}{4Dt} - \beta V(x, \lambda_t)}}{N(t_0, \lambda_{t_0})}, \quad (31)$$

solves (30) with the initial condition

$$g(x, t_0) \equiv P(x, t_0) = P_{\text{GB}}(x, t_0, \lambda_{t_0}). \quad (32)$$

However we also know from the Feynman-Kac formula that (25) with the particular choice made in (28) solves (30). Thus we have

$$\langle \delta(x - x_t) e^{-\Omega_t} \rangle = \frac{e^{-\frac{x^2}{4Dt} - \beta V(x, \lambda_t)}}{N(t_0, \lambda_{t_0})}, \quad (33)$$

which can be rewritten by defining a more general analogue of the Helmholtz free energy difference than (20)

$$\Delta G := -k_B T \ln \left(\frac{N(t, \lambda_t)}{N(t_0, \lambda_{t_0})} \right), \quad (34)$$

as

$$\langle \delta(x - x_t) e^{-(\Omega_t - \beta \Delta G)} \rangle = P_{\text{GB}}(x, t, \lambda_t). \quad (35)$$

Integration over x and using (28) gives a work integral fluctuation theorem

$$\left\langle e^{-\beta \left[W_t - \int_{t_0}^t \left(\frac{k_B T}{2\tau} + \frac{x_\tau F(x_\tau, \lambda_\tau)}{2\tau} \right) d\tau - \Delta G \right]} \right\rangle = 1. \quad (36)$$

which by applying the Jensen's inequality yields

$$\langle W_t \rangle \geq \Delta G + \left\langle \int_{t_0}^t \left(\frac{k_B T}{2\tau} + \frac{x_\tau F(x_\tau, \lambda_\tau)}{2\tau} \right) d\tau \right\rangle. \quad (37)$$

The Fluctuation theorem given by Eq.(36) is the central result of this article. Its physical meaning will be discussed in the next section.

V. A POSSIBLE PHYSICAL INTERPRETATION

Let us investigate the terms appearing in the exponent of the fluctuation theorem Eq.(36) in more detail. One major difference with respect to the Jarzynski equality is the additional trajectory dependent term

$$\int_{t_0}^t \left(\frac{k_B T}{2\tau} + \frac{x_\tau F(x_\tau, \lambda_\tau)}{2\tau} \right) d\tau. \quad (38)$$

Another minor difference is that the time dependence of ΔG is not only due to the protocol but also explicitly due to the Gaussian term in the normalization constant. It is clear that both of these discrepancies are a mathematical consequence of the non-equilibrium initial PDF. Using the Feynman-Kac derivation scheme, as presented in the previous section, one could in principle derive an integral fluctuation theorem similar to (36) for any kind

of non-equilibrium initial PDF. However, $P_{\text{GB}}(x, t_0, \lambda_{t_0})$ being the long time asymptotic density lets us expect that for a sufficiently slow protocol i.e. in the quasi-static limit $P(x, \tau, \lambda_\tau) = P_{\text{GB}}(x, \tau, \lambda_\tau)$ for $\tau \geq t_0$. We will support this claim later with numerical evidence. Let us now calculate

$$\langle \Omega_t \rangle_{\text{GB}} = \int_{t_0}^t d\tau \langle f(x, \tau, \lambda_\tau) \rangle_{\text{GB}} \quad (39)$$

$$= \int_{t_0}^t d\tau \int_{-\infty}^{\infty} dx f(x, \tau, \lambda_\tau) \frac{e^{-\frac{x^2}{4D\tau} - \beta V(x, \lambda_\tau)}}{N(\tau, \lambda_\tau)} \quad (40)$$

$$= - \int_{t_0}^t d\tau \frac{1}{N(\tau, \lambda_\tau)} \int_{-\infty}^{\infty} dx (\partial_t - L) e^{-\frac{x^2}{4D\tau} - \beta V(x, \lambda_\tau)} \quad (41)$$

$$= \Delta G. \quad (42)$$

Here $\langle \dots \rangle_{\text{GB}}$ denotes the expectation value with respect to $P_{\text{GB}}(x, \tau, \lambda_\tau)$ or in other words the expectation value in the quasi-static limit. Note, from line (40) to (41) Eq.(30) respectively Eq.(27) was used. Writing Ω_t explicitly using Eq.(28) we get

$$\langle W_t \rangle_{\text{GB}} = \Delta G + \left\langle \int_{t_0}^t \left(\frac{k_B T}{2\tau} + \frac{x_\tau F(x_\tau, \lambda_\tau)}{2\tau} \right) d\tau \right\rangle_{\text{GB}}. \quad (43)$$

