Covariant nonequilibrium thermodynamics from Ito-Langevin dynamics

Mingnan Ding [1] and Xiangjun Xing [1,2,3,*]

[1]*Wilczek Quantum Center, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China*
[2]*T.D. Lee Institute, Shanghai Jiao Tong University, Shanghai 200240, China*
[3]*Shanghai Research Center for Quantum Sciences, Shanghai 201315, China*

Using the recently developed covariant Ito-Langevin dynamics, we develop a nonequilibrium thermodynamic theory for small systems coupled to multiplicative noises. The theory is based on Ito calculus, and is fully covariant under time-independent nonlinear transformation of variables. Assuming instantaneous detailed balance, we derive expressions for various thermodynamic functions, including work, heat, entropy production, and free energy, both at ensemble level and at trajectory level, and prove the second law of thermodynamics for arbitrary nonequilibrium processes. We relate time-reversal asymmetry of path probability to entropy production, and derive its consequences such as fluctuation theorem and nonequilibrium work relation. For Langevin systems with additive noises, our theory is equivalent to the common theories of stochastic energetics and stochastic thermodynamics. Using concrete examples, we demonstrate that whenever kinetic coefficients or metric tensor depend on system variables, the common theories of stochastic thermodynamics and stochastic energetics should be replaced by our theory.

I. INTRODUCTION

One of the main themes of modern nonequilibrium physics is to study irreversible processes in small systems, where fluctuations play a dominant role and stochastic descriptions become necessary. Huge progresses have been achieved in this course in the past few decades [1–3]. Thermodynamic quantities such as heat, work, and entropy production, were defined at the level of dynamic trajectory [4,5]; the deep connection between entropy production and time-reversal asymmetry was discovered [5,6]; a hierarchy of exact identities known as *fluctuation theorems* [1,7,8] and *work relations* [3], have been established. Terms such as *stochastic energetics* [4] and *stochastic thermodynamics* [1] have been coined to describe this very diversified and important field, which combines stochastic methods with mechanical and thermodynamic concepts. More recently serious efforts have been spent to generalize these theories to strongly coupled systems [9–13], system driven by multiplicative noises [14,15], systems in moving frames [16], systems with position-dependent temperature [17–19], as well as quantum systems [20,21]. Nonetheless, there does not yet exist a single unified theoretical formulation that is capable of addressing all these issues. Additionally, even though formulation of Langevin dynamics in curved space was attempted by Graham [22] more than

forty years ago, stochastic thermodynamics of Langevin systems with curved state spaces have not been developed.

The present paper is the third of a sequel dedicated to a general theory of nonequilibrium statistical physics and thermodynamics using the formalism nonlinear Langevin dynamics and Fokker-Planck dynamics. In the first paper [23], a covariant theory of nonlinear Langevin dynamics and Fokker-Planck dynamics was developed using Ito calculus, which are applicable to system driven by multiplicative noises. One distinguishing feature of the covariant Langevin equation is the appearance of derivatives of kinetic coefficients, i.e., the *spurious drift*, which are important to guarantee detailed balance. Another feature is that differential of slow variables transform not as usual vectors, but according to Ito's formula. In the second paper [13], we and collaborator discussed the proper definitions of thermodynamic variables for small Hamiltonian systems that are strongly coupled to environment, such that the interaction is not negligible comparing with the system Hamiltonian. It was demonstrated that, if the dynamics of bath variables is much faster than that of the system variables, the weak coupling theory of stochastic thermodynamics is also applicable in the strong coupling regime, as long as we identify the fluctuating internal energy as the Hamiltonian of mean force. This theory should be contrasted with other strong coupling theories, where various thermodynamic quantities need to be defined differently than those in the weak coupling regime. It was also demonstrated that, at both trajectory and ensemble levels, heat is negative the change of environmental entropy, conditioned on the system variables.

In the present paper, we shall combine the results of the previous two papers [13,23] to formulate a covariant nonequilibrium thermodynamic theory for small classical systems that are coupled to generally multiplicative noises. It is further

*xxing@sjtu.edu.cn

assumed that the space of slow variable may be curved, or curvilinear coordinates are used, so that the metric tensor plays a serious role. As in the previous two papers [13,23], we shall assume that our system is in contact with a single bath with fixed temperature, and is driven by certain time-dependent control parameter. The system dynamics is described by Ito-Langevin dynamics with white Gaussian noises. If the the control parameter is fixed, the system converges to a thermal equilibrium, which exhibits detailed balance. (This excludes systems driven by nonconservative forces, and hence do not have a globally defined Hamiltonian.) Within this setting, we define nonequilibrium work and heat in terms of system variables, both at trajectory level and at ensemble level, and show that the total entropy always increases monotonically, regardless of the initial state of the system. We also relate time-reversal asymmetry of path probability to entropy production, and prove Crook's fluctuation theorem and Jarzynski equality.

The main message of this paper is that if the metric tensor or the kinetic coefficients depend on system variables, spurious drift shows up in the Langevin equation, and the commonly accepted theories of stochastic energetics and stochastic thermodynamics (which hereafter shall be referred to as *the common theories*) need to be replaced by our covariant theory. This is due to two important reasons. Firstly, in the common theories, Langevin equations are formulated on the idea of force balance. These equations are generally incorrect when spurious drift play a role. Secondly, in the common theories, heat is defined, *à la* Sekimoto [4], as the work by friction and random forces. This definition is generally inapplicable to nonlinear Langevin dynamics where friction and reactive force can not be clearly identified. More seriously, in the presence of spurious drift, Sekimoto's definition of heat leads to violation of the second law of thermodynamics. By strong contrast, our theory is free of these pathologies. In Sec. VI, we discuss three systems where the common theories fail, whereas our theory gives correct results. In these examples, the spurious drifts are caused by, respectively, curvilinear coordinates, curved space, and multiplicative noises.

There are many more sophisticated models of stochastic dynamics where spurious drift shows up, such as rotational diffusion of anisotropic objects and spin dynamics as described by stochastic Landau-Lifshitz-Gilbert equation [24]. More interestingly, quantum Markov processes can be described as Brownian motion in Hilbert space [25], which is also curved due to normalization. Strong interactions between a system and its environment may also lead to state-dependent noise correlations, which also leads to spurious drift. Hence it is fair to say that in nonequilibrium statistical physics, spurious drifts are rule of thumb rather than exceptions. We will explore the above mentioned problems in the future.

The remaining of this paper is organized as follows. In Sec. II, we adapt the covariant Langevin dynamics developed in Ref. [23] to systems whose slow variables form a Riemannian manifold. In Sec. III we define thermodynamic quantities and establish the first and second laws at the ensemble level. In Sec. IV, we define thermodynamic quantities at the level of dynamic trajectory, establish the connection between entropy production and time-reversal asymmetry of path probability, derive Crooks fluctuation theorem and Jarzynski equality. In

Sec. V we discuss the connections between our theory and the classical irreversible thermodynamics as well as Hamiltonian dynamics. We also show that our theory is consistent with the common theories in the case of classical point particles coupled to additive noises. In Sec. VI we discuss three concrete examples of Langevin dynamics with spurious drift, where the common theories fail but our covariant theory give the correct results. Finally in Sec. VII we draw conclusive remarks and outline future directions. In Sec. A we provide a detailed derivation of the short time transition probability for nonlinear Langevin dynamics with multiplicative noises.

II. COVARIANT LANGEVIN DYNAMICS ON RIEMANNIAN MANIFOLD

In this section, we shall reformulate the covariant Langevin theory developed in Ref. [23] for systems whose slow variables form a Riemannian manifold.

A. Brief review of covariant Langevin dynamics

The covariant form of nonlinear Langevin equation with multiplicative Gaussian white noises was derived in Ref. [23],

$$dx^i(t) + (L^{ij}\partial_j U - \partial_j L^{ij})dt = b^{i\alpha}dW_\alpha(t), \qquad (2.1)$$

where $\boldsymbol{x} = (x^1, \cdots, x^n)$ are the slow variables, $\partial_j = \partial/\partial x^i$, and $dW_\alpha(t)$ are the Wiener noises, which satisfy

$$dW_\alpha(t)dW_\beta(t) = \delta_{\alpha\beta}dt. \qquad (2.2)$$

In the left-hand side (LHS) of Eq. (2.1), $L^{ij}(\boldsymbol{x})$ are *the kinetic coefficients*, and $b^{i\alpha}(\boldsymbol{x})$ *the noise amplitudes*, both generally depending on the slow variables $\boldsymbol{x}$. The term $-\partial_j L^{ij}dt$ is called *spurious drift*, and shows up when kinetic coefficients depend on $\boldsymbol{x}$, i.e., the noises are multiplicative. Throughout the paper, repeated indices are summed over, unless otherwise specified. In the right-hand side (RHS) of Eq. (2.1), the product $b^{i\alpha}(\boldsymbol{x})dW_\alpha(t)$ is interpreted in Ito's sense. For definitions of Ito product and Stratonovich product of stochastic variables, we refer to Ref. [26]. The matrix $L^{ij}(\boldsymbol{x})$ is decomposed into a symmetric part $B^{ij}(\boldsymbol{x})$ and an antisymmetric part $Q^{ij}(\boldsymbol{x})$,

$$L^{ij}(\boldsymbol{x}) = B^{ij}(\boldsymbol{x}) + Q^{ij}(\boldsymbol{x}), \qquad (2.3a)$$

$$B^{ij}(\boldsymbol{x}) = B^{ji}(\boldsymbol{x}), \qquad (2.3b)$$

$$Q^{ij}(\boldsymbol{x}) = -Q^{ji}(\boldsymbol{x}). \qquad (2.3c)$$

The matrix $b^{i\alpha}(\boldsymbol{x})$ is related to $B^{ij}(\boldsymbol{x})$ via

$$b^{i\alpha}(\boldsymbol{x})b^{j\alpha}(\boldsymbol{x}) = 2B^{ij}(\boldsymbol{x}) = L^{ij}(\boldsymbol{x}) + L^{ji}(\boldsymbol{x}). \qquad (2.3d)$$

All eigenvalues of $B^{ij}(\boldsymbol{x})$ are non-negative, otherwise either normalization or positivity of probability is violated.

In Ref. [23] we define $p(\boldsymbol{x}, t)d^n\boldsymbol{x}$ to be the probability that the slow variables take value in the infinitesimal volume $d^n\boldsymbol{x}$ centered at $\boldsymbol{x}$. Hence $p(\boldsymbol{x}, t)$ is the probability density function (pdf) of slow variables. The function $U(\boldsymbol{x})$ in Eq. (2.1) is called *the generalized potential*, which is related to the steady state pdf $p^{\mathrm{SS}}(\boldsymbol{x})$ via

$$p^{\mathrm{SS}}(\boldsymbol{x}) = e^{-U(\boldsymbol{x})}. \qquad (2.4)$$

If the system has a well-defined Hamiltonian and is in contact with a single heat bath (as we will assume in this paper),

the steady state is the thermal equilibrium state. If the slow variable space is infinite, $U(x)$ must be bounded from below and diverge sufficiently fast as $x \to \infty$, in order for $p^{\text{SS}}(x)$ to be normalizable.

The covariant form of Fokker-Planck equation associated with the Langevin dynamics (2.1) is given by

$$\partial_t \, p(x,t) = L_{\text{FP}} \, p(x,t)$$
$$= \partial_i L^{ij} (\partial_j + (\partial_j U)) \, p(x,t). \quad (2.5)$$

As demonstrated in Ref. [23], under nonlinear transformation of variables (NTV) $x = (x^i) \to x' = (x'^a)$, $L^{ij}(x)$ and $b^{i\alpha}(x)dW_\alpha$ transform respectively as contravariant tensor and vector, whereas $e^{-U(x)}d^n x$ and $p(x,t)d^n x$ transform as scalars. In the proof of covariance, Eq. (2.2) plays an essential role. Qualitatively speaking, Wiener noise $dW_\alpha(t)$ scales as $\sqrt{dt}$, and hence according to Eq. (2.1), dx receives contributions from noises of order $\sqrt{dt}$ and from the systematic forces (the second term on LHS) of order dt. As a consequence, to calculate the variation of a function $f(x)$ up to order dt, we must expand $f(x + dx)$ up to order dx^2,

$$df(x) = f(x + dx) - f(x)$$
$$= \partial_j f(x) \, dx^j + \tfrac{1}{2} \partial_i \partial_j f(x) \, dx^i dx^j. \quad (2.6)$$

The quadratic terms $dx^i dx^j$ can be further simplified using Langevin equation (2.1) and Eq. (2.2),

$$dx^i dx^j = \sum_\alpha b^{i\alpha} b^{j\alpha} dt + O(dt^{3/2}), \quad (2.7)$$

where the neglected terms vanish in the continuum limit. This leads to *Ito's formula* [26],

$$df(x) = \partial_j f(x) \, dx^j + B^{ij} \partial_i \partial_j f(x) dt. \quad (2.8)$$

It is important to note that Eq. (2.8) should be understood as an equality of stochastic variables, which holds in probability, not just in average.

B. Langevin dynamics on Riemannian manifold

The covariant Langevin equation (2.1) does not involve metric tensor, hence the range of applicability is not limited to Riemannian manifold. On the other hand, the pdf $p(x,t)$ and $e^{-U(x)}$ transform as densities, but not as scalars. For many physical systems, there is a natural volume measure $dv(x)$ defined in the manifold of slow variables, such that the volume of a region Ω is given by

$$\text{Vol}[\Omega] = \int_\Omega dv(x). \quad (2.9)$$

The volume measure must be invariant under NTV, which means that in the new coordinates we have $dv'(x') \equiv dv(x)$.

For example, if the manifold of slow variables is a Riemann manifold with a metric tensor g_{ij}, it has a natural volume measure

$$dv(x) = \sqrt{g(x)} \, d^n x, \quad (2.10)$$

where $g = \det(g_{ij})$ is the determinant of the covariant metric tensor. As another example, we consider a classical Hamiltonian system coupled to a heat bath, the set of slow variables consists of canonical coordinates and momenta $x = (q, p)$. It

is well known that Poisson brackets and Liouville volume measure are preserved by all canonical variable transformations in the phase space. Hence the volume measure, i.e., *the Liouville measure*, $d^n x = \prod_i dp_i dq^i$, is invariant under all canonical variable transformations.

In the remaining of this paper, we shall always assume that the slow variable manifold is Riemannian with metric tensor $g_{ij}(x)$, and with an invariant volume measure $\sqrt{g(x)} \, d^n x$. To apply the theory to Hamiltonian systems, we only need to let $\sqrt{g(x)} = 1$, and restrict the nonlinear transformations of variables to canonical transformations. Throughout the paper we shall use the following simplified notation for integration over the invariant measure:

$$\int_x \equiv \int \sqrt{g(x)} d^n x = \int \sqrt{g'(x')} d^n x'. \quad (2.11)$$

It is then convenient to define an invariant pdf $\wp(x)$ and invariant generalized potential $\mathscr{U}(x)$ via

$$\wp(x)\sqrt{g(x)} \, d^n x = p(x)d^n x, \quad (2.12)$$

$$e^{-\mathscr{U}(x)}\sqrt{g(x)} \, d^n x = e^{-U(x)} d^n x. \quad (2.13)$$

Hence we have the following relations:

$$p(x) = \sqrt{g(x)} \, \wp(x), \quad (2.14)$$

$$U(x) = \mathscr{U}(x) - \log \sqrt{g(x)}. \quad (2.15)$$

Using these, we can rewrite the Langevin equation (2.1) and Fokker-Planck equation (2.5) as

$$dx^i + \left(L^{ij} \partial_j \mathscr{U} - \frac{1}{\sqrt{g}} \partial_j \sqrt{g} L^{ij} \right) dt = b^{i\alpha} dW_\alpha, \quad (2.16)$$

$$\partial_t \wp = \frac{1}{\sqrt{g}} \partial_i \sqrt{g} L^{ij} (\partial_j + (\partial_j \mathscr{U})) \wp = \mathscr{L}_{\text{FP}} \wp. \quad (2.17)$$

We shall call the term $-(1/\sqrt{g})(\partial_j \sqrt{g} L^{ij})dt$ in Eq. (2.16) the *spurious drift*, which shows up whenever the kinetic coefficients L^{ij} or the metric determinant g depends on the slow variables x.

