Covariant formulation of nonlinear Langevin theory with multiplicative Gaussian white noises

Mingnan Ding,[1] Zhanchun Tu,[2] and Xiangjun Xing[1,3,4,*]

[1]*Wilczek Quantum Center, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China*

[2]*Department of Physics, Beijing Normal University, Beijing 100875, China*

[3]*T. D. Lee Institute, Shanghai Jiao Tong University, Shanghai 200240, China*

[4]*Shanghai Research Center for Quantum Sciences, Shanghai 201315, China*

(Received 28 March 2020; accepted 12 August 2020; published 9 September 2020)

The multidimensional nonlinear Langevin equation with multiplicative Gaussian white noises in Ito's sense is made covariant with respect to nonlinear transform of variables. The formalism involves no metric or affine connection, works for systems with or without detailed balance, and is substantially simpler than previous theories. Its relation with deterministic theory is clarified. The unitary limit and Hermitian limit of the theory are examined. Some implications on the choices of stochastic calculus are also discussed.

DOI: 10.1103/PhysRevResearch.2.033381

I. INTRODUCTION

Nonlinear Langevin theory with multiplicative noises [1–4] is widely used to describe dynamics out of equilibrium. Yet development of this theory turns out to be very challenging and full of controversies. There have been extensive and long-lasting discussions on the choice of stochastic calculus [5–10], relation between deterministic and stochastic description [11–13], as well as the covariance of theory under nonlinear transform of variables (NTV) [15–18]. Another related issue is discretization scheme for its path integral representation [19–26]. To date, nonlinear Langevin theory with multiplicative noises has not yet been properly understood. While earlier works mostly focus on processes with detailed balance (DB), more recently there have been many efforts trying to develop nonlinear Langevin dynamics lacking DB [27–33]. Nonetheless, various conceptual issues frequently come back.

We think that covariance (with respect to NTV) and DB are two issues of great importance, not only to Langevin theory, but also to nonequilibrium statistical mechanics at large. The common theme of nonequilibrium statistical mechanics is dynamics of slow variables. But nonlinear functions of slow variables are also slow variables. (Of course, we need to assume that these functions are not fast themselves, i.e., they are not "crazy.") This implies that statistical mechanics must be covariant under NTV. When solving high-dimensional problems, NTV is extremely useful and even indispensable. Also, approximations must be covariant or they lead to inconsistency. More importantly, DB must be covariant under NTV, or an equilibrium system would be transformed into a nonequilibrium one, which is total nonsense. Finally, it is by now clear that entropy production can be quantified in terms of DB violation. Hence, DB plays a role in nonequilibrium statistical physics much like the speed of light in relativity. Even though processes without DB do exist, those with DB are special and must be invariant under NTV.

Covariance of nonlinear Langevin equation and of associated Fokker-Planck theory was first addressed by Stratonovich [14] and studied in more detail by Graham (first in Stratonovich's sense [15] and later in Ito's sense [16]). Graham's theory involves metric and affine connections and is very complicated, which hampered its application. Soon after, Grabert, Graham, and Green [18] (GGG) greatly simplified the covariant formulation of the Fokker-Planck theory. The same problem was also addressed by Hänggi [4]. Later, Ramshaw [34] showed how to derive GGG's Fokker-Planck theory from Langevin theory, but no conclusive result has been established about covariance of nonlinear Langevin theory. For a more recent work on covariant formulation of Langevin theory involving metric and affine connections, see Ref. [35].

In this work, we discuss a covariant formulation of nonlinear Langevin theory which involves no metric tensor or affine connection. Rules of transformation for parameters and physical variables are clearly demonstrated. The theory cannot be simplified further without loss of generality or physical meanings. Fully consistent deterministic limit can be obtained as either the thermodynamic limit or the low-temperature limit. While our formalism is inspired by the previous works, it is substantially simpler, more general, and more clear in physical meanings. It works for systems with both even and odd variables, with or without detailed balance. We believe that it is of considerable value to the general theory of nonlinear Langevin dynamics with multiplicative noises.

The remains of this work is organized as follows. In Sec. II, we derive the covariant form of Langevin theory and Fokker-Planck theory and explicitly demonstrate their covariance

*xxing@sjtu.edu.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

under general nonlinear transformation of variables. We also discuss the deterministic limit, unitary limit, and Hermitian limit of our theory, which are again covariant. In Sec. III, we discuss time-reversal symmetry of the theory and formulate the conditions of detailed balance in a fully covariant fashion. In Sec. V, we conclude this work with some comments on the general issue of stochastic calculus.

II. COVARIANT FORM

In the traditional *Langevin approach* [2], one starts from deterministic equation for slow variables and adds noises to obtain a Langevin equation. While this approach works perfectly for linear systems, it leads to many controversies in the nonlinear case [2,11,12,18]. Here we will start with Langevin theory as a purely phenomenological theory and re-express it in terms of observable quantities, such that it becomes fully covariant. After understanding the deterministic limit of our theory, we will see why the conventional approach is so difficult.

Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be the slow variables, whose dynamic evolution obey a stochastic differential equation. The most general nonlinear Langevin equation is either of the following two equivalent forms:

$$dx_i = F_i(\boldsymbol{x}, t)dt + b_{i\alpha}(\boldsymbol{x})dW_\alpha(t), \tag{2.1a}$$

$$\dot{x}_i = F_i(\boldsymbol{x}, t) + b_{i\alpha}(\boldsymbol{x})\xi_\alpha(t), \tag{2.1b}$$

where repeated indices are summed over. In the above $W_\alpha(t), \alpha = 1, \ldots, m$ are m-dimensional Wiener process, whereas $\xi_\alpha(t) = dW_\alpha(t)/dt$ are normalized Gaussian white noises with correlations given by

$$\langle \xi_\alpha(t)\xi_\beta(t')\rangle = \delta_{\alpha\beta}\delta(t - t'). \tag{2.2}$$

We expect $m \gg n$, since there are in general many more fast variables (which are modeled as white noises in this theory) than slow variables. In this work, we shall assume that $F_i(\boldsymbol{x}, t)$ and $b_{i\alpha}(\boldsymbol{x})$ are both time independent. The products $b_{i\alpha}(\boldsymbol{x})dW_\alpha(t)$ and $b_{i\alpha}(\boldsymbol{x})\xi_\alpha(t)$ in Eq. (2.1) are defined in Ito's sense. Finally, while Eq. (2.1a) is the mathematically rigorous formulation of stochastic differential equation and universally used by mathematicians, Eq. (2.1b) is the traditional formulation by Langevin [36] and is still preferred by many physicists. We refer to the classical textbook by Gardiner [1] for a detailed introduction of all relevant formalisms.

The Fokker-Planck equation (FPE) associated with Langevin equation (2.1) can be derived using the standard method [1,3]:

$$\begin{aligned}\partial_t p(\boldsymbol{x}, t) &= -\partial_i(F_i(\boldsymbol{x}, t)p(\boldsymbol{x}, t)) + \partial_i\partial_j(B_{ij}(\boldsymbol{x})p(\boldsymbol{x}, t)) \\ &= -\partial_i j_i(\boldsymbol{x}, t),\end{aligned} \tag{2.3}$$

where the matrix $\boldsymbol{B} = (B_{ij})$ is given by

$$B_{ij}(\boldsymbol{x}) \equiv \tfrac{1}{2}\, b_{i\alpha}(\boldsymbol{x})b_{j\alpha}(\boldsymbol{x}) = B_{ji}(\boldsymbol{x}). \tag{2.4}$$

Since $b_{i\alpha}(\boldsymbol{x})$ are real, the matrix $\boldsymbol{B}$ is non-negative. However, it may have zero eigenvalues. $j_i(\boldsymbol{x}, t)$ is the *probability current* defined as

$$j_i \equiv F_i p - \partial_j(B_{ij}p). \tag{2.5}$$

Equation (2.3) has the form of continuity equation $\partial_t p + \partial_i j_i = 0$, which describes conservation of probability.

We shall assume that the Langevin dynamics has a unique steady state

$$p^S(\boldsymbol{x}) = e^{-U(\boldsymbol{x})}, \tag{2.6}$$

where $U(\boldsymbol{x})$ is called the *generalized potential*. By substituting Eq. (2.6) in Eq. (2.5), we can re-express the steady-state probability current as

$$j_i^S = [F_i - (\partial_j B_{ij}) + B_{ij}(\partial_j U)]e^{-U}. \tag{2.7}$$

Using this result to express F_i in terms of j_i^S, we can rewrite Eq. (2.3) into

$$\partial_t p = -\partial_i\big(j_i^S e^U p\big) + \partial_i B_{ij}(\partial_j + (\partial_j U))p. \tag{2.8}$$

Note that each of the two terms in the right-hand side (RHS) vanishes separately in the steady state.