The equation above shows that in the quasi-static limit inequality (37) becomes an equality. The analogue statement for confined systems is that for sufficiently slow protocols the system stays Boltzmann distributed which leads to $\langle W_t \rangle_{\text{B}} = \Delta F$, where $\langle \dots \rangle_{\text{B}}$ denotes the average with respect to the Boltzmann density (1). However there is a very intriguing difference between these two statements. For cyclic protocols, meaning $\lambda_{t_0} = \lambda_t$, applied to confined systems it is clear that $\langle W_t \rangle_{\text{B}} = 0$ since $\Delta F = 0$. Whereas for cyclic protocols applied to non confined systems it is not obvious from (43) whether $\langle W_t \rangle_{\text{GB}} = 0$. This raises the question if its possible to get $\langle W_t \rangle_{\text{GB}} \leq 0$ or more generally $\langle W_t \rangle \leq 0$. Or in other words is it possible to extract energy in the form of work by applying a cyclic protocol? It is important to realize that due to the never-ending diffusive process a cyclic protocol does not mean that the system itself returns to its initial state. For now we will leave this question open and approach it numerically in the next section. So far we can make the following conclusions:

$$\Delta G + \left\langle \int_{t_0}^t \left(\frac{k_B T}{2\tau} + \frac{x_\tau F(x_\tau, \lambda_\tau)}{2\tau} \right) d\tau \right\rangle \quad (44)$$

is a quantity that puts a lower bound on the average work needed to externally change the potential. In the quasi-static limit this quantity becomes the average work and if negative it is free to use for the external observer. It should be mentioned that due to the protocol dependence of second term it is not something like a free energy in the sense of an thermodynamic potential like the Helmholtz free energy.

Let us now focus on the second term in (44). Since it originates from the Gaussian part of $P_{\text{GB}}(x, \tau, \lambda_\tau)$ we claim, at least in the quasi-static limit, that it can be interpreted as an energy coming from the expansion of the system. And indeed it can be brought into a convenient form resembling pressure-volume work. In order to do that we first need to establish a notion of pressure. The osmotic pressure Π of a Brownian particle confined in a region of size L and inside a force field $F(x)$ is given by [22]

$$\Pi = \frac{1}{L} [k_B T + \langle x F(x) \rangle]. \quad (45)$$

Of course our system is not confined so it is questionable how to make use of the above expression, especially how to choose the size of the system. Nevertheless, choosing the length scale of diffusion $L_\tau = \sqrt{2D\tau}$ as a measure for the size of the system and introducing a quantity

$$p_\tau := \frac{1}{L_\tau} [k_B T + F(x_\tau, \lambda_\tau) x_\tau], \quad (46)$$

which can be seen as an analogue of Π but for a single particle, allows us to rewrite

$$\int_{t_0}^t \left(\frac{k_B T}{2\tau} + \frac{x_\tau F(x_\tau, \lambda_\tau)}{2\tau} \right) d\tau = \int_{L_{t_0}}^{L_t} p_\tau dL_\tau. \quad (47)$$

Here we have substituted $\tau = L_\tau^2/(2D)$ in the integral and used definition (46). Consequently Eq.(36) and inequality (37) can be written as

$$\left\langle e^{-\beta [W_t - \int_{L_{t_0}}^{L_t} p_\tau dL_\tau - \Delta G]} \right\rangle = 1, \quad (48)$$

and

$$\langle W_t \rangle \geq \Delta G + \left\langle \int_{L_{t_0}}^{L_t} p_\tau dL_\tau \right\rangle. \quad (49)$$

We agree that the structure of the integral in (48) could just be a nice coincidence. However let us present another argument. Assume a one dimensional Brownian particle with diffusion coefficient \tilde{D} and temperature \tilde{T} inside a confining potential $\tilde{V}(x, t)$ given by

$$\tilde{V}(x, \tau) = V(x, \tau) + \frac{x^2}{4D\tau} k_B T. \quad (50)$$

Note $V(x, \tau)$ is as before a non confining potential but $\tilde{V}(x, \tau)$ is now enclosed by an additional harmonic potential which opens up with time. In the quasi-static limit the PDF of the system is given by $P_{\text{GB}}(x, \tau)$ and is thus indistinguishable from our non-confined system.

