

Comparison of work fluctuation relations

To cite this article: Jordan Horowitz and Christopher Jarzynski *J. Stat. Mech.* (2007) P11002

View the [article online](#) for updates and enhancements.

You may also like

- [Nonequilibrium fluctuation theorems from equilibrium fluctuations](#)
Saar Rahav and Christopher Jarzynski
- [Rigorous results of nonequilibrium statistical physics and their experimental verification](#)
Lev P Pitaevskii
- [Nuclear-pumped lasers and physical problems in constructing a reactor-laser](#)
A V Karelin, A A Sinyanskii and Sergei I Yakovlenko

Comparison of work fluctuation relations

Jordan Horowitz¹ and Christopher Jarzynski²

¹ Department of Physics, University of Maryland, College Park, MD 20742, USA

² Department of Chemistry and Biochemistry, and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA
E-mail: horowitz@umd.edu and cjarzyns@umd.edu

Received 6 September 2007

Accepted 12 October 2007

Published 9 November 2007

Online at stacks.iop.org/JSTAT/2007/P11002
[doi:10.1088/1742-5468/2007/11/P11002](https://doi.org/10.1088/1742-5468/2007/11/P11002)

Abstract. We compare two predictions regarding the microscopic fluctuations of a system that is driven away from equilibrium: one due to Crooks (1998 *J. Stat. Phys.* **90** 1481) which has gained recent attention in the context of nonequilibrium work and fluctuation theorems, and an earlier, analogous result obtained by Bochkov and Kuzovlev (1977 *Zh. Eksp. Teor. Fiz.* **72** 238). Both results quantify irreversible behavior by comparing probabilities of observing particular microscopic trajectories during thermodynamic processes related by time-reversal, and both are expressed in terms of the *work* performed when driving the system away from equilibrium. By deriving these two predictions within a single, Hamiltonian framework, we clarify the precise relationship between them and discuss how the different definitions of work used by the two sets of authors give rise to different physical interpretations. We then obtain an extended fluctuation relation that contains both the Crooks and the Bochkov–Kuzovlev results as special cases.

Keywords: nonequilibrium fluctuations in small systems

Contents

1. Introduction	2
2. Derivations of earlier results	3
2.1. Setup	3
2.2. Crooks fluctuation relation	5
2.3. Bochkov–Kuzovlev fluctuation relation	7
3. Extended relation	8
3.1. Hamiltonian dynamics	9
3.2. Markovian stochastic dynamics	10
3.3. Discussion	13
4. Conclusion	13
Acknowledgment	13
References	14

1. Introduction

Recent interest in the nonequilibrium thermodynamics of small systems [1] has brought attention to a set of papers written by Bochkov and Kuzovlev in the late 1970s and early 1980s [2]–[5]. In view of evident similarities between results derived in these papers, and predictions obtained over the past decade or so by other authors (see in particular [6]–[10]), it is desirable to clarify the precise relationships between these two sets of results. In an earlier paper [11], one of us has compared the nonequilibrium work relation of [6], $\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}$, with an analogous equality found in [2]–[5], $\langle e^{-\beta W_0} \rangle = 1$. In the present paper we turn our attention to *fluctuation relations* for systems driven away from equilibrium. Specifically, we compare a result due to Crooks [8] ((1) below) with a similar result obtained by Bochkov and Kuzovlev (2).

Briefly, the setting is a system driven away from equilibrium by an externally applied, time-dependent force, X_t . We consider two different protocols (or schedules) for applying this force, X_t^F and X_t^R , related by time-reversal (5), where the superscripts denote ‘forward’ and ‘reverse’. The microscopic evolution of the system in either case is described by a phase space trajectory, γ . Under appropriate assumptions regarding this evolution, Crooks has shown that the probability to observe a given trajectory γ^F during the forward process, and the probability to observe its time-reversed counterpart γ^R during the reverse process, are related by [8]

$$\frac{P_F[\gamma^F]}{P_R[\gamma^R]} = e^{\beta(W - \Delta F)}. \quad (1)$$

Here, W denotes the work performed on the system and ΔF is a free energy difference between the equilibrium ensembles from which initial conditions are sampled, as described in detail below. From this result others are easily derived, including a fluctuation theorem

for the ratio of forward and reverse probability distributions of work values, $\rho_F(W)$ and $\rho_R(W)$ [9], which has recently been confirmed experimentally [12].

In a similar context, Bochkov and Kuzovlev have made an analogous assertion (equation (2.11) of [4] or equation (7) of [3]), which in the notation of the present paper is

$$\frac{P_F^0[\gamma^F]}{P_R^0[\gamma^R]} = e^{\beta W_0}, \quad (2)$$

where W_0 again denotes work.

As the above notation suggests, the two definitions of work, W and W_0 , differ, as do the probability distributions $P_{F/R}$ and $P_{F/R}^0$. Moreover, ΔF appears in (1), but not in (2). To establish the relationship between these two predictions, we first derive them both within a Hamiltonian framework, in section 2. Following that, in section 3 we derive a more general result, containing both (1) and (2) as special cases. This extended relation is obtained for both Hamiltonian dynamics and Markovian stochastic dynamics.

Independently, Seifert [13] has concluded that both (1) and (2) are special cases of a more general result, obtained within a theoretical framework that defines entropy production in out-of-equilibrium systems [14]. Our approach complements Seifert's, by focusing on the different physical interpretations of work appearing in (1) and (2).