It is then easy to see that the steady state is

$$\wp^{\text{SS}}(x) = e^{-\mathscr{U}(x)} \quad (2.18)$$

and satisfies Eq. (2.17). The invariant Fokker-Planck operator $\mathscr{L}_{\text{FP}}$ in Eq. (2.17) is

$$\mathscr{L}_{\text{FP}} = \frac{1}{\sqrt{g}} \partial_i \sqrt{g} L^{ij} (\partial_j + (\partial_j \mathscr{U})), \quad (2.19)$$

which is related to L_{FP} defined in Eq. (2.5) via

$$\mathscr{L}_{\text{FP}} = \frac{1}{\sqrt{g}} L_{\text{FP}} \sqrt{g}. \quad (2.20)$$

Now the transformation laws of various components of Eqs. (2.16) and (2.17) can be obtained from the corresponding results in Ref. [23]. Specifically dx^i transforms according to the Ito's formula,

$$dx' \to dx'^a = \frac{\partial x'^a}{\partial x^i} dx^i + \frac{\partial^2 x'^a}{\partial x^i \partial x^j} B^{ij} dt, \quad (2.21)$$

whilst $\wp(\boldsymbol{x})$ and $\mathscr{U}(\boldsymbol{x})$ transform as scalars, whereas $b^{i\alpha}(\boldsymbol{x})$ and $L^{ij}(\boldsymbol{x})$ transform respectively as contravariant vector and tensor,

$$\wp(\boldsymbol{x}) \rightarrow \wp'(\boldsymbol{x}') = \wp(\boldsymbol{x}), \tag{2.22a}$$

$$\mathscr{U}(x) \rightarrow \mathscr{U}'(\boldsymbol{x}') = \mathscr{U}(\boldsymbol{x}), \tag{2.22b}$$

$$b^{i\alpha}(\boldsymbol{x}) \rightarrow b'^{a\alpha}(\boldsymbol{x}') = \frac{\partial x'_a}{\partial x^i} b^{i\alpha}(\boldsymbol{x}), \tag{2.22c}$$

$$L^{ij}(\boldsymbol{x}) \rightarrow (L')^{ab}(\boldsymbol{x}') = \frac{\partial x'^a}{\partial x^i} L^{ij}(\boldsymbol{x}) \frac{\partial x'^b}{\partial x^j}. \tag{2.22d}$$

The metric transforms as rank two covariant tensor

$$g_{ij} \rightarrow g'_{ab} = \frac{\partial x^i}{\partial x'^a} g_{ij} \frac{\partial x^j}{\partial x'^b}, \tag{2.23}$$

such that $g_{ij} dx^i dx^j = g_{ab} dx'^a dx'^b$ transforms as a scalar. Using Eq. (2.20a) of Ref. [23], we can also see that the invariant Fokker-Planck operator, defined in Eq. (2.19) indeed transforms as a scalar,

$$\mathscr{L}'_{\text{FP}}(\boldsymbol{x}') = \mathscr{L}_{\text{FP}}(\boldsymbol{x}). \tag{2.24}$$

Note that our notations are consistent with usual tensor analysis, where upper indices and lower indices are used respectively for contravariant and covariant objects.

Note that if we choose $L^{ij} = (g^{-1})^{ij}$ to be the contravariant metric tensor and $\mathscr{U}(\boldsymbol{x})$ constant, the Fokker-Planck operator (2.19) becomes the covariant Laplacian in Riemannian manifold: $\Delta = \sqrt{g}^{-1} \partial_i \sqrt{g} (g^{-1})^{ij} \partial_j$, which is known to be invariant under reparameterization.

An important but surprising fact about Eq. (2.16) is that neither dx^i nor $-L^{ij}\partial_j(\mathscr{U} - \log\sqrt{g})dt + \partial_j L^{ij} dt$ transforms as a vector. But the linear combination $dx^i + L^{ij}\partial_j(\mathscr{U} - \log\sqrt{g})dt - \partial_j L^{ij} dt$ appearing in the LHS of Eq. (2.16) does behave as a contravariant vector, and so is the Ito product $b^{i\alpha} dW_\alpha$ in the RHS.

In Ref. [23], $b^{i\alpha}$, L^{ij}, and U are assumed to be independent of time. However it is easy to realize that the formalism still works if these functions are time dependent. In this paper, we shall assume that $b^{i\alpha}$, L^{ij}, B^{ij}, Q^{ij} and $\mathscr{U}$ may depend on a set of parameters λ, which may be externally controlled to vary over time. This is necessary since we aim to study nonequilibrium processes in small systems. When there is no danger of confusion, we will hide their dependence on $\lambda(t)$. We shall always assume that the metric tensor **g** is time independent. There are problems where the metric structure of the slow variables changes with time, such as the diffusion problem on a deformable membrane. Note also that we only consider NTV independent of time. In Ref. [16], Speck *et al.* use time-dependent variable transformation to study stochastic thermodynamics in moving frames. We shall not touch this issue here.

C. Reversibility and detailed balance

As specified in the Introduction, the system being studied has a well defined Hamiltonian and is in contact with a single heat bath. For fixed control parameter, the system converges to a unique thermal equilibrium state. The system dynamics then has time-reversal symmetry, whose implications are discussed here.

As in Ref. [23], we assume that each slow variable has definite time-reversal symmetry. Hence under time reversal the slow variables transform as $\boldsymbol{x} \rightarrow \boldsymbol{x}^*$, where $\boldsymbol{x}^* = (\varepsilon_1 x^1, \varepsilon_2 x^2, ..., \varepsilon_n x^n)$ with $\varepsilon_i = +1, -1$ for even and odd variables respectively. Take Hamiltonian systems as an example, we have $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p})$ and $\boldsymbol{x}^* = (\boldsymbol{q}, -\boldsymbol{p})$, where $\boldsymbol{q}$ are the canonical coordinates and $\boldsymbol{p}$ the canonical momenta. For all problems we know the metric tensor is itself invariant under time reversal,

$$\varepsilon_i g_{ij}(\boldsymbol{x}^*) \varepsilon_j = g_{ij}(\boldsymbol{x}), \tag{2.25a}$$

where no summation is implied about repeated indices. Hence the invariant volume measure $dv(\boldsymbol{x})$ is also invariant under time reversal,

$$g(\boldsymbol{x}^*) = g(\boldsymbol{x}), \quad dv(\boldsymbol{x}^*) = dv(\boldsymbol{x}). \tag{2.25b}$$

The external control parameter λ transforms under time reversal as $\lambda \rightarrow \lambda^*$. For a moment we shall assume that λ does not vary over time. Two typical examples where λ changes sign under time reversal are magnetic field and angular velocity. Note that if $\lambda \neq \lambda^*$, the process with external parameter λ is different from that associated with λ^*. It is customary to call the process with parameter λ the *forward process* and that with parameter λ^* the *backward process*. The backward process is also described by Langevin equation (2.16) and Fokker Planck equation (2.17), but with $\mathscr{U}(\boldsymbol{x}, \lambda)$, $b^{i\alpha}(\boldsymbol{x}, \lambda)$, $L^{ij}(\boldsymbol{x}, \lambda)$ replaced by $\mathscr{U}(\boldsymbol{x}, \lambda^*)$, $b^{i\alpha}(\boldsymbol{x}, \lambda^*)$, $L^{ij}(\boldsymbol{x}, \lambda^*)$. Accordingly, the steady-state distributions of the forward and backward processes are

$$\wp_F(\boldsymbol{x}_1) = e^{-\mathscr{U}(\boldsymbol{x}_1, \lambda)}, \tag{2.26a}$$

$$\wp_B(\boldsymbol{x}_1^*) = e^{-\mathscr{U}(\boldsymbol{x}_1^*, \lambda^*)}. \tag{2.26b}$$

Since $(\lambda^*)^* = \lambda$, the backward of the backward process is the forward process itself.

Let $\wp_F(\boldsymbol{x}_2, t_2; \boldsymbol{x}_1, t_1)$ and $\wp_B(\boldsymbol{x}_2, t_2; \boldsymbol{x}_1, t_1)$ be the two-time joint pdfs of the forward and backward processes respectively, both assumed in the stationary regime. Here and below the subscripts in boldface variables $\boldsymbol{x}_1, \boldsymbol{x}_2 \cdots$ denote time, not the components of slow variables. These pdfs must have time-translational symmetry, i.e., they can only depend on the time difference $t_2 - t_1$. A stochastic dynamics is said to to be *reversible* if the following relation is satisfied:

$$\wp_F(\boldsymbol{x}_2, t_2; \boldsymbol{x}_1, t_1) dv(\boldsymbol{x}_2) dv(\boldsymbol{x}_1)$$
$$= \wp_B(\boldsymbol{x}_1^*, -t_1; \boldsymbol{x}_2^*, -t_2) dv(\boldsymbol{x}_2^*) dv(\boldsymbol{x}_1^*). \tag{2.27a}$$

Because of Eq. (2.25b), the volume elements in two sides can be dropped, which leads to

$$\wp_F(\boldsymbol{x}_2, t_2; \boldsymbol{x}_1, t_1) = \wp_B(\boldsymbol{x}_1^*, -t_1; \boldsymbol{x}_2^*, -t_2). \tag{2.27b}$$

Because of stationarity, these joint pdfs only depend on $t_2 - t_1$, and Eq. (2.27b) is equivalent to

$$\wp_F(\boldsymbol{x}_2, t_2; \boldsymbol{x}_1, t_1) = \wp_B(\boldsymbol{x}_1^*, t_2; \boldsymbol{x}_2^*, t_1). \tag{2.27c}$$

Note also that this condition is symmetric with respect to the exchange of the forward and backward processes, hence the forward process is reversible if and only if the backward is so. Furthermore, a process may be reversible even if it is not the same as the backward process, i.e. if $\lambda \neq \lambda^*$. Stationarity

together with Eqs. (2.27) imply that the system is in thermal equilibrium, i.e., the steady states are actually thermal equilibrium states.

Our definition of reversibility may appear confusing because in quantum mechanics, it is often said that magnetic field (and rotation of frame) breaks time-reversal symmetry. In this paper, however, we shall use the term "irreversible" to describe dissipative processes that produce entropy. Magnetic field and rotation of frame only shift the equilibrium state, but do not lead to a dissipative nonequilibrium state. Hence, according to our terminology, they do not change the reversible nature of the dynamic process.

Integrating Eq. (2.27b) over x_2, we obtain $p_F(x_1, t_1) = p_B(x_1^*, -t_2)$. But these one-time pdfs are given by Eq. (2.26). Hence reversibility implies

$$\mathscr{U}(x^*, \lambda^*) = \mathscr{U}(x, \lambda). \tag{2.28}$$

Because the processes are Markovian, the two-time pdfs can be written as products of transition probabilities, i.e., conditional pdfs, and one-time pdf's,

$$p_F(x_2, t_2; x_1, t_1) = p_F(x_2, t_2|x_1, t_1)p_F(x_1),$$
$$p_B(x_1^*, -t_1; x_2^*, -t_2) = p_B(x_1^*, -t_1|x_2^*, -t_2)p_B(x_2^*). \tag{2.29}$$

Combining these with Eqs. (2.26), we obtain another implication of reversibility,

$$\frac{p_F(x_2, t_2|x_1, t_1)}{p_B(x_1^*, -t_1|x_2^*, -t_2)} = e^{-\mathscr{U}(x_2, \lambda) + \mathscr{U}(x_1, \lambda)}. \tag{2.30}$$

It is important to note that RHS is independent of time, whereas in LHS, the conditional pdfs only depend on time difference $\Delta t = t_2 - t_1$, as well as on the control parameter λ.

Generalizing the proof presented in Ref. [23], we can show that Eqs. (2.28) and (2.30) are equivalent to

$$\mathscr{U}(x^*, \lambda^*) = \mathscr{U}(x, \lambda), \tag{2.31a}$$

$$\varepsilon_i L^{ij}(x^*, \lambda^*)\varepsilon_j = L^{ji}(x, \lambda), \tag{2.31b}$$

$$\int_x e^{-\mathscr{U}(x, \lambda)} = 1. \tag{2.31c}$$

Note that Eq. (2.31b) is equivalent to

$$\varepsilon_i B^{ij}(x^*, \lambda^*)\varepsilon_j = B^{ij}(x, \lambda), \tag{2.31d}$$

$$\varepsilon_i Q^{ij}(x^*, \lambda^*)\varepsilon_j = -Q^{ij}(x, \lambda). \tag{2.31e}$$

Equations (2.31) are called *detailed balance conditions*. For Markov processes, Eqs. (2.27), (2.33), and (2.31) are all equivalent. With detailed balance satisfied, the steady state $e^{-U(x, \lambda)}$ then describes thermal equilibrium. An alternative proof of these conditions using Green's function will be presented in a separate paper [27]. Note also that Eq. (2.31c), even though not shown in most papers, is important. It can be shown that if $e^{-U(x, \lambda)}$ is not normalizable, the system converges to either a dissipative nonequilibrium steady state, or some other state that constantly changing with time. These will be studied in a separate paper.

Because of the Markovian property, we can construct N-time joint pdfs in the stationary regime from initial pdfs and transition probabilities, both for the forward process and for the backward process,

$$p_F(x_N, t_N; \cdots; x_1, t_1; x_0, t_0) = p_F(x_N, t_N|x_{N-1}, t_{N-1}) \cdots p_F(x_1, t_1|x_0, t_0) e^{-\mathscr{U}(x_0, \lambda)}, \tag{2.32a}$$

$$p_B(x_0^*, t_N; \cdots; x_{N-1}^*, t_1; x_N^*, t_0) = p_B(x_0^*, t_N|x_1^*, t_{N-1}) \cdots p_B(x_{N-1}^*, t_1|x_N^*, t_0) e^{-\mathscr{U}(x_N^*, \lambda^*)}. \tag{2.32b}$$

It is understood that $t_N > t_{N-1} > \cdots > t_1 > t_0$ and that time propagates from right to left in these expressions. Taking the ratio of Eqs. (2.32a) and (2.32a), and using Eqs. (2.30) and (2.28) repeatedly, we find

$$p_F(x_N, \ldots, x_1, x_0) = p_B(x_0^*, x_1^*, \ldots, x_N^*). \tag{2.33}$$

It is useful to think of $(x_N, \ldots, x_1, x_0)$ as a discretized dynamic trajectory, and $(x_0^*, x_1^*, \ldots, x_N^*)$ as its time reversal. Equation (2.33) then says that, for stationary reversible Markov process, the probability of dynamic trajectory is invariant under time reversal of both the path and the dynamic protocol, i.e., reversal of the parameter λ. In other words, a trajectory in the forward process is equally probable as the backward trajectory in the backward process. Combining Eqs. (2.32) and (2.33), we may also obtain

$$\frac{p_F(x_N|x_{N-1}) \cdots p_F(x_1|x_0)}{p_B(x_0^*|x_1^*) \cdots p_B(x_{N-1}^*|x_N^*)} = e^{\mathscr{U}(x_0, \lambda) - \mathscr{U}(x_N, \lambda)}. \tag{2.34}$$

We note however that Eqs. (2.31) and (2.34) are derived by assuming that the parameter λ is fixed. If λ change over time, neither the forward process and the backward process can be

stationary, yet the detailed balance conditions, Eqs. (2.31), are still satisfied. Equations (2.33) and (2.34), however, will be replaced by more complicated relations, see Eqs. (4.31) and (4.30).

III. THERMODYNAMICS AT ENSEMBLE LEVEL

We shall now study the nonequilibrium thermodynamics associated with the covariant Langevin dynamics (2.16) and Fokker-Planck dynamics (2.17), assuming that the external parameter $\lambda(t)$ is tuned externally as a function of time, and that detailed balance is satisfied for every fixed λ. We start with the setting discussed in Ref. [13], such that at the microscopic level, the total Hamiltonian of the system and the bath is

$$H_{\text{tot}} = H_X(x; \lambda) + H_B(y; x), \tag{3.1}$$

where x, y are variables of the system and the bath separately. Unlike in Ref. [13], however, here we further assume that the external parameter λ is coupled to $H_X(x; \lambda)$ but not to $H_B(y; x)$. The decomposition of the total Hamiltonian in

Eq. (3.1) is such that $H_{\mathbf{X}}(\boldsymbol{x}; \lambda)$ is the Hamiltonian of mean force (HMF) of the system. Hence the marginal equilibrium pdf of $\boldsymbol{x}$ is

$$p_{\mathbf{X}}^{\mathrm{EQ}}(\boldsymbol{x}) = \frac{1}{Z_{\mathbf{X}}} e^{-\beta H_{\mathbf{X}}} = e^{-\beta H_B + \beta F(\lambda)}. \tag{3.2}$$

where $F(\lambda)$ is the equilibrium free energy

$$F(\lambda) = -T \log \int_{\boldsymbol{x}} e^{-\beta H_{\mathbf{X}}(\boldsymbol{x}, \lambda)}. \tag{3.3}$$

We refer the readers to Ref. [13] for detailed construction of this decomposition. In the limit of time-scale separation, where the dynamics of $\boldsymbol{y}$ (bath variables) is much faster than that of $\boldsymbol{x}$ (system variables), we expect that, after taking into account fluctuations of $\boldsymbol{y}$, the distribution of $\boldsymbol{x}$ evolves according to the Fokker-Planck equation (2.17). If we follow the dynamic trajectories of $\boldsymbol{x}$, then the dynamics is described by the nonlinear Langevin equation (2.16), whereas the fast variables behave as Gaussian white noises. A mathematical derivation from deterministic unitary microscopic dynamics to mesoscopic Langevin dynamics or Fokker-Planck dynamics can be achieved using projection operator methods.

