Multi-scale projection operator method and coarse-graining of covariant Fokker-Planck theory

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*

(Received 18 October 2022; accepted 24 February 2023; published 20 March 2023)

A multi-scale projection operator method is developed and applied to study coarse-graining of covariant Fokker-Planck theory and the associated Ito-Langevin dynamics. Explicit expressions for the renormalized kinetic coefficients are obtained. It is also proved that the property of detailed balance is preserved by coarse-graining. These results demonstrate that covariant Fokker-Planck dynamics and covariant Ito-Langevin dynamics emerge naturally as a consequence of coarse-graining of microscopic dynamics. We illustrate our theory using several concrete examples.

DOI: 10.1103/PhysRevResearch.5.013193

I. INTRODUCTION

One of the most important tasks of nonequilibrium statistical physics is to derive macroscopic, irreversible dynamics from microscopic, reversible dynamics, and thereby obtain a first principle understanding of macroscopic irreversible phenomena. For this purpose, one must integrate out a vast number of *fast variables* whose dynamics is fast comparing with experimentally relevant timescales, a procedure usually called *coarse-graining*. As a result, one obtains an effective, low-dimensional dynamic theory for the *slow variables*. The effective slow dynamics is parameterized by a set of macroscopic transport coefficients, which are functions of parameters in the microscopic dynamics. Coarse-graining is also frequently applied to stochastic models, which yields lower-dimensional models that can be more efficiently studied. More generally, classical thermodynamics may be understood as a coarse-graining of statistical mechanics, where all variables, except a few thermodynamic variables, are integrated out. Renormalization group transformation may also be understood as a special case of coarse-graining, where the theories before and after coarse-graining are related to each other by a scaling transformation.

One popular coarse-graining method, known as the *projection operator method*, was developed by Nakajima [1], Mori [2], Zwanzig [3–6], and many others [7–11]. Whilst the method was initially developed for classical systems, it has also been generalized to quantum systems [12–17]. It yields integrodifferential equations for slow variables with colored noises, which are called *generalized Langevin equations*. With a further Markov approximation, generalized Langevin equations can be reduced to usual Langevin equations with white noises. Such an approximation is, however, justified only if there is a wide separation of timescales between fast and slow variables. Otherwise, one has to tackle much more challenging integro-differential equations. The situation was greatly improved in the "time-convolutionless projection operator formalism" [18–20], where the probability density function (pdf) of slow variables obeys differential equation even without timescale separation. Whilst this formalism has been successfully adapted to the quantum systems [21,22], its application to classical systems has been rather limited.

In recent decades, methods of stochastic processes have been successfully applied to study basic problems in nonequilibrium statistical mechanics. Many exact results such as fluctuation theorems and work identities have been discovered, and their relations with time-reversal symmetry have been clarified [23–31]. These exact results, which capture important and universal features of nonequilibrium fluctuations, have been derived using either Hamiltonian dynamics, or stochastic dynamics at different levels, such as Fokker-Planck dynamics, Langevin dynamics, or master equations. Study of these important issues leads to the emergence of a new field called *stochastic thermodynamics* [32–40]. More recently, it was discovered that, in some situations [41–46], entropy production is not invariant under coarse-graining. Whilst it is generally believed that this is due to the nonequilibrium nature of some hidden variables, which are not taken into account in the theory, the precise mechanism is far from clear. Entropy production is a key concept in nonequilibrium statistical physics, yet the above results seem to impose serious challenge on the objectivity of this quantity in nonequilibrium settings. Stochastic thermodynamics is a diversified field, where different models and methods are often applied to the same problem, resulting theories at different levels of coarse-graining. Regardless of many recent studies [35,42,45–60], it is not yet fully understood how thermodynamic quantities

*dmnphy@sjtu.edu.cn

†xxing@sjtu.edu.cn

at different levels of description are related to each other.

The preset paper is the fifth of a sequel dedicated to a unified theory of thermodynamics and stochastic thermodynamics for nonlinear Langevin systems. In the first paper [61], Langevin equation was formulated in a covariant fashion using Ito-calculus. The most salient feature of this theory is that it is parameterized by a generalized potential and a kinetic matrix, both behaving as tensors under nonlinear transform of variables. The detailed balance conditions are also formulated in a covariant fashion. In the second paper [62], stochastic thermodynamics was formulated using covariant Ito-Langevin equation for systems in contact with a single heat bath and driven by conservative forces. All thermodynamic quantities are expressed in terms of the generalized potential, the kinetic matrix, and the probability distribution of system variables. Using concrete examples, it was demonstrated that for systems with multiplicative noises or with curvilinear coordinates, this theory should be used in replacement of the conventional theory of stochastic thermodynamics and stochastic energetics. In the third paper [63], it was shown that the theory remains applicable even if the system is strongly coupled to the bath, and there is no need to revise definitions of thermodynamic variables, as long as the Hamiltonian of mean force is identified as the fluctuating internal energy of the system. In the fourth paper [64], the theory was generalized to systems driven by nonconservative forces. This results in a unified theory of thermodynamics and stochastic thermodynamics for classical nonequilibrium systems, which exhibit nonequilibrium steady states. The theory also provides a statistical mechanical foundation of Glansdorff-Prigogine stability theory of nonequilibrium steady state [65] as well as the steady-state thermodynamics of Sasa and Tasaki [66].

In the present paper and another paper, which will follow immediately [67], we will construct a systematic coarse-graining method for stochastic thermodynamics of nonlinear Langevin systems. Starting from a *fine-grained theory* with both slow and fast variables, we integrate out fast variables and obtain a *coarse-grained theory*, which is an effective theory for slow dynamics. The task of the present paper is to find the generalized potential and the kinetic matrix of the coarse-grained theory in terms of the corresponding quantities of the fine-grained theory. *Accomplishment of this task automatically proves that the covariant Langevin dynamics is the natural formulation of nonlinear-Langevin dynamics.* Whilst the generalized potential of the coarse-grained theory can be obtained from the fine-grained theory using elementary probability theory, calculation of the kinetic matrix of coarse-grained theory is much more nontrivial, and should be deemed an important and challenging problem in nonequilibrium statistical mechanics. In the next paper [67], we will apply these results to derive relations between thermodynamic quantities, in particular, entropy productions, in the fine-grained theory and the coarse-grained theory.

Besides working out the effective dynamics of the slow variables, we would also like to know how the probability distribution of fast variables evolve over time in the middle of nonequilibrium processes. The latter information is much needed when we try to relate thermodynamic quantities at different levels of coarse-graining. For this reason, we find that the convolution-less projection operator theory [18–20] is not directly applicable. Hence we will develop a systematic multi-scale expansion method, which yields at every order simultaneously the pdf of fast variables and the effective FPO of the slow variables. Our multi-scale projection operator method is a systematic generalization of the method developed in Ref. [68]. Multi-scale expansion methods have been applied to study coarse-graining of Fokker-Planck dynamics and Langevin dynamics, as well as some associated questions of stochastic thermodynamics [45,46,52,53], such as overdamped limit, anomalous entropy production, and Martingale structure. These works do not use covariant theories of Langevin dynamics and Fokker-Planck dynamics, and hence do not yield results that we seek, i.e., the relation between the kinetic coefficients of the fine-grained theory and those of the coarse-grained theory.

The remaining of this paper is organized as follows. In Sec. II, we combine projection operator theory with multi-scale analysis to develop a systematic expansion for the effective slow dynamics of a general Fokker-Planck theory, with the ratio ϵ between fast and slow timescales treated as a small parameter. At each order of ϵ, we obtain the probability distribution of fast variables and the renormalized FPO of the slow variables. These results show that, starting from the first order in ϵ, fast variables are driven away from the steady state due to their coupling with the slow variables, and it is this deviation that causes renormalization of the FPO of the effective slow dynamics. In Sec. III, we apply the result to covariant Fokker-Planck theory, and obtain, with certain assumption made on the kinetic matrix, explicit expression for the first-order renormalized kinetic matrix of the effective slow dynamics. We also show that the coarse-grained theory inherits the detailed balance of the fine-grained theory.

In the latter part of this paper, we demonstrate the power of our coarse-graining method using several examples. In Sec. IV, we consider a Hamiltonian system interacting with a large bath, which is also a Hamiltonian system. By integrating out the bath variables, we obtain underdamped Langevin equations, which describe the Brownian motion of the system, with friction coefficients expressed as integrated correlation functions of collision force acting on the Brownian particle. Unlike in the traditional projection operator formalism, there is no need to take further Markov approximation. More importantly, we show that the noises acting on the Brownian particle are generically multiplicative if the bath lack translational symmetry. This happens, e.g., whenever the Brownian particle is near an interface. In the presence of multiplicative noises, our theory automatically yields the covariant form of Ito-Langevin dynamics. In Sec. V, we specialize to the case where the bath consists of a large number of harmonic oscillators. Our coarse-graining method, whilst yielding identical results as the classical projection operator method [5], demonstrates clearly the connection between timescale separation and Markov approximation. In Sec. VI, we study the overdamped limit of Langevin equations for Brownian motion. For the case of spatially varying friction coefficient, our analysis yields the correct form of spurious drift, which was derived previously using more complex method [69]. For the case of spatially independent friction coefficient, we go to third order in the multi-scale expansion, and also work out the short-scale

details of the effective slow dynamics. In Sec. VII, we study the Brownian motion of a rod-like particle in two dimension in the limit where rotational dynamics is much faster than translation dynamics. We show that at the first order in ϵ, the renormalized FPO of translational motion is a fourth-order differential operators. Finally, in Sec. VIII we draw concluding remarks and project future research directions. In the Appendix, we discuss the mathematical structure of covariant Fokker-Planck theory, which are useful for systematic understanding of projection operator formalism.

II. MULTI-SCALE PROJECTION OPERATOR THEORY

A. Fokker-Planck dynamics of the joint system

We consider a set of slow variables $\boldsymbol{x}$ that is interacting with a set of fast variables $\boldsymbol{y}$. The fast variables are in contact with a stochastic environment, which is modeled as a set of independent Wiener noises. The slow variables may or not be in contact with a stochastic environment. The timescales of fast variables $\boldsymbol{y}$ and slow variables $\boldsymbol{x}$ are respectively τ_y and τ_x, which satisfy the assumption of timescale separation,

$$\epsilon = \frac{\tau_y}{\tau_x} \ll 1. \tag{2.1}$$

Note that we are not assuming that the environment is an equilibrium system. Neither do we assume that the fluctuations of fast variables are small.

The dynamics of the joint system consisting of slow and fast variables, collectively denoted as $\mathbf{z} = (\boldsymbol{x}, \boldsymbol{y})$, can be described by Fokker-Planck equation. We consider the following generalized Fokker-Planck type of dynamic equation for the pdf $p(\mathbf{z}, t)$:

$$\partial_t p(\mathbf{z}, t) = \mathscr{L} p(\mathbf{z}, t) = (\mathscr{L}_S(t) + \epsilon^{-1}\mathscr{L}_\mathbf{Y}) p(\mathbf{z}, t), \tag{2.2}$$

where $\mathscr{L}$ is the *Fokker-Planck operator* (FPO) of the joint dynamics. Note here we assume that $\mathscr{L}_\mathbf{Y}$ is a differential operator of the fast variables $\boldsymbol{y}$ only, which may depend parametrically on $\boldsymbol{x}$. $\mathscr{L}_S$, with the subscript S meaning slow, is a differential operator of both $\boldsymbol{x}$ and $\boldsymbol{y}$. The factor ϵ^{-1} in front of $\mathscr{L}_\mathbf{Y}$ is introduced to make the timescale of $\boldsymbol{y}$ of order $O(\epsilon)$, whereas the timescale of $\boldsymbol{x}$ is of order $O(1)$, so that Eq. (2.1) holds [76]. We shall further assume that $\mathscr{L}_\mathbf{Y}$ is independent of time, whilst $\mathscr{L}_S$ may depend on time, due to the variation of some external control parameter. For usual Fokker-Planck dynamics, $\mathscr{L}_S$, $\mathscr{L}_\mathbf{Y}$ are both of second order. Our theory developed in this section, however, is applicable even if these operators are of higher order.

It is useful to study the following reduced Fokker-Planck dynamics of the fast variables:

$$\partial_t p = \epsilon^{-1}\mathscr{L}_\mathbf{Y} p, \tag{2.3}$$

where the slow variables $\boldsymbol{x}$ serve only as fixed parameters. $\epsilon^{-1}\mathscr{L}_\mathbf{Y}$ may be understood as the FPO of the fast variables conditioned on the slow variables. (Recall $\mathscr{L}_\mathbf{Y}$ does not involve derivative over $\boldsymbol{x}$.) Equation (2.3) can be obtained from Eq. (2.2) by neglecting $\mathscr{L}_S$. Hence it is a good approximation of the joint dynamics in the short timescales $\Delta t \ll \tau_x \sim O(1)$, since within this timescales, the slow variables barely change. We shall assume that for any given $\boldsymbol{x}$, Eq. (2.3) has a unique steady state,

$$\mathscr{L}_\mathbf{Y}\, e^{-U_\mathbf{Y}(\boldsymbol{y};\boldsymbol{x})} = 0, \tag{2.4}$$

$$\int_y e^{-U_\mathbf{Y}} = 1. \tag{2.5}$$

This means that $e^{-U_\mathbf{Y}}$ is the unique eigenfunction of $\mathscr{L}_\mathbf{Y}$ with eigenvalue zero. The operator $\mathscr{L}_\mathbf{Y}$ is not invertible.

Fokker-Planck dynamics must preserve the total probability, which means that if we integrate Eqs. (2.3) over $\boldsymbol{y}$, we must get zero for arbitrary p,

$$\int_y \mathscr{L}_\mathbf{Y} = 0. \tag{2.6}$$

Similarly, probability conservation of Eq. (2.2) implies

$$\int_{x,y} \mathscr{L} = 0. \tag{2.7}$$

The main goal of this section is to integrate out the fast variables $\boldsymbol{y}$ and obtain the renormalized slow dynamics. The main results are as follows. In the timescales much longer than τ_y, the marginal pdf $p_\mathbf{X}(\boldsymbol{x}, t)$ for the slow variable $\boldsymbol{x}$ satisfies a renormalized Fokker-Planck equation,

$$\frac{d}{dt} p_\mathbf{X} = \mathscr{L}^R_\mathbf{X}\, p_\mathbf{X}, \tag{2.8}$$

where $\mathscr{L}^R_\mathbf{X}$ is the *renormalized FPO*, expanded as

$$\mathscr{L}^R_\mathbf{X} = \mathscr{L}^{(0)}_\mathbf{X} + \epsilon\mathscr{L}^{(1)}_\mathbf{X} + \epsilon^2\mathscr{L}^{(2)}_\mathbf{X} + \cdots. \tag{2.9a}$$

$$\mathscr{L}^{(0)}_\mathbf{X} \equiv \int_y \mathscr{L}_S\, e^{-U_\mathbf{Y}}. \tag{2.9b}$$

$$\mathscr{L}^{(1)}_\mathbf{X} \equiv -\int_y \mathscr{L}_S \mathscr{L}^{-\perp}_\mathbf{Y} \mathscr{L}_S e^{-U_\mathbf{Y}}. \tag{2.9c}$$

The operator $L^{-\perp}_\mathbf{Y}$ is defined in Eq. (2.24), Sec. II C. The third term in the expansion (2.9a) is presented in Sec. II D 4. The renormalized dynamics of slow variables in the short timescale $t \leqslant \tau_y$ is discussed in Sec. II E.

We note that Eq. (2.2) is not the most general form of Fokker-Planck equation with separation of timescales. It is possible to generalize our theory to the more general and complicated case, where the FPO of the joint dynamics has the following form:

$$\mathscr{L} = \mathscr{L}_{S,0} + \epsilon^{-1/2}\mathscr{L}_{S,-1/2} + \epsilon^{-1}\mathscr{L}_\mathbf{Y}. \tag{2.10}$$

This will be explored in a future publication.

B. Projection operators

Fokker-Planck theory studies the dynamic evolution of probability density functions $p(\boldsymbol{x}, \boldsymbol{y})$. We first define the real Hilbert space $\mathscr{H}_\text{dist}$ spanned by all probability density functions $p(\boldsymbol{x}, \boldsymbol{y})$. An element $\psi(\boldsymbol{x}, \boldsymbol{y})$ in the space $\mathscr{H}_\text{dist}$ will be called a *distribution*. The *slow projection operator* $\mathscr{P}_\text{S}$ is defined as

$$\mathscr{P}_\text{S}\psi(\boldsymbol{x}, \boldsymbol{y}) \equiv e^{-U_\mathbf{Y}(\boldsymbol{y};\boldsymbol{x})} \int_{y'} \psi(\boldsymbol{x}, \boldsymbol{y}'), \tag{2.11}$$

which has two defining properties: (1) $\mathscr{P}_{\rm S}$ preserves the marginal pdf of $\boldsymbol{x}$,

$$\int_y \mathscr{P}_{\rm S}\psi(\boldsymbol{x},\boldsymbol{y}) = \int_y \psi(\boldsymbol{x},\boldsymbol{y}), \tag{2.12}$$

and (2) the fast variables achieve steady state conditioned on the slow variables $\boldsymbol{x}$ in the state $\mathscr{P}_{\rm S}\psi(\boldsymbol{x},\boldsymbol{y})$. It is easy to prove that $\mathscr{P}_{\rm S}$ is involutive, i.e., $\mathscr{P}_{\rm S}\mathscr{P}_{\rm S} = \mathscr{P}_{\rm S}$, and is a Hermitian operator according to the inner-product structure defined in Sec. III B. Hence $\mathscr{P}_{\rm S}$ is a legitimate projection operator. It projects the Hilbert space $\mathscr{H}_{\rm dist}$ onto the *slow subspace* $\mathscr{H}^{\rm S}_{\rm dist}$, where the fast variables achieve steady state conditioned on slow variables. An distribution in $\mathscr{H}^{\rm S}_{\rm dist}$ will be called a *slow distribution*, and can always be written in the form of $f(\boldsymbol{x})\,e^{-U_{\rm Y}(\boldsymbol{y};\boldsymbol{x})}$ [c.f. Eq. (2.11)]. Hence we have

$$\mathscr{H}^{\rm S}_{\rm dist} \equiv \mathscr{P}_{\rm S}\mathscr{H}_{\rm dist} = \{f(\boldsymbol{x})\,e^{-U_{\rm Y}(\boldsymbol{y};\boldsymbol{x})}\}. \tag{2.13}$$

The *fast projection operator* $\mathscr{P}_{\rm F}$ is defined as the orthogonal complement of $\mathscr{P}_{\rm S}$,

$$\mathscr{P}_F \equiv I - \mathscr{P}_{\rm S}, \tag{2.14}$$

where I is the identity operator. When acting on a generic distribution $\psi(\boldsymbol{x},\boldsymbol{y})$, it yields

$$\mathscr{P}_F\psi(\boldsymbol{x},\boldsymbol{y}) = \psi(\boldsymbol{x},\boldsymbol{y}) - e^{-U_{\rm Y}(\boldsymbol{y};\boldsymbol{x})}\int_{y'}\psi(\boldsymbol{x},\boldsymbol{y}'), \tag{2.15}$$

which can be understood as the fluctuations around the slow distribution $\mathscr{P}_S\psi(\boldsymbol{x},\boldsymbol{y})$. Using Eq. (2.5) we easily see

$$\int_y \mathscr{P}_F\psi(\boldsymbol{x},\boldsymbol{y}) = 0.$$

It then follows that $\mathscr{P}_F$ projects the Hilbert space $\mathscr{H}_{\rm dist}$ onto the fast subspace $\mathscr{H}^{\rm F}_{\rm dist}$, defined as,

$$\mathscr{H}^{\rm F}_{\rm dist} \equiv \mathscr{P}_{\rm F}\mathscr{H}_{\rm dist} = \left\{\psi(\boldsymbol{x},\boldsymbol{y})\middle|\int_y \psi(\boldsymbol{x},\boldsymbol{y}) = 0\right\}. \tag{2.16}$$

Any distribution $\psi(\boldsymbol{x},\boldsymbol{y})$ can be decomposed into a slow component $\psi_{\rm S}(\boldsymbol{x},\boldsymbol{y})$ and a fast component $\psi_{\rm F}(\boldsymbol{x},\boldsymbol{y})$,

$$\begin{aligned}\psi(\boldsymbol{x},\boldsymbol{y}) &= \mathscr{P}_{\rm S}\psi(\boldsymbol{x},\boldsymbol{y}) + \mathscr{P}_{\rm F}\psi(\boldsymbol{x},\boldsymbol{y})\\ &= \psi_{\rm S}(\boldsymbol{x},\boldsymbol{y}) + \psi_{\rm F}(\boldsymbol{x},\boldsymbol{y}).\end{aligned} \tag{2.17}$$

The Hilbert space of distribution $\mathscr{H}_{\rm dist}$ is decomposed as

$$\mathscr{H}_{\rm dist} = \mathscr{H}^{\rm S}_{\rm dist} \oplus \mathscr{H}^{\rm F}_{\rm dist}. \tag{2.18}$$

Below we summarize the useful properties of $\mathscr{P}_{\rm S}$, $\mathscr{P}_{\rm F}$:

$$\mathscr{P}_{\rm S}^2 = \mathscr{P}_{\rm S}, \quad \mathscr{P}_{\rm F}^2 = \mathscr{P}_{\rm F}, \tag{2.19a}$$

$$\mathscr{P}_{\rm S} + \mathscr{P}_{\rm F} = I, \tag{2.19b}$$

$$\mathscr{P}_{\rm S}^\dagger = \mathscr{P}_{\rm S}, \quad \mathscr{P}_{\rm F}^\dagger = \mathscr{P}_{\rm F}, \tag{2.19c}$$

$$\mathscr{P}_{\rm S}\mathscr{P}_{\rm F} = \mathscr{P}_{\rm F}\mathscr{P}_{\rm S} = 0. \tag{2.19d}$$

For any function $f(\boldsymbol{x})$ of slow variables, we have

$$\mathscr{P}_{\rm S}(f(\boldsymbol{x})\,e^{-U_{\rm Y}}) = f(\boldsymbol{x})\,e^{-U_{\rm Y}}, \tag{2.20a}$$

$$\mathscr{P}_{\rm F}(f(\boldsymbol{x})\,e^{-U_{\rm Y}}) = 0. \tag{2.20b}$$

C. Solubility condition

Because $\mathscr{L}_{\mathbf{Y}}$ does not contain $\boldsymbol{x}$ derivative, it always returns zero when acting on a slow distribution,

$$\mathscr{L}_{\mathbf{Y}} f(\boldsymbol{x})e^{-U_{\rm Y}} = f(\boldsymbol{x})\mathscr{L}_{\mathbf{Y}}e^{-U_{\rm Y}} = 0. \tag{2.21}$$

Since the function $f(\boldsymbol{x})$ is arbitrary, we have the operator identity $\mathscr{L}_{\mathbf{Y}}\mathscr{P}_S = 0$. Reciprocally, if $\mathscr{L}_{\mathbf{Y}}\psi(\boldsymbol{x},\boldsymbol{y}) = 0$, then $\psi(\boldsymbol{x},\boldsymbol{y})$ must be a slow distribution, since by assumption $\mathscr{L}_{\mathbf{Y}}$ has no steady state other than $e^{-U_{\rm Y}}$. The slow subspace $\mathscr{H}^S_{\rm dist}$ is therefore the kernel of the operator $\mathscr{L}_{\mathbf{Y}}$. Additionally, combination of Eqs. (2.11) and (2.6) implies the operator identity $\mathscr{P}_S\mathscr{L}_{\mathbf{Y}} = 0$. Summarizing we find

$$\mathscr{P}_S\mathscr{L}_{\mathbf{Y}} = \mathscr{L}_{\mathbf{Y}}\mathscr{P}_S = 0, \tag{2.22}$$

Using Eqs. (2.19b) and (2.22), we have

$$\begin{aligned}\mathscr{L}_{\mathbf{Y}} &= (\mathscr{P}_S + \mathscr{P}_F)\mathscr{L}_{\mathbf{Y}}(\mathscr{P}_S + \mathscr{P}_F)\\ &= \mathscr{P}_F\mathscr{L}_{\mathbf{Y}}\mathscr{P}_F.\end{aligned} \tag{2.23}$$

The operator $\mathscr{L}_{\mathbf{Y}}$ is not invertible, since it has a zero eigenvalue with eigenfunction $e^{-U_{\rm Y}}$. But we can define an operator $\mathscr{L}_{\mathbf{Y}}^{-\perp}$, which behaves as the inverse of $\mathscr{L}_{\mathbf{Y}}$ when confined in the fast subspace $\mathscr{H}^{\rm F}_{\rm dist}$,

$$\mathscr{L}_{\mathbf{Y}}^{-\perp} \equiv \mathscr{P}_F\mathscr{L}_{\mathbf{Y}}^{-1}\mathscr{P}_F, \tag{2.24}$$

$$\mathscr{L}_{\mathbf{Y}}^{-\perp}L_{\mathbf{Y}} = L_{\mathbf{Y}}\mathscr{L}_{\mathbf{Y}}^{-\perp} = \mathscr{P}_F. \tag{2.25}$$

When acting on a distribution $\psi = \psi_{\rm S} + \psi_{\rm F}$, $\mathscr{L}_{\mathbf{Y}}^{-\perp}$ returns a fast distribution,

$$\mathscr{L}_{\mathbf{Y}}^{-\perp}(\psi_{\rm S} + \psi_{\rm F}) = \mathscr{L}_{\mathbf{Y}}^{-\perp}\psi_{\rm F} \in \mathscr{H}^F_{\rm dist}. \tag{2.26}$$

Note that by definition (2.24) $\mathscr{L}_{\mathbf{Y}}^{-\perp}\psi_{\rm S} = 0$.