2. Derivations of earlier results

In this section we derive (1) and (2) within a single framework, using Hamilton's equations to model the microscopic evolution of the system of interest. We introduce and specify the notation for this framework in section 2.1, and then in sections 2.2 and 2.3 we obtain the desired results. These side-by-side derivations clarify the physical interpretation of the differences between (1) and (2), as highlighted in a brief summary at the end of this section. We note that Crooks originally obtained (1) by modeling the evolution of the system as a stochastic Markov process. The Hamiltonian derivation of this result presented here is similar to the analyses of [15, 16].

2.1. Setup

Consider a classical system with N degrees of freedom, described by coordinates $\mathbf{q} = (q_1, \dots, q_N)$ and conjugate momenta $\mathbf{p} = (p_1, \dots, p_N)$, and let $z = (\mathbf{q}, \mathbf{p})$ denote a point in its phase space. We will be interested in the evolution of this system, in the presence of an externally controlled, time-dependent force X_t . In this section we model this evolution using Hamilton's equations, assuming a time-dependent Hamiltonian of the form

$$H(z; X_t) = H_0(z) - X_t \alpha(z). \quad (3)$$

Here α is the coordinate conjugate to the external force X , and the 'bare' Hamiltonian H_0 denotes the energy of the system in the absence of this force. For simplicity, we further assume that H is time-reversal-invariant for any fixed value of the external force, that is,

$$H(z^*; X) = H(z; X), \quad (4)$$

where the asterisk denotes a reversal of momenta, $\mathbf{p} \rightarrow -\mathbf{p}$.

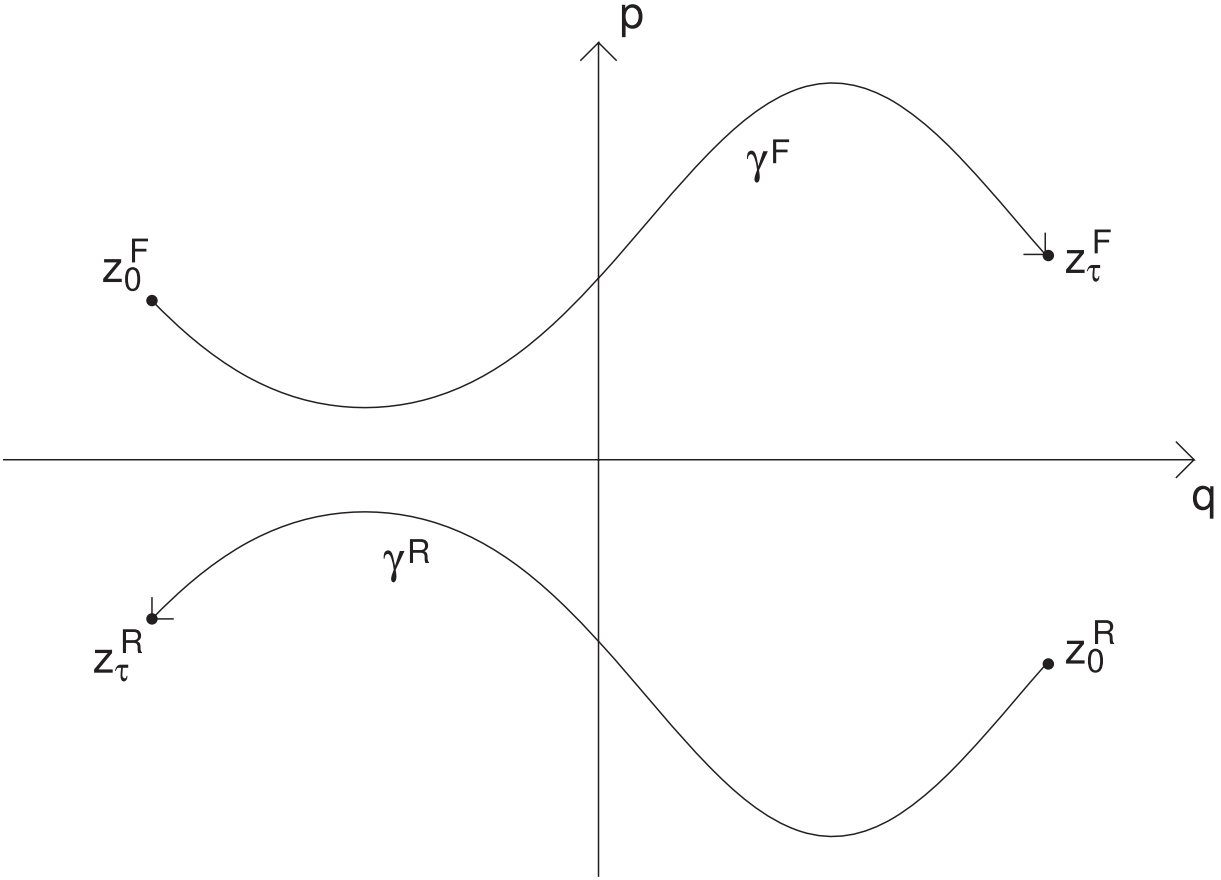


Figure 1. Schematic depiction of a forward trajectory γ^F and its conjugate twin γ^R . The two are related by a reversal of momenta and the direction of time, i.e. $z_t^R = z_{\tau-t}^{F*}$ for $0 \leq t \leq \tau$.