The average work in the confined system yields

$$\langle \tilde{W}_t \rangle_{\text{GB}} = \left\langle \int_{t_0}^t \frac{\partial \tilde{V}(x_\tau, \tau)}{\partial \tau} d\tau \right\rangle_{\text{GB}} \quad (51)$$

$$= \left\langle \int_{t_0}^t \partial_\tau V(x_\tau, \tau) d\tau \right\rangle_{\text{GB}} - k_B T \left\langle \int_{t_0}^t \frac{x_\tau^2}{4D\tau^2} d\tau \right\rangle_{\text{GB}} \quad (52)$$

$$= - \int_{t_0}^t d\tau \left\langle \partial_\tau V(x, \tau) - k_B T \frac{x^2}{4D\tau^2} \right\rangle_{\text{GB}} \quad (53)$$

$$= \langle W_t \rangle_{\text{GB}} - \left\langle \int_{t_0}^t \left(\frac{k_B T}{2\tau} + \frac{x_\tau F(x_\tau, \tau)}{2\tau} \right) d\tau \right\rangle_{\text{GB}} \quad (54)$$

$$= \langle W_t \rangle_{\text{GB}} - \left\langle \int_{L_{t_0}}^{L_t} p_\tau dL_\tau \right\rangle_{\text{GB}}, \quad (55)$$

where Eq. (30) was used to get from line (53) to line (54), note the vanishing boundary terms. The above calculation shows that in the quasi-static limit the work done by opening the harmonic potential coincides with the path dependent part of the expansion energy of the non-confined system. The main difference between the two forms of energy is that in the confined system work is assumed to be externally controllable. In the non-confined system a part of the explicit time dependence comes from the inherent diffusive expansion and is hence not assumed to be externally controllable.

VI. EXAMPLES

As an example for our theory we choose the inverted Gaussian potential well

$$V(x, \lambda_\tau) = -A(\lambda_\tau) e^{-\frac{(x-B(\lambda_\tau))^2}{2}}, \quad (56)$$

whose depth $A(\lambda_\tau)$ or location $B(\lambda_\tau)$ is changed in time by the protocol λ_τ . A convenient way to show the integral fluctuation theorem Eq.(36) is by showing it indirectly via verifying Eq.(35). We proceed in the following manner. An ensemble of $n_{\text{trajectory}}$ trajectories is generated using the standard Euler-Maruyama method with an time increment of $\Delta\tau$ and initial position $x = 0$. At $\tau = t_0$ a PDF is constructed and checked if it has converged to $P_{\text{GB}}(x, t_0, \lambda_{t_0})$. At the end of the protocol which is at $\tau = t$, the PDF is checked again to make sure that it is now different from $P_{\text{GB}}(x, t, \lambda_t)$, which should be the case for sufficiently fast protocols. The PDF's are simply constructed as histograms from the ensemble. To verify Eq.(35) we have to recall that expectation values for stochastic processes are path integrals, namely we can write: $\langle \delta(x - x_t) e^{-\Omega[x_\tau] - \beta\Delta G} \rangle = \int \mathcal{D}[x_\tau] \delta(x - x_t) e^{-\Omega[x_\tau] - \beta\Delta G} p[x_\tau]$, where $p[x_\tau] \mathcal{D}[x_\tau]$ is a measure for the probability to observe a trajectory x_τ .