When carrying out multi-scale analysis, we will repeatedly solve the following linear equation:

$$\mathscr{L}_{\mathbf{Y}}p(\boldsymbol{x},\boldsymbol{y}) = \psi(\boldsymbol{x},\boldsymbol{y}). \tag{2.27}$$

Let $p_S = \mathscr{P}_{\rm S}p$ and $p_F = \mathscr{P}_{\rm F}p$ be the slow component and fast component of p. Since $\mathscr{L}_{\mathbf{Y}}p_S = 0$, we see that p_S cannot be determined from Eq. (2.27): it is arbitrary. Hence Eq. (2.27) can be rewritten into

$$\mathscr{L}_{\mathbf{Y}}p_F = \psi. \tag{2.28}$$

We apply $\mathscr{P}_S$ to Eq. (2.28) from the left. The left-hand side (l.h.s.) vanishes because of Eq. (2.22). Hence we find

$$0 = \mathscr{P}_S\psi = \psi_S, \tag{2.29a}$$

or equivalently,

$$\int_y \psi(\boldsymbol{x},\boldsymbol{y}) = 0. \tag{2.29b}$$

This imposes a necessary condition in order for Eq. (2.27) to be solvable. Equations (2.29) shall be called the *solubility condition*. Now taking the fast component of Eq. (2.28), we find

$$\mathscr{L}_{\mathbf{Y}}p_F = \psi_{\rm F}, \tag{2.30}$$

whose solution determines the fast component of p,

$$p_F = \mathscr{L}_{\mathbf{Y}}^{-\perp}\psi_F. \tag{2.31}$$

With the solubility condition (2.29) satisfied, the general solution to Eq. (2.27) is

$$p(\mathbf{z}) = \phi_S + \mathscr{L}_{\mathbf{Y}}^{-\perp}\psi_F, \tag{2.32}$$

where ϕ_S is an arbitrary slow distribution. Two terms in the right-hand side (r.h.s.) of Eq. (2.32) are, respectively, the homogeneous solution and the particular solution of Eq. (2.27).

In summary, the slow component of Eq. (2.27) gives the solubility condition, whilst the fast component of the equation determines the fast component of p, finally the slow component of p remains arbitrary.

D. Multi-scale expansion of effective slow dynamics

We aim to derive the effective dynamics of the slow variables from the Fokker-Planck dynamics for the joint system, Eq. (2.2). We shall first carry out a multi-scale expansion of the solution to Eq. (2.2), and obtain a hierarchy of equations for each terms in the expansion. We then project these equations into the slow subspace and obtain effective Fokker-Planck equations for the slow components of distributions at each order.

Pedagogical discussion of multi-scale analysis can be found in the classical monograph on asymptotic analysis by Bender and Orszag [70]. Our method is generalization of the method developed in Ref. [68], which studies the overdamped limit of harmonic oscillator with friction and noise. Multi-scale techniques have also been applied to more general (but also more abstract) coarse-graining problems, both by mathematicians [52,71–73] and by physicists [45,46,53].

We introduce a set of timescales $t_1 = \epsilon t, t_2 = \epsilon^2 t, \cdots$, and treat the pdf as a function of $\boldsymbol{x}, \boldsymbol{y}$ and all these timescales. We consider the following multi-scale expansion:

$$\begin{aligned} p(\boldsymbol{x},\boldsymbol{y};t,t_1,t_2,\cdots) = {} & p^{(0)}(\boldsymbol{x},\boldsymbol{y};t,t_1,t_2,\cdots) \\ & + \epsilon\, p^{(1)}(\boldsymbol{x},\boldsymbol{y};t,t_1,t_2,\cdots) \\ & + \epsilon^2\, p^{(2)}(\boldsymbol{x},\boldsymbol{y};t,t_1,t_2,\cdots) \\ & + \cdots . \end{aligned} \tag{2.33}$$

The dependence of $p^{(l)}(\mathbf{z};t,t_1,t_2,\cdots)$ on $t,t_1,t_2,\cdots$ will be constructed such that the expansion exists, i.e., the functions $p^{(k)}(\boldsymbol{x},\boldsymbol{y};t,t_1,t_2,\cdots)$ can be found at each order. When writing down Eq. (2.33), we are already assuming that the joint pdf does not depend on the fast-time variable $t_{-1} = t/\epsilon$. This may sound strange since the fast dynamics as determined by Eq. (2.3) has a characteristic timescale of $O(\epsilon)$. Indeed, assuming the dynamics starting from the initial time t_0 with generic initial conditions, if we can resolve the dynamical details of for $p(\boldsymbol{x},\boldsymbol{y},t)$ in the short period $t_0 < t < t_0 + O(\epsilon)$, we would see that $p(\boldsymbol{x},\boldsymbol{y};t)$ also depends on t_{-1}. In longer time, however, this fast dynamics quickly decay, and the long-scale asymptotic dynamics of $p(\boldsymbol{x},\boldsymbol{y};t)$ only depends on the slower-time variables $t_0 = t,\ t_1 = \epsilon t, t_2 = \epsilon^2 t, \cdots$. Qualitatively speaking, in longer timescales, fast variables are slaved by the slow variables, so that we do not need to explicitly keep track of the fast variables. This point will be explained in greater details in Sec. II E.

The time derivative in l.h.s. of Eq. (2.2) now should be understood as the *full-time derivative*, which can be expressed in terms of partial derivatives with respect to $t,t_1,t_2,\cdots$ via the chain rule

$$\begin{aligned} \frac{\partial}{\partial t} \to \frac{d}{dt} &= \frac{\partial}{\partial t} + \frac{dt_1}{dt}\frac{\partial}{\partial t_1} + \frac{dt_2}{dt}\frac{\partial}{\partial t_2} + \cdots \\ &= \partial_t + \epsilon\,\partial_{t_1} + \epsilon^2\partial_{t_2} + \cdots . \end{aligned} \tag{2.34}$$

Substituting Eqs. (2.33) and (2.34) into Eq. (2.2), expanding both sides in terms of ϵ, and comparing coefficients at each order of ϵ, we obtain

$$\mathscr{L}_{\mathbf{Y}} p^{(0)} = 0, \tag{2.35a}$$

$$\mathscr{L}_{\mathbf{Y}} p^{(1)} = \partial_t p^{(0)} - \mathscr{L}_S p^{(0)}, \tag{2.35b}$$

$$\mathscr{L}_{\mathbf{Y}} p^{(2)} = \partial_{t_1} p^{(0)} + \partial_t\, p^{(1)} - \mathscr{L}_S p^{(1)}, \tag{2.35c}$$

$$\cdots\cdots = \cdots\cdots\cdots .$$

These equations have the same structure as Eq. (2.27). We will solve them order by order using the method developed in Sec. II C. As demonstrated in Sec. II C, the slow component of every $p^{(k)}$ remains arbitrary. For convenience, we impose the convention for all $k \geqslant 1$, $p^{(k)}$ has no slow component,

$$\mathscr{P}_S p^{(k)} = p_S^{(k)} = 0, \quad p^{(k)} = p_F^{(k)}. \tag{2.36}$$

Equivalently we have

$$\int_{\boldsymbol{y}} p^{(k)} = 0, \quad \forall k \geqslant 1. \tag{2.37}$$

This means that the slow component of p is exclusively contained in $p^{(0)}$. As we will see, this convention will greatly simplify the multi-scale analysis. The marginal pdf of $\boldsymbol{x}$ can be obtained from Eq. (2.33) by integrating out $\boldsymbol{y}$, or projecting out the slow component,

$$p_{\mathbf{X}}(\boldsymbol{x};t,\cdots) = \sum_{k=0}^{\infty} \epsilon^k \int_{\boldsymbol{y}} p^{(k)}(\boldsymbol{x},\boldsymbol{y};t,\cdots). \tag{2.38}$$

Only the first term in the r.h.s. survives since $p^{(k)}$ for $k > 0$ have no slow component. Hence we have

$$p_{\mathbf{X}}(\boldsymbol{x};t,\cdots) = \int_{\boldsymbol{y}} p^{(0)} = p_S^{(0)}\, e^{U_{\mathbf{Y}}}, \tag{2.39}$$

$$p_S^{(0)}(\boldsymbol{x};t,\cdots) = p_{\mathbf{X}}(\boldsymbol{x};t,\cdots)\, e^{-U_{\mathbf{Y}}}. \tag{2.40}$$

1. Zeroth order

At the zeroth order, the r.h.s. of Eq. (2.35a) vanishes, hence the solubility condition is trivial. Hence the general solution to Eq. (2.35a) is

$$p^{(0)} = p_S^{(0)} = p_{\mathbf{X}}\, e^{-U_{\mathbf{Y}}}, \tag{2.41}$$

where $p_{\mathbf{X}} = p_{\mathbf{X}}(\boldsymbol{x};t,t_1,\cdots)$ remains an arbitrary function of $\boldsymbol{x},t,t_1$. Its dependence on various time-scales will be determined as we go to higher orders in perturbation analysis.

2. First order

The solubility condition of Eq. (2.35b) is

$$\begin{aligned} 0 &= \int_{\boldsymbol{y}} (\partial_t p^{(0)} - \mathscr{L}_S p^{(0)}) \\ &= \int_{\boldsymbol{y}} \partial_t p_{\mathbf{X}} e^{-U_{\mathbf{Y}}} - \int_{\boldsymbol{y}} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}. \end{aligned} \tag{2.42}$$

where used was Eq. (2.41). The first term in r.h.s. of Eq. (2.42) can be simplified as

$$\int_y \partial_t p_{\mathbf{X}}\, e^{-U_{\mathbf{Y}}} = \partial_t p_{\mathbf{X}} \int_y e^{-U_{\mathbf{Y}}} = \partial_t p_{\mathbf{X}}, \tag{2.43}$$

where in the last equality we have used Eq. (2.5). Let us define an operator $\mathscr{L}_{\mathbf{X}}^{(0)}$ of slow variables,

$$\mathscr{L}_{\mathbf{X}}^{(0)} \equiv \int_y \mathscr{L}_S\, e^{-U_{\mathbf{Y}}}. \tag{2.44}$$

It acts only on $\boldsymbol{x}$ but not on $\boldsymbol{y}$, since $\boldsymbol{y}$ are already integrated out. We can then rewrite Eq. (2.42) into

$$\partial_t\, p_{\mathbf{X}} = \mathscr{L}_{\mathbf{X}}^{(0)} p_{\mathbf{X}}, \tag{2.45}$$

which determines the t dependence of $p_{\mathbf{X}}$.

Now take the fast component of Eq. (2.35b). Using Eq. (2.41) and noticing that $e^{-U_{\mathbf{Y}}}$ is independent of time, we see that the fast component of $\partial_t p^{(0)}$ vanishes, hence the fast component of Eq. (2.35b) becomes

$$\mathscr{L}_{\mathbf{Y}} p_F^{(1)} = -\mathscr{P}_F \mathscr{L}_S p_{\mathbf{X}} e^{-U_{\mathbf{Y}}}, \tag{2.46}$$

whose solution is

$$\begin{aligned} p^{(1)} = p_F^{(1)} &= \mathscr{L}_{\mathbf{Y}}^{-\perp}(-\mathscr{P}_F \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}), \\ &= -\mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}, \end{aligned} \tag{2.47}$$

where we have used the convention that the slow component of $p^{(1)}$ vanishes. Note that in the last step we have used Eq. (2.24) as well as $\mathscr{P}_F \mathscr{P}_F = \mathscr{P}_F$. Hence $p^{(1)}$ is fully determined by $p_{\mathbf{X}}$ at the same instant.

3. Second order

Using Eq. (2.47) and $p_S^{(1)} = 0$, we see

$$\begin{aligned} \mathscr{P}_S \partial_t\, p^{(1)} &= -\mathscr{P}_S \partial_t \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} \\ &= -\mathscr{P}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \partial_t \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} \\ &= 0, \end{aligned} \tag{2.48}$$

where in the last step we used $\mathscr{P}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} = 0$. Further using Eqs. (2.41) and (2.47), the solubility condition of Eq. (2.35c) becomes

$$0 = \int_y \partial_{t_1} p_{\mathbf{X}}\, e^{-U_{\mathbf{Y}}} + \int_y \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}. \tag{2.49}$$

Defining another operator

$$\mathscr{L}_{\mathbf{X}}^{(1)} \equiv -\int_y \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S e^{-U_{\mathbf{Y}}}, \tag{2.50}$$

we can rewrite Eq. (2.49) into

$$\partial_{t_1}\, p_{\mathbf{X}} = \mathscr{L}_{\mathbf{X}}^{(1)} p_{\mathbf{X}}, \tag{2.51}$$

which determines the t_1 dependence of $p_{\mathbf{X}}$.

Let us now take the fast component of Eq. (2.35c). Again we can easily see that $\partial_{t_1} p^{(0)}$ has no fast component, and hence we find

$$\mathscr{L}_{\mathbf{Y}} p_F^{(2)} = \partial_t p_F^{(1)} - \mathscr{P}_F \mathscr{L}_S p_F^{(1)}, \tag{2.52}$$

whose solution is

$$\begin{aligned} p_F^{(2)} &= \mathscr{L}_{\mathbf{Y}}^{-\perp}\big[\partial_t\big(-\mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}\big) \\ &\quad + \mathscr{P}_F \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}\big]. \\ &= \mathscr{L}_{\mathbf{Y}}^{-\perp}\big[-\mathscr{L}_{\mathbf{Y}}^{-\perp}(\partial_t \mathscr{L}_S)\, e^{-U_{\mathbf{Y}}} - \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S e^{-U_{\mathbf{Y}}} \mathscr{L}_{\mathbf{X}}^{(0)} \\ &\quad + \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}}\big] p_{\mathbf{X}}, \end{aligned} \tag{2.53}$$

which fully determines $p_F^{(2)}(t)$ in terms of $p_{\mathbf{X}}(t)$.

Inspecting the structure of Eq. (2.51) we can now understand why it is necessary to introduce the slow timescales $t_1, t_2, \cdots$ in the multi-scale expansion Eq. (2.33). Without these slow timescales, Eq. (2.51) would lead to $0 = \mathscr{L}_{\mathbf{X}}^{(1)} p_{\mathbf{X}}$, an inconsistency.

4. Effective slow dynamics

This above process can be continued indefinitely. For example, the solubility condition at the third order is

$$\partial_{t_2}\, p_{\mathbf{X}} = \mathscr{L}_{\mathbf{X}}^{(2)} p_{\mathbf{X}}, \tag{2.54}$$

where the operator $\mathscr{L}_{\mathbf{X}}^{(2)}$ is given by

$$\begin{aligned} \mathscr{L}_{\mathbf{X}}^{(2)} \equiv &\int_y \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S e^{-U_{\mathbf{Y}}} \\ &- \int_y \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S \mathscr{L}_{\mathbf{X}}^{(0)} e^{-U_{\mathbf{Y}}} \\ &- \int_y \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_{\mathbf{Y}}^{-\perp} (\partial_t \mathscr{L}_S) e^{-U_{\mathbf{Y}}}, \end{aligned} \tag{2.55}$$

which determines the t_2 dependence of the marginal pdf $p_{\mathbf{X}}$. Note that the time derivative of $\mathscr{L}_S$ appears in the last term. Taking the fast component of the third-order equation, we find that the fast component $p_F^{(3)}$ is completely determined by $p_{\mathbf{X}}$ at the same instant. The results become more and more complex as one go to higher and higher orders, so we stop here.

The sequence of equations (2.45), (2.51), and (2.54), as well as their higher analogues, are mathematically equivalent to the following renormalized Fokker-Planck equation of the slow variable pdf $p_{\mathbf{X}}$:

$$\frac{d}{dt} p_{\mathbf{X}} = \mathscr{L}_{\mathbf{X}}^{R}\, p_{\mathbf{X}}, \tag{2.56}$$

where d/dt is the full time derivative defined in Eq. (2.34), and $\mathscr{L}_{\mathbf{X}}^{R}$ is the *renormalized FPO*, expanded as

$$\begin{aligned} \mathscr{L}_{\mathbf{X}}^{R} &= \mathscr{L}_{\mathbf{X}}^{(0)} + \epsilon \mathscr{L}_{\mathbf{X}}^{(1)} + \epsilon^2 \mathscr{L}_{\mathbf{X}}^{(2)} + \cdots. \\ &= \int_y \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} - \epsilon \int_y \mathscr{L}_S \mathscr{L}_{\mathbf{Y}}^{-\perp} \mathscr{L}_S e^{-U_{\mathbf{Y}}} + \cdots. \end{aligned} \tag{2.57}$$

To see this, we insert Eqs. (2.34) and (2.57) into Eq. (2.56), expand both sides of the equation in terms of ϵ, and compare coefficients of each power, we obtain Eqs. (2.45), (2.51), and (2.54), as well as all higher-order conditions. Hence by carrying our the multi-scale expansion, we obtain the renormalized FPO of the effective slow dynamics. These results have already been summarized in the end of Sec. II A.

$\mathscr{L}_{\mathbf{X}}^{(1)}$ and higher-order terms in Eq. (2.57) characterize renormalization of the effective slow dynamics due to the deviation of fast variables from conditional equilibrium/steady

state. As we have seen above, in the limit $\epsilon \to 0$, the fast component of p vanishes, which means that the fast variables remain strictly in conditional equilibrium/steady state. The FPO of the effective slow dynamics is then precisely given by $\mathscr{L}_{\mathbf{X}}^{(0)}$, the zeroth-order term in Eq. (2.57).

In usual Fokker-Planck dynamics, $\mathscr{L}_S$ is second-order differential operator in slow variables $\boldsymbol{x}$. As one can see from Eqs. (2.9) and (2.55), higher-order terms in the expansion (2.9a) involve more and more powers of $\mathscr{L}_S$. As a consequence, if we truncate the expansion at a finite order, we may get differential operators in $\boldsymbol{x}$, which are of order higher than two. Then according to the Pawula theorem [74], the pdf $p_{\mathbf{X}}$ is no longer necessarily positive definite. Note, however, the violation of positivity is controlled by small parameter $\epsilon = \tau_y/\tau_x$. As long as ϵ is sufficiently small, the truncated theory may still provides accurate approximation to the renormalized dynamics of slow variables. An example is discussed in Sec. VII where the first-order renormalized slow dynamics involves fourth-order differential operators in slow variables.

E. Including short-timescale details

To take care of the fast dynamics happening in short timescale, we introduce a fast timescale,

$$t_{-1} \equiv \epsilon^{-1} t. \tag{2.58}$$

There is no need to introduce even faster timescales $\epsilon^{-2}t$ etc. The expansion of the pdf, Eq. (2.33) now becomes

$$\begin{aligned} p(\boldsymbol{x},\boldsymbol{y};t,t_1,t_2,\cdots) = {} & p^{(0)}(\boldsymbol{x},\boldsymbol{y};t_{-1},t,t_1,\cdots) \\ & + \epsilon\, p^{(1)}(\boldsymbol{x},\boldsymbol{y};t_{-1},t,t_1,\cdots) \\ & + \epsilon^2\, p^{(2)}(\boldsymbol{x},\boldsymbol{y};t_{-1},t,t_1,\cdots) \\ & + \cdots . \end{aligned} \tag{2.59}$$

The full-time derivative Eq. (2.34) should be replaced by

$$\frac{d}{dt} = \epsilon^{-1}\partial_{t_{-1}} + \partial_t + \epsilon\partial_{t_1} + \epsilon^2\partial_{t_2} + \cdots. \tag{2.60}$$

Inserting these back into Eq. (2.2), expanding in terms of ϵ, and comparing coefficients order by order, we obtain

$$-\partial_{t_{-1}} p^{(0)} + \mathscr{L}_{\mathbf{Y}} p^{(0)} = 0, \tag{2.61a}$$

$$-\partial_{t_{-1}} p^{(1)} + \mathscr{L}_{\mathbf{Y}} p^{(1)} = \partial_t p^{(0)} - \mathscr{L}_S p^{(0)}, \tag{2.61b}$$

$$-\partial_{t_{-1}} p^{(2)} + \mathscr{L}_{\mathbf{Y}} p^{(2)} = \partial_{t_1} p^{(0)} + \partial_t p^{(1)} - \mathscr{L}_S p^{(1)}. \tag{2.61c}$$

The main difference with Eqs. (2.35) is the appearance of $\partial_{t_{-1}}$ term on the l.h.s., which brings additional details in a very short time window after the initial time.

Let us first try to solve a generic equation

$$-\partial_{t_{-1}} p + \mathscr{L}_{\mathbf{Y}} p = \psi, \tag{2.62}$$

where neither $\mathscr{L}_{\mathbf{Y}}$ nor ψ depends on t_{-1}. First of all, taking the slow component of this equation, we obtain

$$-\partial_{t_{-1}} p_S = \psi_S. \tag{2.63}$$

If ψ_S is nonzero, we would obtain

$$p_S = -\psi_S\, t_{-1} + p_S^0, \tag{2.64}$$

which increases linearly with t_{-1}. This is a secular term, which will eventually make the expansion breaking down. Hence we must impose the solubility condition

$$\psi_S = 0, \tag{2.65}$$

i.e., the r.h.s. of Eq. (2.62) must be a fast distribution in order for the solution containing no secular term. With this condition satisfied, the general solution to Eq. (2.62) is then given by

$$\begin{aligned} p(t_{-1}) = {} & \phi(\boldsymbol{x}) e^{-U_{\mathbf{Y}}} + e^{t_{-1}\mathscr{L}_{\mathbf{Y}}}\phi_F^0 \\ & - \int_0^{t_{-1}} e^{(t_{-1}-t'_{-1})\mathscr{L}_{\mathbf{Y}}}\psi_F\, dt'_{-1}, \end{aligned} \tag{2.66}$$

where ϕ_F^0 is a fast distribution. The first two terms are the homogeneous solution, whereas the last term is the particular solution. According to our assumption, ψ is independent of t_{-1}, hence can be pulled out of the integral in the r.h.s. of Eq. (2.66). Further noticing that $\mathscr{L}_{\mathbf{Y}}$ is also independent of t_{-1}, we can make variable transformation: $t_{-1} - t'_{-1} \to s$, and rewrite Eq. (2.66) as

$$p(t_{-1}) = \phi(\boldsymbol{x}) e^{-U_{\mathbf{Y}}} + e^{t_{-1}\mathscr{L}_{\mathbf{Y}}}\phi_F^0 - \left[\int_0^{t_{-1}} e^{s\mathscr{L}_{\mathbf{Y}}} ds\right]\psi_F. \tag{2.67}$$

Let us define a new operator

$$\hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}(t_{-1}) \equiv -\mathscr{P}_F\left[\int_0^{t_{-1}} e^{s\mathscr{L}_{\mathbf{Y}}} ds\right]\mathscr{P}_F, \tag{2.68}$$

which has the following limiting properties:

$$\lim_{t_{-1}\to 0} \hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}(t_{-1}) = 0, \tag{2.69}$$

$$\lim_{t_{-1}\to \infty} \hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}(t_{-1}) = \mathscr{L}_{\mathbf{Y}}^{-\perp}. \tag{2.70}$$

The solution Eq. (2.67) then can be rewritten as

$$p(t_{-1}) = \phi(\boldsymbol{x}) e^{-U_{\mathbf{Y}}} + e^{t_{-1}\mathscr{L}_{\mathbf{Y}}}\phi_F^0 + \hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}(t_{-1})\psi_F. \tag{2.71}$$

Equation (2.71) satisfies the initial condition

$$p(t_{-1} = 0) = \phi(\boldsymbol{x}) e^{-U_{\mathbf{Y}}} + \phi_F^0. \tag{2.72}$$

Hence ϕ_F^0 is the fast component of $p(t_{-1})$ at the initial time. Since $\mathscr{L}_{\mathbf{Y}}$ is a negative operator for fast distributions, Eq. (2.71) says that the initial fast component ϕ_F^0 decays rapidly, with characteristic timescale t_{-1}. As $t_{-1} \to +\infty$, Eq. (2.71) approaches to the following asymptotic behavior:

$$p(t_{-1}) \to \phi(\boldsymbol{x}) e^{-U_{\mathbf{Y}}} + \mathscr{L}_{\mathbf{Y}}^{-\perp}\psi_F, \tag{2.73}$$

where the neglected terms are exponentially small in t_{-1}. Equation (2.73) has the same form as Eq. (2.32).

