CrossMark

Universal Form of Stochastic Evolution for Slow Variables in Equilibrium Systems

Masato Itami[1] (ORCID) · **Shin-ichi Sasa**[2]

Abstract Nonlinear, multiplicative Langevin equations for a complete set of slow variables in equilibrium systems are generally derived on the basis of the separation of time scales. The form of the equations is universal and equivalent to that obtained by Green. An equation with a nonlinear friction term for Brownian motion turns out to be an example of the general results. A key method in our derivation is to use different discretization schemes in a path integral formulation and the corresponding Langevin equation, which also leads to a consistent understanding of apparently different expressions for the path integral in previous studies.

Keywords Nonlinear Langevin equation · Onsager theory · Large deviation theory · Path integral formulation

1 Introduction

We start with a brief review of Onsager theory [1–6]. Let $X = (X^i)_{i=1}^N$ be a complete set of unconstrained thermodynamic extensive variables of an isolated system. Onsager theory formulates the deterministic dynamics of the thermodynamic variables in relaxation processes to the equilibrium state. The time evolution is simply expressed as

$$\frac{\mathrm{d}X^i}{\mathrm{d}t} = \sum_j L^{ij} \frac{\partial S(X)}{\partial X^j}, \tag{1}$$

✉ Masato Itami
itami@fukui.kyoto-u.ac.jp

Shin-ichi Sasa
sasa@scphys.kyoto-u.ac.jp

1   Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan

2   Department of Physics, Kyoto University, Kyoto 606-8502, Japan

where $S(X)$ is the thermodynamic entropy of the system, $\partial S(X)/\partial X_j$ corresponds to the thermodynamic force, and L^{ij} is called the Onsager coefficient. The important consequence of Onsager theory is the reciprocity

$$L^{ij} = L^{ji}. \tag{2}$$

This nontrivial result was derived by studying fluctuations in equilibrium. Concretely, the fluctuation is assumed to be described by a Langevin equation,

$$\frac{\mathrm{d}X^i}{\mathrm{d}t} = \sum_j L^{ij} \frac{\partial S(X)}{\partial X^j} + \sum_j l^{ij} \xi^j, \tag{3}$$

with the Gaussian white noise satisfying $\langle \xi^i(t)\xi^j(t') \rangle = \delta(t - t')$. This assumption means that the most probable regression process for a given fluctuation is equivalent to the relaxation dynamics, which is referred to as the regression hypothesis. Furthermore, according to equilibrium statistical mechanics, the stationary probability density of X is

$$P_{\mathrm{eq}}(X) = \frac{1}{Z} \exp[S(X)], \tag{4}$$

where Z is the normalization constant. The time-reversibility of microscopic systems then provides the nontrivial relation

$$\sum_k l^{ik} l^{jk} = 2L^{ij}, \tag{5}$$

which leads to (2). The relation (5) is referred to as the fluctuation–dissipation relation of the second kind.

It should be noted that (1) is a nonlinear equation for X because $S(X)$ is not necessarily a quadratic function. In the argument above, L^{ij} is assumed to be independent of X. Because dependence of L^{ij} on X is expected in general cases, it is natural to consider generalized forms of (1) and (3). One approach assumes (3) with $L^{ij}(X)$ and $l^{ij}(X)$ as the starting equation, where a multiplication rule for $l^{ij}(X)$ and ξ^j is specified. Once a stochastic system is defined, the stationary distribution for X is determined. We then find that the stationary distribution is not given by (4) for any multiplication rule for $l^{ij}(X)$ and ξ^j. This means that there is no consistent description of (3)–(5) when dependence of L^{ij} on X is considered. The important thing here is that a generalization of (3) with $L^{ij}(X)$ is not obvious at all.

We can now describe the dynamics of X on the basis of a Hamiltonian system consisting of atoms and molecules. Suppose that X is a complete set of slow variables for the system. Examples include a complete set of unconstrained extensive variables in thermodynamics. Then, by using a separation of time scales, we may study the time evolution of X from the microscopic mechanical description. The result is that (3) is replaced by

$$\frac{\mathrm{d}X^i}{\mathrm{d}t} = \mathcal{J}^i_{\mathrm{rev}}(X) + \sum_j L^{ij}(X) \frac{\partial S(X)}{\partial X^j} + \sum_j \frac{\partial L^{ij}(X)}{\partial X^j} + \sum_j l^{ij}(X) \cdot \xi^j, \tag{6}$$

where the multiplication of $l^{ij}(X)$ and ξ^j is interpreted as Itô-type, and $\mathcal{J}^i_{\mathrm{rev}}$ is the so-called reversible term that does not contribute to changes in entropy.

There is a long history of studies on (6). In [7], Green derived the Fokker–Planck equation corresponding to (6) by combining phenomenological arguments with microscopic considerations. This was the genesis of (6). After this paper, many further studies were performed. For example, Green's derivation was improved in [8] where fewer assumptions were used. More formal studies under the microscopic description re-derived the Fokker–Planck equation using a projection-operator method [9,10] and a nonequilibrium statistical operator method

[11]. In another direction, the Fokker–Planck equation corresponding to (6) was also derived from a general Fokker–Planck equation by imposing a detailed balance condition [12,13]. Finally, the Langevin equation (6) was derived directly from Liouville's equation using a nonlinear projection operator method [14]. Thus (6) was well established by 1975.

However, the result (6) is less well known nowadays. There may be two reasons. First, Graham attempted to develop a co-variant description of nonlinear Onsager theory. Although this theory is complicated, many papers in this direction followed [15–19]. Unfortunately, we do not find a final answer in this direction, but more importantly, we consider such a generalization to not be necessary at all. Equation (6) is sufficient to be general and universal. Second, when X is a set of extensive variables in thermodynamics, the third term becomes higher-order than the second term from the estimates $S = O(\Omega)$ and $X^i = O(\Omega)$ for the system size Ω. Therefore, if one combines the system size expansion [20,21] in deriving the equation for slow variables, the third term does not appear for thermodynamically normal systems. Although this argument is correct, we emphasize that (6) can be a starting point for all systems, including small systems, once we identify a complete set of slow variables under equilibrium conditions. That is, in our opinion, (6) should be recognized as a fundamental equation for slow variables.

The main purpose of this paper is to re-derive (6) with particular emphasis of the separation of time scales and a universal asymptotic form of the probability density for time-averaged fluxes [22]. We first assume a complete set of slow variables. Let τ_{macro} be the shortest time scale of the slow variables and τ_{micro} be the largest time scale of the other dynamical variables. Then, from the separation of time scales $\tau_{\text{micro}} \ll \tau_{\text{macro}}$, we can find Δt such that $\tau_{\text{micro}} \ll \Delta t \ll \tau_{\text{macro}}$. This Δt plays two crucial roles in the derivation of the equation for slow variables. First, because $\tau_{\text{micro}} \ll \Delta t$, we can consider the central limit theorem for the time averaged flux as a universal form of the asymptotic behavior of the transition probability of the slow variables during a time interval Δt. The time reversibility in microscopic Hamiltonian systems provides a restriction on the transition probability. Second, because $\Delta t \ll \tau_{\text{macro}}$, this universal form of the transition probability leads to the path integral form of a stochastic system. This stochastic system is nothing but (6). This concept is quite natural and general. Indeed, one can interpret the arguments of Onsager and Green through this concept. Nevertheless, as far as we know, there is no explicit presentation of the derivation of (6) with the universal asymptotic form of the probability density for time-averaged fluxes and the path integral formulation under $\tau_{\text{micro}} \ll \Delta t \ll \tau_{\text{macro}}$. We thus expect that this paper will be instructive for understanding the universal form (6), and will also be useful for deriving the equation for slow variables even in systems out of equilibrium.

Here, we point out the difference between our and previous approaches. Our final goal is to establish a firm connection between a Langevin equation and a microscopic mechanical system. The previous studies [9–11,14] using a projection operator method or a nonequilibrium statistical operator method have the same motivation as ours. Their methods use some physical approximation (such as a markovian approximation) just before obtaining a Langevin equation. The validity of the approximation depends on observation time scales and details of a system, and their formulation is based on only an identity, which is useful but far from a physical principle. Thus, their assumptions are out of scope of the theories. Then, we aim to achieve our goal with physical principles. From this motivation, we use the central limit theorem with the separation of time scales for connecting a microscopic mechanical and mesoscopic stochastic description in a mathematically and physically clear way. This paper also differs from another type of derivation of the Langevin equation on the basis of arguments within stochastic processes [12,13]. Their derivation is self-contained and elegant, but arguments relating to microscopic descriptions are out of scope of their theory. As

a technical remark, we note that they used the Kolmogorov forward and backward equations for restricting the form of the Langevin equation by imposing a detailed balance condition, while we directly use the transition probabilities. Although we do not completely achieve our aim, we believe that it is important to show the outline of our approach even without a rigorous proof. Our approach is not simply another derivation of known results, but provides a new direction of future studies.

