Microscopic reversibility and heat for thermostatted systems

Takaaki Monnai*
Waseda Institute for Advanced Study, 1-6-1 Nishiwaseda Shinjuku-ku, Tokyo 169-8050, Japan
(Received 25 December 2012; published 11 April 2013)

In order to test the universality of a symmetry for the trajectory obtained for Hamiltonian dynamics, we investigate the case of Nose-Hoover thermostatted dynamics with the use of a clear separation between the system and reservoir. Remarkably, the same symmetry as the Hamiltonian dynamics holds despite the presence of the dissipation, which causes the phase volume contraction. As a nontrivial application of the symmetry, we further show that the microscopic reversibility for open systems holds just as in the Hamiltonian dynamics. This bridges the first and second laws of thermodynamics under the proper definition of the work and heat.

DOI: 10.1103/PhysRevE.87.042107 PACS number(s): 05.70.Ln, 05.40.−a

I. INTRODUCTION

Time reversibility is one of the most fundamental laws of Hamiltonian dynamics and electromagnetic theory [1,2]. The reversibility is also the case for Nose-Hoover thermostatted dynamics, which is widely used for molecular dynamics simulations [3,4]. We consider the thermostatted dynamics, since it is dissipative and has a clear separation between the system and reservoir. Indeed, recently, the thermostatted dynamics has been made consistent with the first and second law of thermodynamics [5] by the separation between the system and reservoir, i.e., regarding the thermostat as a reservoir. Then, as explained in Sec. IV, there is a well-defined heat flowing to the system as a functional of the system state trajectory.

Still, it is unclear whether the nonequilibrium thermodynamic relations valid for Hamiltonian dynamics also hold for the thermostatted case. Actually, the thermostatted dynamics cannot be a special limiting case of the Hamiltonian dynamics, since the friction coefficient takes both positive and negative values to maintain the kinetic temperature constant [3,6]. Reference [5] shows that the first and second laws of thermodynamics are well-defined also for thermostatted dynamics. This is achieved by regarding the thermostat as a reservoir and assuming that the total energy is defined so that the stationary state is described by a corresponding canonical distribution function. Given the definition of the heat, it is nontrivial whether a rigorous theorem of microscopic reversibility obtained in Refs. [7,8] for Hamiltonian dynamics holds also for dissipative thermostatted dynamics. Actually, the weak coupling plays an essential role for the derivation. However, this is not the case for the thermostatted dynamics due to the clear separation of system and reservoir.

In this paper, we test the universality of the microscopic reversibility obtained in Refs. [7,8] by exploring the case of Nose-Hoover dynamics. Especially, we show that the corresponding relation, and as a special case, Crooks relation [9–11], exactly holds also for thermostatted dynamics Eqs. (13) and (24) under the same nontrivial definitions of the heat and work as in Ref. [5]. This is remarkable, since the dynamics is here dissipative and coupling strength is arbitrary, which is in marked contrast to Hamiltonian dynamics.

This paper is organized as follows. In Sec. II, we explain the time evolution and conservation of the probability. In Sec. III, we derive a rigorous theorem, Eq. (13), based on the time reversal symmetry. In Sec. IV, Crooks relation is derived as well, based on the separation of system and reservoir and Eq. (13). Section V is devoted to a summary.

II. TIME EVOLUTION

We consider an N-particle system in a D-dimensional time-dependent potential in contact with a Nose-Hoover thermostat [3,12,15]. We briefly summarize fundamental aspects of the thermostatted dynamics. The equations of motion are given as

$$\begin{aligned}\dot{q}_i(t) &= \frac{p_i(t)}{m_i}\\ \dot{p}_i(t) &= -\frac{\partial}{\partial q_i(t)}V(\{q_i(t)\},t) - \zeta(t)p_i(t)\\ \dot{\zeta}(t) &= \frac{1}{\tau^2}\frac{\sum_{i=1}^{DN}\frac{p_i(t)^2}{m_i} - DNk_BT}{k_BT},\end{aligned} \tag{1}$$

where q_i and p_i stand for the coordinates and momenta of ith degrees of freedom, and m_i is the mass. The friction coefficient ζ mimics the reservoir variables, which maintain the kinetic energy per unit particle around the equipartition value $\frac{Dk_BT}{2}$ with a finite relaxation time τ. Here, k_B is the Boltzmann constant, and T is the absolute temperature. We assume that the potential $V(\{q_i(t)\},t)$ is explicitly time dependent by an external forcing.

