Jarzynski equality: Connections to thermodynamics and the second law

Benoit Palmieri and David Ronis*
Department of Chemistry, McGill University, 801 Sherbrooke Ouest, Montréal, Québec, Canada H3A 2K6
(Received 6 September 2006; published 31 January 2007)

The one-dimensional expanding ideal gas model is used to compute the exact nonequilibrium distribution function. The state of the system during the expansion is defined in terms of local thermodynamics quantities. The final equilibrium free energy, obtained a long time after the expansion, is compared against the free energy that appears in the Jarzynski equality. Within this model, where the Jarzynski equality holds rigorously, the free energy change that appears in the equality does not equal the actual free energy change of the system at any time of the process. More generally, the work bound that is obtained from the Jarzynski equality is an upper bound to the upper bound that is obtained from the first and second laws of thermodynamics. The cancellation of the dissipative (nonequilibrium) terms that result in the Jarzynski equality is shown in the framework of response theory. This is used to show that the intuitive assumption that the Jarzynski work bound becomes equal to the average work done when the system evolves quasistatically is incorrect under some conditions.

DOI: 10.1103/PhysRevE.75.011133 PACS number(s): 05.20.−y, 05.70.Ln, 05.40.−a

I. INTRODUCTION

Only a few relations are exactly satisfied for processes in systems far from thermodynamic equilibrium. Of the most recent of these relations is the Jarzynski equality, which relates the nonequilibrium average work done by a driving force on a system initially at equilibrium to the free energy difference between two equilibrium states of the system. This equality was first derived classically by Jarzynski [1–3] and later extended to stochastic system by Crooks [4]. Quantum mechanical Jarzynski equalities have also been investigated by Mukamel [5] and Esposito and Mukamel [6]. The Jarzynski equality has also been extensively tested numerically and analytically for various models; e.g., Marathe and Dhar [7] studied spin systems, while Oostenbrink and van Gunsteren [8] studied the redistribution of charges, creation and annihilation of neutral particles, and conformational changes in molecules. In each case the Jarzynski equality was confirmed. The Jarzynski equality has also been extensively tested for ideal gas expansions by Pressé and Silbey [9], Lua and Grosberg [10], and Bena, Van den Broeck, and Kawai [11]. Finally, the Jarzynski equality has been verified experimentally by Liphardt *et al.* [12] by stretching single RNA molecules.

On the other hand, the validity of the Jarzynski equality is still under debate. For example, see the objections of Cohen and Mauzerall in Refs. [13,14]; in particular, one of their concerns involves the temperature appearing in the Jarzynski equality,

$$\langle e^{\beta W}\rangle = \frac{Z_b(\beta)}{Z_a(\beta)} \equiv e^{-\beta \Delta A_J}, \tag{1.1}$$

where W is the work done by the system for a process that brings the systems from an equilibrium initial state with work parameter a to a final state with work parameter b, $\langle \cdots \rangle$ is a nonequilibrium average, $\beta \equiv 1/(k_B T)$, and ΔA_J is obtained, as above, by the ratio of two canonical partition functions defined at the same temperature, but for the two different work parameters a and b. Their claim was that the temperature in the last equation is defined without foundation since the temperature is undetermined during irreversible processes and often differs from the initial temperature. Although this statement is correct, it is equally clear that the Jarzynski equality is exact provided that the system is canonically distributed initially and that β is defined in terms of the initial equilibrium temperature of the system. In fact, the derivation by Jarzynski in Ref. [3] is fairly general and, at least in the original presentation, is based solely on Liouville's theorem, when the free energy change describes the system plus any bath, although in the event that the bath is explicitly included in the dynamics several complications arise—specifically, in extracting the system free energy change from the total and with the identification of work and heat exchange between system and bath [15].

In this work, we accept the Jarzynski relation as a mathematical identity and note that it has been recognized as a tool for calculating free energies [8,16]. Here, we study the connections between the Jarzynski equality and nonequilibrium thermodynamics. We will focus on systems that are governed by Hamiltonian dynamics. For such systems, the process driven by external forces is always adiabatic since the Hamiltonian presumably contains everything, and provided Liouville's equation is valid, the derivation of the identity in Ref. [3] shows unambiguously that $\langle e^{\beta W}\rangle$ is equal to the ratio of two partition functions defined at the same temperature. As above, we call this ratio $e^{-\beta \Delta A_J}$. One of the questions we will answer in this paper is the following: In general, how is ΔA_J related to the true free energy changes at the end of the process, where, as pointed out in Ref. [13], the temperature, if it can be defined, has usually changed. Moreover, it is sometimes believed that the Jarzynski equality can be used to prove thermodynamics bounds [11] (i.e., the second law) from mechanics. We will show here that this is not true. These questions will be addressed both generally using thermodynamics or response theory and within the context of some of the simple numerical models considered in the literature.

*Author to whom correspondence should be addressed. Electronic address: David.Ronis@McGill.ca.

We will also show how the Jarzynski equality works in the context of response theory. In particular, we will show that the nonequilibrium average of the work equals $-\Delta A_J$ plus so-called dissipative terms. These dissipative terms are then successively canceled by the higher-order cumulants of the work. Response theory also provides a framework which will then be used to derive the specific conditions under which $-\Delta A_J$ becomes the true upper bound to the average work. More specifically, we show that when the process is quasistatic *and* leaves the basic quantities that define the ensemble (e.g., temperature, density, chemical potential, etc.) unchanged, the Jarzynski work bound becomes the lowest upper bound to the work; in that case, the dissipative terms vanish. For other kinds of processes, in particular for adiabatic ones like the one-dimensional (1D) expanding ideal gas, the Jarzynski work bound is an upper bound to the thermodynamic work upper bound, no matter how slowly the process is carried out. This fact has already been noted numerically by Oberhofer, Dellago, and Geissler [16] when they compared numerical schemes for calculating equilibrium free energies based on the Jarzynski equality to those computed using the Widom insertion method in a soft-sphere liquid [17]. This was also observed by Jarzynski [2] for the isolated harmonic oscillator model where the natural frequency is increased as a function of time. Here, we show this results from general thermodynamic considerations or within the context of response theory.

The paper is divided as follows. In Sec. II, we define the one-dimensional gas model and we calculate the full nonequilibrium distribution function when the gas is expanding. Within this exactly solvable model, we characterize the actual state of the system during and after the expansion in terms of local thermodynamic quantities and we compare the final-state free energy with ΔA_J. In Sec. III, we compare the work bound obtained from the Jarzynski relation, invoking the Gibbs-Bogoliubov-Jensen-Peierls inequality, against the work bound imposed by thermodynamics and we show that Jarzynski's bound is less restrictive. Within the ideal expanding gas model, we calculate the average work as a function of the rate of the expansion and we compare the final free energy difference with ΔA_J. In Sec. IV, we use response theory to show how the Jarzynski equality works and, in particular, how the nonequilibrium terms cancel. We also explain how the dissipative terms can contribute to the work even if the process is carried out quasistatically. We also argue that this contribution of the dissipative terms could be used to explain why the average work does not equal $-\Delta A_J$ in Fig. 3 A of Liphardt *et al.* [12] when the work is performed very slowly. Section V contains a discussion and some concluding remarks.

II. ONE-DIMENSIONAL EXPANDING IDEAL GAS

In this section, we consider the one-dimensional expanding ideal gas. As has been shown by many authors [9–11], this model is fully consistent with the Jarzynski equality with

$$\Delta A_J = -k_B T_i \ln(L_f/L_i), \tag{2.1}$$

where T_i is the initial temperature of the gas, k_B is Boltzmann's constant, and L_i and L_f are the initial and final lengths of the box confining the gas, respectively. Here, ΔA_J represents the free energy difference between two states at the same temperature, but having different lengths. To follow the expansion of the gas, we work with the same model as in Refs. [10,11], which were inspired by the earlier work of Jepsen [18] and of Lebowitz and Percus [19]. In this model, the gas is initially at equilibrium in a box of length $L=L_i$. This box is closed from the left by a hard wall and from the right by an infinitely massive piston. At time $t=0$, the piston starts to move to the right with velocity V. Here, we will investigate the nonequilibrium process and see how it is related to ΔA_J. As mentioned above, this model as already been studied in detail in Refs. [10,11]; here, we use it to study several issues that were missed in these references—in particular, how the quasistatic limit arises, how the work relates to the maximum work predicted by thermodynamics, and what, if anything, the work distribution is really telling us about the actual state of the system. The ideal gas model here is used to raise questions, which will be investigated in more general terms in Secs. III and IV.

The complete knowledge of the system for $t>0$ can be obtained from the nonequilibrium distribution function $f(x,u;t)$, where, as usual, x is the position and u the velocity of the gas. For the expansion of the ideal gas, this distribution function can be obtained by solving the Liouville or Boltzmann equation [20,21] with no external forces and, of course, no collisions—i.e.

$$\frac{\partial f(x,u;t)}{\partial t} + u\frac{\partial f(x,u;t)}{\partial x} = 0, \tag{2.2}$$

with the boundary conditions

$$f(L_i + Vt, u; t) = f(L_i + Vt, 2V - u; t) \tag{2.3}$$

for $u>V$ and

$$f(0,u;t) = f(0,-u;t) \tag{2.4}$$

for all u. These boundary conditions account for the change in velocity of a gas particle hitting the piston at $x=L_i+Vt$ [cf. Eq. (2.3)] or the stationary wall at $x=0$ [cf. Eq. (2.4)]. As in the general presentation of the Jarzynski relation, we assume that the initial equilibrium distribution function is canonical—specifically,

$$f_0(x,u) = \frac{1}{L_i}\left(\frac{\beta}{2\pi}\right)^{1/2} e^{-\beta u^2/2}\Theta(L_i - x)\Theta(x), \tag{2.5}$$

where $\beta=1/T_i$, $\Theta(x)$ is the Heaviside step function, and where, henceforth, we set Boltzmann's constant (k_B) and the mass of the gas particles to 1. The product of step functions guarantees that the gas is initially confined between $x=0$ and $x=L_i$.