The remainder of this paper is organized as follows. In Sect. 2, as preliminaries for the argument, we review a path integral formulation for a discrete-time Langevin equation. As described above, the central limit theorem for time-averaged fluxes is closely related to the path integral formulation of stochastic processes. The technical difficulty in the argument arises from its complicated expression, which may be entirely associated with the ill-defined nature of the multiplication of some quantities. To make the argument as clear as possible, we study a path integral form for discrete-time Langevin equations while keeping the time interval dt finite. We then find relations between different expressions of the path integral forms as first derived by Wissel [23]. This recovers each of the correct but apparently different expressions for the path integral in [24–26]. In Sect. 3, we consider as a special case a nonlinear Langevin equation for the momentum $\boldsymbol{P}$ of a Brownian particle of mass M in a homogeneous environment of temperature T,

$$
\begin{aligned}
\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} &= -\frac{\gamma(\boldsymbol{P})}{M}\boldsymbol{P} + \sqrt{2\gamma(\boldsymbol{P})T} \odot \boldsymbol{\xi} \\
&= -\frac{\gamma(\boldsymbol{P})}{M}\boldsymbol{P} + T\nabla\gamma(\boldsymbol{P}) + \sqrt{2\gamma(\boldsymbol{P})T} \cdot \boldsymbol{\xi},
\end{aligned}
\tag{7}
$$

where $\odot$ denotes multiplication with the anti-Itô rule. Note that the Itô and anti-Itô rule will be explained in Sect. 2.1. Although (7) is well known, it has never been recognized as an example of (6) to the best of our knowledge. Indeed, we can derive (7) by using the central limit theorem with the separation of time scales. Then, in Sect. 4, we derive the general formula (6). Throughout this paper, the Boltzmann constant k_B is set to unity.

2 Preliminaries: Path Integral Formulation of Discrete Time Stochastic Systems

2.1 Model and Path Integral Formulation

Let $\boldsymbol{x} = (x^1, \ldots, x^N)$ be a collection of dynamical variables. We study the time evolution of $\boldsymbol{x}$ for a fixed time interval dt. We denote $x^i(ndt)$ by x_n^i, and we assume that $dx_n^i \equiv x_{n+1}^i - x_n^i$ satisfies

$$
dx_n^i = f^i(\boldsymbol{x}_n)\,dt + \sum_j g^{ij}(\boldsymbol{x}_n)\,dB_n^j,
\tag{8}
$$

where f^i and g^{ij} are smooth functions of $\boldsymbol{x}$, $\boldsymbol{B}(t)$ is a standard N-dimensional Wiener process [27], and $dB_n^i \equiv B^i(ndt + dt) - B^i(ndt)$ is a Gaussian white noise with mean zero and covariance $\langle dB_n^i dB_m^j \rangle = \delta^{ij}\delta_{nm}dt$. Because the short time interval dt can be considered to consist of shorter time intervals, we may use $dB_n^i dB_m^j = \delta^{ij}\delta_{nm}dt$ and ignore any $o(dt)$ terms in the Taylor expansion, which is known as Itô's lemma. Note that we obtain the Itô stochastic differential equation from (8) in the limit $dt \to 0$.

Denoting the probability density of finding $\boldsymbol{\chi}$ at time t by

$$P(\boldsymbol{\chi}, t) = \langle \delta (\boldsymbol{\chi} - \boldsymbol{x}_t) \rangle, \tag{9}$$

and using Itô's lemma and (8), we obtain the Fokker–Planck equation [27]

$$\frac{\partial}{\partial t} P(\boldsymbol{\chi}, t) = -\sum_i \frac{\partial}{\partial \chi^i} \left[f^i(\boldsymbol{\chi}) P(\boldsymbol{\chi}, t) \right] + \sum_{i,j} \frac{\partial}{\partial \chi^i} \frac{\partial}{\partial \chi^j} \left[G^{ij}(\boldsymbol{\chi}) P(\boldsymbol{\chi}, t) \right], \tag{10}$$

in the limit $dt \to 0$ with

$$G^{ij}(\boldsymbol{x}) = \frac{1}{2} \sum_k g^{ik}(\boldsymbol{x}) g^{jk}(\boldsymbol{x}). \tag{11}$$

We introduce a parameter α satisfying $0 \le \alpha \le 1$, and define

$$\bar{x}_n^i \equiv \alpha x_{n+1}^i + (1 - \alpha) x_n^i. \tag{12}$$

The purpose of this section is to express the transition probability $\mathcal{P}(\boldsymbol{x}_{n+1}|\boldsymbol{x}_n)$ from $\boldsymbol{x}_n$ to $\boldsymbol{x}_{n+1}$ in terms of $d\boldsymbol{x}_n$ and $\bar{\boldsymbol{x}}_n$. For any function $A(\boldsymbol{x})$ in the remainder of Sect. 2, we abbreviate $A(\bar{\boldsymbol{x}}_n)$ and $\partial A(\bar{\boldsymbol{x}}_n)/\partial \bar{x}_n^i$ to A and $\partial^i A$, respectively. We present the expression for $\mathcal{P}(\boldsymbol{x}_{n+1}|\boldsymbol{x}_n)$ and derive it in the next subsection. The transition probability is

$$\mathcal{P}(\boldsymbol{x}_{n+1}|\boldsymbol{x}_n) = \frac{1}{\sqrt{(4\pi dt)^N \det \mathsf{G}}} \exp \left[-\frac{dt}{4} \sum_{i,j} \Delta_n^i \left(G^{-1} \right)^{ij} \Delta_n^j \right.$$

$$\left. - \alpha \sum_i \partial^i f^i dt + \alpha^2 \sum_{i,j} \partial^i \partial^j G^{ij} dt \right], \tag{13}$$

with

$$\Delta_n^i \equiv \frac{dx_n^i}{dt} - f^i(\bar{\boldsymbol{x}}_n) + 2\alpha \sum_k \partial^k G^{ik}(\bar{\boldsymbol{x}}_n), \tag{14}$$

where $(G^{-1})^{ij}$ is the ij component of the inverse of the matrix $\mathsf{G} = (G^{ij})$. In [23], this expression for the transition probability was derived from the Fokker–Planck equation (10). Note that (8) itself does not depend on α. In Appendix, we check the normalization condition for the transition probability (13) for $N = 1$, and derive the Fokker–Planck equation (10) from the transition probability (13) in the limit $dt \to 0$.

Next, we consider the continuous-time limit of the transition probability. To avoid divergence of the prefactor $1/\sqrt{(4\pi dt)^N \det \mathsf{G}}$ in the transition probability (13), we rewrite (13) as

$$\mathcal{P}(\boldsymbol{x}_{n+1}|\boldsymbol{x}_n) = \int \frac{d^N \bar{\boldsymbol{p}}_n}{(2\pi)^N} \exp \left[dt \mathcal{L} \left(\boldsymbol{x}_{n+1}, \bar{\boldsymbol{p}}_n | \boldsymbol{x}_n \right) \right], \tag{15}$$

with

$$\mathcal{L} \left(\boldsymbol{x}_{n+1}, \bar{\boldsymbol{p}}_n | \boldsymbol{x}_n \right) = -\sum_{i,j} \bar{p}_n^i G^{ij}(\bar{\boldsymbol{x}}_n) \bar{p}_n^j - i \sum_i \bar{p}_n^i \Delta_n^i$$

$$- \alpha \sum_i \partial^i f^i(\bar{\boldsymbol{x}}_n) + \alpha^2 \sum_{i,j} \partial^i \partial^j G^{ij}(\bar{\boldsymbol{x}}_n), \tag{16}$$

where $\bar{\boldsymbol{p}}_n = (\bar{p}_n^1, \ldots, \bar{p}_n^N)$ is interpreted as the conjugate momentum of $\bar{\boldsymbol{x}}_n$. Using (15) repeatedly in each step and taking the limit $dt \to 0$, we can obtain the path integral for

the Langevin equation. Note that the Stratonovich convention ($\alpha = 1/2$) in (15) may be convenient for a perturbative analysis based on the path integral formulation because the Stratonovich convention conserves the chain rule of differential calculus.

Finally, we compare (13) and (15) with some previous studies [24–26]. Instead of (8), we consider

$$dx_n^i = \tilde{F}^i(\tilde{\boldsymbol{x}}_n)\, dt + \sum_j g^{ij}(\tilde{\boldsymbol{x}}_n)\, dB_n^j, \tag{17}$$

with

$$\tilde{x}_n^i \equiv \tilde{\alpha} x_{n+1}^i + (1 - \tilde{\alpha}) x_n^i, \tag{18}$$

where $\tilde{F}^i$ is some smooth function, and $\tilde{\alpha}$ is a parameter satisfying $0 \leq \tilde{\alpha} \leq 1$. Here, $\tilde{\alpha} = 0$, $1/2$, and 1 correspond to the Itô, Stratonovich, and anti-Itô convention, respectively, in the limit $dt \to 0$. Using Itô's lemma, we can rewrite (17) as

$$dx_n^i = \tilde{F}^i(\boldsymbol{x}_n)\, dt + \tilde{\alpha} \sum_{j,k} g^{kj}(\boldsymbol{x}_n) \frac{\partial g^{ij}(\boldsymbol{x}_n)}{\partial x_n^k} dt + \sum_j g^{ij}(\boldsymbol{x}_n)\, dB_n^j. \tag{19}$$

Thus, we can obtain the transition probability $\mathcal{P}(\boldsymbol{x}_{n+1}|\boldsymbol{x}_n)$ for (17) by using (13) with

$$f^i = \tilde{F}^i + \tilde{\alpha} \sum_{j,k} g^{kj} \partial^k g^{ij}. \tag{20}$$

Note that $\tilde{\alpha}$ may be different from α. When $\alpha = \tilde{\alpha}$ (15) with (20) is equivalent to the results given in [24,25]. When $(\alpha, \tilde{\alpha}) = (1/2, 1)$ (13) with (20) is equivalent to the results given in [26].