We have assumed that X appears *linearly* in the definition of the Hamiltonian, $H = H_0 - X\alpha$, as in the papers by Bochkov and Kuzovlev [2]–[5]. As mentioned by those authors, this assumption can be relaxed to allow for nonlinear coupling $H(z; X)$, as in [6]–[10]. However, for clarity of exposition, we follow [2]–[5] and stay with the linear assumption throughout this paper.

We will use the term *process* to denote the following sequence of events. First, the system is prepared in a state of thermal equilibrium at inverse temperature β . (This step will be discussed in more detail below.) Subsequently, from $t = 0$ to τ the system evolves in time under Hamilton's equations, as the external force is applied according to a schedule, or *protocol*, X_t . We will use the notation z_t to specify the phase space coordinates of the system at a particular time t during this interval, and γ to denote the entire trajectory from $t = 0$ to τ . Such a trajectory represents a single *realization*—i.e. one possible microscopic history—of the process in question. By performing this process repeatedly—always preparing the system in equilibrium, and always applying the same protocol X_t —we effectively generate independent samples $\gamma_1, \gamma_2, \dots$ from a statistical ensemble of possible realizations of the process.

We will consider two different processes, which we label *forward* (F) and *reverse* (R). During the forward process the force is switched from $X_0^F = A$ to $X_\tau^F = B$ using the

forward protocol X_t^F , while during the reverse process the force is switched from B back to A , using the time-reversed protocol, i.e.

$$X_t^R = X_{\tau-t}^F. \quad (5)$$

Our assumption of time-reversal invariance then implies that solutions of Hamilton's equations come in *conjugate pairs* [15, 16]: for every trajectory $\gamma^F = \{z_t^F\}$ that represents a microscopic realization of the forward process, the conjugate twin $\gamma^R = \{z_t^R\}$ is a realization of the reverse process, where

$$z_t^R = z_{\tau-t}^{F*}, \quad (6)$$

as illustrated in figure 1. While we restrict our analysis to Hamiltonian dynamics, we note that conjugate pairing of trajectories holds also for non-Hamiltonian time-reversible deterministic dynamics, as exploited by Evans to derive nonequilibrium work relations for deterministically thermostatted systems [17].

In the derivations that follow, we will make use of two different definitions of the *work* performed on the system during such a process. The first definition, which appears in the papers by Bochkov and Kuzovlev (e.g. equation (2.9) of [4]), is the integral of force versus displacement familiar from introductory mechanics texts [18]:

$$W_0 \equiv - \int_0^\tau \dot{\alpha} \frac{\partial H}{\partial \alpha} dt = \int_0^\tau X \dot{\alpha} dt = H_0(z_\tau) - H_0(z_0). \quad (7)$$

By contrast, recent papers (see e.g. [6, 8]) have used a different definition:

$$W \equiv \int_0^\tau \dot{X} \frac{\partial H}{\partial X} dt = - \int_0^\tau \dot{X} \alpha dt = H(z_\tau; X_\tau) - H(z_0; X_0) \quad (8)$$

which traces its origins to discussions of the statistical foundations of thermodynamics [19]–[21]. In both (7) and (8) the properties of Hamilton's equations have been invoked to rewrite the integral as a net change in the value of a Hamiltonian, either H_0 or H (see [11]). At the end of this section we briefly discuss the distinction between these two definitions of work.

2.2. Crooks fluctuation relation

For the situation that we have just described, we now derive (1). In this equation, $P_F[\gamma^F]$ denotes the probability distribution to observe a trajectory γ^F , when performing the forward process. This distribution must be defined with respect to a *measure* on the space of trajectories γ^F . Since Hamiltonian evolution is deterministic, it is natural to use the Liouville phase space measure, applied to the initial conditions of the trajectory. Thus, if we imagine a narrow ‘tube’ of Hamiltonian trajectories evolving from $t = 0$ to τ , from a small patch of initial conditions in phase space, then we equate the path-space measure of this tube of trajectories with the Liouville measure of the patch of initial conditions: $d\gamma^F = dz_0^F = d^N \mathbf{q}_0 d^N \mathbf{p}_0$. Similarly, for the reverse process the distribution $P_R[\gamma^R]$ is defined with respect to the measure $d\gamma^R = dz_0^R$. Because Hamiltonian evolution preserves phase space volume [22], the measure of a given set of forward trajectories is the same as that of the conjugate set of reverse trajectories:

$$d\gamma^F = d\gamma^R. \quad (9)$$

In Crooks' formulation [8], the initial conditions for the forward process are sampled from the equilibrium (canonical) distribution corresponding to the initial value of the external force, $X_0^F = A$:

$$P_F[\gamma^F] = p_A^{\text{eq}}(z_0^F) = \frac{1}{Z(A)} \exp[-\beta H(z_0^F; A)]. \quad (10)$$

Similarly for the reverse process,

$$P_R[\gamma^R] = p_B^{\text{eq}}(z_0^R) = \frac{1}{Z(B)} \exp[-\beta H(z_0^R; B)]. \quad (11)$$