Since the steady-state current is conserved, $\partial_i j_i^S = 0$, it is always possible to write it in terms of an antisymmetric matrix function $\boldsymbol{Q}$ with components $Q_{ij} = -Q_{ji}$, such that

$$j_i^S = \partial_j(Q_{ij}e^{-U}). \tag{2.9}$$

Such a parametrization was first used by Graham [15], later by Eyink *et al.* [37], and also by Xing [30]. Note that $\boldsymbol{Q}$ is generally not unique. Using Eq. (2.9) in Eq. (2.7), we can express F_i as

$$F_i = -L_{ij}\partial_j U + \partial_j L_{ij}, \tag{2.10}$$

where $\boldsymbol{L} = (L_{ij})$ is the matrix of *kinetic coefficients* with

$$L_{ij} = B_{ij} + Q_{ij}, \quad B_{ij} = B_{ji}, \quad Q_{ij} = -Q_{ji}. \tag{2.11}$$

Substituting Eq. (2.10) back into Eqs. (2.1), we obtain the following standard forms of nonlinear Langevin equation:

$$dx_i = -L_{ij}(\boldsymbol{x})\partial_j U(\boldsymbol{x})dt + \partial_j L_{ij}(\boldsymbol{x})dt + b_{i\alpha}(\boldsymbol{x})dW_\alpha(t), \tag{2.12a}$$

$$\dot{x}_i = -L_{ij}(\boldsymbol{x})\partial_j U(\boldsymbol{x}) + \partial_j L_{ij}(\boldsymbol{x}) + b_{i\alpha}(\boldsymbol{x})\xi_\alpha(t), \tag{2.12b}$$

Substituting Eq. (2.10) back into Eqs. (2.3) and (2.5), we obtain the following standard form of FPE:

$$\partial_t p = \mathcal{L}_{\rm FP} p = -\partial_i j_i, \tag{2.13a}$$

where the Fokker-Planck operator $\mathcal{L}_{\rm FP}$ and the probability current $\boldsymbol{j}$ are given respectively by

$$\mathcal{L}_{\rm FP} \equiv \partial_i L_{ij}(\partial_j + (\partial_j U)), \tag{2.13b}$$

$$j_i = -L_{ij}(\partial_j + (\partial_j U))p + \partial_j(Q_{ij}p). \tag{2.13c}$$

Note that Eq. (2.13a) has a formal solution $p(t) = e^{t\mathcal{L}_{\rm FP}}p(0)$. Since $\mathcal{L}_{\rm FP}e^{-U} = 0$, $p^S(\boldsymbol{x}) = e^{-U(\boldsymbol{x})}$ is indeed the steady-state solution.

As one can see from Eq. (2.13b), the Fokker-Planck operator depends on $b_{i\alpha}(\boldsymbol{x})$ only through the combination $B_{ij}(\boldsymbol{x}) = b_{i\alpha}(\boldsymbol{x})b_{j\alpha}(\boldsymbol{x})/2$, as defined in Eq. (2.4). Consequently, we can apply $\boldsymbol{x}$-dependent right rotation on $b_{i\alpha}(\boldsymbol{x})$ without changing the dynamics of $\boldsymbol{x}$. This freedom was first pointed out by Graham [16].

A. Proof of covariance

First, we define the notions of covariance and contravariance. Suppose $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ are the slow variables, and we have a Langevin dynamics as formulated above. Consider a time-independent NTV $\boldsymbol{x} \to \boldsymbol{y} = (y_1, \ldots, y_n)$ and let $J \equiv \det(\partial y_a/\partial x_i)$ be the Jacobian. As we discussed in the introduction, $\boldsymbol{y} = (y_1, \ldots, y_n)$ are also slow variables, and hence the Langevin theory can also be formulated in terms of $\boldsymbol{y}$, and two formulations must be mathematically equivalent. As in general relativity theory, vectors and tensors appearing in two formulations are related to each other via linear transformations, whose coefficients are generically nonlinear functions of slow variables. Unlike in general relativity theory, however, metric tensor plays no role in our Langevin theory, because there is no notion of distance in the manifold of slow variables.

Suppose a vector $z_i(\boldsymbol{x})$ and a tensor $A_{ij}(\boldsymbol{x})$ are transformed to $z'_a(\boldsymbol{y})$ and $A'_{ab}(\boldsymbol{y})$ in the new coordinate system. We call them *covariant* if they transform as

$$z'_a(\boldsymbol{y}) = \frac{\partial x_i}{\partial y_a} z_i, \tag{2.14a}$$

$$A'_{ab}(\boldsymbol{y}) = \frac{\partial x_i}{\partial y_a} A_{ij}(\boldsymbol{x}) \frac{\partial x_j}{\partial y_b}. \tag{2.14b}$$

We call them *contravariant* if they transform as

$$z'_a(\boldsymbol{y}) = \frac{\partial y_a}{\partial x_i} z_i, \tag{2.15a}$$

$$A'_{ab}(\boldsymbol{y}) = \frac{\partial y_a}{\partial x_i} A_{ij}(\boldsymbol{x}) \frac{\partial y_b}{\partial x_j}. \tag{2.15b}$$

Usually the indices of contravariant vectors and tensors are displayed as superscripts instead of subscripts, whereas those of covariant vectors and tensors are displayed as subscripts. Here to unclutter the notations, we display all indices as subscripts and indicate explicitly how they transform under NTV.

We call a function $f(\boldsymbol{x})$ a *scalar* if it transforms as

$$f(\boldsymbol{x}) \to f'(\boldsymbol{y}) = f(\boldsymbol{x}). \tag{2.16a}$$

We call a function $\phi(\boldsymbol{x})$ a *density* if it transforms as

$$\phi(\boldsymbol{x}) \to \phi'(\boldsymbol{y}) = J^{-1}\phi(\boldsymbol{x}), \tag{2.17}$$

which also implies $\phi'(\boldsymbol{y})d^n\boldsymbol{y} = \phi(\boldsymbol{x})d^n\boldsymbol{x}$.

Below we will demonstrate that the nonlinear Langevin theory and Fokker-Planck theory, Eqs. (2.12) and (2.13a) respectively, keep their form under nonlinear transform of variables, but with all vectors and tensors replaced by their appropriately transformed versions [see Eqs. (2.18) below]. This is what we mean by the *covariance of the nonlinear Langevin theory and Fokker-Planck theory*.

While the old theory is characterized by $U(\boldsymbol{x})$, $L_{ij}(\boldsymbol{x})$, $b_{i\alpha}(\boldsymbol{x})$ as well as a probability density $p(\boldsymbol{x})$, the new theory is characterized by $U'(\boldsymbol{y})$, $L'_{ab}(\boldsymbol{y})$, $b'_{a\alpha}(\boldsymbol{y})$ as well as the probability density $p'(\boldsymbol{y})$. We shall directly write down the rules of transform between two theories:

$$p'(\boldsymbol{y}) = J^{-1}\, p(\boldsymbol{x}), \tag{2.18a}$$

$$U'(\boldsymbol{y}) = U(\boldsymbol{x}) + \log J, \tag{2.18b}$$

$$b'_{a\alpha}(\boldsymbol{y}) = \frac{\partial y_a}{\partial x_i} b_{i\alpha}(\boldsymbol{x}), \tag{2.18c}$$

$$L'_{ab}(\boldsymbol{y}) = \frac{\partial y_a}{\partial x_i} L_{ij}(\boldsymbol{x}) \frac{\partial y_b}{\partial x_j}. \tag{2.18d}$$

Hence, pdf $p(\boldsymbol{x})$ transform as a density, whereas $b_{i\alpha}$ and L_{ij} transform as respectively contravariant vector and tensor of rank 2. Note that the transform of generalized potential U is such that the steady-state $e^{-U(\boldsymbol{x})}$ transforms a density $e^{-U(\boldsymbol{x})}d\boldsymbol{x} = e^{-U'(\boldsymbol{y})}d\boldsymbol{y}$. Note also that Eq. (2.18d) implies that $\boldsymbol{B}$ and $\boldsymbol{Q}$ transform as

$$B'_{ab}(\boldsymbol{y}) = \frac{\partial y_a}{\partial x_i} B_{ij}(\boldsymbol{x}) \frac{\partial y_b}{\partial x_j}, \tag{2.18e}$$

$$Q'_{ab}(\boldsymbol{y}) = \frac{\partial y_a}{\partial x_i} Q_{ij}(\boldsymbol{x}) \frac{\partial y_b}{\partial x_j}. \tag{2.18f}$$

Hence, $\boldsymbol{B}$ and $\boldsymbol{Q}$ do not mix under NTV. The Fokker-Planck operator and probability current in the transformed theory are obtained analogous to Eqs. (2.13b) and (2.13c):

$$\mathcal{L}'_{\rm FP} = \partial'_a L'_{ab}(\partial'_b + (\partial'_b U')), \tag{2.18g}$$

$$j'_a = -L'_{ab}(\partial'_b + (\partial'_b U'))p' + \partial'_b(Q'_{ab}p'), \tag{2.18h}$$

where we have used the shorthand $\partial'_a = \partial/\partial y_a$.