Let us now try to solve Eqs. (2.61). As in Sec. II D, we adopt the convention that for $k \geqslant 1$, i.e., $p_S^{(k)} = 0$, i.e., the slow part of p is completely allocated to the zeroth order. Furthermore, we shall assume that at the initial time, the pdf as given by Eq. (2.59) only has slow component. These impose the following initial conditions on $p^{(k)}$:

$$p^{(0)}(t_{-1} = t = \cdots = 0) = p_X(0) e^{-U_Y}. \tag{2.74}$$

$$p^{(k)}(t_{-1} = t = \cdots = 0) = 0, \quad k \geqslant 1. \tag{2.75}$$

Let us now inspect Eq. (2.61a). The solubility condition is trivially satisfied. Its solution is given by the first two terms of

the r.h.s. of Eq. (2.71). Because of the initial condition (2.74), there can be no fast component of $p^{(0)}$ at the initial time, hence we have

$$p^{(0)} = p_S^{(0)} = p_X(t, t_1, ...)e^{-U_Y}, \tag{2.76}$$

i.e., there is no fast component of $p^{(0)}$ in any time.

Let us now move on to Eq. (2.61b). The solubility condition is

$$\mathscr{P}_S(\partial_t p^{(0)} - \mathscr{L}_S p^{(0)}) = 0. \tag{2.77}$$

Applying Eq. (2.76), we obtain

$$\partial_t p_{\mathbf{X}} = \int_{\mathbf{y}} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} \equiv \mathscr{L}_{\mathbf{X}}^{(0)} p_{\mathbf{X}}, \tag{2.78}$$

which is identical to Eq. (2.45).

Taking the fast component of Eq. (2.61b), we obtain

$$-\partial_{t_{-1}} p_F^{(1)} + \mathscr{L}_{\mathbf{Y}} p_F^{(1)} = -\mathscr{P}_F \mathscr{L}_S p^{(0)}, \tag{2.79}$$

which as the same form as Eq. (2.62), with the r.h.s. independent of t_{-1}. Using the initial condition (2.75) and convention that $p^{(1)}$ has no slow component, we fix the solution,

$$p_F^{(1)} = -\hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}(t_{-1})\mathscr{P}_F \mathscr{L}_S p^{(0)}. \tag{2.80}$$

Now let us go to Eq. (2.61c). Since $p^{(1)}$ has no slow component, the solubility condition is

$$\partial_{t_1} p_S^{(0)} = \mathscr{P}_S \mathscr{L}_S p^{(1)} = \mathscr{P}_S \mathscr{L}_S p_F^{(1)}. \tag{2.81}$$

Inserting Eqs. (2.76) and (2.80) into the preceding equation, we obtain

$$\partial_{t_1} p_X = \int_{\mathbf{y}} \mathscr{P}_S \mathscr{L}_S p_F^{(1)} = \mathscr{L}_{\mathbf{X}}^{(1)} p_X, \tag{2.82}$$

$$\mathscr{L}_{\mathbf{X}}^{(1)} \equiv -\int_{\mathbf{y}} \mathscr{L}_S \hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}(t_{-1}) \mathscr{L}_S e^{-U_{\mathbf{Y}}}, \tag{2.83}$$

where t_{-1} is related to t via Eq. (2.58). Note that Eq. (2.82) is formally identical to Eq. (2.51). The operator defined in Eq. (2.83) is almost the same as Eq. (2.50), with the only difference that $\hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}(t/\epsilon)$ replacing $\mathscr{L}_{\mathbf{Y}}^{-\perp}$.

It was pointed out in end of Sec. II D that the physical significance of $\mathscr{L}_{\mathbf{X}}^{(1)}$ and higher-order terms is the renormalization of the effective slow dynamics due to the deviation of fast variables from conditional equilibrium/steady state. Since the system starts from conditional equilibrium/steady state, it takes time for the fast variables develop these deviations. The relevant timescale is precisely the that of the fast dynamics. This is the physical origin of t_{-1} dependence of $\mathscr{L}_{\mathbf{X}}^{(1)}$ we discussed here. In Sec. VI B, we will illustrate this point using a concrete example.

F. Comparisons with previous theories

In the traditional projection operator theory [5], one first formally solves Eq. (2.2) for the fast component of $p(\boldsymbol{x}, \boldsymbol{y})$ in terms of the slow components, and then plugging it back into the slow component of the same equation. This results in an integral (in time) equation for the effective slow dynamics. Additional Markov approximation is needed in order to transform this integral equation into a differential equation. The fact that $p_{\mathbf{X}}$ obeys an integral equation instead of a differential equation means that the effective slow dynamics is non-Markovian, or hidden Markov. This is generally understood as an inevitable cost for integrating out dynamic variables from a Markov process. The associated Langevin equation is characterized by colored noises with correlation functions delocalized in time. These equations are commonly known as *generalized Langevin equations* and have been the subjects of intensive studies in the past decades [3–5,11,15,17,40,73]. Since the fast variables are invisible in the effective slow dynamics, distribution of $\boldsymbol{x}$ at the present time is no longer sufficient to fully determine the the evolution of $p_{\mathbf{X}}$ in all future time.

By strong contrast, our multi-scale projection analysis yields Eq. (2.56), which is a first-order *differential equation* (with respect to time) for the marginal pdf of slow variables. This seems to suggest that the effective slow dynamics is Markovian. Note that as we go to higher and higher order in the multi-scale expansion (2.57), we obtain more accurate expression for the renormalized FP operator $\mathscr{L}_{\mathbf{X}}^{R}$, without changing the first-order nature of time derivative in the l.h.s. of Eq. (2.56). Hence our multi-scale coarse-graining leads to great conceptual simplification as it provides directly a Markov description of the slow dynamics. To resolve this apparent paradox, we note that it is known [39] that the pdf of a non-Markovian process can always be expressed as the solution to differential equation. In the present case, by inspecting Eqs. (2.47) and (2.53) (and their higher-order analogues), we see that at each order, the fast components of pdf are completely determined by the slow component, i.e., by the marginal pdf of the slow variables. The reason underlying these results was already pointed out in Sec. II D: in the long timescales, the fast variables is completely slaved to the slow variables, and hence the history of the former is no longer relevant. The simplification, however, does not come without a cost. As we go to higher order in the expansion (2.57), there are more and more derivatives in slow variables, and hence we need to specify more and more boundary conditions in order to fix the solution to Eq. (2.56).

1. The approach of Laplace transform

An interesting coarse-graining method based on Laplace transform was discussed in Gardiner's classic monograph [75], for the particular problem of overdamped limit of one dimensional Brownian motion. Here we reformulate this method using our notations for a more general problem, so that a systematic comparison with our own method can be carried out.

The FP equation in Gardiner's theory has the form

$$\partial_t p = (\gamma \mathscr{L}_Y + \mathscr{L}_S)p, \tag{2.84}$$

where $\gamma = 1/\epsilon$ is a large parameter. We shall further assume that both $\mathscr{L}_Y$ and $\mathscr{L}_S$ are time independent.

We decompose the pdf into fast and slow components,

$$p = p_F + p_S = \mathscr{P}_F p + \mathscr{P}_S p, \tag{2.85}$$

and further using Eq. (2.22), which we rewrite below

$$\mathscr{P}_S \mathscr{L}_Y = \mathscr{L}_Y \mathscr{P}_S = 0. \tag{2.86}$$

We can rewrite Eq. (2.84) into

$$\partial_t p_S = \mathscr{P}_S(\gamma\mathscr{L}_Y + \mathscr{L}_S)(p_F + p_S)$$
$$= \mathscr{P}_S\mathscr{L}_S p_F + \mathscr{P}_S\mathscr{L}_S p_S; \tag{2.87a}$$
$$\partial_t p_F = \mathscr{P}_F(\gamma\mathscr{L}_Y + \mathscr{L}_S)(p_F + p_S)$$
$$= \gamma\mathscr{P}_F\mathscr{L}_Y p_F + \mathscr{P}_F\mathscr{L}_S p_F + \mathscr{P}_F\mathscr{L}_S p_S$$
$$= \gamma\mathscr{L}_Y p_F + \mathscr{P}_F\mathscr{L}_S p_F + \mathscr{P}_F\mathscr{L}_S p_S. \tag{2.87b}$$

For the overdamped limit problem with constant friction coefficient, there is an additional relation $\mathscr{P}_S\mathscr{L}_S\mathscr{P}_S = 0$, which makes the second term in Eq. (2.87a) vanish.

Now recall the Laplacian transform and its inverse are

$$\tilde{f}(s) = \int_0^\infty e^{-st} f(t)dt; \tag{2.88}$$
$$\int_0^\infty e^{-st}\frac{df}{dt}dt = s\tilde{f}(s) - f(0). \tag{2.89}$$

Applying the transform to Eqs. (2.87), we obtain (this is the place where we need $\mathscr{L}_Y$ and $\mathscr{L}_S$ to be time independent)

$$s\tilde{p}_S(s) - p_S(0) = \mathscr{P}_S\mathscr{L}_S\tilde{p}_F + \mathscr{P}_S\mathscr{L}_S\tilde{p}_S; \tag{2.90}$$
$$s\tilde{p}_F(s) - p_F(0) = (\gamma\mathscr{L}_Y + \mathscr{P}_F\mathscr{L}_S)\tilde{p}_F + \mathscr{P}_F\mathscr{L}_S\tilde{p}_S. \tag{2.91}$$

As before, we assume that at the initial time $p(0)$ has no fast component, i.e., $p_F(0) = 0$, then we obtain from Eq. (2.91),

$$\tilde{p}_F = (s - \gamma\mathscr{L}_Y - \mathscr{P}_F\mathscr{L}_S)^{-1}\mathscr{P}_F\mathscr{L}_S\tilde{p}_S. \tag{2.92}$$

Substitute this back into Eq. (2.90), we obtain

$$s\tilde{p}_S(s) = p_S(0) + \mathscr{P}_S\mathscr{L}_S\tilde{p}_S$$
$$+ \mathscr{P}_S\mathscr{L}_S(s - \gamma\mathscr{L}_Y - \mathscr{P}_F\mathscr{L}_S)^{-1}\mathscr{P}_F\mathscr{L}_S\tilde{p}_S. \tag{2.93}$$

If γ is large, we may take the approximation

$$(s - \gamma\mathscr{L}_Y - \mathscr{P}_F\mathscr{L}_S)^{-1} \approx -\gamma\mathscr{L}_Y^{-\perp} \tag{2.94}$$

and obtain

$$s\tilde{p}_S(s) = p_S(0) + \mathscr{P}_S\mathscr{L}_S\tilde{p}_S - \gamma^{-1}\mathscr{P}_S\mathscr{L}_S\mathscr{L}_Y^{-\perp}\mathscr{L}_S\tilde{p}_S. \tag{2.95}$$

This is in fact the Laplace transform of our first-order multi-scale expansion without taking into account the short-timescale details, see Eqs. (2.44) and (2.50).

Note that this method of Laplace transform is no longer applicable if $\mathscr{L}_S$ depend on time.

2. *Time-convolutionless projection operator formalism*

A projection operator method without time convolution was developed by Chaturvedi and Shibata [18–20] more than forty years ago. The formalism is based on interaction picture of Fokker-Planck dynamics, which makes it readily adaptable to quantum systems [21,22]. Nonetheless, it has not been widely applied to classical problems, probably due to its very compact and abstract structure. Even though the notations in the theory of Chaturvedi and Shibata are very different from those in the present theory, there seems close connections. If initially the fast variables are in conditional equilibrium/steady state, their formalisms yield equivalent results up to the second order. This connection will be discussed in great detail in Sec. VI B, where we study the overdamped limit of a Brownian motion. This connection may persist to infinite order.

III. COARSE-GRAINING OF COVARIANT FOKKER-PLANCK DYNAMICS

It must be admitted that Eqs. (2.56) and (2.57), which are the main results of the preceding section, are no more than mathematical formulas. They can acquire physical meaning only if the operators $\mathscr{L}_S$ and $\mathscr{L}_\mathbf{Y}$ have physical meanings. This is achieved in the present section by specializing to the covariant theory of Fokker-Planck dynamics [61]. Within this theory, the FPO is of second order and is parameterized by a generalized potential and a matrix of kinetic coefficients, both bearing clear physical meanings [77]. Applying Eq. (2.57), we will establish the connection between detailed balance property of the joint Fokker-Planck dynamics and that of the effective slow dynamics. Additionally, further assuming that the kinetic submatrix L^{ij} is independent of the fast variables $\boldsymbol{y}$, we will find that the first-order renormalized FPO of the effective slow dynamics is also a second-order operator with covariant form. The renormalized kinetic coefficients are explicitly expressed as functions of those of the joint dynamics. In a future publication, these results will be applied to study coarse-graining of stochastic thermodynamics, a theory of nonequilibrium statistical mechanics tailor-designed for small systems.

A. Covariant Fokker-Planck dynamics of the joint system

In Ref. [61] it was shown that any Langevin equation can be formulated in the following covariant form:

$$dz^\mu + (L^{\mu\nu}\partial_\nu U - \partial_\nu L^{\mu\nu})dt = b^{\mu\alpha}dW_\alpha, \tag{3.1}$$

where $\partial_\mu = \partial/\partial z^\mu$, and $L^{\mu\nu}(\mathbf{z})$ is the matrix of kinetic coefficients, whilst $U(\mathbf{z}) = U(\boldsymbol{x},\boldsymbol{y})$ is called the *generalized potential,* related to the steady-state distribution via $p^{\rm SS}(\mathbf{z}) = e^{-U(\mathbf{z})}$, which is assumed to be unique, and is not necessary a thermal equilibrium state. dW_α are Wiener noises, whose dimension is irrelevant. The product in the r.h.s. is understood in Ito's sense. Note that repeated indices are all summed over. The associated covariant Fokker-Planck equation is given by

$$\partial_t\, p(\mathbf{z},t) = \mathscr{L}(\mathbf{z})p(\mathbf{z},t), \tag{3.2a}$$
$$\mathscr{L}(\mathbf{z}) = \partial_\mu L^{\mu\nu}(\partial_\nu + (\partial_\nu U)), \tag{3.2b}$$

where $\mathscr{L}(\mathbf{z})$ is the *covariant FPO*, $p(\mathbf{z},t)$ is the pdf of variables $\mathbf{z}$.

As shown in Ref. [61], the noise amplitudes $b^{\mu\alpha}(\mathbf{z})$ are related to the symmetric part of kinetic matrix $L^{\mu\nu}(\mathbf{z})$ via

$$\sum_\alpha b^{\mu\alpha}b^{\nu\alpha} = L^{\mu\nu} + L^{\nu\mu} = 2B^{\mu\nu}. \tag{3.3}$$

The symmetric matrix $B^{\mu\nu}$ must be semipositive definite. The probability current is given by

$$j^\mu = -L^{\mu\nu}(\partial_\nu + (\partial_\nu U))p + \partial_\nu(Q^{\mu\nu}p), \tag{3.4}$$

where $Q^{\mu\nu}$ is the antisymmetric part of $L^{\mu\nu}$, so that

$$L^{\mu\nu} = B^{\mu\nu} + Q^{\mu\nu}. \tag{3.5}$$

Hence the Fokker-Planck equation (3.2a) can be rewritten as conservation of probability,

$$\partial_t p = -\partial_\mu j^\mu. \tag{3.6}$$

Note that in general $Q^{\mu\nu} \neq 0$, and hence the probability current does not vanish in the steady state $p^{\rm SS}(\mathbf{z}) = e^{-U(\mathbf{z})}$.

As in Sec. II, we write $\mathbf{z} = (\boldsymbol{x}, \boldsymbol{y})$, where $\boldsymbol{x}, \boldsymbol{y}$ are respectively the slow and fast variables. We assume that the noises acting on the fast variables and those acting on the slow variables are completely decoupled. We further introduce the small parameter ϵ by writing the matrix $b^{\mu\alpha}$ into the following block-diagonal form:

$$\boldsymbol{b} = \begin{pmatrix} b^{i\alpha} & 0 \\ 0 & b^{a\alpha}/\sqrt{\epsilon} \end{pmatrix}. \tag{3.7}$$

This also implies that the noises acting on the fast variables are much larger than those acting on the slow variables. In the above equation indices $i, j, k, \cdots$, and $a, b, c, \cdots$ refer, respectively, to the slow variables $\boldsymbol{x}$ and fast variables $\boldsymbol{y}$. The greek letters μ, ν are used exclusively for $\mathbf{z}$, and the index α is used exclusively for noise components. These conventions will be used throughout the paper. Substituting Eq. (3.7) back into Eq. (3.3) we find

$$\begin{aligned} \boldsymbol{L} + \boldsymbol{L}^T &= \begin{pmatrix} b^{i\alpha} b^{i\alpha} & 0 \\ 0 & b^{a\alpha} b^{b\alpha}/\epsilon \end{pmatrix} \\ &= 2\boldsymbol{B} = \begin{pmatrix} B^{ij} & 0 \\ 0 & B^{ab}/\epsilon \end{pmatrix}. \end{aligned} \tag{3.8}$$

We further assume that the antisymmetric matrix $Q^{\mu\nu}$ has the following form:

$$\boldsymbol{Q} = \begin{pmatrix} Q^{ij} & Q^{ia} \\ Q^{jb} & Q^{ab}/\epsilon \end{pmatrix}, \tag{3.9}$$

which implies that the conservative dynamics of the fast variables is much faster than that of slow variables. Substituting the preceding two equations into Eq. (3.5), we see that the kinetic matrix $\boldsymbol{L}$ can be written as

$$\begin{aligned} \boldsymbol{L} = \boldsymbol{B} + \boldsymbol{Q} &= \begin{pmatrix} B^{ij} + Q^{ij} & Q^{ia} \\ Q^{jb} & (B^{ab} + Q^{ab})/\epsilon \end{pmatrix} \\ &= \begin{pmatrix} L^{ij} & L^{ia} \\ L^{jb} & L^{ab}/\epsilon \end{pmatrix}. \end{aligned} \tag{3.10}$$

The Langevin equation (3.1) can then be rewritten as

$$\begin{aligned} &dx^k + (L^{kj}\partial_j U - \partial_j L^{kj})dt \\ &\quad + (L^{ka}\partial_a U - \partial_a L^{ka})dt = b^{k\alpha} dW_\alpha, \end{aligned} \tag{3.11a}$$

$$\begin{aligned} &dy^a + (L^{aj}\partial_j U - \partial_j L^{aj})dt \\ &\quad + \epsilon^{-1}(L^{ab}\partial_b U - \partial_b L^{ab})dt = \epsilon^{-1/2} b^{a\alpha} dW_\alpha, \end{aligned} \tag{3.11b}$$

whereas the FPO (3.2b) can be written in the form of Eq. (2.2),

$$\mathscr{L} = \mathscr{L}_S + \epsilon^{-1}\mathscr{L}_{\mathbf{Y}}, \tag{3.12a}$$

$$\mathscr{L}_S \equiv \mathscr{L}_{\mathbf{X}} + \mathscr{L}_{\mathbf{XY}} + \mathscr{L}_{\mathbf{YX}}, \tag{3.12b}$$

$$\mathscr{L}_{\mathbf{X}} \equiv \partial_i L^{ij}(\partial_j + (\partial_j U)), \tag{3.12c}$$

$$\mathscr{L}_{\mathbf{XY}} \equiv \partial_i L^{ia}(\partial_a + (\partial_a U)), \tag{3.12d}$$

$$\mathscr{L}_{\mathbf{YX}} \equiv \partial_a L^{ai}(\partial_i + (\partial_i U)), \tag{3.12e}$$

$$\mathscr{L}_{\mathbf{Y}} \equiv \partial_a L^{ab}(\partial_b + (\partial_b U)). \tag{3.12f}$$

Hence the parameter ϵ indeed plays the role of τ_y/τ_x. Note that $\mathscr{L}_{\mathbf{Y}}$ involves derivatives over $\boldsymbol{y}$ but not over $\boldsymbol{x}$, as we assumed in Sec. II A. Note also that Eqs. (2.6) and (2.7) are automatically satisfied by Eqs. (3.12). The kinetic matrices $L^{ij}, L^{ia}, L^{ai}, L^{ab}$ may depend on both $\boldsymbol{y}$ and $\boldsymbol{x}$. We shall see below that the renormalized dynamics of the slow variables is simpler if the submatrix L^{ij} [which enters $\mathscr{L}_{\mathbf{X}}$ in Eq. (3.12c)] is independent of $\boldsymbol{y}$.

Let us further define $U_{\mathbf{X}}(\boldsymbol{x})$ and $U_{\mathbf{Y}}(\boldsymbol{y};\boldsymbol{x})$,

$$U_{\mathbf{X}}(\boldsymbol{x}) \equiv -\log \int_{\boldsymbol{y}} e^{-U(\boldsymbol{x},\boldsymbol{y})}, \tag{3.13}$$

$$U_{\mathbf{Y}}(\boldsymbol{y};\boldsymbol{x}) \equiv U(\boldsymbol{x},\boldsymbol{y}) - U_{\mathbf{X}}(\boldsymbol{x}). \tag{3.14}$$

Note that $U_{\mathbf{X}}(\boldsymbol{x})$ depends on $\boldsymbol{x}$ but not on $\boldsymbol{y}$. Defined as above, $e^{-U_{\mathbf{X}}(\boldsymbol{x})}$ is the marginal pdf of $\boldsymbol{x}$ in the steady state, and $e^{-U_{\mathbf{Y}}(\boldsymbol{x},\boldsymbol{y})}$ is the conditional pdf of $\boldsymbol{y}$ given $\boldsymbol{x}$ in the steady state,

$$p^{\rm SS}_{\mathbf{X}}(\boldsymbol{x}) = e^{-U_{\mathbf{X}}(\boldsymbol{x})}, \tag{3.15}$$

$$p^{\rm SS}_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{y}|\boldsymbol{x}) = e^{-U_{\mathbf{Y}}(\boldsymbol{x};\boldsymbol{y})}, \tag{3.16}$$

$$e^{-U(\boldsymbol{x},\boldsymbol{y})} = e^{-U_{\mathbf{X}}(\boldsymbol{x}) - U_{\mathbf{Y}}(\boldsymbol{y};\boldsymbol{x})}, \tag{3.17}$$

which satisfy the following normalization conditions:

$$\int_{\boldsymbol{x}} e^{-U_{\mathbf{X}}(\boldsymbol{x})} = \int_{\boldsymbol{y}} e^{-U_{\mathbf{Y}}(\boldsymbol{y};\boldsymbol{x})} = 1. \tag{3.18}$$

This explicitly verifies Eq. (2.5). Since $U_{\mathbf{X}}$ is independent of $\boldsymbol{y}$, Eqs. (3.12d) and (3.12f) can be rewritten as

$$\mathscr{L}_{\mathbf{Y}} = \partial_a L^{ab}(\partial_b + (\partial_b U_{\mathbf{Y}})), \tag{3.19}$$

$$\mathscr{L}_{\mathbf{XY}} = \partial_i L^{ia}(\partial_a + (\partial_a U_{\mathbf{Y}})), \tag{3.20}$$

both of which annihilate the conditional steady state,

$$\mathscr{L}_{\mathbf{Y}} e^{-U_{\mathbf{Y}}} = 0, \tag{3.21}$$

$$\mathscr{L}_{\mathbf{XY}} e^{-U_{\mathbf{Y}}} = 0. \tag{3.22}$$

Hence Eq. (3.16) is the steady state of the conditional fast dynamics, as we have required in Eq. (2.4).

The following operator identity can be easily proved:

$$(\partial_j + (\partial_j U))e^{-U_{\mathbf{Y}}} = e^{-U_{\mathbf{Y}}}(\partial_j + (\partial_j U_{\mathbf{X}})). \tag{3.23}$$

Recall that the slow and fast projection operators are defined in Eqs. (2.11) and (2.15). Using Eq. (3.17), we can easily verify we see

$$\mathscr{P}_S\, e^{-U} = e^{-U}, \quad \mathscr{P}_F\, e^{-U} = 0. \tag{3.24}$$

which means that the joint steady state pdf e^{-U} is a slow distribution. In fact, every operator defined in Eq. (3.12) annihilates the joint steady state,

$$\begin{aligned} \mathscr{L}_{\mathbf{X}} e^{-U} &= \mathscr{L}_{\mathbf{XY}} e^{-U} = \mathscr{L}_{\mathbf{YX}} e^{-U} \\ &= \mathscr{L}_{\mathbf{Y}} e^{-U} = \mathscr{L}_S e^{-U} = 0. \end{aligned} \tag{3.25a}$$

These properties are peculiar to the covariant theory of Fokker-Planck theory. Finally, since $\mathscr{L}_{\mathbf{Y}}^{-\perp}\psi$ is a fast distribution for any $\psi(\boldsymbol{x},\boldsymbol{y})$, we also have

$$\int_{\boldsymbol{y}} \mathscr{L}_{\mathbf{Y}}^{-\perp}\psi(\boldsymbol{x},\boldsymbol{y}) = 0. \tag{3.25b}$$

B. More on Hilbert spaces and projections

We need to define more mathematical notations. We define three types of inner products,

$$(\phi,\psi) = \int_{\boldsymbol{x},\boldsymbol{y}} \phi(\boldsymbol{x},\boldsymbol{y})\psi(\boldsymbol{x},\boldsymbol{y})e^{U}, \tag{3.26a}$$

$$\langle A,B\rangle = \int_{\boldsymbol{x},\boldsymbol{y}} A(\boldsymbol{x},\boldsymbol{y})B(\boldsymbol{x},\boldsymbol{y})e^{-U}, \tag{3.26b}$$

$$\langle A,\phi\rangle = \int_{\boldsymbol{x},\boldsymbol{y}} A(\boldsymbol{x},\boldsymbol{y})\phi(\boldsymbol{x},\boldsymbol{y}), \tag{3.26c}$$

where (ϕ,ψ) is the inner product between two distributions, $\langle A,B\rangle$ is the inner product between two observables, whereas $\langle A,\phi\rangle$ is the mixed product between observable and distribution. The precise meanings of distributions and observables are explained in Appendix A 1. Roughly speaking, distributions are analogues of probability density functions, but are not necessary positive. They converge to zero sufficiently fast as $\boldsymbol{x},\boldsymbol{y}\to\infty$, so that the inner product in Eq. (3.26a) is finite. By contrast, observables are functions are $\boldsymbol{x},\boldsymbol{y}$, which make the inner product in Eq. (3.26b) convergent.