In these equations

$$Z(X) = \int dz e^{-\beta H(z; X)} \quad (12)$$

is the classical partition function associated with the canonical distribution at a fixed value of X ; the corresponding free energy is given by the expression

$$F(X) = -\beta^{-1} \ln Z(X). \quad (13)$$

Combining (10), (11) and (13) we get, for a conjugate pair of trajectories γ^F and γ^R ,

$$\frac{P_F[\gamma^F]}{P_R[\gamma^R]} = e^{-\beta \Delta F} e^{\beta [H(z_0^R; B) - H(z_0^F; A)]}, \quad (14)$$

where $\Delta F \equiv F(B) - F(A)$. We now use (4), (6) and (8) to rewrite the quantity inside the last exponent above:

$$\begin{aligned} H(z_0^R; B) - H(z_0^F; A) &= H(z_\tau^{F*}; B) - H(z_0^F; A) \\ &= H(z_\tau^F; B) - H(z_0^F; A) \\ &= W, \end{aligned} \quad (15)$$

where W represents the work (as defined by (8)) performed on the system during the realization γ^F . Thus we finally obtain

$$\frac{P_F[\gamma^F]}{P_R[\gamma^R]} = e^{\beta(W - \Delta F)}. \quad (16)$$

This result was first derived by Crooks [8], who modeled the evolution of the system as a stochastic discrete-time Markov process, rather than using Hamiltonian dynamics. One consequence of (16) is the Crooks fluctuation theorem for the ratio of forward and reverse work distributions [9]:

$$\frac{\rho_F(W)}{\rho_R(-W)} = e^{\beta(W - \Delta F)}. \quad (17)$$

Often (16) and (17) are expressed in terms of the dissipated work, $W_d = W - \Delta F$ [23].

2.3. Bochkov–Kuzovlev fluctuation relation

In deriving (16), we imagined that initial conditions were sampled from the canonical distributions associated with the Hamiltonians $H(z; A)$ and $H(z; B)$, for the forward and reverse processes, respectively. Now let us consider a different situation, in which—for both processes—initial conditions are sampled from the canonical distributions corresponding to the bare Hamiltonian $H_0(z)$. We suppose, however, that the protocols X_t^F and X_t^R are the same as above, namely from A to B and from B to A , where these values remain arbitrary. If we use $P_F^0[\gamma^F]$ and $P_R^0[\gamma^R]$ to denote the corresponding distributions of trajectories, then we have

$$P_F^0[\gamma^F] = \frac{1}{Z(0)} \exp[-\beta H_0(z_0^F)], \quad (18)$$

$$P_R^0[\gamma^R] = \frac{1}{Z(0)} \exp[-\beta H_0(z_0^R)] \quad (19)$$

in place of (10) and (11). Taking the ratio then gives us

$$\frac{P_F^0[\gamma^F]}{P_R^0[\gamma^R]} = e^{\beta[H_0(z_0^R) - H_0(z_0^F)]}. \quad (20)$$

By analogy with (15), we have

$$H_0(z_0^R) - H_0(z_0^F) = H_0(z_\tau^F) - H_0(z_0^F) = W_0, \quad (21)$$

where W_0 is the work (as defined by (7)) performed during the realization γ^F . Thus we finally arrive at

$$\frac{P_F^0[\gamma^F]}{P_R^0[\gamma^R]} = e^{\beta W_0}, \quad (22)$$

which corresponds to equation (2.11) of [4].

Related to (22) is a fluctuation theorem for the work distributions (cf (17))

$$\frac{\rho_F^0(W_0)}{\rho_R^0(-W_0)} = e^{\beta W_0}, \quad (23)$$

where $\rho_{F/R}^0(W_0)$ are the work distributions associated with the trajectory distributions in (18) and (19). Note that the quantity W_0 that appears in (22) and (23) is, in general, different from the dissipated work W_d .

As highlighted by these derivations, a principal difference between (16) and (22) (equivalently, (1) and (2)) is the choice of the equilibrium distributions from which initial conditions are drawn. If we sample from canonical distributions corresponding to the Hamiltonians $H(A)$ and $H(B)$, then we obtain (16), while if we sample initial conditions from the bare canonical distribution, corresponding to H_0 , then we arrive at (22).

It is useful to spell out the physical interpretation associated with this distinction. In Crooks' formulation (16), we imagine that prior to $t = 0$ we allow the system to come to equilibrium with a thermal reservoir at inverse temperature β , while holding the external force fixed at either A or B . We then let the system evolve as the external force is manipulated according to the forward or reverse protocol. To obtain Bochkov and Kuzovlev's result (22), we instead imagine that the external force is absent ($X = 0$) as

the system equilibrates with the reservoir. The system thus settles into an equilibrium state corresponding to the bare Hamiltonian H_0 . Just prior to $t = 0$ we suddenly turn the force on to A or B , subsequently switching it continuously according to the forward or reverse protocol.

The two results involve not only different distributions of initial conditions, but also different definitions of work, W_0 and W . The right sides of (7) and (8) suggest a physical interpretation of this distinction. Namely, in one case we define the *internal energy* of the system to be given by the value of the bare Hamiltonian H_0 ; then W_0 is viewed as the work performed by the application of an external force X that affects the evolution of the system in a fixed energy landscape H_0 . In the other case, the system's internal energy is taken to be the full Hamiltonian H ; we then view W as the work associated with externally controlled manipulations of the energy landscape $H = H_0 - X\alpha$. See [11] for a more detailed discussion, and illustration, of this distinction.

