Dynamic dependence of nonequilibrium work limits the validity of the Jarzynski equality

Amilcare Porporato and Salvatore Calabrese

1Department of Civil and Environmental Engineering and Princeton Environmental Institute, Princeton University, Princeton, NJ, USA.
1(AP) aporpora@princeton.edu; (SC) sc58@princeton.edu

November 18, 2019

Abstract

The Jarzynski equality (JE) is analyzed in regard to its validity for both quasi-static transformations in the thermodynamic limit and Hamiltonian evolutions of the work protocol. In the first case, we show that the JE holds for isothermal transformations only; in the second case, we show that the work done (linked to the final state Hamiltonian) depends on the temporal dynamics of the work protocol (including its speed), thus precluding the possibility of identifying it with the free energy or any other thermodynamic state function. Even in the case of thermodynamic limit and infinitesimally slow transformations of the Hamiltonian (adiabatic invariance) following states of thermodynamic equilibrium, the resulting work expression does not default to the JE, but to the work relation for adiabatic thermodynamic transformations, \( W = \Delta U \), where \( W \) is work and \( U \) is internal energy.

1 Introduction

The non-equilibrium work relation obtained more than twenty years ago by Jarzynski [1] has become a very popular result in modern statistical mechanics. It is often cited as an exact result, valid for systems driven arbitrarily away from equilibrium [2, 3]. The Jarzynski equality (JE) states that, for a system initially in thermodynamic equilibrium with a bath at temperature \( T \), the statistics of the work \( W \) done on a system obey the relation

\[
\langle e^{-\beta W} \rangle = e^{-\beta \Delta F},
\]

where the brackets \( \langle \cdot \rangle \) denote an average over the initial canonical ensemble with parameter \( \beta = 1/T \) and \( \Delta F \) is the difference in Helmholtz free energy, \( F = U - TS \), with \( U \) internal energy and \( S \) entropy. \( \Delta F \) is calculated between the initial state and a final state obtained by letting the system equilibrate again with the same bath after the work protocol has been completed (note that no work is done during this final equilibration). The protocol consists in varying a work parameter, \( \lambda(t) \), typically an extensive quantity of the system such as the volume of a piston, for a given duration \( \tau \) of the protocol, \( 0 \leq t \leq \tau \). Note that in the case of a Dirac delta distribution of the work \( W \) (e.g., macroscopic systems, in which the fluctuations are negligible) relation [1] simply becomes \( W = \Delta F \).

The JE has aroused considerable interest, especially for its promise as a means to estimate the thermodynamic properties of a system from work measurements with arbitrary protocols. Clearly, this would be especially relevant for small systems in which thermodynamic fluctuations are not negligible, including nanosystems (e.g., proteins) [4, 5]. On the other hand, the JE is also rather surprising, because it expresses general non-equilibrium work statistics as a function only of equilibrium quantities considered in isothermal conditions, a fact that is striking for its extremely broad applicability. Perhaps for this reason, it has attracted some criticism and debate [6, 7, 8, 9, 11].

13 Nov 2019
The JE has been derived using different approaches, following either a Hamiltonian evolution or a stochastic one. In the first approach, the system is isolated during the work protocol (i.e., energy can only be exchanged as work and the process is adiabatic – a part from the possible subsequent thermal equilibration), while in the stochastic approach the transformation is basically isothermal (at least in the overdamped case and for infinitely slow protocols [2]). Here we analyze the applicability of the JE considering the Hamiltonian approach, which has been the subject of most debate, in particular in relation to the definition of work [7]. For concreteness we focus also on the thermodynamic limit considering the case of the compression of a gas in a piston, where the work parameter is the volume \( V \), which is varied from an initial volume \( V_0 \) at temperature \( T \) to a final volume \( V_f \).