For a generic operator $\mathscr{O}$ acting on distributions, we define $\mathscr{O}^\dagger$, $\mathscr{O}^+$, $(\mathscr{O}^\dagger)^+$ using inner products,

$$(\psi,\mathscr{O}^\dagger\phi) = (\mathscr{O}\psi,\phi), \tag{3.27a}$$

$$\langle\mathscr{O}^+A,\psi\rangle = \langle A,\mathscr{O}\psi\rangle, \tag{3.27b}$$

$$\langle(\mathscr{O}^\dagger)^+A,\psi\rangle = \langle A,\mathscr{O}^\dagger\psi\rangle. \tag{3.27c}$$

The following properties can be easily seen:

$$(\mathscr{O}_1\mathscr{O}_1)^\dagger = \mathscr{O}_2^\dagger\mathscr{O}_1^\dagger, \tag{3.28}$$

$$(\mathscr{O}_1\mathscr{O}_1)^+ = \mathscr{O}_2^+\mathscr{O}_1^+, \tag{3.29}$$

$$((\mathscr{O}_1\mathscr{O}_1)^\dagger)^+ = (\mathscr{O}_1^\dagger)^+(\mathscr{O}_2^\dagger)^+. \tag{3.30}$$

Note that $\mathscr{O}$, $\mathscr{O}^\dagger$ are operators acting on distributions, whereas $\mathscr{O}^+$, $(\mathscr{O}^\dagger)^+$ are operators acting on observables. It is easy to verify that the projection operators $\mathscr{P}_S$ and $\mathscr{P}_F$ defined in Sec. III B, which act on distributions, are both Hermitian,

$$\mathscr{P}_S^\dagger = \mathscr{P}_S, \quad \mathscr{P}_F^\dagger = \mathscr{P}_F. \tag{3.31}$$

The operators $\mathscr{L}^\dagger$, $\mathscr{L}^+$, $(\mathscr{L}^\dagger)^+$ associated with the FPO $\mathscr{L}$, defined in Eq. (3.2b), can be calculated using the definitions (3.27),

$$\mathscr{L} = \partial_\mu L^{\mu\nu}(\partial_\nu + (\partial_\nu U)), \tag{3.32a}$$

$$\mathscr{L}^\dagger = \partial_\mu L^{\nu\mu}(\partial_\nu + (\partial_\nu U)), \tag{3.32b}$$

$$\mathscr{L}^+ \equiv (\partial_\mu - \partial_\mu U)L^{\nu\mu}\partial_\nu, \tag{3.32c}$$

$$(\mathscr{L}^\dagger)^+ \equiv (\partial_\mu - \partial_\mu U)L^{\mu\nu}\partial_\nu. \tag{3.32d}$$

Similarly we can work out the operators $\mathscr{L}_{\mathbf{Y}}^\dagger$, $\mathscr{L}_{\mathbf{Y}}^+$, $(\mathscr{L}_{\mathbf{Y}}^\dagger)^+$ associated with the FPO $\mathscr{L}_{\mathbf{Y}}$,

$$\mathscr{L}_{\mathbf{Y}} \equiv \partial_a L^{ab}(\partial_b + (\partial_b U_{\mathbf{Y}})), \tag{3.33a}$$

$$\mathscr{L}_{\mathbf{Y}}^\dagger = \partial_a L^{ba}(\partial_b + (\partial_b U_{\mathbf{Y}})), \tag{3.33b}$$

$$\mathscr{L}_{\mathbf{Y}}^+ = (\partial_a - (\partial_a U_{\mathbf{Y}}))L^{ba}\partial_b, \tag{3.33c}$$

$$(\mathscr{L}_{\mathbf{Y}}^\dagger)^+ = (\partial_a - (\partial_a U_{\mathbf{Y}}))L^{ab}\partial_b. \tag{3.33d}$$

It is also easy to verify the following relations:

$$\mathscr{L}_{\mathbf{Y}}^+ = e^{U_{\mathbf{Y}}}\mathscr{L}_{\mathbf{Y}}^\dagger e^{-U_{\mathbf{Y}}}, \tag{3.34a}$$

$$(\mathscr{L}_{\mathbf{Y}}^\dagger)^+ = e^{U_{\mathbf{Y}}}\mathscr{L}_{\mathbf{Y}} e^{-U_{\mathbf{Y}}}. \tag{3.34b}$$

It is useful to define the various inner products associated with the conditional fast dynamics Eq. (2.3),

$$(\phi,\psi)_{\mathbf{Y}} = \int_{\boldsymbol{y}} \phi(\boldsymbol{x},\boldsymbol{y})\psi(\boldsymbol{x},\boldsymbol{y})e^{U_{\mathbf{Y}}}, \tag{3.35a}$$

$$\langle A,B\rangle_{\mathbf{Y}} = \int_{\boldsymbol{y}} A(\boldsymbol{x},\boldsymbol{y})B(\boldsymbol{x},\boldsymbol{y})e^{-U_{\mathbf{Y}}}, \tag{3.35b}$$

$$\langle A,\phi\rangle_{\mathbf{Y}} = \int_{\boldsymbol{y}} A(\boldsymbol{x},\boldsymbol{y})\phi(\boldsymbol{x},\boldsymbol{y}). \tag{3.35c}$$

Note that only $\boldsymbol{y}$ are integrated over in the above expressions. Using Eqs. (3.33) and (3.26), we can further prove

$$(\mathscr{L}_{\mathbf{Y}}^\dagger\psi,\phi)_{\mathbf{Y}} = (\psi,\mathscr{L}_{\mathbf{Y}}\phi)_{\mathbf{Y}}, \tag{3.36a}$$

$$\langle\mathscr{L}_{\mathbf{Y}}^+A,\psi\rangle_{\mathbf{Y}} = \langle A,\mathscr{L}_{\mathbf{Y}}\psi\rangle_{\mathbf{Y}}, \tag{3.36b}$$

$$\langle(\mathscr{L}_{\mathbf{Y}}^\dagger)^+A,\psi\rangle_{\mathbf{Y}} = \langle A,\mathscr{L}_{\mathbf{Y}}^\dagger\psi\rangle_{\mathbf{Y}}. \tag{3.36c}$$

It is also useful to define the time-correlation functions of the conditional fast dynamics,

$$\begin{aligned}\langle A(t)B(0)\rangle_{\mathbf{Y}} &\equiv \int_{\boldsymbol{y}} (e^{t\mathscr{L}_{\mathbf{Y}}^+/\epsilon}A)B\, e^{-U_{\mathbf{Y}}} \\ &= \langle e^{t\mathscr{L}_{\mathbf{Y}}^+/\epsilon}A,B\rangle_{\mathbf{Y}},\end{aligned} \tag{3.37}$$

where the slow variables are fixed but the fast variables evolve according to Eq. (2.3). Integrating this correlation function over time, we find that the result is of order ϵ,

$$\int_0^\infty dt\ \langle A(t)B(0)\rangle_{\mathbf{Y}} = \epsilon\int_0^\infty dt\ \langle e^{t\mathscr{L}_{\mathbf{Y}}^+}A,B\rangle_{\mathbf{Y}}, \tag{3.38}$$

where we have rescaled the time variable in the r.h.s.. Note that $e^{t\mathscr{L}_{\mathbf{Y}}^+/\epsilon}$ can be understood as evolution operator in Heisenberg picture. For detailed discussion on Heisenberg vs Schrödinger picture of Fokker-Planck dynamics, see Appendix A 6.

Using Eq. (3.27b), we define operators $\mathscr{P}_{\mathrm{S}}^+$, $\mathscr{P}_{\mathrm{F}}^+$ via

$$\langle\mathscr{P}_{\mathrm{S}}^+A,\psi\rangle = \langle A,\mathscr{P}_{\mathrm{S}}\psi\rangle, \tag{3.39a}$$

$$\langle\mathscr{P}_{\mathrm{F}}^+A,\psi\rangle = \langle A,\mathscr{P}_{\mathrm{F}}\psi\rangle. \tag{3.39b}$$

These are slow and fast projection operators in the space of observables. It is easy to verify that $\mathscr{P}_{\mathrm{S}}^+$, $\mathscr{P}_{\mathrm{F}}^+$ are involutive and Hermitian, and hence are legitimate projection operators.

Straightforward calculation leads to

$$\mathscr{P}_{\rm S}^{+}A(\boldsymbol{x},\boldsymbol{y})=\int_{\boldsymbol{y}'}A(\boldsymbol{x},\boldsymbol{y}')e^{-U_{\rm Y}(\boldsymbol{y}';\boldsymbol{x})}, \tag{3.40a}$$

$$\mathscr{P}_{\rm F}^{+}A(\boldsymbol{x},\boldsymbol{y})=A(\boldsymbol{x},\boldsymbol{y})-\int_{\boldsymbol{y}'}A(\boldsymbol{x},\boldsymbol{y}')e^{-U_{\rm Y}(\boldsymbol{y}';\boldsymbol{x})}. \tag{3.40b}$$

Hence $\mathscr{P}_{\rm S}^{+}$ carries out conditional steady-state average over fast variables, whereas $\mathscr{P}_{\rm F}^{+}$ projects out the fluctuations away from conditional steady state. Note that $\mathscr{P}_{\rm S}$, $\mathscr{P}_{\rm S}^{+}$, $\mathscr{P}_{\rm F}$, $\mathscr{P}_{\rm F}^{+}$ are all integral operators, which act only on the fast variable $\boldsymbol{y}$, but not on the slow variables $\boldsymbol{x}$.

An arbitrary observable $A(\boldsymbol{x},\boldsymbol{y})$ can be decomposed into a slow component and a fast component,

$$\begin{aligned}A(\boldsymbol{x},\boldsymbol{y}) &= \mathscr{P}_{\rm S}^{+}A(\boldsymbol{x},\boldsymbol{y})+\mathscr{P}_{\rm F}^{+}A(\boldsymbol{x},\boldsymbol{y})\\ &= A_{\rm S}(\boldsymbol{x})+A_{\rm F}(\boldsymbol{x},\boldsymbol{y}).\end{aligned} \tag{3.41}$$

Note that the slow component $A_{\rm S}(\boldsymbol{x})$ is independent of the fast variables $\boldsymbol{y}$, whereas the conditional average of the fast component $A_{\rm F}(\boldsymbol{x},\boldsymbol{y})$ vanishes identically. Correspondingly the Hilbert space of observables $\mathscr{H}_{\rm obs}$ is decomposed into the subspace of slow observables and that of the fast observables,

$$\mathscr{H}_{\rm obs}=\mathscr{H}_{\rm obs}^{\rm S}\oplus\mathscr{H}_{\rm obs}^{\rm F}, \tag{3.42a}$$

$$\mathscr{H}_{\rm obs}^{\rm S}=\{A(\boldsymbol{x})\}, \tag{3.42b}$$

$$\mathscr{H}_{\rm obs}^{\rm F}=\left\{A(\boldsymbol{x},\boldsymbol{y})\middle|\int_{\boldsymbol{y}}A(\boldsymbol{x},\boldsymbol{y})e^{-U_{\rm Y}(\boldsymbol{y};\boldsymbol{x})}=0\right\}, \tag{3.42c}$$

Using Eqs. (3.40) and (2.11), (2.14), we can directly verify

$$\mathscr{P}_{\rm S}^{+}=e^{U_{\rm Y}}\mathscr{P}_{\rm S}e^{-U_{\rm Y}}, \tag{3.43a}$$

$$\mathscr{P}_{\rm F}^{+}=e^{U_{\rm Y}}\mathscr{P}_{\rm F}\,e^{-U_{\rm Y}}. \tag{3.43b}$$

These relations are clearly analogous to Eq. (3.34a). Combining these with Eqs. (2.19) we obtain

$$(\mathscr{P}_{\rm S}^{+})^2=\mathscr{P}_{\rm S}^{+},\quad(\mathscr{P}_{\rm F}^{+})^2=\mathscr{P}_{\rm F}^{+}, \tag{3.44a}$$

$$\mathscr{P}_{\rm S}^{+}+\mathscr{P}_{\rm F}^{+}=I, \tag{3.44b}$$

$$(\mathscr{P}_{\rm S}^{+})^{\dagger}=\mathscr{P}_{\rm S}^{+},\quad(\mathscr{P}_{\rm F}^{+})^{\dagger}=\mathscr{P}_{\rm F}^{+}, \tag{3.44c}$$

$$\mathscr{P}_{\rm S}^{+}\mathscr{P}_{\rm F}^{+}=\mathscr{P}_{\rm F}^{+}\mathscr{P}_{\rm S}^{+}=0. \tag{3.44d}$$

C. Preservation of detailed balance

Stochastic processes with detailed balance play an important role in nonequilibrium statistical physics, because they (in the steady regime) correspond to reversible processes with no entropy production. Any process violating detailed balance is irreversible with positive entropy production. This essential connection between breaking of detailed balance and dissipation can be deemed as the cornerstone of stochastic thermodynamics and more generally of nonequilibrium statistical physics.

As in the previous papers of this sequel [61,62,64], we always choose each component z^{μ} of system variables with definite symmetry under time reversal,

$$z^{\mu}\to z^{*\mu}=\varepsilon_{\mu}z^{\mu}, \tag{3.45}$$

where $\varepsilon_{\mu}=1,-1$ for even and odd variables respectively. The FPOs may further depend on certain external parameter λ, which is transformed into λ^{*} under time reversal. For odd parameter, such as magnetic field or angular velocity, we have $\lambda^{*}=-\lambda$. We call the Markov process generated by $\mathscr{L}(\mathbf{z},\lambda)$ the *forward process*, and that generated by $\mathscr{L}(\mathbf{z},\lambda^{*})$ the *backward process*. A Fokker-Planck dynamics is said to have *time-reversal symmetry*, if for any t_1,t_2, the two-point joint pdfs of the forward and backward processes satisfy the following relation:

$$P_F(\mathbf{z}_2,t_2;\mathbf{z}_1,t_1)=P_B(\mathbf{z}_1^{*},t_2;\mathbf{z}_2^{*},t_1). \tag{3.46}$$

As explained in Appendix A 8, this relation automatically implies that the both the forward and the backward processes are in the steady regime. In Ref. [61], it was shown that time-reversal symmetry is equivalent to the following *conditions of detailed balance*:

$$U(\mathbf{z},\lambda)=U(\mathbf{z}^{*},\lambda^{*}), \tag{3.47a}$$

$$\varepsilon_{\mu}L^{\mu\nu}(\mathbf{z}^{*},\lambda^{*})\varepsilon_{\nu}=L^{\nu\mu}(\mathbf{z},\lambda). \tag{3.47b}$$

In Appendix A 8, it is shown that these conditions can be expressed in an equivalent but compact form

$$\mathscr{L}^{\dagger}(\mathbf{z},\lambda)=\mathscr{L}(\mathbf{z}^{*},\lambda^{*}), \tag{3.48}$$

where $\mathscr{L}^{\dagger}$ is defined in Eq. (3.32b).

Assuming that the joint dynamics of fast and slow variables possessing time-reversal symmetry, we would like to know whether this symmetry is inherited by the effective slow dynamics. For this purpose, we only need to integrate out the fast variables $\boldsymbol{y}$ directly from the relation (3.46). We obtain the following relation for the two-point joint pdf of the effective slow dynamics:

$$P_F(\boldsymbol{x}_2,t_2;\boldsymbol{x}_1,t_1)=P_B(\boldsymbol{x}_1^{*},t_2;\boldsymbol{x}_2^{*},t_1). \tag{3.49}$$

This shows that time-reversal symmetry is preserved by coarse-graining. Following the same logic as one goes from Eq. (3.46) to Eqs. (3.47) and (3.48), we then see that the effective slow dynamics also satisfies the conditions of detailed balance,

$$\left(\mathscr{L}_{\mathbf{X}}^{R}\right)^{\dagger}(\boldsymbol{x},\lambda)=\mathscr{L}_{\mathbf{X}}^{R}(\boldsymbol{x}^{*},\lambda^{*}), \tag{3.50a}$$

or equivalently

$$U_{\mathbf{X}}(\boldsymbol{x},\lambda)=U_{\mathbf{X}}(\boldsymbol{x}^{*},\lambda^{*}), \tag{3.50b}$$

$$\varepsilon_i(L^{R})^{ij}(\boldsymbol{x}^{*},\lambda^{*})\varepsilon_j=(L^{R})^{ji}(\boldsymbol{x},\lambda). \tag{3.50c}$$

In Appendix A 9 we outline a direct proof of Eq. (3.50a) using the explicit (perturbative) form of renormalized FPO $\mathscr{L}_{\mathbf{X}}^{R}$.

D. Renormalization of covariant FPO

$\mathscr{L}_S$ as given by Eq. (3.12b) is in general a second-order partial differential operator of $\boldsymbol{x}$, the first-order correction as given by Eq. (2.50) is a fourth-order differential. Hence the resulting first-order renormalized FPO of the effective slow dynamics is also a fourth-order operator, which does not correspond to Langevin dynamics with usual Gaussian white noises. This point was also pointed out in Ref. [53]. In this section, we will assume that the submatrix of kinetic

coefficients L^{ij} [which parameterizes $\mathscr{L}_{\mathbf{X}}$, c.f. Eq. (3.12c)] is independent of the fast variables $\boldsymbol{y}$. With this assumption, we will find that the first-order renormalized FPO is a second-order differential operator in $\boldsymbol{x}$, and has the covariant form Eq. (3.2b). Several equivalent expressions will be derived for the kinetic matrix of the effective slow dynamics expressed. One particularly useful form involves integrated correlation functions of the conditional fast dynamics. In Sec. VII, we will discuss an example where L^{ij} depends on fast variables $\boldsymbol{y}$. We will show explicitly that the first-order renormalized FPO contains fourth-order derivative of slow variables. This implies that the effective slow dynamics is not usual Markov process with Gaussian white noises.

We use Eqs. (2.44) and (3.12b) to calculate $\mathscr{L}^{(0)}_{\mathbf{X}} p_{\mathbf{X}}$,

$$\mathscr{L}^{(0)}_{\mathbf{X}} p_{\mathbf{X}} = \int_y (\mathscr{L}_{\mathbf{X}} + \mathscr{L}_{\mathbf{XY}} + \mathscr{L}_{\mathbf{YX}}) e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}. \tag{3.51}$$

Using Eqs. (3.12d) and (3.12e) it is easy to see that

$$\mathscr{L}_{\mathbf{XY}} e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} = 0, \tag{3.52}$$

$$\int_y \mathscr{L}_{\mathbf{YX}} \cdots = 0. \tag{3.53}$$

Hence Eq. (3.51) can be simplified to

$$\begin{aligned} \mathscr{L}^{(0)}_{\mathbf{X}} p_{\mathbf{X}} &= \int_y \partial_i L^{ij} (\partial_j + \partial_j U) e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} \\ &= \int_y \partial_i L^{ij} e^{-U_{\mathbf{Y}}} (\partial_j + \partial_j U_{\mathbf{X}})\, p_{\mathbf{X}}, \end{aligned} \tag{3.54}$$

where in the second step we used Eq. (3.23). Since L^{ij}, $U_{\mathbf{X}}$, $p_{\mathbf{X}}$ are all independent of $\boldsymbol{y}$, the integral over $\boldsymbol{y}$ can be trivially carried out, yielding

$$\mathscr{L}^{(0)}_{\mathbf{X}} = \partial_i L^{ij} (\partial_j + (\partial_j U_{\mathbf{X}})). \tag{3.55}$$

Now let us calculate the first-order renormalization. Using Eq. (3.12b) in Eq. (2.50), we have

$$\begin{aligned} \mathscr{L}^{(1)}_{\mathbf{X}} p_{\mathbf{X}} &= -\int_y (\mathscr{L}_{\mathbf{X}} + \mathscr{L}_{\mathbf{XY}} + \mathscr{L}_{\mathbf{YX}}) \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} \\ &= -\partial_i \Big[L^{ik} \partial_k \int_y \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} + L^{ik} \int_y (\partial_k U) \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} + \int_y L^{ia} \partial_a \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} \\ &\quad + \int_y L^{ia} (\partial_a U_{\mathbf{Y}}) \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} \Big], \end{aligned} \tag{3.56}$$

where in the second step we have pulled L^{ij} out of the integral, since they are independent of $\boldsymbol{y}$. Inside the square bracket in Eq. (3.56), the first term vanishes because of Eq. (3.25b). The second term can be simplified as

$$L^{ik} \int_y (\partial_k U_{\mathbf{Y}}) \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} + L^{ik} (\partial_k U_{\mathbf{X}}) \int_y \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} = L^{ik} \int_y (\partial_k U_{\mathbf{Y}}) \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}, \tag{3.57}$$

again using Eq. (3.25b). The third term in the square bracket in Eq. (3.56) becomes $-\int_y (\partial_a L^{ia}) \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}$ upon integration by parts. Summarizing, we can rewrite Eq. (3.56) as

$$\begin{aligned} \mathscr{L}^{(1)}_{\mathbf{X}} p_{\mathbf{X}} &= -\partial_i \int_y \big[L^{ik} (\partial_k U_{\mathbf{Y}}) - (\partial_a L^{ia}) + L^{ia} (\partial_a U_{\mathbf{Y}}) \big] \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} \\ &= \partial_i \int_y e^{U_{\mathbf{Y}}} \big[L^{ik} (\partial_k e^{-U_{\mathbf{Y}}}) + (\partial_a L^{ia} e^{-U_{\mathbf{Y}}}) \big] \mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}. \end{aligned} \tag{3.58}$$

Let us now calculate $\mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}}$. Using Eq. (3.12b), (3.12c), and (3.12e), we have

$$\begin{aligned} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} &= (\mathscr{L}_{\mathbf{X}} + \mathscr{L}_{\mathbf{YX}} + \mathscr{L}_{\mathbf{XY}}) p_{\mathbf{X}} e^{-U_{\mathbf{Y}}} \\ &= \partial_l L^{lj} (\partial_j + (\partial_j U)) p_{\mathbf{X}} e^{-U_{\mathbf{Y}}} + \partial_b L^{bj} (\partial_j + (\partial_j U)) p_{\mathbf{X}} e^{-U_{\mathbf{Y}}} \\ &= \partial_l L^{lj}\, e^{-U_{\mathbf{Y}}} (\partial_j + (\partial_j U_{\mathbf{X}})) p_{\mathbf{X}} + \partial_b L^{bj}\, e^{-U_{\mathbf{Y}}} (\partial_j + (\partial_j U_{\mathbf{X}})) p_{\mathbf{X}}, \end{aligned} \tag{3.59}$$

where in the second step we used Eqs. (3.52), in the last equality we have used Eq. (3.23). Note that ∂_l, ∂_j, ∂_b act on everything to the right. Expanding the derivatives, we further obtain

$$\mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} = e^{-U_{\mathbf{Y}}} \partial_l (L^{lj} (\partial_j + (\partial_j U_{\mathbf{X}})) p_{\mathbf{X}}) + [L^{lj} (\partial_l e^{-U_{\mathbf{Y}}}) + (\partial_b L^{bj} e^{-U_{\mathbf{Y}}})] (\partial_j + (\partial_j U_{\mathbf{X}})) p_{\mathbf{X}}. \tag{3.60}$$

Let $\mathscr{L}^{-\perp}_{\mathbf{Y}}$ acting on Eq. (3.60) from left, the first term in r.h.s. disappears, again because L^{lj} are independent of $\boldsymbol{y}$. Hence we find

$$\mathscr{L}^{-\perp}_{\mathbf{Y}} \mathscr{L}_S\, e^{-U_{\mathbf{Y}}} p_{\mathbf{X}} = \mathscr{L}^{-\perp}_{\mathbf{Y}} \big[L^{lj} (\partial_l e^{-U_{\mathbf{Y}}}) + (\partial_b L^{bj} e^{-U_{\mathbf{Y}}}) \big] (\partial_j + (\partial_j U_{\mathbf{X}})) p_{\mathbf{X}}.$$

Inserting this into Eq. (3.58), we see that $\mathscr{L}^{(1)}_{\mathbf{X}}$ can be written as a covariant FPO,

$$\mathscr{L}^{(1)}_{\mathbf{X}} = \partial_i \delta L^{ij}(\boldsymbol{x}) (\partial_j + \partial_j U_{\mathbf{X}}), \tag{3.61}$$

where $\delta L^{ij}(\boldsymbol{x})$ can be understood as the first-order renormalization of kinetic coefficients due to fluctuations of fast variables, and is defined by

$$\delta L^{ij}(\boldsymbol{x}) \equiv \int_{\boldsymbol{y}} e^{U_{\mathbf{Y}}}[L^{ik}(\partial_k e^{-U_{\mathbf{Y}}}) + (\partial_a L^{ia} e^{-U_{\mathbf{Y}}})]\mathscr{L}_{\mathbf{Y}}^{-\perp}[(\partial_l e^{-U_{\mathbf{Y}}})L^{lj} + (\partial_b L^{bj} e^{-U_{\mathbf{Y}}})] \tag{3.62a}$$

$$= ([L^{ik}(\partial_k e^{-U_{\mathbf{Y}}}) + (\partial_a L^{ia} e^{-U_{\mathbf{Y}}})],\ \mathscr{L}_{\mathbf{Y}}^{-\perp}[(\partial_l e^{-U_{\mathbf{Y}}})L^{lj} + (\partial_b L^{bj} e^{-U_{\mathbf{Y}}})])_{\mathbf{Y}}, \tag{3.62b}$$

where $(\ ,\ )_{\mathbf{Y}}$ is the inner product defined in Eqs. (3.35).

