Jarzynski on work and free energy relations: The case of variable volume

Original
Jarzynski on work and free energy relations: The case of variable volume / Ciccotti, G., Rondoni, L.. - In: AICHE JOURNAL. - ISSN 0001-1541. - STAMPA. - 67:1(2021), pp. 1-6. [10.1002/aic.17082]

Publisher:
John Wiley and Sons

Published
DOI:10.1002/aic.17082

AIChE Journal
DOI 10.1002/aic.17082

THERMODYNAMICS AND MOLECULAR-SCALE PHENOMENA

Jarzynski on work and free energy relations: the case of variable volume

G. Ciccotti[1,2,3,6] L. Rondoni[4,5,7]

[1] Istituto per le Applicazioni del Calcolo "Mauro Picone" IAC-CNR, Roma, Italia
[2] Dipartimento di Fisica, Universit'a di Roma "La Sapienza", Roma, Italia
[3] School of Physics, University College Dublin (UCD), Dublin, Ireland
[4] Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, 10125, Italia
[5] INFN, Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italia
[6] ORCID: 0000-0003-0764-9657
[7] ORCID: 0000-0002-4223-6279

September 28, 2020

Abstract

Derivations of the Jarzynski equality (JE) appear to be quite general, and applicable to any particle system, whether deterministic or stochastic, under equally general perturbations of an initial equilibrium state at given temperature T. At the same time, the definitions of the quantities appearing in the JE, in particular the work, have been questioned. Answers have been given, but a deeper understanding of the range of phenomena to which the JE applies is necessary, both conceptually and in order to interpret the experiments in which it is used. In fact, domains in which the JE is not applicable have been identified. To clarify

Received: Jun 19, 2019; Revised: Sep 22, 2020; Accepted: Sep 27, 2020

the issue, we scrutinize the applicability of the JE to a Hamiltonian particle system in a variable volume. We find that, in this case, the standard interpretation of the terms appearing in the JE is not adequate.

1 Introduction

At the end of the twentieth century, relations were developed to describe the fluctuations of the energy dissipation of particle systems in a nonequilibrium steady state, now generically known as Fluctuation Relations [1]. They have been considered both as the apex of almost a century of investigations into the consequences of the fluctuations of the microscopic constituents of macroscopic objects [2] and, at the same time, they have opened a new line of research on the properties of small (mesoscopic) systems. The literature on the subject is vast. References [3, 4, 5] summarize those developments, which continued with the derivation of different kinds of Fluctuation Relation, concerning transient rather than steady states.

Among the last ones, the relation known as the Jarzynski Equality (JE) [6] has become popular in different fields, biophysics in particular. Indeed, it is meant to relate the free energy variations of a system to the work done on it when it is driven away, in an arbitrary fashion, from a canonical equilibrium state. This is particularly interesting in situations in which quasi-static thermodynamic transformations are not possible, either because the time scale of the experiment is too short, or because signals are heavily affected by fluctuations.

The JE can be suitably illustrated following Ref.[7]. To do that, let us consider an N-particle system S, $N \geq 1$, interacting with an environment E, and denote by S+E the combined system. Let $\Gamma = (x, y)$ represent the microscopic state of S+E in the phase space $\mathcal{M}$, with $x = \left(q^{(S)}, p^{(S)}\right)$ the canonical coordinates and momenta of S, and $y = \left(q^{(E)}, p^{(E)}\right)$ those of E. Let the dynamics be determined by a time dependent Hamiltonian:

$$\mathcal{H}(\Gamma;\lambda) = H_S(x;\lambda) + H_E(y) + H_{\text{int}}(x,y) \tag{1}$$

where the parameter λ varies according to a given protocol $\lambda = \lambda(t)$, during the time interval $[0,\tau]$. The variation of λ is due to unspecified external agents that, changing λ, perform a given work W on S. The number of particles N can indifferently be large or small, and the time dependence of λ can be chosen without limitations.

