Entropy production and volume contraction in thermostated Hamiltonian dynamics

John D. Ramshaw

Department of Physics, Portland State University, Portland, Oregon 97207, USA (Received 30 August 2017; published 15 November 2017)

Patra *et al.* [Int. J. Bifurcat. Chaos **26**, 1650089 (2016)] recently showed that the time-averaged rates of entropy production and phase-space volume contraction are equal for several different molecular dynamics methods used to simulate nonequilibrium steady states in Hamiltonian systems with thermostated temperature gradients. This equality is a plausible statistical analog of the second law of thermodynamics. Here we show that those two rates are identically equal in a wide class of methods in which the thermostat variables **z** are determined by ordinary differential equations of motion (i.e., methods of the Nosé-Hoover or integral feedback control type). This class of methods is defined by three relatively innocuous restrictions which are typically satisfied in methods of this type.

DOI: 10.1103/PhysRevE.96.052122

I. INTRODUCTION AND SUMMARY

The relation between the time-averaged rates of entropy production and phase-space volume contraction in thermostated Hamiltonian systems has been discussed by several authors [1-6], but for various different systems and from various different viewpoints which make it difficult to compare the results and assess whatever more general significance they may possess. Counterexamples have established that those two rates are not universally identical [2-4], but Patra, Hoover, Hoover, and Sprott (PHHS) [6] observe that they are indeed equal in several quite different molecular dynamics methods with thermostated temperature gradients. PHHS interpret this equivalence as a statistical analog of the second law of thermodynamics, and suggest on that basis that it should be regarded as a desirable property which such systems should ideally possess in order to faithfully represent the nonequilibrium thermodynamic behavior of real physical systems.

The analysis of PHHS suggests that the time-averaged rates of entropy production and phase-space volume contraction may in fact be identical for a broad class of methods whose precise boundaries have not previously been delineated. Our purpose here is to confirm this suspicion by establishing sufficient conditions under which those two rates are identically equal for methods with thermostated temperature gradients in which the thermostat variables are determined by ordinary differential equations of motion; i.e., methods of the Nosé-Hoover or integral feedback control type [1,7,8]. These conditions are relatively unrestrictive and are commonly satisfied in typical methods of this type. They can be summarized as follows:

(A) The Hamiltonian state variables \mathbf{x} and the thermostat variables \mathbf{z} are statistically independent in thermal equalibrium, so that the equilibrium probability distribution in (\mathbf{x}, \mathbf{z}) -space factors into the product of the desired canonical distribution in \mathbf{x} and a distribution in \mathbf{z} .

(B) The equilibrium distribution in **z** is independent of the specified thermostat temperature.

(C) The local nonequilibrium temperature gradient does not contribute to the local phase-space volume contraction rate.

Most methods of the present type which have previously been explored appear to satisfy these conditions, including those considered by PHHS [6].

II. THE EQUILIBRIUM FORMULATION

The present development presumes that the unmodified (i.e., pre-thermostated) system of interest is of the compact generalized Hamiltonian form [9-13]

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} H, \tag{1}$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is the phase point, $f \equiv df/dt$, *t* is the time, $H(\mathbf{x})$ is the Hamiltonian function, $\nabla_x \equiv \partial/\partial \mathbf{x}$, and $\mathbf{A}(\mathbf{x})$ is an antisymmetric matrix which satisfies the condition

$$\nabla_{x} \cdot \mathbf{A} = \mathbf{0}. \tag{2}$$

The antisymmetry of **A** implies that *H* is a constant of the motion, while Eq. (2) further implies that Eq. (1) generates an incompressible or volume-preserving flow in phase space [12,13]. This formulation encompasses both canonical and generalized Hamiltonian systems in a phase space **x** of arbitrary dimensionality *n*, so it applies to both few- and many-particle systems with either integral or half-integral degrees of freedom (i.e., even or odd *n*). In canonical systems with *m* degrees of freedom, n = 2m and $\mathbf{x} = (\mathbf{q}, \mathbf{p})$, where $\mathbf{q} = (q_1, \ldots, q_m)$ and $\mathbf{p} = (p_1, \ldots, p_m)$ are the canonical coordinates and momenta [12,13].

The generalized Hamiltonian dynamics above was recently adopted as the basis of a general formalism for thermostated Hamiltonian dynamics with only a single thermostat variable z [13]. However, most methods in the literature employ more than one thermostat variable, so the singly thermostated formalism is insufficiently general for present purposes. The present development is accordingly based on a more general formulation in which the number of thermostat variables $\mathbf{z} = (z_1, z_2, ...)$ is arbitrary and the thermostated equations of motion are of the form

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} H + \mathbf{V}(\mathbf{x}, \mathbf{z}), \tag{3}$$

$$\dot{\mathbf{z}} = \mathbf{W}(\mathbf{x}, \mathbf{z}). \tag{4}$$

The singly thermostated Hamiltonian systems considered in Ref. [13] are of the general form of Eqs. (3) and (4), so the present development and results apply to that class of systems as a special case.