2 Thermodynamic Limit

Before delving into the details of the derivation, it is useful to recall some basic facts regarding quasi-static thermodynamic transformations (i.e., following a path of equilibrium states), for which the connection between work and thermodynamic potentials is well-known [11]. Indeed, if \( \Delta U \) applies independently of the size of the fluctuations and the speed of the protocol, it should also hold for quasi-static transformations of systems in the thermodynamic limit, the fluctuations of which are negligible. We focus here on the isothermal and the adiabatic cases. As is well known [12], the thermodynamic work (also called reversible mechanical work) \( W \) is obtained simply by integrating the \( pdV \) term in the Gibbs equation, or equivalently by computing the area under the respective compression curves in the \( pV \)-diagram (pressure-volume); see Fig. 1. Such a work is the minimum work required (or extractable) during a transformation between two equilibrium states and is governed by the maximum work theorem [11]. This reversible work can be calculated also by integrating the infinitesimal work, \( dW = dU - TdS \), along the sequence of equilibrium states. In the isothermal case, \( W = \Delta U - T\Delta S \), so that going from point A to B at the same temperature \( T \), the work is equal to

\[
W_{(i)} = \Delta F, \tag{2}
\]

just like in the deterministic counterpart of \( \Delta U \). For this transformation, an amount of heat equivalent to the work done on the system is simultaneously released to the bath, \( -Q_{(i)} = W_{(i)} \).

For the case of a quasi-static adiabatic compression, corresponding to the path going from A to C in Fig. 1a, no heat is released (\( Q = \int TdS = 0 \)) and the work done on the system is

\[
W_{(a)} = U_C - U_A = \Delta U. \tag{3}
\]

Because the adiabatic compressibility is smaller than the isothermal one ([11], page 210), the adiabatic work is always (i.e., for any substance) larger than the isothermal work (i.e., \( W_{(a)} - W_{(i)} > 0 \), and the final temperature is higher than the initial one, \( T_C > T \). If the adiabatically compressed system is left to equilibrate again to temperature \( T \) from \( T_C \), the extra amount of work also is released into the bath as heat. Alone, these simple considerations on the isothermal and adiabatic transformations (i.e., protocols) show that (i) the work done to bring the system from state A to B does depend on the specific protocol (e.g., isothermal or adiabatic) and (ii) only for an isothermal transformation does the work equal the Helmholtz free energy difference between states A and B.

In a general non-equilibrium compression performed in finite time, the system is brought away from equilibrium and the work done is larger than the work done in the corresponding ideal cases. The work done depends on the temporal dynamics of the protocol (the faster the piston is moved, the larger the work and the energy dissipation). If the system is kept isolated during the compression, the final temperature and pressure that are reached once the system has equilibrated internally are higher than the temperature and pressure at the point C in Fig. 1a (e.g., point D in the diagram). With reference to a fluid system, with increasingly faster protocols, the work and the subsequent dissipation both increase. This is clear when considering, for example, a fluid compressed in a piston: for slow compression the dissipation is due to viscous dissipation at low Reynolds numbers, while for faster protocols the dissipation is enhanced by turbulence and vorticity generation as well as by possible generation of shock waves, etc.
3 Hamiltonian Dynamics

The above observations on systems in the thermodynamic limit also should appear from the microscopic dynamics. We consider therefore the JE derivation based on Hamiltonian dynamics, following [14]. The system is assumed to consist of particles evolving according to Hamiltonian dynamics and the work is performed according to a given protocol, defined by a change in one of the Hamiltonian parameters, $\lambda(t)$. This implies that the system is isolated (no heat exchange) and that the evolution is deterministic (i.e., for given initial conditions and protocol evolution, Newton’s equations of motion lead to the same final state at the end of the protocol). Thus the statistics of the work originate only from choosing different initial conditions of the microstates, sampled from an initial canonical distribution, but not from any randomness – internal or external – during the evolution. We focus on the case where system and bath are disconnected during the protocol [14], because this allows us to make our points more directly; however, the conclusions also apply in the case in which they remain in contact (e.g., [15, 3]), as will be reported elsewhere.