Rather than solving Liouville's equation, it is easier to get a more direct solution. The derivation follows the ideas of Refs. [10,22,23], and the solution is expressed as an infinite sum over n, where n is the number of collisions a gas particle makes with the piston. The details of the derivation are given in the Appendix, and the final expression for $f(x,u;t)$ is

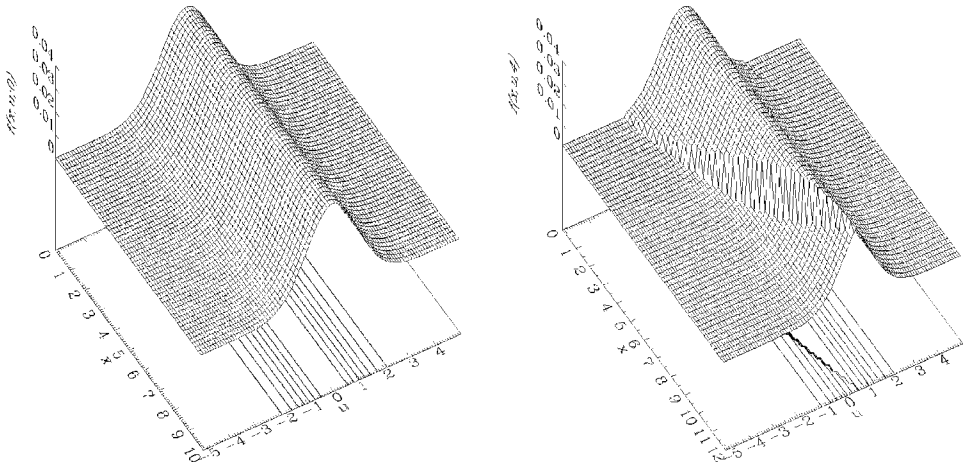

FIG. 1. The initial equilibrium distribution function at $t=0$ is shown on the left and the nonequilibrium distribution function at $t=4$ is shown on the right. Here, $L_i=10$, $V=0.5$, and $T_i=1$ and the moving piston is on the right. The contours are drawn when $f(x,u)=0.005$, 0.01, 0.015, 0.02, 0.025, 0.03, and 0.035.

$$f(x,u;t) = \Theta(L_i + Vt - x)\Theta(x)\Big\{ f_0(x - ut, u) + f_0(-x + ut, -u) + \sum_{n=1}^{\infty} [f_0(-x + ut + 2nL_i, 2nV - u) + f_0(x - ut + 2nL_i, 2nV + u) + f_0(x - ut - 2nL_i, u - 2nV) + f_0(-x + ut - 2nL_i, -u - 2nV)] \Big\}, \tag{2.6}$$

where the first and second terms on the right-hand side represent, respectively, free streaming and a single collision with the back wall. The four terms in the sum account for, respectively, an initially positive velocity resulting in n collisions with the piston followed by free streaming, an initially positive velocity resulting in n collisions with the piston followed by a collision with the back wall, an initially negative velocity resulting in n collisions with the piston followed by free streaming, and an initially negative velocity resulting in n collisions with the piston followed by a collision with the back wall. The two step functions in Eq. (2.6) account for the fact that the gas is confined to the $[0, L_i + Vt]$ interval during the expansion. This expression for $f(x,u;t)$ can easily be shown to obey Liouville's equation and satisfies both boundary conditions; cf. Eqs. (2.3) and (2.4) and the initial condition, Eq. (2.5). Note that, when $V \to 0$—i.e., the process is carried out reversibly—it is easy to show (by converting the sums into integrals) that Eq. (2.6) reduces to the appropriate equilibrium distribution function; it is uniform in x and Gaussian in u with a reduced temperature $T_i[L_i/(L_i+Vt)]^2$, and is in complete agreement with thermodynamics.

The initial equilibrium distribution function and the nonequilibrium distribution function ($t=4$) for a piston of initial length $L_i=10$, piston velocity $V=0.5$, and initial temperature $T_i=1$ are shown in Fig. 1. As expected, the distribution function is affected by the expansion only in the vicinity of the moving piston. We next show some examples of what we can extract from this distribution function.

While the nonequilibrium process is occurring, the state of the system can be described, at least in part, in terms of local thermodynamic quantities [24], perhaps generalized in various ways, if at all. For example, the local density, velocity, energy per particle, temperature, and entropy per particle are well known (see, e.g., Ref. [21]) and are defined by

$$\rho(x,t) \equiv \int_{-\infty}^{\infty} du f(x,u;t), \tag{2.7}$$

$$U(x,t) \equiv \int_{-\infty}^{\infty} du \frac{f(x,u;t)}{\rho(x,t)} u, \tag{2.8}$$

$$\frac{E(x,t)}{k_B T_i} \equiv \int_{-\infty}^{\infty} du \frac{f(x,u;t)u^2}{2\rho(x,t)}, \tag{2.9}$$

$$k_B T(x,t) \equiv \int_{-\infty}^{\infty} du \frac{f(x,u;t)[u - U(x,t)]^2}{\rho(x,t)}, \tag{2.10}$$

and

$$\frac{S(x,t)}{k_B} \equiv -\int_{-\infty}^{\infty} du \frac{f(x,u;t)\ln f(x,u;t)}{\rho(x,t)}, \tag{2.11}$$

respectively. Note that

$$\frac{E(x,t)}{k_B} = \frac{1}{2}[T(x,t) + U^2(x,t)]. \tag{2.12}$$

Clearly, all but the entropy per particle will be given in terms of sums of error functions upon the substitution of Eq. (2.6) into Eqs. (2.7)–(2.10).

We also can evaluate the boundary conditions obeyed by the thermodynamic variables at the piston ($x=L_i+Vt$). The simplest boundary condition is the one for the velocity field at the boundary, $U(L_i+Vt,t)=V$, as expected from mass conservation. We show examples of the local thermodynamic quantities profiles as a function of time and position in Fig. 2.

Before going on, we clarify a few points about the unit conventions that we use. We have set the mass of the particle and Boltzmann's constant to unity; therefore, all quantities that have the units of energy, the work (W) and the free energy (A), and temperature all have the same units. It also

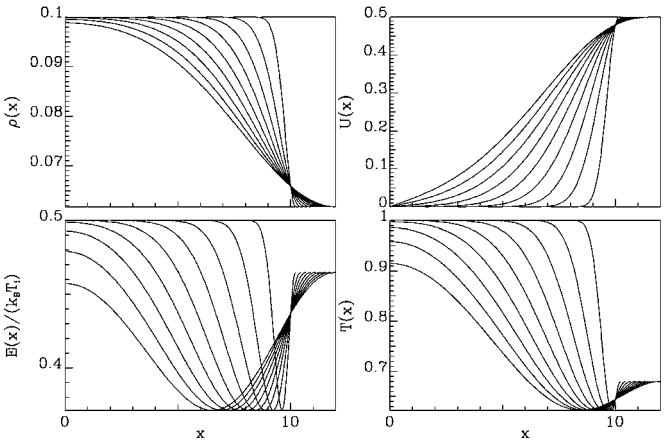

FIG. 2. The local thermodynamics quantities defined by Eqs. (2.7)–(2.10) are shown as a function of time from $t=0$ to $t=4$ for an expansion where $L_i=10$ and $V=0.5$. At $t=0$, all quantities are uniform for $0<x<L_i$ and zero elsewhere. The curves are equally spaced at time intervals of 4/9. Larger values of t have larger nonuniform regions.

follows that the particle's velocity u and the piston velocity V have units of $\sqrt{\text{energy}}$. We use this convention for what follows.

At this stage, we consider the case where the piston stops moving—e.g., when $t=4$. At this particular time, it is quite clear that the system is far from equilibrium (as shown from the inhomogeneous character of the quantities displayed in Fig. 2). We define a local Helmholtz free energy per particle as

$$A(x) = E(x) - T(x)S(x). \tag{2.13}$$

We expect that the system will eventually come to equilibrium, provided we wait long enough, and when this happens, the free energy, and all other thermodynamic quantities, will be uniform in position. In particular, we expect that there the final free energy per particle is

$$A_f = -T_f \ln[L_f(2\pi T_f)^{1/2}], \tag{2.14}$$

where T_f and L_f are the final temperature and length, respectively, and where Planck's constant h has been set to 1. The final temperature is easily obtained from the total final energy [cf. Eq. (2.12)] or equivalently from $E_f = \int_0^{L_f} E(x,4)\rho(x,4)$, which is conserved when the piston is at rest and which is related to the temperature at equilibrium in the usual manner (i.e., $E_f = T_f/2$).

Perhaps surprisingly at first glance, the free energy of the system, as computed from thermodynamics—i.e., Eq. (2.14)—is not the one obtained from $A = \int_0^{L_f} dx \rho(x,\infty)A(x,\infty)$, even at infinite time after the piston has stopped moving. This happens because the entropy, as defined above in Eq. (2.11), is the fine-grained entropy per particle. As is well known [25,26], the total fine-grained entropy is a constant of the motion, but the final entropy of the system should equal

$$S_f = \frac{1}{2} + \ln[L_f(2\pi T_f)^{1/2}], \tag{2.15}$$

which is obviously different from the initial entropy, unless the expansion is done reversibly and adiabatically.

This well-known paradoxical result is resolved by introducing a coarse-grained entropy [25,26]. There are many ways to do this, and in some of the works on the Jarzynski equality, this is done by averaging out so-called bath degrees of freedom [15]. The model here is too simple to allow for this sort of coarse graining, and instead we consider an older approach due to Kirkwood [27] and define a time-averaged distribution function as

$$\bar{f}(x,u;t) \equiv \frac{1}{\tau}\int_t^{t+\tau} ds f(x,u;s), \tag{2.16a}$$

and then use it to define a coarse grained entropy as

$$\frac{\bar{S}(t)}{k_B} \equiv -\int_0^{L_f}\int_{-\infty}^{\infty} dx du \bar{f}(x,u;t)\ln \bar{f}(x,u;t). \tag{2.16b}$$

Note that this time-averaging procedure only makes sense at long times after the end of the expansion, $t\to\infty$, where the fine-grained distribution is approximately uniform (or, more generally, slowly varying in time). Also, this time averaging of the distribution function preserves the important property that the energy is conserved after the expansion. Finally, it can be shown that the time-averaged distribution is uniform in position as $\tau\to\infty$.