Let us abbreviate the set of all the variables as $\Gamma = (\{q_i\},\{p_i\},\zeta)$ and similarly the set of positions and momenta as $\Gamma_s = (\{q_i\},\{p_i\})$. The time evolution of an arbitrary dynamical variable $A(\Gamma(t))$ is described by the Liouvillian operator $L(t)$ as

$$\begin{aligned}&\frac{d}{dt}A[\Gamma(t)] = L(t)A[\Gamma(t)];\\ &L(t) \equiv \sum_{i=1}^{DN}\left[\dot{q}_i(t)\frac{\partial}{\partial q_i(t)} + \dot{p}_i(t)\frac{\partial}{\partial p_i(t)}\right] + \dot{\zeta}(t)\frac{\partial}{\partial\zeta(t)}.\end{aligned} \tag{2}$$

*monnai@suou.waseda.jp

On the other hand, the distribution function $\rho(\Gamma,t)$ of the total state evolves as

$$\begin{aligned}\frac{\partial}{\partial t}\rho(\Gamma,t) &= -\sum_{i=1}^{DN}\left\{\frac{\partial}{\partial q_i(t)}[\dot{q}_i(t)\rho(\Gamma,t)]+\frac{\partial}{\partial p_i(t)}[\dot{p}_i(t)\rho(\Gamma,t)]\right\}\\ &\quad -\frac{\partial}{\partial\zeta(t)}[\dot{\zeta}(t)\rho(\Gamma,t)]\\ &= -[L(t)-DN\zeta]\rho(\Gamma,t). \end{aligned} \tag{3}$$

From Eqs. (2) and (3), the time evolution of the probability distribution is given as

$$\begin{aligned}\frac{d}{dt}\rho[\Gamma(t),t] &= \frac{\partial}{\partial\Gamma(t)}\{\dot{\Gamma}(t)\rho[\Gamma(t),t]\}+\frac{\partial}{\partial t}\rho[\Gamma(t),t]\\ &= \{L(t)-[L(t)-DN\zeta(t)]\}\,\rho[\Gamma(t),t]. \end{aligned} \tag{4}$$

And the distribution at time t is thus expressed as

$$\rho[\Gamma(t),t] = e^{DN\int_0^t\zeta(s)ds}\rho[\Gamma(0),0], \tag{5}$$

with the phase volume contraction factor $DN\int_0^t\zeta(s)ds$. In Sec. IV, we will give an additional remark on the phase volume contraction in the context of thermodynamics.

We regard Γ_s and ζ as the system and the reservoir variables [6]. The energies of the system H_s and the reservoir H_r are defined as [3,12]

$$\begin{aligned}H_s(\Gamma_s,t) &= \sum_{i=1}^{DN}\frac{p_i^2}{2m_i}+V(\{q_i\},t)\\ H_r(\zeta) &= \frac{k_BT\tau^2}{2}\zeta^2. \end{aligned} \tag{6}$$

The total energy is then expressed by the sum

$$H(\Gamma,t) = H_s(\Gamma_s,t)+H_r(\zeta). \tag{7}$$

Note that the energies H_s and H_r are not actual Hamiltonians, i.e., time evolution operators. Indeed, the reservoir energy H_r is defined so that the canonical state $\rho_{\rm eq}(\Gamma)\propto e^{-\frac{H(\Gamma)}{k_BT}}$ gives a stationary distribution provided that the potential V does not explicitly depend on time.