We start from a system in a specific initial condition of microstates $x_0$, defining the initial value of the Hamiltonian $H_0$, and perform work on it by changing $\lambda$. The work is given by the conservation of energy (Eq. (10) of [14])

$$W = H_f - H_0.$$  \hspace{1cm} (4)

If we start from a condition of thermodynamic equilibrium, then the initial Hamiltonian can be identified with the internal energy of the system, $H_0 = U_0$. However, $H_f$, the energy at the end of the protocol, in general cannot be associated with the internal energy (or any equilibrium thermodynamic potential) of the system, because for protocols completed in finite times such an energy typically also contains kinetic components linked to organized motion (e.g., mean kinetic energy of the center of mass) and spatial inhomogeneities in the particle density.

The previous expression [4] can be written in terms of the rate of change of the Hamiltonian, following Landau and Lifshitz ([16, Eq. (40.5)] and Eq. (49.2)),

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda} \dot{\lambda},$$  \hspace{1cm} (5)

so that the work is

$$W = \int_0^\tau \frac{\partial H}{\partial \lambda} \dot{\lambda} dt.$$  \hspace{1cm} (6)
This expression for the work has been used in the Hamiltonian derivations of the JE (although this expression has been questioned by Rubi and Vilar [7] and subsequently defended by [8, 14] and [9, 17]). It is crucial to realize, however, that in general (and in contrast with e.g., Eq. (10) of [14]), the work $W$ does not depend only on the initial and final points of the protocol, but also on its temporal evolution. The reason for this is that $\frac{\partial H}{\partial \lambda}$ is not only a function of $\lambda$ but also of the configurational and momentum coordinates, as clearly stated in the classical text by Landau and Lifshitz. Quoting verbatim [10] on p. 155: “... the rate of change of the energy of the system is $\frac{dE}{dt} = \frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda} \frac{d\lambda}{dt}$. The expression on the right depends not only on the slowly varying quantity $\lambda$ but also on the rapidly varying quantities $q$ and $p$.”

This can be seen explicitly from a simple extension of the pedagogic case presented by Sokolov [13], consisting of a particle interacting with a moving piston; see Fig. 1b. The Hamiltonian of the system is $H = 1/2mv^2 + \Theta(x - x^w)1/2kx^2$, where $m$ is the particle mass, $v$ is its velocity, $x$ is its position, $x^w$ is the position of the wall, $k$ is its stiffness, and $\Theta(\cdot)$ is the Heaviside function. At time $t = 0$, the wall starts moving from position $x^w_0$ at a constant speed $v^w$, and the work at every instant in time is computed as $W(t) = H(t) - H_0$. The protocol ends at time $t = \tau$, when the wall reaches the final position $x^w_f$. As can be seen in Fig. 1b, for a given initial position and momentum of the particle (and hence Hamiltonian $H_0$), the Hamiltonian at the end of the protocol, $H_f = H(t = \tau)$, is determined by the rate of change of the protocol, hence by the time $\tau$ needed to move from $x^w_0$ to $x^w_f$. For a very slow motion of the piston ($\tau^* \gg 1$), the final Hamiltonian $H_f$ changes only slightly, because the particle has time to move far away from the piston, while for an infinitely fast protocol ($\tau^* \rightarrow 0$), the particle instantaneously acquires energy because it remains in proximity of the piston.

Going back to the derivation, when integrating the work done between $0$ and $\tau$, (5) produces different values of work depending on the temporal evolution (for example the speed) with which the protocol is followed, and not only on its initial ($\lambda(t = 0) = A$) and final value ($\lambda(t = \tau) = B$); this is because the specifics of the temporal evolution of the protocol also determine the evolution of $q$ and $p$ within the time varying term $\frac{\partial H}{\partial \lambda}$. For example, in the Eq. (10) of [14] the work is incorrectly expressed as a function of the initial and final conditions only (thus implying that nonequilibrium work is a state function), while instead $x_f$ depends also on the history of the protocol (i.e., $\lambda(t)$, $\dot{\lambda}(t)$) and not only on $x_0$.