Because $\mathscr{L}_{\mathbf{Y}}$ is invertible and negative in the fast subspace $\mathscr{H}^F_{\rm dist}$, the following operator identity can be established:

$$\mathscr{L}_{\mathbf{Y}}^{-\perp} = \mathscr{P}_{\rm F}\mathscr{L}_{\mathbf{Y}}^{-1}\mathscr{P}_{\rm F} = -\mathscr{P}_{\rm F}\int_0^\infty dt\, e^{t\,\mathscr{L}_{\mathbf{Y}}}\mathscr{P}_{\rm F}. \tag{3.63}$$

Using this, we can rewrite Eq. (3.62b) as

$$\delta L^{ij}(\boldsymbol{x}) = -\int_0^\infty dt\ ([L^{ik}(\partial_k e^{-U_{\mathbf{Y}}}) + (\partial_a L^{ia} e^{-U_{\mathbf{Y}}})],\ \mathscr{P}_{\rm F}e^{t\,\mathscr{L}_{\mathbf{Y}}}\mathscr{P}_{\rm F}[(\partial_l e^{-U_{\mathbf{Y}}})L^{lj} + (\partial_b L^{bj} e^{-U_{\mathbf{Y}}})])_{\mathbf{Y}} \tag{3.64a}$$

$$= -\int_0^\infty dt\ (\mathscr{P}_{\rm F}[L^{ik}(\partial_k e^{-U_{\mathbf{Y}}}) + (\partial_a L^{ia} e^{-U_{\mathbf{Y}}})],\ e^{t\,\mathscr{L}_{\mathbf{Y}}}\mathscr{P}_{\rm F}[(\partial_l e^{-U_{\mathbf{Y}}})L^{lj} + (\partial_b L^{bj} e^{-U_{\mathbf{Y}}})])_{\mathbf{Y}}, \tag{3.64b}$$

we we have used the fact that $\mathscr{P}_{\rm F}$ is Hermitian with respect to the inner product $(\ ,\ )_{\mathbf{Y}}$.

Furthermore, using the definition of $\mathscr{P}_{\rm F}$, Eq. (2.15), we can explicitly verify

$$\mathscr{P}_{\rm F}(e^{-U_{\mathbf{Y}}}\partial_k L^{ik}) = \mathscr{P}_{\rm F}(e^{-U_{\mathbf{Y}}}\partial_l L^{lj}) = 0. \tag{3.65}$$

Hence we have

$$\mathscr{P}_{\rm F}[L^{ik}(\partial_k e^{-U_{\mathbf{Y}}}) + (\partial_a L^{ia} e^{-U_{\mathbf{Y}}})] = \mathscr{P}_{\rm F}[(\partial_k L^{ik} e^{-U_{\mathbf{Y}}}) + (\partial_a L^{ia} e^{-U_{\mathbf{Y}}})] = \mathscr{P}_{\rm F}(\partial_\mu L^{i\mu} e^{-U_{\mathbf{Y}}}), \tag{3.66a}$$

$$\mathscr{P}_{\rm F}[(\partial_l e^{-U_{\mathbf{Y}}})L^{lj} + (\partial_b L^{bj} e^{-U_{\mathbf{Y}}})] = \mathscr{P}_{\rm F}[(\partial_l e^{-U_{\mathbf{Y}}} L^{lj}) + (\partial_b L^{bj} e^{-U_{\mathbf{Y}}})] = \mathscr{P}_{\rm F}(\partial_\nu L^{\nu j} e^{-U_{\mathbf{Y}}}), \tag{3.66b}$$

where repeated indices μ, ν mean summations over components of $\mathbf{z} = (\boldsymbol{x}, \boldsymbol{y})$. Hence Eq. (3.64) can be rewritten into

$$\delta L^{ij}(\boldsymbol{x}) = -\int_0^\infty dt([\mathscr{P}_{\rm F}(\partial_\mu L^{i\mu} e^{-U_{\mathbf{Y}}})],\ e^{t\,\mathscr{L}_{\mathbf{Y}}}[\mathscr{P}_{\rm F}(\partial_\nu L^{\nu j} e^{-U_{\mathbf{Y}}})])_{\mathbf{Y}} \tag{3.67a}$$

$$= -\int_0^\infty dt \int_{\boldsymbol{y}} e^{U_{\mathbf{Y}}}[\mathscr{P}_{\rm F}(\partial_\mu L^{i\mu} e^{-U_{\mathbf{Y}}})]e^{t\,\mathscr{L}_{\mathbf{Y}}}[\mathscr{P}_{\rm F}(\partial_\nu L^{\nu j} e^{-U_{\mathbf{Y}}})]. \tag{3.67b}$$

Further using the following identity [which follows from Eq. (3.36b)]

$$\int_{\boldsymbol{y}} \psi\, e^{t\,\mathscr{L}_{\mathbf{Y}}}\phi = \int_{\boldsymbol{y}} (e^{t\,\mathscr{L}^+_{\mathbf{Y}\mathbf{Y}}}\psi)\,\phi, \tag{3.68}$$

as well as Eqs. (3.43b), we can rewrite Eq. (3.67b) as

$$\delta L^{ij}(\boldsymbol{x}) = -\int_0^\infty dt\,\langle e^{t\,\mathscr{L}^+_{\mathbf{Y}}}\mathscr{P}^+_{\rm F}(\partial_\mu L^{i\mu} - L^{i\mu}\partial_\mu U_{\mathbf{Y}}),\ \mathscr{P}^+_{\rm F}(\partial_\nu L^{\nu j} - L^{\nu j}\partial_\nu U_{\mathbf{Y}})\rangle_{\mathbf{Y}} \tag{3.69a}$$

$$= -\int_0^\infty dt\,\langle e^{t\,\mathscr{L}^+_{\mathbf{Y}}}\mathscr{P}^+_{\rm F}(\partial_\mu L^{i\mu} - L^{i\mu}\partial_\mu U),\ \mathscr{P}^+_{\rm F}(\partial_\nu L^{\nu j} - L^{\nu j}\partial_\nu U)\rangle_{\mathbf{Y}}, \tag{3.69b}$$

where in the second equality we have used

$$\mathscr{P}^+_{\rm F}L^{i\mu}\partial_\mu U_{\mathbf{X}} = \mathscr{P}^+_{\rm F}L^{ik}\partial_k U_{\mathbf{X}} = 0, \tag{3.70a}$$

$$\mathscr{P}^+_{\rm F}L^{\nu j}\partial_\nu U_{\mathbf{X}} = \mathscr{P}^+_{\rm F}L^{lj}\partial_l U_{\mathbf{X}} = 0, \tag{3.70b}$$

which follows directly from Eq. (3.40b).

Now with $\mathscr{L}^+$, $(\mathscr{L}^\dagger)^+$ defined Eqs. (3.32), it is easy to show

$$\mathscr{L}^+ x^i = (\partial_\mu L^{i\mu}) - L^{i\mu}(\partial_\mu U), \tag{3.71a}$$

$$(\mathscr{L}^\dagger)^+ x^i = (\partial_\mu L^{\mu i}) - L^{\mu i}(\partial_\mu U). \tag{3.71b}$$

Hence we can rewrite Eq. (3.69b) as

$$\delta L^{ij}(\boldsymbol{x}) = -\int_0^\infty dt\,\langle e^{t\,\mathscr{L}^+_{\mathbf{Y}}}\mathscr{P}^+_{\rm F}\mathscr{L}^+ x^i,\ \mathscr{P}^+_{\rm F}(\mathscr{L}^\dagger)^+ x^j\rangle_{\mathbf{Y}}. \tag{3.72}$$

The inner product in Eq. (3.72) has a clear physical meaning. It is the correlation function of two observables $\mathscr{P}_{\rm F}^{+}\mathscr{L}^{+}x^{i}$ and $\mathscr{P}_{\rm F}^{+}(\mathscr{L}^{\dagger})^{+}x^{j}$ of the conditional fast dynamics. (Recall that $\mathscr{P}_{\rm F}^{+}\mathscr{L}^{+}A$ is the fluctuation of observable A around its conditional average.)

Finally, using Eqs. (2.15) and (3.71), we can explicitly verify

$$\mathscr{P}_{\rm F}^{+}\mathscr{L}^{+}x^{i} = -L^{ik}(\partial_k U_{\mathbf{Y}}) + (\partial_a L^{ia}) - L^{ia}(\partial_a U_{\mathbf{Y}}), \tag{3.73a}$$

$$\mathscr{P}_{\rm F}^{+}(\mathscr{L}^{\dagger})^{+}x^{i} = -L^{li}(\partial_l U_{\mathbf{Y}}) + (\partial_b L^{bi}) - L^{bi}(\partial_b U_{\mathbf{Y}}). \tag{3.73b}$$

Hence Eq. (3.72) can also be rewritten into

$$\delta L^{ij}(\boldsymbol{x}) = -\int_0^\infty dt \int_{\boldsymbol{y}} e^{-U_{\mathbf{Y}}}(e^{t\,\mathscr{L}_{\mathbf{Y}}^{+}}[L^{ik}(\partial_k U_{\mathbf{Y}}) - \partial_a L^{ia} + L^{ia}(\partial_a U_{\mathbf{Y}})])[L^{lj}(\partial_l U_{\mathbf{Y}}) - \partial_b L^{bj} + L^{bj}(\partial_b U_{\mathbf{Y}})]. \tag{3.74}$$

Equations (3.62), (3.64), (3.67), (3.69), (3.72), and (3.74) are all equivalent to each other.

Summing up Eqs. (3.55) and (3.61), we obtain the first-order renormalized FPO of effective slow dynamics, which assumes the covariant form

$$\mathscr{L}_{\mathbf{X}}^{R} = \partial_i(L^{ij} + \epsilon\,\delta L^{ij})(\partial_j + (\partial_j U_{\mathbf{X}})). \tag{3.75}$$

Several important implications can be inferred. Firstly it shows that if the effective slow dynamics is described by covariant Fokker-Planck theory, as long as the joint dynamics is so. Hence the covariant Fokker-Planck theory is the natural formalism for effective slow dynamics of classical nonequilibrium processes. Secondly, we have supplied explicit formulas for the renormalized kinetic coefficients of the effective slow dynamics. These formulas are highly nontrivial and have never been obtained previously. Furthermore, as we have shown above, if the joint dynamics has detailed balance, the effective slow dynamics also has detailed balance. In fact it can be explicitly verified, using, e.g., Eq. (3.74), that the first-order correction to kinetic coefficients satisfy detailed balance conditions,

$$\varepsilon_i \delta L^{ij}(\boldsymbol{x}^*, \lambda^*)\varepsilon_{ji} = \delta L^{ij}(\boldsymbol{x}, \lambda). \tag{3.76}$$

Finally, if we want to take into account short timescale details as in Sec. II E, we should replace the operator $\mathscr{L}_{\mathbf{Y}}^{-\perp}$, which is given by Eq. (3.63), by $\hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}(t_{-1})$ as given by Eq. (2.68) with $t_{-1} = t/\epsilon$. This amounts to replacing the upper bound of the integrals in Eqs, (3.64), (3.67), (3.69), (3.72), and (3.74) by t/ϵ. For example, Eq. (3.72) should be replaced by

$$\delta L^{ij}(\boldsymbol{x};t) = -\int_0^{t/\epsilon} ds \langle e^{s\,\mathscr{L}_{\mathbf{Y}}^{+}}\mathscr{P}_{\rm F}^{+}\mathscr{L}^{+}x^{i},\, \mathscr{P}_{\rm F}^{+}(\mathscr{L}^{\dagger})^{+}x^{j}\rangle_{\mathbf{Y}}. \tag{3.77}$$

IV. HAMILTONIAN SYSTEMS

Consider a Hamiltonian systems with canonical variables $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p})$ interacting with another Hamiltonian system canonical variables $\boldsymbol{y} = (\mathbf{Q}, \mathbf{P})$. A typical example is a colloid interacting with a large number of fluid molecules. $\mathbf{q}, \mathbf{p}$ are then the coordinate and momentum of the colloid, whereas $\mathbf{Q} = (\mathbf{Q}_1, \mathbf{Q}_2, \cdots)$, and $\mathbf{P} = (\mathbf{P}_1, \mathbf{P}_2, \cdots)$ are the shorthands for collections of all coordinates and momenta of fluid particles. The total Hamiltonian is given by

$$H^{\rm tot} = H_{\mathbf{X}}(\boldsymbol{q}, \boldsymbol{p}) + H_{\mathbf{Y}}(\mathbf{Q}, \mathbf{P}, \boldsymbol{q}), \tag{4.1}$$

$$H_{\mathbf{X}}(\boldsymbol{q}, \boldsymbol{p}) = \frac{\boldsymbol{p}^2}{2m} + V_{\mathbf{X}}(\boldsymbol{q}), \tag{4.2}$$

$$H_{\mathbf{Y}}(\mathbf{Q}, \mathbf{P}, \boldsymbol{q}) = \frac{\mathbf{P}^2}{2\mu} + V_{\rm B}(\mathbf{Q}, \boldsymbol{q}), \tag{4.3}$$

where $V_{\mathbf{X}}(\boldsymbol{q})$ is the potential energy of the colloid, which may also be externally controlled, whereas $V_{\rm B}(\mathbf{Q}, \boldsymbol{q})$ is the interaction between fluid particles and also between fluid particles and the colloid.

The dynamics of a generic many-body Hamiltonian system is chaotic and ergodic, which means that the pdf of the total system converges to an equiprobability distribution on the energy hypersurface. It is also well known that if the system is large, an equiprobability distribution is statistically equivalent to a canonical distribution,

$$p(\mathbf{Q}, \mathbf{P}, \boldsymbol{q}, \boldsymbol{p}) = e^{-\beta(H^{\rm tot} - F^{\rm tot})} = e^{-U}, \tag{4.4}$$

$$p_{\mathbf{X}}(\boldsymbol{q}, \boldsymbol{p}) = e^{-\beta(H_{\mathbf{X}} - F_{\mathbf{X}})} = e^{-U_{\mathbf{X}}}, \tag{4.5}$$

where $F^{\rm tot}, F_{\mathbf{X}}$ are the equilibrium free energies of the total system and of the large particle, which are related to the generalized potentials $U, U_{\mathbf{X}}$ via

$$U = \beta(H^{\rm tot} - F^{\rm tot}), \tag{4.6}$$

$$U_{\mathbf{X}} = \beta(H_{\mathbf{X}} - F_{\mathbf{X}}). \tag{4.7}$$

Likewise, the conditional equilibrium pdf of $\boldsymbol{y}$ given $\boldsymbol{x}$ is

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Q}, \mathbf{P}|\boldsymbol{q}, \boldsymbol{p}) = e^{-\beta(H_{\mathbf{Y}} - F_{\mathbf{Y}})} = e^{-U_{\mathbf{Y}}}, \tag{4.8}$$

$$U_{\mathbf{Y}} = \beta(H_{\mathbf{Y}} - F_{\mathbf{Y}}) = U - U_{\mathbf{X}}. \tag{4.9}$$

Note that $p_{\mathbf{Y}|\mathbf{X}}$ and $U_{\mathbf{Y}}$ are independent of $\boldsymbol{p}$. As in Ref. [63] and in Sec. III A, the total Hamiltonian (4.1) is split such that the equilibrium distribution of the colloidal variables has the Gibbs canonical form (4.5).

The Hamiltonian equations of the combined system can be written as the following matrix form:

$$\begin{pmatrix} d\boldsymbol{q} \\ d\boldsymbol{p} \\ d\mathbf{Q} \\ d\mathbf{P} \end{pmatrix} + \begin{pmatrix} 0 & -T & 0 & 0 \\ T & 0 & 0 & 0 \\ 0 & 0 & 0 & -T \\ 0 & 0 & T & 0 \end{pmatrix} \begin{pmatrix} \partial U/\partial \boldsymbol{q} \\ \partial U/\partial \boldsymbol{p} \\ \partial U/\partial \mathbf{Q} \\ \partial U/\partial \mathbf{P} \end{pmatrix} dt = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{4.10}$$

This has the same form as the covariant Langevin equation (3.1), but with vanishing noise amplitudes. Accordingly the kinetic matrix $L^{\mu\nu}$ only has antisymmetric part,

$$\mathbf{L} = \begin{pmatrix} 0 & -T & 0 & 0 \\ T & 0 & 0 & 0 \\ 0 & 0 & 0 & -T \\ 0 & 0 & T & 0 \end{pmatrix}. \tag{4.11}$$

Strictly speaking, 0 and T in the above matrix mean submatrices with various dimensions. But we will not worry about these notational subtleties as they make no difference in the final results. Note that $\mathbf{L}$ is a constant matrix, so the spurious drift $\partial_\mu L^{\mu\nu}$ vanishes identically. The FPO associated with Eq. (4.10) can be obtained,

$$\begin{aligned} \mathscr{L} &= -\frac{\partial H}{\partial \boldsymbol{p}}\frac{\partial}{\partial \boldsymbol{q}} + \frac{\partial H}{\partial \boldsymbol{q}}\frac{\partial}{\partial \boldsymbol{p}} - \frac{\partial H}{\partial \mathbf{P}}\frac{\partial}{\partial \mathbf{Q}} + \frac{\partial H}{\partial \mathbf{Q}}\frac{\partial}{\partial \mathbf{P}} \\ &= -\dot{\boldsymbol{q}}\frac{\partial}{\partial \boldsymbol{q}} - \dot{\boldsymbol{p}}\frac{\partial}{\partial \boldsymbol{p}} - \dot{\mathbf{Q}}\frac{\partial}{\partial \mathbf{Q}} - \dot{\mathbf{P}}\frac{\partial}{\partial \mathbf{P}}, \end{aligned} \tag{4.12}$$

which is a first-order partial differential operator, known the *Liouville operator* in classical mechanics. This operator can be decomposed in the form of Eq. (3.12a),

$$\mathscr{L}_S = -\frac{\partial H}{\partial \boldsymbol{p}}\frac{\partial}{\partial \boldsymbol{q}} + \frac{\partial H}{\partial \boldsymbol{q}}\frac{\partial}{\partial \boldsymbol{p}}, \tag{4.13}$$

$$\mathscr{L}_{\mathbf{Y}} = -\frac{\partial H_{\mathbf{Y}}}{\partial \mathbf{P}}\frac{\partial}{\partial \mathbf{Q}} + \frac{\partial H_{\mathbf{Y}}}{\partial \mathbf{Q}}\frac{\partial}{\partial \mathbf{P}}. \tag{4.14}$$

The operators $\mathscr{L}^\dagger$, $\mathscr{L}^+$, $(\mathscr{L}^\dagger)^+$ can be easily calculated. They are different from $\mathscr{L}$ only by a sign,

$$\mathscr{L} = -\mathscr{L}^\dagger = -\mathscr{L}^+ = (\mathscr{L}^\dagger)^+. \tag{4.15}$$

Similarly we have

$$\mathscr{L}^+_{\mathbf{Y}} = -\mathscr{L}_{\mathbf{Y}}. \tag{4.16}$$

We apply $\mathscr{L}^+$ on q^i, ith component of $\boldsymbol{q}$.

$$\mathscr{L}^+ q^i = -\mathscr{L} q^i = \frac{\partial H}{\partial \boldsymbol{p}^i} = \frac{p_i}{m}. \tag{4.17}$$

We use Eqs. (3.40) to project out the slow component of $\mathscr{L}^+ q^i$,

$$\mathscr{P}^+_S \mathscr{L}^+ q^i = \left\langle \frac{p_i}{m} \right\rangle_Y = \frac{p_i}{m}, \tag{4.18}$$

where the second equality follows from the fact that the conditional equilibrium distribution of the fast variables is independent of the momentum $\boldsymbol{p}$ of the colloid. The fast component of $\mathscr{L}^+ q^i$ is then

$$\mathscr{P}^+_F \mathscr{L}^+ q^i = \mathscr{P}^+_F (\mathscr{L}^\dagger)^+ q^i = 0. \tag{4.19}$$

Let us now calculate $\mathscr{L}^+ p^i$,

$$\begin{aligned} \mathscr{L}^+ p^i &= -\mathscr{L} p^i = \frac{\partial H}{\partial q^i} \\ &= \frac{\partial V_{\mathbf{X}}}{\partial q^i} + \frac{\partial V_{\mathrm{B}}}{\partial q^i}, \end{aligned} \tag{4.20}$$

The physical meaning of this term is negative the force acting on the colloid, which consists of two parts. The first part is due to the external potential, whereas the part is due to the interaction with all fluid particles. The slow component of $\mathscr{L}^+ p^i$ can be similarly calculated,

$$\mathscr{P}^+_S \mathscr{L}^+ p^i = \left\langle \frac{\partial V_{\mathbf{X}}}{\partial q^i} + \frac{\partial V_{\mathrm{B}}}{\partial q^i} \right\rangle_{\mathbf{Y}} = \frac{\partial V_{\mathbf{X}}}{\partial q^i}, \tag{4.21}$$

which is solely due to the external potential. The fast component is

$$\mathscr{P}^+_F \mathscr{L}^+ p^i = -\mathscr{P}^+_F (\mathscr{L}^\dagger)^+ p^i = \frac{\partial V_{\mathrm{B}}}{\partial q^i} = -f^i_{\mathrm{coll}}. \tag{4.22}$$

which is solely due to interaction with fluid particles. In another word, $-\mathscr{P}^+_F \mathscr{L}^+ p^i$ is the collision force acting on the colloid. In a homogeneous fluid, the conditional average of collision force vanishes identically, due to the translational symmetry of the environment.

Substituting the above results into Eq. (3.72), we obtain the first-order renormalization of kinetic coefficients,

$$T\,\gamma^{ij} \equiv \delta L^{p^i p^j} = \int_0^\infty dt \left\langle e^{t\mathscr{L}^+_{\mathbf{Y}}} \frac{\partial V_{\mathrm{B}}}{\partial q^i}, \frac{\partial V_{\mathrm{B}}}{\partial q^j} \right\rangle_{\mathbf{Y}} \tag{4.23}$$

$$= \int_0^\infty dt \left\langle f^i_{\mathrm{coll}}(t), f^j_{\mathrm{coll}}(0) \right\rangle_{\mathbf{Y}}, \tag{4.24}$$

which is the integrated correlation function of collision force acting on the colloid. There is no renormalization of other kinetic coefficients. The corresponding nonlinear Langevin equation can be written as

$$dq^i = \frac{\partial H_{\mathbf{X}}}{\partial p^i} dt, \tag{4.25}$$

$$dp^i = -\frac{\partial H_{\mathbf{X}}}{\partial q^i} dt - \gamma^{ij}\frac{\partial H_{\mathbf{X}}}{\partial p^j} dt + b^{i\alpha} dW_\alpha, \tag{4.26}$$

where γ^{ij} and $b^{i\alpha}$ are related to each other via

$$T\,\gamma^{ij} = 2\, b^{i\alpha} b^{j\alpha}. \tag{4.27}$$

In the above, γ^{ij} is the matrix of friction coefficients of the effective slow dynamics, and reflects the symmetry of the background fluid. If the fluid is homogeneous and isotropic, γ^{ij} is independent of $\boldsymbol{q}$ (and of $\boldsymbol{p}$), and is proportional to the identity matrix. By contrast, the fluid is a nematic liquid crystal with broken orientational order, γ^{ij} is anisotropic. If the colloid is confined near the boundary of the fluid, γ^{ij} depends on the position $\boldsymbol{q}$, which means that the noises acting on the colloid are multiplicative.