III. MICROSCOPIC REVERSIBILITY

In this section, we calculate the ratio between the probability functionals to have time forward and its corresponding reversed trajectories. First, let us define the probability to have a specific trajectory of the system state expressed by a sequence of the state $\{\Gamma_s(0),\Gamma_s(\Delta t),\ldots,\Gamma_s(j\Delta t),\ldots,\Gamma_s[(N-1)\Delta t],\Gamma_s(N\Delta t)\}$. Here we consider the snapshot of the trajectory with a short time step $\Delta t=\frac{\mathcal{T}}{N}$ for a technical reason, and later the continuous limit $N\to\infty$ is taken [13]. The initial state $\Gamma(0)$ is chosen from the initial distribution $\rho(\Gamma_s,\zeta,0)=\rho_s(\Gamma_s,0)\rho_r^{\rm eq}(\zeta)$. Here the system distribution $\rho_s(\Gamma_s,0)$ is arbitrary and we denote the canonical distribution of the reservoir as $\rho_r^{\rm eq}(\zeta)=\frac{1}{Z_r}e^{-\frac{H_r(\zeta)}{k_BT}}$. The equilibrium distribution is used for the reservoir variable by an analogy to the thermodynamic large reservoir. The system state trajectory $\{\Gamma_s(t)\}$ fluctuates from one realization to the other due to the interaction with the reservoir. The probability to have the above-mentioned trajectory is defined as

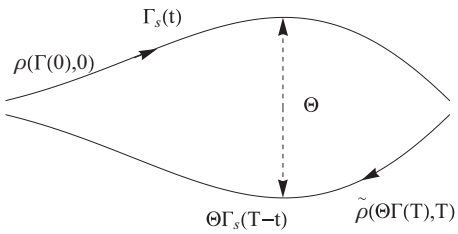

FIG. 1. Time-forward and -reversed trajectories $\Gamma_s(t)$ and $\Theta\Gamma_s(\mathcal{T}-t)$ are schematically shown. Initial distributions are $\rho[\Gamma(0),0]$ and $\tilde{\rho}[\Theta\Gamma(\mathcal{T}),\mathcal{T}]$.

$$\begin{aligned}&P_F(\{\Gamma_s(0),\Gamma_s(\Delta t),\ldots,\Gamma_s(j\Delta t),\ldots,\Gamma_s(N\Delta t)\})\\ &\quad=\int d\Gamma\rho(\Gamma,0)\prod_{j=0}^{N}\delta[\Gamma_s(j\Delta t)-(\mathcal{V}_{j\Delta t}\Gamma)_s], \end{aligned} \tag{8}$$

where the time evolution operator during $0\leqslant t\leqslant j\Delta t$ is abbreviated as $\mathcal{V}_{j\Delta t}=\mathrm{T}\{e^{\int_0^{j\Delta t}L(s)ds}\}$ [7]. The Dirac delta specifies the system state at each time $t=j\Delta t$.

We also consider the corresponding time reversed process as in Fig. 1. And for this purpose let Θ be the time-reversal operator acting on the state. In the absence of the magnetic field, Θ just reverses the momenta and the friction coefficient as $\Theta(\{q_i\},\{p_i\},\zeta)=(\{q_i\},-\{p_i\},-\zeta)$ [18]. If there is a magnetic field, we also reverse its direction. Then the motion of charged particles is also reversible.

Since the Nose-Hoover dynamics is time-reversal symmetric, the reversed trajectory $\{\Theta\Gamma_s(N\Delta t),\Theta\Gamma_s[(N-1)\Delta t],\ldots,\Theta\Gamma_s(\Delta t),\Theta\Gamma_s(0)\}$ is also a solution of the equations of motion under the reversed time dependence of the potential $V(\{q_i\},\mathcal{T}-t)$. As an initial distribution of the reversed dynamics, we choose $\tilde{\rho}(\Theta\Gamma,\mathcal{T})=\tilde{\rho}_s(\Theta\Gamma_s,\mathcal{T})\rho_r^{\rm eq}(\Theta\zeta)$. Again, the system distribution $\tilde{\rho}_s(\Theta\Gamma_s,\mathcal{T})$ is arbitrary. The tilde indicates that this distribution is statistically independent from that of the forward process $\rho(\Gamma,0)$. Then the probability to have the reversed trajectory is defined as