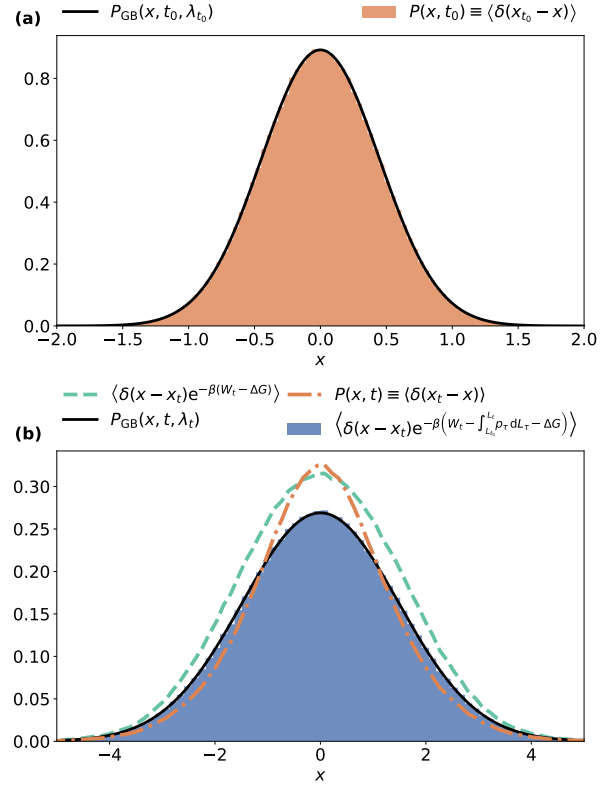


FIG. 1. Numerical results for Brownian particle inside a potential given by (56), parameters are chosen as follows: $D = k_B T = \gamma = 1$, $A(\tau) = \theta(\tau - t_0) \sin(\frac{\tau - t_0}{t - t_0} \pi)$, $B(\tau) = 1$, $t_0 = 0.1$, $t = 1$, $\Delta\tau = 10^{-3}$, $n = 10^6$. (a) Shows a comparison at time $\tau = t_0$ of the analytic expression (11) for asymptotic long time density $P_{\text{GB}}(x, t_0, \lambda_{t_0})$ (black solid line) with a histogram (orange filled histogram) constructed from an ensemble of numerically generated trajectories representing the PDF $P(x, t_0)$. (b) Shows a comparison at $\tau = t$ of the analytic expression (11) for the asymptotic long time density $P_{\text{GB}}(x, t, \lambda_t)$ (black solid line) with three different histograms. Each of these histograms is constructed from an ensemble of numerically generated trajectories. The dash dotted orange line represents the regular PDF $P(x, t)$. The blue filled histogram represents the left-hand side of Eq. (35), the path probabilities are thus re-scaled by $\exp(-\beta [W_t - \int_{L_{t_0}}^{L_t} p_\tau dL_\tau - \Delta G])$. The green dashed line represents a histogram where path probabilities are re-scaled by $\exp(-\beta [W_t - \Delta G])$, emphasizing the relevance of the "pressure-volume" term".

Plugging this into Eq. (35) yields

$$\int \mathcal{D}[x_\tau] \delta(x - x_t) e^{-\Omega[x_\tau] - \beta\Delta G} p[x_\tau] = \frac{e^{-\frac{x^2}{4D\tau} - \beta V(x, \lambda_t)}}{N(t_0, \lambda_{t_0})}. \quad (57)$$

The form of (57) allows us to interpret $e^{-\Omega[x_\tau] - \beta\Delta G}$ as an additional weight on the path-probability. Therefore, if we multiply the increment that one particle adds to the height of a bin in the PDF's histogram by $e^{-\Omega[x_\tau] - \beta\Delta G}$ we get an histogram representing the left hand side of Eq.(35). And if Eq.(35) is correct this histogram should match $P_{\text{GB}}(x, t, \lambda_t)$. The results for different cases of

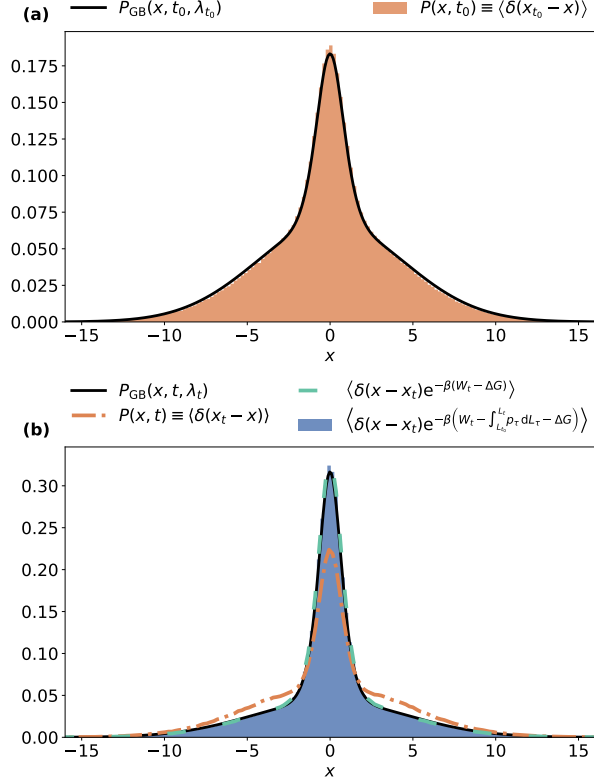


FIG. 2. Same scenario as described in Fig.1 with parameters chosen as follows: $D = k_B T = \gamma = 1$, $A(\tau) = \theta(\tau - t_0) \frac{\tau - t_0}{t - t_0} + 1$, $B(\tau) = 1$, $t_0 = 10$, $t = 11$, $\Delta\tau = 10^{-3}$, $n = 10^6$.