2.2 Derivation

We derive the transition probability (13) from (8) without using the Fokker–Planck equation (10). Using Itô's lemma, we first rewrite (8) in terms of $\bar{\boldsymbol{x}}_n$ as

$$dx_n^i = F^i dt + \sum_j g^{ij} dB_n^j, \tag{21}$$

with

$$F^i = f^i - \alpha \sum_{j,k} g^{kj} \partial^k g^{ij}. \tag{22}$$

Because dB_n^i is the Gaussian white noise with covariance $\langle dB_n^i dB_m^j \rangle = \delta^{ij} \delta_{nm} dt$, the probability density of $\{dB_n^i\}$ is given by

$$P\left(\left\{dB_n^i\right\}\right) = \frac{1}{(\sqrt{2\pi dt})^N} \exp\left[-\frac{1}{2dt} \sum_i \left(dB_n^i\right)^2\right]. \tag{23}$$

Using (8), $\boldsymbol{x}_{n+1}$ is uniquely determined by $\boldsymbol{x}_n$ and $\{dB_n^i\}$. Thus, we have

$$\mathcal{P}(\boldsymbol{x}_{n+1}|\boldsymbol{x}_n) = P\left(\left\{dB_n^i\right\}\right) |\det \mathcal{J}|, \tag{24}$$

where $\mathcal{J} = (J^{ij})$ is the Jacobian matrix defined by

$$J^{ij} \equiv \frac{\partial(dB_n^i)}{\partial x_{n+1}^j}. \tag{25}$$

We next calculate the determinant of the Jacobian matrix $\mathcal{J}$. Differentiating both sides of (21) with respect to x_{n+1}^l, we obtain

$$\sum_j g^{ij} J^{jl} = \mu^{il}, \tag{26}$$

where the matrix $\mathcal{M} = (\mu^{il})$ is given by

$$\mu^{il} = \delta^{il} - \alpha \partial^l F^i dt - \alpha \sum_j \partial^l g^{ij} dB_n^j. \tag{27}$$

Denoting the identity matrix of size N by I_N, we define the matrix $\widetilde{\mathcal{M}} = (\tilde{\mu}^{ij})$ by $\widetilde{\mathcal{M}} \equiv I_N - \mathcal{M}$. Using Itô's lemma, the determinant of the matrix $\mathcal{M}$ is

$$
\begin{aligned}
\det \mathcal{M} &= \det \left[\exp(\log \mathcal{M}) \right] \\
&= \exp \left[\mathrm{Tr}(\log \mathcal{M}) \right] \\
&= \exp \left[-\mathrm{Tr}\,\widetilde{\mathcal{M}} - \frac{1}{2}\mathrm{Tr}\,\widetilde{\mathcal{M}}^2 \right] \\
&= \exp \left[-\alpha \sum_i \partial^i F^i dt - \alpha \sum_{i,j} \partial^i g^{ij} dB_n^j - \frac{\alpha^2}{2} \sum_{i,j,k} \partial^i g^{jk} \partial^j g^{ik} dt \right].
\end{aligned} \tag{28}
$$

Because (26) leads to

$$\det \mathcal{G} \det \mathcal{J} = \det \mathcal{M}, \tag{29}$$

with $\mathcal{G} = (g^{ij})$, we obtain

$$\det \mathcal{J} = \frac{1}{\det \mathcal{G}} \exp \left[-\alpha \sum_i \partial^i F^i dt - \alpha \sum_{i,j} \partial^i g^{ij} dB_n^j - \frac{\alpha^2}{2} \sum_{i,j,k} \partial^i g^{jk} \partial^j g^{ik} dt \right]. \tag{30}$$

Substituting (23) and (30) into (24), we obtain

$$
\begin{aligned}
\mathcal{P}(x_{n+1}|x_n) &= \frac{1}{(\sqrt{2\pi dt})^N |\det \mathcal{G}|} \exp \left[-\frac{1}{2dt} \sum_i \left[dB_n^i + \alpha \sum_k \partial^k g^{ki} dt \right]^2 \right. \\
&\qquad \left. -\alpha \sum_i \partial^i F^i dt + \frac{\alpha^2}{2} \sum_{i,j,k} \left[\partial^i g^{ik} \partial^j g^{jk} - \partial^i g^{jk} \partial^j g^{ik} \right] dt \right] \\
&= \frac{1}{(\sqrt{2\pi dt})^N |\det \mathcal{G}|} \exp \left[-\frac{dt}{2} \sum_i \left[\sum_j (g^{-1})^{ij} \right. \right. \\
&\qquad \left. \left(\frac{dx_n^j}{dt} - F^j + \alpha \sum_{k,l} g^{jl} \partial^k g^{kl} \right) \right]^2 - \alpha \sum_i \partial^i F^i dt \\
&\qquad \left. + \frac{\alpha^2}{2} \sum_{i,j,k} \left[\partial^i g^{ik} \partial^j g^{jk} - \partial^i g^{jk} \partial^j g^{ik} \right] dt \right],
\end{aligned} \tag{31}
$$

where $(g^{-1})^{ij}$ is the ij component of the inverse of the matrix $\mathcal{G}$. Using (22), $\det \mathsf{G} = (\det \mathcal{G})^2 / 2^N$, and

$$g^{ik} \partial^j g^{jk} - g^{jk} \partial^j g^{ik} = \partial^j \left(g^{ik} g^{jk} \right) - 2 g^{jk} \partial^j g^{ik}, \tag{32}$$

we can rewrite (31) as (13).

3 Example: Brownian Motion

3.1 Setup

In this section, we study the motion of a single Brownian particle in a fluid (heat bath). We derive (7) under the assumption that the relaxation time of the momentum τ_{macro} is much larger than the time scales of the other degrees of freedom. We describe the system as a Hamiltonian system. The system consists of N bath particles of mass m and a Brownian particle of mass M in a cube of side length L. For simplicity, periodic boundary conditions are assumed. Let $(\boldsymbol{r}_i, \boldsymbol{p}_i)$ $(1 \leq i \leq N)$ be the position and momentum of the ith bath particle, and $(\boldsymbol{R}, \boldsymbol{P})$ be those of the Brownian particle. The collection of the positions and momenta of all particles is denoted by $\Gamma = (\boldsymbol{r}_1, \boldsymbol{p}_1, \ldots, \boldsymbol{r}_N, \boldsymbol{p}_N, \boldsymbol{R}, \boldsymbol{P})$, which represents the microscopic state of the system. For any state Γ, we denote its time reversal by Γ^*, namely, the state obtained by reversing all the momenta, and denote the time reversal of $\boldsymbol{P}$ by $\boldsymbol{P}^* = -\boldsymbol{P}$. For convenience, we denote the microscopic state excluding the momentum of the Brownian particle by $\tilde{\Gamma} = (\boldsymbol{r}_1, \boldsymbol{p}_1, \ldots, \boldsymbol{r}_N, \boldsymbol{p}_N, \boldsymbol{R})$.