Let S+E be initially in thermodynamic equilibrium with a bath B at temperature T, and take $\lambda(0) = \alpha$ and $\lambda(\tau) = \omega$. At time 0, an initial condition, $\Gamma_0 = (x_0, y_0) \in \mathcal{M}$ is picked at random from the corresponding canonical distribution,

$$f_\alpha(\Gamma) = \frac{1}{Q_\alpha} e^{-\beta\mathcal{H}(\Gamma;\alpha)} \tag{2}$$

where $\beta = 1/k_{_B}T$, $k_{_B}$ is the Boltzmann constant, the Hamiltonian is evaluated at $\lambda(0) = \alpha$, and

$$Q_\lambda = \int e^{-\beta\mathcal{H}(\Gamma;\lambda)} \mathrm{d}\Gamma \tag{3}$$

defines the partition function for an equilibrium system at temperature T, and Hamiltonian $\mathcal{H}(\Gamma;\lambda)$ for any λ. At time $t = 0$, S+E is isolated from B, and the external action is switched on, making λ change from its initial value $\lambda(0) = \alpha$ to its final value $\lambda(\tau) = \omega$, while S+E evolves as determined by the time dependent Hamiltonian $\mathcal{H}(\Gamma;\lambda(t))$. Denoting by $S^t : \mathcal{M} \to \mathcal{M}$ the corresponding evolution operator, the final microscopic state is $S^\tau\Gamma_0 = (x_\tau(\Gamma_0), y_\tau(\Gamma_0))$, while the final value of the energy of S+E is $\mathcal{H}(S^\tau\Gamma_0;\omega)$.

The same procedure is repeated a large number of times, and each time S+E is found in a different unknown initial condition $\Gamma_0 \in \mathcal{M}$. As the Hamiltonian is given and the protocol λ is deterministically fixed, one may say that the work for each trajectory only depends on Γ_0. It is still useful to make explicit its functional dependence on the trajectory and on the protocol as follows:

$$W = W\left[S^t\Gamma_0; 0 \leq t \leq \tau\right] \tag{4}$$

Jarzynski now introduces the following definitions:

$$H_S^*(x;\lambda) = H_S(x;\lambda) - \frac{1}{\beta}\ln\frac{\int \mathrm{d}y e^{-\beta[H_E(y)+H_{\mathrm{int}}(x,y)]}}{\int \mathrm{d}y e^{-\beta H_E(y)}}\,, \tag{5}$$

$$F_S(\lambda) = -\beta \ln Z_\lambda\,; \quad Z_\lambda = \int \mathrm{d}x\, e^{-\beta H^*(x;\lambda)} \tag{6}$$

and the quantity

$$\begin{aligned} W_J\left[S^t\Gamma_0; 0 \leq t \leq \tau\right] &\equiv \int_0^\tau \mathrm{d}t\,\dot{\lambda}\frac{\partial H_S}{\partial \lambda}(x_t(\Gamma_0);\lambda(t)) & (7)\\ &= \mathcal{H}(S^\tau\Gamma_0;\omega) - \mathcal{H}(\Gamma_0;\alpha) & (8)\end{aligned}$$

which is interpreted as work. Here, the second line follows from the fact that $\mathcal{H}$ depends explicitly on time only through λ, which in turn only appears in H_S:

$$\frac{\partial H_S}{\partial t} = \dot{\lambda}\frac{\partial H_S}{\partial \lambda} = \dot{\lambda}\frac{\partial \mathcal{H}}{\partial \lambda} = \frac{\partial \mathcal{H}}{\partial t} = \frac{\mathsf{d}\mathcal{H}}{\mathsf{d}t} \tag{9}$$

Given the above definitions, the JE states that

$$\left\langle e^{-\beta W_J}\right\rangle_\alpha = e^{-\beta[F_S(\omega)-F_S(\alpha)]} \tag{10}$$

where the angular brackets represent a canonical average with respect to the *initial* ensemble f_α, and $F_S(\lambda)$ is interpreted as the free energy of S when S+E is in canonical equilibrium at temperature T and parameter λ.

Although at time τ the system is typically not in equilibrium at temperature

T with $\lambda = \omega$, Eq.(10) is expected to compute the free energy variation of S. If possible, this would be extremely useful in those cases in which thermalization is not feasible, which are frequent in mesoscopic scale experiments.

The generality in which the JE is claimed to hold is striking: its validity is stated to bridge all imaginable space and time scales, as long as a classical Hamiltonian description and the canonical ensemble are valid. As this is rarely the case for any law of physics, various critical investigations of its validity have ensued. For instance, Ref.[8] questions the validity of the experimental tests, while Refs.[9, 10], question the definition of work. Replies were provided [7, 17] and various answers exist, but the debate appears to have left some issues unresolved, see *e.g.* Refs.[11, 12, 13]. For instance, it is acknowledged that the definitions of the various physical quantities given by Jarzynski may not correspond in general to the usual notions of mechanical and thermodynamical quantities. At the same time, this is not considered a problem, since the JE is mainly intended to describe small systems; see *e.g.* Chapter 6 of Ref.[14], which argues that the macroscopic and the mesoscopic notions of heat may as well be different, hence all other thermodynamic quantities may. Other authors argue that the relevant notion of work is W_J since, once averaged, it could represent the thermodynamic work done by S on the steering bodies [17].