Two important results can be inferred here. Firstly covariant Fokker-Planck theory and the associated covariant Ito-Langevin dynamics emerges naturally as the coarse-grained effective theory of unitary Hamiltonian dynamics, in the limit of timescale separation. Secondly, multiplicative noises naturally arise in systems with heterogeneous background.

V. THE HARMONIC BATH MODEL

In this section, we apply the results developed in Sec. IV to the famous toy model of harmonic oscillator bath. We will see that our theory supplies a more natural perspective to the problem.

A. Conventional treatment

Here we use notations in accordance with Zwanzig [5]. The total system consists of a Brownian particle with canonical position x and momentum p, and a bath of harmonic oscillators with coordinates and momenta $\{q_i, p_i\}$. Throughout this section, p refers to the momentum of the Brownian particle, not the pdf of system variables. The total Hamiltonian is decomposed as

$$H^{\text{tot}} = H_{\mathbf{X}} + H_{\mathbf{Y}}, \tag{5.1}$$

$$H_{\mathbf{X}} = \frac{p^2}{2m} + V(x), \tag{5.2}$$

$$H_{\mathbf{Y}} = \sum_j^N \left[\frac{1}{2} p_j^2 + \frac{\omega_j^2}{2} \left(q_j - \frac{\gamma_j}{\omega_j^2} x \right)^2 \right]. \tag{5.3}$$

The Hamiltonian equations are

$$\frac{dx}{dt} = \frac{p}{m}, \tag{5.4a}$$

$$\frac{dp}{dt} = -V'(x) + \sum_j^N \left(\gamma_j q_j - \frac{\gamma_j^2}{\omega_j^2} x \right), \tag{5.4b}$$

$$\frac{dq_i}{dt} = p_i, \tag{5.4c}$$

$$\frac{dp_i}{dt} = -\omega_i^2 q_i + \gamma_i x. \tag{5.4d}$$

In the traditional treatment, one solves for the motion of oscillators in terms of the Brownian particle,

$$q_j(t) = q_j(0)\cos(\omega_j t) + p_j(0)\frac{\sin(\omega_j t)}{\omega_j} + \gamma_j \int_0^t ds \frac{\sin\left(\omega_j (t-s)\right)}{\omega_j} x(s), \tag{5.5}$$

where $q_j(0)$, $p_j(0)$ are the initial data of the oscillators. Integrating by parts, we can rewrite $q_j(t)$ as

$$q_j(t) - \frac{\gamma_j}{\omega_j^2} x(t) = \left(q_j(0) - \frac{\gamma_j}{\omega_j^2} x(0) \right) \cos(\omega_j t) + p_j(0)\frac{\sin(\omega_j t)}{\omega_j} - \gamma_j \int_0^t ds \frac{p(s)}{m} \frac{\cos\left(\omega_j (t-s)\right)}{\omega_j^2}. \tag{5.6}$$

Substituting this back to Eq. (5.4b), we obtain

$$\frac{dp(t)}{dt} = -V'(x(t)) - \int_0^t ds K(s) \frac{p(t-s)}{m} + F_p(t), \tag{5.7}$$

where $F_p(t)$ behaves as colored noises and is defined as

$$F_p(t) = \sum_j \gamma_j p_j(0) \frac{\sin(\omega_j t)}{\omega_j} + \sum_j \gamma_j \left(q_j(0) - \frac{\gamma_j}{\omega_j^2} x(0) \right) \cos(\omega_j t), \tag{5.8}$$

whilst $K(t)$ behaves as the kernel for time-lagged friction,

$$K(t) = \sum_j \frac{\gamma_j^2}{\omega_j^2} \cos(\omega_j t). \tag{5.9}$$

We assume that the oscillator bath is initially in equilibrium conditioned on the Brownian particle, which means that $p_j(0)$, $q_j(0)$ are Gaussian distributed with variances given by

$$\left\langle p_j(0)^2 \right\rangle = k_B T, \tag{5.10}$$

$$\left\langle \left(q_j(0) - \frac{r_j}{\omega_j} x(0) \right)^2 \right\rangle = \frac{T}{\omega_j^2}, \tag{5.11}$$

$$\langle p_j(0) q_i(0) \rangle = 0. \tag{5.12}$$

We can directly verify

$$\langle F_p(t) F_p(t') \rangle = T K(t-t'), \tag{5.13}$$

which is generally known as *the second Fluctuation Theorem Relation*. It can be understood as the condition of detailed balance for non-Markovian Langevin dynamics.

If the frequencies of oscillators are densely distributed, we may approximate the summation in Eq. (5.9) as an integral,

$$K(t) = \int \rho(\omega) \frac{\gamma(\omega)^2}{\omega^2} \cos(\omega t) d\omega. \tag{5.14}$$

By properly choosing the functions $\rho(\omega)$, $\gamma(\omega)$, one may let $K(t)$ approaching a δ function, then Eq. (5.7) becomes the usual underdamped Langevin equation with white noises [5]. In reality, however, we can never fine tune the functions $\rho(\omega)$, $\gamma(\omega)$. Hence the above argument is too restrictive, and does not explain why Langevin dynamics with white noises is so successful in so many different physical systems.

In fact, as we have learned in this paper, the effective dynamics of the Brownian particle approaches the Markov limit if its dynamics is much slower than that of the oscillators. To see this point, we introduce a large parameter α and scale up the oscillator frequencies ω_i and the coefficients γ_i simultaneously,

$$(\omega_j, \gamma_j) \to (\alpha\omega_j, \alpha\gamma_j). \tag{5.15}$$

The kernel $K(t)$ then becomes

$$K(t) \to K_\alpha(t) = \sum_{\omega_j} \frac{\gamma_j^2}{\omega_j^2} \cos(\alpha\omega_j t) = K(\alpha t). \tag{5.16}$$

In the limit $\alpha \to +\infty$, the width of $K_\alpha(t)$ reduces to zero, and the kernel converges to

$$\alpha^{-1} \left[\int_0^{+\infty} dt' \int \rho(\omega) \frac{\gamma(\omega)^2}{\omega^2} \cos(\omega t') d\omega \right] \delta(t) \tag{5.17}$$

in the sense of generalized function. Consequently, $F_p(t)$ behaves as white noises.

B. Multi-scale coarse-graining

Let us apply the rescaling Eq. (5.15) to Hamiltonian equations (5.4) and further introduce rescaled coordinates $\tilde{q} \equiv \alpha q$, we obtain

$$\frac{dx}{dt} = \frac{p}{m}; \tag{5.18a}$$

$$\frac{dp}{dt} = -V'(x) + \sum_j^N \left(\gamma_j \tilde{q}_j - \frac{\gamma_j^2}{\omega_j^2} x \right); \tag{5.18b}$$

$$\frac{d\tilde{q}_i}{dt} = \alpha p_i; \tag{5.18c}$$

$$\frac{dp_i}{dt} = -\alpha \omega_i^2 \tilde{q}_i + \alpha \gamma_i x. \tag{5.18d}$$

These equations are of the form as Eqs. (3.11), if α is identified as $1/\epsilon$, $\{\tilde{q}_i, p_i\}$ are identified as fast variables, and the kinetic matrix identified as

$$\mathbf{L} = \begin{pmatrix} 0 & -T & 0 & 0 \\ T & 0 & 0 & 0 \\ 0 & 0 & 0 & -\alpha T \\ 0 & 0 & \alpha T & 0 \end{pmatrix}, \tag{5.19}$$

and finally, the generalized potentials identified as

$$U = U_{\mathbf{X}} + U_{\mathbf{Y}}, \tag{5.20}$$

$$T\, U_{\mathbf{X}} = \frac{p^2}{2m} + V(x), \tag{5.21}$$

$$T\, U_{\mathbf{Y}} = \sum_j^N \left[\frac{1}{2} p_j^2 + \frac{\omega_j^2}{2} \left(\tilde{q}_j - \frac{\gamma_j}{\omega_j^2} x \right)^2 \right]. \tag{5.22}$$

Hence it becomes clear that the results in the preceding section are fully applicable. In particular, if we apply Eq. (4.23) and work out the correlation function of the conditional fast dynamics, we obtain

$$\begin{aligned} \delta L^{pp} &= \int_0^\infty dt \sum_i \cos(\omega_i t) \frac{\gamma_i^2}{\omega_i^2} \\ &\to \int_0^{+\infty} dt \int \rho(\omega) \frac{\gamma(\omega)^2}{\omega^2} \cos(\omega t) d\omega, \end{aligned} \tag{5.23}$$

which is consistent with Eq. (5.17).

VI. OVERDAMPED LIMIT OF BROWNIAN MOTION

In this section we discuss two aspects of the overdamped limit of Brownian motion. In the first part, we discuss the case where the friction coefficient varies with position. Naive treatment, e.g., throwing out the internal term in the underdamped Langevin equation, fails in this case, because it does not capture the spurious drift. In the second part, we study the influence of the fast scale details on the effective dynamics of the coordinate.

A. Position-dependent friction

We consider a Brownian particle in a heterogeneous background with position dependent friction coefficient. The underdamped Langevin equations are

$$dx = pdt, \tag{6.1a}$$

$$dp = -V'(x)dt - \frac{\gamma(x)pdt}{\epsilon} + \sqrt{\frac{2\gamma(x)T}{\epsilon}} dW(t), \tag{6.1b}$$

where we have set the mass to unity and have divided the friction $\gamma(x)$ by a small parameter ϵ. We are interested in the overdamped limit of the dynamics where $\epsilon \to 0$. It was shown in a preceding paper [62] that this limit is very subtle. The naive treatment, where one simply delete the inertial term dp in Eq. (6.1b) is erroneous. Here we use our multi-scale projection operator theory to carry out a systematic expansion in terms of ϵ. To the first order, we rederive the result of Sec. VI C of Ref. [62], and also in Ref. [69]. These earlier results were obtained using very different methods. To the third order, we discover higher-order derivatives in the renormalized FPO.

Equations (6.1) can be written in the covariant form (3.11) with $\mathbf{L}$ and $e^{-U_{\mathbf{Y}}}$ given by

$$\boldsymbol{L} = \begin{pmatrix} 0 & -T \\ T & \epsilon^{-1}\gamma(x)T \end{pmatrix}, \tag{6.2}$$

$$e^{-U_{\mathbf{Y}}} = \sqrt{\frac{\beta}{2\pi}} e^{-\beta p^2/2}. \tag{6.3}$$

Various FPOs in Eq. (3.12) are

$$\mathscr{L}_{\mathbf{X}} = 0; \tag{6.4a}$$

$$\mathscr{L}_{\mathbf{XY}} = -\partial_x T(\partial_p + \beta p); \tag{6.4b}$$

$$\mathscr{L}_{\mathbf{YX}} = \partial_p T(\partial_x + \beta V'); \tag{6.4c}$$

$$\mathscr{L}_{\mathbf{Y}} = T\gamma(x)\partial_p^2 + \gamma(x)\partial_p p. \tag{6.4d}$$

Using the first-order result Eq. (3.72), we obtain

$$\begin{aligned} \delta L &= \int_0^\infty dt \int_p (-T\partial_p U_{\mathbf{Y}}) e^{t\mathscr{L}_{\mathbf{Y}}} e^{-U_{\mathbf{Y}}} (-T\partial_p U_{\mathbf{Y}}) \\ &= \int_0^\infty dt \int_p p\, e^{t\mathscr{L}_{\mathbf{Y}}} (p e^{-U_{\mathbf{Y}}}) = \frac{1}{\beta\gamma(x)}. \end{aligned} \tag{6.5}$$

Hence the renormalized covariant FPO of the overdamped theory is

$$\mathscr{L}_{\mathrm{FP}}^{od} = \epsilon\, \partial_x \frac{1}{\beta\gamma(x)} (\partial_x + \beta V'(x)), \tag{6.6}$$

which agrees with the results in Ref. [69] and Ref. [62].

We can also apply the multi-scale expansion developed in Sec. II to higher orders. We will simply present the results without supplying details. Due to symmetry reasons, there is no correction to the second order. To the third order, the renormalized FPO is

$$\begin{aligned} \mathscr{L}_{\mathrm{FP}}^{od} &= \epsilon\, \partial_x \frac{T}{\gamma} (\partial_x + \beta V') + \epsilon^3 \partial_x \frac{T}{\gamma^3} V''(\partial_x + \beta V') \\ &\quad + \epsilon^3 \partial_x \frac{T}{\gamma} \left(\partial_x \frac{1}{\gamma} \right) (\partial_x + \beta V') \frac{1}{\gamma} (\partial_x + \beta V'). \end{aligned} \tag{6.7}$$

The last term contains third-order differential operator.

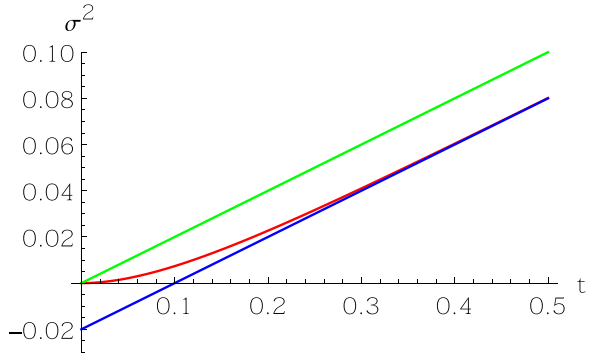

FIG. 1. The variance $\sigma^2(t)$ of coordinate, starting from initial condition Eq. (6.9). Red: Exact result Eq. (6.10) of the underdamped theory. Blue: long time asymptotics of Eq. (6.10); green: overdamped result Eq. (6.14). The temperature and friction coefficient are set as $T = 1$ and $\gamma = 10$.

B. Short-time analysis of overdamped theory of Brownian motion

In this part, we discuss the short-time details of the overdamped Langevin equation, for the case where γ is independent of position. We shall first discuss the simpler case of free particle, i.e., $V = 0$, and then come back to the case $V \neq 0$. The mass is again set to unity. We also absorb $1/\epsilon$ into γ, so γ becomes a large parameter.

The underdamped Fokker-Planck equation, popularly known as *the Kramers equation*, is given by

$$\partial_t \rho(x, p, t) = \big(T\gamma\, \partial_p^2 + \gamma\, \partial_p p - p\, \partial_x\big)\rho(x, p, t). \tag{6.8}$$

Note that we use ρ instead of p to denote the pdf, to avoid confusion with the momentum variable. We are interested in the initial conditions where x is completely fixed, and p is Maxwellian,

$$\rho(x, p, 0) = \sqrt{\frac{\beta}{2\pi}} e^{-\beta p^2/2}\delta(x). \tag{6.9}$$

The exact solution to Eq. (6.8) is known,

$$\rho(x, p, t) = \frac{\gamma \exp\big(-\frac{\gamma^2x^2+2p^2(\gamma t-1)-2\gamma px+2pe^{-\gamma t}(\gamma x+p)}{2T(2\gamma t-3+4e^{-\gamma t}-e^{-2\gamma t})}\big)}{2\pi T\sqrt{2\gamma t - 3 + 4e^{-\gamma t} - e^{-2\gamma t}}}. \tag{6.10a}$$

This result captures both the long-timescale properties and the short-timescale properties correctly. From this, we can easily derive the marginal pdf of x,

$$\rho_X(x, t) = \frac{1}{\sqrt{2\pi\sigma(t)^2}} e^{-x^2/2\sigma(t)^2}; \tag{6.10b}$$

$$\sigma^2(t) = \frac{2T}{\gamma^2}(\gamma t - 1 + e^{-\gamma t}). \tag{6.10c}$$

Hence x is a Gaussian variable with variance $\sigma^2(t)$ monotonically increasing with time, which is plotted in Fig. 1. Its short-time asymptotics is

$$\sigma^2(t) \sim T\left(t^2 - \frac{\gamma t^3}{3}\right), \quad \gamma t \ll 1; \tag{6.11}$$

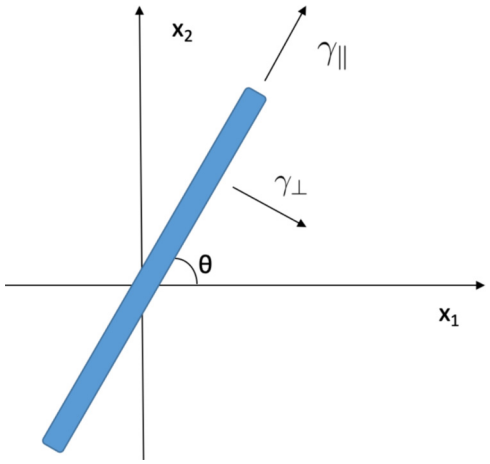

FIG. 2. Schematics of Brownian rod.

which clearly corresponds to ballistic motion. The long-time asymptotics is

$$\sigma^2(t) \sim T\left(\frac{2t}{\gamma} - \frac{2}{\gamma^2}\right), \quad \gamma t \gg 1, \tag{6.12}$$

which corresponds to normal diffusion.

For the case we are studying, the overdamped FP operator (6.7) becomes the usual Laplacian. The overdamped FPE then becomes

$$\partial_t \rho_X = \frac{T}{\gamma}\partial_x^2 \rho. \tag{6.13}$$

It yields a Gaussian solution with variance, which is proportional to time,

$$\sigma^2(t) = \frac{2Tt}{\gamma}. \tag{6.14}$$

At $t = 0$, the variance vanishes identically, hence $\rho_X(0) = \delta(x)$. As illustrated in Fig. 1, this differs from Eq. (6.12) by a constant $-2T/\gamma^2$. The difference is evidently due to the ballistic motion in the short timescales, which is captured by the underdamped theory but not the overdamped theory. To capture this effect in the overdamped theory, we may use a "renormalized initial condition" $\sigma_R^2(0) = -2T/\gamma^2$. The value of $\sigma_R^2(0)$, however, can be determined by comparing the long-time asymptotics of two theories.

To capture the short-time behaviors in the overdamped theory, we need to use the formalism developed in Sec. II E, where the fast timescale t_{-1} is retained. The operator $\hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp}$, defined in Eq. (2.68), can be calculated exactly, which is presented below without supplying derivations. Let $f(p)$ be an arbitrary function, we have

$$\hat{\mathscr{L}}_{\mathbf{Y}}^{-\perp} f(p') = \int G(p, t_{-1}|p', t')f(p')dp'; \tag{6.15}$$

$$G(p, t_{-1}|p', t') = \frac{1}{\sqrt{2\pi T(1 - e^{-2\gamma(t_{-1}-t')})}} \exp -\frac{1}{2T}\left[\frac{p - p'e^{-\gamma(t_{-1}-t')}}{1 - e^{-2\gamma(t_{-1}-t')}}\right]^2. \tag{6.16}$$

We substitute this back into Eq. (2.83), and obtain the first-order renormalized Fokker-Planck equation,

$$\partial_t \rho_X(x) = \epsilon(1 - e^{-\gamma t/\epsilon})\frac{T}{\gamma}\partial_x^2 \rho_X. \tag{6.17}$$

It is interesting to note that Eq. (6.10b) is actually the exact solution of Eq. (6.17). Hence at least for the case of free Brownian particle, our projection operator theory developed in Sec. II E captures the short-time details correctly. Equation (6.17) was also derived in [18] using the time-convolutionless projection operator formalism.

We can in fact extend the multi-scale method to more general case, where the potential is nonzero, and the initial momentum distribution is an arbitrary Gaussian,

$$\rho(x, p, 0) = \sqrt{\frac{1}{2\pi\sigma(0)^2}}e^{-p^2/2\sigma(0)^2}\delta(x), \tag{6.18}$$

After a much more complicated analysis, we find the following FPE for the marginal pdf $p_X(x, t)$:

$$\begin{aligned}\partial_t \rho_X &= \epsilon\frac{1 - e^{-\gamma t/\epsilon}}{\gamma}\partial_x(\sigma^2(t/\epsilon)\partial_x + V')\rho_X;\\ \sigma(t/\epsilon)^2 &= T(1 - e^{-2\gamma t/\epsilon}(1 - \sigma(0)^2/T)).\end{aligned} \tag{6.19}$$

This equation is homogeneous, and give correct results both in short timescale and in long timescale, up to the first order in ϵ. Such a result has never been derived previously, to the best of our knowledge.

VII. BROWNIAN ROD

In Sec. III, we assumed that the kinetic matrix for the slow variables L^{ij} in Eq. (3.12c) does not depend on the fast variables $\mathbf{Y}$. This assumption makes it possible to derive an explicit expression for the first-order renormalized FPO. In this section, we discuss a simple example where this assumption does not hold, and hence the results of Sec. III are no longer applicable. The method developed in Sec. II, however, remains applicable. We will see that the resultant first-order renormalized FPO of the slow variables contains higher-order derivatives at the first order.

We study the Brownian dynamics of a two dimensional symmetric rod in the overdamped regime, as illustrated in Fig. 2. The coefficient of friction is $\gamma_\parallel$ along its long axis and $\gamma_\perp$ in the perpendicular direction. Let (x_1, x_2) be the coordinates of the center of the mass of the rod, and θ the angle between the rod axis and x_1 coordinate axis. The overdamped Langevin equations are

$$\gamma_{ij}(\theta)dx_j + \partial_i V(\boldsymbol{x})dt = b^{i\alpha}dW_\alpha, \tag{7.1a}$$

$$\gamma_\theta\, d\theta = b^{\theta\alpha}dW_\alpha, \tag{7.1b}$$

where $\partial_i \equiv \partial_{x_i}$ are the spatial derivatives. The matrix of spatial friction coefficients γ_{ij} is given by

$$(\gamma_{ij}) = \begin{pmatrix} r\gamma + \Delta\gamma\cos(2\theta)/2 & -\Delta\gamma\sin(2\theta)/2 \\ -\Delta\gamma\sin(2\theta)/2 & \bar{\gamma} - \Delta\gamma\cos(2\theta)/2 \end{pmatrix}, \tag{7.2}$$

with $\bar{\gamma} = (\gamma_\parallel + \gamma_\perp)/2$ and $\Delta\gamma = \gamma_\parallel - \gamma_\perp$. The rotational friction coefficient is assumed to be constant. The amplitudes of noises satisfy $b^{i\alpha}b^{j\alpha} = 2T\gamma_{ij}$ and $b^{\theta\alpha}b^{\theta\alpha} = 2T\gamma_\theta$. It is assumed that the potential $V(x_1, x_2)$ is independent of θ. The corresponding FPO is

$$\mathscr{L} = T\left(\gamma_{ij}^{-1}\partial_i(\partial_j + \beta(\partial_j V)) + \gamma_\theta^{-1}\partial_\theta^2\right). \tag{7.3}$$

We consider the limit where γ_θ is much smaller than $\gamma_\parallel$ and $\gamma_\perp$ so that the rotational dynamics is much faster than the translational dynamics. (We do not worry how such a limit can be realized in an experimental system. This will be addressed in a future publication.) For convenience we introduce a small parameter ϵ such that the FPO (7.3) assumes the form of Eq. (2.2),

$$\mathscr{L} = \mathscr{L}_S + \epsilon^{-1}\mathscr{L}_{\mathbf{Y}}. \tag{7.4a}$$

$$\mathscr{L}_{\mathbf{Y}} = T\gamma_\theta^{-1}\partial_\theta^2 \tag{7.4b}$$

$$\mathscr{L}_S = T\gamma_{ij}^{-1}\partial_i(\partial_j + \beta(\partial_j V)). \tag{7.4c}$$

Note that the conditional steady state $e^{-U_{\mathbf{Y}}}$ is flat. The methods developed in Sec. II then becomes applicable.

Using Eqs. (2.44) and (2.50), we obtain the renormalized FPO to the first order,

$$\begin{aligned}\mathscr{L}_{\mathbf{X}}^R &= \int_\theta \mathscr{L}_S\, e^{-U_\theta} - \epsilon\int_\theta \mathscr{L}_S\mathscr{L}_{\mathbf{Y}}^{-\perp}\mathscr{L}_S\, e^{-U_\theta}\\ &= T\frac{\gamma_\parallel + \gamma_\perp}{2\gamma_\parallel\gamma_\perp}(\partial_1(\partial_1 + \beta(\partial_1 V)) + \partial_2(\partial_2 + \beta(\partial_2 V)))\\ &\quad + \frac{\epsilon T\gamma_\theta\Delta\gamma^2}{32\gamma_\parallel^2\gamma_\perp^2}[(\partial_1(\partial_1 + \beta(\partial_1 V)) + \partial_2(\partial_2 + \beta(\partial_2 V)))^2\\ &\quad + \partial_k\delta L_{kl}(\partial_l + \beta(\partial_l V)) + \beta^2((\partial_1 V)\partial_2 - (\partial_2 V)\partial_1)^2].\end{aligned} \tag{7.5}$$

where the matrix δL_{ij} is defined as

$$(\delta L_{ij}) = 2\beta\begin{pmatrix} -\partial_2^2 V & \partial_1\partial_2 V \\ \partial_1\partial_2 V & -\partial_1^2 V \end{pmatrix}. \tag{7.6}$$

For the case $V = 0$, Eq. (7.5) reduces to

$$\mathscr{L}_{\mathbf{X}}^R = T\frac{\gamma_\parallel + \gamma_\perp}{2\gamma_\parallel\gamma_\perp}\left(\partial_1^2 + \partial_2^2\right) + \frac{\epsilon T\gamma_\theta\Delta\gamma^2}{32\gamma_\parallel^2\gamma_\perp^2}\left(\partial_1^2 + \partial_2^2\right)^2.$$

Note that there are fourth-order derivatives at the first order of ϵ.