In the special case of cyclic processes, defined by the condition $X_0 = X_\tau$, we have $\Delta F = 0$ and $W = W_0$ [11]. In this situation (16) and (22) are equivalent. If we further restrict ourselves to time symmetric protocols ($X_t = X_{\tau-t}$), then there is no distinction between the forward and reverse processes. In this case, (16) and (22) can be viewed as special cases of a more general result derived by Evans and Searles (see equation (4.15) of [24], as well as the recent review by Sevick *et al* [25]).

3. Extended relation

In section 2 we used Hamiltonian dynamics to model the evolution of the system. By contrast, as noted above, Crooks modeled this evolution as a discrete-time stochastic Markov process to represent a system in contact with a thermal reservoir [8]. If we were to carry out an analysis similar to that of section 2, but in a Markovian rather than a Hamiltonian framework, then we would arrive at the same conclusion, namely that both (1) and (2) are valid, the difference between them relating to the choice of canonical ensemble from which to sample initial conditions. Rather than showing this explicitly, in this section we derive a more general result, (36) below, that contains both (1) and (2) as special cases. We will present both Hamiltonian and Markovian derivations of this result.

Consider a Hamiltonian of the form

$$H(z; \lambda_t, X_t) = H_0(z; \lambda_t) - X_t \alpha(z), \quad (24)$$

where now H_0 itself is parameterized by a variable λ . Thus the full Hamiltonian H is made time-dependent by manipulating both the force X , and a *work parameter*, λ . From a mathematical perspective, this division of the parameters into an external force and an internal work parameter is somewhat arbitrary, particularly if λ appears linearly in H_0 (as in (49) below). While in specific situations the distinction may be motivated by physical considerations or by the questions being addressed, our aim here is to develop a general formalism that encompasses both (1) and (2). For this purpose, the form of H given in (24) is useful.

For the bare Hamiltonian H_0 , we have a parameter-dependent canonical distribution

$$\tilde{p}_0^{\text{eq}}(z; \lambda) = \frac{1}{\tilde{Z}(\lambda)} e^{-\beta H_0(z; \lambda)}, \quad (25)$$

for which the partition function and free energy are given by

$$\tilde{Z}(\lambda) = \int dz e^{-\beta H_0(z;\lambda)} \quad \text{and} \quad \tilde{F}(\lambda) = -\beta^{-1} \ln \tilde{Z}(\lambda). \quad (26)$$

Throughout this section we will take the internal energy of the system to be defined by the value of the bare Hamiltonian H_0 , and we will consider processes during which the system evolves in time as both λ and X are switched according to pre-determined protocols. Thus we picture a system that evolves in a time-dependent energy landscape H_0 , while also coupled to a time-dependent external force X . For a given realization of such a process, we take the work performed on the system to be given by

$$\tilde{W} = \int_0^\tau \dot{\lambda} \frac{\partial H}{\partial \lambda} dt + \int_0^\tau X \dot{\alpha} dt, \quad (27)$$

or by the discrete-time analog of this expression, (39) below. This is a hybrid definition whose two terms are similar to the expressions for W and W_0 introduced earlier. As we will see, for processes during which the system is thermodynamically isolated the value of \tilde{W} is equal to the net change in H_0 , (33); whereas, if the system remains in contact with a reservoir as λ and X are varied, then the net change in H_0 is equal to the sum of \tilde{W} and a quantity Q that represents the heat absorbed by the system, (42). In either case, (27) provides a definition of work that is faithful to the first law of thermodynamics and consistent with our definition of internal energy, H_0 .

As before, we will compare two processes (F , R), characterized by conjugate protocols for the externally controlled parameters:

$$X_t^R = X_{\tau-t}^F, \quad \lambda_t^R = \lambda_{\tau-t}^F. \quad (28)$$

For the forward process, we let A and B denote the initial and final values of X , as before, and we let 0 and 1 denote the initial and final values of λ . For the reverse process, then, X is switched from B to A , and λ from 1 to 0. We will assume that initial conditions for the forward and reverse processes are sampled from $p_0^{\text{eq}}(z; \lambda = 0)$ and $p_0^{\text{eq}}(z; \lambda = 1)$, respectively (see (25)). The physical interpretation of this assumption is similar to that discussed in the context of the Bochkov–Kuzovlev results, near the end of section 2. Namely, we imagine that the external force is absent ($X = 0$) as the system equilibrates with the reservoir prior to time $t = 0$. The work parameter is held fixed at either $\lambda = 0$ (for the forward process) or $\lambda = 1$ (for the reverse process) during this equilibration stage. Then, immediately prior to $t = 0$ the external force is turned on to either A or B , and subsequently both λ and X are varied according to the appropriate protocol.

3.1. Hamiltonian dynamics

With the set-up just described, we now suppose that the microscopic history of the system during the time interval $0 \leq t \leq \tau$ (for either process) is described by a Hamiltonian trajectory z_t . Such an evolution has the property that the total rate of change of the Hamiltonian is equal to the partial derivative of the Hamiltonian function with respect

to time (see e.g. equations (8)–(35) of [22]). Thus

$$\frac{d}{dt}H(z_t; \lambda_t, X_t) = \dot{\lambda}_t \frac{\partial H}{\partial \lambda} + \dot{X}_t \frac{\partial H}{\partial X} \quad (29)$$

$$= \dot{\lambda}_t \frac{\partial H_0}{\partial \lambda} - \dot{X} \alpha \quad (30)$$

$$= \dot{\lambda}_t \frac{\partial H_0}{\partial \lambda} + X \dot{\alpha} - \frac{d}{dt}(X \alpha). \quad (31)$$

Adding $(d/dt)(X \alpha)$ to both sides, we get

$$\frac{d}{dt}H_0 = \dot{\lambda}_t \frac{\partial H_0}{\partial \lambda} + X \dot{\alpha}. \quad (32)$$

The first two terms on the right are the integrands appearing in our definition of \tilde{W} (27). Thus, integrating both sides of (32) with respect to time, we see that \tilde{W} is the net change in the value of the bare Hamiltonian:

$$\tilde{W} = H_0(z_\tau; \lambda_\tau) - H_0(z_0; \lambda_0). \quad (33)$$

This identity will now be put to use much as (7) and (8) were in section 2.