The equilibrium probability distribution $\rho(\mathbf{x}, \mathbf{z})$ in the combined phase space (\mathbf{x}, \mathbf{z}) must satisfy the stationary Liouville equation

$$\nabla_{x} \cdot [\rho(\mathbf{A} \cdot \nabla_{x} H + \mathbf{V})] + \nabla_{z} \cdot (\rho \mathbf{W}) = 0, \qquad (5)$$

where $\nabla_z \equiv \partial/\partial z$. In order to describe a system in thermal equilibrium with a heat reservoir, the reduced probability distribution in x-space must have the canonical form

$$\int d\mathbf{z} \,\rho(\mathbf{x}, \mathbf{z}) \,=\, \rho_c(\mathbf{x}) \,\equiv\, Q^{-1} \exp\{-\beta H(\mathbf{x})\}, \qquad (6)$$

where $\beta = 1/T$, the specified thermostat temperature *T* is defined in energy units for convenience, and $Q(\beta) = \int d\mathbf{x} \exp\{-\beta H(\mathbf{x})\}$ is the canonical partition function. The simplest way to enforce Eq. (6) is to impose Condition (A), whereby \mathbf{x} and \mathbf{z} are required to be statistically independent in thermal equilibrium [13]; i.e.,

$$\rho(\mathbf{x}, \mathbf{z}) = \rho_c(\mathbf{x}) \,\sigma(\mathbf{z}). \tag{7}$$

The distribution $\sigma(\mathbf{z})$ has no physical significance so its functional form can be specified at will, subject to the requirements that $\sigma > 0$ is normalized and that functions V and W can be found such that Eq. (5) is satisfied.

Combining Eqs. (2) and (5)–(7) we obtain, after a little algebra,

$$-\beta \mathbf{V} \cdot \nabla_{x} H + \nabla_{x} \cdot \mathbf{V} + \mathbf{W} \cdot \nabla_{z} \log \sigma + \nabla_{z} \cdot \mathbf{W} = 0.$$
(8)

A particular thermostated Hamiltonian model or method is defined by the form of the functions V(x,z), W(x,z), and $\sigma(z)$, which must be chosen in such a way that Eq. (8) is identically satisfied. For present purposes it is unnecessary to know the mathematical form of those functions; it suffices that they exist and satisfy Eq. (8). However, the presence of β in Eq. (8) implies that W(x,z) must depend parametrically on β , and in general V(x,z) will likewise depend parametrically on β . For the time being, we shall allow for the possibility that $\sigma(z)$ also depends on β . Thus the quantities V, W, and σ will henceforth be denoted as $V(x,z,\beta)$, $W(x,z,\beta)$, and $\sigma(z,\beta)$, and it must be remembered in what follows that the differential operators ∇_x and ∇_z in Eq. (8) represent partial derivatives with β held constant.

III. THE NONEQUILIBRIUM FORMULATION

The equilibrium formulation of Sec. II will be extended to nonequilibrium steady states by the common expedient [5,6,14–17] of replacing the thermostat temperature *T* by a nonuniform state-dependent temperature $T(\mathbf{x})$. We emphasize that (a) the present development is restricted to nonequilibrium thermostats of this particular type, which can be thought of as representing an artificial continuous distribution of thermal reservoirs which somehow interpenetrate the entire system; and (b) such a distribution is physically unrealistic, especially in many-particle systems, and one has no assurance that the results thereby obtained remain valid for other more realistic methods of producing nonequilibrium steady states; e.g., specifying temperatures only at the boundaries of the system (e.g., Ref. [18]).