$$\begin{aligned}&P_R(\{\Theta\Gamma_s(\mathcal{T}),\Theta\Gamma_s(\mathcal{T}-\Delta t),\ldots,\Theta\Gamma_s(\Delta t),\Theta\Gamma_s(0)\})\\ &\quad=\int d\Gamma\tilde{\rho}(\Theta\Gamma,\mathcal{T})\prod_{j=0}^{N}\delta\left[\Theta\Gamma_s(j\Delta t)-\left(\Theta\mathcal{V}_{j\Delta t}\mathcal{V}_{N\Delta t}^{-1}\Gamma\right)_s\right]. \end{aligned} \tag{9}$$

The constraint for $j=N$ specifies the initial state as $\Theta\Gamma_s(N\Delta t)=(\Theta\Gamma)_s$. It is remarked that the evolution operator from the initial state $\Theta\Gamma(N\Delta t)$ to $\Theta\Gamma(j\Delta t)$ is given by

$$\begin{aligned}&\Theta\mathcal{V}_{j\Delta t}\mathcal{V}_{N\Delta t}^{-1}\\ &\quad= e^{L[(j+1)\Delta t]\Delta t}e^{L[(j+2)\Delta t]\Delta t}\cdots e^{L[(N-1)\Delta t]\Delta t}\\ &\qquad\times e^{L(N\Delta t)\Delta t}\Theta+O\left[\left(\frac{\Delta t}{\mathcal{T}}\right)^2\right], \end{aligned} \tag{10}$$

where we used the property $\Theta L(t)=-L(t)\Theta$. In the continuous limit $\Delta t\to 0$, the right-hand side is expressed by the time-ordered product. Here is a technical remark. If the probability of the reversed trajectory happens to be zero, then that of the forward trajectory under the forward protocol is also vanishing due to the reversibility.

The forward probability is rewritten into an expression similar to Eq. (9)

$$\begin{aligned} P_F(\{\Gamma_s(0),\Gamma_s(\Delta t),\ldots,\Gamma_s(j\Delta t),\ldots,\Gamma_s(N\Delta t)\}) &= \int d\Gamma \rho(\Gamma,0)\prod_{j=0}^{N}\delta[\Gamma_s(j\Delta t)-(\mathcal{V}_{j\Delta t}\Gamma)_s] \\ &= \int d\Gamma'\left|\frac{\partial\Gamma}{\partial\Gamma'}\right|\rho\left(\mathcal{V}_{N\Delta t}^{-1}\Gamma',0\right)\prod_{j=0}^{N}\delta\left[\Gamma_s(j\Delta t)-\left(\mathcal{V}_{j\Delta t}\mathcal{V}_{N\Delta t}^{-1}\Gamma'\right)_s\right] \\ &= \int d\Gamma'\rho(\Gamma',N\Delta t)\prod_{j=0}^{N}\delta\left[\Theta\Gamma_s(j\Delta t)-\left(\Theta\mathcal{V}_{j\Delta t}\mathcal{V}_{N\Delta t}^{-1}\Gamma\right)_s\right], \end{aligned} \tag{11}$$

where in the second equality, we changed the variable from Γ to $\Gamma' = \mathcal{V}_{N\Delta t}\Gamma$; the conservation of the probability Eq. (5) is used in the third equality. Indeed, it is straightforward to verify $|\frac{\partial\Gamma}{\partial\Gamma'}| = e^{DN\int_0^{\mathcal{T}}\zeta(t)dt}$, and this term provides a factor $e^{DN\int_0^{\mathcal{T}}\zeta(t)dt}$ in Eq. (5) for $t=\mathcal{T}$. This point is an important step for the derivation of the microscopic reversibility of dissipative open systems. And Θ is inserted in the Dirac delta in the third equality.