$A(\tau)$ and $B(\tau)$ are displayed in Fig.(1-3). Let us briefly discuss them.

In Fig.1 $A(\tau) = \theta(t_0 - \tau) \sin(\frac{\tau - t_0}{t - t_0} \pi)$ and $B(\tau) = 1$, which means the particle freely diffuses during the initial relaxation meaning $P(x, t_0) = P_{GB}(x, t_0, \lambda_{t_0})$ is exact for an arbitrary small t_0 and the derivation of Eq.(35) is exact as well. As such we can see this case as a test of the numerical procedure more than a test of the analytical results. Fig. 1 (a) shows the expected agreement of the PDF with $P_{GB}(x, t_0, \lambda_{t_0})$. Fig.1 (b) shows that in the end of the protocol $P(x, t, \lambda_t) \neq P_{GB}(x, t, \lambda_t)$ furthermore it clearly verifies Eq.(35) and shows the significance of $\int_{L_{t_0}}^{L_t} p_\tau dL_\tau$.

In Fig.2 $A(\tau) = \theta(t_0 - \tau) \frac{\tau - t_0}{t - t_0} + 1$ and $B(\tau) = 1$. Contrary to the previous case there is a potential during the initial relaxation. This means t_0 has to be chosen sufficiently big such that $P_{GB}(x, t_0, \lambda_{t_0}) \approx P(x, t, \lambda_t)$. Fig.2(a) shows that for the particularly chosen parameters $t_0 = 10$ suffices. As before Fig.2(b) verifies Eq.(35) however $\int_{L_{t_0}}^{L_t} p_\tau dL_\tau$ seems to be negligible.

In Fig.3 $A(\tau) = 5$ and $B(\tau) = 5\theta(\tau - t_0) \frac{\tau - t_0}{t - t_0}$, again there is a potential during the initial relaxation which we chose to be $5 k_B T$ deep. Instead of changing the amplitude $A(\tau)$ we are now changing the location of the

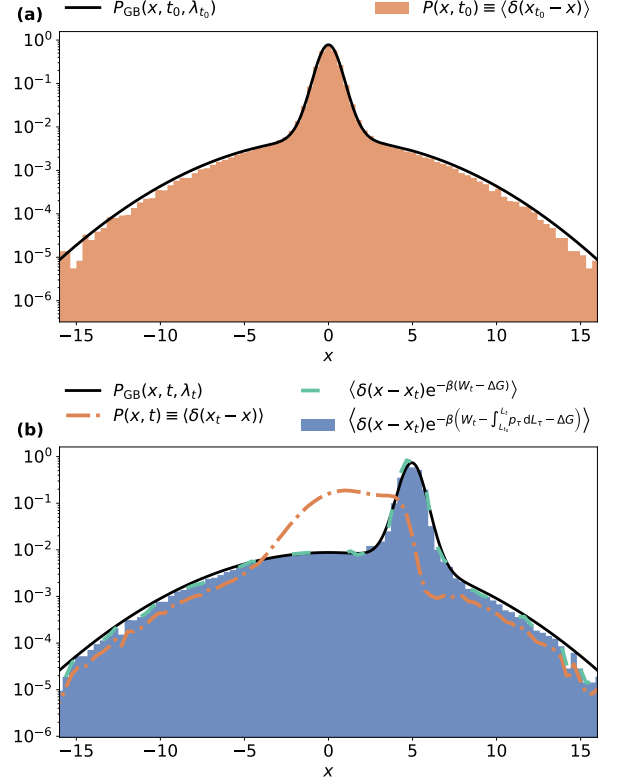


FIG. 3. Same scenario as described in Fig.1 with parameters chosen as follows: $D = k_B T = \gamma = 1$, $A(\tau) = 5$, $B(\tau) = 5\theta(\tau - t_0) \frac{\tau - t_0}{t - t_0}$, $t_0 = 10$, $t = 11$, $\Delta\tau = 10^{-3}$, $n = 10^6$. Note the logarithmically scaled y axis.