In Appendix, we prove that under the above transform Eqs. (2.12) are transformed into

$$dy_a = -L'_{ab}\partial'_b U' dt + \partial'_b L'_{ab} dt + b'_{a\alpha} dW_\alpha, \tag{2.19a}$$

$$\dot{y}_a = -L'_{ab}\partial'_b U' + \partial'_b L'_{ab} + b'_{a\alpha}\xi_\alpha. \tag{2.19b}$$

Ito's formula plays a crucial role in proof of Eq. (2.19). Moreover, we also prove that the transformed FP operator $\mathcal{L}'_{\rm FP}$ [Eq. (2.18g)] and probability current j'_a [Eq. (2.18h)] are related to those in the old theory $\mathcal{L}_{\rm FP}$, j_i [which are respectively defined in Eqs. (2.13b) and (2.13c)] via

$$\mathcal{L}'_{\rm FP} = J^{-1}\mathcal{L}_{\rm FP} J, \tag{2.20a}$$

$$j'_a = J^{-1}(\partial y_a/\partial x_i) j_i, \tag{2.20b}$$

$$\partial'_a j'_a = J^{-1}\partial_i j_i. \tag{2.20c}$$

It then follows that

$$\begin{aligned} \partial_t p' &= J^{-1}\, \partial_t p = J^{-1}\mathcal{L}_{\rm FP} p \\ &= J^{-1}\mathcal{L}_{\rm FP} J J^{-1} p = \mathcal{L}'_{\rm FP} p' \\ &= -\partial'_a j'_a, \end{aligned} \tag{2.20d}$$

where we have used Eqs. (2.18a), (2.13a), and (2.20a). Hence, we obtain the transformed FPE:

$$\partial_t p' = \mathcal{L}'_{\rm FP} p' = -\partial'_a j'_a. \tag{2.20e}$$

The covariance of our theory is evident by comparing Eqs. (2.19) and (2.20e) with Eqs. (2.12) and (2.13a). We also note that Eq. (2.20c) implies that steady solutions keep their identity during NTV.

Comparing with the earlier works [15–18,35], our formalism does not involve affine connection or metric tensor, and establishes the covariance of Langevin theory and Fokker Planck theory simultaneously. (The essence of metric is to

define the notion of distance in the space of slow variables, which is not needed for study of statistical mechanics. Indeed, there were discussions on the *natural choice* of metric in earlier works [15–18], which did not lead to a clear-cut conclusion.) This makes our formalism much simpler and more useful. Note also that our formalism is applicable regardless of DB.

B. Spurious drift

The term $\partial_j L_{ij}(\boldsymbol{x})$ in the RHS of Eq. (2.12) may appear unpleasant because it spoils the linear relation between the deterministic force and the *thermodynamic forces* $\partial_i U$. Furthermore, because $\partial_j L_{ij}(\boldsymbol{x})$ generally contains a constant part, it also makes the most probable value of $\boldsymbol{x}$ different from the minimum of $U(\boldsymbol{x})$. We will call this term the *spurious drift*. A part of this term $\partial_j B_{ij}$ has received lots of attention [2,11,19,38]. The other part $\partial_j Q_{ij}$, though less well known, is also important to guarantee the covariance of the theory.

One might wish to remove the term $\partial_j L_{ij}(\boldsymbol{x})$ by choosing a different stochastic calculus, i.e., different interpretation of the dot product $b_{i\alpha}(\boldsymbol{x})\xi_\alpha(t)$. As pointed out by Hänggi [11] for a one-dimensional case long ago, in general neither Stratonovich nor kinetic can achieve the purpose. For high-dimensional cases, such a stochastic calculus does not exist, because the antisymmetric matrix $\boldsymbol{Q}(\boldsymbol{x})$ appears in $\partial_j L_{ij}(\boldsymbol{x})$ but not in the noise terms. One might also attempt to absorb the term $\partial_j L_{ij}(\boldsymbol{x})$ by redefining the function $U(\boldsymbol{x})$, but then $U(\boldsymbol{x})$ would lose its contact with the stationary state; i.e., it is no longer equal to $-\log p^S(\boldsymbol{x})$. Finally, one might wish to make NTV such that the new symmetric matrix B'_{ab} becomes independent of $\boldsymbol{y}$. This is generically impossible [15] for the same reason that Riemann curvature cannot be transformed away. Furthermore, even if B'_{ab} becomes constant, Q'_{ab} are generically not. Hence, Eq. (2.12) cannot be simplified further without loss of generality or physical significance.

C. Deterministic limit

When solving nonlinear Langevin dynamics, one may wish to take the *drift approximation* by throwing out the noise term in Eq. (2.12) and obtaining a deterministic equation. Such an approximation generically destroys the covariance and therefore leads to inconsistency. (One can straightforwardly repeat our proof of covariance in SI and demonstrate this point.) However, if we delete both the spurious drift and the noise term in Eq. (2.12), we obtain a deterministic equation

$$\dot{x}_i = -L_{ij}(\boldsymbol{x})\partial_j U(\boldsymbol{x}), \tag{2.21a}$$

which describes the irreversible dynamics of slow variables. More precisely, $\boldsymbol{x}$ relaxes toward the minimum of $U(\boldsymbol{x})$. The relaxation may be oscillatory if $\boldsymbol{Q} \neq 0$. This equation is covariant not under the transform as specified by Eqs. (2.18) but under the following *revised rules of transform* for U:

$$U'(\boldsymbol{y}) = U(\boldsymbol{x}). \tag{2.21b}$$

The transform of L_{ij} remains the same as in Eq. (2.18d). Proof of covariance is elementary. Anticipating Eqs. (3.16), we easily see that the deterministic theory (2.21) satisfies detailed balance if and only if the stochastic theory does so.

But how to reconcile the different transformation rules for $U(\boldsymbol{x})$, Eq. (2.18b) in the Langevin theory and Eq. (2.21b) in the deterministic theory? In generic cases, this difference indicates there is no consistent way of taking deterministic limit of a stochastic theory. However, there are two important cases where the inconsistency goes away. The first case corresponds to the *thermodynamic limit*, where both $\boldsymbol{x}$ and $U(\boldsymbol{x})$ are extensive, whereas spurious drift $\partial_j L_{ij}$, the Jacobian J, and the noises are all subextensive. If we throw out these subextensive terms, Eq. (2.12) reduces to Eq. (2.21a), and Eq. (2.18b) reduces to Eq. (2.21b), and hence we obtain the deterministic theory Eqs. (2.21). The second case corresponds to the *low-temperature limit*, where $U = F/T$, with F being the free energy and $T \to 0$. In this case, we have $\boldsymbol{x}$ independent of T, whereas L_{ij} is linear in T. The noise amplitudes $b_{i\alpha}$ then scale as $\sqrt{T}$ according to Eq. (2.4). In this low-temperature limit, again Eq. (2.12) reduces to Eq. (2.21a), and Eq. (2.18b) reduces to Eq. (2.21b), and hence we obtain the deterministic theory Eqs. (2.21).

This discussion naturally leads us to the converse question: If we are given a deterministic dynamics $\dot{x}_i = F_i(\boldsymbol{x})$ together with a generalized potential $U(\boldsymbol{x})$, can we construct a covariant stochastic theory such that $e^{-U(\boldsymbol{x})}$ is the steady state? This has been one of the perplexing questions in the *traditional Langevin approach* [2]. We are now imposing it in a more general setting that is independent of detailed balance (DB). The answer is *yes*, if we are willing to accept different transformation rules for U in the deterministic and stochastic theories. The matrix of kinetic coefficients can be found by solving the equations $L_{ij}\partial_i U = -F_i$. Inserting them back to Eq. (2.12), we find the desired covariant nonlinear Langevin theory. Without first understanding the covariant theory Eq. (2.12), however, it is very difficult to take such a leap from deterministic to stochastic.

D. Unitary limit and Hermitian limit

Following Qian [31], we define an inner product of two functions

$$(\phi, \psi)_U \equiv \int d^n\boldsymbol{x}\, e^{U(\boldsymbol{x})}\, \phi(\boldsymbol{x})\psi(\boldsymbol{x}) = (\psi, \phi)_U, \tag{2.22}$$

where $U(\boldsymbol{x})$ is the generalized potential of the Langevin theory we aim to study. Hermitian conjugate of operator $\boldsymbol{O}$ is defined as

$$(\phi, \boldsymbol{O}^\dagger\psi)_U \equiv (\boldsymbol{O}\phi, \psi)_U = (\psi, \boldsymbol{O}\phi)_U. \tag{2.23}$$

Because we only consider real functions and operators, the inner product is symmetric. Using integration by parts, we easily see that the Hermitian conjugate of ∂_i is

$$\partial_i^\dagger = -\partial_i - (\partial_i U). \tag{2.24}$$

Clearly this also means

$$-\partial_i = [\partial_i + (\partial_i U)]^\dagger. \tag{2.25}$$

For many cases, $U(\boldsymbol{x}) \to +\infty$ as $\boldsymbol{x} \to \infty$, so the choice of weight function $e^{U(\boldsymbol{x})}$ of this inner product defined in Eq. (2.22) is quite unconventional. In order for the norm of a function (ϕ, ϕ) to be finite, $\phi(\boldsymbol{x})$ needs to decay to zero at least

as fast as $e^{-U(\boldsymbol{x})/2}$. This condition is satisfied by all physically relevant distribution functions which decay as $e^{-U(\boldsymbol{x})}$.