Comparing Eqs. (9) and (11), the evolution operator $\Theta\mathcal{V}_{j\Delta t}\mathcal{V}_{N\Delta t}^{-1}$ is common and we have

$$\begin{aligned} &\frac{P_F(\{\Gamma_s(0),\Gamma_s(\Delta t),\ldots,\Gamma_s[(N-1)\Delta t],\Gamma_s(N\Delta t)\})}{P_R(\{\Theta\Gamma_s(N\Delta t),\ldots,\Theta\Gamma_s(0)\})} \\ &= \int d\Gamma_s d\zeta\,\frac{\rho(\Gamma,N\Delta t)}{\tilde{\rho}(\Theta\Gamma,N\Delta t)}\frac{\tilde{\rho}(\Theta\Gamma,N\Delta t)\prod_{j=0}^{N}\delta\left[\Theta\Gamma_s(j\Delta t)-\left(\Theta\mathcal{V}_{j\Delta t}\mathcal{V}_{N\Delta t}^{-1}\Gamma\right)_s\right]}{\int d\Gamma'\tilde{\rho}(\Theta\Gamma',N\Delta t)\prod_{j=0}^{N}\delta[\Theta\Gamma_s(j\Delta t)-(\Theta\mathcal{V}_{j\Delta t}\mathcal{V}_{N\Delta t}^{-1}\Gamma')_s]} \\ &= \int d\zeta\,\frac{\rho(\Gamma,N\Delta t)}{\tilde{\rho}(\Theta\Gamma,N\Delta t)}\frac{\chi(\zeta)}{\int d\zeta'\chi(\zeta')} \equiv \left\langle\frac{\rho(\Gamma,N\Delta t)}{\rho(\Theta\Gamma,N\Delta t)}\right\rangle_R, \end{aligned} \tag{12}$$

as in Ref. [7]. In the second equality, we integrate over the system variables and introduced the characteristic function $\chi(\zeta) \equiv \tilde{\rho}_s[\Theta\Gamma_s(\mathcal{T}),\mathcal{T}]\rho_r^{\rm eq}(\zeta)\prod_{j=0}^{N-1}\delta[\Theta\Gamma_s(j\Delta t)-(\Theta\mathcal{V}_{j\Delta t}\mathcal{V}_{N\Delta t}^{-1}\Gamma)_s]$, where Γ abbreviates the set $[\Gamma_s(N\Delta t),\zeta]$. Actually, $\chi(\zeta)$ is a function of ζ and gives a finite contribution only when the time-reversed trajectory is just equal to $\{\Theta\Gamma_s(N\Delta t),\Theta\Gamma_s[(N-1)\Delta t],\ldots,\Theta\Gamma_s(\Delta t),\Theta\Gamma_s(0)\}$ and otherwise zero. Here, $\langle\rangle_R$ shows the average with respect to the normalized probability density $\frac{\chi(\zeta)}{\int d\zeta'\chi(\zeta')}$. Such a probability is given by the characteristic function with which the reversed trajectory $\{\Theta\Gamma_s(\mathcal{T}),\Theta\Gamma(\mathcal{T}-\Delta t),\ldots,\Theta\Gamma_s(\Delta t),\Theta\Gamma_s(0)\}$ is observed under the reversed protocol [14]. Note that the average is not over the probability of the reversed trajectory. We remark that multiple reservoir states could amount to the same trajectory $\{\Gamma_s(t)\}$, and the average is necessary. In the continuous limit $\Delta t\to 0$ and $N\Delta t\to\mathcal{T}$, we finally have a universal expression of the microscopic reversibility just as in the Hamiltonian dynamics

$$\frac{P_F[\Gamma_s(t)]}{P_R[\Theta\Gamma_s(\mathcal{T}-t)]} = \left\langle\frac{\rho(\Gamma,\mathcal{T})}{\tilde{\rho}(\Theta\Gamma,\mathcal{T})}\right\rangle_R. \tag{13}$$