Comparing Eqs. (3.2) with (2.4), we find

$$\mathscr{U}(\boldsymbol{x}, \lambda) = \beta H_{\mathbf{X}}(\boldsymbol{x}, \lambda) - \beta F(\lambda). \tag{3.4}$$

The condition of detailed balance Eq. (2.31a) is then translated into

$$H_{\mathbf{X}}(\boldsymbol{x}^*, \lambda^*) = H_{\mathbf{X}}(\boldsymbol{x}, \lambda), \quad F(\lambda^*) = F(\lambda). \tag{2.31a$'$}$$

Since both Langevin equation (2.16) and Fokker-Planck equation (2.17) depend on $\mathscr{U}$ only through $\partial_i \mathscr{U}$, we see that the additive constant $\beta F(\lambda)$ in Eq. (3.4) can not be determined from the dynamics of slow variables. Instead it has to obtained from study of the statistical mechanics of the total system consisting of both the slow variables and fast variables.

Thermodynamic quantities are also defined in the same way as in Ref. [13]. We identify $H_{\mathbf{X}}(\boldsymbol{x}, \lambda)$ as the *fluctuating internal energy* of the system, and its ensemble average as the *internal energy*. Nonequilibrium entropy and free energy of the system are defined as

$$S[p] \equiv -\int_{\boldsymbol{x}} p(\boldsymbol{x}) \log p(\boldsymbol{x}), \tag{3.5a}$$

$$F[p(\boldsymbol{x})] \equiv \int_{\boldsymbol{x}} p(\boldsymbol{x})(H_{\mathbf{X}}(\boldsymbol{x}, \lambda) + T \log p(\boldsymbol{x}))$$
$$= \langle H_{\mathbf{X}}(\boldsymbol{x}, \lambda) \rangle - T S[p(\boldsymbol{x})]. \tag{3.5b}$$

Now consider an infinitesimal process where time evolves by dt, the parameter λ changes by $d\lambda$, and the pdf of slow variables changes by $dp = \mathscr{L}_{\mathrm{FP}} p \, dt$ according to Eq. (2.17). We define the ensemble averaged differential work and heat as

$$đW \equiv \int_{\boldsymbol{x}} p \, d_\lambda H_{\mathbf{X}} = \int_{\boldsymbol{x}} p \, (\partial_\lambda H_{\mathbf{X}}) d\lambda, \tag{3.6a}$$

$$đQ \equiv \int_{\boldsymbol{x}} H_{\mathbf{X}} \, dp = \int_{\boldsymbol{x}} H_{\mathbf{X}} (\mathscr{L}_{\mathrm{FP}} p) \, dt, \tag{3.6b}$$

where $d_\lambda H_{\mathbf{X}} = (\partial_\lambda H_{\mathbf{X}}) d\lambda$ is the differential of $H_{\mathbf{X}}$ due to variation of λ. It is shown in Ref. [13] that $\beta đQ$ as defined

in (3.6b) equals to negative the variation of environmental entropy, conditioned on the slow variables $\boldsymbol{x}$. Note that even though the kinetic matrices B^{ij}, Q^{ij} may also depend on λ, only the λ dependence of $H_{\mathbf{X}}$ contributes to the work $đW$.

It is important to note that work and heat are defined in Eqs. (3.6) in terms of Hamiltonian of mean force $H_{\mathbf{X}}$, but not in terms of the generalized potential $\mathscr{U}$. Nonetheless, $H_{\mathbf{X}}$ and $T \mathscr{U}$ differ from each other only by an additive constant $F(\lambda)$, see Eq. (3.4). Because of the particular form of the operator $\mathscr{L}_{\mathrm{FP}}$ [Eq. (2.19)], one easily see via integration by parts that the heat can be equivalently defined in terms of $\mathscr{U}$ as

$$đQ \equiv T \int_{\boldsymbol{x}} \mathscr{U} \, dp = T \int_{\boldsymbol{x}} \mathscr{U} (\mathscr{L}_{\mathrm{FP}} p) \, dt. \tag{3.6b$'$}$$

By contrast, the work cannot be rewritten in terms of $\mathscr{U}$ because the constant $F(\lambda)$ in Eq. (3.4) depends on λ. The fact that work cannot be fully determined using the information contained in the Langevin equation should not surprise us. It is in fact shared by the commonly accepted theories of stochastic energetics and stochastic thermodynamics as well. To explain this issue more clearly, let us consider adding a time-dependent constant potential $V(t) = V_0 t/T$ to a Hamiltonian system, where V_0 is independent of coordinates and momenta. It of course means that the external agent does work of the same amount to the system. Yet introduction of V_0 does not affect the dynamic equations of the system, since $\partial_q V(t) = \partial_p V(t) = 0$. As will be shown in Sec. V C, our definitions of heat and work reduce to those of Sekimoto for the case of additive noises.

The differential of the internal energy is

$$d\langle H_{\mathbf{X}}(\boldsymbol{x}, \lambda) \rangle = \int_{\boldsymbol{x}} p \, dH_{\mathbf{X}} + \int_{\boldsymbol{x}} H_{\mathbf{X}} dp$$
$$= đW + đQ, \tag{3.7}$$

which has the form of *the first law of thermodynamics* at the ensemble level. We can similarly obtain a differential relation for the nonequilibrium free energy $F[p]$,

$$dF = đW + đQ - T \, dS. \tag{3.8}$$

In Ref. [13], it is demonstrated that the differential of the total entropy (the joint Gibbs-Shannon entropy of the system variables and bath variables) is

$$dS^{\mathrm{tot}} = dS - \beta đQ$$
$$= -dt \int_{\boldsymbol{x}} (\log p + U) \mathscr{L}_{\mathrm{FP}} p, \tag{3.9}$$

where in the last equality we have used Eqs. (3.6b) and (3.4), as well as

$$dS[p] = -\int_{\boldsymbol{x}} \log p(\boldsymbol{x}) \, dp(\boldsymbol{x}),$$
$$= -dt \int_{\boldsymbol{x}} \log p(\boldsymbol{x}) \mathscr{L}_{\mathrm{FP}} p. \tag{3.10}$$

Now inserting Eqs. (2.19) and (2.11) into Eq. (3.9), and integrating by parts [28], we further have

$$\frac{dS^{\mathrm{tot}}}{dt} = \int_{\boldsymbol{x}} ((\partial_j + \partial_i \mathscr{U}) p) \frac{B^{ij}}{p} ((\partial_j + \partial_j \mathscr{U}) p), \tag{3.11}$$

which is manifestly positive, since the matrix B^{ij} is so. Combining this with Eqs. (3.8) and (3.9), we obtain

$$dS^{\text{tot}} = dS - \beta đQ \geqslant 0, \quad (3.12a)$$

$$dS^{\text{tot}} = đW - dF \geqslant 0. \quad (3.12b)$$

Whilst Eq. (3.12a) is called the Clausius inequality, Eq. (3.12b) is usually known as the *principle of minimal work*, which says that the minimal work needed for a process is the change of the system free energy. In classical thermodynamics, these inequalities are valid only for processes starting from and ending at equilibrium states. In the present theory, however, they are valid for arbitrary nonequilibrium processes.

Note that all thermodynamic variables have been constructed using $H_{\mathbf{X}} = T\mathscr{U} + F$, p, as well as the invariant volume measure, which transform as scalars under NTV. Consequently, all thermodynamic variables and relations discussed in this section are invariant under NTV. It is also important to note that the definitions of thermodynamic variables are contingent on the identification of fluctuation internal energy as the HMF $H_{\mathbf{X}}$. If the system is driven by nonconservative forces, there is no unambiguous definition of internal energy. Then definitions of heat and work will have to been reevaluated very carefully. This will be discussed in a future paper.

IV. THERMODYNAMICS AT TRAJECTORY LEVEL

A. Work and heat at trajectory level

Consider a small time step dt along a particular trajectory in the forward process, where the slow variables evolve from $\mathbf{x}$ at time t to $\mathbf{x} + d\mathbf{x}$ at time $t + dt$, while λ changes to $\lambda + d\lambda$. In Ref. [13], work at trajectory level was defined as the change of total energy of the universe, averaged over fast variables. It was then proved that the differential work is also the differential of $H_{\mathbf{X}}$ due to the change of λ. By contrast, heat at trajectory level was defined as negative the energy change of the environment, again averaged over fast variables. It was then proved that the heat is also the differential of $H_{\mathbf{X}}$ due to the change of $\mathbf{x}$. In the setting of nonlinear Langevin dynamics, the environmental variables are not explicitly displayed. Hence work and heat at trajectory level are defined in terms of system variables as

$$đW \equiv d_\lambda H_{\mathbf{X}} = (\partial_\lambda H_{\mathbf{X}})\,d\lambda$$
$$= H_{\mathbf{X}}(\mathbf{x}, \lambda + d\lambda) - H_{\mathbf{X}}(\mathbf{x}, \lambda), \quad (4.1a)$$

$$đ\mathscr{Q} \equiv d_x H_{\mathbf{X}} \equiv H_{\mathbf{X}}(\mathbf{x} + d\mathbf{x}, \lambda) - H_{\mathbf{X}}(\mathbf{x}, \lambda). \quad (4.1b)$$

Now the differential of the fluctuating internal energy can be decomposed into work and heat at trajectory level,

$$dH_{\mathbf{X}} = H_{\mathbf{X}}(\mathbf{x} + d\mathbf{x}, \lambda + d\lambda) - H_{\mathbf{X}}(\mathbf{x}, \lambda)$$
$$\approx H_{\mathbf{X}}(\mathbf{x}, \lambda + d\lambda) - H_{\mathbf{X}}(\mathbf{x}, \lambda)$$
$$+ H_{\mathbf{X}}(\mathbf{x} + d\mathbf{x}, \lambda) - H_{\mathbf{X}}(\mathbf{x}, \lambda) + O(dt^{3/2}) \quad (4.2)$$

$$= đW + đ\mathscr{Q} + O(dt^{3/2}). \quad (4.3)$$

For typical trajectories, the work is of order dt whereas the heat is of order $d\mathbf{x} \sim dt^{1/2}$. Hence neglected terms in the above equation are of order $dt^{3/2}$ and hence do not contribute to the continuous limit. Hence we arrive at the first law at trajectory level,

$$dH_{\mathbf{X}} = đW + đ\mathscr{Q} = d_\lambda H_{\mathbf{X}} + d_x H_{\mathbf{X}}. \quad (4.4)$$

Both $d_\lambda H_{\mathbf{X}}$ and $d_x H_{\mathbf{X}}$ transform as scalars under NTV.

The definition of heat, Eq. (4.1b), can be expanded in terms of $d\mathbf{x}$. In view of Eq. (2.7), however, we should expand up to $d\mathbf{x}^2$,

$$đ\mathscr{Q} = (\partial_i H_{\mathbf{X}})dx^i + \frac{1}{2}(\partial_i\partial_j H_{\mathbf{X}})dx^i dx^j$$
$$= (\partial_i H_{\mathbf{X}})dx^i + B^{ij}(\partial_i\partial_j H_{\mathbf{X}})dt, \quad (4.5)$$

where we have also used Eqs. (2.3d).

In order for the thermodynamics at trajectory level to be consistent with the thermodynamics at ensemble level, we need to show that the ensemble averages of differential work and heat at trajectory level, defined in Eqs. (4.1), equal to the average work and heat we defined in Eqs. (3.6). Here ensemble average means averaging both over the pdf $p(\mathbf{x}, t)$ and over the noise $dW_\alpha(t)$, which acts on the system during $(t, t + dt)$. Note from Eq. (4.2) that $đW$ does not depend on noise. Hence to calculate its ensemble average, we only need to multiply Eq. (4.2) by $p(\mathbf{x}, t)\sqrt{g(\mathbf{x})}$ and integrate over $\mathbf{x}$. We find Eq. (3.6a) as expected,

$$\langle đW \rangle = \int_x p(x, t)(\partial_\lambda H_{\mathbf{X}})d\lambda = dW. \quad (4.6)$$

To calculate the ensemble average of $d\mathscr{Q}$, we express dx^i in Eq. (4.5) using the Langevin equation (2.16), use Eq. (3.4) to express $\mathscr{U}$ in terms of $H_{\mathbf{X}}$, and further average over noise. The noise term disappears up on averaging because of Ito calculus, and we obtain

$$[B^{ij}(\partial_i\partial_j H_{\mathbf{X}} - \beta(\partial_i H_{\mathbf{X}})(\partial_j H_{\mathbf{X}}))$$
$$+ (\partial_i H_{\mathbf{X}})\partial_j L^{ij} + L^{ij}(\partial_i H_{\mathbf{X}})(\partial_j \log \sqrt{g})]dt.$$

We further multiply the above equation by $p(\mathbf{x}, t)\sqrt{g(\mathbf{x})}$, then integrate over $\mathbf{x}$, and find

$$\langle đ\mathscr{Q} \rangle = dt \int_x p[B^{ij}(\partial_i\partial_j H_{\mathbf{X}} - \beta(\partial_i H_{\mathbf{X}})(\partial_j H_{\mathbf{X}}))$$
$$+ (\partial_i H_{\mathbf{X}})\partial_j L^{ij} + L^{ij}(\partial_i H_{\mathbf{X}})(\partial_j \log \sqrt{g})]. \quad (4.7)$$

Further carrying out a few integrations by parts, we find that $đ\mathscr{Q}$ is indeed Eq. (3.6b), heat at ensemble level

$$\langle đ\mathscr{Q} \rangle = dt \int_x H_{\mathbf{X}} \frac{1}{\sqrt{g}}\partial_i \sqrt{g} L^{ij}(\partial_j + (\partial_j\mathscr{U}))p$$

$$= dt \int_x H_{\mathbf{X}} \mathcal{L}_{\text{FP}} p. \quad (4.8)$$

B. Forward and backward trajectories, protocols, and processes

Following the common terminology in stochastic thermodynamics, we shall use *protocol* to denote a generic time-dependent external control parameter $\lambda(t)$, and *trajectory* to denote a generic dynamic path $\mathbf{x}(t)$. We use the terms *backward protocol* for $\tilde{\lambda}(t) = \lambda^*(-t)$, and *backward*

trajectory for $\tilde{x}(t) = x^*(-t)$. To make comparison, we shall also call $\lambda(t)$ and $x(t)$ respectively the *forward protocol* and the *forward trajectory*. We call the Langevin dynamics with the forward protocol $\lambda(t)$ the *forward process*, and that with the backward protocol $\tilde{\lambda}(t)$ the *backward process*. Note that if the forward process is defined in the time interval (t_I, t_F), then the backward process is defined in $(-t_F, -t_I)$. In many previous papers, both the forward process and backward process are defined in the time intervals $(0, T)$. The backward trajectory and protocol are then defined as $x^*(T - t)$ and $\lambda^*(T - t)$. Evidently, the choices of time interval is only a matter of convenience. Strictly speaking, to define a Langevin process, we need to specify the pdf of slow variables at the initial time. If the control parameter λ is fixed, and the initial pdf is the equilibrium pdf, then the process is stationary, as we discussed in Sec. II C. In this section, we shall not assume the system starting from equilibrium.