We decompose $\mathcal{L}_{\rm FP}$ into a part $\mathcal{L}^H$ linear in B_{ij} and another part $\mathcal{L}^A$ linear in Q_{ij}:

$$\mathcal{L}_{\rm FP} = \mathcal{L}^H + \mathcal{L}^A = -\partial_i L_{ij}\partial_j^\dagger, \tag{2.26a}$$

$$\mathcal{L}^H = -\partial_i B_{ij}\partial_j^\dagger = \mathcal{L}^{H\dagger}, \tag{2.26b}$$

$$\mathcal{L}^A = -\partial_i Q_{ij}\partial_j^\dagger = -\mathcal{L}^{A\dagger}. \tag{2.26c}$$

Hence $\mathcal{L}^H$ is Hermitian, whereas $\mathcal{L}^A$ is *anti-Hermitian*. Since B_{ij} is non-negative, the Hermitian operator $\mathcal{L}^H$ is also non-negative, which means that none of its eigenvalues is negative.

Likewise, the current (2.13c) can be also decomposed:

$$j_i = j_i^H + j_i^A, \tag{2.27a}$$

$$j_i^H = -B_{ij}(\partial_j + (\partial_j U))p = B_{ij}\partial_j^\dagger p, \tag{2.27b}$$

$$j_i^A = Q_{ij}\partial_j^\dagger p + \partial_j(Q_{ij}p). \tag{2.27c}$$

Since $\boldsymbol{B}$ and $\boldsymbol{Q}$ do not mix under NTV, these decompositions of $\mathcal{L}_{\rm FP}$ and j_i are covariant. Note that at the steady state $\partial_i^\dagger p^S = -(\partial_i + \partial_i U)e^{-U} = 0$, and hence j_i^H vanishes, but j_i^A does not need to vanish at stationarity.

In general, the evolution operator $e^{t\mathcal{L}_{\rm FP}}$ of the Fokker-Planck theory is neither unitary nor Hermitian. There are, however, two particularly nice limits of the theory. If $B_{ij} \to 0$, then $\mathcal{L}_{\rm FP} \to \mathcal{L}^A$, which is anti-Hermitian. The resulting evolution operator $e^{t\mathcal{L}^A}$ is deterministic and unitary with imaginary eigenvalues, corresponding to oscillatory dynamics. A simple example of this is Hamiltonian dynamics, where $\mathcal{L}^A$ is just the Liouville operator. Note, however, this unitary theory is very different from the deterministic irreversible theory shown in Eq. (2.21). On the other hand, if $Q_{ij} \to 0$, then $\mathcal{L}_{\rm FP} \to \mathcal{L}^H$ and the evolution operator $e^{t\mathcal{L}^H}$ is Hermitian, with real and nonpositive engeivatlues. The system relaxes toward steady state without oscillation. If neither $\mathcal{L}^H$ nor $\mathcal{L}^A$ vanishes, then we may try to treat either $\mathcal{L}^H$ or $\mathcal{L}^A$ as perturbation. The fact that $\mathcal{L}^H, \mathcal{L}^A$ are respectively Hermitian and anti-Hermitian makes the expansion particularly convenient.

III. TIME REVERSAL AND DETAILED BALANCE

Detailed balance is a reflection of time-reversal symmetry of the microscopic dynamics. To discuss time-reversal symmetry, we choose slow variables such that each component has definite time parity $\varepsilon_i = 1$, or -1. Under time reversal, we have $x_i \to \varepsilon_i x_i$. The vector $\boldsymbol{x}$ is time reversed to $\boldsymbol{x}^*$ with components $\varepsilon_i x_i$. For Hamiltonian systems, we have $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p})$, where $\boldsymbol{q}, \boldsymbol{p}$ are respectively canonical coordinates and momenta. Hence, we have $\boldsymbol{x}^* = (\boldsymbol{q}, -\boldsymbol{p})$. As a consequence, the integral measure $d^n\boldsymbol{x}$ is also invariant under time reversal, i.e., $d^n\boldsymbol{x} = d^n\boldsymbol{x}^*$. A stationary Markov process is said to be reversible, or satisfy detailed balance, if its steady-state two-time joint PDF satisfy [1,39]

$$p_2^S(\boldsymbol{x}_1, t; \boldsymbol{x}_0, 0) = p_2^S(\boldsymbol{x}_0^*, t; \boldsymbol{x}_1^*, 0). \tag{3.1}$$

(Here we assume that there is no magnetic field or other external field which breaks time-reversal symmetry explicitly, so that the subtle difference between reversibility and detailed balance does not arise.) The steady state then becomes the *equilibrium state*.

Suppose the system start from a state $\boldsymbol{x}_0$ at $t = 0$, the initial probability density is just $p(\boldsymbol{x}, 0) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$, whereas the probability density $p(\boldsymbol{x}_1, t|\boldsymbol{x}_0, 0)$ at time t, conditioned on the initial position $\boldsymbol{x}_0$, is given by

$$\begin{aligned} p(\boldsymbol{x}_1, t|\boldsymbol{x}_0, 0) &= e^{tL_{\rm FP}(\boldsymbol{x}_1)}\delta(\boldsymbol{x}_1 - \boldsymbol{x}_0) \\ &= \int d^n\boldsymbol{x}\, \delta(\boldsymbol{x} - \boldsymbol{x}_1)e^{tL_{\rm FP}(\boldsymbol{x})}\delta(\boldsymbol{x} - \boldsymbol{x}_0). \end{aligned} \tag{3.2}$$

Note that if we integrate $p(\boldsymbol{x}_1, t|\boldsymbol{x}_0, 0)$ over $\boldsymbol{x}_1$, we obtain unity, as it should be. Now because of the Markovian property, the steady-state two-time joint PDF $p_2^S(\boldsymbol{x}_1, t; \boldsymbol{x}_0, 0)$ can be obtained as

$$p_2^S(\boldsymbol{x}_1, t; \boldsymbol{x}_0, 0) = p(\boldsymbol{x}_1, t|\boldsymbol{x}_0, 0)\, e^{-U(\boldsymbol{x}_0)}. \tag{3.3}$$

A. Time reversal

Starting from Eq. (3.2), we can show

$$\begin{aligned} &e^{U(\boldsymbol{x}_1)}p(\boldsymbol{x}_1, t|\boldsymbol{x}_0, 0) \\ &= \int d^n\boldsymbol{x}\; e^{U(\boldsymbol{x})}\delta(\boldsymbol{x} - \boldsymbol{x}_1)e^{tL_{\rm FP}(\boldsymbol{x})}\delta(\boldsymbol{x} - \boldsymbol{x}_0) \\ &= [\delta(\boldsymbol{x} - \boldsymbol{x}_1), e^{tL_{\rm FP}(\boldsymbol{x})}\delta(\boldsymbol{x} - \boldsymbol{x}_0)]_U \\ &= [\delta(\boldsymbol{x} - \boldsymbol{x}_0), e^{tL_{\rm FP}^\dagger(\boldsymbol{x})}\delta(\boldsymbol{x} - \boldsymbol{x}_1)]_U \\ &= [\delta(\boldsymbol{x}^* - \boldsymbol{x}_0^*), e^{tL_{\rm FP}^\dagger(\boldsymbol{x})}\delta(\boldsymbol{x}^* - \boldsymbol{x}_1^*)]_U \\ &= \int d^n\boldsymbol{x}\, e^{U(\boldsymbol{x})}\delta(\boldsymbol{x}^* - \boldsymbol{x}_0^*)e^{tL_{\rm FP}^\dagger(\boldsymbol{x})}\delta(\boldsymbol{x}^* - \boldsymbol{x}_1^*) \\ &= \int d^n\boldsymbol{x}\, e^{U(\boldsymbol{x}^*)}\delta(\boldsymbol{x} - \boldsymbol{x}_0^*)e^{tL_{\rm FP}^\dagger(\boldsymbol{x}^*)}\delta(\boldsymbol{x} - \boldsymbol{x}_1^*). \end{aligned} \tag{3.4}$$

In the second and third equalities, we have used respectively definitions (2.22) and (2.23). In the fourth equality, we have used the time-reversal invariance property of the δ function $\delta(\boldsymbol{x}) = \delta(\boldsymbol{x}^*)$. In the sixth equality, we have transformed the dummy variable $\boldsymbol{x} \to \boldsymbol{x}^*$. This does not change the integral, since the volume measure $d^n\boldsymbol{x}$ is invariant under time reversal. Note that the weight function is also transformed from $U(\boldsymbol{x})$ to $U(\boldsymbol{x}^*)$.