In the case of a very slow variation of $\lambda$ (i.e., adiabatically in the sense of classical mechanics – see [10] p. 154 and [18] p. 549), and further assuming that the interactions among particles are such that the system follows a sequence of Hamiltonian equilibrium states, indeed one can identify the final Hamiltonian with the internal energy of the system at the end of the adiabatic work protocol, $H_f = U_C$. As a consequence,

$$W = H_f - H_0 = W_0 = \Delta U,$$

(7)

which agrees with the expression for the quasi-static adiabatic (in the sense of thermodynamics) compression of the piston (see previous equation (3)). Thus, with an infinitely slow protocol and simultaneous internal equilibration, which ensures compatibility with thermodynamics, the definition of work is equivalent to the one used in reversible thermodynamics [17], corresponding to the usual $pdV$ term in the Gibb's equation. As noticed by Peliti [17] this is compatible with the case described by Tolman (19 on page 542 of his treatise) of adiabatic transformation; however, this is so only for infinitesimally slow transformations with sufficiently chaotic interactions among particles and walls to ensure internal thermodynamic equilibrium.

Keeping in mind the previous observations, we can now proceed with the derivation considering the statistics of work (still following [14]). The work distribution comes about by assuming that the system is initially in thermodynamic equilibrium with a bath, so that the initial condition can be sampled randomly from a canonical distribution with inverse temperature $\beta$. Starting from different initial conditions, chosen at random, the deterministic protocol $\lambda(t)$ then produces different values of the work $W$. Multiplying the latter random variable by $\beta$, exponentiating, and averaging over the initial canonical distribution, one can write

$$\langle e^{-\beta W} \rangle = \int dx_0 \rho^{can}(x_0) e^{-\beta(H_f - H_0)} = \frac{1}{Z_A} \int dx_0 e^{-\beta H_f},$$

(8)

being $\rho^{can}(x_0) = \frac{1}{Z_A} e^{-\beta H_0}$. For a given protocol, because of the deterministic Hamiltonian dynamics, the initial condition is related one-to-one to the final condition. Therefore, the distribution inside the integral can be transformed using the fact that the phase-space volume is conserved during this transformation, so that

$$\langle e^{-\beta W} \rangle = \frac{1}{Z_A} \int dx_f e^{-\beta H_f(x_f)}.$$

(9)
At this point, if \( \int dx_f e^{-\beta H_f(x_f)} \) were equal to \( Z_B \) and recalling that \( \beta F = -\ln Z^{-1} \), one would then readily obtain the JE (1). This is not possible, however, for two reasons, both of which should already be clear from the previous discussion.

The first is related to the fact that, even if the protocol is infinitesimally slow and the system proceeds in a sequence of equilibrium states, the temperature at the end is higher than \( T = 1/\beta \) and equal to \( T_C = 1/\beta_C \) and \( H_f = U_C \), so that

\[
\langle e^{-\beta W} \rangle = \frac{1}{Z_A} \int dx_f e^{-\beta U_C},
\]

(10)
in which \( \int dx_f e^{-\beta U_C} \) is the partition function for neither the state C nor B. Additionally, if the system equilibrates with a bath at the original temperature \( T \), releasing heat, the work done is unchanged (and therefore the left hand side of the equation) but the end state has lower energy, \( H_e < H_f \), which invalidates the equation (moreover one might foresee difficulty in doing exact calculations due to the stochastic interactions with the thermal bath). One can recover formally the JE for a system in the thermodynamic limit (negligible fluctuations) if, after this equilibration, the work is corrected by the heat returned to the bath, \( W' = W - Q = \Delta F \) (W’ is the isothermal work \( W_{(i)} \), see Section 2).