Corresponding to every infinitesimal step in the forward process, which we discussed in Sec. IV A, there is a step in the backward process, where the system goes from $x^* + dx^*$ to x^*, and the parameter from $\lambda^* + d\lambda^*$ to λ^*. The resulting variation of H_X can be analogously decomposed into work $(d\mathcal{W})_{\text{bw}}$ and heat $(d\mathcal{Q})_{\text{bw}}$,

$$(dH_X)_{\text{bw}} = H_X(x^*, \lambda^*) - H_X(x^* + dx^*, \lambda^* + d\lambda^*)$$
$$= (d\mathcal{Q})_{\text{bw}} - (d\mathcal{W})_{\text{bw}} + O(dt^{3/2}), \quad (4.9)$$

$$(d\mathcal{W})_{\text{bw}} \equiv H_X(x^*, \lambda^*) - H_X(x^*, \lambda^* + d\lambda^*), \quad (4.10)$$

$$(d\mathcal{Q})_{\text{bw}} \equiv H_X(x^*, \lambda^*) - H_X(x^* + dx^*, \lambda^*). \quad (4.11)$$

Here the subscript bw denotes backward. Using the detailed balance condition Eq. (3.6b'), we see that in the backward step, $(dH_X)_{\text{bw}}, (d\mathcal{Q})_{\text{bw}}, (d\mathcal{Q})_{\text{bw}}$ are the negatives of the corresponding quantities in the forward step,

$$(dH_X)_{\text{bw}} = -(dH_X)_{\text{fw}}, \quad (4.12a)$$

$$(d\mathcal{W})_{\text{bw}} = -(d\mathcal{W})_{\text{fw}}, \quad (4.12b)$$

$$(d\mathcal{Q})_{\text{bw}} = -(d\mathcal{Q})_{\text{fw}}, \quad (4.12c)$$

where quantities in RHS with subscript fw refer to the forward process and are given in Eqs. (4.1) and (4.3).

Let us now consider a trajectory of finite duration. The total heat and work along the forward trajectory $x(t)$ in the forward process can be obtained by integrating $d\mathcal{Q}$ and $d\mathcal{W}$ along the path,

$$\mathcal{W}[x(t), \lambda(t)] = \int_{t_I}^{t_F} d\mathcal{W}, \quad (4.13a)$$

$$\mathcal{Q}[x(t), \lambda(t)] = \int_{t_I}^{t_F} d\mathcal{Q}, \quad (4.13b)$$

The integrated first law takes the form

$$\Delta H_X = H_X(x(t_F), \lambda(t_F)) - H_X(x(t_I), \lambda(t_I))$$
$$= \mathcal{W}[x(t), \lambda(t)] + \mathcal{Q}[x(t), \lambda(t)]. \quad (4.14)$$

Note that we have displayed the dependence of $\mathcal{W}, \mathcal{Q}$ on trajectory and protocol.

Let $\mathcal{Q}[\tilde{x}(t), \tilde{\lambda}(t)]$ and $\mathcal{W}[\tilde{x}(t), \tilde{\lambda}(t)]$ be the total heat and work along the backward trajectory $\tilde{x}(t)$ of the backward process, which can be obtained by integrating $(d\mathcal{Q})_{\text{bw}}, (d\mathcal{Q})_{\text{bw}}$ along the backward trajectory. Using Eqs. (4.12) and (4.13), we easily find

$$\mathcal{Q}[\tilde{x}(t), \tilde{\lambda}(t)] = -\mathcal{Q}[x(t), \lambda(t)], \quad (4.15a)$$

$$\mathcal{W}[[\tilde{x}(t), \tilde{\lambda}(t)] = -\mathcal{W}[x(t), \lambda(t)]. \quad (4.15b)$$

The meaning of Eqs. (4.12) is that work and heat are both odd under simultaneous reversal of trajectory and dynamic protocols.

C. Path probability

We shall now define probability density of dynamic trajectory. To circumvent the difficulty associated with infinite dimensional space of dynamic trajectories, we discretize the time variable, i.e., use the trick of time slicing. We divide the interval (t_I, t_F) into $N \gg 1$ infinitesimal steps with duration $dt = (t_F - t_I)/N$. In kth step, the system goes from $x_{k-1} = x(t_{k-1})$ at time t_{k-1} to $x_k = x(t_k)$ at time t_k, whereas the parameter changes from $\lambda_{k-1} = \lambda(t_{k-1})$ to $\lambda_k = \lambda(t_k)$. We can now approximate the forward trajectory $x(t)$ as a discrete sequence of states, which we denote as γ,

$$\gamma : \quad \begin{pmatrix} x_N \\ t_N \end{pmatrix} \leftarrow \cdots \leftarrow \begin{pmatrix} x_k \\ t_k \end{pmatrix} \leftarrow \cdots \leftarrow \begin{pmatrix} x_0 \\ t_0 \end{pmatrix}, \quad (4.16a)$$

where $t_N = t_F, t_0 = t_I$, and the arrow indicates propagation of time. We similarly approximate the backward trajectory $\tilde{x}(t) = x^*(-t)$ as the reversed discrete sequence, and denote it as $\tilde{\gamma}$,

$$\tilde{\gamma} : \quad \begin{pmatrix} x_N^* \\ -t_N \end{pmatrix} \rightarrow \cdots \rightarrow \begin{pmatrix} x_k^* \\ -t_k \end{pmatrix} \rightarrow \cdots \rightarrow \begin{pmatrix} x_0^* \\ -t_0 \end{pmatrix}, \quad (4.16b)$$

where time propagates from left to right. Note that the initial state of the backward trajectory x_N^* is time reversal of the final state of the forward trajectory, whereas the final state of the backward trajectory x_0^* is time-reversal of the initial state of the forward trajectory.

We further define the invariant volume measure of the forward trajectory as

$$D\gamma \equiv dv_N \cdots dv_1 dv_0$$
$$= dv(x_N) \cdots dv(x_1) dv(x_0), \quad (4.17)$$

where $dv(x)$ is defined in Eq. (2.10). We then define the differential probability of the discretized path γ in the forward process as the N-time joint pdf,

$$p_F[\gamma] D\gamma \equiv p_F(x_N, t_N; \cdots; x_1, t_1; x_0, t_0)$$
$$\times dv_N \cdots dv_1 dv_0. \quad (4.18)$$

Similar to Eqs. (2.32), the N-time joint pdf can be decomposed as

$$p_F(\boldsymbol{x}_N, t_N; \cdots; \boldsymbol{x}_0, t_0) = p_F(\boldsymbol{x}_N, t_N | \boldsymbol{x}_{N-1}, t_{N-1})$$
$$\times \cdots \cdots \cdots \cdots$$
$$\times p_F(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_0, t_0)$$
$$\times p_F(\boldsymbol{x}_0, t_0). \quad (4.19)$$

Unlike in Eqs. (2.32), however, here the initial pdf $p_{F0}(\boldsymbol{x}_0, t_0)$ remains arbitrary.

Let us further introduce the notation $\gamma_0 = \boldsymbol{x}_0$ to denote the initial state of the forward trajectory, and $p_F(\gamma_0) = p_F(\boldsymbol{x}_0, t_0)$ its pdf. The product of N transition probabilities in Eq. (4.19) can be understood as the conditional pdf of the forward trajectory, given its initial state $\gamma_0 = \boldsymbol{x}_0$,

$$p_F[\gamma|\gamma_0] = p_F(\boldsymbol{x}_N, t_N | \boldsymbol{x}_{N-1}, t_{N-1}) \cdots p_F(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_0, t_0). \quad (4.20)$$

The unconditional path pdf can then be expressed as

$$p_F[\gamma] = p_F[\gamma|\gamma_0] \, p_F(\gamma_0). \quad (4.21)$$

The same things can be done for the discretized backward trajectory Eq. (4.16b) of the backward process. The invariant volume measure of the backward trajectory is

$$D\tilde{\gamma} \equiv dv_0^* dv_1^* \cdots dv_N^*$$
$$= dv(\boldsymbol{x}_0^*) dv(\boldsymbol{x}_1^*) \cdots dv(\boldsymbol{x}_N^*) = D\gamma, \quad (4.22)$$

where we have used the symmetry Eq. (2.25b). The counterpart of Eqs. (4.18) is

$$p_B[\tilde{\gamma}]D\tilde{\gamma} \equiv p_B(\boldsymbol{x}_0^*, -t_0; \boldsymbol{x}_1^*, -t_1 \cdots; \boldsymbol{x}_N^*, -t_N)dv_0^* dv_1^* \cdots dv_N^*. \quad (4.23)$$

The counterpart of Eq. (4.19) is

$$p_B(\boldsymbol{x}_0^*, -t_0; \cdots \boldsymbol{x}_N^*, -t_N) = p_B(\boldsymbol{x}_0^*, -t_0 | \boldsymbol{x}_1^*, -t_1)$$
$$\times \cdots \cdots \cdots \cdots$$
$$\times p_B(\boldsymbol{x}_{N-1}^*, -t_{N-1} | \boldsymbol{x}_N^*, -t_N)$$
$$\times p_B(\boldsymbol{x}_N^*, -t_N). \quad (4.24)$$

Furthermore introducing the notation $\tilde{\gamma}_0 = \boldsymbol{x}_N^*$ for the initial state of the backward trajectory, and $p_B(\tilde{\gamma}_0)$ we have the counterparts for Eqs. (4.20) and (4.21),

$$p_B[\tilde{\gamma}|\tilde{\gamma}_0] = p_B(\boldsymbol{x}_0^*, -t_0 | \boldsymbol{x}_1^*, -t_1) \cdots$$
$$p_B(\boldsymbol{x}_{N-1}^*, -t_{N-1} | \boldsymbol{x}_N^*, -t_N), \quad (4.25)$$

$$p_B[\tilde{\gamma}] = p_F[\tilde{\gamma}|\tilde{\gamma}_0] \, p_B(\tilde{\gamma}_0). \quad (4.26)$$

D. Detailed fluctuation theorem

Recall that in Sec. II C we have derived Eq. (2.30) by assuming that λ is fixed. If $t_2 - t_1 = dt$ is small, $d\boldsymbol{x} = \boldsymbol{x}_2 - \boldsymbol{x}_1 \sim \sqrt{dt}$ is also small for typically trajectories. Hence the exponent in the RHS of Eq. (2.30) can be rewritten as

$$-d_{\boldsymbol{x}}\mathscr{U}(\boldsymbol{x}, \lambda) = -\beta \, d_{\boldsymbol{x}} H_{\mathbf{X}}(\boldsymbol{x}, \lambda) = -\beta d\mathscr{Q}, \quad (4.27)$$

where we have used Eqs. (3.4) and (4.1b). Hence Eq. (2.30) can be rewritten into

$$\frac{p_F(\boldsymbol{x} + d\boldsymbol{x}, t + dt | \boldsymbol{x}, t)}{p_B(\boldsymbol{x}^*, -t | \boldsymbol{x}^* + d\boldsymbol{x}^*, -t - dt)} = e^{-\beta d\mathscr{Q}}. \quad (4.28)$$

Now consider the case that $\lambda(t)$ is varied continuously, both in the forward process and in the backward process. There comes the issue now when should $\lambda(t)$ be evaluated in Eq. (4.27). The answer is very simple. In the time step dt, $d\lambda \sim dt$, but $d\mathscr{Q}$ is already of order $dt^{1/2}$. Hence change of time where $\lambda(t)$ is evaluated only leads to correction of Eq. (4.27) at the order of $dt^{3/2}$, which is negligible in the continuum limit. It does not matter when we evaluate $\lambda(t)$.

In Appendix, we explicitly calculate the short time transition probabilities $p_F(\boldsymbol{x} + d\boldsymbol{x}, t + dt | \boldsymbol{x}, t)$ using the covariant Langevin equation (2.16), and verify Eq. (4.28). The formula of short-time transition probability obtained in Appendix is useful for construction of path-integral representation of Langevin dynamics in curved space or driven by multiplicative noises.

Let us apply Eq. (4.28) to a pair of evolution steps in the forward and backward processes, as shown in Eqs. (4.16). The heat associated with the forward step is $\mathscr{U}(\boldsymbol{x}_{k+1}, \lambda(\tau_k)) - \mathscr{U}(\boldsymbol{x}_k, \lambda(\tau_k)) = d\mathscr{Q}_k$, where τ_k is an arbitrary point between t_k, t_{k+1}, whose precise value does not matter for the reason explained above. We then obtain

$$\frac{p_F(\boldsymbol{x}_{k+1}, t_{k+1} | \boldsymbol{x}_k, t_k)}{p_B(\boldsymbol{x}_k^*, -t_k | \boldsymbol{x}_{k+1}^*, -t_{k+1})} = e^{-\beta d\mathscr{Q}_k}. \quad (4.29)$$

We can now calculate the ratio of Eqs. (4.20) and (4.25). Using Eq. (4.29) for every step, we find

$$\frac{p_F[\gamma|\gamma_0]}{p_B[\tilde{\gamma}|\tilde{\gamma}_0]} = e^{-\beta \mathscr{Q}[\gamma, \lambda]} = e^{\beta \mathscr{Q}[\tilde{\gamma}, \tilde{\lambda}]}, \quad (4.30)$$

where $Q[\gamma, \lambda]$, $\mathscr{Q}[\tilde{\gamma}, \tilde{\lambda}]$ are respectively the total heat absorbed by the system along the forward/backward trajectory in the forward/backward process, which depends both on the trajectory $\gamma, \tilde{\gamma}$ and on the protocol $\lambda, \tilde{\lambda}$. Equations (4.30), or its short time version Eq. (4.28), play a fundamental role in the theory of stochastic thermodynamics, has often been called the *detailed fluctuation theorem*. Note that this result is a consequence of detailed balance, and is valid for arbitrary trajectory γ. In some previous papers, when studying general nonlinear Langevin dynamics, Eq. (4.30) or (4.28) was often treated as definition of heat. Such a definition however must be carefully dealt with, as its connection to change of bath entropy is not clear. In our covariant, however, the equivalence of Eqs. (4.30) and (4.1b) is explicitly established, and in Ref. [13], it was demonstrated that $-\beta d_{\boldsymbol{x}} H_{\mathbf{X}}$ is indeed the entropy change of environment.

Combining Eqs. (4.30), (4.21), and (4.26), we obtain

$$\log \frac{p_F[\gamma]}{p_B[\tilde{\gamma}]} = -\beta \mathscr{Q}[\gamma, \lambda] + \log p_F(\gamma_0) - \log p_B(\tilde{\gamma}_0),$$
$$(4.31)$$

which is valid for arbitrary initial pdfs of the forward and backward processes. Since $-\mathscr{Q}[\gamma, \lambda]$ is the energy transfer from the system to the heat bath, $-\beta \mathscr{Q}[\gamma, \lambda]$ is the entropy change of the heat bath, conditioned on the slow

variables (assuming, of course, that the bath is always in equilibrium). Seifert calls $-\log \wp(\boldsymbol{x})$ *the stochastic entropies* of the system at state $\boldsymbol{x}$, whose ensemble average gives the usual Gibbs-Shannon entropy. If one further choose $\wp_B(\tilde{\gamma}_0) = \wp_B(\boldsymbol{x}^*, -t_N) = \wp_F(\boldsymbol{x}, t_N)$, which means that the initial pdf of the backward process is chosen to be the time reversal of the final pdf of the forward process, then $\log \wp_F(\gamma_0) - \log \wp_B(\tilde{\gamma}_0)$ is the change of the system stochastic entropy along the trajectory γ, and the RHS of Eq. (4.31) may be interpreted as the change of total entropy of the universe along the dynamic trajectory γ in the forward process.

E. Crooks and Jarzynski

The derivations of Crooks fluctuation theorem and Jarzynski equality (integrated fluctuation theorem) from the "detailed fluctuation theorem" Eq. (4.31), are standard and have been discussed in many papers. Nonetheless, to make the paper self-contained, we briefly present these derivations using our own notations. Suppose in the forward protocol, the control parameter starts from $\lambda(t_I) = \lambda_I$ and ends at $\lambda(t_F) = \lambda_F$. In the backward protocol, then the control parameter starts from $\tilde{\lambda}(-t_F) = \lambda_F^*$ and ends at $\tilde{\lambda}(-t_I) = \lambda_I^*$.

Furthermore, we assume that both in the forward process and in the backward process, the system starts from thermal equilibrium with respect to the control parameter at that moment. Let a generic trajectory γ have initial state $\gamma_0 = \boldsymbol{x}_I$ and final state $\boldsymbol{x}_F$, then $\tilde{\gamma}_0 = \boldsymbol{x}_F^*$, and we have

$$\wp_F(\gamma_0) = e^{\beta F(\lambda_I) - \beta H_{\mathbf{X}}(\boldsymbol{x}_I, \lambda_I)}, \tag{4.32a}$$

$$\begin{aligned}\wp_B(\tilde{\gamma}_0) &= e^{\beta F(\lambda_F^*) - \beta H_{\mathbf{X}}(\boldsymbol{x}_F^*, \lambda_F^*)} \\ &= e^{\beta F(\lambda_F) - \beta H_{\mathbf{X}}(\boldsymbol{x}_F, \lambda_F)},\end{aligned} \tag{4.32b}$$

where in the last equality we have used detailed balance, Eq. (3.6b′). Substituting Eqs. (4.32a) and (4.32b) back into Eq. (4.31), and using Eq. (4.14), we obtain

$$\log \frac{\wp_F[\gamma]}{\wp_B[\tilde{\gamma}]} = \beta \mathscr{W}[\gamma, \lambda] - \beta \Delta F, \tag{4.33}$$

where $\Delta F \equiv F(\lambda_F) - F(\lambda_I)$ is the difference of Gibbs free energy between the final equilibrium state and the initial equilibrium state.