In this respect, one difficulty might be that such bodies are macroscopic objects, that abide by the usual mechanical and thermodynamical laws. In addition, measuring the thermodynamic work through external bodies leads to the difficulty that the external work is measurable through S only when the transformations are (thermodynamically) reversible. Only in that case, in fact, do the values of the internal and the external work differ merely by a sign, and knowledge of the external work amounts to knowledge of the internal work, which is the one contributing to the variations of the state of S.

In certain cases, such as those described by Langevin equations for a single or very few degrees of freedom, that are of interest in many experiments at the mesoscopic scale [23], the coincidence of different quantities may overcome these and other difficulties. However, for the full generality attributed to the JE, further investigation is desired. As a matter of fact, some limitations on its applicability have already been reported in special situations, such as those in which the system does not return to its initial state, although λ evolves in a closed loop with $\lambda(0) = \lambda(\tau)$, cf. Ref.[15] and references therein.[1]

Improving our understanding of these facts may help better define the range of applicability of the JE, and to interpret experiments [21, 23, 22]. Therefore, in this paper, we discuss some of the issues raised by the JE formalism, investigating the simple, but also paradigmatic example of an object with a variable volume. The following section discusses the derivation of the JE. Then we analyse in detail our specific example, and the last section summarizes our conclusions.

2 Derivation of the JE

In his paper [7], replying to Cohen and Mauzerall [8], who had criticized the interpretation of protein stretching experiments in terms of the JE, Jarzynski developed a derivation of his equality that can be considered totally general, for Hamiltonian particle systems driven by external (steering) fields. We begin recalling this derivation.

The universe treated in [7] consists of the system of interest S, its environment E, and a super-environment, which is a bath B at temperature T. The external driving mechanisms are supposed to act directly on the energy of S,

---

[1] Plastic deformation seems to belong to a similar class of phenomena. Here, the system is permanently deformed by external objects that move from an initial configuration, $\lambda(0)$, and return to it at the end of the cycle, $\lambda(\tau) = \lambda(0)$, cf. *e.g.* Refs.[18, 19].

varying the parameter λ, and to have no effect on E and B. Thinking of protein stretching experiments, S is the protein, E is the solution in which S is immersed, B is the laboratory, and the external device is the optical tweezers that pull a lattice bead to which one end of the protein is attached. Furthermore, the whole system S+E is initially assumed to be in equilibrium with B, so that the statistics of its *initial* micro-states is canonical.

The Hamiltonian $\mathcal{H}$ is then decomposed in three parts, as in Eq(1): 1. $H_S(x;\lambda)$, whose average is thought to represent the internal energy of S, which can be modified by external agents modifying the parameter λ; 2. $H_E(y)$, the internal energy of E, supposed to be insensitive to external actions; 3. $H_{\mathrm{int}}(x,y)$, the interaction energy of S and E.

Note that identifying (the average of) H_S with the thermodynamic internal energy is not correct in general, since that requires the average of the interaction energy H_{int} to be negligible with respect to H_S. All thermodynamic fields are defined within cells that are small compared to the macroscopic scales, but so large with respect to the molecular scales that bulk quantities dominate over surface contributions. Therefore, whenever H_{int} is not sufficiently small, neither H_S nor its average can be taken as genuine properties of the system S. This may have an impact on the applicability of Jarzynski's definitions, possibly even admitting that consistency with macroscopic quantities is not necessary in dealing with small systems [14]. Consistently with Sekimoto's argument, Ref.[17] argues that the natural extension of the elementary thermodynamic work dW to fluctuating instantaneous work, is the work performed by S on the external steering bodies, and is given by Eq.(7), *i.e.* that $W \equiv W_J$. Therefore, one can eventually write:

$$\left\langle e^{-\beta W}\right\rangle_{\alpha} = \left\langle e^{-\beta W_J}\right\rangle_{\alpha} = \int \mathrm{d}\Gamma e^{-\beta W(\Gamma)} \frac{e^{-\beta \mathcal{H}(\Gamma;\alpha)}}{Q_{\alpha}}$$
$$= \frac{Q_{\omega}}{Q_{\alpha}} = e^{-\beta(F_{\mathrm{S+E}}(\omega) - F_{\mathrm{S+E}}(\alpha))} \tag{11}$$