The parameter $\beta = 1/T$ is therefore a specified constant in the equilibrium case, while in the nonequilibrium case $\beta(\mathbf{x}) = 1/T(\mathbf{x})$ is a specified function of \mathbf{x} . In both cases, we presume and require that the motion remains confined to a finite region of the combined phase space (\mathbf{x}, \mathbf{z}) , so that all nonsingular state functions of (\mathbf{x}, \mathbf{z}) remain bounded for all time and their time averages are well defined. The time average of an arbitrary function F(t) is defined by

$$\langle F(t) \rangle \equiv \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^{\tau} dt F(t).$$
 (9)

Thus if F(t) is bounded for $0 \leq t < \infty$, then

$$\langle \dot{F} \rangle \equiv \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt \, \frac{dF(t)}{dt} = \lim_{\tau \to \infty} \frac{1}{\tau} [F(\tau) - F(0)] = 0,$$
(10)

i.e., the time average of the time derivative of any bounded function of t vanishes. Of course, any state function $F(\mathbf{x}, \mathbf{z})$ can be regarded as an implicit function of t via the functions $\mathbf{x}(t)$ and $\mathbf{z}(t)$ determined by Eqs. (3) and (4); i.e., $F(t) = F(\mathbf{x}(t), \mathbf{z}(t))$. It then follows that $\dot{F} = dF/dt = \dot{\mathbf{x}} \cdot \nabla_x F + \dot{\mathbf{z}} \cdot \nabla_z F$.

The replacement $\beta \rightarrow \beta(\mathbf{x})$ is to be made in the functions $\mathbf{V}(\mathbf{x}, \mathbf{z}, \beta)$ and $\mathbf{W}(\mathbf{x}, \mathbf{z}, \beta)$ in the dynamical equations of motion (3) and (4), which thereby become

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} H + \widehat{\mathbf{V}}(\mathbf{x}, \mathbf{z}), \tag{11}$$

$$\dot{\mathbf{z}} = \widehat{\mathbf{W}}(\mathbf{x}, \mathbf{z}),\tag{12}$$

where $\widehat{\mathbf{V}}(\mathbf{x}, \mathbf{z}) \equiv \mathbf{V}(\mathbf{x}, \mathbf{z}, \beta(\mathbf{x}))$ and $\widehat{\mathbf{W}}(\mathbf{x}, \mathbf{z}) \equiv \mathbf{W}(\mathbf{x}, \mathbf{z}, \beta(\mathbf{x}))$. The replacement $\beta \rightarrow \beta(\mathbf{x})$ modifies the velocity field $(\dot{\mathbf{x}}, \dot{\mathbf{z}})$ in phase space, and thereby the trajectories $(\mathbf{x}(t), \mathbf{z}(t))$ which the system traces out in time. Now for any particular model or method of interest, Eq. (8) is an identity which is valid for arbitrary values of the variables (\mathbf{x}, \mathbf{z}) and the parameter β , so it remains valid when β is replaced by $\beta(\mathbf{x})$ therein. As noted above, however, the differential operators ∇_x and ∇_z in Eq. (8) represent, and must be interpreted as, partial derivatives at constant β , so all terms in Eq. (8) must be expressed as functions of $(\mathbf{x}, \mathbf{z}, \beta)$ by evaluating those derivatives *prior* to replacing β by $\beta(\mathbf{x})$. However, inspection of Eq. (8) shows that $\nabla_x \cdot \mathbf{V}$ is actually the only term therein for which the operations ∇_x and $\beta \rightarrow \beta(\mathbf{x})$ do not commute, so it is the only term affected by these considerations. Thus β can simply be replaced by $\beta(\mathbf{x})$ in the other terms as they stand, but the operator ∇_x in the term $\nabla_x \cdot \mathbf{V}$ does not, and cannot be allowed to, operate on the x dependence of $\beta(\mathbf{x})$. The net result of replacing β by $\beta(\mathbf{x})$ in Eq. (8) is therefore

$$-\beta(\mathbf{x})\widehat{\mathbf{V}}\cdot\nabla_{x}H + [\nabla_{x}\cdot\mathbf{V}(\mathbf{x},\mathbf{z},\beta)]_{\beta=\beta(\mathbf{x})} +\widehat{\mathbf{W}}\cdot\nabla_{z}\log\widehat{\sigma}+\nabla_{z}\cdot\widehat{\mathbf{W}}=0,$$
(13)

where $\widehat{\sigma}(\mathbf{z}, \mathbf{x}) \equiv \sigma(\mathbf{z}, \beta(\mathbf{x}))$.

The second term in Eq. (13) is related to $\nabla_x \cdot \widehat{\mathbf{V}}$ by the chain rule, which implies

$$\nabla_{x} \cdot \widehat{\mathbf{V}} = [\nabla_{x} \cdot \mathbf{V}(\mathbf{x}, \mathbf{z}, \beta)]_{\beta = \beta(\mathbf{x})} + \mathbf{\Lambda} \cdot \nabla_{x} \beta, \qquad (14)$$

where

$$\mathbf{\Lambda}(\mathbf{x}, \mathbf{z}) \equiv \left[\frac{\partial \mathbf{V}(\mathbf{x}, \mathbf{z}, \beta)}{\partial \beta}\right]_{\beta = \beta(\mathbf{x})}.$$
 (15)