Adopting notation similar to that of section 2, we let $\tilde{P}_F[\gamma^F]$ and $\tilde{P}_R[\gamma^R]$ denote the probability distributions of realizations of the forward and reverse processes; as before, we use the Liouville measure of initial conditions to define a measure on the space of possible trajectories, for both processes. We then have, by analogy with (10) and (11),

$$\tilde{P}_F[\gamma^F] = \tilde{p}_0^{\text{eq}}(z_0^F; 0) = \frac{1}{\tilde{Z}(0)} \exp[-\beta H_0(z_0^F; 0)] \quad (34)$$

$$\tilde{P}_R[\gamma^R] = \tilde{p}_0^{\text{eq}}(z_0^R; 1) = \frac{1}{\tilde{Z}(1)} \exp[-\beta H_0(z_0^R; 1)], \quad (35)$$

since initial conditions are sampled from the canonical ensembles associated with H_0 . Taking the ratio of these two distributions for a conjugate pair of trajectories, and making use of (26) and (33), we obtain

$$\frac{\tilde{P}_F[\gamma^F]}{\tilde{P}_R[\gamma^R]} = e^{\beta(\tilde{W} - \Delta \tilde{F})}, \quad (36)$$

where $\Delta \tilde{F} = \tilde{F}(1) - \tilde{F}(0)$.

3.2. Markovian stochastic dynamics

We now derive the same result, (36), for a system that remains in weak contact with a thermal reservoir during the interval $0 \leq t \leq \tau$, and whose evolution in this situation is modeled as a discrete-time Markov process. Specifically, we imagine that the microscopic history of the system is described by a sequence z_0, z_1, \dots, z_N , representing phase space points visited at times t_0, t_1, \dots, t_N , with $t_n = n\tau/N$. With each time increment, both the phase space point z and the parameters $\pi \equiv (\lambda, X)$ are updated with a combination of a stochastic step and a switching step, as follows.

Stochastic step. During a stochastic step, starting from, say, z_n at time t_n , the system makes a random jump to a new point in phase space, z_{n+1} , according to the transition probability $P(z_n \rightarrow z_{n+1}; \lambda, X)$, where (λ, X) are the fixed values of the external parameters during this step. This transition probability is assumed to satisfy detailed balance (a requirement that may be reduced, at the expense of using a modified dynamics for the reverse process [23]), expressed as

$$\frac{P(z_n \rightarrow z_{n+1}; \lambda, X)}{P(z_n \leftarrow z_{n+1}; \lambda, X)} = \frac{e^{-\beta H(z_{n+1}; \lambda, X)}}{e^{-\beta H(z_n; \lambda, X)}}. \quad (37)$$

The direction of the arrow indicates the sense of the transition, e.g. from z_{n+1} to z_n in the denominator. Since each stochastic step, $z_n \rightarrow z_{n+1}$, is meant to represent effects of the interaction with the thermal reservoir, any change in the system's energy during such a step will be interpreted as heat exchanged with the reservoir; see (38) below.

Detailed balance (37) represents an assumption about the dynamics of the system. Among other things it implies that the equilibrium distribution is conserved when the parameters λ and X are held fixed. Since the equilibrium distribution is determined by all the physical forces acting on the system, both those that we choose to view as internal and those that we treat as external, detailed balance is expressed in terms of the full Hamiltonian, H , rather than H_0 .

Switching step. During a switching step the external parameters are updated, e.g. from $\pi_n = (\lambda_n, X_n)$ to $\pi_{n+1} = (\lambda_{n+1}, X_{n+1})$, according to a given protocol, with the phase space point held fixed.

The evolution of the system from $t_0 = 0$ to $t_N = \tau$ consists of an alternating sequence of stochastic and switching steps. For the forward process we assume that the switching step comes first, giving us a progression of the form $(z_n^F, \pi_n^F) \rightarrow (z_n^F, \pi_{n+1}^F) \rightarrow (z_{n+1}^F, \pi_{n+1}^F)$ at each time step. For the reverse process, the stochastic steps are $(z_n^R, \pi_n^R) \rightarrow (z_{n+1}^R, \pi_n^R) \rightarrow (z_{n+1}^R, \pi_{n+1}^R)$.