where $[F_{\mathrm{S+E}}(\omega) - F_{\mathrm{S+E}}(\alpha)]$ is supposed to be the free energy variation of S+E. This is not yet the desired quantity. One wishes to compute the free energy variation of S alone, supposedly $[F_S(\omega) - F_S(\alpha)]$. Thanks to the definition (6), this is obtained observing that the denominator of:

$$Z_{\lambda} = \frac{\mathcal{Q}_{\lambda}}{\int \mathrm{d}y \exp(-\beta H_E(y))} \tag{12}$$

does not depend on λ, so that

$$\frac{Q_{\omega}}{Q_{\alpha}} = \frac{Z_{\omega}}{Z_{\alpha}} \tag{13}$$

which is the JE (10).

Given the Jarzynski framework outlined above, it is worth stressing that Eq.(10) and its applicability rest on the following assumptions:

1) the quantity F_S defined by Eq.(6) does represent the free energy of S,

2) W_J represents the microscopic observable generating the macroscopic work.

In Ref.[6], point 1) is considered justified in equilibrium situations, such as those represented by the canonical distribution with $\lambda = \omega$. As argued by Peliti, [16, 17], referring to Gibbs' book, [30], *e.g.* page 42, point 2) is justified when W_J is the thermodynamic work in quasi-static transformations. However, as noted also by Peliti, Gibbs only considers analogies between thermodynamic and mechanical quantities, and the thermodynamic ones need averaging with respect to the proper ensemble, cf. *e.g.* page 45 of [30]. Moreover, Gibbs' approach requires quasi-static transformations, which exclude the irreversible ones.

Finally, if the quantity W_J provides the thermodynamic work only after averaging, it cannot be exponentiated and then averaged as if averaging some function of work. This does not forbid averaging the exponential of the non-averaged W_J, but this is not the exponential of the work that contributes to the variations of the properties of S. The right work is the one performed by the external agents on S.

Even granting that W_J is the proper quantity to consider, at least for the state variation of a small system S, it remains that the correspondingly small free energy difference concerning this variation is referred to the variation of free energy of S+E, which is huge compared to that of S alone. The most obvious reason is that obtaining a small number from the difference of two very large numbers is problematic in general.

As numerous experiments have nevertheless verified the JE [20, 21, 22], further analysis seems to be necessary. Below we consider a simple, but paradigmatic example, in which the JE as developed above does not seem to hold.

3 A critical case: S in a variable volume

To exemplify some of the difficulties associated with the derivation illustrated above, let us consider some critical questions related to a specific model. We begin observing that the quantity W_J defined by Eq.(8) does not always represent a real work. In fact, while the JE is of particular interest when quasi-static transformations are impossible, W_J expresses a real work contributing to the variations of the properties of S only in quasi-static transformations, that is to say only when it equals, apart from a sign, the work done on S.

Take for instance a system S with a variable volume $\lambda = V_S(t)$ which defines the protocol. Let V_{tot} denote the total volume available to S+E. The volume V_{tot} can be finite, variously subdivided between S and E, or infinitely large depending on the case at hand. For example, suppose that S alone is within the volume V_S, so that E occupies the rest of V_{tot}, $V_E = V_{\mathrm{tot}} - V_S$. Then, taking the $V_{\mathrm{tot}} \to \infty$ limit, H_E becomes volume independent, while H_{int} continues to depend on V_S. Would the JE apply in this case, it would allow the calculation of the free energy of a dense gas, in quite a simple fashion, unlike the complicated currently known perturbative methods, that start from the analytically computable ideal gas contribution. The general result can be summarized as follows.

1) The variations of V_S affect not only S, but also E, either directly via a V_S-dependence in the energy of E, expressed by $H_E = H_E(y; V_{\mathrm{tot}} - V_S)$, or indirectly, through the term $H_{\mathrm{int}} = H_{\mathrm{int}}(x, y; V_S, V_{\mathrm{tot}} - V_S)$. Here, indeed, even if the V_S dependence of H_E can be eliminated by taking the $V_{\mathrm{tot}} \to \infty$ limit, H_{int} remains V_S dependent.