One may be attempted to think of Eq. (4.33) as the entropy increase of the universe along the trajectory γ, from t_I to t_F. But this is not correct in general. The entropy change of the universe along γ can be written as

$$-\beta \mathscr{Q}[\gamma, \lambda] - \log \wp_F(\boldsymbol{x}_F, t_F) + \log \wp_F(\boldsymbol{x}_I, t_I), \tag{4.34}$$

where the first term is the entropy change of the environment, conditioned on the system variables, and the rest two terms are the change of stochastic entropy of the system. For a genetic dynamic protocol, the system is off equilibrium at t_F, and $\wp_F(\boldsymbol{x}_F, t_F)$ is not Gibbs-Boltzmann. Hence Eq. (4.34) is not the same as $\beta \mathscr{W}[\gamma, \lambda] - \beta \Delta F$. Nonetheless, if the protocol is such that at t_F the system already equilibrates, then Eq. (4.34) equals to $\beta \mathscr{W}[\gamma, \lambda] - \beta \Delta F$, and can be understood as the total entropy change of the universe as the system evolves along the trajectory γ.

Note that the initial conditions of the forward and backward processes, Eqs. (4.32), are chosen such that they are symmetric with respect to reversal. In another word, the backward of the backward process is the forward process. Defining two pdfs as follows:

$$p_F(\sigma) \equiv \int D\gamma \, \wp_F(\gamma) \, \delta \left(\sigma - \log \frac{\wp_F[\gamma]}{\wp_B[\tilde{\gamma}]} \right), \quad (4.35a)$$

$$p_B(\sigma) \equiv \int D\tilde{\gamma} \, \wp_B(\tilde{\gamma}) \, \delta \left(\sigma - \log \frac{\wp_B[\tilde{\gamma}]}{\wp_F[\gamma]} \right). \quad (4.35b)$$

If the protocols are such that, both in the forward process and in the backward process, the system already equilibrates in the final state, then $p_F(\sigma)$, $p_B(\sigma)$ can be understood as the pdf of entropy production in the forward and backward processes, respectively.

Using a theorem proved by van der Broeck and Cleuven [29], one can prove

$$\frac{p_F(\sigma)}{p_B(-\sigma)} = e^{-\sigma}. \tag{4.36}$$

But according to Eq. (4.33), $\log \wp_F[\gamma]/\wp_B[\tilde{\gamma}]$ is linearly related to the work $\mathscr{W}[\gamma, \lambda]$ along γ, hence we see that the probability distribution of work $\mathscr{W}[\gamma]$ obeys the Crooks fluctuation theorem,

$$\frac{p_F(w)}{p_B(-w)} = e^{-\beta \Delta F + \beta w}. \tag{4.37}$$

Multiplying both sides by $p_B(-w)e^{-\beta w}$ and integrating over w, we obtain Jarzynski work relation,

$$\langle e^{-\beta w} \rangle_F = \int dw \, e^{-\beta w} p_F(w) = e^{-\beta \Delta F}. \tag{4.38}$$

V. CONNECTIONS WITH SOME PREVIOUS THEORIES

In this section, we show that both classical irreversible thermodynamics and Hamiltonian dynamics can be understood as special limits of our covariant theory. We also show that for weakly damped classical mechanical systems, our theory is consistent with the common theories of stochastic energetics and stochastic thermodynamics.

A. Deterministic limit and classical irreversible thermodynamics

As discussed in Ref. [23], there are two possible scenarios where a deterministic limit of our covariant Langevin dynamics can be obtained: (i) the thermodynamic limit, where $\mathscr{U}$ becomes extensive, whereas both the spurious drift and the noise terms are subextensive; (ii) the low-temperature limit, where $\mathscr{U} = \beta H_{\mathbf{X}} - \beta F$ becomes large comparing with the spurious drift and noises. In either case, the spurious drift and noises can be dropped and the covariant Langevin equation (2.16) reduces to a deterministic equation

$$\dot{x}^i = -L^{ij} \partial_j \beta H_{\mathbf{X}}(\boldsymbol{x}, \lambda), \tag{5.1}$$

where $H_{\mathbf{X}}$ is the Hamiltonian of mean force as defined in Eq. (3.4). Up to an additive constant, $-\beta H_{\mathbf{X}}(\boldsymbol{x}, \lambda)$ can be understood as the entropy of the universe, including the system and all its environment. Hence $-\beta \partial_j H_{\mathbf{X}} = \partial S^{\text{tot}}/\partial x^j$ is

the *affinity* corresponding to the thermodynamic variable x^i, whereas $\dot{x}^i$ is the associated *flux*. Equation (5.1) then becomes the basic equation of irreversible thermodynamics [30], which linearly relates affinities to fluxes, and dictates how thermodynamic variables $\boldsymbol{x}$ relax towards equilibrium. Correspondingly, L^{ij} are precisely the kinetic coefficients in the classical irreversible thermodynamics, which are usually assumed to be constants. The detailed balance conditions (2.31) on the kinetic coefficients then become the famous *Onsager-Casimir reciprocal relations*.

Note however, in classical irreversible thermodynamics, $\dot{x}^i$ transforms as usual contravariant vectors, whereas in the covariant Langevin theory, dx^i transforms according to Ito formula, Eq. (2.21). The difference between these two transformations is due to the last term in the RHS of Eq. (2.21), which is proportional to B^{ij}. It becomes negligible in the deterministic limit, for the same reason that the spurious drift becomes negligible.

For the particular case of a Hamiltonian system weakly damped by the environment, we have $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p}) = (q^1, \ldots, q^n, p^1, \ldots, p^n)$, and

$$H_{\mathbf{X}}(\boldsymbol{q}, \boldsymbol{p}) = K(\boldsymbol{p}) + V(\boldsymbol{q}) - F, \tag{5.2}$$

$$\boldsymbol{Q} = \begin{pmatrix} 0 & -\mathbf{I}_n \\ \mathbf{I}_n & 0 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{\gamma} \end{pmatrix}, \tag{5.3}$$

where $K(\boldsymbol{p}), V(\boldsymbol{q})$ are respectively the kinetic and potential energies, whilst $\boldsymbol{\gamma}$ is a $n \times n$ symmetric positive constant matrix. Equation (5.1) then reduces to

$$\dot{q}^i = \frac{\partial K}{\partial p^i}, \tag{5.4}$$

$$\dot{p}^i = -\frac{\partial V}{\partial q^i} - \gamma^{ij} \frac{\partial K}{\partial p^i}. \tag{5.5}$$

These equations are identical to Eqs. (121.2) and (121.5) of Landau and Lifshitz [31].

B. Unitary dynamics and Hamiltonian dynamics

As shown in Ref. [23], the unitary limit of our covariant Langevin dynamics is reached by the limit $B^{ij} = 0$, where Eq. (2.16) becomes

$$dx^i + \left(Q^{ij} \partial_j \mathscr{U} - \frac{1}{\sqrt{g}} \partial_j \sqrt{g} Q^{ij} \right) dt = 0. \tag{5.6}$$

In this limit, there is no energy dissipation, and the dynamics becomes deterministic and unitary. The precise meaning of unitarity is discussed in detail in Ref. [23].

Recall in Riemann manifold covariant derivatives of tensor fields are defined as

$$\nabla_k A^{ij} = \partial_k A^{ij} + \Gamma^i_{mk} A^{mj} + \Gamma^j_{mk} A^{im}, \tag{5.7}$$

where Γ^i_{mk} are the Christoffel symbols

$$\Gamma^i_{kl} = \frac{1}{2} g^{im} (\partial_l g_{mk} + \partial_k g_{ml} - \partial_m g_{kl}), \tag{5.8}$$

which have the following properties:

$$\Gamma^j_{kj} = \partial_k \log \sqrt{g}. \tag{5.9}$$

Using these results we can easily show that

$$\nabla_j Q^{ij} = \partial_j Q^{ij} + \frac{1}{\sqrt{g}} Q^{ij} \partial_j \sqrt{g}$$

$$= \frac{1}{\sqrt{g}} \partial_j \sqrt{g} Q^{ij}. \tag{5.10}$$

Hence Eq. (5.6) can be written as

$$\dot{x}^i = v^i(\boldsymbol{x}) = -Q^{ij} \nabla_j \mathscr{U} + \nabla_j Q^{ij}. \tag{5.11}$$

For reason, which will become clear below, we assume

$$(\nabla_i Q^{ij})(\nabla_j \mathscr{U}) = 0. \tag{5.12}$$

As a special example, we consider Hamiltonian dynamics with canonical variables. Q^{ij} is just the constant symplectic matrix, and the Christoffel symbols are zero, hence $\nabla_i Q^{ij} = 0$, and Eq. (5.12) is trivially satisfied.

Recall that $\nabla_j Q^{ij}$ is a contravariant vector field, and for any contra-variant vector field we have

$$\nabla_i v^i = \frac{1}{\sqrt{g}} \partial_i \sqrt{g} v^i. \tag{5.13}$$

Combining this with Eq. (5.10), we find

$$\nabla_i \nabla_j Q^{ij} = 0. \tag{5.14}$$

Now take the covariant divergence of Eq. (5.11) and use the Leibnitz rule of covariant derivatives as well as $\nabla_i \nabla_j \varphi = \nabla_j \nabla_i \varphi$, we find

$$\nabla_i \dot{x}^i = \nabla_i v^i(\boldsymbol{x}) = -(\nabla_i Q^{ij})(\nabla_j \mathscr{U}) = 0, \tag{5.15}$$

where the last step follows from our assumption Eq. (5.12). Hence the flow defined by Eq. (5.11) is incompressible, i.e., the space volume of slow variable is conserved by the unitary dynamics. The conservation of phase space volume in Hamiltonian dynamics is known as the Liouville theorem.

We can also show that the generalized potential $\mathscr{U}$ is conserved by the dynamics,

$$\frac{d\mathscr{U}}{dt} = (\partial_i \mathscr{U}) \dot{x}^i$$

$$= -(\nabla_i \mathscr{U}) Q^{ij} (\nabla_j \mathscr{U}) + (\nabla_i \mathscr{U})(\nabla_j Q^{ij}) = 0. \tag{5.16}$$

Similarly we can prove that the Gibbs-Shannon entropy $S[\wp]$ is also conserved. Taking time derivative of the entropy, we find

$$\frac{dS[\wp]}{dt} = -\int_{\boldsymbol{x}} (\log \wp \, \mathscr{L}_{\text{FP}} \wp). \tag{5.17}$$

Using Eq. (2.19), together with the antisymmetry of Q^{ij}, as well as a few integrations by parts, we can show that

$$\frac{dS[\wp]}{dt} = -\int_{\boldsymbol{x}} \wp \, (\nabla_i Q^{ij})(\nabla_j \mathscr{U}) = 0. \tag{5.18}$$

Hence we find that all following properties hold about the unitary dynamics Eq. (5.6), if $(\nabla_i \mathscr{U})(\nabla_j Q^{ij}) = 0$:

(1) The flow defined by the unitary dynamics Eq. (5.11) is incompressible.

(2) The generalized potential $\mathscr{U}$ is conserved by the unitary dynamics Eq. (5.11).

(3) The Gibbs-Shannon entropy is conserved by the unitary dynamics Eq. (5.11).

Of course, all these properties are destroyed once the noises and dissipations are turned on.

C. Stochastic energetics and stochastic thermodynamics

The common theories of stochastic energetics and stochastic thermodynamics have two cornerstones. Firstly, Langevin equations are formulated using the "conventional Langevin approach" [32], where one adds frictions and noises (random forces) to the otherwise unitary dynamic equations. The resulting equations are understood as balance of forces, or, more generally, balance of thermodynamic forces, and the variances of random forces are chosen such that the steady states correspond to thermal equilibrium states in the absence of time-dependence driving forces. Secondly, heat is defined, *à la* Sekimoto [4], as the work done to the system by the friction and random forces, and other thermodynamic variables are defined such that the first and second laws of thermodynamics are valid for arbitrary nonequilibrium processes.

When dealing with systems whose dynamics are described by Eq. (2.16), however, both these cornerstones are problematic. Firstly, the Langevin equation (2.16) cannot be naively understood as condition of force balance. In fact, if one writes down force balance equation in the presence of multiplicative noises, one would obtain an equation similar to Eq. (2.16) but with the spurious drift missing, and with detailed balance unintentionally broken. This is in fact a well-known pathology of the "conventional Langevin approach" [32–36]. Secondly, whenever spurious drift shows up, Sekimoto's definition of heat differs from ours, and leads to violation of the second law of thermodynamics. Hence whenever the metric tensor or kinetic coefficients depend on system variables, the common theories need to be replaced by the theory developed in this paper. In Sec. VI we supply three concrete examples to illustrate these points.

It is important to emphasize that the conceptually all Langevin equations are effective dynamic equations of slow variables obtained via coarse-graining of fast variables. This applies both to our covariant theory and to the "conventional Langevin approach". The difference is how the Langevin equations are formulated. In contrast with the ad hoc way of the "conventional Langevin approach", in our covariant theory, Langevin equations are constructed using tensor objects, which automatically guarantees detailed balance.

For classical point particles coupled to additive noises, our theory is in fact fully consistent with the common theories. We first use one-dimensional system to illustrate the main point. Following the conventional Langevin approach, we add noise and friction to the Hamiltonian equations, and obtain

$$\frac{dq}{dt} = \frac{p}{m}, \tag{5.19a}$$

$$\frac{dp}{dt} = -\partial_q V(q, \lambda) - \frac{\gamma p}{m} + \eta(t), \tag{5.19b}$$

where λ is the control parameter, and the noise correlation function is

$$\langle \eta(t)\eta(t') \rangle = 2\gamma T \delta(t - t'), \tag{5.20}$$

such that the system converges to an equilibrium Gibbs-Boltzmann distribution with energy $E = p^2/2m + V(q)$ and

with temperature T. As shown in Sec. 4 B of Ref. [23], these Langevin equations can also be written in the form of covariant Langevin equations.

In the common theories, heat is defined as the work done by the friction and random force, which can be written as the following Stratonovich product:

$$d\mathcal{Q}_{\text{SE}} \equiv \left(-\gamma \frac{dq}{dt} + \xi(t)\right) \circ dq(t), \tag{5.21}$$

where the subscript SE denotes Sekimoto. The Stratonovich product $\circ$ has the special property that for any function $f(q, p)$, the usual calculus rule holds

$$df(q, p) = \partial_q f \circ dq + \partial_p f \circ dp. \tag{5.22}$$

For a detailed discussion on Stratonovich product of stochastic variables, see chapter 4 of Gardiner's book [26]. Using the Langevin equation (5.19b), Sekimoto (see Ch. 4 of Ref. [4]) further showed

$$d\mathcal{Q}_{\text{SE}} = (\dot{p} + \partial_q V) \circ dq$$
$$= d\left(\frac{p^2}{2m} + V(q, \lambda)\right) - \frac{\partial H}{\partial \lambda} d\lambda. \tag{5.23}$$

But this is precisely the heat defined in Eq. (4.1b), with $H = p^2/2m + V$ playing the role of $H_{\mathbf{X}}$,

$$d\mathcal{Q} = dH_{\mathbf{X}} - d_\lambda H_{\mathbf{X}} = d_x H_{\mathbf{X}}, \tag{5.24}$$

where $x = (q, p)$ is the shorthand for the canonical variables. The work according to Sekimoto is then given by

$$d\mathcal{W}_{\text{SE}} = dH_{\mathbf{X}} - d\mathcal{Q}_{\text{SE}} = d_\lambda H_{\mathbf{X}}, \tag{5.25}$$

which is also identical to our definitions of work, Eqs. (4.1a).