In Fig. 3, we show the time-averaged distribution function for $t=1000$ and $\tau=100$ with 2000 discretization points in the average and we compare it with the fine-grained distribution at the same time. Note that the structure appearing in the fine-grained distribution is averaged out after coarse graining. More quantitatively, the coarse-grained entropy equals $\bar{S}=3.811$. This should be compared with $S_f=3.812$ and $S_i=3.722$ as obtained from Eqs. (2.15) and (2.11), respectively.

We conclude these comments on fine- and coarse-grained entropy by pointing out that, within this model, the system never quite comes back to equilibrium, even a long time after the expansion (this is well known and has been pointed out in Ref. [28]). If it did, the final fine-grained distribution would be completely uniform in position and defect free. This does not happen here because this one-dimensional model does not contain any mechanism that will randomize the velocities efficiently and does not have a strong separation of time scales.

Returning to the Jarzynski equality, the final free energy appearing in Eq. (1.1), A_J, can be obtained from Eq. (2.14) with $T_i=T_f$ and the final equilibrium free energy can be obtained from Eq. (2.13), provided we use a coarse-grained entropy, or from Eq. (2.14) with the appropriate final temperature (for this model $T_f=2E_f$). For the expansion parameters defined above, $\Delta A=0.466$ as obtained from Eqs. (2.14) or (2.13) and $\Delta A_J=-0.1823$ (see Fig. 5 of the next section for a more detailed comparison between ΔA and ΔA_J when the piston speed is varied). These differences are explained

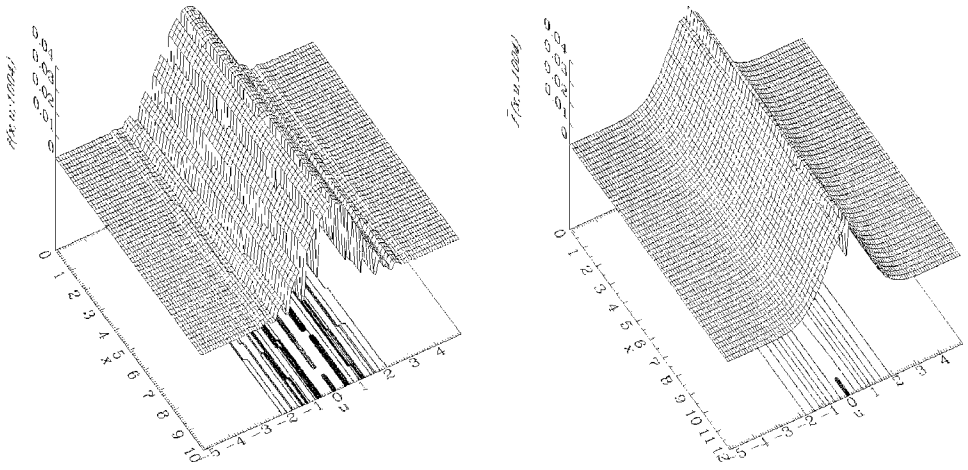

FIG. 3. The fine-grained distribution function (on the left) is compared with the time-averaged distribution function that is later used to calculate the coarse-grained entropy according to Eq. (2.16b). Here, $t=1000$ and $\tau=100$. The contours are drawn when $f(x,u)=0.005$, 0.01, 0.015, 0.02, 0.025, 0.03, and 0.035.

by the fact that the expansion of the gas is not isothermal. In fact, for the parameters described above, the temperature drops from 1.0 to 0.832.

Thus, while the Jarzynski equality is rigorously correct, the above discussion shows that the equality does not provide much information about the nonequilibrium state of the system, neither locally nor thermodynamically. Most significantly, ΔA_J is not the free energy change predicted by thermodynamics. Even if local thermodynamic equilibrium can only be invoked approximately during the expansion process [cf. Eq. (2.13)], the nonequilibrium distribution function is a complete description of the nonequilibrium state and it carries a lot of information that is hidden in the Jarzynski equality. Still, the Jarzynski equality holds. Recall here that it relates the nonequilibrium average of $e^{\beta W}$ to the equilibrium free energy difference between two states at the same temperature, but with different lengths.

III. WORK BOUNDS

As pointed out by Jarzynski in his original papers [1,2], the Gibbs-Bogoliubov-Jensen-Peierls inequality

$$\langle e^{\beta W}\rangle \geqslant e^{\beta\langle W\rangle} \tag{3.1}$$

combined with the Jarzynski equality automatically implies that the average work is bounded by

$$\langle W\rangle \leqslant -\Delta A_J, \tag{3.2}$$

strongly reminiscent of the usual bound for work in isothermal process found in thermodynamics. Again, recall that we are using the convention that $\langle W\rangle$ is the work done by the system on the surroundings. We now compare this bound against the bound obtained from the laws of thermodynamics. This is an important question, since the degree to which the two bounds differ disproves the contention that Eq. (3.2) is essentially a proof of the second law of thermodynamics from mechanics (this is claimed in Refs. [11,28] and in the review section of Ref. [29]).

On the other hand, the laws of thermodynamics also provide bounds for the average work; specifically, the first and second laws imply that

$$\langle W\rangle \leqslant -\Delta A - \int (dT\,S - dN\mu_{op}), \tag{3.3}$$

where μ_{op} is the opposing chemical potential and N is the number of particles. This becomes the usual work bound in terms of the Helmholtz free energy change for isothermal and constant-N processes. In what follows, we will consider processes which conserve the number of particles, as this is appropriate for most of the numerical studies of the Jarzynski equality, and drop the dN term in Eq. (3.3). In general the $\int(dT\,S - dN\mu_{op})$ term is not a state function and greatly reduces the utility of Eq. (3.3). Nonetheless, one can obtain a useful bound for the work by noting that in an adiabatic expansion,

$$\langle W\rangle = -\Delta E \tag{3.4a}$$

$$= -\Delta A - S_f\Delta T - T_i\Delta S = -\Delta A - S_i\Delta T - T_f\Delta S. \tag{3.4b}$$

Since $\Delta S \geqslant 0$ for a spontaneous adiabatic process, Eqs. (3.4a) and (3.4b) imply that

$$\langle W\rangle \leqslant -\max_{\alpha=i,f}(\Delta A + S_\alpha\Delta T), \tag{3.5}$$

where the inequality becomes an equality for reversible processes. Although in principle ΔT can have any sign in an adiabatic expansion, it must be negative for ideal gases, and thus, Eq. (3.5) implies that

$$\langle W\rangle \leqslant -\Delta A - S_i\Delta T \equiv W_{rev}. \tag{3.6}$$

It is straightforward to relate the Jarzynski bound to that given in Eq. (3.5); specifically, it is easy to show that

$$\Delta A + S_f\Delta T = \Delta A_J + \int_{T_i}^{T_f} dT\frac{C_V(T,V_f)}{T}(T - T_i), \tag{3.7}$$

where $C_V(T,V_f)$ is the constant-volume heat capacity at the final volume, V_f. The last integral is strictly positive for $\Delta T \neq 0$, and this in turn means that the Jarzynski bound is greater than the one implied by thermodynamics; cf. Eq. (3.5).

This general result reproduces the expected work bounds for the isolated harmonic oscillator model with an increasing

natural frequency (ω_i that increases to ω_f) discussed by Jarzynski (cf. Fig. 2 of Ref. [2]) and for the above ideal gas expansion. In both cases, when the work is done reversibly and adiabatically $S_f=S_i$. This guarantees $-(\Delta A+S_f\Delta T)$ to be the true upper bound to the work from Eq. (3.5). For the harmonic oscillator model, $\Delta A_J=T_i\ln(\omega_f/\omega_i)$, while for the ideal gas model, $\Delta A_J=-T_i\ln(L_f/L_i)$. In the two cases C_v is independent of temperature and the correction to ΔA_J in Eq. (3.7) becomes

$$\int_{T_i}^{T_f} dT\frac{C_V(T,V_f)}{T}(T-T_i)=C_VT_i\left[\frac{T_f}{T_i}-1-\ln\left(\frac{T_f}{T_i}\right)\right]. \tag{3.8}$$

In both models, when the work is done reversibly and adiabatically it is easy to show that

$$\frac{T_f}{T_i}=\begin{cases}\dfrac{\omega_f}{\omega_i} & \text{harmonic oscillator},\\ \left(\dfrac{L_i}{L_f}\right)^2 & \text{ideal gas}.\end{cases} \tag{3.9}$$

When this is used in Eq. (3.8), it is easy to see that these terms are, as expected, positive. Moreover, when they are used in Eq. (3.7), the ΔA_J term exactly cancels and one is left with the usual thermodynamic work bound

$$W_{rev}=T_i\begin{cases}1-\dfrac{\omega_f}{\omega_i} & \text{harmonic oscillator},\\ 1-\dfrac{L_i^2}{L_f^2} & \text{ideal gas};\end{cases} \tag{3.10}$$

hence, in both cases,

$$W_{rev}\leqslant-\Delta A_J, \tag{3.11}$$

where the equality only holds in the trivial case where no work is done on the system (i.e., $\omega_f=\omega_i$ or $L_f=L_i$). These two special cases were used to illustrate the more general result given in Eq. (3.7) and to highlight the fact that Eq. (3.7) is in quantitative agreement with the harmonic model described by Jarzynski in Ref. [2]. In summary, the Jarzynski equality alone does not guarantee Eq. (3.3) to be satisfied *and, hence, is not a proof of the second law*.

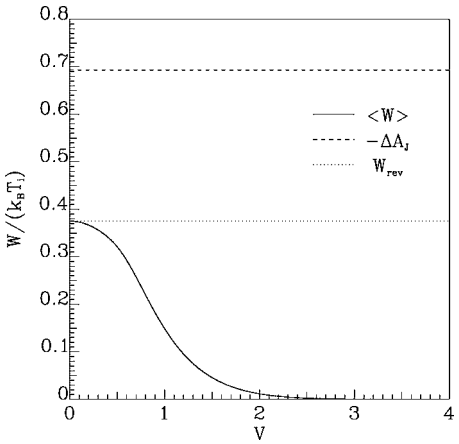

FIG. 4. The average work W is obtained from Eq. (3.13) with $T=1$, $L_i=10$, and $Vt=10$. This is compared with the Jarzynski bound, which is independent of the piston velocity.

