

Microscopic description of the equality between violation of fluctuation-dissipation relation and energy dissipation

Hiroshi Teramoto* and Shin-ichi Sasa†

Department of Pure and Applied Sciences, University of Tokyo, Komaba, Tokyo 153-8902, Japan

(Received 22 September 2005; published 9 December 2005; publisher error corrected 13 December 2005)

In systems far from equilibrium, the fluctuation-dissipation relation is violated due to the lack of detailed balance. Recently, for a class of Langevin equations, it has been proved that this violation is related to energy dissipation as an equality [Phys. Rev. Lett. **95**, 130602 (2005)]. We provide a microscopic description of this equality by studying a nonequilibrium colloidal system on the basis of classical mechanics with some physical assumptions.

DOI: [10.1103/PhysRevE.72.060102](https://doi.org/10.1103/PhysRevE.72.060102)

PACS number(s): 05.40.-a, 05.20.Gg

The construction of nonequilibrium statistical mechanics that is useful for systems far from equilibrium is a fundamental problem in theoretical physics. Let us recall that linear response theory had been formulated for the microscopic description of universal relations established by Einstein, Nyquist, and Onsager [1]. Thus, it is a significant step to provide a microscopic description of formulas derived phenomenologically for systems far from equilibrium.

Recently, an interesting equality relating the violation of the fluctuation-dissipation relation (FDR) with energy dissipation has been found for Langevin equations under non-equilibrium conditions [2,3]. Here, note that the FDR is a fundamental relation proved in a linear response regime around equilibrium states [1], but it is violated in systems far from equilibrium [4]. Recent studies have revealed that the idea of effective temperature is useful to characterize the FDR violation for glassy systems [5–7] and steady state systems [8–10]. In contrast to these studies, this equality claims that the FDR violation is characterized by an energetic quantity.

The equality has been proved for a wide class of Langevin systems including many-body systems, time-dependent potential systems, and systems in contact with many heat reservoirs [3]. We then search for such an equality in more general nonequilibrium systems that are not necessarily described by a Langevin equation. To this end, we start with investigating a nonequilibrium colloidal system on the basis of classical mechanics. Because it is highly expected that the motion of a colloidal particle is described by a Langevin equation, the FDR violation should be related to the energy dissipation even if we describe the system on the basis of classical mechanics. Here, the energy dissipation is given by the energy transfer from the center of mass of the colloidal particle to the other mechanical degrees of freedom in the classical mechanical description. In this paper, we derive an expression of the FDR violation for the classical mechanical description of the system far from equilibrium. By considering the physical conditions of the system, we rederive the equality reported in Ref. [2] from the expression of the FDR violation.

Model. Specifically, we study a system of one colloidal particle suspended in a three-dimensional liquid confined in a region where $-L/2 \leq y \leq L/2$ and $-L/2 \leq z \leq L/2$. Periodic boundary conditions are imposed in the x direction. Let $Z = (\mathbf{R}, \mathbf{P})$ be the position and momentum of the center of mass of the colloidal particle. The colloidal particle is driven by an external field $\mathbf{E}(t) = (E_0 + \epsilon f^p(t), 0, 0)$ and is subject to a spatially periodic potential $U_0(\mathbf{R})$, where E_0 is a constant force to realize a nonequilibrium steady state and $\epsilon f^p(t)$ is a small probe force to investigate the response of the system in the steady state.

The center of mass of the colloidal particle interacts with the other mechanical degrees of freedom of the colloid-liquid system. Their dynamical degrees of freedom are represented by $Y = (\mathbf{r}_1, \mathbf{p}_1, \dots, \mathbf{r}_N, \mathbf{p}_N)$. Furthermore, we introduce thermostated walls as a network of boundary particles, whose dynamical degrees of freedom are represented by $B = (\mathbf{u}_1, \mathbf{g}_1, \dots, \mathbf{u}_{N'}, \mathbf{g}_{N'})$. These boundary particles are localized around the boundaries given by $y = \pm L/2$ and $z = \pm L/2$, and interact with both the colloidal particle and the molecules in the liquid.