Combining Eqs. (13) and (14), we obtain

$$-\beta(\mathbf{x})\widehat{\mathbf{V}}\cdot\nabla_{x}H + \nabla_{x}\cdot\widehat{\mathbf{V}} - \mathbf{\Lambda}\cdot\nabla_{x}\beta$$
$$+\widehat{\mathbf{W}}\cdot\nabla_{z}\log\widehat{\sigma} + \nabla_{z}\cdot\widehat{\mathbf{W}} = 0.$$
(16)

The local phase-space volume contraction rate Δ is simply the negative of the divergence of the velocity field in the combined phase space (**x**,**z**). According to Eqs. (11) and (12), that velocity field is ($\dot{\mathbf{x}}, \dot{\mathbf{z}}$) = ($\mathbf{A} \cdot \nabla_{\mathbf{x}} H + \widehat{\mathbf{V}}, \widehat{\mathbf{W}}$), so

$$\Delta = -\nabla_{x} \cdot (\mathbf{A} \cdot \nabla_{x} H + \widehat{\mathbf{V}}) - \nabla_{z} \cdot \widehat{\mathbf{W}}$$
$$= -\nabla_{x} \cdot \widehat{\mathbf{V}} - \nabla_{z} \cdot \widehat{\mathbf{W}}, \qquad (17)$$

where use has again been made of Eq. (2). Equation (16) can therefore be rewritten as

$$-\beta(\mathbf{x})\widehat{\mathbf{V}}\cdot\nabla_{\mathbf{x}}H + \widehat{\mathbf{W}}\cdot\nabla_{\mathbf{z}}\log\widehat{\sigma} = \Delta + \mathbf{\Lambda}\cdot\nabla_{\mathbf{x}}\beta.$$
(18)

We now observe that the first two terms in Eq. (18) are related to the total time derivatives of H and $\hat{\sigma}$. It follows from Eqs. (11) and (12) that

$$\dot{H} = \dot{\mathbf{x}} \cdot \nabla_{x} H = (\mathbf{A} \cdot \nabla_{x} H + \widehat{\mathbf{V}}) \cdot \nabla_{x} H = \widehat{\mathbf{V}} \cdot \nabla_{x} H$$
(19)

and

$$\frac{d\log\widehat{\sigma}}{dt} = \dot{\mathbf{z}} \cdot \nabla_z \log\widehat{\sigma} + \dot{\mathbf{x}} \cdot \nabla_x \log\widehat{\sigma} = \widehat{\mathbf{W}} \cdot \nabla_z \log\widehat{\sigma} + \lambda \dot{\mathbf{x}} \cdot \nabla_x \beta, \qquad (20)$$

where

$$\lambda(\mathbf{z}, \mathbf{x}) \equiv \left[\frac{\partial \log \sigma(\mathbf{z}, \beta)}{\partial \beta}\right]_{\beta = \beta(\mathbf{x})}.$$
 (21)

Combining Eqs. (18)–(20), we obtain

$$-\beta(\mathbf{x})\dot{H} + \frac{d\log\widehat{\sigma}}{dt} = \Delta + \mathbf{\Lambda}\cdot\nabla_{x}\beta + \lambda\dot{\mathbf{x}}\cdot\nabla_{x}\beta.$$
(22)

The terms proportional to $\nabla_x \beta$ in Eq. (22) spoil what would otherwise be a remarkably simple relationship between \dot{H} and Δ . We shall therefore remove the second such term by imposing Condition (B), which requires σ to be independent of β so that $\lambda = 0$. Note that Condition (B) is automatically satisfied if the thermostat variables z are defined to be dimensionless, since the only way to nondimensionalize any dependence of σ on β or T is to introduce a corresponding dependence on H and thereby on x, which would violate Condition (A). This observation further encourages the guiding principle that thermostat variables should be dimensionless [6,13].

Equations (14) and (17) show that the term $\mathbf{\Lambda} \cdot \nabla_x \beta$ represents the contribution of $\nabla_x \beta$ to $-\Delta$. Fortunately, this term vanishes identically in most thermostated models of canonical Hamiltonian systems, in which $\mathbf{x} = (\mathbf{q}, \mathbf{p})$. The reason is that such models typically possess the following two simplifying features: (i) $\dot{\mathbf{q}}$ contains no thermostat terms, in accordance with the fact that the time evolution of **q** is essentially kinematical in nature, so it seems conceptually incongruous to tamper with it [13]. Thus **V** is usually of the form $(\mathbf{0}, \mathbf{V}_p)$, so that **A** is of the form $(\mathbf{0}, \mathbf{A}_p)$. (ii) $\beta(\mathbf{x})$ is invariably presumed to depend only on **q**, since it would also be incongruous for the thermostat temperatures to depend on the momenta **p**. This implies that $\nabla_x \beta$ is of the form $(\nabla_q \beta, \mathbf{0})$, where $\nabla_q \equiv \partial/\partial \mathbf{q}$. It then follows from (i) and (ii) that