We now define the heat absorbed by the system from the reservoir, for a given realization of the forward process, as a sum of energy changes during the stochastic steps:

$$Q = \sum_{n=0}^{N-1} \delta Q_n = \sum_{n=0}^{N-1} [H(z_{n+1}; \pi_{n+1}) - H(z_n; \pi_{n+1})]. \quad (38)$$

For the same realization, the work performed on the system is defined as the discrete-time version of (27):

$$\begin{aligned} \tilde{W} &= \sum_{n=0}^{N-1} \delta \tilde{W}_n \\ &= \sum_{n=0}^{N-1} [H_0(z_n; \lambda_{n+1}) - H_0(z_n; \lambda_n)] + \sum_{n=0}^{N-1} [(\alpha_{n+1} - \alpha_n) X_{n+1}], \end{aligned} \quad (39)$$

introducing the convenient shorthand $\alpha_n \equiv \alpha(z_n)$. On the right side of (39), the first sum represents energy changes due to the switching of λ , while the second sum is associated with the force X acting on the coordinate α . From these definitions, the identity

$$\begin{aligned} \delta H_n &\equiv H(z_{n+1}; \pi_{n+1}) - H(z_n; \pi_n) \\ &= \delta Q_n + \delta \tilde{W}_n + \alpha_n X_n - \alpha_{n+1} X_{n+1} \end{aligned} \quad (40)$$

is easily verified, by inspection. Combining (38)–(40) now gives us the following expressions for the net changes in H and H_0 , from $t = 0$ and τ :

$$H(z_N; \pi_N) - H(z_0; \pi_0) = Q + \tilde{W} + X_0 \alpha_0 - X_N \alpha_N \quad (41)$$

$$H_0(z_N; \lambda_N) - H_0(z_0; \lambda_0) = Q + \tilde{W}. \quad (42)$$

The latter is analogous to (33), but now includes heat exchange with the reservoir.

With these preliminaries out of the way, we analyze the probability distributions of trajectories, for the forward and reverse processes. A trajectory is specified as a sequence of $N + 1$ points in phase space, e.g. $\gamma = (z_0, z_1, \dots, z_N)$. A natural measure on the space of such trajectories is given by $d\gamma = dz_0 dz_1 \dots dz_N$. The probability of obtaining a particular trajectory, $\tilde{P}[\gamma]$, is equal to the probability of sampling its initial conditions z_0 from the appropriate equilibrium distribution, multiplied by the conditional probability of generating the subsequent sequence of steps, $\tilde{P}[\gamma|z_0]$. Thus for the forward process we have

$$\tilde{P}_F[\gamma^F] = \tilde{p}_0^{\text{eq}}(z_0^F; \lambda_0^F = 0) \cdot \tilde{P}_F[\gamma^F|z_0^F] \quad (43)$$

$$\tilde{P}_F[\gamma^F|z_0^F] = P(z_0^F \rightarrow z_1^F; \pi_1^F) \cdots P(z_{N-1}^F \rightarrow z_N^F; \pi_N^F). \quad (44)$$

For the reverse process we have analogous equations, except that $\lambda_0^R = 1$, and the transition probabilities are of the form $P(z_n^R \rightarrow z_{n+1}^R; \pi_n^R)$, reflecting the convention that the stochastic step precedes the switching step, during each time interval of the reverse process. In writing the conditional probability as the product of transition rates in (44), we have made use of the assumed Markov nature of the dynamics.

For a conjugate pair of trajectories γ^F and γ^R , the ratio of conditional probabilities is

$$\frac{\tilde{P}_F[\gamma^F|z_0^F]}{\tilde{P}_R[\gamma^R|z_0^R]} = \frac{P(z_0^F \rightarrow z_1^F; \pi_1^F) \cdots P(z_{N-1}^F \rightarrow z_N^F; \pi_N^F)}{P(z_0^F \leftarrow z_1^F; \pi_1^F) \cdots P(z_{N-1}^F \leftarrow z_N^F; \pi_N^F)}, \quad (45)$$

where we have used the relations $z_n^R = z_{N-n}^F$ and $\pi_n^R = \pi_{N-n}^F$ to rewrite $P_R[\gamma^R|z_0^R]$ in terms of z_n^F 's and π_n^F 's. Following Crooks [8], we now use detailed balance to express the right side of (45) as a product of factors $\exp(-\beta \delta Q_n)$ (see (37) and (38)), which gives us

$$\frac{\tilde{P}_F[\gamma^F|z_0^F]}{\tilde{P}_R[\gamma^R|z_0^R]} = e^{-\beta Q}. \quad (46)$$

To get the ratio of unconditional probabilities $P_F[\gamma^F]/P_R[\gamma^R]$, we multiply the numerator and denominator of both sides by the appropriate initial distributions, as per (43):

$$\frac{\tilde{P}_F[\gamma^F]}{\tilde{P}_R[\gamma^R]} = \frac{\tilde{p}_0^{\text{eq}}(z_0^F; 0)}{\tilde{p}_0^{\text{eq}}(z_0^R; 1)} \cdot e^{-\beta Q} = e^{\beta[H_0(z_N^F; 1) - H_0(z_0^F; 0) - \Delta \tilde{F} - Q]}, \quad (47)$$

using $z_0^R = z_N^F$. With (42) this becomes

$$\frac{\tilde{P}_F[\gamma^F]}{\tilde{P}_R[\gamma^R]} = e^{\beta(\tilde{W} - \Delta \tilde{F})}, \quad (48)$$

as in the Hamiltonian case (36).