All the interaction potentials including $U_0(\mathbf{R})$ are represented by a Hamiltonian $H(Z, Y, B)$. Note that a constant force E_0 cannot be described by the Hamiltonian because E_0 is regarded as a nonpotential force due to the periodic boundary condition in the x direction. The result presented below does not depend on the details of the system. We assume that in the equilibrium case ($E_0 = 0$), the statistical properties of a variable set (Z, Y, B) are described by a canonical distribution with temperature T of the thermostated walls. We also assume that there exists a steady distribution when $E_0 \neq 0$. In order to control the temperature of a finite system, we introduce a Nose-Hoover thermostat only for the boundary particles [11]. The thermostat may introduce an unphysical effect, but we expect that the influence of the thermostat vanishes in the limit $L, N \rightarrow \infty$ with N/L^3 fixed. Furthermore, when the mathematical rigor is not critically taken into account, the analysis developed below can be applied to a Hamiltonian system without a thermostat in this limit.

In this mechanical system, the motion of all the particles is described by

$$\frac{d\mathbf{P}}{dt} = \mathbf{E} - \frac{\partial H}{\partial \mathbf{R}}, \quad (1)$$

*Electronic address: teramoto@mns2.c.u-tokyo.ac.jp

†Electronic address: sasa@jiro.c.u-tokyo.ac.jp

$$\frac{d\mathbf{p}_i}{dt} = -\frac{\partial H}{\partial \mathbf{r}_i}, \quad (2)$$

$$\frac{d\mathbf{g}_i}{dt} = -\frac{\partial H}{\partial \mathbf{u}_i} - \lambda_i \mathbf{g}_i, \quad (3)$$

$$\tau \frac{d\lambda_i}{dt} = \frac{\mathbf{g}_i^2}{m_i^b} - 3T, \quad (4)$$

with $\mathbf{P} = M d\mathbf{R}/dt$, $\mathbf{p}_i = m_i d\mathbf{r}_i/dt$, and $\mathbf{g}_i = m_i^b d\mathbf{u}_i/dt$, where M , m_i , and m_i^b are the masses of the corresponding variables. We express these equations as

$$\frac{d\Gamma}{dt} = \Phi(\Gamma, t) + \epsilon \Phi_1(t), \quad (5)$$

where $\Gamma = (Z, Y, B, \lambda)$ with $\lambda = (\lambda_1, \dots, \lambda_{N'})$.

Formal analysis. We consider the time-dependent distribution function $f(\Gamma, t)$ with an initial condition that $f(\Gamma, t_0) = f_c(\Gamma)$, where

$$f_c(\Gamma) = \frac{1}{\Xi} e^{-\beta H_0(\Gamma) - \beta \tau \sum_{i=1}^{N'} \lambda_i^2/2}, \quad (6)$$

with $H_0(\Gamma) \equiv H(\Gamma) - U_0(\mathbf{R})$. Ξ is the normalization constant. The initial condition represents the canonical distribution with respect to $H_0(\Gamma)$, not to $H(\Gamma)$. This choice is rather technical, but we found that it gives the simplest derivation of the equality reported in Ref. [2]. The distribution function $f(\Gamma, t)$ satisfies the equation

$$\frac{\partial f(\Gamma, t)}{\partial t} = -\frac{\partial}{\partial \Gamma} (\Phi(\Gamma) f(\Gamma, t)) - \epsilon \frac{\partial}{\partial \Gamma} (\Phi_1(t) f(\Gamma, t)). \quad (7)$$

Note that $f_c(\Gamma)$ is the stationary solution of (7) when $\mathbf{E}(t) = 0$ and $U_0(\mathbf{R}) = 0$.