$$\mathbf{\Lambda} \cdot \nabla_{\mathbf{x}} \boldsymbol{\beta} = 0. \tag{23}$$

This simplification provides a further incentive to refrain from introducing thermostat terms into $\dot{\mathbf{q}}$. We shall accordingly impose Condition (C), whereby Δ is required to be independent of $\nabla_x \beta$ so that Eq. (23) is satisfied. The net result of Conditions (B) and (C) is that Eq. (22) now reduces to

$$-\beta(\mathbf{x})\dot{H} + \frac{d\log\widehat{\sigma}}{dt} = \Delta.$$
(24)

All that remains is to take the time average of Eq. (24), which yields our final result:

$$-\langle \beta \dot{H} \rangle = \langle \Delta \rangle, \tag{25}$$

where use has been made of Eq. (10). The local rate of energy transfer from the system to the reservoir is simply $-\dot{H}$, and $T = 1/\beta$ is the corresponding local reservoir temperature. Thus $-\beta \dot{H} = -\dot{H}/T$ is the local rate at which the reservoir entropy increases due to the energy transfer; i.e., the local entropy production rate. Equation (25) therefore states and demonstrates that the time-averaged rates of entropy production and volume contraction are identically equal for all systems of the general form of Eqs. (3) and (4) that satisfy Conditions (A), (B), and (C). The methods recently analyzed by PHHS [6] are all of this generic type, so Eq. (25) applies to each of those methods, thereby confirming the results of PHHS. This development illustrates the utility of a general formulation, which allows one to derive correspondingly general relations that automatically apply to all of its special cases, thereby making it unnecessary to consider them individually.

Finally, we remark that Conditions (B) and (C) are actually stronger than necessary, since the weaker conditions $\langle \lambda \dot{\mathbf{x}} \cdot \nabla_x \beta \rangle = \langle \mathbf{\Lambda} \cdot \nabla_x \beta \rangle = 0$ would have sufficed, but it seems difficult to envision models for which those conditions are satisfied while Conditions (B) and (C) are not.

ACKNOWLEDGMENT

I am grateful to Bill Hoover for many stimulating discussions.

- W. G. Hoover, *Computational Statistical Mechanics* (Elsevier, Amsterdam, 1991).
- [4] H. van Beijeren and J. R. Dorfman, Physica A 279, 21 (2000).
- [2] E. G. D. Cohen and L. Rondoni, Chaos 8, 357 (1998).
- [3] D. Daems and G. Nicolis, Phys. Rev. E 59, 4000 (1999).
- [5] W. G. Hoover and C. G. Hoover, *Time Reversibility, Computer Simulation, Algorithms, Chaos*, 2nd ed. (World Scientific, Singapore, 2012).

- [6] P. K. Patra, W. G. Hoover, C. G. Hoover, and J. C. Sprott, Int. J. Bifurcat. Chaos 26, 1650089 (2016).
- [7] W. G. Hoover, *Molecular Dynamics*, Lecture Notes in Physics, Vol. 258 (Springer-Verlag, Berlin, 1986).
- [8] D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids* (Academic Press, San Diego, CA, 1990).
- [9] L. A. Pars, A Treatise on Analytical Dynamics (Heinemann, London, 1965).
- [10] J. S. Langer, Ann. Phys. (NY) 54, 258 (1969).
- [11] R. Zwanzig, J. Stat. Phys. 9, 215 (1973).

- [12] J. D. Ramshaw and K. Lindenberg, J. Stat. Phys. 45, 295 (1986).
- [13] J. D. Ramshaw, Phys. Rev. E 92, 052138 (2015).
- [14] W. G. Hoover, J. C. Sprott, and P. K. Patra, Phys. Lett. A 379, 2935 (2015).
- [15] W. G. Hoover, C. G. Hoover, and J. C. Sprott, Molec. Simul. 42, 1300 (2016).
- [16] H. A. Posch and W. G. Hoover, Phys. Rev. E 55, 6803 (1997).
- [17] J. C. Sprott, W. G. Hoover, and C. G. Hoover, Phys. Rev. E 89, 042914 (2014).
- [18] B. L. Holian, W. G. Hoover, and H. A. Posch, Phys. Rev. Lett. 59, 10 (1987).