This fundamental relation contains the generalized detailed balance [7], nonequilibrium work relation [15], and an internal relation between the microscopic reversibility and the heat [9]. Here, $P_F[\Gamma_s(t)]$ and $P_R[\Theta\Gamma_s(\mathcal{T}-t)]$ are the probability functionals to have trajectories $\{\Gamma_s(t)\}$ and $\{\Theta\Gamma_s(\mathcal{T}-t)\}$ under the forward and reversed forcing protocols. Equation (13) expresses how easily the forward process occurs by an amount of asymmetry between $\rho(\Gamma,\mathcal{T})$ and $\tilde{\rho}(\Theta\Gamma,\mathcal{T})$. It is remarked that Eq. (13) holds for arbitrary initial distribution functions; i.e., the reservoir state would deviate from equilibrium and especially the generalized detailed balance can be obtained similarly as in Ref. [7]. But the analysis in the next section requires initial canonical distribution of the reservoir.

IV. MICROSCOPIC REVERSIBILITY AND THERMODYNAMIC RELATIONS

In this section, we derive the Crooks relation, which connects the ratio between the forward and reversed *conditional* probabilities to the heat $Q[\Gamma_s(t)]$. This provides another verification of the universality of the thermodynamic relations for Hamiltonian systems. The conditional probabilities to have the forward and reversed trajectories given that the initial states are $\Gamma_s(0)$ and $\Theta\Gamma_s(\mathcal{T})$ are defined as

$$P_F[\Gamma_s(t)|\Gamma_s(0)] \equiv \frac{1}{\rho_s[\Gamma_s(0),0]}P_F[\Gamma_s(t)], \tag{14}$$

and

$$\begin{aligned} &P_R[\Theta\Gamma_s(\mathcal{T}-t)|\Theta\Gamma_s(\mathcal{T})] \\ &\quad\equiv \frac{1}{\tilde{\rho}_s[\Theta\Gamma_s(\mathcal{T}),\mathcal{T}]}P_R[\Theta\Gamma_s(\mathcal{T}-t)], \end{aligned} \tag{15}$$

where the marginal distributions defined in the previous section are used. Note that $P_F[\Gamma_s(t)|\Gamma_s(0)]$ is defined as a continuous limit of

$$\begin{aligned} &P_F(\{\Gamma_s(\Delta t),\ldots,\Gamma_s(j\Delta t),\ldots,\Gamma_s(N\Delta t)\}|\Gamma_s(0)) \\ &\quad= \frac{1}{\rho_s[\Gamma_s(0),0]}P_F(\{\Gamma_s(0),\Gamma_s(\Delta t),\ldots, \\ &\qquad\Gamma_s[(N-1)\Delta t],\Gamma_s(N\Delta t)\}). \end{aligned} \tag{16}$$

The joint probability $P_F[\Gamma_s(t)]$ is divided by the probability $\rho_s[\Gamma_s(0),0]$ that the initial state $\Gamma_s(0)$ was selected.

The conservation of the probability Eq. (5) amounts to

$$\rho[\Gamma_s(\mathcal{T}),\zeta,\mathcal{T}] = e^{DN\int_0^{\mathcal{T}}\zeta(t)dt}\rho[\Gamma_s(0),\mathcal{V}^{-1}\zeta,0]. \quad (17)$$

Here, $\mathcal{V}^{-1} \equiv \lim_{\Delta t\to 0} V_{N\Delta t}^{-1}$ is the inverse of the evolution operator from $t=0$ to $t=\mathcal{T}$. Then with the use of the conservation of probability along the trajectory [16], the ratio in Eq. (13) between $\rho[\Gamma_s(\mathcal{T}),\zeta,\mathcal{T}]$ and $\tilde{\rho}[\Theta\Gamma_s(\mathcal{T}),\Theta\zeta,\mathcal{T}]$ is calculated as

$$\begin{aligned}\frac{\rho(\Gamma_s(\mathcal{T}),\zeta,\mathcal{T})}{\tilde{\rho}(\Theta\Gamma_s(\mathcal{T}),\Theta\zeta,\mathcal{T})} &= \frac{\rho(\Gamma_s(0),\mathcal{V}^{-1}\zeta,0)}{\tilde{\rho}(\Theta\Gamma_s(\mathcal{T}),\Theta\zeta,\mathcal{T})}e^{DN\int_0^{\mathcal{T}}\zeta(t)dt} \\ &= \frac{\rho_s(\Gamma_s(0),0)}{\tilde{\rho}_s(\Theta\Gamma_s(\mathcal{T}),\mathcal{T})}e^{\frac{H_r[\zeta(\mathcal{T})]-H_r[\zeta(0)]+DNk_BT\int_0^{\mathcal{T}}\zeta(t)dt}{k_BT}}. \quad (18)\end{aligned}$$