2) The quantity

$$W_J = \int_0^\tau \dot{V}_S \frac{\partial H_S}{\partial V_S} \mathrm{d}t \tag{14}$$

cannot be taken by itself as mechanical work. Moreover, the virial theorem implies that in equilibrium it represents only a contribution to the thermodynamic work done by S on the external bodies, that are responsible for the variations of V_S. The reason is the neglect of the contribution of H_{int} to the virial.

In principle, the Hamiltonian does not contain the volume, since that is neither part of the kinetic energy, nor of the potential energy. The procedure to introduce the volume in $\mathcal{H}$ follows from the rescaling of the configuration coordinates of the system of interest, a procedure attributed to H.S. Green, cf. [27, 28].

In this procedure, we consider first the Hamiltonian of the N-particle system in a cube V_S of side L, contained in the total very large (virtually infinite) volume V_{tot} hosting also E. One can write:

$$\mathcal{H}(q,p) = \mathcal{H}(L\rho^{(S)}, q^{(E)}, \pi^{(S)}/L, p^{(E)}) = K_S\left(\pi^{(S)}/L\right) + K_E\left(p^{(E)}\right)$$
$$+U_S\left(L\rho^{(S)}\right) + U_E\left(q^{(E)}\right) + U_{\mathrm{int}}\left(L\rho^{(S)}, q^{(E)}\right) \tag{15}$$

where $\rho^{(S)} = q^{(S)}/L$ is the set of dimensionless configuration coordinates of S, $\pi^{(S)} = p^{(S)}L$ the set of the associated momenta, K_S and K_E are the kinetic energies of S and E, respectively, U_S and U_E are corresponding potential energies, and $U_{\mathrm{int}} = H_{\mathrm{int}}$. In our case

$$H_S(\rho^{(S)}, \pi^{(S)}; L = V_S^{1/3}) = \sum_{i=1}^{N_S} \frac{{\pi_i^{(S)}}^2}{2mL^2} + U_S\left(L\rho^{(S)}\right) \tag{16}$$

so that

$$\frac{\partial H_S}{\partial V_S} = \frac{\partial H_S}{\partial L}\cdot\frac{\partial L}{\partial V_S} = \frac{1}{3L^2}\left[-\sum_{i=1}^{N_S}\frac{{p_i^{(S)}}^2}{mL} + \sum_{i=1}^{N_S}\rho_i^{(S)}\frac{\partial U_S}{\partial q_i^{(S)}}\right] \tag{17}$$

$$= \frac{1}{3V_S}\sum_{i=1}^{N_S}\left[-\frac{{p_i^{(S)}}^2}{m} + q_i^{(S)}\frac{\partial U_S}{\partial q_i^{(S)}}\right] = -\frac{1}{3V_S}\sum_{i=1}^{N_S}\left(\frac{{p_i^{(S)}}^2}{m} + q_i^{(S)}\cdot F_i\right) \tag{18}$$

hence

$$\frac{\partial H_S}{\partial V_S}\cdot\dot{V}_S = -\frac{1}{3}\frac{\dot{V}_S}{V_S}\sum_{i=1}^{N_S}\left(\frac{{p_i^{(S)}}^2}{m} + q_i^{(S)}\cdot F_i\right) \tag{19}$$

For an infinitesimal quasi-static transformation one can write

$$\left\langle \frac{1}{3V_S}\sum_{i=1}^{N_S}\left(\frac{{p_i^{(S)}}^2}{m} + q_i^{(S)}\cdot F_i^{\mathrm{int}}\right)\right\rangle_{\mathrm{eq}} = P \tag{20}$$

where F_i^{int} is the *total internal* force acting on the i-th particle of S, and the subscript "eq" denotes averaging with respect to the actual equilibrium ensemble describing the stationary state. As shown below, averaging over the equilibrium distribution one can indeed write:

$$\left\langle \frac{\partial H_S}{\partial V_S} \cdot \dot{V}_S \, \mathrm{d}t \right\rangle_{\mathrm{eq}} = \left\langle \frac{\partial H_S}{\partial V_S} \right\rangle_{\mathrm{eq}} \mathrm{d}V_S = P\mathrm{d}V_S \tag{21}$$

which is the thermodynamic work done by the internal forces, if averaging is performed with respect to the initial equilibrim ensemble.[2] Therefore, the average of $\mathrm{d}W_J$ represents at most the contribution of the internal interactions[3] to the thermodynamic work $P\mathrm{d}V_S$ done by S on the external bodies that change V_S. However, taking only $\mathrm{d}W_J$, instead of the total energy of S, the averaging yields at most a part of the thermodynamic work done by S on the external bodies, since W_J misses the contribution from H_{int}.