We define a new Fokker-Planck operator $\tilde{L}_{\rm FP}(\boldsymbol{x})$ as

$$\begin{aligned} \tilde{L}_{\rm FP}(\boldsymbol{x}) &\equiv L_{\rm FP}^\dagger(\boldsymbol{x}^*) = \partial_i^* L_{ji}(\boldsymbol{x}^*)\{\partial_j^* + [\partial_j^* U(\boldsymbol{x}^*)]\} \\ &= \frac{\partial}{\partial x_i^*}L_{ji}(\boldsymbol{x}^*)\left\{\frac{\partial}{\partial x_j^*} + \left[\frac{\partial}{\partial x_j^*}U(\boldsymbol{x}^*)\right]\right\}, \end{aligned} \tag{3.5}$$

where $x_i^* = \varepsilon_i x_i$ and $\partial_i^* = \varepsilon_i \partial/\partial x_i$. Further defining two functions $\tilde{L}_{ij}(\boldsymbol{x})$ and $\tilde{U}(\boldsymbol{x})$ via

$$\tilde{L}_{ij}(\boldsymbol{x}) \equiv \varepsilon_i L_{ji}(\boldsymbol{x}^*)\varepsilon_j, \quad \tilde{U}(\boldsymbol{x}) \equiv U(\boldsymbol{x}^*), \tag{3.6}$$

we can write $\tilde{L}_{\rm FP}(\boldsymbol{x})$ in the following form:

$$\tilde{L}_{\rm FP}(\boldsymbol{x}) = \partial_i \tilde{L}_{ij}(\partial_j + \partial_j \tilde{U}(\boldsymbol{x})), \tag{3.7}$$

which has the standard form of Fokker-Planck operator [cf. Eq. (2.13b)], but with $\boldsymbol{L}, U$ replaced by $\tilde{\boldsymbol{L}}$ and $\tilde{U}$. If we recall the decomposition (2.11), Eq. (3.6) also implies

$$\tilde{B}_{ij}(\boldsymbol{x}) = \varepsilon_i B_{ij}(\boldsymbol{x}^*)\varepsilon_j, \quad \tilde{Q}_{ij}(\boldsymbol{x}) = -\varepsilon_i Q_{ij}(\boldsymbol{x}^*)\varepsilon_j. \tag{3.8}$$

Let us study the steady state of the Langevin process defined by $\tilde{L}_{\rm FP}(\boldsymbol{x})$, which for now shall be referred as the *tilde process*. The steady-state distribution is

$$\tilde{p}^S(\boldsymbol{x}) = e^{-\tilde{U}(\boldsymbol{x})} = e^{-\tilde{U}(\boldsymbol{x}^*)} = p^S(\boldsymbol{x}^*), \tag{3.9}$$

which is the time reversal of the steady state of the original process $e^{-U(\boldsymbol{x})}$. The steady-state probability current of the tilde process can be obtained from Eq. (2.9):

$$\tilde{j}_i^S(\boldsymbol{x}) = \partial_j(\tilde{Q}_{ij}(\boldsymbol{x})e^{-\tilde{U}(\boldsymbol{x})}), \tag{3.10}$$

which is related to the original steady-state current via

$$\tilde{j}_i^S(\boldsymbol{x}^*) = -\varepsilon_i j_i^S(\boldsymbol{x}). \tag{3.11}$$

Equation (3.4) can now be rewritten as

$$\begin{aligned} e^{U(\boldsymbol{x}_1)}p(\boldsymbol{x}_1, t|\boldsymbol{x}_0, 0) &= \int_{\boldsymbol{x}} e^{\tilde{U}(\boldsymbol{x})}\delta(\boldsymbol{x}-\boldsymbol{x}_0^*)e^{t\tilde{L}_{\rm FP}(\boldsymbol{x})}\delta(\boldsymbol{x}-\boldsymbol{x}_1^*) \\ &= e^{\tilde{U}(\boldsymbol{x}_0^*)}\tilde{p}(\boldsymbol{x}_0^*, t|\boldsymbol{x}_1^*, 0), \end{aligned} \tag{3.12}$$

where $\tilde{p}(\cdot|\cdot)$ is the transition probability density of the tilde process. Further multiplying both sides by $e^{-\tilde{U}(\boldsymbol{x}_1^*)-U(\boldsymbol{x}_0)}$, we obtain

$$p(\boldsymbol{x}_1, t|\boldsymbol{x}_0, 0)e^{-U(\boldsymbol{x}_0)} = \tilde{p}(\boldsymbol{x}_0^*, t|\boldsymbol{x}_1^*, 0)e^{-\tilde{U}(\boldsymbol{x}_1^*)}. \tag{3.13}$$

According to Eq. (3.3), this can be further written as an equality between the steady-state two-time joint pdf of the original process and the tilde process:

$$p_2^S(\boldsymbol{x}_1, t; \boldsymbol{x}_0, 0) = \tilde{p}_2^S(\boldsymbol{x}_0^*, t; \boldsymbol{x}_1^*, 0). \tag{3.14}$$

Equations (3.9), (3.11), and (3.14) demonstrate the physical significance of the tilde process defined by $\tilde{L}_{\rm FP}$: It is the *macroscopic time reversal* of original process corresponding to $L_{\rm FP}$, because all macroscopic properties are reversed.

B. Covariant formulation of detailed balance

Recall that detailed balance is defined by the condition (3.1). Combining this with Eq. (3.14), we find

$$p_2^S(\boldsymbol{x}_1, t; \boldsymbol{x}_0, 0) = \tilde{p}_2^S(\boldsymbol{x}_1, t; \boldsymbol{x}_0, 0),$$

which in turn implies

$$\tilde{\mathcal{L}}_{\rm FP}(\boldsymbol{x}) = \mathcal{L}_{\rm FP}(\boldsymbol{x}). \tag{3.15}$$

In view of Eqs. (3.7), this is further equivalent to

$$\tilde{U}(\boldsymbol{x}) \equiv U(\boldsymbol{x}^*) = U(\boldsymbol{x}), \tag{3.16a}$$

$$\tilde{L}_{ij}(\boldsymbol{x}) \equiv \varepsilon_i L_{ji}(\boldsymbol{x}^*)\varepsilon_j = L_{ij}(\boldsymbol{x}). \tag{3.16b}$$

Equations (3.16b) are further equivalent to

$$\tilde{B}_{ij}(\boldsymbol{x}) \equiv \varepsilon_i B_{ij}(\boldsymbol{x}^*)\varepsilon_j = B_{ij}(\boldsymbol{x}), \tag{3.16c}$$

$$\tilde{Q}_{ij}(\boldsymbol{x}) \equiv -\varepsilon_i Q_{ij}(\boldsymbol{x}^*)\varepsilon_j = Q_{ij}(\boldsymbol{x}). \tag{3.16d}$$

Equations (3.16) are the necessary and sufficient conditions for *detailed balance* of Langevin dynamics (2.12). In the setting of linear response theory, these conditions are better known as Onsager-Casimir reciprocal symmetry of the kinetic coefficients L_{ij} [40,41]. They agree with Eq. (2.72) of Ref. [37] and also with those derived by Gardiner [1], Sec. 5.3.5. Note, however, Gardiner's (5.3.53 iii) is much more complicated and less transparent.

Combining Eqs. (3.11) and (3.16d), we find that for reversible Langevin dynamics, the stationary probability current transforms under time reversal as

$$j_i^S(\boldsymbol{x}^*) = -\epsilon_i\, j_i^S(\boldsymbol{x}). \tag{3.17}$$

We can now explicitly show that the conditions of DB (3.16) are covariant under NTVs that respect time-reversal symmetry. But the latter, we mean that the each y_a of the new variables $\boldsymbol{y} = (y_1, \ldots, y_n)$ also has definite time-parity ε_a, and further satisfy $y_a^*(\boldsymbol{x}^*) = y_a(\boldsymbol{x})$. But this also means $\partial y_a(\boldsymbol{x})/\partial x_i = \varepsilon_a \partial y_a(\boldsymbol{x}^*)/\partial x_i\, \varepsilon_i$, and hence the Jacobian is invariant under time reversal: $J(\boldsymbol{x}^*) = J(\boldsymbol{x})$. Using these together with Eqs. (2.18b) and (2.18d), we can easily see that Eqs. (3.16a) and (3.16b) reduce to

$$\tilde{U}'(\boldsymbol{y}) \equiv U'(\boldsymbol{y}^*) = U'(\boldsymbol{y}), \tag{3.18a}$$

$$\tilde{L}'_{ab}(\boldsymbol{y}) \equiv \varepsilon_a L_{ba}(\boldsymbol{y}^*)\varepsilon_b, \tag{3.18b}$$

which are just the conditions of DB in the new variables. Hence, DB keeps its identity during NTV, as we expected.

C. Monotonic decrease of free energy

We can show that a functional of pdf $p(\boldsymbol{x}, t)$ monotonically decreases as a function of time. For reversible Markov process which satisfies DB, this result can be understood as a reflection of the second law of thermodynamics, which dictates that the total entropy of an isolated system can only increase. We will call this functional *free energy*, which is defined as

$$F[p(t)] \equiv T \int d\boldsymbol{x}\, p(\boldsymbol{x}, t)[U(\boldsymbol{x}) + \log p(\boldsymbol{x}, t)], \tag{3.19}$$

where T (temperature) is just a constant of proportionality and plays no significant role here. The time derivative of $F[p(t)]$ can be calculated using the Fokker-Planck equation:

$$\begin{aligned} \frac{1}{T}\frac{dF}{dt} &= \int d\boldsymbol{x}[U(\boldsymbol{x}) + \log p(\boldsymbol{x}, t)](\partial_t p) \\ &= -\int d\boldsymbol{x}\,[U(\boldsymbol{x}) + \log p(\boldsymbol{x}, t)]\partial_i L_{ij}\partial_j^\dagger p \\ &= \int d\boldsymbol{x}\,\{\partial_i[U(\boldsymbol{x}) + \log p(\boldsymbol{x}, t)]\}L_{ij}\partial_j^\dagger p \\ &= \int d\boldsymbol{x}\, p^{-1}((\partial_i U) + (\partial_i)p)L_{ij}\partial_j^\dagger p \\ &= -\int d\boldsymbol{x}\, p^{-1}(\partial_i^\dagger p)L_{ij}(\partial_j^\dagger p) \\ &= -\int d\boldsymbol{x}\, p^{-1}(\partial_i^\dagger p)B_{ij}(\partial_j^\dagger p) \leqslant 0, \end{aligned} \tag{3.20}$$

where in the last line, we have used the facts that the antisymmetric part of L_{ij} does not contribute to the quadratic form and that the matrix B_{ij} is semipositive definite. Hence, the free energy decreases over time.