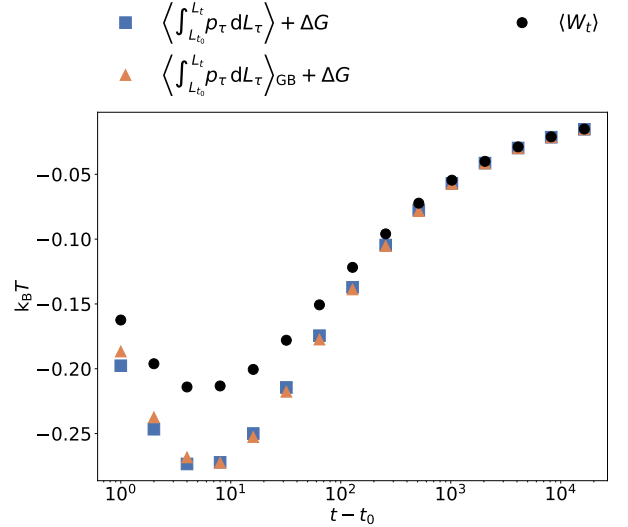


FIG. 4. Average work (black dots) and expression (44) (blue squares) vs. duration of the protocol $t - t_0$. The orange triangles are displaying the semi-analytical calculation of the right-hand-side of Eq. 43. Parameters are chosen as follows: $D = k_B T = \gamma = 1$, $A(\tau) = \theta(\tau - t_0) \sin(\frac{\tau - t_0}{t - t_0} \pi)$, $B(\tau) = 1$, $t_0 = 0.5$, $\Delta\tau = 10^{-3}$, $n = 10^5$.

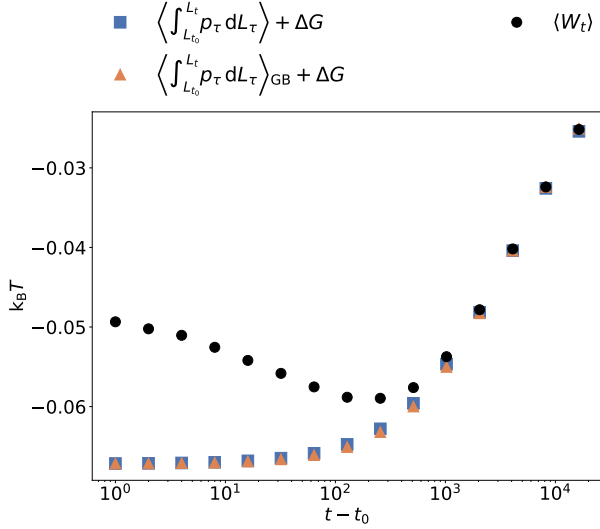


FIG. 5. Average work (black dots) and expression (44) (blue squares) vs. duration of the protocol $t - t_0$. The orange triangles are displaying the right-hand-side of Eq. 43. Parameters are chosen as follows: $D = k_B T = \gamma = 1$, $A(\tau) = \theta(\tau - t_0) \frac{\tau - t_0}{t - t_0} + 1$, $B(\tau) = 1$, $t_0 = 10^3$, $\Delta\tau = 10^{-3}$, $n = 10^5$

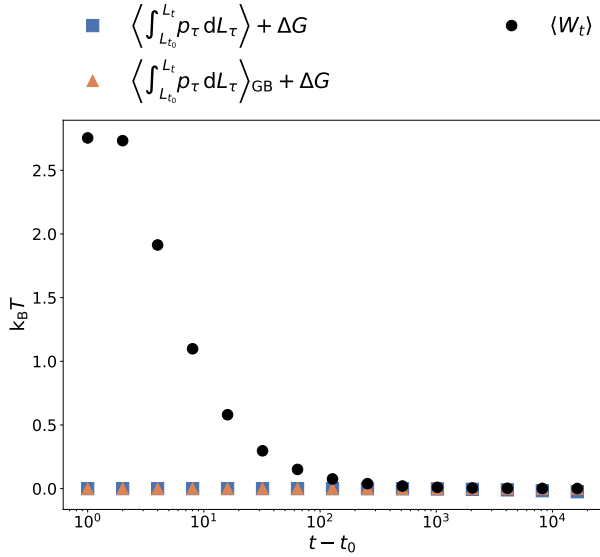


FIG. 6. Average work (black dots) and expression (44) (blue squares) vs. duration of the protocol $t - t_0$. The orange triangles are displaying the right-hand-side of Eq. 43. Parameters are chosen as follows: $D = k_B T = \gamma = 1$, $A(\tau) = 5$, $B(\tau) = 5\theta(\tau - t_0) \frac{\tau - t_0}{t - t_0}$, $t_0 = 10^4$, $\Delta\tau = 10^{-3}$, $n = 10^5$

potential. As in the previous case Fig.3(a) shows good agreement of $P_{GB}(x, t_0, \lambda_{t_0})$ with $P(x, t, \lambda_t)$ and Fig.3(b) verifies Eq.(35). However also in this case $\int_{L_{t_0}}^{L_t} p_\tau dL_\tau$ seems to be negligible.