The 1D gas model can further be used to calculate the average work done for any expansion rate or piston velocity V. Again, the average work, like the nonequilibrium distribution function [Eq. (2.6)], can be expressed as a sum over the number of collisions with the piston,

$$\langle W\rangle=\frac{1}{L_i}\left(\frac{2\beta}{\pi}\right)^{1/2}\int_0^{L_i}dx_0\sum_{n=1}^{\infty}\left(\int_{(2n-1)(L_i/t+V)-x/t}^{(2n+1)(L_i/t+V)-x/t}du_0e^{-\beta u_0^2/2}\times(nu_0V-n^2V^2)+\int_{(2n-1)(L_i/t+V)+x/t}^{(2n+1)(L_i/t+V)+x/t}du_0e^{-\beta u_0^2/2}\times(nu_0V-n^2V^2)\right), \tag{3.12}$$

where x_0 and u_0 are the initial position and velocities of the gas particle. This is equivalent to the expression for $\langle W\rangle$ obtained from the work distribution $P(W)$, as defined by Eq. (13) in Lua and Grosberg [10] [$\langle W\rangle=\int_0^\infty dWWP(W)$]. Simple but lengthy manipulations transform this expression into

$$\begin{aligned}\langle W\rangle=&\frac{Vt}{\beta L_i}\sum_{n=1}^{\infty}\left\{\operatorname{erfc}\left(\frac{\alpha}{t}[(2n+1)(L_i+Vt)-L_i]\right)-\operatorname{erfc}\left(\frac{\alpha}{t}[(2n+1)(L_i+Vt)+L_i]\right)\right\}+\frac{Vt}{\beta L_i}\left\{\operatorname{erfc}(\alpha V)-\operatorname{erfc}\left(\frac{\alpha}{t}(2L_i+Vt)\right)\right\}\\&+\frac{V^2t}{L_i}\left(\frac{2}{\beta}\right)^{1/2}\sum_{n=1}^{\infty}(2n+1)\left\{\frac{\alpha}{t}[(2n+1)(L_i+Vt)-L_i]\operatorname{erfc}\left(\frac{\alpha}{t}[(2n+1)(L_i+Vt)-L_i]\right)-\frac{\alpha}{t}[(2n+1)(L_i+Vt)\right.\\&\left.+L_i]\operatorname{erfc}\left(\frac{\alpha}{t}[(2n+1)(L_i+Vt)+L_i]\right)-\pi^{-1/2}e^{-(\alpha^2/t^2)[(2n+1)(L_i+Vt)-L_i]^2}+\pi^{-1/2}e^{-(\alpha^2/t^2)[(2n+1)(L_i+Vt)+L_i]^2}\right\}\\&+\frac{V^2t}{L_i}\left(\frac{2}{\beta}\right)^{1/2}\left\{\alpha V\operatorname{erfc}(\alpha V)-\frac{\alpha}{t}(2L_i+Vt)\operatorname{erfc}\left(\frac{\alpha}{t}(2L_i+Vt)\right)-\pi^{-1/2}e^{-\alpha^2V^2}+\pi^{-1/2}e^{-(\alpha^2/t^2)(2L_i+Vt)^2}\right\},\end{aligned} \tag{3.13}$$

where $\alpha \equiv (\beta/2)^{1/2}$ and where $\mathrm{erfc}(x)$ is the complementary error function. When the expansion is done reversibly, we have $\alpha V \ll 1$, $t \gg 1$ and $\alpha L_i/t \ll 1$. In this limit, the sums in Eq. (3.13) can be replaced by integrals with the result that

$$\langle W \rangle \sim \frac{Vt(Vt+2L_i)}{2\beta(L_i+Vt)^2} + O(V), \tag{3.14}$$

which, after some simple manipulations, agrees with Eq. (3.10). For the rest of this section, we will consider the case where $L_i=10$ and $Vt=10$ (the length of the box doubles).

For finite piston velocities, the average work is calculated using Eq. (3.13), keeping enough terms such that the error falls within a small tolerance. In Fig. 4, the average work is shown as a function of the velocity of the piston for $T_i=1$, $L_i=10$, and $Vt=10$. As $V \to 0$, it is seen that $\langle W \rangle$ tends to W_{rev}, which, for these parameters, equals 3/8. This figure clearly shows that W_{rev} (the straight dotted line in Fig. 4) is the smallest upper bound to the process while the Jarzynski bound $-\Delta A_J$ sits above (straight dashed line in Fig. 4). Note that, for large V (the tail region in Fig. 4), the Jarzynski equality still holds even if $\langle W \rangle$ is very small. This seemingly paradoxical results was investigated in Ref. [10]. Also note that there is no contradiction between the results shown in Fig. 4 and the results of Lua and Grosberg [10]. In fact, Fig. 6 of Ref. [10] shows that the average work becomes equal to $-\Delta A_J$ for small piston velocities. The problem there is that they kept t and L_i constant and reduced V. There, as V tends to zero, $L_f - L_i$ vanishes and there is no expansion; *this is not what is meant by the quasistatic limit*. As stated after Eq. (3.11), the two work bounds trivially agree in that limit.

ΔA is compared with ΔA_J as a function of the piston velocity in Fig. 5. Recall that the final temperature of the system is determined from the work. The data show that ΔA does not have a definite sign. Further, for very small V (i.e., a nearly reversible process), ΔA exhibits the largest differences from ΔA_J. On the other hand, for large V, the two free energy differences agree with each other. This result is easily explained in terms of temperature changes. When the process is slow, maximum work is done, and hence, the temperature changes the most. On the other hand, when the piston is pulled very quickly, only a small fraction of the particles can collide with it, very little work is done, and the temperature does not change. In this case, the free energy that appears in the Jarzynski equality describes the final state appropriately.

In a recent article by Baule, Evans, and Olmsted [28], it was shown that the ideal gas expansion, within an isothermal model where the gas is effectively coupled to a thermostat, does satisfy the Jarzynski equality and that $-\Delta A_J = W_{rev}$ in this case (note that this model does not fall into any category of model for which the Jarzynski equality has been derived rigorously).

We conclude this section with a short remark on the experimental consequences of our observations. In many experiments, in particular in the famous experiment by Liphardt *et al.* [12], it is often assumed that $-\Delta A_J$ is equal to W_{rev}. Then, many realizations of the work are performed irreversibly and $\langle e^{\beta W} \rangle$ is computed and compared to $e^{-\beta \Delta A_J}$. We have shown that, in general, this is not exactly true. In fact, when the temperature change is significant, W_{rev} and $-\Delta A_J$ can be very different. In the case of the experiment of Liphardt *et al.*, the work was done on a single RNA molecule in solution, in which case, naively, the temperature change should be small. This will be further investigated below.

IV. JARZYNSKI RELATION AND RESPONSE THEORY

In the previous sections we showed that, even though the Jarzynski equality is exactly satisfied for the 1D expanding gas, ΔA_J, in general, does not characterize the actual state of the system at any time during or after the expansion process. Further, the bound that we get from the Jarzynski equality tells us less than what we already know from thermodynamics. Therefore, one could wonder why the Jarzynski equality works. In this section, we show how, in the context of response theory, the terms that give rise to nonequilibrium effects cancel when $\langle e^{\beta W} \rangle$ is computed and examine the average work done.

Consider a classical system that evolves under the Hamiltonian

$$H(t) = H_0 + H_1(t), \tag{4.1}$$

where H_0 and $H_1(t)$ are, respectively, explicitly time independent and time dependent. In response theory, $H_1(t)$ is treated as a perturbation that has the general form

$$H_1(t) = -\sum_j \int d\mathbf{r} A_j(\mathbf{r}, X^N) F_j(\mathbf{r}, t) \equiv -\mathsf{A}(t) * \mathsf{F}(t), \tag{4.2}$$

where X^N is the phase point, the $A_j(\mathbf{r}, X^N)$'s are local observables at position $\mathbf{r}$ that depend implicitly on time through the motion of the particles, and the $F_j(\mathbf{r}, t)$'s are the external fields. We assume that the $F_j(\mathbf{r}, t)$'s vanish for $t \leqslant 0$. By using this form for the Hamiltonian, the right-hand side of the Jarzynski equality becomes

$$e^{-\beta \Delta A_J} = \langle e^{\beta \mathsf{A} * \mathsf{F}(t)} \rangle_0, \tag{4.3}$$

where

$$\langle (\ldots) \rangle_0 = \frac{\int dX^N e^{-\beta H_0} (\cdots)}{\int dX^N e^{-\beta H_0}} \tag{4.4}$$

is a canonical average with respect to H_0. This can be expanded in powers of the external fields as follows:

$$-\beta \Delta A_J = \sum_{n=1}^{\infty} \frac{\beta^n}{n!} \langle\langle \mathsf{A}^n \rangle\rangle_0 (*)^n \mathsf{F}(t)^n, \tag{4.5}$$

where $\langle\langle (\cdots) \rangle\rangle_0$ are cumulant averages [30]. We now show how, to second order in the external fields, the Jarzynski equality is satisfied.