The Hamiltonian of the system is given by

$$H(\Gamma) = \tilde{H}(\tilde{\Gamma}) + \frac{|\boldsymbol{P}|^2}{2M}, \tag{33}$$

with

$$\tilde{H}(\tilde{\Gamma}) = \sum_{i=1}^{N} \left[\frac{|\boldsymbol{p}_i|^2}{2m} + \sum_{j>i} \Phi_{\text{int}}\left(|\boldsymbol{r}_i - \boldsymbol{r}_j|\right) + \Phi_{\text{B}}\left(|\boldsymbol{r}_i - \boldsymbol{R}|\right) \right], \tag{34}$$

where Φ_{int} is a short-range interaction potential between two bath particles, and Φ_{B} is that between a bath particle and the Brownian particle. Then the Hamiltonian satisfies the time-reversal symmetry

$$H(\Gamma^*) = H(\Gamma). \tag{35}$$

For the Hamiltonian equations with a given state Γ at $t = 0$, Γ_t denotes the solution at time t. In this setup, energy is conserved, that is,

$$H(\Gamma_t) = H(\Gamma), \tag{36}$$

and Liouville's theorem that

$$\left| \frac{\partial \Gamma_t}{\partial \Gamma} \right| = 1, \tag{37}$$

holds. The total force acting on the Brownian particle is given by

$$\boldsymbol{F}(\Gamma) = -\frac{\partial H(\Gamma)}{\partial \boldsymbol{R}}$$
$$= -\sum_{i=1}^{N} \frac{\partial \Phi_{\text{B}}(|\boldsymbol{r}_i - \boldsymbol{R}|)}{\partial \boldsymbol{R}}. \tag{38}$$

For convenience, we abbreviate $A(\Gamma_t)$ to A_t for any function A. The equation of motion for the Brownian particle is

$$\frac{\mathrm{d}\boldsymbol{P}_t}{\mathrm{d}t} = \boldsymbol{F}_t. \tag{39}$$

We assume that the system in equilibrium is at temperature T. Suppose that the momentum of the Brownian particle is $\boldsymbol{P}_i$ at an initial time. Then the other mechanical state $\tilde{\Gamma}$ is sampled according to the probability density

$$\tilde{P}_{\text{eq}}(\tilde{\Gamma}) = \exp\left[-\frac{\tilde{H}(\tilde{\Gamma}) - \tilde{\Psi}_{\text{eq}}}{T}\right], \tag{40}$$

where $\tilde{\Psi}_{\text{eq}}$ is the normalization constant. The Hamiltonian equation determines the value of $\boldsymbol{P}$ at time t. By taking the average over initial realizations of $\tilde{\Gamma}$, we determine the probability density of $\boldsymbol{P} = \boldsymbol{P}_{\text{f}}$ at time t for a given $\boldsymbol{P}_{\text{i}}$ at time 0 in the form

$$\mathcal{P}_t\left(\boldsymbol{P}_{\text{f}}|\boldsymbol{P}_{\text{i}}\right) \equiv \int d\Gamma\, \tilde{P}_{\text{eq}}(\tilde{\Gamma})\delta\left(\boldsymbol{P} - \boldsymbol{P}_{\text{i}}\right)\delta\left(\boldsymbol{P}_t - \boldsymbol{P}_{\text{f}}\right). \tag{41}$$

It should be noted that

$$\int d^3\boldsymbol{P}_{\text{f}}\mathcal{P}_t\left(\boldsymbol{P}_{\text{f}}|\boldsymbol{P}_{\text{i}}\right) = 1. \tag{42}$$

When we describe the motion of the Brownian particle, the position $\boldsymbol{R}$ should be treated in the same manner as $\boldsymbol{P}$ because $d\boldsymbol{R}/dt = \boldsymbol{P}$. For simplicity, we consider space translational symmetric systems, so that we do not need to specify the position at the initial time. If one considers an external potential acting on the Brownian particle, then $\mathcal{P}_t(\boldsymbol{P}_{\text{f}}|\boldsymbol{P}_{\text{i}})$ should be replaced by $\mathcal{P}_t(\boldsymbol{P}_{\text{f}}, \boldsymbol{R}_{\text{f}}|\boldsymbol{P}_{\text{i}}, \boldsymbol{R}_{\text{i}})$.

Here, the most important property of $\mathcal{P}_t$ is

$$\mathcal{P}_t\left(\boldsymbol{P}_{\text{f}}|\boldsymbol{P}_{\text{i}}\right) P_{\text{MB}}\left(\boldsymbol{P}_{\text{i}}\right) = \mathcal{P}_t\left(\boldsymbol{P}_{\text{i}}^*|\boldsymbol{P}_{\text{f}}^*\right) P_{\text{MB}}\left(\boldsymbol{P}_{\text{f}}\right), \tag{43}$$

where we denote the Maxwell–Boltzmann distribution by

$$P_{\text{MB}}(\boldsymbol{P}) = (2\pi MT)^{-3/2} \exp\left[-\frac{|\boldsymbol{P}|^2}{2MT}\right]. \tag{44}$$

Property (43) is called the detailed balance condition. Using microscopic reversibility $(\Gamma^*)_{-t} = (\Gamma_t)^*$ (35)–(37), we can obtain the detailed balance condition (43).

3.2 Assumptions

Let τ_{micro} be the correlation time of the force acting on the Brownian particle, and τ_{macro} be the relaxation time of the momentum of the Brownian particle. We have the separation of time scales represented by $\tau_{\text{micro}} \ll \tau_{\text{macro}}$ because we assume that the relaxation time of the momentum is much larger than the time scales of the other degrees of freedom.

We define the time-averaged total force acting on the Brownian particle as

$$\bar{\boldsymbol{F}}(\Gamma) \equiv \frac{1}{\Delta t} \int_0^{\Delta t} ds\, \boldsymbol{F}(\Gamma_s), \tag{45}$$

where Δt is a finite time interval that satisfies $\tau_{\text{micro}} \ll \Delta t \ll \tau_{\text{macro}}$. We define the conditional probability density of $\bar{\boldsymbol{F}}$ given $\boldsymbol{P}_{\text{i}}$ by

$$\widetilde{\mathcal{P}}\left(\bar{\boldsymbol{F}}|\boldsymbol{P}_{\text{i}}\right) \equiv \int d\Gamma\, \tilde{P}_{\text{eq}}(\tilde{\Gamma})\delta\left(\boldsymbol{P} - \boldsymbol{P}_{\text{i}}\right)\delta(\bar{\boldsymbol{F}}(\Gamma) - \bar{\boldsymbol{F}}). \tag{46}$$

Considering $\Delta t \gg \tau_{\text{micro}}$, we may employ the central limit theorem, according to which and the isotropic property in equilibrium we have the following Gaussian form of $\widetilde{\mathcal{P}}(\bar{\boldsymbol{F}}|\boldsymbol{P}_{\text{i}})$:

$$\widetilde{\mathcal{P}}\left(\bar{\boldsymbol{F}}|\boldsymbol{P}_{\text{i}}\right) = C \exp\left[-\Delta t \mathcal{I}\left(\bar{\boldsymbol{F}}|\boldsymbol{P}_{\text{i}}\right)\right], \tag{47}$$

with

$$\mathcal{I}\left(\bar{\boldsymbol{F}}|\boldsymbol{P}_{\text{i}}\right) = \frac{|\bar{\boldsymbol{F}} - \boldsymbol{\mathcal{F}}(\boldsymbol{P}_{\text{i}})|^2}{4T\gamma(\boldsymbol{P}_{\text{i}})}, \tag{48}$$

where C is the normalization constant given by

$$C = \left[\frac{\Delta t}{4\pi T \gamma (\boldsymbol{P}_\mathrm{i})} \right]^{3/2}, \tag{49}$$

$\mathcal{F}(\boldsymbol{P}_\mathrm{i})$ is the most probable value of $\bar{\boldsymbol{F}}$, and $2T\gamma(\boldsymbol{P})$ is the dispersion of fluctuations of $\bar{\boldsymbol{F}}$. The dispersion $\gamma(\boldsymbol{P})$ is assumed to be positive for any $\boldsymbol{P}$. Because of the space-reflection symmetry, $\mathcal{F}$ and γ satisfy

$$\mathcal{F}(\boldsymbol{P}^*) = -\mathcal{F}(\boldsymbol{P}), \tag{50}$$

$$\gamma(\boldsymbol{P}^*) = \gamma(\boldsymbol{P}), \tag{51}$$

respectively. Assumptions (47) with (48) is essential for our derivation of a nonlinear Langevin equation for the Brownian particle. Note that we may prove this assumption when bath particles collisions with a Brownian particle can be regarded as independent.

3.3 Derivation

Using the equation of motion (39), $\bar{\boldsymbol{F}}(\Gamma)$ defined in (45) can be rewritten as

$$\bar{\boldsymbol{F}}(\Gamma) = \frac{\boldsymbol{P}_{\Delta t} - \boldsymbol{P}}{\Delta t}. \tag{52}$$

Thus, by changing variables from $\bar{\boldsymbol{F}}(\Gamma)$ to $\boldsymbol{P}_{\Delta t}$ in (47) with (48), we obtain

$$\begin{aligned}
\mathcal{P}_{\Delta t}\left(\boldsymbol{P}_\mathrm{f} | \boldsymbol{P}_\mathrm{i}\right) &= \widetilde{\mathcal{P}}\left(\bar{\boldsymbol{F}} | \boldsymbol{P}_\mathrm{i}\right) / (\Delta t)^3 \\
&= \left[4\pi \Delta t T \gamma\left(\boldsymbol{P}_\mathrm{i}\right)\right]^{-3/2} \exp\left[-\frac{\Delta t}{4T\gamma(\boldsymbol{P}_\mathrm{i})} \left| \frac{\boldsymbol{P}_\mathrm{f} - \boldsymbol{P}_\mathrm{i}}{\Delta t} - \mathcal{F}(\boldsymbol{P}_\mathrm{i}) \right|^2 \right].
\end{aligned} \tag{53}$$

When we compare (53) with (13) for $\alpha = 0$, we can describe the discrete time evolution of $\boldsymbol{P}$ as the discrete stochastic system (8). Now, taking the limit $\Delta t/\tau_\mathrm{macro} \to 0$, we obtain the Langevin equation

$$\frac{\mathrm{d}\boldsymbol{P}_t}{\mathrm{d}t} = \mathcal{F}(\boldsymbol{P}_t) + \sqrt{2T\gamma(\boldsymbol{P}_t)} \cdot \boldsymbol{\xi}_t, \tag{54}$$

where $\boldsymbol{\xi}_t$ is the zero-mean Gaussian white noise with covariance $\langle \xi_t^a \xi_s^b \rangle = \delta^{ab}\delta(t-s)$ and $\cdot$ denotes multiplication with the Itô rule.