We first set

$$f(\Gamma, t) = f_c(\Gamma) e^{\beta \phi(\Gamma, t)}. \quad (8)$$

Substituting this expression into (7), we obtain

$$\begin{aligned} \frac{\partial \phi(\Gamma, t)}{\partial t} = & -\Phi(\Gamma) \frac{\partial \phi(\Gamma, t)}{\partial \Gamma} + \mathbf{F}(\mathbf{R}) \cdot \frac{\partial H_0}{\partial \mathbf{P}} + \epsilon f^{\text{p}}(t) \frac{\partial H_0}{\partial P_x} \\ & - \epsilon f^{\text{p}}(t) \frac{\partial \phi}{\partial P_x}, \end{aligned} \quad (9)$$

where $\mathbf{F}(\mathbf{R})$ is defined as

$$\mathbf{F}(\mathbf{R}) \equiv E_0 \mathbf{e}_x - \frac{\partial U_0(\mathbf{R})}{\partial \mathbf{R}}. \quad (10)$$

We solve (9) formally as

$$\phi(\Gamma, t) = \int_{t_0}^t ds e^{\Lambda(s-t)} \left[\mathbf{F}(\mathbf{R}) \cdot \frac{\partial H_0}{\partial \mathbf{P}} + \epsilon f^{\text{p}}(s) \frac{\partial H_0}{\partial P_x} - \epsilon f^{\text{p}}(s) \frac{\partial \phi}{\partial P_x} \right], \quad (11)$$

where we have defined

$$\Lambda \equiv \Phi(\Gamma) \frac{\partial}{\partial \Gamma}. \quad (12)$$

In the argument below, we fix t and denote the solution of (5) with $\epsilon=0$, which satisfies $\Gamma(t)=\Gamma$, as $\Gamma(s)$ with $t_0 \leq s \leq t$. That is, $\Gamma(s)$ is regarded as a function of Γ . Then, for an arbitrary function $A(\Gamma)$, we obtain

$$\frac{dA(\Gamma(s))}{ds} = \Lambda A(\Gamma') \Big|_{\Gamma'=\Gamma(s)}. \quad (13)$$

Therefore, noting $A(\Gamma(t))=A(\Gamma)$, we write

$$e^{\Lambda(s-t)} A(\Gamma) = A(\Gamma(s)), \quad (14)$$

where the right-hand side is regarded as a function of Γ according to the rule mentioned above. Using this expression, we rewrite (11) as

$$\begin{aligned} \phi(\Gamma, t) = & \int_{t_0}^t ds \left[\mathbf{F}(\mathbf{R}(s)) \cdot \mathbf{V}(s) + \epsilon f^{\text{p}}(s) \frac{\partial H_0(\Gamma(s))}{\partial P_x(s)} \right. \\ & \left. - \epsilon f^{\text{p}}(s) \frac{\partial \phi(\Gamma(s), t)}{\partial P_x(s)} \right], \end{aligned} \quad (15)$$

where we have defined $\mathbf{V}(s) \equiv \mathbf{P}(s)/M$.

Now, let us expand $\phi(\Gamma, t)$ as

$$\phi(\Gamma, t) = \phi^{(0)}(\Gamma, t) + \epsilon \phi^{(1)}(\Gamma, t) + O(\epsilon^2). \quad (16)$$

Substituting this expression into (15), we arrange the terms according to the powers of ϵ . From the terms independent of ϵ , we obtain

$$\phi^{(0)}(\Gamma, t) = W(\Gamma, t_0, t; t) \quad (17)$$

with

$$W(\Gamma, t_0, s; t) = \int_{t_0}^s d\mathbf{R}(s') \cdot \mathbf{F}(\mathbf{R}(s')), \quad (18)$$

where $W(\Gamma, t_0, s; t)$ represents the accumulated work done by \mathbf{F} during the time interval $[t_0, s]$ for the trajectory satisfying $\Gamma(t)=\Gamma$. The expression of the distribution function (8) with (17) and (18) is similar to that proposed by Zubarev [12] and McLennan [13]. Next, the terms proportional to ϵ yield

$$\phi^{(1)}(\Gamma, t) = \int_{t_0}^t ds f^{\text{p}}(s) \left(V_x(s) - \frac{\partial W(\Gamma, t_0, s; t)}{\partial P_x(s)} \right). \quad (19)$$