VIII. CONCLUSIONS

In the present paper, we have shown that the covariant formulation of Fokker-Planck theory naturally emerges as a consequence of coarse-graining of microscopic unitary dynamics. Furthermore, even though the fast variables necessarily deviate from equilibrium, detailed balance is preserved by the process of coarse-graining. The kinetic matrix of the effective slow dynamics is expressed in terms of integrated correlation functions of fast variables, conditioned on the slow variables. These results not only establish the connection between reversible dynamics at the microscopic level and stochastic irreversible dynamics at the mesoscopic level, but also demonstrate that the covariant Fokker-Planck theory (and the associated covariant Ito-Langevin dynamics) is the universal description for all continuous classical nonequilibrium small systems. In the future, we will apply this method

to study effective slow dynamics of many concrete nonequilibrium classical systems, including those with biological significances.

ACKNOWLEDGMENT

M.D. thanks Chenxing Post Doctoral Incentive Project for financial support. X.X. acknowledges support from NSFC Grant No. 1674217, and Shanghai Municipal Science and Technology Major Project (Grant No. 2019SHZDZX01). X.X. also thanks Wang Haitao, Liu Fei, and Lan Yueheng, as well as some other members of Online Club Nanothermodynamica for interesting discussions.

APPENDIX: MATHEMATICAL STRUCTURE OF FOKKER-PLANCK DYNAMICS

In this Appendix, we discuss some mathematical structures of Fokker-Planck dynamics, which will be useful for our analysis. We shall assume that the Fokker Planck operator $\mathscr{L}$ is fixed, with $e^{-U(\mathbf{z})}$ the steady-state pdf.

1. Hilbert spaces of distributions and observables

Here we elaborate on the inner products defined previously in Eqs. (3.26).

In the study of Fokker-Planck dynamics, there are two distinct types of functions. We define an inner product of two functions $\psi(\mathbf{z})$, $\phi(\mathbf{z})$ as [c.f. Eq. (3.26a)]

$$(\psi(\mathbf{z}), \phi(\mathbf{z})) \equiv \int d\mathbf{z}\, \psi(\mathbf{z})\, \phi(\mathbf{z})\, e^{U(\mathbf{z})}. \tag{A1}$$

A *distribution* $\psi(\mathbf{z})$ is a function, which satisfies

$$(\psi, \psi) = \int_{\mathbf{z}} \psi(\mathbf{z})^2\, e^{U(\mathbf{z})} < \infty. \tag{A2}$$

All distributions form a real Hilbert space, which we will call $\mathscr{H}_{\rm dist}$. According to this definition, a distribution is generically neither normalized nor positive. The set of all positive normalized distribution (which has the meaning of pdf) form a convex subset of $\mathscr{H}_{\rm dist}$.

We define another inner product between functions [c.f. Eq. (3.26b)],

$$\langle A(\mathbf{z}), B(\mathbf{z})\rangle \equiv \int d\mathbf{z}\, A(\mathbf{z}) B(\mathbf{z})\, e^{-U(\mathbf{z})}. \tag{A3}$$

An observable $A(\mathbf{z})$ is a function, which satisfies

$$(A, A) = \int_{\mathbf{z}} A(\mathbf{z})^2\, e^{-U(\mathbf{z})} < \infty. \tag{A4}$$

All observable form a real Hilbert space, which we will call $\mathscr{H}_{\rm obs}$. Physical quantities such as Hamiltonian, momenta, coordinates, etc., are all observables.

We need to distinguish $\mathscr{H}_{\rm obs}$ and $\mathscr{H}_{\rm dist}$ because the phase space of classical Hamiltonian dynamics is unbounded. (Momenta are always unbounded. Furthermore, coordinates are also unbounded if we consider free boundary conditions.) In general, the effective potential $U(\mathbf{z})$ is bounded from below but not from above. Hence $\mathscr{H}_{\rm dist}$ is a subspace of $\mathscr{H}_{\rm obs}$ but not vice versa. There is a 1–1 correspondence between two Hilbert spaces,

$$A(\mathbf{z}) \in \mathscr{H}_{\rm obs} \quad \longleftrightarrow \quad A(\mathbf{z}) e^{-U(\mathbf{z})} \in \mathscr{H}_{\rm dist}, \tag{A5}$$

which preserves the inner product

$$\langle A, B\rangle = (A\, e^{-U}, B\, e^{-U}). \tag{A6}$$

It is also useful to define the "mixed product" between an observable and a distribution as [c.f. Eq. (3.26c)]

$$\langle A, \psi\rangle \equiv \int d\mathbf{z}\, A(\mathbf{z}) \psi(\mathbf{z}), \tag{A7}$$

we see that each distribution $\phi(\mathbf{z}) \in \mathscr{H}_{\rm dist}$ can be understood as a linear functional on $\mathscr{H}_{\rm obs}$ and each $A(\mathbf{z}) \in \mathscr{H}_{\rm obs}$ can be understood as a linear functional on $\mathscr{H}_{\rm dist}$. In this sense, $\mathscr{H}_{\rm obs}$ and $\mathscr{H}_{\rm dist}$ are *dual spaces* to each other. The following relations can be easily proved:

$$\langle A, B\rangle = \langle A, B\, e^{-U}\rangle = (A\, e^{-U}, B\, e^{-U}), \tag{A8a}$$

$$(\phi, \psi) = \langle \phi\, e^{U}, \psi\rangle = \langle \phi\, e^{U}, \psi\, e^{U}\rangle, \tag{A8b}$$

$$\langle A, \psi\rangle = \langle A, \psi\, e^{U}\rangle = (A\, e^{-U}, \psi). \tag{A8c}$$

If $p(\mathbf{z})$ is a pdf, then Eq. (A7) becomes the average of $A(\mathbf{z})$ in the state described by the pdf $p(\mathbf{z})$,

$$\langle A, p\rangle \equiv \int d\mathbf{z}\, A(\mathbf{z}) p(\mathbf{z}). \tag{A9}$$

2. Fokker-Planck and related operators

Here we elaborate on the † conjugate and + conjugate of operators defined previously in Eqs. (3.27).

The covariant FPO is [c.f. Eq. (3.2b)]

$$\mathscr{L} \equiv \partial_\mu L^{\mu\nu}(\partial_\nu + (\partial_\nu U)) = \partial_\mu L^{\mu\nu} e^{-U} \partial_\nu e^{U}, \tag{A10}$$

which is understood as an operator in the space of distribution $\mathscr{H}_{\rm dist}$. According to Eqs. (3.27a), its Hermitian conjugate $\mathscr{L}^\dagger$ is defined using the inner product as

$$(\psi, \mathscr{L}^\dagger \phi) = (\mathscr{L}\psi, \phi). \tag{A11}$$

Hence both $\mathscr{L}$ and $\mathscr{L}^\dagger$ are operators in $\mathscr{H}_{\rm dist}$. Simple calculation leads to [c.f. Eq. (3.32b)]

$$\mathscr{L}^\dagger = \partial_\mu L^{\nu\mu}(\partial_\nu + (\partial_\nu U)) = \partial_\mu L^{\nu\mu} e^{-U} \partial_\nu e^{U}. \tag{A12}$$

Note that the steady-state pdf e^{-U} is a common eigenfunction of $\mathscr{L}$ and $\mathscr{L}^\dagger$ with eigenvalue zero,

$$\mathscr{L}\, e^{-U} = \mathscr{L}^\dagger e^{-U} = 0. \tag{A13}$$

Generically $L^{\mu\nu}$ is neither symmetric or anti-symmetric, hence $\mathscr{L}$ is neither Hermitian nor anti-Hermitian, and may not have a complete set of eigenfunctions.

The matrix $L^{\mu\nu}$ can be decomposed as follows:

$$L^{\mu\nu} = B^{\mu\nu} + Q^{\mu\nu}, \tag{A14a}$$

$$B^{\mu\nu} = B^{\nu\mu}, \tag{A14b}$$

$$Q^{\mu\nu} = -Q^{\nu\mu}. \tag{A14c}$$

Hence $\mathscr{L}$, $\mathscr{L}^\dagger$, as given by Eqs. (A10) and (A12), can be decomposed into a Hermitian part and an anti-Hermitian part,

$$\mathscr{L} = \partial_\mu B^{\mu\nu}(\partial_\nu + \partial_\nu U) + \partial_\mu Q^{\mu\nu}(\partial_\nu + \partial_\nu U), \tag{A15a}$$

$$\mathscr{L}^\dagger = \partial_\mu B^{\mu\nu}(\partial_\nu + \partial_\nu U) - \partial_\mu Q^{\mu\nu}(\partial_\nu + \partial_\nu U). \tag{A15b}$$

According to Eqs. (3.27b) and (3.27c), $\mathscr{L}^+$, $(\mathscr{L}^\dagger)^+$ as operators in $\mathscr{H}_{\rm obs}$ are defined as

$$\langle \mathscr{L}^+ A, p\rangle = \langle A, \mathscr{L} p\rangle, \tag{A16a}$$

$$\langle (\mathscr{L}^\dagger)^+ A, p\rangle = \langle A, \mathscr{L}^\dagger p\rangle. \tag{A16b}$$

Simple calculation leads to [c.f. Eqs. (3.32c) and (3.32d)]

$$\mathscr{L}^+ \equiv (\partial_\mu - \partial_\mu U)L^{\nu\mu}\partial_\nu, \tag{A16c}$$

$$(\mathscr{L}^\dagger)^+ \equiv (\partial_\mu - \partial_\mu U)L^{\mu\nu}\partial_\nu. \tag{A16d}$$

It is easy to see that constant functions are common eigenfunctions of $\mathscr{L}^+$ and $(\mathscr{L}^\dagger)^+$ with eigenvalue zero,

$$\mathscr{L}^+ 1 = (\mathscr{L}^\dagger)^+ 1 = 0. \tag{A17}$$

Through direct calculation, it is easy to verify

$$\mathscr{L}^+ = e^U \mathscr{L}^\dagger e^{-U}, \tag{A18a}$$

$$(\mathscr{L}^\dagger)^+ = e^U \mathscr{L} e^{-U}. \tag{A18b}$$

Using Eqs. (A8) and (A16a), we further prove

$$\begin{aligned}\langle \mathscr{L}^+ A, B\rangle &= \langle \mathscr{L}^+ A, e^{-U}B\rangle = \langle A, \mathscr{L}\, e^{-U}B\rangle \\ &= \langle A, e^{-U}e^U \mathscr{L}\, e^{-U}B\rangle \\ &= \langle A, e^{-U}(\mathscr{L}^\dagger)^+ \mathscr{L} B\rangle \\ &= \langle A, (\mathscr{L}^\dagger)^+ B\rangle.\end{aligned} \tag{A19}$$

Finally using Eqs. (A16c) and (A16d) it is straightforward to verify

$$\mathscr{L}^+ x^\mu = -L^{\mu\nu}(\partial_\nu U) + (\partial_\nu L^{\mu\nu}), \tag{A20a}$$

$$(\mathscr{L}^\dagger)^+ x^\mu = -L^{\nu\mu}(\partial_\nu U) + (\partial_\nu L^{\nu\mu}). \tag{A20b}$$

Using Eq. (A20a), the covariant Langevin equation (3.1) can be written into the following compact form:

$$dx^\mu - (\mathscr{L}^+ x^\mu)dt = b^{\mu\alpha}dW_\alpha. \tag{A21}$$

3. Decomposition of $\mathscr{H}_{\rm dist}$

It is known that two operators that are Hermitian conjugates of each other, such as $\mathscr{L}$ and $\mathscr{L}^\dagger$, share the same set of eigenvalues. On the other hand, from Eq. (A18) one easily see that $\mathscr{L}$ and $(\mathscr{L}^\dagger)^+$, as well as $\mathscr{L}^\dagger$ and $\mathscr{L}^+$, also share the same set of eigenvalues. Hence all four operators $\mathscr{L}$, $\mathscr{L}^\dagger$, $\mathscr{L}^+$, $(\mathscr{L}^\dagger)^+$ share the same set of eigenvalues. They have a common eigenvalue zero. These operators cannot have positive eigenvalue. Otherwise, the solution to the Fokker-Planck equation would blow up, which leads to violation of either positivity or normalization, both of which are necessary for $p(\mathbf{z}, t)$ to be interpreted as the pdf of $\mathbf{z}$.

According to our assumption, the zero eigenvalue of $\mathscr{L}$ is nondegenerate, i.e., $\mathscr{L}$ has no stationary state other than e^{-U}. This excludes all systems with spontaneous broken symmetry. It then follows that the *kernel* of $\mathscr{L}$ is one dimensional, and is expanded by e^{-U}. We shall use $\mathscr{H}^0_{\rm dist}$ to denote this subspace of $\mathscr{H}_{\rm dist}$,

$$\begin{aligned}{\rm Ker}\mathscr{L} = \mathscr{H}^0_{\rm dist} &= \{\psi \in \mathscr{H}_{\rm dist} | \mathscr{L}\psi = 0\} \\ &= \{c\, e^{-U} | c \text{ is real}\}.\end{aligned} \tag{A22}$$

Now define the subspace $\mathscr{H}^\perp_{\rm dist}$ as the orthogonal complement of $\mathscr{H}^0_{\rm dist}$,

$$\mathscr{H}^\perp_{\rm dist} = \{\psi \in \mathscr{H}_{\rm dist} | (\psi, e^{-U}) = 0\}, \tag{A23}$$

where the inner product is defined in Eq. (A1). Hence the space $\mathscr{H}_{\rm dist}$ is the direct sum of $\mathscr{H}^0_{\rm dist}$ and $\mathscr{H}^\perp_{\rm dist}$,

$$\mathscr{H}^0_{\rm dist} \oplus \mathscr{H}^\perp_{\rm dist} = \mathscr{H}_{\rm dist}. \tag{A24}$$

Any distribution $\psi(\mathbf{z})$ can be uniquely decomposed into a part in $\mathscr{H}^0_{\rm dist}$ and another part in $\mathscr{H}^\perp_{\rm dist}$,

$$\psi(\mathbf{z}) = c\, e^{-U(\mathbf{z})} + \psi^\perp. \tag{A25}$$

Note that $c = (\psi, e^{-U})$. If ψ is a normalized distribution, $c = 1$, and $\psi^\perp$ describes fluctuation away from the steady state.

Now consider a nonvanishing $\psi^\perp \in \mathscr{H}^\perp_{\rm dist}$, we have $\mathscr{L}\psi^\perp \neq 0$, since $\psi^\perp$ is not a stationary state. On the other hand, using Eqs. (A12) and (A13), we easily see

$$(\mathscr{L}\psi^\perp, e^{-U}) = (\psi^\perp, \mathscr{L}^\dagger e^{-U}) = (\psi^\perp, 0) = 0, \tag{A26}$$

which means $\mathscr{L}\psi^\perp$ is orthogonal to e^{-U}, i.e., $\mathscr{L}\psi^\perp \in \mathscr{H}^\perp_{\rm dist}$. Hence $\mathscr{H}^\perp_{\rm dist}$ is an invariant subspace of $\mathscr{L}$. The kernel $\mathscr{H}^0_{\rm dist}$ is of course also an invariant subspace of $\mathscr{L}$.

Let $\mathscr{P}$, $\mathscr{P}_\perp$ be the projection operators onto $\mathscr{H}^0_{\rm dist}$ and $\mathscr{H}^\perp_{\rm dist}$ respectively. For any $\psi(\mathbf{z}) \in \mathscr{H}_{\rm dist}$, we have

$$\begin{aligned}\mathscr{P}\psi(\mathbf{z}) &= (\psi, e^{-U})\, e^{-U(\mathbf{z})} \\ &= e^{-U(\mathbf{z})}\int_{\mathbf{z}'} \psi(\mathbf{z}'),\end{aligned} \tag{A27a}$$

$$\begin{aligned}\mathscr{P}_\perp\psi(\mathbf{z}) &= \psi(\mathbf{z}) - \mathscr{P}\psi(\mathbf{z}) \\ &= \psi(\mathbf{z}) - e^{-U(\mathbf{z})}\int_{\mathbf{z}'} \psi(\mathbf{z}').\end{aligned} \tag{A27b}$$

For any ψ given in Eq. (A25), we have

$$\mathscr{P}\psi(\mathbf{z}) = c(\psi)\, e^{-U(\mathbf{z})}, \tag{A28a}$$

$$\mathscr{P}_\perp\psi(\mathbf{z}) = \psi^\perp(\mathbf{z}). \tag{A28b}$$

In another word, $\mathscr{P}$ projects out the component of steady-state distribution, whereas $\mathscr{P}_\perp$ projects out fluctuations away from steady state.

For any two distributions ψ, ϕ, it is easy to verify

$$(\phi, \mathscr{P}\psi) = (\mathscr{P}\phi, \psi) = (\mathscr{P}\phi, \mathscr{P}\psi), \tag{A29a}$$

$$(\phi, \mathscr{P}_\perp\psi) = (\mathscr{P}_\perp\phi, \psi) = (\mathscr{P}_\perp\phi, \mathscr{P}_\perp\psi), \tag{A29b}$$

which imply that $\mathscr{P}$, $\mathscr{P}_\perp$ are both Hermitian operators. Summarizing, $\mathscr{P}$, $\mathscr{P}_\perp$ satisfy all standard properties of projection operators,

$$\mathscr{P}^2 = \mathscr{P}, \quad \mathscr{P}_\perp^2 = \mathscr{P}_\perp, \tag{A30a}$$

$$\mathscr{P} + \mathscr{P}_\perp = I, \tag{A30b}$$

$$\mathscr{P}^\dagger = \mathscr{P}, \quad \mathscr{P}_\perp^\dagger = \mathscr{P}_\perp, \tag{A30c}$$

$$\mathscr{P}\mathscr{P}_\perp = \mathscr{P}_\perp\mathscr{P} = 0. \tag{A30d}$$

4. Solubility condition

Using Eq. (A30b), $\mathscr{L}$ can be decomposed as

$$\begin{aligned}\mathscr{L} &= (\mathscr{P} + \mathscr{P}_\perp)\mathscr{L}(\mathscr{P} + \mathscr{P}_\perp) \\ &= \mathscr{P}\mathscr{L}\mathscr{P} + \mathscr{P}\mathscr{L}\mathscr{P}_\perp + \mathscr{P}_\perp\mathscr{L}\mathscr{P} + \mathscr{P}_\perp\mathscr{L}\mathscr{P}_\perp.\end{aligned} \tag{A31}$$

Since $\mathscr{H}^0_{\rm dist}$, $\mathscr{H}^\perp_{\rm dist}$ are both invariant subspaces of $\mathscr{L}$, two cross terms in Eq. (A31) vanish identically: $\mathscr{P}\mathscr{L}\mathscr{P}_\perp =$

$\mathscr{P}_\perp \mathscr{L} \mathscr{P} = 0$. Also since $\mathscr{L} e^{-U} = 0$, we have $\mathscr{P} \mathscr{L} \mathscr{P} = 0$. Hence we have

$$\mathscr{L} = \mathscr{P}_\perp \mathscr{L} \mathscr{P}_\perp \equiv \mathscr{L}^\perp. \tag{A32}$$

Further recall that $\mathscr{L}$ is nonsingular in $\mathscr{H}^\perp_{\rm dist}$, we can define an operator $\mathscr{L}^{-\perp}$ as

$$\mathscr{L}^{-\perp} \equiv \mathscr{P}_\perp \mathscr{L}^{-1} \mathscr{P}_\perp. \tag{A33}$$

When acting on the distribution (A25), it returns

$$\mathscr{L}^{-\perp}(c\, e^{-U} + \psi^\perp) = \mathscr{L}^{-1} \psi^\perp \in \mathscr{H}^\perp_{\rm dist}. \tag{A34}$$

Let us now consider a linear equation,

$$\mathscr{L}(\mathbf{z}) p(\mathbf{z}) = \psi(\mathbf{z}) = c(\psi)\, e^{-U(\mathbf{z})} + \psi^\perp(\mathbf{z}). \tag{A35}$$

It is a partial differential equation, since $\mathscr{L}$ is an partial differential operator. It has solutions if and only if the following *solubility condition* is satisfied:

$$c = (\psi, e^{-U}) = \int_{\mathbf{z}} \psi(\mathbf{z}) = 0, \tag{A36a}$$

or equivalently

$$\mathscr{P} \psi = 0, \tag{A36b}$$

i.e., ψ must be orthogonal to the steady-state distribution e^{-U}. With this condition satisfied, we have $\psi = \psi^\perp$, and the general solution to Eq. (A35) can be written as

$$p(\mathbf{z}) = c'\, e^{-U(\mathbf{z})} + \mathscr{L}^{-\perp}(\mathbf{z}) \psi^\perp(\mathbf{z}), \tag{A37}$$

where the first term in r.h.s. is the homogeneous solution with c' an arbitrary constant, whereas the second term the particular solution.

5. Decomposition of $\mathscr{H}_{\rm obs}$

There is also a decomposition of the Hilbert space $\mathscr{H}_{\rm obs}$ of observables. We define projection operators $\mathscr{P}^+$, $\mathscr{P}^+_\perp$ using the following relations:

$$(\mathscr{P}^+ A, \psi) = (A, \mathscr{P} \psi), \tag{A38a}$$

$$(\mathscr{P}^+_\perp A, \psi) = (A, \mathscr{P}_\perp \psi). \tag{A38b}$$

These operators act on observables, rather than on distributions. It is straightforward to verify

$$\mathscr{P}^+ A(\mathbf{z}) \equiv (A, e^{-U}) = \int_{\mathbf{z}'} A(\mathbf{z}') e^{-U(\mathbf{z}')}, \tag{A39a}$$

$$\mathscr{P}^+_\perp A(\mathbf{z}) \equiv (I - \mathscr{P}^+) A(\mathbf{z}) = A(\mathbf{z}) - (A, e^{-U}). \tag{A39b}$$

Hence $\mathscr{P}^+ A(\mathbf{z})$ gives the steady-state average of $A(\mathbf{z})$ (which is a constant independent of $\mathbf{z}$), whereas $\mathscr{P}^+_\perp A(\mathbf{z})$ gives the fluctuations. The subspaces $\mathscr{H}^0_{\rm obs}$ and $\mathscr{H}^\perp_{\rm obs}$ are defined as

$$\mathscr{H}^0_{\rm obs} = \mathscr{P}^+ \mathscr{H}_{\rm obs} = \{c | c \text{ is real}\}, \tag{A40a}$$

$$\mathscr{H}^\perp_{\rm obs} = \mathscr{P}^+_\perp \mathscr{H}_{\rm obs} = \{\psi \in \mathscr{H}_{\rm obs} | (\psi, e^{-U}) = 0\}. \tag{A40b}$$

Combining Eqs. (A27), (A39), and (A39), we can verify the following operator identities:

$$\mathscr{P} = e^{-U} \mathscr{P}^+ e^U, \quad \mathscr{P}_\perp = e^{-U} \mathscr{P}^+_\perp e^U, \tag{A41a}$$

$$\mathscr{P}^+ = e^U \mathscr{P} e^{-U}, \quad \mathscr{P}^+_\perp = e^U \mathscr{P}_\perp e^{-U}. \tag{A41b}$$

These relations are clearly the analogues of Eqs. (A18). (Recall that $\mathscr{P}$, $\mathscr{P}_\perp$ are both Hermitian.) Using these we easily establish the following properties:

$$(\mathscr{P}^+)^2 = \mathscr{P}^+, \quad (\mathscr{P}^+_\perp)^2 = \mathscr{P}^+_\perp, \tag{A42a}$$

$$\mathscr{P}^+ + \mathscr{P}^+_\perp = I, \tag{A42b}$$

$$(\mathscr{P}^+)^\dagger = \mathscr{P}^+, \quad (\mathscr{P}^+_\perp)^\dagger = \mathscr{P}^+_\perp, \tag{A42c}$$

$$\mathscr{P}^+ \mathscr{P}^+_\perp = \mathscr{P}^+_\perp \mathscr{P}^+ = 0, \tag{A42d}$$

which are the counterparts of Eqs. (A30).