Next, we express $\mathcal{F}$ in terms of γ from the detailed balance condition (43). Using (13) with $\alpha = 1/2$, we can rewrite the transition probability (53) in terms of $\boldsymbol{P}_\mathrm{m} \equiv (\boldsymbol{P}_\mathrm{f} + \boldsymbol{P}_\mathrm{i})/2$ as

$$\begin{aligned}
\mathcal{P}_{\Delta t}\left(\boldsymbol{P}_\mathrm{f} | \boldsymbol{P}_\mathrm{i}\right) &= \left[4\pi \Delta t T \gamma\left(\boldsymbol{P}_\mathrm{m}\right)\right]^{-3/2} \exp\left[-\frac{\Delta t}{4T\gamma(\boldsymbol{P}_\mathrm{m})} \left| \frac{\boldsymbol{P}_\mathrm{f} - \boldsymbol{P}_\mathrm{i}}{\Delta t} - \widetilde{\mathcal{F}}(\boldsymbol{P}_\mathrm{m}) \right|^2 \right. \\
&\quad \left. - \frac{\Delta t}{2} \sum_{a=x,y,z} \frac{\partial \mathcal{F}^a(\boldsymbol{P}_\mathrm{m})}{\partial P_\mathrm{m}^a} + \frac{\Delta t}{4} \sum_{a=x,y,z} T \frac{\partial^2 \gamma(\boldsymbol{P}_\mathrm{m})}{\partial P_\mathrm{m}^a \partial P_\mathrm{m}^a} \right],
\end{aligned} \tag{55}$$

with

$$\widetilde{\mathcal{F}}^a\left(\boldsymbol{P}_\mathrm{m}\right) \equiv \mathcal{F}^a\left(\boldsymbol{P}_\mathrm{m}\right) - T \frac{\partial \gamma(\boldsymbol{P}_\mathrm{m})}{\partial P_\mathrm{m}^a}, \tag{56}$$

where the superscript a represents the indices in Cartesian coordinates (x, y, z), and we have ignored all $o(\Delta t)$ terms because these terms are irrelevant in the limit $\Delta t/\tau_\mathrm{macro} \to 0$.

Note that the Stratonovich convention ($\alpha = 1/2$) is convenient when using (43) because it has the property that the forward and backward paths are evaluated at the same points [24]. Substituting (55) into (43) with (44), we obtain

$$\widetilde{\mathcal{F}}(\boldsymbol{P}_{\mathrm{m}}) = -\frac{\gamma(\boldsymbol{P}_{\mathrm{m}})}{M}\boldsymbol{P}_{\mathrm{m}}, \tag{57}$$

which is called the fluctuation–dissipation relation of the second kind. From (54), (56), and (57), we have

$$\begin{aligned}
\frac{\mathrm{d}P_t^a}{\mathrm{d}t} &= -\frac{\gamma(\boldsymbol{P}_t)}{M}P_t^a + T\frac{\partial\gamma(\boldsymbol{P}_t)}{\partial P_t^a} + \sqrt{2T\gamma(\boldsymbol{P}_t)}\cdot\xi_t^a \\
&= -\frac{\gamma(\boldsymbol{P}_t)}{M}P_t^a + \sqrt{2T\gamma(\boldsymbol{P}_t)}\odot\xi_t^a,
\end{aligned} \tag{58}$$

which is equivalent to (7). From (58), $\gamma(\boldsymbol{P}_t)$ can be interpreted as a nonlinear friction coefficient.

4 Generalization

4.1 Motivation

We consider fluctuations of a system in equilibrium. There is a special set of variables whose time scales are well separated from those of the other dynamical degrees of freedom. We refer to such a set as a complete set of slow variables and denote it by $X = (X^1, X^2, \ldots, X^N)$. For the example in the previous section, $X = (\boldsymbol{R}, \boldsymbol{P})$. As a different example, one may consider fluctuations in a thermodynamically isolated system separated into two regions by a freely movable diabatic wall. In this case, unconstrained thermodynamic extensive variables, the energy and the volume in one region are assumed to form a complete set of slow variables. Furthermore, hydrodynamic fluctuations, which are long-wavelength fluctuations of locally conserved quantities, in an equilibrium liquid are another example of a complete set of slow variables. For simplicity, we assume that the Hamiltonian of the microscopic mechanical system is symmetric with respect to the time-reversal operation. For such a system, the probability density of X is denoted by

$$P_{\mathrm{eq}}(X) = \frac{1}{Z}\exp[S(X)], \tag{59}$$

where Z is the normalization constant, and $S(X^*) = S(X)$ for the time reversal X^* of X. The physical interpretation of S depends on the system being studied. For example, $S(X)$ corresponds to entropy when thermodynamic fluctuations in an isolated system are considered. For other cases, $S(X)$ should be read from the form of the stationary distribution. Suppose that the system is in equilibrium. We then expect that the time evolution of X can be described by a Langevin equation. In this section, we derive the equation by generalizing the arguments in the previous section.

4.2 Basic Concept

On the basis of a microscopic mechanical description, we can define the conditional probability density of $X = X_{\mathrm{f}}$ at time t, denoted by $\mathcal{P}_t(X_{\mathrm{f}}|X_{\mathrm{i}})$, provided that $X = X_{\mathrm{i}}$ at time 0. There are two important properties of this probability density. First, following the central

limit theorem, we assume the Gaussian form of form of the probability density for the time averaged flux written as $(X_{\Delta t} - X)/\Delta t$. The result (53) in the previous section becomes

$$
\mathcal{P}_{\Delta t}\left(X_{\mathrm{f}}|X_{\mathrm{i}}\right) = \frac{1}{\sqrt{(4\pi\,\Delta t)^N \det \mathsf{L}(X_{\mathrm{i}})}} \exp\left[-\frac{\Delta t}{4}\sum_{i,j}\left(L^{-1}\right)^{ij}(X_{\mathrm{i}})\right.
$$
$$
\left.\times\left[\frac{X_{\mathrm{f}}^i - X_{\mathrm{i}}^i}{\Delta t} - \mathcal{J}^i(X_{\mathrm{i}})\right]\left[\frac{X_{\mathrm{f}}^j - X_{\mathrm{i}}^j}{\Delta t} - \mathcal{J}^j(X_{\mathrm{i}})\right]\right], \tag{60}
$$

where $\mathcal{J}^i$ is the most probable value of the time averaged flux, and $2L^{ij}$ is the dispersion matrix. We assume that the matrix $\mathsf{L} = (L^{ij})$ is positive definite. This means that each X_{f}^i is not uniquely determined by X_{i}. If X_{f}^i is uniquely determined by X_{i}, such as $X_{\mathrm{f}}^i = X_{\mathrm{i}}^i + \mathcal{J}^i(X_{\mathrm{i}})\Delta t$, then we multiply the right-hand side of (60) by $\delta(X_{\mathrm{f}}^i - X_{\mathrm{i}}^i - \mathcal{J}^i(X_{\mathrm{i}})\Delta t)$ and consider the submatrix formed by deleting the ith row and ith column of L. Second, from the reversibility of microscopic Hamiltonian systems, we can obtain

$$
\mathcal{P}_t\left(X_{\mathrm{f}}|X_{\mathrm{i}}\right) P_{\mathrm{eq}}\left(X_{\mathrm{i}}\right) = \mathcal{P}_t\left(X_{\mathrm{i}}^*|X_{\mathrm{f}}^*\right) P_{\mathrm{eq}}\left(X_{\mathrm{f}}\right). \tag{61}
$$

Then, by substituting (60) into (61), we obtain a possible form of $\mathcal{J}^i(X)$ and a symmetry property of L. For convenience, we denote X^* by $\boldsymbol{\epsilon}X = (\epsilon^1 X^1,\ \epsilon^2 X^2, \ldots, \epsilon^N X^N)$, where $\epsilon^i = +1$ or -1 for X^i. We decompose $\mathcal{J}^i$ into two parts,

$$
\mathcal{J}^i(X) = \mathcal{J}_{\mathrm{rev}}^i(X) + \mathcal{J}_{\mathrm{irr}}^i(X), \tag{62}
$$

with

$$
\mathcal{J}_{\mathrm{rev}}^i(X) \equiv \frac{\mathcal{J}^i(X) - \epsilon^i\,\mathcal{J}^i(X^*)}{2}, \tag{63}
$$
$$
\mathcal{J}_{\mathrm{irr}}^i(X) \equiv \frac{\mathcal{J}^i(X) + \epsilon^i\,\mathcal{J}^i(X^*)}{2}, \tag{64}
$$

which satisfy $\mathcal{J}_{\mathrm{rev}}^i(X^*) = -\epsilon^i\,\mathcal{J}_{\mathrm{rev}}^i(X)$ and $\mathcal{J}_{\mathrm{irr}}^i(X^*) = \epsilon^i\,\mathcal{J}_{\mathrm{irr}}^i(X)$. We also define the matrix $\mathsf{L}_{\mathrm{T}} = (L_{\mathrm{T}}^{ij})$ by

$$
L_{\mathrm{T}}^{ij}(X) = \epsilon^i\epsilon^j L^{ij}(X^*). \tag{65}
$$

Note that $\det \mathsf{L}_{\mathrm{T}}(X) = \det \mathsf{L}(X^*)$.