The nontrivial factor in the exponent is the integral of the friction, which is absent in the Hamiltonian dynamics. From Eqs. (18) and (24), the ratio of the conditional probabilities is

$$\begin{aligned}&\frac{P_F[\Gamma_s(t)|\Gamma_s(0)]}{P_R[\Theta\Gamma_s(\mathcal{T}-t)|\Theta\Gamma_s(\mathcal{T})]} \\ &= \frac{\tilde{\rho}_s[\Theta\Gamma_s(\mathcal{T}),\mathcal{T}]}{\rho_s[\Gamma_s(0),0]} \\ &\quad\times\left\langle\frac{\rho_s[\Gamma_s(0),0]}{\tilde{\rho}_s[\Theta\Gamma_s(\mathcal{T}),\mathcal{T}]}e^{\frac{H_r[\zeta(\mathcal{T})]-H_r[\zeta(0)]+DNk_BT\int_0^{\mathcal{T}}\zeta(t)dt}{k_BT}}\right\rangle_R \\ &= \left\langle e^{\frac{H_r[\zeta(\mathcal{T})]-H_r[\zeta(0)]+DNk_BT\int_0^{\mathcal{T}}\zeta(t)dt}{k_BT}}\right\rangle_R. \quad (19)\end{aligned}$$

In the second equality, $\frac{\rho_s[\Gamma_s(0),0]}{\tilde{\rho}_s[\Theta\Gamma_s(\mathcal{T}),\mathcal{T}]}$ is extracted from the average, since it does not depend on the trajectory. The exponent of Eq. (19) is regarded as the heat divided by k_BT [5]. For completeness, let us summarize the definition of the work $W[\Gamma_s(t)]$ and heat $Q[\Gamma_s(t)]$ in Ref. [5]. We use the same definition of the work done on the system as Refs. [15–17] by the energy change due to the explicit time dependence of Hamiltonian

$$W[\Gamma_s(t)] \equiv \int_0^{\mathcal{T}}\frac{\partial H_s(\Gamma_s(t),t)}{\partial t}dt. \quad (20)$$

Since only the system Hamiltonian is explicitly time dependent, this is also equal to

$$\begin{aligned}&W[\Gamma_s(t)] \\ &= \int_0^{\mathcal{T}}\frac{dH(\Gamma(t),t)}{dt}dt - \int_0^{\mathcal{T}}\frac{\partial H(\Gamma(t),t)}{\partial\Gamma(t)}\dot{\Gamma}(t)dt \\ &= H(\Gamma(\mathcal{T}),\mathcal{T}) - H(\Gamma(0),0) + DNk_BT\int_0^{\mathcal{T}}\zeta(t)dt. \quad (21)\end{aligned}$$

Thus, by requiring the first law of thermodynamics, the heat is equal to

$$Q[\Gamma_s(t)] = -\{H_r[\zeta(\mathcal{T})] - H_r[\zeta(0)]\} - DNk_BT\int_0^{\mathcal{T}}\zeta(t)dt. \quad (22)$$