If $\boldsymbol{B}$ is positive definite, the above inequality is sufficient to guarantee that the system converges to a unique equilibrium state with minimal free energy. Furthermore, using Eq. (2.27), we can also express Eq. (3.20) in terms of the Hermitian

probability current j^H:

$$\frac{1}{T}\frac{dF}{dt} = -\int d\boldsymbol{x}\, P^{-1} j_i^H B_{ij}^{-1} j_j^H. \tag{3.21}$$

Hence, we find that the Hermitian current j^H but not the anti-Hermitian current j^A contributes to entropy production [cf. Eqs. (2.27)]. It is then appropriate to call $\boldsymbol{Q}$ the reactive couplings and $\boldsymbol{B}$ the dissipative couplings. This is consistent with Eqs. (3.16d) and (3.16c), which say that $\boldsymbol{Q}$ couple even variables to odd variables, while $\boldsymbol{B}$ couple variables with the same signature of time reversal. For systems without DB, however, both $\mathcal{L}^H$ and $\mathcal{L}^A$ may contribute to dissipation in general and the term *reactive* or *conservative* cannot be applied to $\boldsymbol{Q}$.

IV. APPLICATIONS

In this section, we discuss two simple applications of our theory. We will start from the simple and well-known case of linear response theory, and then discuss a slightly more complicated case of weakly damped classical Hamiltonian system. Both systems, however, are reversible and have additive noises. More complicated cases (without detailed balance and with multiplicative noises) will be discussed in future publications.

A. Linear response theory

The simplest case is that all kinetic coefficients L_{ij} are constants, and the generalized potential is quadratic:

$$U(\boldsymbol{x}) = \tfrac{1}{2} s_{ij} x_i x_j. \tag{4.1a}$$

The Langevin equations then become linear:

$$\dot{x}_i + (B_{ij} + Q_{ij}) s_{jk} x_k = \eta_i, \tag{4.1b}$$

$$\langle \eta_i(t_1)\eta_j(t_2)\rangle = 2B_{ij}\delta(t_1 - t_2), \tag{4.1c}$$

where the noises are related to ξ_α defined in Eq. (2.1) via

$$\eta_j(t) = b_{j\alpha}\xi_\alpha(t). \tag{4.1d}$$

The detailed balance conditions (3.16) guarantee that B_{ij} couple only variables with same time-reversal symmetry, and Q_{ij} only couple variables with opposite time-reversal symmetry. These are precisely the reciprocal symmetry discovered by Onsager [40] and Casimir [41] long ago.

B. Weakly damped Hamiltonian system

The slightly more complex case is that L_{ij} remain constants, but $U(\boldsymbol{x})$ becomes an arbitrary function bound from below. One of the simplest realizations of this case is a one-dimensional classical Harmonic system damped by weak ambient noises. The slow variables are $\boldsymbol{x} = (q, p)$, where q, p are the canonical coordinate and momentum. Let $H(q, p) = p^2/2m + V(q)$ be the Hamiltonian. The steady state is the equilibrium state with Gibbs-Boltzmann distribution $p^S(q, p) = e^{-\beta H(q,p)+\beta F(T)}$, where $F(T)$ is the free energy. Hence, the generalized potential is

$$U = \beta H - \beta F = \beta[p^2/2m + V(q) - F]. \tag{4.2}$$

The Langevin equations are

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m}, \tag{4.3a}$$

$$\dot{p} = -\gamma\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q} + \eta(t) = -\gamma\dot{q} - \partial_q V(\boldsymbol{q}) + \eta(t). \tag{4.3b}$$

These equations can be rewritten as the standard form:

$$\begin{pmatrix}\dot{q}\\ \dot{p}\end{pmatrix} + \begin{pmatrix}0 & -T\\ T & T\gamma\end{pmatrix}\begin{pmatrix}\partial_q U\\ \partial_p U\end{pmatrix} = \begin{pmatrix}0\\ \eta(t)\end{pmatrix}. \tag{4.4}$$

From this, we read off the matrix of kinetic coefficients:

$$\boldsymbol{L} = \begin{pmatrix}0 & -T\\ T & T\gamma\end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix}0 & 0\\ 0 & T\gamma\end{pmatrix}, \quad \boldsymbol{Q} = \begin{pmatrix}0 & -T\\ T & 0\end{pmatrix}. \tag{4.5}$$

The detailed balance conditions (3.16) can be easily verified. The noise variance is given by the Einstein relation

$$\langle \eta(t_1)\eta(t_2)\rangle = 2T\gamma\,\delta(t_1 - t_2). \tag{4.6}$$

Note that L_{ij} are proportional, whereas U is antiproportional, to temperature T, as we have claimed in Sec. II C. As a consequence, Eqs. (4.3) depend on temperature only through the noise variance (assuming that the friction coefficient γ is independent of T). If we take the zero-temperature limit, we obtain a set of deterministic irreversible equations:

$$\dot{q} = \frac{p}{m}, \tag{4.7a}$$

$$\dot{p} = -\gamma\dot{q} - \partial_q V(\boldsymbol{q}). \tag{4.7b}$$

This is the deterministic limit we discussed in Sec. II C. As discussed there, the same set of equations can also be obtained from Eqs. (4.3) in the thermodynamic limit, where the particle is very massive, so that the noise can be neglected.

As discussed in Sec. II D, if we let the symmetric part of kinetic coefficients go to zero, $\gamma \to 0$, we obtain a unitary and deterministic dynamics. The Langevin equations then become

$$\dot{q} = \frac{p}{m}, \tag{4.8a}$$

$$\dot{p} = -\partial_q V(\boldsymbol{q}). \tag{4.8b}$$

But these are just the Hamiltonian dynamics without damping. Hence, the unitary limit corresponds to the limit of vanishing friction and noise.

V. CONCLUSION

We conclude our work with a few comments on the choices of stochastic calculus. It is known [1,3] that the Ito-Langevin equation (2.1) is mathematically equivalent to the Stratonovich-Langevin equation

$$\dot{x}_i = F_i^S(\boldsymbol{x}, t) + b_{i\alpha}(\boldsymbol{x}) \circ \xi_\alpha(t), \tag{5.1}$$

if we impose the following relation between F_i^S and F_i:

$$F_i^S = F_i - \tfrac{1}{2} b_{j\alpha}\partial_j b_{i\alpha}. \tag{5.2}$$

Here, in Eq. (5.1) the product $b_{i\alpha}(\boldsymbol{x}) \circ \xi_\alpha(t)$ is interpreted in the sense of Stratonovich. Hence, our covariant Langevin

equation (2.12) can also be represented as a Stratonovich-Langevin equation. One can further show that $F_i^S(\boldsymbol{x},t)$ in Eq. (5.1) transforms as contravariant vector under NTV. In fact, this simple transformation law for $F_i^S(\boldsymbol{x},t)$ has been deemed as a major advantage of Stratonovich over Ito. Of course, in this work we have demonstrated that Ito-Langevin is also fully covariant as long as it is parameterized in terms of U, $\boldsymbol{L}$, so Ito-Langevin is at least as convenient as Stratonovich-Langevin.

Stratonovich-Langevin, however, has inconvenient features. Its drift term $F_i^S(\boldsymbol{x},t)$, though covariant, is connected to observables $U(\boldsymbol{x})$ and $\boldsymbol{L}$ in a more complex way, and hence conditions of DB become obscure. More importantly, as pointed out by Graham [16], $F_i^S(\boldsymbol{x},t)$ is changed by $\boldsymbol{x}$-dependent right rotation of noises $b_{i\alpha}(\boldsymbol{x})$, and hence has lower symmetry than Ito-Langevin. Qualitatively, it indicates that the Stratonovich-Langevin equation is sensitive to details of fast variables, a rather strange feature. Finally, as is well known, the Stratonovich-Langevin equation (5.1) "looks into the future," which makes numerical studies very inconvenient and violates the principle of causality. Quite obviously, all these comments apply to other non-Ito schemes as well. Our conclusion is that the formalism developed here has higher symmetry and clearer physical meanings, as well as much simpler behaviors under general NTV, and hence is a more natural formalism for nonlinear Langevin dynamics with multiplicative white noises.

ACKNOWLEDGMENTS

The authors acknowledge support from NSFC via Grants No. 11674217 (X.X.) and No. 11675017 (Z.C.T.), as well as Shanghai Municipal Science and Technology Major Project (Grant No. 2019SHZDZX01). X.X. also is thankful for additional support from a Shanghai talent program.