Let us show this last point by computing the pressure of our isolated system. The canonical partition function takes the form:

$$\begin{aligned} \mathcal{Q}^{(S)} = \int_{[0,L]^{3N_S}} \mathrm{d}q^{(S)} \int_{\mathbb{R}^{3N_S}} \mathrm{d}p^{(S)} \, e^{-\beta H_S(q,p)} = \\ L^{3N_S} \int_{[0,1]^{3N_S}} \mathrm{d}\rho^{(S)} \int_{\mathbb{R}^{3N_S}} \mathrm{d}p^{(S)} \, e^{-\beta H_S(L\rho^{(S)},p^{(S)})} \end{aligned} \tag{22}$$

For the pressure in the volume $V = L^3$, we then obtain:

$$\begin{aligned} P = k_{\scriptscriptstyle B} T \frac{\partial}{\partial V} \ln \mathcal{Q} = k_{\scriptscriptstyle B} T \frac{\mathrm{d}L}{\mathrm{d}V} \frac{\partial}{\partial L} \ln \mathcal{Q} = \frac{k_{\scriptscriptstyle B} T}{3L^2} \frac{\partial}{\partial L} \ln \mathcal{Q} \\ = \frac{k_{\scriptscriptstyle B} T N}{V} - \frac{1}{3V\mathcal{Q}} \times \\ \int \mathrm{d}q^{(S)} \mathrm{d}p^{(S)} \, e^{-\beta H_S(q^{(S)},p^{(S)})} q^{(S)} \frac{\partial U_S(q^{(S)})}{\partial q^{(S)}} \end{aligned} \tag{23}$$

$$= \frac{k_{\scriptscriptstyle B} T N}{V} - \frac{1}{3V} \left\langle q^{(S)} \frac{\partial U_S}{\partial q^{(S)}} \right\rangle_{\mathrm{eq}} \tag{24}$$

---

[2] For slow infinitesimal transformations the use of the initial ensemble is allowed. For evolving states, one would have to modify the ensemble as the process unfolds. However that would be possible only for quasi-static transformation that justify the use of varying ensembles to compute the shfting time averages.

[3] The interaction energy H_{int} has been neglected.

Note that, here, no time dependence is involved. Multiplying Eq.(24) by dV yields indeed thermodynamic work. However, it is only in equilibrium that one can speak of macroscopic work here. The quantity in the exponent which is averaged in Eq.(10) cannot be claimed to represent macroscopic work already at $t = 0$, and moreover it is altered by the time evolution. Therefore, the average in Eq.(10) does not lead to any clearly measurable quantity. In fact, in general, the initial equilibrium of S+E at temperature T is driven for a time τ and turns into a non-equilibrium state, hence the expression to be averaged after exponentiation in Eq.(10), represents neither a virial in terms of the initial conditions nor macroscopic work, or any other quantity simply related to a thermodynamic property. In any case, it clearly does not represent the macroscopic work done by S on the external devices, *i.e.* the work that is measured by the operator. The difficulties are even more substantial if the volume of E also changes. In all cases, extracting the free energy variation of S alone seems to be problematic.

4 Concluding remarks

In this paper, we have provided evidence of where the definitions adopted in the derivation of the JE may fail, at least for protocols concerning macroscopic systems. Although consistency between macroscopic mechanics and thermodynamics is not strictly required in some of the specialized literature concerning small systems, our findings may still have partial relevance for them. Indeed, even accepting for small systems the validity of the definitions of Eqs.(7,8), it is not clear at which scale the standard definition of work should replace W_J. Moreover, the experimental measurements are performed on macroscopic objects. Note that protocol realisations may well end in a nonequilibrium state. Indeed, the appeal of the JE lies precisely in the possibility of obtaining the free energy difference between initial and final equilibrium states, with no need to reach the final equilibrium. We summarised a standard derivation of the JE to show how this can be the case.

The differences between mechanical and thermodynamical work, besides the fact that certain quantities only represent partial works, may also play a role in the results. In particular, we have noted that in our case W_J is thermodynamic measurable work only in equilibrium. This does not prevent to compute by simulation the average of the exponential of W_J. However, an experimentalist cannot measure such a quantity. Furthermore, in the likely event that work cannot be exclusively done on the system of interest, but also on its environment, one should check that this fraction of work remains negligible with respect to the quantity of interest. As a matter of fact, the protocol may in general affect the interaction energy H_{int} as well.