One nice property of the thermostatted dynamics is that there is a clear separation of the system and reservoir, and heat is defined as a functional of the system state trajectory without a kind of weak coupling condition. The strong interaction to the thermostat is advantageous for actual computational purpose. It is nontrivial that Eqs. (13) and (24) hold despite such a property of the thermostat. Note that the so-obtained thermodynamic entropy change of the reservoir reasonably consists of the energy change divided by the temperature and *phase volume contraction*

$$\begin{aligned}\Delta S_r &\equiv -\frac{Q[\Gamma_s(t)]}{T} \\ &= \frac{H_r[\zeta(\mathcal{T})] - H_r[\zeta(0)]}{T} + DNk_B\int_0^{\mathcal{T}}\zeta(t)dt, \quad (23)\end{aligned}$$

where the second term is caused by the friction, and proportional to the number of degrees of freedom DN. Equations (19) and (22) amount to the microscopic reversibility for thermostatted system with an arbitrary relaxation time of kinetic energy

$$\begin{aligned}\frac{P_F[\Gamma_s(t)|\Gamma_s(0)]}{P_R[\Theta\Gamma_s(\mathcal{T}-t)|\Theta\Gamma_s(\mathcal{T})]} &= \langle e^{-\beta Q[\Gamma_s(t)]}\rangle_R \\ &= e^{-\beta Q[\Gamma_s(t)]}, \quad (24)\end{aligned}$$

where $Q[\Gamma_s(t)]$ is fully determined by the system state trajectory, and thus extracted from the average $\langle\rangle_R$ [6]. Indeed, in the first law $W[\Gamma_s(t)] + Q[\Gamma_s(t)] = H_s[\Gamma_s(\mathcal{T}),\mathcal{T}] - H_s[\Gamma_s(0),0]$ the work done $W[\Gamma_s(t)]$ and the energy change of the system are determined by the system state trajectory.

V. SUMMARY

We have derived a rigorous universal relation Eq. (13) based on the microscopic reversibility for thermostatted dynamics and conservation of probability. Such an expression of the microscopic reversibility formally contains generalized detailed balance [7], nonequilibrium work relation, and Crooks relation for the heat. Remarkably, the expression is exactly the same as Hamiltonian dynamics despite the presence of the dissipation. Especially, the clear separation between the system and reservoir is advantageous to provide thermodynamic relations for open systems. This allows coexistence of the thermodynamic properties of open systems and the strong interaction to the reservoir. The latter is essential for the computational purposes. These facts are nontrivial, since the thermostatted dynamics would not be derived from Hamiltonian dynamics as a coarse graining or other special limiting cases.

Indeed, the lack of Hamiltonian amounts to the ambiguity of the definition of the total energy, and therefore the definition of the heat is nontrivial. Then we showed that the definition of the reservoir "Hamiltonian" in Refs. [3,12,15] actually makes the thermostatted dynamics consistent with thermodynamics. Especially, we derived exactly the Crooks relation under the operationally natural definition of the work and heat calculated from the first law of thermodynamics [5]. In this way, we showed that such a definition of the heat is thermodynamically meaningful.

ACKNOWLEDGMENT

This work is financially supported by JSPS Research Program under Grant No. 22·7744.

[1] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, Oxford, 2001).
[2] D. Andrieux, P. Gaspard, T. Monnai, and S. Tasaki, New. J. Phys. **11**, 043014 (2009).
[3] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[4] J. M. Thijssen, *Computational Physics* (Cambridge University Press, Cambridge, 1999).
[5] M. Esposito, and T. Monnai, J. Phys. Chem. B **115**, 5144 (2011).
[6] The friction coefficient mimics a large reservoir, which is always in the same equilibrium state.
[7] T. Monnai, J. Stat. Phys. **149**, 1058 (2012).
[8] T. Monnai, J. Phys. A **45**, 051001 (2012).
[9] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999); **61**, 2361 (2000).
[10] A. Imparato and L. Peliti, Phys. Rev. E **74**, 026106 (2006).
[11] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, J. Stat. Mech. (2006) P08001.
[12] B. L. Holian, Phys. Rev. A **33**, 1152 (1986).
[13] In Eq. (8), the product of $N\delta$ functions is well-defined for finite N.
[14] This condition is equivalent to having the corresponding forward trajectory under the forward protocol.
[15] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997); Phys. Rev. E **56**, 5018 (1997).
[16] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007).
[17] K. Sekimoto, Prog. Theor. Phys. Suppl. **130**, 17 (1998).
[18] C. Jarzynski, J. Stat. Phys. **98**, 77 (2000).