We now consider a collection of point particles moving in a fluid and interacting with each other via certain potential $u(\mathbf{r}_i - \mathbf{r}_j)$. The slow variables consist of the Cartesian coordinates $\mathbf{r}_i$ and canonical momenta $\boldsymbol{p}_i$ of all particles., i.e., $x = (\mathbf{r}_1, \boldsymbol{p}_1, \ldots, \mathbf{r}_n, \boldsymbol{p}_n)$. There is also an external potential field $\psi_i(\mathbf{r}, \lambda)$ acting on each particle, with λ the external control parameter. The total Hamiltonian is then given by

$$H = \sum_i \left(\frac{\boldsymbol{p}_i^2}{2m_i} + \psi_i(\mathbf{r}_i, \lambda)\right) + \sum_{i<j} u(\mathbf{r}_i - \mathbf{r}_j). \tag{5.26}$$

The system is weakly coupled to a heat bath T, so that the equilibrium state is

$$p^{\text{EQ}}(\mathbf{r}, \boldsymbol{p}) = e^{\beta F - \beta H}. \tag{5.27}$$

Hence the generalized potential is

$$\mathcal{U} = \beta(H - F). \tag{5.28}$$

The nonlinear Langevin equations are given by

$$d\mathbf{r}_i = \frac{\boldsymbol{p}_i}{m_i} dt, \tag{5.29a}$$

$$d\boldsymbol{p}_i = -\left[\sum_j \nabla_i u(\mathbf{r}_i - \mathbf{r}_j) + \nabla_i \psi_i(\mathbf{r}_i)\right] dt$$
$$- \gamma_i \dot{\mathbf{r}}_i dt + \sqrt{2\gamma_i T} d\mathbf{W}_i, \tag{5.29b}$$

where γ_i is the friction coefficient of ith particle, and $d\mathbf{W}_i$ are vector-valued Wiener noises. Note that $\nabla_i = \partial/\partial \boldsymbol{x}_i$ is the

gradient operator in Cartesian coordinates. Equations (5.29) can be written in the covariant form Eq. (2.16), with $g = 1$, and the following nonvanishing kinetic coefficients:

$$-Q^{r_i,p_i} = Q^{p_i,r_i} = T, \tag{5.30a}$$

$$B^{p_i,p_i} = \gamma_i T, \tag{5.30b}$$

which satisfy condition (2.31b). There is no spurious drift since all kinetic coefficients are constants. The remaining conditions of detailed balance are

$$\psi_i(\mathbf{r}_i, \lambda) = \psi_i(\mathbf{r}_i^*, \lambda^*), \tag{5.31a}$$

$$u(\mathbf{r}) = u(\mathbf{r}^*). \tag{5.31b}$$

In the limit $\gamma_i \to 0$, Eqs. (5.29) reduce to the well-known Hamiltonian equations.

Thermodynamic quantities at ensemble level, such as system entropy, free energy etc, can be straightforwardly obtained using the results in Sec. III. Here we focus on the work and heat at trajectory level, which can be written down using Eqs. (4.2) and (4.5),

$$đ\mathcal{W} = d_\lambda H = \sum_i (\partial_\lambda \psi_i) d\lambda, \tag{5.32}$$

$$\begin{aligned} đ\mathcal{Q} &= d_x H \\ &= \sum_i \left[-\frac{\gamma_i}{m_i} \left(\frac{\boldsymbol{p}_i^2}{m_i} - T \right) dt + \sqrt{2\gamma_i T} \ \dot{\mathbf{r}}_i \cdot d\mathbf{W}_i \right], \end{aligned} \tag{5.33}$$

where we have used Ito's formula Eq. (2.8) in the calculation of heat, and the product $\dot{\mathbf{r}}_i \cdot d\mathbf{W}_i$ is understood in Ito's sense. The average heat is obtained by averaging $d\mathcal{Q}$ both over noises and over $\wp$. The average of product $\dot{\mathbf{r}}_i \cdot d\mathbf{W}_i$ vanishes because $\dot{\mathbf{r}}_i$ is independent of $d\mathbf{W}_i$ in Ito calculus. Hence we find

$$đQ = \langle đ\mathcal{Q} \rangle = -\sum_i \frac{2\gamma_i \, dt}{m_i} \int_{\mathbf{r},\boldsymbol{p}} \wp \left(\frac{\boldsymbol{p}_i^2}{2m_i} - \frac{3T}{2} \right). \tag{5.34}$$

Equations (5.32) and (5.34) agree with those of Sec. 4.1 of Ref. [4]. Note, however, Stratonovich's calculus instead of Ito calculus is used in Ref. [4].

If the process is quasistatic, which means that the system remains in equilibrium all the time, then the average of kinetic $\boldsymbol{p}_i^2/2m_i$ equals to $3T/2$ (the equipartition theorem), and the heat vanishes identically, and there is no dissipation of energy. On the other hand, Eq. (5.34) also means that as long as there is energy dissipation due to friction, the average kinetic energy cannot satisfy the equipartition theorem. This implies that the overdamped limit of a Hamiltonian cannot be obtained by simply assuming that the momentum remains in equilibrium, for that would imply no friction.

Using the covariant properties established in this paper, we are able to carry out arbitrary nonlinear transformation of variables, which can be either canonical or noncanonical. This can be very useful when dealing with multidimensional Hamiltonian systems with various kind of nonlinear constraints. More examples will be explored in future papers.

VI. THREE EXAMPLES

In this section, we discuss three examples where the common theories fail due to the presence of spurious drift, whereas our theory yields correct results. In the first example, Langevin equations are formulated in curvilinear coordinates, whereas in the second example, the slow manifold is curved. In the third example, the kinetic coefficient is state dependent. More results will be supplied in the future.

A. Diffusion in polar coordinates

We consider the overdamped limit of two dimensional Brownian motion, subjected to a potential $V(r; \lambda)$ with rotational symmetry and/or with radial boundary condition $r < r_0$. It is then more convenient to use polar coordinates. We first start from the overdamped Langevin equations in Cartesian coordinates,

$$\gamma \, dx + \partial_x V(r; \lambda) \, dt = \sqrt{2T\gamma} \, dW_x, \tag{6.1a}$$

$$\gamma \, dy + \partial_y V(r; \lambda) \, dt = \sqrt{2T\gamma} \, dW_y, \tag{6.1b}$$

which can be readily understood as conditions of force balance. The noise amplitude $\sqrt{2T\gamma}$ is such that the steady-state distribution is the thermal equilibrium state $e^{-\beta V(r)} dx dy$. Equations (6.1) can be written into the matrix form

$$\begin{pmatrix} dx \\ dy \end{pmatrix} + \begin{pmatrix} T/\gamma & 0 \\ 0 & T/\gamma \end{pmatrix} \begin{pmatrix} \partial_x \beta V \\ \partial_y \beta V \end{pmatrix} = \begin{pmatrix} \sqrt{2T/\gamma} \, dW_x \\ \sqrt{2T/\gamma} \, dW_y \end{pmatrix}, \tag{6.2}$$

which is our standard form, Eq. (2.16), with $\mathscr{U} = \beta H_{\mathbf{X}} = \beta V$, $g = 1$, and

$$\boldsymbol{L} = \frac{T}{\gamma} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{b} = \sqrt{\frac{2T}{\gamma}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{6.3}$$

Note that $\boldsymbol{L}$ is a constant matrix, hence there is no spurious drift.

There are two possible ways to obtain Langevin equations in polar coordinates: (i) We may use the transformation rules Eqs. (2.22) to obtain various tensors in polar coordinates and then assemble the Langevin equation in polar coordinates using Eq. (2.16). (ii) Alternatively we may also apply NTV and Ito's formula (2.8) directly to Eqs. (6.1), and obtain the new equations. The results are of course the same. Here we take the first route.

The transformation matrix from Cartesian coordinates to polar coordinates is

$$\frac{\partial(r, \phi)}{\partial(x, y)} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi/r & \cos\phi/r \end{pmatrix}. \tag{6.4}$$

Using Eqs. (2.22), we obtain the transformed kinetic matrix and noise amplitude

$$\boldsymbol{L}' = \frac{T}{\gamma} \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix}, \quad \boldsymbol{b}' = \sqrt{\frac{2T}{\gamma}} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi/r & \cos\phi/r \end{pmatrix}. \tag{6.5}$$

The determinant of the new metric tensor is $g' = r^2$. The generalized potential $\mathcal{U}$ remains the same since it is a scalar. Using these results and further defining new Wiener noises

$$dW_r = \cos\phi \, dW_x + \sin\phi \, dW_y, \quad (6.6a)$$

$$dW_r = -\sin\phi \, dW_x + \cos\phi \, dW_y, \quad (6.6b)$$

we obtain the transformed Langevin equations

$$dr + \frac{1}{\gamma}\left(\partial_r V - \frac{T}{r}\right)dt = \sqrt{\frac{2T}{\gamma}}dW_r, \quad (6.7a)$$

$$d\phi = \sqrt{\frac{2T}{\gamma}}\frac{1}{r}dW_\phi. \quad (6.7b)$$

Note that the term $-T\,dt/\gamma\, r$ in Eq. (6.7a) is the spurious drift, resulting from the r dependence of the metric. The associated Fokker-Planck equation is

$$\partial_t \wp = \frac{T}{\gamma r}\partial_r r(\partial_r \wp + (\partial_r \beta V)\wp) + \frac{T}{\gamma r^2}\partial_\phi^2 \wp. \quad (6.8)$$

The heat and work at trajectory level are

$$d\mathcal{Q} = T d_r \mathcal{U} = d_r V = (\partial_r V)dr + \frac{T(\partial_r^2 V)}{\gamma}dt, \quad (6.9a)$$

$$d\mathcal{W} = T d_\lambda \mathcal{U} = d_\lambda V = (\partial_\lambda V)d\lambda. \quad (6.9b)$$

The total entropy production rate is given by Eq. (3.11). Using the above results, we obtain

$$\frac{dS^{\text{tot}}}{dt} = \int_{r,\phi}\frac{T}{\gamma\,\wp}\left[(\partial_r \wp + (\partial_r \beta V)\wp)^2 + \frac{(\partial_\phi \wp)^2}{r^2}\right], \quad (6.10)$$

which is non-negative, and vanishes identically at equilibrium where $\wp \sim e^{-\beta V}$.

If one try to write down Langevin equations in polar coordinates using the idea of force balance, one would obtain

$$-V' - \gamma\dot{r} + \sqrt{2T\gamma}\frac{dW_r}{dt} = 0, \quad (6.11a)$$

$$-\gamma r\dot{\phi} + \sqrt{2T\gamma}\frac{dW_\phi}{dt} = 0. \quad (6.11b)$$

which, comparing with the correct equations (6.7), misses the spurious drift. It is important to note that noises are additive in this problem, hence there is no issue of stochastic calculus here.

Additionally, if we adopt Sekimoto's definition of heat as the work done by friction and random force, and use the correct Langevin equations (6.7), we would have

$$d\mathcal{Q}_{\text{SE}} = \left(-\gamma\dot{r} + \sqrt{2T\gamma}\frac{dW_r}{dt}\right)\circ dr$$

$$\quad + \left(-\gamma r\dot{\phi} + \sqrt{2T\gamma}\frac{dW_\phi}{dt}\right)\circ rd\phi$$

$$\quad = (V'(r) - T/r)\circ dr$$

$$\quad = d\mathcal{Q} - T\,d\log r, \quad (6.12)$$

which differs from Eq. (6.9a) by $-T\,d\log r$. Using this to calculate the total entropy production, we would obtain

$$\frac{dS_{\text{SE}}^{\text{tot}}}{dt} = \frac{dS}{dt} - \beta\left\langle\frac{d\mathcal{Q}_{\text{SE}}}{dt}\right\rangle$$

$$= \frac{dS^{\text{tot}}}{dt} + \langle d_r \log r\rangle$$

$$= \frac{dS^{\text{tot}}}{dt} - \frac{1}{\gamma}\int_{r,\phi}\frac{\partial_r V}{r}\wp, \quad (6.13)$$

where $\langle\,\cdot\,\rangle$ means average over both probability distribution and noises, which we have calculated using Eqs. (6.11a) and Ito formula $drdr = 2T\,dt/\gamma$ [obtained from Eq. (2.8)]. At thermal equilibrium, $\wp^{\text{EQ}}(r) \propto e^{-\beta V(r)}$, and the extra term in Eq. (6.13) can be calculated,

$$-\frac{1}{\gamma}\int\frac{\partial_r V}{r}\wp = \frac{2\pi T}{\gamma}(\wp^{\text{EQ}}(r_0) - \wp^{\text{EQ}}(0)), \quad (6.14)$$

which is nonvanishing as long as $V(r_0) \neq V(0)$. Since $dS^{\text{tot}}/dt = 0$ at equilibrium, Eq. (6.13) predicts a nonvanishing entropy production rate at equilibrium for $V(r_0) \neq V(0)$, which violates the second law of thermodynamics.

It is important to note that for this particular simple problem, it is possible to interpret the spurious drift as an additional friction force due to geometry of curvilinear coordinates, such that Sekimoto's definition of heat becomes correct. Such a procedure however does not work for general multidimensional problems, because the spurious drift $(1/\sqrt{g}\,\partial_j(\sqrt{g}L^{ij})$ involves both reactive and dissipative parts of L^{ij} and cannot not be interpreted as friction force.

B. Diffusion on a sphere

As a similar problem, we consider diffusion of a molecule on a unit sphere, or the rotational diffusion of an uniaxial particle, subjected to an external potential $V(\theta; \lambda)$ that is invariant under rotation around z axis. We shall again consider the *overdamped limit*, hence the slow variables are the polar and azimuthal angles θ, ϕ. Unlike the preceding example, however, we have no Cartesian coordinates to start with, but have to construct the equations in polar coordinates directly. The metric has the following form:

$$ds^2 = g_{\theta\theta}d\theta^2 + g_{\phi\phi}d\phi^2 = d\theta^2 + \sin^2\theta d\phi^2. \quad (6.15)$$

The covariant and contravariant metric tensors **g** then have the following components:

$$g_{\theta\theta} = 1, \quad g_{\phi\phi} = \sin^2\theta, \quad g_{\theta\phi} = g_{\phi\theta} = 0, \quad (6.16)$$

$$g^{\theta\theta} = 1, \quad g^{\phi\phi} = \sin^{-2}\theta, \quad g^{\theta\phi} = g^{\phi\theta} = 0. \quad (6.17)$$

The corresponding matrix forms are **g** and $\mathbf{g}^{-1}$, given by

$$\mathbf{g} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}, \quad \mathbf{g}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^{-2}\theta \end{pmatrix}. \quad (6.18)$$

The invariant volume measure is then

$$dv = \sqrt{g}d\theta d\phi = \sin\theta d\theta d\phi. \quad (6.19)$$

We shall assume that there is an external potential, which only depends on the polar angle θ, so that the equilibrium pdf is $e^{-\beta V(\theta,\lambda)}\sin\theta d\theta d\phi$. Since we are studying the

overdamped theory, detailed balance demands that the anti-symmetric matrix Q^{ij} vanishes identically. We shall assume that the symmetric matrix B^{ij} has rotational symmetry. It is known that g_{ij} and g^{ij} are the only covariant and contravariant tensors on the sphere with rotational symmetry. Hence we see that $\boldsymbol{B}$ must be proportional to $\mathbf{g}^{-1}$,

$$\boldsymbol{B} = \begin{pmatrix} D & 0 \\ 0 & D\sin^{-2}\theta \end{pmatrix}, \quad (6.20)$$

where D is the diffusion constant.

The Langevin equations can be obtained from Eq. (2.16) via proper choices of the matrix $b^{i\alpha}$,

$$d\theta + D(\beta\partial_\theta V - \cot\theta)dt = \sqrt{2D}\,dW_\theta, \quad (6.21a)$$

$$d\phi = \frac{\sqrt{2D}}{\sin\theta}\,dW_\phi, \quad (6.21b)$$

where the term $-D\cot\theta$ in Eq. (6.21a) is spurious drift.

The associated Fokker-Planck equation is

$$\partial_t \mathit{p} = \frac{D}{\sin\theta}\partial_\theta\sin\theta(\partial_\theta \mathit{p} + \beta(\partial_\theta V)\mathit{p}) + \frac{D}{\sin^2\theta}\partial_\phi^2 \mathit{p}. \quad (6.22)$$

Equations (6.21) and (6.22) agree with the results obtained by Raible and Engel [37]. They also reduce to Eqs. (6.7) and (6.8) as $\sin\theta \sim r \to 0$ with $D = T/\gamma$. Using Eqs. (4.2), (4.5), and (3.4), the differential heat and work at trajectory level are

$$d\mathscr{Q} = (\partial_\theta V)d\theta + D\left(\partial_\theta^2 V\right)dt, \quad (6.23)$$

$$dW = (\partial_\lambda V)d\lambda. \quad (6.24)$$

The total entropy production rate can again be calculated using Eq. (3.11),

$$\frac{dS^{\text{tot}}}{dt} = \int_{\theta,\phi}\frac{D}{\mathit{p}}\left[(\partial_\theta \mathit{p} + (\partial_\theta\beta V)\mathit{p})^2 + \frac{(\partial_\phi \mathit{p})^2}{\sin^2\theta}\right], \quad (6.25)$$

which is non-negative, and vanishes identically at equilibrium where $\mathit{p} \sim e^{-\beta V(\theta;\lambda)}$.