4.3 Result

Direct substitution of (60) into (61) would result in a complicated form, so we use a trick. Considering that (13) holds for any α, as in the previous section, we can rewrite (60) by changing $\alpha = 0$–1/2. The result is

$$
\mathcal{P}_{\Delta t}\left(X_{\mathrm{f}}|X_{\mathrm{i}}\right) = \frac{1}{\sqrt{(4\pi\,\Delta t)^N \det \mathsf{L}(X_{\mathrm{m}})}} \exp\left[-\frac{\Delta t}{4}\sum_{i,j}\left(L^{-1}\right)^{ij}(X_{\mathrm{m}})\right.
$$
$$
\left.\times\left[\frac{X_{\mathrm{f}}^i - X_{\mathrm{i}}^i}{\Delta t} - \widetilde{\mathcal{J}}^i(X_{\mathrm{m}})\right]\left[\frac{X_{\mathrm{f}}^j - X_{\mathrm{i}}^j}{\Delta t} - \widetilde{\mathcal{J}}^j(X_{\mathrm{m}})\right]\right.
$$

$$- \frac{\Delta t}{2} \sum_i \frac{\partial \mathcal{J}_{\mathrm{rev}}^i (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i} - \frac{\Delta t}{2} \sum_i \frac{\partial \mathcal{J}_{\mathrm{irr}}^i (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i}$$

$$+ \frac{\Delta t}{4} \sum_{i,j} \frac{\partial^2 L^{ij} (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i \partial X_{\mathrm{m}}^j} \Bigg], \tag{66}$$

with

$$\widetilde{\mathcal{J}}^i (\boldsymbol{X}_{\mathrm{m}}) \equiv \mathcal{J}_{\mathrm{rev}}^i (\boldsymbol{X}_{\mathrm{m}}) + \mathcal{J}_{\mathrm{irr}}^i (\boldsymbol{X}_{\mathrm{m}}) - \sum_j \frac{\partial L^{ij} (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^j}. \tag{67}$$

We also obtain

$$\mathcal{P}_{\Delta t} \left(X_{\mathrm{i}}^* | X_{\mathrm{f}}^* \right) = \frac{1}{\sqrt{(4\pi \Delta t)^N \det \mathsf{L}_{\mathrm{T}}(\boldsymbol{X}_{\mathrm{m}})}} \exp \Bigg[-\frac{\Delta t}{4} \sum_{i,j} \left(L_{\mathrm{T}}^{-1} \right)^{ij} (\boldsymbol{X}_{\mathrm{m}})$$

$$\times \left[\frac{X_{\mathrm{f}}^i - X_{\mathrm{i}}^i}{\Delta t} - \widetilde{\mathcal{J}}_{\mathrm{T}}^i (\boldsymbol{X}_{\mathrm{m}}) \right] \left[\frac{X_{\mathrm{f}}^j - X_{\mathrm{i}}^j}{\Delta t} - \widetilde{\mathcal{J}}_{\mathrm{T}}^j (\boldsymbol{X}_{\mathrm{m}}) \right]$$

$$+ \frac{\Delta t}{2} \sum_i \frac{\partial \mathcal{J}_{\mathrm{rev}}^i (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i} - \frac{\Delta t}{2} \sum_i \frac{\partial \mathcal{J}_{\mathrm{irr}}^i (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i}$$

$$+ \frac{\Delta t}{4} \sum_{i,j} \frac{\partial^2 L_{\mathrm{T}}^{ij} (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i \partial X_{\mathrm{m}}^j} \Bigg], \tag{68}$$

with

$$\widetilde{\mathcal{J}}_{\mathrm{T}}^i (\boldsymbol{X}_{\mathrm{m}}) \equiv \mathcal{J}_{\mathrm{rev}}^i (\boldsymbol{X}_{\mathrm{m}}) - \mathcal{J}_{\mathrm{irr}}^i (\boldsymbol{X}_{\mathrm{m}}) + \sum_j \frac{\partial L_{\mathrm{T}}^{ij} (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^j}. \tag{69}$$

Substituting (66) and (68) into (61) with (59), we obtain

$$\frac{\Delta t}{4} \sum_{i,j} \left(L^{-1} \right)^{ij} (\boldsymbol{X}_{\mathrm{m}}) \left[\frac{X_{\mathrm{f}}^i - X_{\mathrm{i}}^i}{\Delta t} - \widetilde{\mathcal{J}}^i (\boldsymbol{X}_{\mathrm{m}}) \right] \left[\frac{X_{\mathrm{f}}^j - X_{\mathrm{i}}^j}{\Delta t} - \widetilde{\mathcal{J}}^j (\boldsymbol{X}_{\mathrm{m}}) \right]$$

$$- \frac{\Delta t}{4} \sum_{i,j} \left(L_{\mathrm{T}}^{-1} \right)^{ij} (\boldsymbol{X}_{\mathrm{m}}) \left[\frac{X_{\mathrm{f}}^i - X_{\mathrm{i}}^i}{\Delta t} - \widetilde{\mathcal{J}}_{\mathrm{T}}^i (\boldsymbol{X}_{\mathrm{m}}) \right] \left[\frac{X_{\mathrm{f}}^j - X_{\mathrm{i}}^j}{\Delta t} - \widetilde{\mathcal{J}}_{\mathrm{T}}^j (\boldsymbol{X}_{\mathrm{m}}) \right]$$

$$+ \Delta t \sum_i \frac{\partial \mathcal{J}_{\mathrm{rev}}^i (\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i} - \frac{\Delta t}{4} \sum_{i,j} \frac{\partial^2}{\partial X_{\mathrm{m}}^i \partial X_{\mathrm{m}}^j} \left[L^{ij} (\boldsymbol{X}_{\mathrm{m}}) - L_{\mathrm{T}}^{ij} (\boldsymbol{X}_{\mathrm{m}}) \right]$$

$$+ \frac{1}{2} \log \frac{\det \mathsf{L}(\boldsymbol{X}_{\mathrm{m}})}{\det \mathsf{L}_{\mathrm{T}}(\boldsymbol{X}_{\mathrm{m}})} + \Delta t \sum_i \frac{X_{\mathrm{f}}^i - X_{\mathrm{i}}^i}{\Delta t} \frac{\partial S(\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i} = 0, \tag{70}$$

where we have used

$$S (\boldsymbol{X}_{\mathrm{f}}) - S (\boldsymbol{X}_{\mathrm{i}}) = \Delta t \sum_i \frac{X_{\mathrm{f}}^i - X_{\mathrm{i}}^i}{\Delta t} \frac{\partial S(\boldsymbol{X}_{\mathrm{m}})}{\partial X_{\mathrm{m}}^i} + O \left((\Delta t)^2 \left(\frac{X_{\mathrm{f}}^i - X_{\mathrm{i}}^i}{\Delta t} \right)^2 \right), \tag{71}$$

and where the $O((X_{\mathrm{f}}^i - X_{\mathrm{i}}^i)^2)$ terms in (70) are irrelevant in the limit $\Delta t / \tau_{\mathrm{macro}} \to 0$. Note that $\partial S(\boldsymbol{X}) / \partial X^i$ are called the thermodynamic forces. Because (70) holds for any $X_{\mathrm{f}}^i - X_{\mathrm{i}}^i$ and X_{m}^i, comparing the quadratic terms in $X_{\mathrm{f}}^i - X_{\mathrm{i}}^i$ in (70) yields

$$L^{ij}(X) = L_{\mathrm{T}}^{ij}(X). \tag{72}$$

Comparing the first-order terms in $X_{\mathrm{f}}^i - X_{\mathrm{i}}^i$ in (70) with (72), we also have

$$\mathcal{J}_{\mathrm{irr}}^i(X) = \sum_j L^{ij}(X) \frac{\partial S(X)}{\partial X^j} + \sum_j \frac{\partial L^{ij}(X)}{\partial X^j}, \tag{73}$$

which is called the fluctuation–dissipation relation of the second kind. Comparing the zero-order terms in $X_{\mathrm{f}}^i - X_{\mathrm{i}}^i$ in (70) with (72) and (73), we finally obtain

$$\sum_i \frac{\partial}{\partial X^i} \left[\mathcal{J}_{\mathrm{rev}}^i(X) P_{\mathrm{eq}}(X) \right] = 0. \tag{74}$$

Now, we go back to (60). This is interpreted as the transition probability for the discrete time Langevin equation (8). By taking the limit $\Delta t / t_{\mathrm{macro}} \to 0$, we obtain

$$\frac{\mathrm{d}X_t^i}{\mathrm{d}t} = \mathcal{J}_{\mathrm{rev}}^i(X_t) + \sum_j L^{ij}(X_t) \frac{\partial S(X_t)}{\partial X_t^j} + \sum_j \frac{\partial L^{ij}(X_t)}{\partial X_t^j} + \sum_j l^{ij}(X_t) \cdot \xi_t^j, \tag{75}$$

where we have used (73), and l^{ij} satisfies

$$L^{ij}(X_t) = \frac{1}{2} \sum_k l^{ik}(X_t) l^{jk}(X_t). \tag{76}$$

It should be noted that the third term on the right-hand side of (75) is not eliminated even if we replace $l^{ij}(X_t) \cdot \xi_t^j$ by $l^{ij}(X_t) \odot \xi_t^j$ for general multi-component systems. We emphasize that the result (58) in the previous section is rather accidental.