Furthermore, the average of $A(\Gamma)$ by the distribution function $f(\Gamma, t)$,

$$\langle A(\Gamma(t)) \rangle_{\epsilon, t_0} \equiv \int d\Gamma f(\Gamma, t) A(\Gamma), \quad (20)$$

is expanded in ϵ as

$$\langle A(\Gamma(t)) \rangle_{\epsilon, t_0} = \langle A(\Gamma(t)) \rangle_{t_0}^{(0)} + \epsilon \langle A(\Gamma(t)) \rangle_{t_0}^{(1)} + O(\epsilon^2). \quad (21)$$

Then, from (8), (16), and (19), we obtain

$$\begin{aligned} \langle V_x(t) \rangle_{t_0}^{(1)} = & \beta \int_{t_0}^t ds f^{\text{p}}(s) \langle V_x(t) V_x(s) \rangle_{t_0}^{(0)} - \beta \int_{t_0}^t ds f^{\text{p}}(s) \\ & \times \left\langle V_x(t) \frac{\partial W(\Gamma(t), t_0, s; t)}{\partial P_x(s)} \right\rangle_{t_0}^{(0)}. \end{aligned} \quad (22)$$

Here, we remark on the steady state of the system. It can be expected that $\langle A(\Gamma(t)) \rangle_{\epsilon, t_0}$ becomes an averaged value in the steady state when we take the limit $t_0 \rightarrow -\infty$. However, $f(\Gamma, t)$ itself is divergent in this limit, as seen from (17) and (18). In this paper, we consider the limit for the averaged quantities, but we do not study the singularity of the distribution function itself.

Taking this into consideration, setting $t_0 \rightarrow -\infty$ in (22), we obtain the formula

$$C(t-s) = TR(t-s) + \lim_{t_0 \rightarrow -\infty} \left\langle V_x(t) \frac{\partial W(\Gamma(t), t_0, s; t)}{\partial P_x(s)} \right\rangle_{t_0}^{(0)} \quad (23)$$

for $t > s$. In this formula, the time correlation function $C(t-s)$ and the response function $R(t-s)$ are defined as

$$\langle V_x(t) \rangle_{-\infty}^{(1)} = \int_{-\infty}^t ds R(t-s) f^p(s), \quad (24)$$

$$C(t-s) = \langle V_x(t) V_x(s) \rangle_{-\infty}^{(0)}. \quad (25)$$

We also define $R(t) = 0$ for $t < 0$ from the causality.

To this point, no approximation is involved. Then, in the equilibrium case ($E_0 = 0$), we derive the FDR

$$C(t-s) = TR(t-s) \quad (26)$$

for $t > s$. This derivation is straightforward when we note that $W(\Gamma(t), t_0, s; t) = -U_0[\mathbf{R}(s)] + U_0[\mathbf{R}(t_0)]$ in this case. Here, we have used

$$\lim_{t_0 \rightarrow -\infty} \left\langle V_x(t) \frac{\partial U_0[\mathbf{R}(t_0)]}{\partial P_x(s)} \right\rangle_{t_0}^{(0)} = 0. \quad (27)$$

On the other hand, when $E_0 \neq 0$, the second term on the right-hand side of (23) takes a finite (nonzero) value in general. Thus, (23) provides an expression of the FDR violation in the mechanical description.

Physical consideration. Let us estimate the quantity $\partial W(\Gamma(t), t_0, s; t) / \partial P_x(s)$ in (23) for the system considered in this study. Our estimation is based on the two important assumptions on the time scales. First, let τ_m be the slowest time scale of phenomena that a set of variables (Y, B, λ) exhibits. This time scale τ_m is expected to be much smaller than the relaxation time of the velocity of the colloidal particle, which is denoted as τ_M . Considering this physical expectation, we assume that there exists a time scale Δ_1 satisfying [14]

$$\tau_m \ll \Delta_1 \ll \tau_M. \quad (28)$$

We choose such a time scale Δ_1 and hereinafter fix it. Using this Δ_1 , we define the time-averaged quantity as

$$\overline{A(\Gamma(s))} \equiv \frac{1}{\Delta_1} \int_{s-\Delta_1/2}^{s+\Delta_1/2} ds' A(\Gamma(s')). \quad (29)$$