Similar to the first example, if one follow the conventional Langevin approach, incorrect Langevin equations is obtained. If one use Sekimoto's definition of heat, incorrect result for total entropy production is obtained, and the second law is violated. The derivations are very similar, so we skip them here.

C. Diffusion with position-dependent friction

Consider diffusion of a particle in heterogeneous environment with position-dependent friction coefficient $\gamma(x)$ and in an external potential $V(x)$. (Note this is not the same problem as studied in Refs. [17–19], where temperature varies over space.) The underdamped Langevin equations are

$$dx = \frac{p}{m}dt, \quad (6.26a)$$

$$dp = -\frac{\gamma(x)p}{m}dt - V'(x)dt + \sqrt{2T\gamma(x)}dW, \quad (6.26b)$$

where dW is the Wiener noise and $V'(x)$ is spatial derivative of the potential. These equations can be rewritten as the covariant form Eq. (2.16) with

$$\mathscr{U}(x, p) = \beta(H - F) = \beta\left(\frac{p^2}{2m} + V(x) - F\right), \quad (6.27)$$

$$Q = \begin{pmatrix} 0 & -T \\ T & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & \gamma(x)T \end{pmatrix}, \quad (6.28)$$

where the first columns (rows) of the matrices refer to x and the second columns (rows) refer to p. There is no spurious drift because $\gamma(x)$ depends only on x but not on p. Detailed balance is clearly satisfied.

We are interested in the over-damped limit of this problem, where $\gamma(x)$ is very large (or equivalently, the mass is very small), so that only the dynamics of position becomes relevant. In Ref. [38], it was shown through a very long calculation that the overdamped equation is

$$dx + \frac{V'(x)}{\gamma(x)}dt - \left(\frac{T}{\gamma(x)}\right)'dt = \sqrt{\frac{2T}{\gamma(x)}}dW, \quad (6.29)$$

where the third term in LHS is the spurious drift. The over-damped equation can also be obtained using multiscale perturbation method. The calculation is however still rather complicated. We note that if the friction is independent of position, Eq. (6.29) reduces to the well-known result

$$dx + \frac{V'(x)}{\gamma}dt = \sqrt{\frac{2T}{\gamma}}dW. \quad (6.30)$$

Using our covariant formalism of Langevin dynamics, however, the overdamped equation (6.29) can be obtained almost instantaneously. The covariant Langevin equation, Eq. (2.16), tells us that the overdamped equation must have the following form:

$$dx + B(x)\mathscr{U}_{od}'(x)dt - B'(x)dt = \sqrt{2B(x)}dW. \quad (6.31)$$

Recall that the equilibrium pdf of the overdamped theory is $e^{-\mathscr{U}_{od}(x)}$, and obtained form that of the underdamped theory by integrating out p,

$$e^{-\mathscr{U}_{od}(x)} = \int e^{-\mathscr{U}(x,p)}dp, \quad (6.32)$$

where $\mathscr{U}(x, p)$ is given in Eq. (6.27). Hence we find

$$\mathscr{U}_{od}(x) = \beta V(x) + C. \quad (6.33)$$

Now Eq. (6.31) must reduce to Eq. (6.30) if the friction is independent of x. Hence $B(x) = T/\gamma(x)$, and Eq. (6.31) is indeed identical to Eq. (6.29). This example illustrates the power of our covariant Langevin dynamics.

Using Eq. (4.5) the heat at trajectory level is

$$d\mathscr{Q} = d_x V(x) = V' \circ dx = V'dx + \frac{T}{2\gamma}V''dt, \quad (6.34)$$

where $\circ$ means product in Stratonovich's sense. The total entropy production rate is calculated using Eq. (3.11),

$$
\begin{aligned}
\frac{dS^{\text{tot}}}{dt} &= dS - \beta d\mathcal{Q} \\
&= \int_x \frac{T}{\gamma p} (\partial_x p + (\partial_x \beta V) p)^2,
\end{aligned}
\tag{6.35}
$$

which is non-negative and vanishes at the equilibrium where $p \sim e^{-\beta V}$.

If we follow the conventional Langevin approach to write down an overdamped equation based on force balance, we would obtain

$$
\gamma(x)dx + V'(x)dt = \sqrt{2T\gamma(x)}\, dW,
\tag{6.36}
$$

which misses the spurious drift completely. The heat defined by Sekimoto is

$$
\begin{aligned}
d\mathcal{Q}_{\text{SE}} &= \left(-\gamma\, dx/dt + \sqrt{2T\gamma(x)}dW/dt\right) \circ dx \\
&= (V' + T\gamma'/\gamma) \circ dx \\
&= d_x V(x) + T\, d_x \log \gamma(x) \\
&= d\mathcal{Q} + T\, d_x \log \gamma(x),
\end{aligned}
\tag{6.37}
$$

where in the second equality we have used the correct Langevin equation (6.29). This definition of heat differs from the correct definition (6.34) by a term

$$
T\, d_x \log \gamma(x) = T\frac{\gamma'}{\gamma}dx + \frac{T^2}{2\gamma}\left(\frac{\gamma'}{\gamma}\right)' dt,
\tag{6.38}
$$

where we have used Ito's formula (2.8) to expand $d_x \log \gamma(x)$ to the second order in dx.

For simplicity, let us assume that there is no the external potential, i.e., $V = 0$, and the system is confined in an interval with length L. The equilibrium state then corresponds to a flat distribution $p^{\text{EQ}}(x) = 1/L$. To calculate the ensemble average of Eq. (6.38), we use Eq. (6.29) to replace dx, average over noise, and multiply by $p^{\text{EQ}}(x)$, and finally integrate over x, and obtain

$$
\langle T\, d_x \log \gamma(x) \rangle_{\text{EQ}} = \frac{dt\, T^2}{L}\left(\frac{\gamma'}{\gamma^2}\right)\bigg|_0^L,
\tag{6.39}
$$

which is generically nonvanishing. Hence, using of Eq. (6.37) as definition of heat leads to a nonvanishing total entropy production rate in the equilibrium state,

$$
\frac{dS^{\text{tot}}_{\text{SE}}}{dt} = -\frac{dt\, T}{L}\left(\frac{\gamma'}{\gamma^2}\right)\bigg|_0^L,
\tag{6.40}
$$

which violates the second law of thermodynamics.

VII. CONCLUSIONS

In this paper, we have developed a covariant theory of thermodynamics and stochastic thermodynamics for nonequilibrium small systems. As we have demonstrated using concrete examples, this theory should be used in replacement of the common theories of stochastic energetics and stochastic thermodynamics if the noises are multiplicative, the slow variable space is curved, or curvilinear coordinates are used. In future, we will use the theory to study more sophisticated models of Langevin dynamics with spurious drift. One interesting possibility is Brownian motion of rigid body and the associated theory of stochastic thermodynamics. We will also study small systems that are driven by nonconservative forces, or with temperature gradients. We will also study how the theory transform under coarse-graining, thereby establish the connections between different levels of stochastic thermodynamics, and understand how entropy productions at different levels of theory are related to each other. Finally we shall also try to generalize the theory to quantum systems.

APPENDIX: ALTERNATIVE PROOF OF EQ. (4.28)

In this Appendix, we explicitly calculate the short time transition probability $dv(\boldsymbol{x}_1)p_F(\boldsymbol{x}_1, t + dt|\boldsymbol{x}; t)$ using the Langevin equation (2.16), and prove the formula Eq. (4.28).

1. Short-time transition probability

Let us for now consider the following general form of Ito-Langevin equation:

$$
dx^i - F^i(\boldsymbol{x}, \lambda)dt = b^{i\alpha}(\boldsymbol{x}, \lambda)dW_\alpha,
\tag{A1}
$$

where $F^i(\boldsymbol{x}, \lambda)$ is usually called the *systematic force*. Note that both $F^i(\boldsymbol{x}, \lambda)$ and $b^{i\alpha}(\boldsymbol{x}, \lambda)$ may depend on the control parameter λ, which may vary with time. Consider a transition from $\boldsymbol{x}$ at time t to $\boldsymbol{x}_1$ at time $t + dt$, and let $\alpha \in (0, 1)$, so that $\boldsymbol{x}_\alpha = \boldsymbol{x} + \alpha(\boldsymbol{x}_1 - \boldsymbol{x}) = \boldsymbol{x} + \alpha d\boldsymbol{x}$ is an intermediate point between $\boldsymbol{x}$ and $\boldsymbol{x}_1$. In Ref. [39] [Eq. (68), with $\bar{\alpha} = 0$, or equivalently, Eqs. (37)], we proved that the differential transition probability can be written as

$$
dv(\boldsymbol{x}_1)p_F\boldsymbol{x}_1, t + dt|\boldsymbol{x}; t) = \frac{d\mu(\boldsymbol{x}_1)\, e^{-A^\alpha(\boldsymbol{x}_1, \boldsymbol{x}; dt, \lambda)}}{\sqrt{(4\pi dt)^n \det B^{ij}(\boldsymbol{x}_\alpha, \lambda)}},
\tag{A2a}
$$

$$
\begin{aligned}
A^\alpha(\boldsymbol{x}_1, \boldsymbol{x}; dt, \lambda) &= [dx^i - dt(F^i - 2\alpha\partial_k B^{ik})_\alpha]\frac{B^{-1}_{ij}(\boldsymbol{x}_\alpha, \lambda)}{4dt}[dx^j - dt(F^j - 2\alpha\partial_l B^{jl})_\alpha] \\
&\quad + \alpha\,(\partial_i F^i)_\alpha dt - \alpha^2(\partial_i\partial_j B^{ij})_\alpha dt,
\end{aligned}
\tag{A2b}
$$

where $d\mu(\boldsymbol{x}_1) = d^n\boldsymbol{x}_1$ is an infinitesimal volume element around $\boldsymbol{x}_1$, and $dv(\boldsymbol{x}_1) = \sqrt{g(\boldsymbol{x}_1)}d^n\boldsymbol{x}_1$ is the invariant volume measure, B^{-1}_{ij} is the inverse matrix of $B^{ij} = b^{i\alpha}b^{j\alpha}$, see Eq. (2.3d), and $(\cdots)_\alpha$ means that all functions inside the bracket are evaluated at the intermediate point $\boldsymbol{x}_\alpha$. The matrix B^{ij} is assumed to satisfy the detailed balance condition Eq. (2.31d).

In particular, the $\alpha = 0$ version of short-time transition probability is

$$dv(\boldsymbol{x}_1) p_F \boldsymbol{x}_1, t + dt | \boldsymbol{x}; t) = \frac{d\mu(\boldsymbol{x}_1) e^{-A^0(\boldsymbol{x}_1, \boldsymbol{x}; dt, \lambda)}}{\sqrt{(4\pi dt)^n \det B^{ij}(\boldsymbol{x}, \lambda)}}, \tag{A3}$$

$$A^0(\boldsymbol{x}_1, \boldsymbol{x}; dt, \lambda) = (dx^i - F^i(\boldsymbol{x}, \lambda)dt)\frac{B_{ij}^{-1}(\boldsymbol{x}, \lambda)}{4dt}(dx^j - F^j(\boldsymbol{x}, \lambda)dt). \tag{A4}$$

Note that the linear combination $dx^i - dt F^i(\boldsymbol{x}, t)$ in the action Eq. (A4) is precisely the deterministic part of the Langevin equation (A1). In our covariant formulation of Langevin theory, this linear combination transforms as a contravariant vector under under nonlinear transformation of variables, even though neither dx^i or F^i does so.

The defining features of Eq. (A3) are that $d\boldsymbol{x}$ is Gaussian random variable, and has the following moments (here κ_m denotes cumulant of order m):

$$\langle dx^i \rangle = F^i(\boldsymbol{x}, \lambda)dt, \tag{A5a}$$

$$\langle dx^i dx^i \rangle - \langle dx^i \rangle\langle dx^j \rangle = 2B^{ij}(\boldsymbol{x}, \lambda)dt, \tag{A5b}$$

$$\kappa_m(d\boldsymbol{x}) = 0, \quad \forall\, m \geqslant 3. \tag{A5c}$$

For $\alpha \neq 0$, the action Eq. (A2b) is not quadratic in $d\boldsymbol{x}$. Hence the transition probability Eq. (A2a) is not Gaussian in $d\boldsymbol{x}$. However, in Ref. [39], we proved that all moments of $d\boldsymbol{x}$ obtained from Eq. (A2a) are independent of α up to order dt. This means that in the short-time limit, $dt \to 0$, Eq. (A2a) with different α are equivalent to each other, in the sense that they all generate the same continuous-time Markov process, with the same statistical properties of physical observables. In other words, the non-Gaussian nature of Eq. (A2a) for $\alpha \neq 0$ makes no contribution to the statistical properties of the Markov process in the continuous time limit.

2. Reversible and irreversible forces

The systematic force $F^i(\boldsymbol{x}, \lambda)$ in Eq. (A1) can be decomposed into a reversible part $F^i_{(R)}(\boldsymbol{x}, \lambda)$ and an irreversible part $F^i_{(IR)}(\boldsymbol{x}, \lambda)$, which are respectively defined as

$$F^i_{(R)}(\boldsymbol{x}, \lambda) \equiv \tfrac{1}{2}(F^i(\boldsymbol{x}, \lambda) - \epsilon_i F^i(\boldsymbol{x}^*, \lambda^*)), \tag{A6}$$

$$F^i_{(IR)}(\boldsymbol{x}, \lambda) \equiv \tfrac{1}{2}(F^i(\boldsymbol{x}, \lambda) + \epsilon_i F^i(\boldsymbol{x}^*, \lambda^*)), \tag{A7}$$

$$F^i(\boldsymbol{x}, \lambda) = F^i_{(IR)}(\boldsymbol{x}, \lambda) + F^i_{(R)}(\boldsymbol{x}, \lambda). \tag{A8}$$

Here and below, we shall hide the time dependence of F^i and B^{ij}, to simplify the notations. It then follows from these definitions that (no summation over repeated indices below)

$$\epsilon_i F^i_{(R)}(\boldsymbol{x}^*, \lambda^*) = -F^i_{(R)}(\boldsymbol{x}, \lambda), \tag{A9a}$$

$$\epsilon_i F^i_{(IR)}(\boldsymbol{x}^*, \lambda^*) = F^i_{(IR)}(\boldsymbol{x}, \lambda), \tag{A9b}$$

$$\epsilon_i F^i(\boldsymbol{x}^*, \lambda^*) = F^i_{(IR)}(\boldsymbol{x}, \lambda) - F^i_{(R)}(\boldsymbol{x}, \lambda). \tag{A9c}$$

For the covariant Langevin equation (2.16), we have

$$F^i(\boldsymbol{x}, \lambda) = \partial_j L^{ij}(\boldsymbol{x}, \lambda) - L^{ij}(\boldsymbol{x}, \lambda)\,\partial_j[\mathscr{U}(\boldsymbol{x}, \lambda) - \log\sqrt{g(\boldsymbol{x})}]. \tag{A10a}$$

Using the conditions of detailed balance, Eqs. (2.31), we can show that the reversible and irreversible parts of the systematic force are

$$F^i_{(R)}(\boldsymbol{x}, \lambda) = \partial_j Q^{ij}(\boldsymbol{x}, \lambda) - Q^{ij}(\boldsymbol{x}, \lambda)\,\partial_j[\mathscr{U}(\boldsymbol{x}, \lambda) - \log\sqrt{g(\boldsymbol{x})}], \tag{A10b}$$

$$F^i_{(IR)}(\boldsymbol{x}, \lambda) = \partial_j B^{ij}(\boldsymbol{x}, \lambda) - B^{ij}(\boldsymbol{x}, \lambda)\,\partial_j[\mathscr{U}(\boldsymbol{x}, \lambda) - \log\sqrt{g(\boldsymbol{x})}], \tag{A10c}$$

$$F^i(\boldsymbol{x}, \lambda) = F^i_{(IR)}(\boldsymbol{x}, \lambda) + F^i_{(R)}(\boldsymbol{x}, \lambda). \tag{A10d}$$

Hence the reversible force only involves Q^{ij} whereas the irreversible force only involves B^{ij}. This is, of course, consistent with our understanding that Q^{ij} are the reactive couplings whereas B^{ij} are the dissipative couplings.