In order to verify Eq. (43) and inequality (37) we investigate the same examples as the quasi static limit is

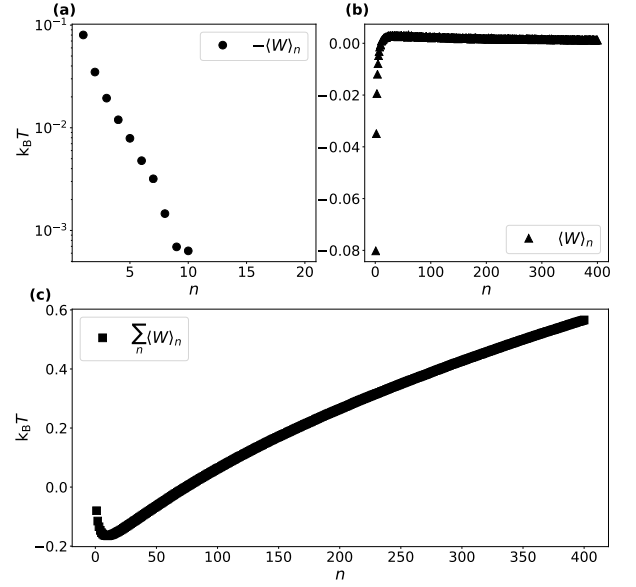


FIG. 7. Behavior of the average work $\langle W \rangle_n$ per cycle n for a cycle duration of 1. (a) Shows the initially exponential behavior. For $n > 9$, $-\langle W \rangle_n < 0$ and hence can not be displayed in the semi-logarithmic plot. (b) Shows that $\langle W \rangle_n$ goes to a value slightly above zero. This results in a linear increasing cumulative sum $\sum_n \langle W \rangle_n$, as can be seen in (c).

approached, simply by making the duration of the protocol successively larger. The results depicted in Fig. (4-6) are in good agreement with Eq. (43) and inequality (37). Note that in Fig. (4-6) the initial relaxation time t_0 was chosen large enough in order to reduce the error from approximating the initial distribution with $P_{GB}(x, t_0)$. It is also important to realize that in Fig.4 and Fig.5 $|\langle W \rangle - \int_{L_{t_0}}^{L_t} p_\tau dL_\tau - \Delta G|$ goes to zero faster than $\langle W \rangle$.

Interestingly the average work for the sinusoidal changing $A(\tau)$ is always negative even though the change of the potential is cyclic, see Fig.4. As already discussed in the previous section this is not possible for confined systems since it would violate the second law of thermodynamics. However, for our non confined systems this is per se not a violation of the second law since the system does not return to its original state. Repeating a cycle n times does not necessarily lead to an infinite energy output, it depends on how the average work per cycle $\langle W \rangle_n$ behaves with n . And indeed as one can see from Fig.7(a,b), $\langle W \rangle_n$ increases exponentially fast and decays to zero from above after a small but positive value has been reached. This behavior leads to a positive total work $\langle W \rangle = \sum_n \langle W \rangle_n > 0$, for large enough n , see Fig. 7(c).

VII. CONCLUSION AND DISCUSSION

We have derived a work fluctuation theorem, see Eq. (36), similar to the Jarzynski equality but applicable to

a Brownian particle inside a potential well with finite depth that is changed in time by an external protocol. Such systems are not able to reach thermal equilibrium which is reflected in the fluctuation theorem by an additional path dependent term (38) besides work. The inequality that results from this fluctuation theorem puts a fundamental lower bound on the work that is needed to change the potential in time. It is expected to become an equality in the quasi-static limit which gives the new term the meaning of an energy that can be extracted from the never ending diffusive spreading of the system.

The only approximation in the derivation of Eq. (36) is done by approximating the PDF at the start of the protocol with the long time asymptotic density $P_{\text{GB}}(x, t_0, \lambda_{t_0})$ given by Eq.(11). Our theory would be exact, if the density at the beginning of the protocol were exactly the Gauss-Boltzmann density $P_{\text{GB}}(x, t_0, \lambda_{t_0})$. This approximation is the better the longer the initial time evolution. So for every finite time evolution t_0 also relation (36) is only an approximation. At first glance this seems to be a disadvantage in comparison to the Jarzynski equality. Here the Boltzmann density, which is an exact solution of the Fokker-Planck equation, is assumed to describe the system at the start of the protocol. However this line of thought is misleading. In Brownian dynamics simulations or an experiment one would need to wait infinitely long for a confined system to reach a state which is exactly described by the Boltzmann density. In that sense assuming that a confined system can be described by the Boltzmann density is as much of an approximation as assuming that a non-confined system can be described by $P_{\text{GB}}(x, t_0, \lambda_{t_0})$. The rate of convergence however might be different.