The second assumption on the time scale of the system is that the relaxation time of the velocity τ_M is much smaller than the typical time scale of the position variation τ_p that is determined by the characteristic length of the potential

$U_0(\mathbf{R})$. That is, we can choose a time scale Δ_2 satisfying

$$\tau_M \ll \Delta_2 \ll \tau_p. \quad (30)$$

Now, we apply the perturbation $\delta V_x(s)$ to the phase space point $\Gamma(s)$. As a result of the perturbation, the trajectory $\Gamma(s')$ with $s' \leq s$ changes to $\Gamma(s') + \delta\Gamma(s')$, and this change yields the additional work $\delta W(\Gamma(t), t_0, s; t)$ from (18). It seems reasonable to assume that the relaxation time of the work rate $\mathbf{V}(s') \cdot \mathbf{F}(\mathbf{R}(s'))$ is of the order of τ_M . Thus, from the second assumption (30), we can estimate

$$\delta W(\Gamma(t), t_0, s; t) \simeq \int_{s-\Delta_2}^s ds' \delta(\mathbf{V}(s') \cdot \mathbf{F}(\mathbf{R}(s'))). \quad (31)$$

Here, assuming $\overline{\mathbf{V} \cdot \mathbf{F}(\Gamma(s'))} \simeq \overline{\mathbf{V}(s')} \cdot \overline{\mathbf{F}(\mathbf{R}(s'))}$, we write

$$\mathbf{V}(s') \cdot \mathbf{F}(\mathbf{R}(s')) = \overline{\mathbf{V}(s')} \cdot \overline{\mathbf{F}(\mathbf{R}(s'))} + \eta(s'), \quad (32)$$

where $\eta(s)$ represents the *fast part* of the work rate fluctuation whose time scale is much smaller than Δ_1 . Using this, we express

$$\delta W(\Gamma, t_0, s; t) \simeq \int_{s-\Delta_2}^s ds' \delta(\overline{\mathbf{V}(s')} \cdot \overline{\mathbf{F}(\mathbf{R}(s'))} + \eta(s')). \quad (33)$$

From the second assumption (30), $\overline{\mathbf{F}(\mathbf{R}(s'))}$ in the integrand can be replaced with $\overline{\mathbf{F}(\mathbf{R}(s))}$, and the estimation

$$\delta \overline{\mathbf{V}(s')} = (e^{-(s-s')/\tau_M} \delta V_x(s), 0) \quad (34)$$

seems reasonable in the time interval $[s-\Delta_2, s]$. Thus, the first term in (33) can be rewritten as

$$\overline{\mathbf{F}(\mathbf{R}(s))} \int_{s-\Delta_2}^s ds' \delta \overline{\mathbf{V}(s')} \simeq F_x[\overline{\mathbf{R}(s)}] \delta V_x(s) \tau_M. \quad (35)$$

From this result, we obtain the following expression:

$$\frac{\delta W(\Gamma, t_0, s; t)}{\partial P_x(s)} \simeq \frac{F_x[\overline{\mathbf{R}(s)}] \tau_M}{M} + \int_{s-\Delta_2}^s ds' \frac{\partial \eta(s')}{\partial P_x(s)}. \quad (36)$$

Recalling the definition of η in (32) and the first assumption (28), we expect that the second term of (36) has negligible correlation with $V_x(t)$. Further, when we assume that a friction force for the colloidal particle is given by $-\gamma \mathbf{V}$, γ is estimated as M/τ_M . Using these, the substitution of (36) into (23) leads to

$$C(t-s) = TR(t-s) + \frac{1}{\gamma} \langle V_x(t) F_x(s) \rangle_{-\infty}^{(0)} \quad (37)$$

for $t > s$.

Let us define the Fourier transform of $C(t-s)$ as

$$\tilde{C}(\omega) \equiv \int_{-\infty}^{\infty} dt' e^{i\omega t'} C(t'). \quad (38)$$

Similarly, we define $\tilde{R}(\omega)$, while $\tilde{R}'(\omega)$ denotes the real part of $\tilde{R}(\omega)$. Then, from (37), we obtain.