The first step is to define the work done by the system in terms of the external fields,

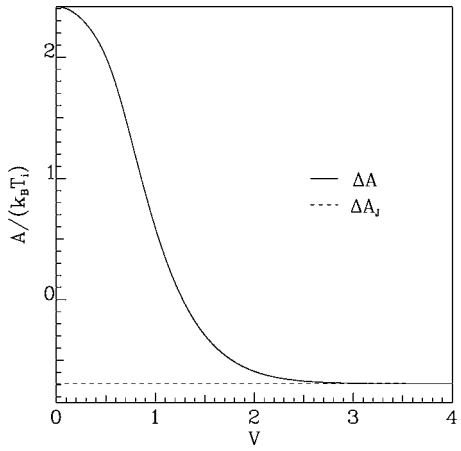

FIG. 5. The real free energy difference ΔA between the two equilibrium states is compared with ΔA_J. The conditions of the expansion are the same as in Fig. 4.

$$W(t) = \int_0^t ds \mathsf{A}(s) * \frac{\partial \mathsf{F}(s)}{\partial s}. \tag{4.6}$$

Next, we rewrite $\langle e^{\beta W(t)} \rangle$ in terms of cumulants—i.e.,

$$\langle e^{\beta W(t)} \rangle = \exp\left(\sum_{n=1}^{\infty} \frac{1}{n!} \beta^n \langle\langle W(t)^n \rangle\rangle \right), \tag{4.7}$$

where the averages are nonequilibrium averages. Response theory provides a formalism that can be used to calculate these nonequilibrium averages to any desired power in the perturbing Hamiltonian [30,31]. In particular, in the linear regime,

$$\langle\langle \mathsf{B}(t) \rangle\rangle = \langle\langle \mathsf{B} \rangle\rangle_0 - \beta \int_0^t ds \langle\langle \dot{\mathsf{B}}(t-s)\mathsf{A} \rangle\rangle_0 * \mathsf{F}(s) + O(\mathsf{F}^2), \tag{4.8}$$

where B is some observable and $\dot{\mathsf{B}}$ is the time derivative of B using the reference Hamiltonian H_0.

We will compare Eqs. (4.5) and (4.7) to second order in the external fields, and since $W(t)$ is already linear in F, the nonlinear response terms can be neglected. Hence, using Eq. (4.8), $\langle\langle W(t) \rangle\rangle$ and $\langle\langle W(t)^2 \rangle\rangle$ are evaluated up to second order in the fields, giving

$$\begin{aligned} \beta \langle\langle W(t) \rangle\rangle &= \beta \int_0^t ds \langle\langle \mathsf{A}(s) \rangle\rangle * \frac{\partial \mathsf{F}(s)}{\partial s} \\ &= \beta \langle\langle \mathsf{A} \rangle\rangle_0 * \mathsf{F}(t) - \beta^2 \int_0^t ds \int_0^s ds' \langle\langle \dot{\mathsf{A}}(s-s')\mathsf{A} \rangle\rangle_0 \\ &\quad \times (*)^2 \mathsf{F}(s') \frac{\partial \mathsf{F}(s)}{\partial s} + O(\mathsf{F}^3) \end{aligned} \tag{4.9}$$

and

$$\begin{aligned} \frac{\beta^2}{2} \langle\langle W(t)^2 \rangle\rangle &= \frac{\beta^2}{2} \int_0^t ds \int_0^t ds' \langle\langle \mathsf{A}(s)\mathsf{A}(s') \rangle\rangle_0 (*)^2 \\ &\quad \times \frac{\partial \mathsf{F}(s)}{\partial s} \frac{\partial \mathsf{F}(s')}{\partial s'} + O(\mathsf{F}^3), \end{aligned} \tag{4.10}$$

where we have used that fact that $\mathsf{F}(t)=0$ for $t \leqslant 0$ and that the equilibrium canonical distribution function is stationary. Equation (4.9) is, of course, an example of the usual fluctuation-dissipation theorem result in a classical system (see, e.g., Ref. [24] or [32]). In Eq. (4.9), the second term of the second line can be integrated by parts using the fact that

$$\langle\langle \dot{\mathsf{A}}(s-s')\mathsf{A} \rangle\rangle_0 = -\frac{\partial}{\partial s'} \langle\langle \mathsf{A}(s-s')\mathsf{A} \rangle\rangle_0, \tag{4.11}$$

and we find that

$$\begin{aligned} \langle\langle W(t) \rangle\rangle &= \langle\langle \mathsf{A} \rangle\rangle_0 * \mathsf{F}(t) + \beta \int_0^t ds \langle\langle \mathsf{AA} \rangle\rangle_0 (*)^2 \mathsf{F}(s) \frac{\partial \mathsf{F}(s)}{\partial s} \\ &\quad - \beta \int_0^t ds \int_0^s ds' \langle\langle \mathsf{A}(s)\mathsf{A}(s') \rangle\rangle_0 (*)^2 \frac{\partial \mathsf{F}(s)}{\partial s} \frac{\partial \mathsf{F}(s')}{\partial s'} \\ &= \langle\langle \mathsf{A} \rangle\rangle_0 * \mathsf{F}(t) + \frac{\beta}{2} \langle\langle \mathsf{AA} \rangle\rangle_0 (*)^2 \mathsf{F}(t)^2 \\ &\quad - \frac{\beta}{2} \int_0^t ds \int_0^t ds' \langle\langle \mathsf{A}(s)\mathsf{A}(s') \rangle\rangle_0 (*)^2 \frac{\partial \mathsf{F}(s)}{\partial s} \frac{\partial \mathsf{F}(s')}{\partial s'}. \end{aligned} \tag{4.12}$$

The last term in this expression is strictly positive and is responsible for any dissipation, and will be denoted as W_d below. The first two terms are just the ones expected using perturbation theory on a quasistatic Hamiltonian and are equal to ΔA_J to $O(\mathsf{F}^2)$, cf. Eq. (4.5).

With these observations, we see that the second cumulant of the work [cf. Eq. (4.10)] is, up to factors of β, just the dissipative part of the work, and thus, finally,

$$\begin{aligned} &\beta \langle\langle W(t) \rangle\rangle + \frac{\beta^2}{2} \langle\langle W(t)^2 \rangle\rangle \\ &\quad = \beta \langle\langle \mathsf{A} \rangle\rangle_0 * \mathsf{F}(t) + \frac{\beta^2}{2} \langle\langle \mathsf{AA} \rangle\rangle_0 (*)^2 \mathsf{F}(t)^2 + O(\mathsf{F}^3), \end{aligned} \tag{4.13}$$

which is in agreement with the Jarzynski relation, again, to second order in the external fields.

Nonlinear response theory could, in principle, be used to prove the Jarzynski equality to all orders in the external fields. Aside from the fact that this becomes messy very quickly, there is no need for such a proof. After all, the equality has already been proved by Jarzynski quite generally in Ref. [3]. The above approach is interesting because it clarifies how the equality works by showing how apparently dissipative terms cancel. In fact, even though $\langle\langle W(t) \rangle\rangle$ itself is a nonequilibrium average that depends very much on how the work is performed, the dissipative terms [the ones containing $\partial \mathsf{F}(s)/\partial s$] in Eq. (4.7) are exactly canceled by the

next term in the work cumulant expansion, Eq. (4.10). Also note that this cancellation only works if the β in Eq. (4.7) equals the one that appears in the initial canonical distribution function. We have checked that, at least to third order in the fields, *all* the (F^3) Jarzynski terms are contained in $\langle W(t)\rangle$; the rest of the terms in $\langle W(t)\rangle$ are dissipative in the sense that they explicitly depend on the rates of change of the F's [e.g., as in the last term on the right-hand side of Eq. (4.12)].

The response theory result seems to contradict what we found thermodynamically or for the ideal gas expansion considered in the preceding section: namely, there, even in the quasistatic or reversible limit, the Jarzynski bound was an upper bound to that predicted by thermodynamics. Here the response theory suggests that the Jarzynski bound is satisfied exactly in the quasistatic limit—i.e., where $\partial \mathsf{F}(s)/\partial s \to 0$.

There is no contradiction for several reasons. The first is trivial. The perturbing potential should be considered as a moving finite step potential with height V_0 that is set to infinity at the end of the calculation (for a discussion of the orders of limits, see Ref. [9]). Response theory assumes that the perturbing potential is small, something that is not the case in Sec. III.

The second reason why there is no contradiction is more subtle and requires us to be more careful defining what is meant by a quasistatic process. A quasistatic process is one that takes place at a rate which is much slower than *all* other dynamical processes taking place in the system. Many-body systems have collective modes corresponding to various mechanically conserved quantities, and these will evolve on arbitrarily long time scales (governed by the wavelength of the mode). As we now show, there are no real quasistatic processes in the sense of the Jarzynski equality unless some addition assumptions are made about the nature of the perturbation applied to the system.

We start by introducing a projection operator [33,34] $\mathcal{P}$ defined as follows:

$$\mathcal{P}\mathsf{A} \equiv \langle\langle \mathsf{AC}\rangle\rangle * \langle\langle \mathsf{CC}\rangle\rangle^{-1} * \mathsf{C}, \tag{4.14}$$

where C is a column vector containing the densities of slowly evolving variables. At the simplest level C must contain the densities of conserved quantities (e.g., number, energy, and momentum densities in a one-component system) but could also contain broken-symmetry variables [32], as well as multilinear products of these fields should mode-coupling effects be important. $\langle\langle(\cdots)\rangle\rangle$ is a cumulant average taken in the reference system, and for the rest of this section, we omit the zero subscript.

Well-known projection operator identities can be used to express $\langle\langle \mathsf{A}(t)\mathsf{A}\rangle\rangle$ in terms of correlations of slow and fast quantities (here, we simply state the result since these techniques are standard). After Laplace transforming in time, we obtain

$$\begin{aligned}\langle\langle \tilde{\mathsf{A}}(s)\mathsf{A}\rangle\rangle = {} & \langle\langle \tilde{\mathsf{A}}^{\ddagger}(s)\mathsf{A}^{\ddagger}\rangle\rangle + (\langle\langle \mathsf{AC}\rangle\rangle - \langle\langle \tilde{\mathsf{A}}^{\ddagger}(s)\dot{\mathsf{C}}^{\ddagger}\rangle\rangle) * \langle\langle \mathsf{CC}\rangle\rangle^{-1} \\ & * \langle\langle \tilde{\mathsf{C}}(s)\mathsf{C}\rangle\rangle * \langle\langle \mathsf{CC}\rangle\rangle^{-1} * (\langle\langle \mathsf{CA}\rangle\rangle \\ & + \langle\langle \tilde{\dot{\mathsf{C}}}^{\ddagger}(s)\mathsf{A}^{\ddagger}\rangle\rangle),\end{aligned} \tag{4.15}$$

where the tilde denotes a Laplace transform in time, s is the Laplace transform frequency, and $\tilde{\mathsf{A}}^{\ddagger}(s)$ is the Laplace transform of

$$\mathsf{A}^{\ddagger}(t) \equiv e^{i(1-\mathcal{P})Lt}(1-\mathcal{P})\mathsf{A}. \tag{4.16}$$

This object is referred to as the dissipative part of A and is orthogonal to the space of slow variables for all times; hence, their correlations should decay on microscopic time (and length) scales. In terms of the Laplace transforms, this means that, for long-time phenomena—i.e., small s—only the frequency dependence of $\langle\langle \tilde{\mathsf{C}}(s)\mathsf{C}\rangle\rangle$ need be considered and will give all the long-time dependence. All the other correlations can be evaluated at $s=0$; indeed, for the standard hydrodynamic variables, the $\langle\langle \mathsf{C}(s)\mathsf{C}\rangle\rangle$ have been studied extensively [32,35].