From a practical point of view, it remains that the free energy variation of S is obtained from the free energy difference of S+E between the initial and final equilibrium states. That can be highly problematic, since the JE procedure aims to extract a generally small quantity from the difference of two large numbers.

While all the above does not seem to have caused difficulties in experimental mesoscopic checks, other experiments, on the same mesoscopic scale, could be affected. It thus seems that the use of the JE should be validated on a case by case basis. Other counterexamples, such as those of "absolute irreversibility" [15] or of indentation [18, 19] also suggest that this is the case. That violations of the JE are generically possible in systems that drastically violate the assumptions of the theory has already been pointed out. For instance, Ref.[23] mentions systems with no Hamiltonian description. We provide here instead a case in which the JE does not apply, even though Jarzynski's framework is preserved.

For instance, ideal or non-ideal gases constitute variable volume systems commonly considered suitable for the application of the JE [29, 31, 32, 33]. In Refs.[29, 32] there is no environment, therefore the varying paramenter λ correctly appears in H_S only, and H_E and H_{int} can be neglected. In such cases, the only question is the experimental measurement of W_J, and the efficiency of the JE as a statistical estimator of the variations of the free energy. In fact, Lua and Grosberg show that while the formula has theoretical meaning for an isolated gas, it has two major drawbacks. First of all, an estimator of the work exponentiated does not represent anymore an exponentiated work, therefore this quantity has not a clear experimental meaning. Secondly, as demonstrated in Ref.[29], the number of repetitions of the protocol needed to verify the JE grows exponentially with the system size. Our contribution concerns cases in which H_E is present and $H_{\text{int}} \neq 0$. Moreover our conclusions go beyond the efficiency and statistical issues.

Consider a case in which the protocol consists in switching on interactions among the particles of an initially ideal gas. In principle, for an expanding such non-ideal gas, the JE yields the absolute energy of the interacting system (the ideal gas contribution is analytically computable), without any difficulty, from a mere canonical average, that does not require complicated perturbative calculations or experiments. We have shown that this is not the case, since the quantity to be averaged after exponentiation is the virial and not W_J.

Acknowledgments

This paper is dedicated to celebrate Keith Gubbins, a real friend and a pioneer in introducing Chemical Engineering to the methods of Statistical Mechanics and Molecular Dynamics, with great mutual advantage for both teeming and fascinating fields. The authors are indebted to Mike Allen for a critical reading of the manuscript and clarifying suggestions. LR is grateful to Javier Varillas

Delgado and Tommaso Vanzan, for enlightening discussions on the validity of the JE. Many thanks also to Luca Peliti, whose lectures in Turin, on Stochastic Thermodynamics and Thermodynamics of Information have been fundamental to appreciate this challenging field. LR acknowledges partial support from MIUR grant Dipartimenti di Eccellenza 2018-2022.

References

[1] D.J. Evans, E.G.D. Cohen and G.P. Morriss, *Probability of second law violations in shearing steady flows*, Phys. Rev. Lett. **71** 2401 (1993)

[2] G. Gallavotti, *Ergodicity, ensembles and beyond*, J. Stat. Phys. **78** 1571 (1995)

[3] D.J. Evans, D.J. Searles, *The fluctuation theorem*, Adv. Phys. **52** 1529 (2002)

[4] L. Rondoni, C. Mejia-Monasterio, *Fluctuation theorems in nonequilibrium statistical mechanics*, Nonlinearity **20** R1 (2007)

[5] B.M.U. Marini, A. Puglisi, L. Rondoni, A. Vulpiani, *Fluctuation–dissipation: response theory in statistical physics*, Phys. Rep. **461** 111 (2008)

[6] C. Jarzynski, *Nonequilibrium equality for free energy differences*, Phys. Rev. Lett. **78** 2690 (1997)

[7] C. Jarzynski, *Nonequilibrium work theorem for a system strongly coupled to a thermal environment*, J. Stat. Mech. P09005 (2004)

[8] E G D Cohen and David Mauzerall, *A note on the Jarzynski equality* J. Stat. Mech. P07006 (2004)

[9] Jose M.G. Vilar, J. Miguel Rubi, *Failure of the Work-Hamiltonian Connection for Free-Energy Calculations*, Phys. Rev. Lett. **100** 020601 (2008)