5 Concluding Remarks

In this paper, we have derived a universal form of nonlinear, multiplicative Langevin equations for slow variables in equilibrium systems. The result is essentially equivalent to that derived by Green in 1952. In contrast to previous studies, we first assume the separation of time scales. Then, by using the central limit theorem, we can formalize the asymptotic form of the probability density for the time-averaged fluxes, which determines the time evolution of the slow variables due to the time-reversal symmetry of fluctuation.

Here, we refer to the large deviation theory. Our assumption (47) means large deviation property for the conditional probability density which is more general assumption than the quadratic form of the large deviation function (48). For instance, when one derives a stochastic time evolution equation with white Poisson noises from a microscopic mechanical system, the large deviation property should be valid. The validity condition for (48) depends on details of a microscopic mechanical system. In equilibrium systems, a large deviation function of thermodynamic variables has all information about fluctuations of the thermodynamic variables, and is expressed in terms of a corresponding thermodynamic function, which can be derived from a microscopic mechanical system by using equilibrium statistical mechanics. A central limit theorem can be derived from the large deviation function, and only gives a corresponding fluctuation–dissipation theorem. In this sense, the large deviation is a more fundamental concept for analyzing fluctuations although physical aspects of the large deviation function for the time-averaged flux remain to be studied.

Before ending this paper, we summarize future problems related to our results. First, although discussions of physical phenomena are out of scope of this paper, it seems interesting to find a system in which the third term on the right-hand side of (75) plays an important role in phenomena. We do not know any such examples explicitly, but there are many cases where transportation coefficients strongly depend on thermodynamic variables. However, it should be noted that the contribution of this term is higher-order in macroscopic systems, as discussed in Sect. 1. We thus seek such systems among small systems or singular systems.

Second, we believe that the result and its derivation method may provide the final answer for formally describing slow variables in equilibrium systems. For example, in principle, fluctuating hydrodynamics in equilibrium systems can be studied explicitly using the same method. In this case, we consider long-wavelength fluctuations of locally conserved quantities, which are called hydrodynamic modes, to be slow variables. Although the formulation may be developed similarly to the ideas in this paper, there will be many technical difficulties in performing concrete calculations. We thus conjecture that a formal derivation of the Navier–Stokes equation from Hamiltonian particle systems [28] may be helpful for completing this problem. Related to the argument of fluctuating hydrodynamics, we recall the assumption in Sect. 3 that the relaxation time of the momentum is much larger than the time scales of the other degrees of freedom. There are cases where this assumption does not hold when the time scale of hydrodynamic modes is comparable with that of the momentum of a Brownian particle, as observed in recent experiments [29–31]. Deriving the Langevin equation describing the motion observed in these experiments is also a challenging task.

Third, the obvious problem we should study in future is formulating the stochastic evolution of slow variables in systems out of equilibrium. If the time scales are separated so that the slow variables are clearly defined, then the concept of large deviation can be used even in nonequilibrium cases. Of course, there are many nonequilibrium phenomena in which a complete set of slow variables is not identified. Although these have interesting phenomena, we do not have a systematic method for studying them. Putting such systems aside, we focus on systems where the slow variables are defined. If the Gaussian form of the large deviation function is effective, then the dynamics of the slow variables will be described by a Langevin equation. Even for this simple class, we do not have a general form, because the symmetry property (61) cannot be used. Rather, one may find that difficulties already appear in writing a deterministic equation for slow variables before considering stochastic systems. Nevertheless, our method will be applied to the Brownian motion in nonuniform temperatures where the system satisfies the detailed balance condition as studied in [32].

Finally, recent work explicitly derived deterministic order parameter equations near the order–disorder transition of the globally coupled XY model and the synchronization–desynchronization transition of the Kuramoto model [33]. The characteristic feature of the derivation method here is to use nonequilibrium identities such as the Jarzynski equality [34] and the Hatano–Sasa equality [35]. Although we need to assume a rather special type of probability distribution at the initial time, the calculation steps are substantially reduced. We thus expect that the unified framework of the approach in [33] and the theory developed in this paper may be the first step in the universal description of the stochastic evolution of slow variables out of equilibrium.

Appendix: Transition Probability for a Discretized Langevin Equation

Normalization Condition

We verify the normalization condition

$$\int dx_{n+1} \mathcal{P}(x_{n+1}|x_n) = 1 + o(dt), \tag{77}$$

for the transition probability

$$\mathcal{P}(x_{n+1}|x_n) = \frac{1}{\sqrt{4\pi \, dt \, G(\bar{x}_n)}} \exp\left[-\frac{dt}{4G(\bar{x}_n)}\left[\frac{dx_n}{dt} - f(\bar{x}_n) + 2\alpha G'(\bar{x}_n)\right]^2\right.$$
$$\left. - \alpha f'(\bar{x}_n)\, dt + \alpha^2 G''(\bar{x}_n)\, dt\right], \tag{78}$$

where primes denote derivatives with respect to the argument, $\bar{x}_n \equiv \alpha x_{n+1} + (1-\alpha)x_n$, and $dx_n \equiv x_{n+1} - x_n$. Changing the variable of integration from x_{n+1} to $z \equiv dx_n/\sqrt{dt}$, we obtain

$$\int dx_{n+1} \mathcal{P}(x_{n+1}|x_n)$$
$$= \int dz \frac{1}{\sqrt{4\pi \, G(\bar{x}_n)}} \exp\left[-\frac{dt}{4G(\bar{x}_n)}\left[\frac{z}{\sqrt{dt}} - f(\bar{x}_n) + 2\alpha G'(\tilde{x}_n)\right]^2\right.$$
$$\left. - \alpha f'(\bar{x}_n)\, dt + \alpha^2 G''(\bar{x}_n)\, dt\right], \tag{79}$$

with $\bar{x}_n = x_n + \alpha z \sqrt{dt}$. In this subsection, for any function $A(\cdot)$, we abbreviate $A(x_n)$ to A. Expanding the integrand on the right-hand side of (79) in powers of $\sqrt{dt}$, we have

$$\frac{\exp\left[-\dfrac{dt}{4G(\bar{x}_n)}\left[\dfrac{z}{\sqrt{dt}} - f(\bar{x}_n) + 2\alpha G'(\bar{x}_n)\right]^2 - \alpha f'(\bar{x}_n)\, dt + \alpha^2 G''(\bar{x}_n)\, dt\right]}{\sqrt{4\pi \, G(\bar{x}_n)}}$$

$$= \frac{e^{-z^2/(4G)}}{\sqrt{4\pi G}}\left\{1 + \sqrt{dt}\left[\frac{f}{2G}z - \frac{3\alpha G'}{2G}z + \frac{\alpha G'}{4G^2}z^3\right] + \frac{dt}{2}\left[\vphantom{\frac{z^4}{6G}}\right.\right.$$
$$-(f - 2\alpha G')\frac{3\alpha G'}{2G^2}\left(z^2 - \frac{z^4}{6G}\right) - \frac{(f - 2\alpha G')^2}{2G}\left(1 - \frac{z^2}{2G}\right)$$
$$- 2\alpha f'\left(1 - \frac{z^2}{2G}\right) + \frac{3(\alpha G')^2}{4G^2}\left(z^2 - \frac{z^4}{G} + \frac{z^6}{12G^2}\right)$$
$$\left.\left. + 2\alpha^2 G''\left(1 - \frac{5z^2}{4G} + \frac{z^4}{8G^2}\right)\right] + o(dt)\right\}. \tag{80}$$

Substituting (80) into (79) and using

$$\int dz \frac{e^{-z^2/(4G)}}{\sqrt{4\pi G}} z^{2n} = (2n-1)!!(2G)^n, \tag{81}$$

$$\int dz \frac{e^{-z^2/(4G)}}{\sqrt{4\pi G}} z^{2n-1} = 0, \tag{82}$$

for $n \in \mathbb{N}$, we obtain the normalization condition (77).