3.3. Discussion

We now argue that (1) and (2) can be viewed as special cases of (36). To establish this, let us suppose that the parameter λ appears linearly in H_0 , so that the full time-dependent Hamiltonian takes the form

$$H(z; \lambda_t, X_t) = \mathcal{H}_0(z) - \lambda_t \eta(z) - X_t \alpha(z). \quad (49)$$

It is natural to view both λ and X as forces, coupled to the coordinates η and α , respectively. However, in keeping with the convention adopted throughout section 3, we treat the term $-\lambda\eta$ as a contribution to the *internal* energy of the system, $H_0 = \mathcal{H}_0 - \lambda\eta$, and X as an *external* force. The work, (27), is now given by

$$\tilde{W} = - \int_0^\tau \dot{\lambda} \eta dt + \int_0^\tau X \dot{\alpha} dt, \quad (50)$$

or by the corresponding discrete-time analog obtained from (39). With these definitions (36) becomes, explicitly,

$$\frac{\tilde{P}_F[\gamma^F]}{\tilde{P}_R[\gamma^R]} = \exp \left[\beta \left(- \int_0^\tau \dot{\lambda} \eta dt + \int_0^\tau X \dot{\alpha} dt - \Delta \tilde{F} \right) \right]. \quad (51)$$

As before, initial conditions are sampled from the canonical ensemble $\tilde{p}_0^{\text{eq}} \propto e^{-\beta H_0}$, associated with the initial value of λ ($\lambda_0^F = 0$ and $\lambda_0^R = 1$).

If we now consider the special case in which $\lambda = 0$ during the entire process (hence $\tilde{p}_0^{\text{eq}} \propto e^{-\beta \mathcal{H}_0}$ and $\Delta \tilde{F} = 0$), then the right side of (51) becomes $\exp(\beta \int X \dot{\alpha} dt)$ and we recover (2). If we instead imagine that $X = 0$ during the entire process, then the right side of (51) becomes $\exp[\beta(-\int \dot{\lambda} \eta dt - \Delta \tilde{F})]$. We then recover (1), with λ and $\eta(z)$ playing the roles assigned to X and $\alpha(z)$ in section 2.2.

To summarize, if we view (1) as a result that applies when a system evolves in a time-dependent energy landscape, and (2) as a result that applies when a system evolves under the influence of a time-dependent external force (as per the discussion at the end of section 2), then (36) is a generalization that applies when both factors are present.

4. Conclusion

Our central aim has been to clarify the relationship between the fluctuation relation obtained by Crooks, (1), and that due to Bochkov and Kuzovlev, (2). These two results share a similar structure, but are not equivalent. The distinction between them is traced to the different definitions of work (W , W_0) used by the authors. One consequence of this difference is that (1) involves (and can be used to estimate) a free energy difference, ΔF , while (2) does not. Moreover, the two results are associated with different definitions of the internal energy of the system, and rely on different assumptions regarding the initial equilibrium preparation of the system. Finally, we have obtained a generalized fluctuation relation (36) that unites (1) and (2).

Acknowledgment

We gratefully acknowledge financial support provided by the University of Maryland.

References

- [1] Bustamante C, Liphardt J and Ritort F, 2005 *Phys. Today* **58** 43
- [2] Bochkov G N and Kuzovlev Y E, 1977 *Zh. Eksp. Teor. Fiz.* **72** 238
- [3] Bochkov G N and Kuzovlev Y E, 1977 *Sov. Phys. JETP* **45** 125
- [4] Bochkov G N and Kuzovlev Y E, 1981 *Physica* **106** 443
- [5] Bochkov G N and Kuzovlev Y E, 1981 *Physica* **106** 480
- [6] Jarzynski C, 1997 *Phys. Rev. Lett.* **78** 2690
- [7] Jarzynski C, 1997 *Phys. Rev. E* **56** 5018
- [8] Crooks G E, 1998 *J. Stat. Phys.* **90** 1481
- [9] Crooks G E, 1999 *Phys. Rev. E* **60** 2721
- [10] Hummer G and Szabo A, 2001 *Proc. Nat. Acad. Sci.* **98** 3658
- [11] Jarzynski C, 2007 *C. R. Physique* **8** 495
- [12] Collin D, Ritort F, Jarzynski C, Smith S B, Tinoco I Jr and Bustamante C, 2005 *Nature* **437** 231
- [13] Seifert U, *Stochastic thermodynamics: principles and perspectives*, 2007 *Preprint* 0710.1187
- [14] Seifert U, 2005 *Phys. Rev. Lett.* **95** 040602
- [15] Cleuren B, Van den Broeck C and Kawai R, 2006 *Phys. Rev. Lett.* **96** 050601
- [16] Jarzynski C, 2006 *Phys. Rev. E* **73** 046105
- [17] Evans D J, 2003 *Mol. Phys.* **101** 1551
- [18] Halliday D and Resnick R, 2005 *Fundamentals of Physics* 7th edn (New York: Wiley)
- [19] Gibbs J, 1902 *Elementary Principles in Statistical Mechanics* (New York: Scribner's)
- [20] Schrödinger E, 1962 *Statistical Thermodynamics* (New York: Cambridge University Press)
- [21] Uhlenbeck G and Ford G, 1963 *Lectures in Statistical Mechanics* (Providence, RI: American Mathematical Society)
- [22] Goldstein H, 1980 *Classical Mechanics* 2nd edn (Reading, MA: Addison-Wesley)
- [23] Crooks G E, 2000 *Phys. Rev. E* **61** 2361
- [24] Evans D J and Searles D J, 2002 *Adv. Phys.* **51** 1529
- [25] Sevick E M, Prabhaker R, Williams S R and Searles D J, *Fluctuation theorems*, 2007 *Preprint* 0709.3888