The second and more important reason is that, in the case the protocol is performed in finite time, the system is driven away from equilibrium. Thus, at the end of the protocol, the particles have acquired organized motion (i.e., not all kinetic energy can be associated with the thermal component of internal energy) and are in an inhomogeneous spatial arrangement. As a result, thermodynamic quantities cannot be defined in a standard way, which means that in (11) the final value of the system energy \( H_f \) (which, we recall, is larger the faster the protocol) cannot be linked either to the internal energy or to the Helmholtz free energy. Only when the system is allowed to re-equilibrate, still remaining adiabatic, is then \( H_f \) equal to \( U_D \) (see Fig. 1b).

4 Conclusion

The fact that \( H_f \) depends on the temporal evolution of the protocol, while \( \Delta F \) does not, undermines the applicability of the JE for nonequilibrium transformations. This is linked to the fact that Helmholtz free energy is a thermodynamic state function, while the nonequilibrium work in general depends on the details of the process. We have explored these concerns considering transformations valid for systems in the thermodynamic limit and a step-by-step discussion of the Hamiltonian derivation of the JE. We are currently addressing the applicability of the JE and related work relations for a system in thermal contact with a bath during the protocol, in the context of both Hamiltonian dynamics [1] and stochastic thermodynamics [2] for small systems [20].

References

[1] Christopher Jarzynski. Nonequilibrium equality for free energy differences. Physical Review Letters, 78(14):2690, 1997.
[2] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. Reports on Progress in Physics, 75(12):126001, 2012.
[3] Christopher Jarzynski. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. Annu. Rev. Condens. Matter Phys., 2(1):329–351, 2011.
[4] Philip Nelson. Biological physics. WH Freeman New York, 2004.
[5] Kazushi Kinbara and Takuzo Aida. Toward intelligent molecular machines: directed motions of biological and artificial molecules and assemblies. Chemical reviews, 105(4):1377–1400, 2005.
[6] EGD Cohen and David Mauzerall. A note on the jarzynski equality. Journal of Statistical Mechanics: Theory and Experiment, 2004(07):P07006, 2004.
[7] Jose MG Vilar and J Miguel Rubi. Failure of the work-hamiltonian connection for free-energy calculations. *Physical review letters*, 100(2):020601, 2008.

[8] J Horowitz and C Jarzynski. Comment on failure of the work-hamiltonian connection for free-energy calculations. *Physical review letters*, 101(9):098901, 2008.

[9] Luca Peliti. Comment on failure of the work-hamiltonian connection for free-energy calculations. *Physical review letters*, 101(9):098903, 2008.

[10] Jose MG Vilar and J Miguel Rubi. Vilar and rubi reply. *Physical Review Letters*, 101(9):098902, 2008.

[11] H.B. Callen. *Thermodynamics and an introduction to thermostatistics*. Wiley India Pvt. Limited, 2006.

[12] J. Kestin. *A Course In Thermodynamics*. Number v. 1. Taylor & Francis, 1979.

[13] Igor M Sokolov. Statistical mechanics of entropic forces: disassembling a toy. *European Journal of Physics*, 31(6):1353, 2010.

[14] C Jarzynski. Nonequilibrium work relations: foundations and applications. *The European Physical Journal B*, 64(3-4):331–340, 2008.

[15] C Jarzynski. Nonequilibrium work theorem for a system strongly coupled to a thermal environment. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(09):P09005, 2004.

[16] LD Landau and EM Lifshitz. *Classical mechanics*. ed: Pergamon Press, Oxford, 1960.

[17] Luca Peliti. On the work–hamiltonian connection in manipulated systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(05):P05002, 2008.

[18] Herbert Goldstein, Charles Poole, and John Safko. *Classical mechanics*. AAPT, 2002.

[19] Richard Chace Tolman. *The principles of statistical mechanics*. Courier Corporation, 1979.

[20] Salvatore Calabrese, Lamberto Rondoni, and Amilcare Porporato. Thermodynamics of fluctuations in small systems interacting with the environment. *arXiv preprint arXiv:1909.09479*, 2019.