Derivation of the Fokker–Planck Equation from the Transition Probability

To establish the equivalence of the Fokker–Planck equation and the path integral formulation, we derive the Fokker–Planck equation

$$\frac{\partial}{\partial t}P(\chi, t) = -\frac{\partial}{\partial \chi}\left[f(\chi)P(\chi, t)\right] + \frac{\partial^2}{\partial \chi^2}\left[G(\chi)P(\chi, t)\right], \tag{83}$$

from the transition probability (78) in the limit $dt \to 0$. To arrive at (83), we evaluate

$$P(x, n+1) = \int dy \mathcal{P}(x|y)P(y, n). \tag{84}$$

Changing the variable of integration from y to $z = (x - y)/\sqrt{dt}$, we obtain

$$P(x, n+1) = \int dz \frac{P(x - z\sqrt{dt}, n)}{\sqrt{4\pi G(\bar{x})}} \exp\left[-\frac{dt}{4G(\bar{x})}\left[\frac{z}{\sqrt{dt}} - f(\bar{x}) + 2\alpha G'(\bar{x})\right]^2\right.$$
$$\left. - \alpha f'(\bar{x})dt + \alpha^2 G''(\bar{x})dt\right], \tag{85}$$

with $\bar{x} = x + (\alpha - 1)z\sqrt{dt}$. In this subsection, we abbreviate $A(x)$ and $A(x, n)$ to A for any function A. Expanding the right-hand side of (85) in powers of $\sqrt{dt}$, we have

$$P(x, n+1) = \int dz \frac{e^{-z^2/(4G)}}{\sqrt{4\pi G}}\left[P - P'z\sqrt{dt} + P''\frac{z^2 dt}{2} + o(dt)\right]$$
$$\times \left\{1 + \sqrt{dt}\left[\frac{f - 2\alpha G'}{2G}z - \frac{(\alpha - 1)G'}{2G}z + \frac{(\alpha - 1)G'}{4G^2}z^3\right]\right.$$
$$+ \frac{dt}{2}\left[-(f - 2\alpha G')\frac{3(\alpha - 1)G'}{2G^2}\left(z^2 - \frac{z^4}{6G}\right)\right.$$
$$-\frac{(f - 2\alpha G')^2}{2G}\left(1 - \frac{z^2}{2G}\right) - 2\alpha f'\left(1 - \frac{(\alpha - 1)z^2}{2\alpha G}\right)$$
$$+ \frac{3[(\alpha - 1)G']^2}{4G^2}\left(z^2 - \frac{z^4}{G} + \frac{z^6}{12G^2}\right)$$
$$\left.\left. + 2\alpha^2 G''\left(1 - \frac{(5\alpha - 1)(\alpha - 1)z^2}{4\alpha^2 G} + \frac{(\alpha - 1)^2 z^4}{8\alpha^2 G^2}\right)\right] + o(dt)\right\}$$
$$= (1 - f'dt + G''dt)P - (fdt - 2G'dt)P' + GdtP'' + o(dt). \tag{86}$$

Thus, we obtain

$$\frac{P(x, n+1) - P(x, n)}{dt} = -\frac{\partial}{\partial x}[f(x)P(x, n)] + \frac{\partial}{\partial x^2}[G(x)P(x, n)] + \frac{o(dt)}{dt}. \tag{87}$$

In the limit $dt \to 0$, we arrive at the Fokker–Planck equation (83).

References

1. Onsager, L.: Reciprocal relations in irreversible processes. I. Phys. Rev. **37**, 405–426 (1931)
2. Onsager, L.: Reciprocal relations in irreversible processes. II. Phys. Rev. **38**, 2265–2279 (1931)
3. Hashitsume, N.: A statistical theory of linear dissipative systems. Prog. Theor. Phys. **8**, 461–478 (1952)
4. Onsager, L., Machlup, S.: Fluctuations and irreversible processes. Phys. Rev. **91**, 1505–1512 (1953)

5. Machlup, S., Onsager, L.: Fluctuations and irreversible process. II. Systems with kinetic energy. Phys. Rev. **991**, 1512–1515 (1953)
6. De Groot, S.R., Mazur, P.: Non-equilibrium Thermodynamics. Dover, New York (1984)
7. Green, M.S.: Markoff random processes and the statistical mechanics of time-dependent phenomena. J. Chem. Phys. **20**, 1281–1295 (1952)
8. Yamamoto, T.: Statistical mechanics of general Brownian motions underlying irreversible processes. Prog. Theor. Phys. **10**, 11–23 (1953)
9. Zwanzig, R.: Memory effects in irreversible thermodynamics. Phys. Rev. **124**, 983–992 (1961)
10. Zwanzig, R., Nordholm, K.S.J., Mitchell, W.C.: Memory effects in irreversible thermodynamics: corrected derivation of transport equations. Phys. Rev. A **5**, 2680–2682 (1972)
11. Bashkirov, A.G., Zubarev, D.N.: The generalized Kramers–Fokker–Planck equation. Physica **48**, 137–145 (1970)
12. Graham, R., Haken, H.: Generalized thermodynamic potential for Markoff systems in detailed balance and far from thermal equilibrium. Z. Phys. **243**, 289–302 (1971)
13. Graham, R., Haken, H.: Fluctuations and stability of stationary non-equilibrium systems in detailed balance. Z. Phys. **245**, 141–153 (1971)
14. Kawasaki, K.: Simple derivations of generalized linear and nonlinear Langevin equations. J. Phys. A **6**, 1289–1295 (1973)
15. Graham, R.: Covariant formulation of non-equilibrium statistical thermodynamics. Z. Phys. B **26**, 397–405 (1977)
16. Grabert, H., Green, M.S.: Fluctuations and nonlinear irreversible processes. Phys. Rev. A **19**, 1747–1756 (1979)
17. Grabert, H., Graham, R., Green, M.S.: Fluctuations and nonlinear irreversible processes. II. Phys. Rev. A **21**, 2136–2146 (1980)
18. Hanggi, P.: Connection between deterministic and stochastic descriptions of nonlinear systems. Phys. Acta **53**, 491–496 (1980)
19. Eyink, G.L., Lebowitz, J.L., Spohn, H.: Hydrodynamics and fluctuations outside of local equilibrium: driven diffusive systems. J. Stat. Phys. **83**, 385–472 (1996)
20. Van Kampen, N.G.: A power series expansion of the master equation. Can. J. Phys. **39**, 551–567 (1961)
21. Van Kampen, N.G.: Stochastic Processes in Physics and Chemistry, 3rd edn. North-Holland, Amsterdam (2007)
22. Oono, Y.: Onsager's principle from large deviation point of view. Prog. Theor. Phys. **89**, 973–983 (1993)
23. Wissel, C.: Manifolds of equivalent path integral solutions of the Fokker–Planck equation. Z. Phys. B **35**, 185–191 (1979)
24. Lau, A.W.C., Lubensky, T.C.: State-dependent diffusion: thermodynamic consistency and its path integral formulation. Phys. Rev. E **76**, 011123 (2007)
25. Aron, C., Barci, D.G., Cugliandolo, L.F., Arenas, Z.G., Lozano, G.S.: Dynamical symmetries of Markov processes with multiplicative white noise. J. Stat. Mech. 053207 (2016)
26. Chernyak, V.Y., Chertkov, M., Jarzynski, C.: Path-integral analysis of fluctuation theorems for general Langevin processes. J. Stat. Mech. P08001 (2006)
27. Gardiner, C.W.: Handbook of Stochastic Methods. Springer, Heidelberg (1983)
28. Sasa, S.: Derivation of hydrodynamics from the Hamiltonian description of particle systems. Phys. Rev. Lett. **112**, 100602 (2014)
29. Huang, R., Chavez, I., Taute, K.M., Lukić, B., Jeney, S., Raizen, M.G., Florin, E.-L.: Direct observation of the full transition from ballistic to diffusive Brownian motion in a liquid. Nat. Phys. **7**, 576–580 (2011)
30. Franosch, T., Grimm, M., Belushkin, M., Mor, F.M., Foffi, G., Forró, L., Jeney, S.: Resonances arising from hydrodynamic memory in Brownian motion. Nature **478**, 85–88 (2011)
31. Kheifets, S., Simha, A., Melin, K., Li, T., Raizen, M.G.: Observation of Brownian motion in liquids at short times: instantaneous velocity and memory loss. Science **343**, 1493–1496 (2014)
32. Polettini, M.: Diffusion in nonuniform temperature and its geometric analog. Phys. Rev. E **87**, 032126 (2013)
33. Sasa, S.: Collective dynamics from stochastic thermodynamics. N. J. Phys. **17**, 045024 (2015)
34. Jarzynski, C.: Nonequilibrium equality for free energy differences. Phys. Rev. Lett. **78**, 2690–2693 (1997)
35. Hatano, T., Sasa, S.: Steady-state thermodynamics of Langevin systems. Phys. Rev. Lett. **86**, 3463–3466 (2001)