We now can analyze the so-called dissipative term in Eq. (4.12). Remember that in this term, there are two implicit spatial integrations and for translationally invariant equilibrium systems it is convenient to switch to a Fourier representation in space, thereby obtaining

$$W_d \equiv -\beta \int_0^t dt_1 \int_0^{t_1} dt_2 \int \frac{d\mathbf{k}}{(2\pi L)^d} \langle\langle \mathsf{A}_{\mathbf{k}}(t_1-t_2)\mathsf{A}_{-\mathbf{k}}\rangle\rangle : \dot{\mathsf{F}}_{-\mathbf{k}}(t_1)\dot{\mathsf{F}}_{\mathbf{k}}(t_2) \tag{4.17a}$$

$$\begin{aligned}\sim -\beta \int_0^t dt_1 \int_0^{t_1} dt_2 \int \frac{d\mathbf{k}}{(2\pi L)^d} & [\langle\langle \mathsf{A}^{\ddagger}_{\mathbf{k}}(t_1-t_2)\mathsf{A}^{\ddagger}_{-\mathbf{k}}\rangle\rangle \\ & + (\langle\langle \mathsf{A}_{\mathbf{k}}\mathsf{C}_{-\mathbf{k}}\rangle\rangle - \langle\langle \tilde{\mathsf{A}}^{\ddagger}_{\mathbf{k}}(0)\dot{\mathsf{C}}^{\ddagger}_{-\mathbf{k}}\rangle\rangle) \\ & \times \langle\langle \mathsf{C}_{\mathbf{k}}\mathsf{C}_{-\mathbf{k}}\rangle\rangle^{-1}\langle\langle \mathsf{C}_{\mathbf{k}}(t_1-t_2)\mathsf{C}_{-\mathbf{k}}\rangle\rangle\langle\langle \mathsf{C}_{\mathbf{k}}\mathsf{C}_{-\mathbf{k}}\rangle\rangle^{-1} \\ & \times (\langle\langle \mathsf{C}_{\mathbf{k}}\mathsf{A}_{-\mathbf{k}}\rangle\rangle + \langle\langle \tilde{\dot{\mathsf{C}}}^{\ddagger}_{\mathbf{k}}(0)\mathsf{A}^{\ddagger}_{-\mathbf{k}}\rangle\rangle)]:\dot{\mathsf{F}}_{-\mathbf{k}}(t_1)\dot{\mathsf{F}}_{\mathbf{k}}(t_2),\end{aligned} \tag{4.17b}$$

where d is the dimension of space and L^d is the volume of the system. The second relation follows by using Eq. (4.15) and ignoring the frequency dependence of the dissipative correlations as discussed above. Within the context of the projection operator approach, this is valid provided that the F's do not evolve on fast (i.e., microscopic) time scales.

We now examine Eq. (4.17b) for several cases. The simplest is when the perturbing fields are spatially uniform—i.e., $\mathsf{F}_{\mathbf{k}}(t) = L^d \Delta(\mathbf{k})\mathsf{F}(t)$, where $\Delta(\mathbf{k})$ is a Kronecker δ [strictly speaking, we have to turn the Fourier integrals in Eq. (4.17b) into Fourier sums by letting $\int d\mathbf{k} \to (2\pi/L)^d \Sigma_{\mathbf{k}}$ to handle this case]. Since $\dot{\mathsf{C}}_{\mathbf{k}}(t) = i\mathbf{k}\cdot\mathbf{J}_{\mathbf{k}}$, where the $\mathbf{J}_{\mathbf{k}}$'s are fluxes, it is easy to show that

$$\begin{aligned}W_d \sim -\beta\langle\langle \tilde{\mathsf{A}}^{\ddagger}_{\mathrm{T}}(0)\mathsf{A}^{\ddagger}_{\mathrm{T}}\rangle\rangle : \int_0^t dt_1 \dot{\mathsf{F}}(t_1)\dot{\mathsf{F}}(t_1) - \frac{\beta}{2}\langle\langle \mathsf{A}_{\mathrm{T}}\mathsf{C}_{\mathrm{T}}\rangle\rangle \\ \times \langle\langle \mathsf{C}_{\mathrm{T}}\mathsf{C}_{\mathrm{T}}\rangle\rangle^{-1}\langle\langle \mathsf{C}_{\mathrm{T}}\mathsf{A}_{\mathrm{T}}\rangle\rangle : \mathsf{F}(t)\mathsf{F}(t),\end{aligned} \tag{4.18}$$

where the subscript "**T**" denotes the total or space integral (i.e., $\mathbf{k}\to 0$ limit) of subscripted quantity and where we have assumed that the decay of the $\langle\langle \mathsf{A}^{\ddagger}_{\mathrm{T}}(t)\mathsf{A}^{\ddagger}_{\mathrm{T}}\rangle\rangle$ correlation func-

tion is on a much faster time scale than any characterizing the $\dot{\mathsf{F}}$'s.

The first term on the right-hand side of Eq. (4.18) is well behaved provided that $\dot{\mathsf{F}}$ is square integrable and in particular should vanish in the quasistatic limit [$\dot{\mathsf{F}}(t)\to 0$, $t\to\infty$, keeping $\dot{\mathsf{F}}(t)t$ constant]; the remaining terms clearly do not vanish and are comparable to the $O(\mathsf{F}^2)$ terms in the Jarzynski free energy difference [cf. Eq. (4.13)]. Moreover, these terms are negative semi-definite, and thus, the response theory also shows that $-\Delta A_J$ is an upper bound to the actual work done, *even when the process is quasistatic*. Indeed, when Eq. (4.18) is used in Eq. (4.12), the latter becomes

$$\langle\langle W(t)\rangle\rangle = \langle\langle \mathsf{A}_{\mathrm{T}}\rangle\rangle_0 \cdot \mathsf{F}(t) + \frac{\beta}{2}\langle\langle \mathsf{A}_{\mathrm{T}}^{\ddagger}\mathsf{A}_{\mathrm{T}}^{\ddagger}\rangle\rangle_0 : \mathsf{F}(t)^2 - \beta\langle\langle \tilde{\mathsf{A}}_{\mathrm{T}}^{\ddagger}(0)\mathsf{A}_{\mathrm{T}}^{\ddagger}\rangle\rangle : \int_0^t dt_1 \dot{\mathsf{F}}(t_1)\dot{\mathsf{F}}(t_1), \quad (4.19)$$

which shows that it is only the parts of A that are orthogonal to the conserved quantities that contribute to the quasistatic work at $O(\mathsf{F}^2)$.

Within the response approach, in order that $-\Delta A_J$ equals the average work done quasistatically, the last terms in Eq. (4.18) must vanish. Since $\langle\langle \mathsf{C}_{\mathrm{T}}\mathsf{C}_{\mathrm{T}}\rangle\rangle$ is positive definite, the only way to eliminate these terms for nonzero $F(t)$ is to have

$$\langle\langle \mathsf{C}_{\mathrm{T}}\mathsf{A}_{\mathrm{T}}\rangle\rangle = \left.\frac{\partial\langle \mathsf{C}_{\mathrm{T}}\rangle_{\mathsf{F}}}{\partial\beta\mathsf{F}}\right|_{\mathsf{F}=0} = \frac{\partial\langle \mathsf{A}_{\mathrm{T}}\rangle}{\partial\beta\Phi} = 0, \quad (4.20)$$

where $\langle\cdots\rangle_{\mathsf{F}}$ is a canonical average using the Hamiltonian $H_0 - \mathsf{A}_T\cdot\mathsf{F}$ and $\beta\Phi$ is a column vector containing the usual conjugate variables to C found in the theory of fluctuations (e.g., $-\beta$ for energy, $\beta\mu$ for number, etc.). In other words, the quasistatic perturbation does not couple to the conserved variables and, hence, in the linear response regime, does not change the latter's averages. For us, this implies that the process must be strictly isochoric and isothermal. It is interesting to note the strong similarity of this criterion to one made by one of us some time ago [36] concerning the independence of the choice ensemble in response treatments of relaxation experiments.

While the preceding analysis at $\mathbf{k}=\mathbf{0}$ is appropriate when discussing the thermodynamics of large systems, it is easy to imagine other types of experiments. One such setup is where the perturbing fields are periodic in space. For a monochromatic (in space) perturbation the relevant time scales to compare with are those contained in $\langle\langle \mathsf{C}_{\mathbf{k}}(t)\mathsf{C}_{-\mathbf{k}}\rangle\rangle$, and while these can be very long, they are finite for nonzero wave vectors. In this case, W_d would vanish in the quasistatic limit *to leading order in the response theory*. We have qualified this last statement because higher-order terms in the response theory for W_d will allow for mode couplings to zero wave vector and may result in problems like those just shown. This will be investigated in a later work.