3. Proof of Eq. (4.28)

To calculate the ratio (4.28), it is most convenient to use $\alpha = 1/2$ version of the transition probability (A2a),

$$dv(\boldsymbol{x}_1) p_F \boldsymbol{x}_1, t + dt | \boldsymbol{x}; t) = \frac{d\mu(\boldsymbol{x}_1) e^{-A^{1/2}(\boldsymbol{x}_1, \boldsymbol{x}; dt, \lambda)}}{\sqrt{(4\pi dt)^n \det B^{ij}(\boldsymbol{x}_{1/2})}}, \tag{A11}$$

$$A^{1/2}(\boldsymbol{x}_1, \boldsymbol{x}; dt, \lambda) = [dx^i - dt(F^i - \partial_k B^{ik})_{1/2}]\frac{B_{ij}^{-1}(\boldsymbol{x}_{1/2})}{4dt}[dx^j - dt(F^j - \partial_l B^{jl})_{1/2}]$$

$$+ \frac{1}{2}(\partial_i F^i)_{1/2}dt - \frac{1}{4}(\partial_i \partial_j B^{ij})_{1/2}dt, \tag{A12}$$

where F^i, B^{ij} depend both on $\boldsymbol{x}_{1/2}$ and on λ. Using Eq. (A8) and introduce a shorthand $\boldsymbol{y} = \boldsymbol{x}_{1/2} = (\boldsymbol{x}_1 + \boldsymbol{x})/2$, we can rewrite Eq. (A12) as

$$A^{1/2}(\boldsymbol{x}_1, \boldsymbol{x}, dt, \lambda) = \left\{dx^i - dt\left[F^i_{(IR)}(\boldsymbol{y}, \lambda) + F^i_{(R)}(\boldsymbol{y}, \lambda) - \partial_k B^{ik}(\boldsymbol{y}, \lambda)\right]\right\}$$

$$\times \frac{B_{ij}^{-1}(\boldsymbol{y}, \lambda)}{4\,dt} \times \left\{dx^j - dt\left[F^j_{(IR)}(\boldsymbol{y}, \lambda) + F^j_{(R)}(\boldsymbol{y}, \lambda) - \partial_l B^{jl}(\boldsymbol{y}, \lambda)\right]\right\}$$

$$+ \frac{1}{2}dt\,\partial_i\left(F^i_{(IR)}(\boldsymbol{y}, \lambda) + F^i_{(R)}(\boldsymbol{y}, \lambda)\right) - \frac{1}{4}dt\,\partial_i\partial_j B^{ij}(\boldsymbol{y}, \lambda). \tag{A13}$$

Now consider the backward process where the system goes from $\boldsymbol{x}_1^*$ to $\boldsymbol{x}^*$, and the parameter is fixed at λ^*. Note that the mid-point of the backward process is $\boldsymbol{y}^* = \boldsymbol{x}^* + d\boldsymbol{x}^*/2$. Let us write down the $\alpha = 1/2$ version of the transition probability for the backward process, which is the counterpart of Eq. (A11):

$$dv(\boldsymbol{x}^*)\mathcal{p}_B(\boldsymbol{x}^*, -t|\boldsymbol{x}_1^*, -t-dt) = \frac{e^{-A^{1/2}(\boldsymbol{x}^*, \boldsymbol{x}_1^*, dt, \lambda^*)}d\mu(\boldsymbol{x}^*)}{\sqrt{(4\pi dt)^n \det B^{ij}(\boldsymbol{y}^*, \lambda^*)}}, \tag{A14}$$

$$A^{1/2}(\boldsymbol{x}^*, \boldsymbol{x}_1^*, dt, \lambda^*) = (-(dx^i)^* - dt(F^i(\boldsymbol{y}^*, \lambda^*) - \partial_k^* B^{ik}(\boldsymbol{y}^*, \lambda^*)))$$

$$\times \frac{B_{ij}^{-1}(\boldsymbol{y}^*, \lambda^*)}{4\,dt} \times (-(dx^j)^* - dt(F^j(\boldsymbol{y}^*, \lambda^*) - \partial_l^* B^{jl}(\boldsymbol{y}^*, \lambda^*)))$$

$$+ \frac{1}{2}dt\,\partial_i^* F^i(\boldsymbol{y}^*, \lambda^*) - \frac{1}{4}dt\,\partial_i^*\partial_j^* B^{ij}(\boldsymbol{y}^*, \lambda^*). \tag{A15}$$

Using $(x^i)^* = \epsilon_i x^i$, we can rewrite the action Eq. (A15) into

$$A^{1/2}(\boldsymbol{x}^*, \boldsymbol{x}_1^*, dt, \lambda^*) = (-dx^i - dt(\epsilon_i F^i(\boldsymbol{y}^*, \lambda^*) - \partial_k \epsilon_i B^{ik}(\boldsymbol{y}^*, \lambda^*)\epsilon_k))$$

$$\times \frac{\epsilon_i B_{ij}^{-1}(\boldsymbol{y}^*, \lambda^*)\epsilon_j}{4\,dt} \times (-dx^j - dt(\epsilon_j F^j(\boldsymbol{y}^*, \lambda^*) - \partial_l \epsilon_j B^{jl}(\boldsymbol{y}^*, \lambda^*)\epsilon_l))$$

$$+ \frac{1}{2}dt\,\partial_i \epsilon_i F^i(\boldsymbol{y}^*, \lambda^*) - \frac{1}{4}dt\,\partial_i\partial_j \epsilon_i B^{ij}(\boldsymbol{y}^*, \lambda^*)\epsilon_j. \tag{A16}$$

Using Eq. (A8) to decompose F^i into $F_{(IR)}^i + F_{(R)}^i$, and further using the symmetry properties, Eqs. (A9), as well as the conditions of detailed balance (2.31d), we can rewrite Eq. (A16) into

$$A^{1/2}(\boldsymbol{x}^*, \boldsymbol{x}_1^*, dt, \lambda^*) = \left\{ -dx^i - dt\left[F_{(IR)}^i(\boldsymbol{y}, \lambda) - F_{(R)}^i(\boldsymbol{y}, \lambda) - \partial_k B^{ik}(\boldsymbol{y}, \lambda)\right]\right\}$$

$$\times \frac{B_{ij}^{-1}(\boldsymbol{y}, \lambda)}{4\,dt} \times \left\{ -dx^j - dt\left[F_{(IR)}^j(\boldsymbol{y}, \lambda) - F_{(R)}^j(\boldsymbol{y}, \lambda) - \partial_l B^{jl}(\boldsymbol{y}, \lambda)\right]\right\}$$

$$+ \frac{1}{2}dt\,\partial_i\left(F_{(IR)}^i(\boldsymbol{y}, \lambda) - F_{(R)}^i(\boldsymbol{y}, \lambda)\right) - \frac{1}{4}dt\,\partial_i\partial_j B^{ij}(\boldsymbol{y}, \lambda). \tag{A17}$$

Now dividing Eq. (A11) by Eq. (A14), taking the logarithm, and using the following symmetry properties, which follow immediately from Eqs. (2.25b) and (2.31d):

$$dv(\boldsymbol{x}_1) = \sqrt{g(\boldsymbol{x}_1)}d\mu(\boldsymbol{x}_1), \tag{A18}$$

$$dv(\boldsymbol{x}^*) = \sqrt{g(\boldsymbol{x}^*)}d^n\boldsymbol{x}^* = \sqrt{g(\boldsymbol{x})}d^n\boldsymbol{x}, \tag{A19}$$

$$d\mu(\boldsymbol{x}) = d\mu(\boldsymbol{x}^*), \tag{A20}$$

$$\det B^{ij}(\boldsymbol{y}^*, \lambda^*) = \det B^{ij}(\boldsymbol{y}, \lambda), \tag{A21}$$

we obtain

$$\log \frac{\mathcal{p}_F(\boldsymbol{x}_1, t+dt|\boldsymbol{x}, t)}{\mathcal{p}_B(\boldsymbol{x}^*, -t|\boldsymbol{x}_1^*, -t-dt)} \cdot \frac{\sqrt{g(\boldsymbol{x}_1)}}{\sqrt{g(\boldsymbol{x})}} = A^{1/2}(\boldsymbol{x}^*, \boldsymbol{x}_1^*, dt, \lambda^*) - A^{1/2}(\boldsymbol{x}_1, \boldsymbol{x}, dt, \lambda). \tag{A22}$$

The RHS can be calculated readily using Eqs. (A13) and (A17):

$$\text{RHS of Eq. (A.22)} = \left(dx^i - dt\,F_{(R)}^i(\boldsymbol{y}, \lambda)\right)B_{ij}^{-1}(\boldsymbol{y}, \lambda)\left(F_{(IR)}^j(\boldsymbol{y}, \lambda) - \partial_l B^{jl}(\boldsymbol{y}, \lambda)\right) - dt\,\partial_i F_{(R)}^i(\boldsymbol{y}, \lambda). \tag{A23}$$

Further using Eqs. (A10b) and (A22), we obtain

$$A^{1/2}(\boldsymbol{x}^*, \boldsymbol{x}_1^*, dt, \lambda^*) - A^{1/2}(\boldsymbol{x}_1, \boldsymbol{x}, dt, \lambda) = -dx^i\,\partial_i[\mathscr{U}(\boldsymbol{y}, \lambda) - \log\sqrt{g(\boldsymbol{y})}]. \tag{A24}$$

Note that $\boldsymbol{y} = \boldsymbol{x}_{1/2} = \boldsymbol{x} + d\boldsymbol{x}/2$. For any smooth function $\Psi(\boldsymbol{x})$, we have

$$dx^i\,\partial_i\Psi(\boldsymbol{x}_{1/2}) = \Psi(\boldsymbol{x} + d\boldsymbol{x}) - \Psi(\boldsymbol{x}) + O(d\boldsymbol{x}^3). \tag{A25}$$

We do not need to worry about $O(d\boldsymbol{x}^3)$ terms, since they scale as $dt^{3/2}$ and do not contribute to the dynamics in the continuum limit. Hence the RHS of Eq. (A24) can be written as

$$-\mathscr{U}(\boldsymbol{x} + d\boldsymbol{x}, \lambda) + \mathscr{U}(\boldsymbol{x}, \lambda) + \log\sqrt{g(\boldsymbol{x} + d\boldsymbol{x})} - \log\sqrt{g(\boldsymbol{x})} = -d_x\mathscr{U}(\boldsymbol{x}, \lambda) + d_x\log\sqrt{g(\boldsymbol{x})}, \tag{A26}$$

where d_x is defined in Eqs. (A22). Substituting these results back into Eq. (A22) we finally obtain

$$\log \frac{p_F(\boldsymbol{x}_1, t + dt | \boldsymbol{x}, t)}{p_B(\boldsymbol{x}^*, -t | \boldsymbol{x}_1^*, -t - dt)} = -d_x \mathscr{U}(\boldsymbol{x}, \lambda) = -\beta d\mathscr{Q}, \tag{A27}$$

which is precisely Eq. (4.28). As explained in Sec. IV D, the time dependence of λ needs not to be worried, since it only contributes only to the order $dt^{3/2}$.

[1] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

[2] D. J. Evans and D. J. Searles, The fluctuation theorem, Adv. Phys. **51**, 1529 (2002).

[3] C. Jarzynski, Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale, Annu. Rev. Condens. Matter Phys. **2**, 329 (2011)

[4] K. Sekimoto, *Stochastic Energetics* (Springer, New York, 2010), Vol. 799.

[5] U. Seifert, Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem, Phys. Rev. Lett. **95**, 040602 (2005).

[6] C. Maes and K. Netočný, Time-reversal and entropy, J. Stat. Phys. **110**, 269 (2003)

[7] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[8] D. J. Searles and D. J. Evans, Fluctuation theorem for stochastic systems, Phys. Rev. E **60**, 159 (1999).

[9] C. Jarzynski, Nonequilibrium work theorem for a system strongly coupled to a thermal environment, J. Stat. Mech. (2004) P09005.

[10] U. Seifert, First and Second Law of Thermodynamics at Strong Coupling, Phys. Rev. Lett. **116**, 020601 (2016).

[11] C. Jarzynski, Stochastic and Macroscopic Thermodynamics of Strongly Coupled Systems, Phys. Rev. X **7**, 011008 (2017).

[12] P. Talkner and P. Hänggi, *Colloquium:* Statistical mechanics and thermodynamics at strong coupling: Quantum and classical, Rev. Mod. Phys. **92**, 041002 (2020).

[13] M. Ding, Z. Tu, and X. Xing, Strong coupling thermodynamics and stochastic thermodynamics from the unifying perspective of time-scale separation, Phys. Rev. Res. **4**, 013015 (2022).

[14] R. E. Spinney and I. J. Ford, Entropy production in full phase space for continuous stochastic dynamics, Phys. Rev. E **85**, 051113 (2012).

[15] R. E. Spinney and I. J. Ford, Nonequilibrium Thermodynamics of Stochastic Systems with Odd and Even Variables, Phys. Rev. Lett. **108**, 170603 (2012).

[16] T. Speck, J. Mehl, and U. Seifert, Role of External Flow and Frame Invariance in Stochastic Thermodynamics, Phys. Rev. Lett. **100**, 178302 (2008).

[17] S. Bo and A. Celani, Multiple-scale stochastic processes: Decimation, averaging and beyond, Phys. Rep. **670**, 1 (2017).

[18] S. Bo, S. H. Lim, and R. Eichhorn, Functionals in stochastic thermodynamics: How to interpret stochastic integrals, J. Stat. Mech. (2019) 084005.

[19] A. Celani, S. Bo, R. Eichhorn, and E. Aurell, Anomalous Thermodynamics at the Microscale, Phys. Rev. Lett. **109**, 260603 (2012).

[20] M. Campisi, P. Hänggi, and P. Talkner, Colloquium: Quantum fluctuation relations: Foundations and applications, Rev. Mod. Phys. **83**, 771 (2011).

[21] M. Esposito, U. Harbola, and S. Mukamel, Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems, Rev. Mod. Phys. **81**, 1665 (2009).

[22] R. Graham, Covariant formulation of non-equilibrium statistical thermodynamics, Z. Phys. B **26**, 397 (1977).

[23] M. Ding, Z. Tu, and X. Xing, Covariant formulation of nonlinear Langevin theory with multiplicative Gaussian white noises, Phys. Rev. Res. **2**, 033381 (2020).

[24] C. Aron, D. G. Barci, L. F. Cugliandolo, Z. González Arenas, and G. S. Lozano, Magnetization dynamics: Path-integral formalism for the stochastic Landau-Lifshitz-Gilbert equation, J. Stat. Mech. (2014) P09008.

[25] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2009).

[26] C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer, Berlin, 2004).

[27] M. Ding and X. Xing, Multi-scale Projection Operator Method and Coarse-graining of Covariant Fokker-Planck Theory (unpublished).

[28] here we assume natural boundary condition or reflection boundary condition, so that boundary terms all vanish.

[29] C. Van den Broeck and B. Cleuren, Comment on "Irreversibility and fluctuation theorem in stationary time series", arXiv:cond-mat/0703213.

[30] H. B. Callen, *Thermodynamics: An Introduction to the Physical Theories of Equilibrium Thermostatics and Irreversible Thermodynamics* (John Wiley, Hoboken, NJ, 1966).

[31] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Vol. 5 (Elsevier, Amsterdam, 2013).

[32] V. Kampen and N. Godfried, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).

[33] P. Hänggi, Connection between deterministic and stochastic descriptions of non-linear systems, Phys. Acta **53**, 491 (1980).

[34] P. Hanggi, Non-linear fluctuations: The problem of deterministic limit and reconstruction of stochastic dynamics, Phys. Rev. A **25**, 1130 (1982).

[35] H. Grabert, R. Graham, and M. S. Green, Fluctuations and nonlinear irreversible processes. II, Phys. Rev. A **21**, 2136 (1980).

[36] N. G. van Kampen, The validity of non-linear Langevin equations, J. Stat. Phys. **25**, 431 (1981).

[37] M. Raible and A. Engel, Langevin equation for the rotation of a magnetic particle, Appl. Organomet. Chem. **18**, 536 (2004).

[38] X. Durang, C. Kwon, and H. Park, over-damped limit and inverse-friction expansion for Brownian motion in an inhomogeneous medium, Phys. Rev. E **91**, 062118 (2015).

[39] M. Ding and X. Xing, Time-slicing path-integral in curved space, Quantum **6**, 694 (2022).