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} [\tilde{C}(\omega) - 2T\tilde{R}'(\omega)] = \frac{1}{\gamma} J, \quad (39)$$

where $J = \langle V_x F_x \rangle_{-\infty}^{(0)}$ is the energy transfer rate from the center of mass of the colloidal particle to the other degrees of freedom. This energy transfer rate is interpreted as the energy dissipation ratio in the description of the center of mass of the colloidal particle. Thus, the equality given in (39) relates the FDR violation with energy dissipation, as presented for a Langevin equation in Ref. [2]. In this manner, we have re-derived the equality on the basis of the classical mechanical description.

Discussion. It might be possible to develop a theory for formalizing the above-mentioned physical consideration. In such a theory, the equality given in (39) might be derived systematically from microscopic dynamics by a calculation technique using the separation of time scales. The construction of the theory is a future research subject.

Related to this subject, we remark that there exist many different expressions of the distribution function. For example, when we assume an initial condition involving H instead of H_0 in (6), we obtain a different expression of the distribution function. However, in this case, we find that it is difficult to relate the obtained expression of the FDR violation with the result for the Langevin equation. Note that the steady distributions for both the initial conditions should be identical when the limit $t_0 \rightarrow -\infty$ is considered. This implies that there is an expression that can be treated in a simple manner. The clarification of this might provide a key step in

the systematic derivation of the equality given in (39).

A more important question is whether FDR violation can be expressed by a form useful for systems in which a Langevin description is not effective. It should be noted that one can derive an expression of FDR violation for any mechanical system in a similar manner to that used for deriving (23). The examples of mechanical systems include electric conduction systems, sheared systems, and heat conduction systems. However, as mentioned above, the obtained expression might have no direct relation with measurable energetic quantities. Thus, it is necessary to find a condition under which the FDR violation takes a physically useful form.

Finally, we remark on a role of the Nosé-Hoover thermostat. This thermostat is considered merely as an illustration of a mechanism to insure that the system reaches a steady state without affecting the formal derivation of the main result of the paper. It is interesting to study systems without the Nosé-Hoover thermostat.

In conclusion, we provide a microscopic description for the equality given in (39) by analyzing the Hamiltonian equation with the Nosé-Hoover thermostat at the boundaries. By examining questions arising from this study, we wish to obtain a deep understanding of nonequilibrium systems.

The authors thank K. Hayashi, T. Harada, and H. Tasaki for useful discussions. This work was supported by a grant from the Ministry of Education, Science, Sports and Culture of Japan (Grant No. 16540337) and a grant from Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists (Grant No. 1711363).

-
- [1] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics* (Springer-Verlag, Berlin, 1991).
- [2] T. Harada and S.-i. Sasa, Phys. Rev. Lett. **95**, 130602 (2005).
- [3] T. Harada and S.-i. Sasa, e-print cond-mat/0510723.
- [4] T. Harada and K. Yoshikawa, Phys. Rev. E **69**, 031113 (2004).
- [5] L. F. Cugliandolo, J. Kuchan, and L. Peliti, Phys. Rev. E **55**, 3898 (1997).
- [6] A. Crisanti and F. Ritort, J. Phys. A **36** R181, (2003).
- [7] L. Berthier and J.-L. Barrat, Phys. Rev. Lett. **89**, 095702 (2002).
- [8] A. B. Kolton *et al.*, Phys. Rev. Lett. **89**, 227001 (2002).
- [9] K. Hayashi and S.-i. Sasa, Phys. Rev. E **69**, 066119 (2004).
- [10] K. Hayashi and S.-i. Sasa, Phys. Rev. E **71**, 046143 (2005).
- [11] S. Nosé, Prog. Theor. Phys. **103**, 117, (1991).
- [12] D. N. Zubarev, *Nonequilibrium Statistical Thermodynamics* (Consultants Bureau, New York, 1974)
- [13] J. A. McLennan, *Introduction to Nonequilibrium Statistical Mechanics* (Prentice Hall, Englewood Cliffs, NJ, 1990).
- [14] This assumption implies that hydrodynamic effects described by the variable Y are not taken into account.