Finally, we consider one last example, one perhaps more appropriate to some of the recent optical tweezers experiments: namely, one where the probe-field is localized—i.e., $\mathsf{F}(\mathbf{r},t)=\delta(\mathbf{r})\mathsf{F}(t)$. We consider the small-wave-number contributions to W_d using hydrodynamic approximations for the $\langle\langle \mathsf{C}_{\mathbf{k}}(t)\mathsf{C}_{-\mathbf{k}}\rangle\rangle$ correlations. The most important terms that result from this analysis are those that decay slowest in t_1-t_2, and these are those corresponding to diffusive modes (e.g., thermal or mass diffusion) with time dependence $e^{-\Gamma_\alpha k^2|t_1-t_2|}$. All other correlations can be evaluated at zero wave vector. With this, the hydrodynamic contribution to W_d can be written as

$$W_d \sim -\beta\sum_\alpha \int_0^t dt_1 \int_0^{t_1} dt_2 \times \frac{\Psi_\alpha^*(t_1)\Psi_\alpha(t_2) S_d \gamma(d/2, \Gamma_\alpha k_c^2|t_1-t_2|)}{2(2\pi)^d(\Gamma_\alpha|t_1-t_2|)^{d/2}}, \quad (4.21)$$

where $S_d \equiv 2\pi^{d/2}/\Gamma(d/2)$ is the area of a d-dimensional unit hypersphere, $\Gamma(x)$ is the gamma function, $\gamma(x,y)$ is the incomplete gamma function, the sum is over all diffusive modes, each with diffusivity Γ_α, k_c is a large-wave-number spherical cutoff necessitated by ignoring the k dependence of the correlation functions, and

$$\Psi_\alpha(t) \equiv \mathsf{u}_\alpha^\dagger \langle\langle \mathsf{C}_{\mathrm{T}}\mathsf{C}_{\mathrm{T}}\rangle\rangle^{-1}\langle\langle \mathsf{C}_{\mathrm{T}}\mathsf{A}\rangle\rangle \dot{\mathsf{F}}(t), \quad (4.22)$$

where a hydrodynamic approximation for $\langle\langle \mathsf{C}_{\mathbf{k}}(t)\mathsf{C}_{-\mathbf{k}}\rangle\rangle$ was used—i.e.,

$$\langle\langle \mathsf{C}_{\mathbf{k}}(t)\mathsf{C}_{-\mathbf{k}}\rangle\rangle \sim L^d \sum_\alpha \mathbf{u}_\alpha e^{(ik\omega_\alpha - \Gamma_\alpha k^2)t}\mathbf{u}_\alpha^\dagger, \quad (4.23)$$

where u_α is the amplitude of the αth mode.

Only purely diffusive modes (those with $\omega_\alpha=0$) are kept. Clearly, this part of the so-called dissipative contribution to the work contains slowly decaying (in-time) contributions, and to get a feel for how important they are, we now assume constant rates; i.e., we assume that the Ψ_α's are constant in time. With this, the integrals in Eq. (4.21) can be done and give

$$W_d \sim -\sum_\alpha \beta|\Psi_\alpha|^2 k_c^d t^2 f(d, k_c^2\Gamma_\alpha t), \quad (4.24)$$

where

$$f(d,x) \equiv \frac{S_d}{(2\pi)^d}\left(\frac{2x^{-d/2}\gamma(d/2,x)}{(d-2)(d-4)} + \frac{e^{-x}-1}{x^2(d-4)} + \frac{2e^{-x}+d-4}{x(d-2)(d-4)}\right). \quad (4.25)$$

At long times, specifically for $k_c^2\Gamma_\alpha t \gg 1$, it follows from this last result that

$$W_d \propto \sum_\alpha |\Psi_\alpha|^2 \begin{cases} t^{2-d/2} \text{ for } d<2, \\ t\ln(t) \text{ for } d=2, \\ t \text{ for } d>2. \end{cases} \quad (4.26)$$

In the quasistatic limit—namely, $\dot{\mathsf{F}}\to 0$, $t\to\infty$ keeping $\dot{\mathsf{F}}t$ constant—this last result implies that W_d vanishes and the Jarzynski bound for the work becomes an equality for localized probes in any spatial dimension. However, note that W_d decays more slowly, all other things being equal, for $d\leqslant 2$.

In summary, response theory shows that the free energy change defined by the Jarzynski equality in general is not the state function that arises in work bounds in thermodynamics, *even for quasistatic processes*. Instead it is an upper bound to the well-known thermodynamic ones. Under special conditions—namely, where there is no coupling between the perturbation and macroscopic state variables—the two bounds become equivalent, but as we have shown, this is probably not the case in macroscopic, nonisothermal [and constant chemical potential, total density, etc., processes; cf. Eq. (4.20) and the discussion that follows]. One important exception to this result is for broad-spectrum (in wave vector) probes—i.e., ones that are spatially localized. Basically, there, the couplings to the very long-wavelength slow modes are weak and a quasistatic limit is possible. In some sense this is similar to the observation in Sec. III for the limit $L_f/L_i \to 1$, although both here and there, the amount of work done becomes small.

V. DISCUSSION

In this article, we have studied the Jarzynski equality both generally and within the context of a well-understood system, the one-dimensional expanding ideal gas. We chose this latter system because it has been shown that the Jarzynski equality is satisfied unambiguously for all choices of system length and piston velocity. The simplicity of this system allowed us to obtain a general form for the full distribution function at any time during and after the expansion. Therefore, the nonequilibrium state of this system is completely known at all times.

We showed that, although the Jarzynski equality is correct and clearly shows that $\langle e^{\beta W} \rangle$ is an invariant, equal to a ratio of equilibrium canonical partition functions, its applicability to nonequilibrium phenomena, and in particular to nonequilibrium thermodynamics, is more subtle. First, at the end of the expansion process, the distribution function is not canonical, and moreover, quantities like the free energy and entropy cannot be defined in terms of their standard equilibrium definitions. Second, if equilibrium is ever attained (remember that our model system never quite reaches equilibrium), the free energy of the final state will, in general, be different than what can be suggested from the Jarzynski equality. This is a consequence of the final temperature that is usually different than the initial one. In fact, in reply to the criticism of Cohen and Mauzerall [13], Jarzynski [3] pointed out that the final free energy that appears in the Jarzynski equality will be the one of the final state, provided we wait long enough and that the overall system contains a large enough bath such that the temperature change can be assumed to be zero. This argument is correct, but should be taken with caution. In fact, there is no limit to the external work done and it is always possible to perform enough work, on a macroscopic region of the system, such that, even with a large bath, the temperature of the full system changes. In the experiment of Liphardt *et al.* [12], where the work is done on a single molecule in solution, it would seem safe to assume that the final and initial temperatures are the same. This also happens, not surprisingly, for our model system: when the system is very large compared to the expansion, $L \gg Vt$, the temperature change tends towards zero [recall Eqs. (3.9)]. Fundamentally, however, the 1D expanding gas model represents an adiabatic process and there must be some temperature change, unless the gas is coupled to a thermostat, in which case, ΔA_J has been shown to be equal to the actual energy change of the system [28].

In Sec. III, we compared the bound implied by the Jarzynski equality against a thermodynamic bound which is obtained from the first and second law of thermodynamics. As shown by Eq. (3.7) or (3.11), the Jarzynski bound is an upper bound to the thermodynamic upper bound to the work and, hence, is does not prove the second law of thermodynamics. The two bounds become close to each other in the limit where the extent of the expansion is small. This is another consequence of the fact that the work done can produce changes in key quantities that define the ensembles in the initial and the final equilibrium states [e.g., temperature; cf. Eq. (4.20) and the subsequent discussion].

We also showed that the statement that $-\Delta A_J = W_{rev}$ can be incorrect even in the quasistatic limit and, in particular, is wrong for our system (see, e.g., Fig. 4). This is an important observation since this assumption is often made in experiments. In the experiment by Liphardt *et al.*, this assumption may be justified because the work is done on a very small part of the system. On the other hand, as we have shown in Eq. (4.19), any coupling to long-wavelength fluctuations of conserved quantities leads to a negative correction to W_d. Liphardt *et al.* measure quantities like $-\Delta A_J - \langle W \rangle_{\text{quasistatic}}$; hence, if their apparatus is not the ideal δ-function coupling considered in the preceding section, but contains some small, even $O(1/N)$, coupling to the conserved quantities, Eq. (4.19) predicts that they should obtain

$$\langle W_{JE} - W_{A,rev} \rangle = -\frac{\beta}{2} \langle\langle \mathsf{A}_\mathrm{T} \mathsf{C}_\mathrm{T} \rangle\rangle \langle\langle \mathsf{C}_\mathrm{T} \mathsf{C}_\mathrm{T} \rangle\rangle^{-1} \langle\langle \mathsf{C}_\mathrm{T} \mathsf{A}_\mathrm{T} \rangle\rangle : \mathsf{F}(t) \mathsf{F}(t) \tag{5.1}$$

in the linear response regime (here we are using the notation and sign convention of Liphardt *et al.* for the work). Thus a quadratic, negative correction should be seen; interestingly, this is exactly what is shown in Fig. 3A of Ref. [12], although there the authors dismiss this as experimental error. If this is the explanation, the coupling must clearly arise from the macroscopic parts of the experimental device, here probably associated with the piezoelectric actuator used to pull the molecule.

A similar observation was made by Oberhofer *et al.* [16] in their numerical study of equilibrium free energies using the schemes based on the Jarzynski equality and on the Widom insertion method in a soft sphere liquid. There, $\langle W_{JE} - W_{A,rev} \rangle$ was also shown to be nonzero and this result was explained in terms of adiabatic invariants. Numerically, they obtained a definite sign for $\langle W_{JE} - W_{A,rev} \rangle$ which agrees with our prediction. The argument is basically an example of more general problems encountered in ergodic theory and in the construction of ensembles when phase space is metrically decomposable [36,37]. The analysis presented in Sec. IV for

systems where the work couples to the densities of conserved quantities offers a quantitative estimate for the difference.