[10] P. Talkner, P. Hänggi, *Open system trajectories specify fluctuating work but not heat*, Phys. Rev. E **94** 022143 (2016)

[11] Jose M.G. Vilar, J. Miguel Rubi, *Vilar and Rubi Reply*, Phys. Rev. Lett. **101** 098902 (2008)

[12] S. Calabrese, A. Porporato, *Dynamic dependence of nonequilibrium work limits the validity of the Jarzynski Equality*, arXiv:1911.06402

[13] P. Talkner, P. Hänggi, *Statistical Mechanics and Thermodynamics at Strong Coupling: Quantum and Classical*, arXiv:1911.11660 (2020)

[14] K. Sekimoto, *Stochastic Energetics*, Lect. Notes Phys. 799, Springer, Berlin (2010)

[15] Y. Murashita, *Absolute Irreversibility in Information Thermodynamics*, Master Thesis, https://arxiv.org/pdf/1506.04470.pdf

[16] L. Peliti, *Comment on "Failure of the Work-Hamiltonian Connection for Free-Energy Calculations"*, Phys. Rev. Lett. **101** 098903 (2008)

[17] L. Peliti, *On the work–Hamiltonian connection in manipulated systems* J. Stat. Mech. P05002 (2008)

[18] J. Varillas Delgado, *A molecular dynamics study of nanocontact plasticity and dislocation avalanches in FCC and BCC crystals*, PhD thesis Department of Materials Science and Engineering, Universitat Politècnica de Catalunya (2019) https://futur.upc.edu/25256062

[19] J. Varillas, J. Ocenasek, J. Torner, J. Alcalá, *Unraveling deformation mechanisms around FCC and BCC nanocontacts through slip trace and pileup topography analyses*, Acta Materialia **125** 431-441 (2017)

[20] C. Bustamante, J. Liphardt, F. Ritort, *The Nonequilibrium Thermodynamics of Small Systems*, Physics Today **58** 43 (2005)

[21] J. Liphardt, S. Dumont, S.B. Smith, I. Tinoco, C. Bustamante,*Equilibrium Information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality*, Science **296** 1932 (2002)

[22] A.M. Monge, M. Manosas, F. Ritort, *Experimental test of ensemble inequivalence and the fluctuation theorem in the force ensemble in DNA pulling experiments*, Phys. Rev. E **98** 032146 (2018)

[23] S. Ciliberto, *Experiments in Stochastic Thermodynamics: Short History and Perspectives*, Phys. Rev. X **7** 021051 (2017)

[24] G. Gallavotti, *Statistical Mechanics: A Short Treatise*, Springer, Berlin (1999)

[25] D.M. Carberry, J.C. Reid, G.M. Wang, E.M. Sevick, D.J. Searles, D.J. Evans, *Fluctuations and Irreversibility: An Experimental Demonstration of a Second-Law-Like Theorem Using a Colloidal Particle Held in an Optical Trap*, Phys. Rev. Lett. **92** 140601 (2004)

[26] Y. Murashita, K. Funo, M. Ueda, *Nonequilibrium equalities in absolutely irreversible processes* Phys. Rev. E **90** 042110 (2014)

[27] T.L. Hill, *Statistical Mechanics*, Dover (1987) (pp. 189-190)

[28] H.C. Andersen, *Molecular dynamics simulations at constant pressure and/or temperature*, J. Chem. Phys. **72** 2384 (1980)

[29] R.C. Lua, A.Y. Grosberg, *Practical Applicability of the Jarzynski Relation in Statistical Mechanics: A Pedagogical Example*, J. Phys. Chem. B **109** 6805 (2005)

[30] J.W. Gibbs, *Elementary Principles in Statistical Mechanics* (New York: Scribner) (1902). Reprinted 1981 (Woodbridge, CT: Ox Bow Press)

[31] C. Jarzynski, *What Is the Microscopic Response of a System Driven Far From Equilibrium?*, in P. Garbaczewski, R. Olkiewicz (eds), Dynamics of Dissipation. Lecture Notes in Physics, vol 597. Springer, Berlin (2002)

[32] G.E. Crooks, C. Jarzynski, *Work distribution for the adiabatic compression of a dilute and interacting classical gas*, Phys. Rev. E **75** 021116 (2007)

[33] C. Jarzynski, *Stochastic and Macroscopic Thermodynamics of Strongly Coupled Systems*, Phys. Rev. X **7** 011008 (2017)