APPENDIX: COVARIANCE OF LANGEVIN AND FOKKER-PLANCK EQUATIONS

Our goal is to prove that under the rules of Eqs. (2.18), the nonlinear Langevin equation, the Fokker-Planck operator, probability current, and FPE transform respectively as Eqs. (2.19) and (2.20). For this purpose, we first need to prove a few useful identities.

1. Jacobi's formula

Let us first establish two very useful identities about Jacobian. Let $\boldsymbol{A}$ be a nonsingular square matrix, with determinant $\det\boldsymbol{A}$, inverse $\boldsymbol{A}^{-1}$, and then *Jacobi's formula* says (here d denotes differential)

$$d\log\det\boldsymbol{A} = \mathrm{Tr}\,\boldsymbol{A}^{-1}d\boldsymbol{A}. \tag{A1}$$

Applying this formula to matrix $\partial y_a/\partial x_i$ which has determinant J and inverse $\partial x_i/\partial y_a$, we have

$$d\log J = J^{-1}dJ = \left(\frac{\partial x_i}{\partial y_a}\right)d\left(\frac{\partial y_a}{\partial x_i}\right). \tag{A2}$$

Hence $\forall\, j$ and we have

$$\begin{aligned} J^{-1}\frac{\partial J}{\partial x_j} &= \left(\frac{\partial x_i}{\partial y_a}\right)\frac{\partial}{\partial x_j}\left(\frac{\partial y_a}{\partial x_i}\right) \\ &= \left(\frac{\partial x_i}{\partial y_a}\right)\left(\frac{\partial^2 y_a}{\partial x_j\partial x_i}\right) = \frac{\partial}{\partial y_a}\frac{\partial y_a}{\partial x_j}, \end{aligned} \tag{A3}$$

from which we further prove a useful identity:

$$\frac{\partial}{\partial y_a}\left(\frac{\partial y_a}{\partial x_i}J^{-1}\right) = 0. \tag{A4}$$

Swapping the roles of $\boldsymbol{x}$ and $\boldsymbol{y}$, we obtain a result that is reciprocal to Eq. (A4):

$$\frac{\partial}{\partial x_i}\left(\frac{\partial x_i}{\partial y_a}J\right) = 0. \tag{A5}$$

Note that in the above when taking partial derivative with respect to x_i (y_a) it is always understood that all other x_j, $j \neq i$ (y_b, $b \neq a$) are fixed. Equations (A4) and (A5) will be very useful below.

2. Proof of Eq. (2.19)

Let us prove Eq. (2.19). Let us rewrite Eqs. (2.12a) and (2.19a) in equivalent forms preferred by mathematicians:

$$dx_i = (-L_{ij}\partial_j U + \partial_j L_{ij})dt + b_{i\alpha}dW_\alpha(t), \tag{A6a}$$

$$dy_a = (-L'_{ab}\partial_b U' + \partial_b L'_{ab})dt + b'_{a\alpha}dW_\alpha(t). \tag{A6b}$$

Here $dW_\alpha(t)$ are differentials of Wiener's processes and obey Ito's rule [3]:

$$dW_\alpha(t)dW_\beta(t) = \delta_{\alpha\beta}dt. \tag{A7}$$

Let $y(\boldsymbol{x})$ be a function of $\boldsymbol{x}$, and so Ito's formula [1,3] relates the differential of $\boldsymbol{y}$ to that of $\boldsymbol{x}$:

$$dy = \frac{\partial y}{\partial x_i}dx_i + \frac{1}{2}\frac{\partial^2 y}{\partial x_i\partial x_j}dx_i dx_j. \tag{A8}$$

We will derive Eq. (A6b) from Eq. (A6a) using rules Eqs. (2.18) and *Ito's formula,* Eq. (A8). We use Eq. (A6a) to rewrite dx_i and dx_j in Eq. (A8) in terms of dt and $dW(t)$. For the quadratic term $dx_i dx_j$, however, we only need to keep terms proportional to $(dW)^2 \sim dt$. Further using Ito's rule (A7), we obtain

$$\begin{aligned} dy_a = &-\frac{\partial y_a}{\partial x_i}L_{ij}(\partial_j U)dt + \frac{\partial y_a}{\partial x_i}\partial_j L_{ij}dt \\ &+ \frac{\partial^2 y_a}{\partial x_i\partial x_j}B_{ij}dt + \frac{\partial y_a}{\partial x_i}b_{i\alpha}dW_\alpha. \end{aligned} \tag{A9}$$

Using Eqs. (2.18b) and (2.18d) as well as the chain rule, the first term in the RHS of Eq. (A9) can be rewritten as $-L'_{ab}(\partial_b U' - \partial_b \log j)dt$. Using Eq. (2.18c), the last term can be rewritten as $b'_{a\alpha}dW_\alpha$. Hence, we have

$$dy_a = (-L'_{ab}\partial_b U' + \partial_b L'_{ab})dt + b'_{a\alpha}dW_\alpha + \Psi\, dt, \tag{A10}$$

$$\Psi \equiv L'_{ab}\partial_b \log j + \frac{\partial y_a}{\partial x_i}\partial_j L_{ij} - \partial_b L'_{ab} + \frac{\partial^2 y_a}{\partial x_i\partial x_j}B_{ij}. \tag{A11}$$

Note that Eq. (A10) differs from Eq. (A6b) only by the term Ψdt, which will be shown to vanish identically. The first term

in Ψ can be calculated using Eq. (A3) and chain rule as well as commutativity of derivatives ∂_i, ∂_j:

$$\begin{aligned} L'_{ab}\partial_b \log j &= L'_{ab}\frac{\partial x_i}{\partial y_c}\frac{\partial}{\partial y_b}\frac{\partial y_c}{\partial x_i} = \frac{\partial y_a}{\partial x_j}L_{jk}\frac{\partial y_b}{\partial x_k}\frac{\partial x_i}{\partial y_c}\frac{\partial}{\partial y_b}\frac{\partial y_c}{\partial x_i} \\ &= \frac{\partial y_a}{\partial x_j}L_{jk}\frac{\partial x_i}{\partial y_c}\frac{\partial}{\partial x_k}\frac{\partial y_c}{\partial x_i} = \frac{\partial y_a}{\partial x_j}L_{jk}\frac{\partial x_i}{\partial y_c}\frac{\partial}{\partial x_i}\frac{\partial y_c}{\partial x_k} \\ &= \frac{\partial y_a}{\partial x_j}L_{jk}\frac{\partial}{\partial y_c}\frac{\partial y_c}{\partial x_k} = \frac{\partial y_a}{\partial x_i}L_{ij}\frac{\partial}{\partial y_b}\frac{\partial y_b}{\partial x_j}. \end{aligned} \tag{A12}$$

Using Eq. (2.18d), negative third term in Ψ can be rewritten as

$$\begin{aligned} \partial_b L'_{ab} &= \frac{\partial}{\partial y_b}\left(\frac{\partial y_a}{\partial x_i}L_{ij}\frac{\partial y_b}{\partial x_j}\right) \\ &= \frac{\partial y_a}{\partial x_i}\frac{\partial y_b}{\partial x_j}\frac{\partial}{\partial y_b}L_{ij} + \frac{\partial}{\partial y_b}\frac{\partial y_a}{\partial x_i}L_{ij}\frac{\partial y_b}{\partial x_j} + \frac{\partial y_a}{\partial x_i}L_{ij}\frac{\partial}{\partial y_b}\frac{\partial y_b}{\partial x_j} \\ &= \frac{\partial y_a}{\partial x_i}\partial_j L_{ij} + \left(\frac{\partial^2 y_a}{\partial x_j \partial x_i}\right)L_{ij} + \frac{\partial y_a}{\partial x_i}L_{ij}\frac{\partial}{\partial y_b}\frac{\partial y_b}{\partial x_j} \\ &= \frac{\partial y_a}{\partial x_i}\partial_j L_{ij} + \left(\frac{\partial^2 y_a}{\partial x_j \partial x_i}\right)B_{ij} + \frac{\partial y_a}{\partial x_i}L_{ij}\frac{\partial}{\partial y_b}\frac{\partial y_b}{\partial x_j}. \end{aligned} \tag{A13}$$

In the final step, we have used the fact that $\frac{\partial^2 y_a}{\partial x_i \partial x_j}$ is symmetric in i, j and hence $\frac{\partial^2 y_a}{\partial x_i \partial x_j}Q_{ij} = 0$.

Substituting Eqs. (A12) and (A13) back into Eq. (A11), we finally see that all terms cancel exactly in the RHS, and hence Ψ vanishes identically. This means Eq. (A10) reduces to Eq. (A6b), as we expected.