A major open question is how Eq. (36) relates to stochastic thermodynamics and one of its main results, the Seifert fluctuation theorem [3]. Considering the simple special case of free Brownian motion it is easy to show that they do not coincide. Furthermore the inequality

implied by Seifert's theorem becomes an equality if the system is time reversible, inequality (37) on the other hand is expected to become an equality if the protocol is quasi-static. Now, for stochastic systems the Jarzynski equality can be seen as a special case of Seifert's fluctuation theorem. Its implied inequality becomes an equality in the quasi-static limit which in this case is also the time reversible limit. The conclusion here would be that for non-confined systems time reversibility is no longer implied by quasi-staticity. Intuitively this is simply a consequence of the never ending diffusive spreading. However, in order to make a more definite statement further investigations are required.

Another question is when the term (38) in the fluctuation theorem (36) becomes irrelevant? It does not appear in the special case of the infinitely fast protocol, see Eq. (23). It also seems to be irrelevant in the numerical examples where the initial relaxation time is much longer than the duration of the protocol, see Fig.2(b) and Fig.3(b). Both of these results point in the direction that (38) is negligible if the initial relaxation time is long compared to the duration of the protocol.

Yet another question is how general these type of work fluctuation theorems are. In principle, the mathematical procedure based on the Feynman-Kac formula can be applied to any long time asymptotic initial PDF. Consequently the difficult part in deriving such a fluctuation theorem is to find this PDF. Some already existing and usable results for further research are presented in [14, 23].

VIII. ACKNOWLEDGEMENT

We are grateful to E.Aghion, E. Barkai, S. Bo and E. Lutz for valuable discussions.

-
- [1] C. Jarzynski, Physical Review Letters **78**, 2690 (1997).
 - [2] U. Seifert, Reports on progress in physics **75**, 126001 (2012).
 - [3] U. Seifert, Physical review letters **95**, 040602 (2005).
 - [4] T. Sagawa and M. Ueda, Physical review letters **104**, 090602 (2010).
 - [5] T. Hatano and S.-i. Sasa, Physical review letters **86**, 3463 (2001).
 - [6] G. E. Crooks, Physical Review E **60**, 2721 (1999).
 - [7] M. Esposito and C. Van den Broeck, Physical review letters **104**, 090601 (2010).
 - [8] L. Dabelow, S. Bo, and R. Eichhorn, Physical Review X **9**, 021009 (2019).
 - [9] C. Jarzynski, Annu. Rev. Condens. Matter Phys. **2**, 329 (2011).
 - [10] I. Neri, Physical Review Letters **124**, 040601 (2020).
 - [11] I. Neri, É. Roldán, S. Pigolotti, and F. Jülicher, Journal of Statistical Mechanics: Theory and Experiment **2019**, 104006 (2019).
 - [12] E. Aghion, D. A. Kessler, and E. Barkai, Physical review letters **122**, 010601 (2019).
 - [13] E. Aghion, D. A. Kessler, and E. Barkai, Chaos, Solitons & Fractals **138**, 109890 (2020).
 - [14] A. Dechant, E. Lutz, E. Barkai, and D. Kessler, Journal of Statistical Physics **145**, 1524 (2011).
 - [15] X. Wang, W. Deng, and Y. Chen, The Journal of chemical physics **150**, 164121 (2019).
 - [16] J. Aaronson, *An introduction to infinite ergodic theory*, 50 (American Mathematical Soc., 1997).
 - [17] N. Leibovich and E. Barkai, Physical Review E **99**, 042138 (2019).
 - [18] P. Meyer and H. Kantz, Physical Review E **96**, 022217 (2017).
 - [19] T. Akimoto, E. Barkai, and G. Radons, Physical Review E **101**, 052112 (2020).

- [20] G. Hummer and A. Szabo, Proceedings of the National Academy of Sciences **98**, 3658 (2001).
- [21] E. Boksenbojm, B. Wynants, and C. Jarzynski, Physica A: Statistical Mechanics and its Applications **389**, 4406 (2010).
- [22] J. F. Brady, The Journal of chemical physics **98**, 3335 (1993).
- [23] L. Defaveri, C. Anteneodo, D. A. Kessler, and E. Barkai, arXiv preprint arXiv:2004.04325 (2020).