We highlight the fact that $-\Delta A_J \neq W_{rev}$ by briefly investigating the Jensen-Peierls-Gibbs-Bogoliubov inequality which can be proved as follows. By defining a function $h(\lambda) \equiv \ln\langle e^{\lambda W}\rangle$ it follows that

$$\frac{dh(\lambda)}{d\lambda} = \frac{\langle We^{\lambda W}\rangle}{\langle e^{\lambda W}\rangle} \equiv \langle W\rangle_\lambda \tag{5.2}$$

and

$$\frac{d^2h(\lambda)}{d\lambda^2} = \frac{\langle W^2e^{\lambda W}\rangle}{\langle e^{\lambda W}\rangle} - \frac{\langle We^{\lambda W}\rangle^2}{\langle e^{\lambda W}\rangle^2} \equiv \langle\langle W^2\rangle\rangle_\lambda, \tag{5.3}$$

where $\langle\langle(\cdots)\rangle\rangle_\lambda$ is a cumulant average. Using the Jarzynski equality and the two averages defined above, it is straightforward to show that

$$-\beta\Delta A_J = \beta\langle W\rangle_{\lambda=0} + \int_0^\beta d\lambda\langle\langle W^2\rangle\rangle_\lambda(\beta-\lambda). \tag{5.4}$$

Because the second term on the right-hand side (we call it the work fluctuation term) is strictly positive, we immediately have $\beta\langle W\rangle_{\lambda=0} \leqslant -\beta\Delta A_J$ ($\langle W\rangle_{\lambda=0}$ is just the usual average work). Hence, $\langle W\rangle = -\Delta A_J$ only when the work fluctuations are zero for all λ. In particular, as stated in Ref. [29], the fluctuations have to be zero in the original canonical ensemble when $\lambda = 0$. These fluctuations can be nonzero even if the work is done quasistatically, as we showed in Sec. IV. In fact, we showed that the fluctuation terms, to lowest order in the external fields, will differ from zero in the quasistatic limit only if the system variables $\mathsf{A}(\mathbf{r},t)$, to which the fields couple, are orthogonal to the $\mathsf{C}(\mathbf{r},t)$'s, the conserved quantities, *or* if the fields have negligible $\mathsf{F}_{k=0}(t)$ components. Recall that, in a case of microscopically localized field, the $\mathsf{F}_{k=0}(t)$ contribution is small and the quasistatic limit is attainable.

We also showed that the true free energy change of the system, for a fixed change in volume, tends to a free energy change that appears in the Jarzynski equality, ΔA_J, when the expansion is very fast (the piston velocity V is large). This result is expected, since, within this model, the average work tends to zero for large piston velocities and produces a negligible temperature change.

Finally, the criticism we wanted to raise in this paper is not against the Jarzynski equality, which we believe to be correct, but rather to how it is often interpreted. We worked on a system that falls in a class described by Hamiltonian dynamics, which is fundamentally adiabatic since the Hamiltonian could, in principle, include everything (in the language of Jarzynski, everything means the system and the bath). We showed that, for that class of systems, the Jarzynski equality can lead to erroneous conclusions. In particular, the free energy change that appears in the equality can have little to do with the actual nonequilibrium thermodynamics free energy computed by standard methods (even when the system equilibrates) and the Jarzynski equality does not contain any information about the nonequilibrium state of the system. Further, we want to reemphasize that the Jarzynski equality cannot be used as a proof of the second law of thermodynamics because the bound that it provides is above the thermodynamical upper bound to the work.

ACKNOWLEDGMENTS

We would like to thank the Natural Sciences and Engineering Research Council of Canada for supporting this work and Irwin Oppenheim and Steve Pressé for useful discussions.

APPENDIX: NONEQUILIBRIUM DISTRIBUTION FUNCTION

As shown in Ref. [10], the positive initial velocity u_0 interval that result in n collisions with the piston at a later time t is

$$(2n-1)(L_i/t+V) - x_0/t < u_0 < (2n+1)(L_i/t+V) - x_0/t, \tag{A1}$$

where x_0 is the initial position of the gas particle. For such a case, the final velocity is

$$u = 2nV - u_0 \tag{A2}$$

and the final position

$$x = -x_0 - u_0t + 2n(L_i + Vt). \tag{A3}$$

When the final position is smaller than zero, it means that there has been a further collision with the hard wall after the last collision with the piston. In such a case, the sign of the final velocity and position in Eqs. (A2) and (A3) is reversed. For negative initial velocities, the interval that leads to n collision is

$$-(2n+1)(L_i/t+V) - x_0/t < u_0 < -(2n-1)(L_i/t+V) - x_0/t. \tag{A4}$$

In this case, the final velocity and position are

$$u = 2nV + u_0 \tag{A5}$$

and

$$x = x_0 + u_0t + 2n(L_i + Vt). \tag{A6}$$

Again, if this results in a negative position, the sign of the last two equations is reversed.

The above results can be used to obtain the final distribution function from an integration over the initial velocities and position weighted by the initial distribution function,

$$f(x,u;t) = \int dx_0 \int du_0 f_0(x_0,u_0) \Bigg\{ \Theta(L_i/t + V - x_0/t - u_0)\Theta(x_0/t + u_0 + L_i/t + V) \times [\Theta(x_0 + u_0 t)\,\delta(u - u_0)\,\delta(x - x_0 - u_0 t) + \Theta(-x_0 - u_0 t)\,\delta(u + u_0)\,\delta(x + x_0 + u_0 t)] + \sum_{n=1}^{\infty} \{\Theta((2n+1)(L_i/t + V) - x_0/t - u_0)\Theta(x_0/t + u_0 - (2n-1)(L_i/t + V)) \times [\Theta(-x_0 - u_0 t + 2n(L_i + Vt))\,\delta(u - (2nV - u_0))\,\delta(x + x_0 + u_0 t - 2n(L_i + Vt)) + \Theta(x_0 + u_0 t - 2n(L_i + Vt))\,\delta(u + (2nV - u_0))\,\delta(x - x_0 - u_0 t + 2n(L_i + Vt))] + \Theta(-(2n-1)(L_i/t + V) - x_0/t - u_0)\Theta(x_0/t + u_0 + (2n+1)(L_i/t + V)) \times [\Theta(x_0 + u_0 t + 2n(L_i + Vt))\,\delta(u - (2nV + u_0))\,\delta(x - x_0 - u_0 t - 2n(L_i + Vt)) + \Theta(-x_0 - u_0 t - 2n(L_i + Vt))\,\delta(u + (2nV + u_0))\,\delta(x + x_0 + u_0 t + 2n(L_i + Vt))]\} \Bigg\}. \tag{A7}$$

After the integrals are performed, Eq. (2.6) is obtained.

[1] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[2] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).
[3] C. Jarzynski, J. Stat. Mech.: Theory Exp. (2005) P09005.
[4] Gavin E. Crooks, J. Stat. Phys. **90**, 1481 (1998).
[5] Shaul Mukamel, Phys. Rev. Lett. **90**, 170604 (2003).
[6] Massimiliano Esposito and Shaul Mukamel, Phys. Rev. E **73**, 046129 (2006).
[7] Rahul Marathe and Abhishek Dhar, Phys. Rev. E **72**, 066112 (2005).
[8] Chris Oostenbrink and Wilfred F. van Gunsteren, Chem. Phys. **323**, 102 (2006).
[9] Steve Pressé and Robert Silbey, J. Chem. Phys. **124**, 054117 (2006).
[10] Rhonald C. Lua and Alexander Y. Grosberg, J. Phys. Chem. A **109**, 6805 (2005).
[11] I. Bena, C. Van den Broeck, and R. Kawai, Europhys. Lett. **71**, 879 (2005).
[12] Jan Liphardt, Sophie Dumont, Steven B. Smith, Ignacio Tinoco, Jr., and Carlos Bustamante, Science **296**, 1832 (2002).
[13] E. G. D. Cohen and David Mauzerall, J. Stat. Mech.: Theory Exp. (2004) P07006.
[14] E. G. D. Cohen and D. Mauzerall, Mol. Phys. **103**, 2923 (2005).
[15] The motivation of a subdivision of the mechanics into system and bath is primarily to introduce a heat bath, whereby isothermal processes may be realized [3,6]. Nonetheless, this should suffice to coarse grain the system reduced distribution. On the other hand, some of the analysis seems to have been overinterpreted in these approaches. By associating the mechanical work, which is assumed to couple directly to the system, with the thermodynamic work, and analyzing the change in the system's energy terms identified with heat are found. This ignores the possibility of worklike interactions between the system and bath (e. g., mass transport, volume expansion, polarization effects, etc.).
[16] Harald Oberhofer, Christoph Dellago, and Phillip L. Geissler, J. Phys. Chem. B **109**, 6902 (2005).
[17] B. Widom, J. Chem. Phys. **39**, 2808 (1963).
[18] D. W. Jepsen, J. Math. Phys. **23**, 405 (1965).
[19] J. L. Lebowitz and J. K. Percus, Phys. Rev. **155**, 122 (1967).
[20] Carlo Cercignani, *Theory and Application of the Boltzmann Equation* (Elsevier, New York, 1975).
[21] P. Résibois and M. De Leener, *Classical Kinetic Theory of Fluids* (Wiley, New York, 1977).
[22] J. Piasecki, J. Stat. Phys. **104**, 1145 (2001).
[23] V. Balakrishnan, I. Bena, and C. Van den Broeck, Phys. Rev. E **65**, 031102 (2002).
[24] S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (Dover, New York, 1984).
[25] W. Pauli and M. Fierz, Z. Phys. **106**, 572 (1937).
[26] R. C. Tolman, *The Principles of Statistical Mechanics* (Oxford University Press, New York, 1938).
[27] John G. Kirkwood, J. Chem. Phys. **15**, 72 (1947).
[28] A. Baule, R. M. L. Evans, and P. D. Olmsted, Phys. Rev. E **74**, 061117 (2006).
[29] Sanghyun Park and Klaus Schulten, J. Chem. Phys. **120**, 5946 (2004).
[30] R. Kubo, Proc. Phys. Soc. Jpn. **17**, 1100 (1962).
[31] D. N. Zubarev, *Nonequilibrium Statistical Thermodynamics* (Consultants Bureau, New York, 1974).
[32] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Benjamin, Reading, MA, 1975).
[33] R. Zwanzig, in *Lectures in Theoretical Physics* (Interscience, New York, 1961), Vol. 3.
[34] H. Mori, Prog. Theor. Phys. **34**, 423 (1965).
[35] Bruce J. Berne and Robert Pecora, *Dynamics Light Scattering* (Dover, New York, 2000).
[36] D. Ronis and I. Oppenheim, Physica A **86**, 475 (1977).
[37] A. Ya. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 2002).