3. Proof of Eq. (2.20a)

First, using Eq. (A5), we immediately obtain the *operator identity*:

$$\frac{\partial}{\partial y_a} = J^{-1}\frac{\partial}{\partial x_l}\frac{\partial x_l}{\partial y_a}J. \tag{A14}$$

Also using the chain rule, (A5) as well as Eq. (2.18b), we have

$$\begin{aligned} \left(\frac{\partial}{\partial y_b} + \frac{\partial U'}{\partial y_b}\right) &= \frac{\partial x_k}{\partial y_b}\left(\frac{\partial}{\partial x_k} + \frac{\partial U'}{\partial x_k}\right) \\ &= \frac{\partial x_k}{\partial y_b}\left(\frac{\partial}{\partial x_k} + \frac{\partial U}{\partial x_k} + J^{-1}\frac{\partial j}{\partial x_k}\right) \\ &= J^{-1}\frac{\partial x_k}{\partial y_b}\left(\frac{\partial}{\partial x_k} + \frac{\partial U}{\partial x_k}\right)J. \end{aligned} \tag{A15}$$

Now take the product of Eq. (A14), Eq. (2.18d), and Eq. (A15) consecutively. On the left-hand side (LHS), we obtain $\mathcal{L}'_{\rm FP}$ according to Eq. (2.18g). On the RHS, we find

$$J^{-1}\frac{\partial}{\partial x_i}L_{ij}\left(\frac{\partial}{\partial x_i} + \frac{\partial U}{\partial x_i}\right)J = J^{-1}\mathcal{L}_{\rm FP}J. \tag{A16}$$

Hence, we obtain Eq. (2.20a) as an operator identity.

4. Proof of Eqs. (2.20b) and (2.20c)

Taking the product Eq. (2.18d) × Eq. (A15) and acting on Eq. (2.18a), we obtain

$$-L'_{ab}(\partial'_a + (\partial'_b U'))p' = -J^{-1}\frac{\partial y_a}{\partial x_i}L_{ij}(\partial_j + (\partial_j U))p. \tag{A17}$$

Using Eqs. (2.18f) and (2.18a), we can also show

$$\begin{aligned} \partial'_b(Q'_{ab}p') &= \frac{\partial}{\partial y_b}\left(\frac{\partial y_a}{\partial x_i}Q_{ij}\frac{\partial y_b}{\partial x_j}J^{-1}p\right) \\ &= \frac{\partial y_a}{\partial x_i}\frac{\partial y_b}{\partial x_j}J^{-1}\frac{\partial}{\partial y_b}Q_{ij}p \\ &\quad + \frac{\partial y_a}{\partial x_i}Q_{ij}p\frac{\partial}{\partial y_b}\frac{\partial y_b}{\partial x_j}J^{-1} \\ &\quad + Q_{ij}p\frac{\partial y_b}{\partial x_j}J^{-1}\frac{\partial}{\partial y_b}\frac{\partial y_a}{\partial x_i}. \end{aligned} \tag{A18}$$

In the RHS, the second term vanishes because of Eq. (A4). The third term vanishes because it can be rewritten as $Q_{ij}pJ^{-1}\frac{\partial^2 y_a}{\partial x_i \partial x_j}$, which again vanishes because of the antisymmetry of Q_{ij}. Hence, we find that

$$\partial'_b(Q'_{ab}p') = J^{-1}\frac{\partial y_a}{\partial x_i}L_{ij}\partial_j(Q_{ij}p). \tag{A19}$$

Adding up Eqs. (A17) and (A19), and using Eqs. (2.13c) and (2.18h), we obtain Eq. (2.20b):

$$j'_a = J^{-1}(\partial y_a/\partial x_i)j_i. \tag{2.16b}$$

Taking the partial derivative ∂'_a of both sides of Eq. (2.20b) and using Eq. (A4), we easily find Eq. (2.20c):

$$\partial'_a j'_a = J^{-1}\partial_i j_i. \tag{2.16c}$$

At steady state, both sides vanish. Hence, a steady state is transformed into a steady state. Notice that, using Eq. (2.9), we can also rewrite Eq. (A19) as

$$j'^S_a = J^{-1}\frac{\partial y_a}{\partial x_i}j^S_i. \tag{A20}$$

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

[2] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).

[3] K. Jacobs, *Stochastic Processes for Physicists: Understanding Noisy Systems* (Cambridge University Press, Cambridge, UK, 2010).

[4] P. Hänggi and H. Thomas, Stochastic processes: Time evolution, symmetries and linear response, Phys. Rep. **88**, 207 (1982).

[5] N. G. Van Kampen, Itô versus Stratonovich, J. Stat. Phys. **24**, 175 (1981).

[6] Y. L. Klimontovich, Ito, Stratonovich, and kinetic forms of stochastic equations, Phys. A (Amsterdam, Neth.) **163**, 515 (1990).

[7] R. Kupferman, G. A. Pavliotis, and A. M. Stuart, Ito versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise, Phys. Rev. E **70**, 036120 (2004).

[8] G. Germano, M. Politi, E. Scalas, and R. L. Schilling, Stochastic calculus for uncoupled continuous-time random walks, Phys. Rev. E **79**, 066102 (2009).

[9] I. M. Sokolov, Ito, Stratonovich, Hänggi, and all the rest: The thermodynamics of interpretation, Chem. Phys. **375**, 359 (2010).

[10] R. Mannella and P. V. E. McClintock, Itô versus Stratonovich: 30 years later, Fluct. Noise Lett. **11**, 1240010 (2012).

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

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

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

[14] R. L. Stratonovich, *Conditional Markov Processes and their Application to the Theory of Optimal Control* (Elsevier, New York, 1968).

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

[16] R. Graham, Covariant stochastic calculus in the sense of Itô, Phys. Lett. A **109**, 209 (1985).

[17] H. Grabert and M. S. Green, Fluctuations and non-linear irreversible processes, Phys. Rev. A **19**, 1747 (1979).

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

[19] A. W. C. Lau and T. C. Lubensky, State-dependent diffusion: Thermodynamic consistency and its path integral formulation, Phys. Rev. E **76**, 011123 (2007).

[20] F. Langouche, D. Roekaerts, and E. Tirapegui, *Functional Integration and Semiclassical Expansions*, Mathematics and Its Applications (Springer Science & Business Media, Berlin, 2013), Vol. 10.

[21] P. Arnold, Langevin equations with multiplicative noise: Resolution of time discretization ambiguities for equilibrium systems, Phys. Rev. E **61**, 6091 (2000).

[22] P. Arnold, Symmetric path integrals for stochastic equations with multiplicative noise, Phys. Rev. E **61**, 6099 (2000).

[23] H. Calisto and E. Tirapegui, Comment on symmetric path integrals for stochastic equations with multiplicative noise, Phys. Rev. E **65**, 038101 (2002).

[24] M. V. Moreno, Z. G. Arenas, and D. G. Barci, Langevin dynamics for vector variables driven by multiplicative white noise: A functional formalism, Phys. Rev. E **91**, 042103 (2015).

[25] D. G. Barci, Z. G. Arenas, and M. V. Moreno, Path integral approach to nonequilibrium potentials in multiplicative Langevin dynamics, EPL **113**, 10009 (2016).

[26] L. F. Cugliandolo, V. Lecomte, and F. Van Wijland, Building a path-integral calculus, J. Phys. A **52**, 50LT01 (2018).

[27] P. Ao, Potential in stochastic differential equations: Novel construction, J. Phys. A: Math. Gen. **37**, L25 (2004).

[28] C. Kwon, P. Ao, and D. J. Thouless, Structure of stochastic dynamics near fixed points, Proc. Natl. Acad. Sci. USA **102**, 13029 (2005).

[29] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, Path-integral analysis of fluctuation theorems for general Langevin processes, J. Stat. Mech.: Theory Exp. (2006) P08001.

[30] J. Xing, Mapping between dissipative and Hamiltonian systems, J. Phys. A **43**, 375003 (2010).

[31] H. Qian, A decomposition of irreversible diffusion processes without detailed balance, J. Math. Phys. **54**, 053302 (2013).

[32] H. Qian, The zeroth law of thermodynamics and volume-preserving conservative system in equilibrium with stochastic damping, Phys. Lett. A **378**, 609 (2014).

[33] J. Wang, Landscape and flux theory of non-equilibrium dynamical systems with application to biology, Adv. Phys. **64**, 1 (2015).

[34] J. D. Ramshaw, Augmented Langevin approach to fluctuations in non-linear irreversible processes, J. Stat. Phys. **38**, 669 (1985).

[35] M. Polettini, Generally covariant state-dependent diffusion, J. Stat. Mech.: Theory Exp. (2013) P07005.

[36] D. S. Lemons and A. Gythiel, Paul Langevin's 1908 paper "On the theory of Brownian motion" ["sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) **146**, 530 (1908)] Paul langevin's 1908 paper 'on the theory of brownian motion' Am. J. Phys. **65**, 1079 (1997).

[37] G. L. Eyink, J. L. Lebowitz, and H. Spohn, Hydrodynamics and fluctuations outside of local equilibrium: Driven diffusive systems, J. Stat. Phys. **83**, 385 (1996).

[38] M. Itami and S. Shin-ichi, Universal form of stochastic evolution for slow variables in equilibrium systems, J. Stat. Phys. **167**, 46 (2017).

[39] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1996), pp. 63–95.

[40] L. Onsager, Reciprocal relations in irreversible processes. I, Phys. Rev. **37**, 405 (1931).

[41] H. B. G. Casimir, On Onsager's principle of microscopic reversibility, Rev. Mod. Phys. **17**, 343 (1945).