The following identities can also be easily verified:

$$(A, \mathscr{P} \psi) = (\mathscr{P}^+ A, \mathscr{P} \psi) = (\mathscr{P}^+ A, \psi), \tag{A43a}$$

$$(A, \mathscr{P}_\perp \psi) = (\mathscr{P}^+_\perp A, \mathscr{P}_\perp \psi) = (\mathscr{P}^+_\perp A, \psi), \tag{A43b}$$

$$(\mathscr{P}^+_\perp A, \mathscr{P} \psi) = (\mathscr{P}^+ A, \mathscr{P}_\perp \psi) = 0. \tag{A43c}$$

6. Schrödinger and Heisenberg pictures

The relation between distributions and observables in Fokker-Planck theory is similar to that between wave functions and observables in quantum mechanics. In particular, the dynamic evolution can be ascribed either to distributions or to observables. In the *Schrödinger picture*, observables remain unchanged whereas distributions evolve according to the Fokker-Planck equation (3.2a), whose solution can be formally written as

$$p(\mathbf{z}, t) = e^{\mathscr{L}(\mathbf{z}) t} p(\mathbf{z}, 0) = e^{\mathscr{L}(\mathbf{z}) t} p(\mathbf{z}). \tag{A44}$$

In the *Heisenberg picture*, by contrast, distributions remain unchanged whereas observables evolve according to the conjugate Fokker-Planck equation,

$$\partial_t A(\mathbf{z}, t) = \mathscr{L}^+ A(\mathbf{z}, t), \tag{A45}$$

whose formal solution is given by

$$A(\mathbf{z}, t) = e^{\mathscr{L}^+ t} A(\mathbf{z}, 0) = e^{\mathscr{L}^+ t} A(\mathbf{z}). \tag{A46}$$

Because of Eqs. (A16a), the average of an observable A at time t can be calculated either in Schrödinger picture or in Heisenberg picture,

$$\begin{aligned}(A, e^{\mathscr{L} t} p) &= \int_{\mathbf{z}} A(\mathbf{z}) p(\mathbf{z}, t) \\ &= (e^{\mathscr{L}^+ t} A, p) = \int_{\mathbf{z}} A(\mathbf{z}, t)\, p(\mathbf{z}) = \langle A(t) \rangle.\end{aligned} \tag{A47}$$

There are, however, important differences between Fokker-Planck theory and quantum mechanics. In the former, observables are functions of $\mathbf{z}$, whereas in the later, observables are operators acting on wave functions. Also in Fokker-Planck theory, average of an observable is linear in $p(\mathbf{z}, t)$ whereas in quantum mechanics, the corresponding quantity is bilinear in wave function and its complex conjugate. We shall use calligraphic letters $\mathscr{L}$, $\mathscr{P}$, … etc to denote operators and capital Roman letters $A(\mathbf{z})$, $B(\mathbf{z})$, … denote observables in Fokker-Planck theory.

7. Green's functions and correlation functions

We define the Green's functions for $\mathscr{L}$ and $\mathscr{L}^+$,

$$\partial_t G(\mathbf{z},\mathbf{z}';t) = \mathscr{L}(\mathbf{z})G(\mathbf{z},\mathbf{z}';t), \tag{A48a}$$
$$G(\mathbf{z},\mathbf{z}';0) = \delta(\mathbf{z}-\mathbf{z}); \tag{A48b}$$
$$\partial_t G^+(\mathbf{z},\mathbf{z}';t) = \mathscr{L}^+(\mathbf{z})G^+(\mathbf{z},\mathbf{z}';t), \tag{A48c}$$
$$G^+(\mathbf{z},\mathbf{z}';0) = \delta(\mathbf{z}-\mathbf{z}). \tag{A48d}$$

The physical significance of $G(\mathbf{z},\mathbf{z}';t)$ is the conditional pdf that the system variables take value $\mathbf{z}$ at time t, given that it starts from $\mathbf{z}'$ at $t=0$.

These Green's functions can be formally written as

$$G(\mathbf{z},\mathbf{z}';t) = e^{t\mathscr{L}(\mathbf{z})}\delta(\mathbf{z}-\mathbf{z}'), \tag{A49a}$$
$$G^+(\mathbf{z},\mathbf{z}';t) = e^{t\mathscr{L}^+(\mathbf{z})}\delta(\mathbf{z}-\mathbf{z}'). \tag{A49b}$$

The solution to FPE, Eq. (A44), can then be written as

$$\begin{aligned} e^{\mathscr{L}(\mathbf{z})t}p(\mathbf{z}) &= \int_{\mathbf{z}'} e^{\mathscr{L}(\mathbf{z})t}\delta(\mathbf{z}-\mathbf{z}')p(\mathbf{z}') \\ &= \int_{\mathbf{z}'} G(\mathbf{z},\mathbf{z}';t)p(\mathbf{z}') \end{aligned} \tag{A50}$$

Consider the following equality:

$$\begin{aligned} G(\mathbf{z}'',\mathbf{z}';t) &= \langle\delta(\mathbf{z}-\mathbf{z}''), G(\mathbf{z},\mathbf{z}';t)\rangle \\ &= \langle\delta(\mathbf{z}-\mathbf{z}''), e^{t\mathscr{L}(\mathbf{z})}\delta(\mathbf{z}-\mathbf{z}')\rangle \\ &= \langle e^{t\mathscr{L}^+(\mathbf{z})}\delta(\mathbf{z}-\mathbf{z}''), \delta(\mathbf{z}-\mathbf{z}')\rangle. \\ &= \langle G^+(\mathbf{z},\mathbf{z}'';t), \delta(\mathbf{z}-\mathbf{z}')\rangle \\ &= G^+(\mathbf{z}',\mathbf{z}'';t). \end{aligned} \tag{A51}$$

Hence we obtain the following symmetry property:

$$G(\mathbf{z}'',\mathbf{z}';t) = G^+(\mathbf{z}',\mathbf{z}'';t). \tag{A52}$$

The average of observable $A(\mathbf{z})$, Eq. (A47) can be rewritten in terms of Green's function as

$$\begin{aligned} \langle A(t)\rangle &= \int_{\mathbf{z}}\int_{\mathbf{z}'} A(\mathbf{z})G(\mathbf{z},\mathbf{z}';t)p(\mathbf{z}') \\ &= \int_{\mathbf{z}}\int_{\mathbf{z}'} [G^+(\mathbf{z}',\mathbf{z};t)A(\mathbf{z})]p(\mathbf{z}'). \end{aligned} \tag{A53}$$

The correlation function of two observables, A at time t and B at time $t=0$, can be expressed as

$$\begin{aligned} \langle A(t)B(0)\rangle &= \int_{\mathbf{z}}\int_{\mathbf{z}'} A(\mathbf{z})G(\mathbf{z},\mathbf{z}';t)B(\mathbf{z}')p(\mathbf{z}'). \qquad \text{(A54)} \\ &= \int_{\mathbf{z}} A(\mathbf{z})e^{\mathscr{L}(\mathbf{z})t}B(\mathbf{z})p(\mathbf{z}) \\ &= \int_{\mathbf{z}} [e^{\mathscr{L}^+(\mathbf{z})t}A(\mathbf{z})]B(\mathbf{z})p(\mathbf{z}), \qquad \text{(A55)} \end{aligned}$$

where $p(\mathbf{z})$ is the initial pdf. If the system start from steady state, $p(\mathbf{z}) = e^{-U(\mathbf{z})}$, and $\langle A(t)B(0)\rangle$ become the steady-state correlation functions:

$$\begin{aligned} \langle A(t)B(0)\rangle_{\rm SS} &= \int_{\mathbf{z}} e^{-U(\mathbf{z})}[e^{\mathscr{L}^+(\mathbf{z})t}A(\mathbf{z})]B(\mathbf{z}) \\ &= \langle e^{\mathscr{L}^+(\mathbf{z})t}A(\mathbf{z}), B(\mathbf{z})\rangle \\ &= \langle A(\mathbf{z};t), B(\mathbf{z})\rangle. \end{aligned} \tag{A56}$$

8. Time-reversal symmetry and detailed balance

A detailed discussion on detailed balance, i.e., time-reversal symmetry of Fokker-Planck dynamics is given in Ref. [61], as well as in Sec. II A of Ref. [62]. Here we present a compact derivation of the detailed balance conditions, Eqs. (3.47) and (3.48).

As shown in Eq. (3.46), a Fokker-Planck dynamics has time-reversal symmetry if

$$P_F(\mathbf{z}_2,t_2;\mathbf{z}_1,t_1) = P_B(\mathbf{z}_1^*,t_2;\mathbf{z}_2^*,t_1). \tag{A57}$$

Note that integrating over $\mathbf{z}_2$, we obtain $p_F(\mathbf{z}_1,t_1) = p_B(\mathbf{z}_1^*,t_2)$, which implies that each side of the equation must be independent of t_1,t_2. *This is possible only if both processes are in the stationary regime*, and hence $p_F(\mathbf{z}_1,t_1) = p_B(\mathbf{z}_1^*,t_2) = e^{-U(\mathbf{z}_1,\lambda)} = e^{-U(\mathbf{z}_1^*,\lambda^*)}$. Thus we obtain

$$U(\mathbf{z}_1,\lambda) = U(\mathbf{z}_1^*,\lambda^*). \tag{A58}$$

Since the process is Markovian, the two-point joint pdf $P_F(\mathbf{z}_2,t_2;\mathbf{z}_1,t_1)$ can be written as a product of the marginal pdf $p_F(\mathbf{z}_1) = e^{-U(\mathbf{z}_1,\lambda)}$ and the conditional pdf, which is precisely the Green's function $G(\mathbf{z}_2,\mathbf{z}_1;\Delta t)$, where $\Delta t = t_2 - t_1$. Using Eqs. (A49a), (A52), and (A49b) successively, we can further write l.h.s. of Eq. (A57) as

$$\begin{aligned} P_F(\mathbf{z}_2,t_2;\mathbf{z}_1,t_1) &= G_F(\mathbf{z}_2,\mathbf{z}_1;\Delta t)e^{-U(\mathbf{z}_1,\lambda)} \\ &= G_F^+(\mathbf{z}_1,\mathbf{z}_2;\Delta t)e^{-U(\mathbf{z}_1,\lambda)} \\ &= e^{-U(\mathbf{z}_1,\lambda)}\,e^{\Delta t\,\mathscr{L}^+(\mathbf{z}_1,\lambda)}\delta(\mathbf{z}_1-\mathbf{z}_2). \end{aligned} \tag{A59}$$

Similarly, we can write the r.h.s. of Eq. (A57) as

$$\begin{aligned} P_B(\mathbf{z}_1^*,t_2;\mathbf{z}_2^*,t_1) &= G_B(\mathbf{z}_1^*,\mathbf{z}_2^*;\Delta t)e^{-U(\mathbf{z}_2^*,\lambda^*)} \\ &= e^{t\mathscr{L}(\mathbf{z}_1^*,\lambda^*)}\delta(\mathbf{z}_1^*-\mathbf{z}_2^*)e^{-U(\mathbf{z}_2^*,\lambda^*)} \\ &= e^{t\mathscr{L}(\mathbf{z}_1^*,\lambda^*)}\delta(\mathbf{z}_1-\mathbf{z}_2)e^{-U(\mathbf{z}_2,\lambda)} \\ &= e^{t\mathscr{L}(\mathbf{z}_1^*,\lambda^*)}e^{-U(\mathbf{z}_1,\lambda)}\delta(\mathbf{z}_1-\mathbf{z}_2), \end{aligned}$$

where in the third equality we have used $\delta(\mathbf{z}^*) = \delta(\mathbf{z})$, and $U(\mathbf{z}_1,\lambda) = U(\mathbf{z}_1^*,\lambda^*)$. In the last equality, we have used the property of delta function to replace $e^{-U(\mathbf{z}_1,\lambda)}$ by $e^{-U(\mathbf{z}_2,\lambda)}$. Note that $e^{t\mathscr{L}(\mathbf{z}_1^*,\lambda^*)}$ now acts on the whole product to its right.

Now substituting the preceding two equations into Eq. (A57), and get rid of the delta function, we find an operator identity,

$$e^{-U(\mathbf{z},\lambda)}\,e^{\Delta t\,\mathscr{L}^+(\mathbf{z},\lambda)} = e^{\Delta t\mathscr{L}(\mathbf{z}^*,\lambda^*)}e^{-U(\mathbf{z},\lambda)}. \tag{A61}$$

We can expand both sides in terms of Δ. To the zeroth-order we obtain Eq. (A58). To the first order we obtain

$$e^{-U(\mathbf{z},\lambda)}\mathscr{L}^+(\mathbf{z},\lambda) = \mathscr{L}(\mathbf{z}^*,\lambda^*)e^{-U(\mathbf{z},\lambda)}, \tag{A62}$$

which is Eq. (A62). Further using Eq. (A18), we can rewrite Eq. (A62) into an equivalent form

$$\mathscr{L}^\dagger(\mathbf{z},\lambda) = \mathscr{L}(\mathbf{z}^*,\lambda^*), \tag{A63}$$

or equivalent as

$$(\mathscr{L}^\dagger)^+(\mathbf{z},\lambda) = \mathscr{L}^+(\mathbf{z}^*,\lambda^*). \tag{A64}$$

Substituting the concrete forms of operators, Eqs. (A10) and (A12) into Eq. (A63), we obtain the conditions of detailed balance,

$$U(\mathbf{z},\lambda) = U(\mathbf{z}^*,\lambda^*), \tag{A65a}$$

$$\varepsilon_\mu L^{\mu\nu}(\mathbf{z}^*,\lambda^*)\varepsilon_\nu = L^{\nu\mu}(\mathbf{z},\lambda). \tag{A65b}$$

These results were derived in Ref. [61] using different method. To make sure that e^{-U} can be realized physically, of course, it must also be normalizable,

$$\int_{\mathbf{z}} e^{-U} = 1. \tag{A65c}$$

In fact, Eqs. (A62), (A63), (A64), and (A65) are all equivalent, and any of them can be used to define detailed balance.

9. Preservation of detailed balance by coarse-graining

Here we outline an equivalent proof for preservation of detailed balance by coarse-graining. Assuming that the joint dynamics has detailed balance, which means that Eqs. (A65) and (A62) are valid, we only need to show

$$e^{-U_{\mathbf{X}}(\boldsymbol{x},\lambda)}\big(\mathscr{L}^R_{\mathbf{X}}(\boldsymbol{x},\lambda)\big)^+ e^{U_{\mathbf{X}}(\boldsymbol{x},\lambda)} = \mathscr{L}^R_{\mathbf{X}}(\boldsymbol{x}^*,\lambda^*). \tag{A66}$$

This is clearly the analogue of Eq. (A62) for the coarse-grained dynamics.

Recall that the kinetic matrix of the joint system can be decomposed via Eq. (3.7), whereas the generalized potential can be decomposed via Eq. (3.14). We shall further assume that the control parameter λ does not appear in $U_{\mathbf{Y}}(\boldsymbol{y};\boldsymbol{x})$. Equations (A65) can then be rewritten into the following equivalent forms:

$$U_{\mathbf{X}}(\boldsymbol{x},\lambda) = U_{\mathbf{X}}(\boldsymbol{x}^*,\lambda^*), \tag{A67a}$$

$$\varepsilon_i L^{ij}(\mathbf{z}^*,\lambda^*)\varepsilon_j = L^{ji}(\mathbf{z},\lambda), \tag{A67b}$$

$$U_{\mathbf{Y}}(\boldsymbol{y};\boldsymbol{x}) = U_{\mathbf{Y}}(\boldsymbol{y}^*;\boldsymbol{x}^*), \tag{A67c}$$

$$\varepsilon_a L^{ab}(\mathbf{z}^*,\lambda^*)\varepsilon_b = L^{ba}(\mathbf{z},\lambda), \tag{A67d}$$

$$\varepsilon_i L^{ia}(\mathbf{z}^*,\lambda^*)\varepsilon_a = L^{ai}(\mathbf{z},\lambda), \tag{A67e}$$

$$\varepsilon_a L^{ai}(\mathbf{z}^*,\lambda^*)\varepsilon_i = L^{ia}(\mathbf{z},\lambda). \tag{A67f}$$

Equations (A62) also implies that $\mathscr{L}^+_S(\mathbf{z},\lambda)$ and $\mathscr{L}^+_{\mathbf{Y}}(\mathbf{z},\lambda)$ have the same symmetry,

$$e^{-U(\mathbf{z},\lambda)}\mathscr{L}^+_S(\mathbf{z},\lambda) = \mathscr{L}_S(\mathbf{z}^*,\lambda^*)e^{-U(\mathbf{z},\lambda)}, \tag{A68a}$$

$$e^{-U(\mathbf{z},\lambda)}\mathscr{L}^+_{\mathbf{Y}}(\mathbf{z},\lambda) = \mathscr{L}_{\mathbf{Y}}(\mathbf{z}^*,\lambda^*)e^{-U(\mathbf{z},\lambda)}. \tag{A68b}$$

Using Eqs. (3.43) and (2.24), we can further show that $\mathscr{L}^{-\perp}_{\mathbf{Y}}(\mathbf{z},\lambda)$ also as the same symmetry,

$$e^{-U(\mathbf{z},\lambda)}\big(\mathscr{L}^{-\perp}_{\mathbf{Y}}(\mathbf{z},\lambda)\big)^+ = \mathscr{L}^{-\perp}_{\mathbf{Y}}(\mathbf{z}^*,\lambda^*)e^{-U(\mathbf{z},\lambda)}. \tag{A68c}$$

We will now show that the first operator $\mathscr{L}^0_{\mathbf{X}}$ in the expansion (2.57) satisfy detailed balance condition, if the joint dynamics satisfy detailed balance. Firstly, using Eq. (2.44), we have

$$\begin{aligned} &e^{-U_{\mathbf{X}}(\boldsymbol{x},\lambda)}\big(\mathscr{L}^{(0)}_{\mathbf{X}}(\boldsymbol{x},\lambda)\big)^+ e^{U_{\mathbf{X}}(\boldsymbol{x},\lambda)} \\ &= e^{-U_{\mathbf{X}}(\boldsymbol{x},\lambda)}\left[\int_{\boldsymbol{y}} \mathscr{L}_S(\mathbf{z},\lambda)e^{-U_{\mathbf{Y}}(\mathbf{z},\lambda)}\right]^+ e^{U_{\mathbf{X}}(\boldsymbol{x},\lambda)} \\ &= e^{-U_{\mathbf{X}}(\boldsymbol{x},\lambda)}\int_{\boldsymbol{y}} e^{-U_{\mathbf{Y}}(\mathbf{z},\lambda)}\mathscr{L}^+_S(\mathbf{z},\lambda)e^{U_{\mathbf{X}}(\boldsymbol{x},\lambda)}, \end{aligned} \tag{A69}$$

where in the last equality, we have used $(AB)^+ = B^+A^+$ and $(e^{-U_{\mathbf{Y}}})^+ = e^{-U_{\mathbf{Y}}}$. Further using Eqs. (3.14), (A68a), and (A67c), we have

$$\begin{aligned} \text{Eq. (A.69)} &= \int_{\boldsymbol{y}} e^{-U(\mathbf{z},\lambda)}\mathscr{L}^+_S(\mathbf{z},\lambda)e^{U(\mathbf{z},\lambda)}e^{-U_{\mathbf{Y}}(\mathbf{z})} \\ &= \int_{\boldsymbol{y}} \mathscr{L}_S(\mathbf{z}^*,\lambda^*)e^{-U_{\mathbf{Y}}(\mathbf{z}^*)} \\ &= \int_{\boldsymbol{y}^*} \mathscr{L}_S(\mathbf{z}^*,\lambda^*)e^{-U_{\mathbf{Y}}(\mathbf{z}^*)} \\ &= \mathscr{L}^{(0)}_{\mathbf{X}}(\boldsymbol{x}^*,\lambda^*), \end{aligned} \tag{A70}$$

where in the third equality we time reversed the dummy variable $\boldsymbol{y}$, and in the last step we used Eq. (2.44). Summarizing we have

$$e^{-U_{\mathbf{X}}(\boldsymbol{x},\lambda)}\big(\mathscr{L}^{(0)}_{\mathbf{X}}(\boldsymbol{x},\lambda)\big)^+ e^{U_{\mathbf{X}}(\boldsymbol{x},\lambda)} = \mathscr{L}^{(0)}_{\mathbf{X}}(\boldsymbol{x}^*,\lambda^*), \tag{A71}$$

which is the detailed balance condition for $\mathscr{L}^{(0)}_{\mathbf{X}}$, the zeroth-order term in the expansion (2.57).

Using the same method, we can also prove

$$e^{-U_{\mathbf{X}}(\boldsymbol{x},\lambda)}\big(\mathscr{L}^{(1)}_{\mathbf{X}}(\boldsymbol{x},\lambda)\big)^+ e^{U_{\mathbf{X}}(\boldsymbol{x},\lambda)} = \mathscr{L}^{(1)}_{\mathbf{X}}(\boldsymbol{x}^*,\lambda^*), \tag{A72}$$

which is the detailed balance condition for $\mathscr{L}^{(0)}_{\mathbf{X}}$, the first-order term in the expansion (2.57). The same proof may be carried out for higher-order terms, even though in most cases, we only need the first two terms in the renormalized FPO. Summarizing, we find that the first-order renormalized FPO of the effective slow dynamics satisfies detailed balance condition,

$$e^{-U_{\mathbf{X}}(\boldsymbol{x},\lambda)}\big(\mathscr{L}^R_{\mathbf{X}}(\boldsymbol{x},\lambda)\big)^+ e^{U_{\mathbf{X}}(\boldsymbol{x},\lambda)} = \mathscr{L}^R_{\mathbf{X}}(\boldsymbol{x}^*,\lambda^*). \tag{A73}$$

Hence detailed balance property is preserved by coarse-graining.

[1] S. Nakajima, On quantum theory of transport phenomena: Steady diffusion, Prog. Theor. Phys. **20**, 948 (1958).

[2] H. Mori, Transport, collective motion, and Brownian motion, Prog. Theor. Phys. **33**, 423 (1965).

[3] R. Zwanzig, Nonlinear generalized Langevin equations, J. Stat. Phys. **9**, 215 (1973).

[4] R. Zwanzig, Memory effects in irreversible thermodynamics, Phys. Rev. **124**, 983 (1961).

[5] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, Oxford, 2001).

[6] R. Zwanzig, Ensemble method in the theory of irreversibility, J. Chem. Phys. **33**, 1338 (1960).

[7] H. Grabert, *Projection Operator Techniques in Nonequilibrium Statistical Mechanics*, Vol. 95 (Springer, New York, 2006).

[8] J. Xing and K. S. Kim, Application of the projection operator formalism to non-Hamiltonian dynamics, J. Chem. Phys. **134**, 044132 (2011).

[9] M. Te Vrugt and R. Wittkowski, Mori-Zwanzig projection operator formalism for far-from-equilibrium systems with time-dependent Hamiltonians, Phys. Rev. E **99**, 062118 (2019).

[10] M. H. Lee, Derivation of the generalized Langevin equation by a method of recurrence relations, J. Math. Phys. **24**, 2512 (1983).

[11] J. M. Deutch and R. Silbey, Exact generalized Langevin equation for a particle in a harmonic lattice. Phys. Rev. A **3**, 2049 (1971).

[12] A. O. Caldeira and A. J. Leggett, Path integral approach to quantum Brownian motion, Physica A **121**, 587 (1983).

[13] A. O. Caldeira and A. J. Leggett, Quantum tunnelling in a dissipative system, Ann. Phys. (NY) **149**, 374 (1983).

[14] A. Rivas and S. F. Huelga, *Open Quantum Systems* (Springer, Berlin, 2012).

[15] G. W. Ford, J. T. Lewis, and R. F. O'Connell, Quantum Langevin equation, Phys. Rev. A **37**, 4419 (1988).

[16] D. Chruściński and A. Kossakowski, Non-Markovian Quantum Dynamics: Local versus Nonlocal, Phys. Rev. Lett. **104**, 070406 (2010).

[17] E. Cortes, B. J. West, and K. Lindenberg, On the generalized Langevin equation: Classical and quantum mechanical, J. Chem. Phys. **82**, 2708 (1985).

[18] S. Chaturvedi and F. Shibata, Time-convolutionless projection operator formalism for elimination of fast variables. Applications to Brownian motion, Z. Phys. B **35**, 297 (1979).

[19] F. Shibata and T. Arimitsu, Expansion formulas in nonequilibrium statistical mechanics, J. Phys. Soc. Jpn. **49**, 891 (1980).

[20] C. Uchiyama and F. Shibata, Unified projection operator formalism in nonequilibrium statistical mechanics, Phys. Rev. E **60**, 2636 (1999).

[21] H. P. Breuer, E. M. Laine, J. Piilo, and B. Vacchini, Colloquium: Non-Markovian dynamics in open quantum systems, Rev. Mod. Phys. **88**, 021002 (2016).

[22] H. P. Breuer, B. Kappler, F. Petruccione, The time-convolutionless projection operator technique in the quantum theory of dissipation and decoherence, Ann. Phys. (NY) **291**, 36 (2001).

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

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

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

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

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

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

[29] G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in stationary states, J. Stat. Phys. **80**, 931 (1995).

[30] J. L. Lebowitz and H. Spohn, A Gallavott-Cohen-type symmetry in the large deviation functional for stochastic dynamics, J. Stat. Phys. **95**, 333 (1999).

[31] T. Hatano and S.-i. Sasa, Steady-State Thermodynamics of Langevin Systems, Phys. Rev. Lett. **86**, 3463 (2001).

[32] L. Peliti and S. Pigolotti, *Stochastic Thermodynamics: An Introduction* (Princeton University Press, Princeton, 2021).

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

[34] M. Esposito, Stochastic thermodynamics under coarse graining, Phys. Rev. E **85**, 041125 (2012).

[35] T. Leonard, B. Lander, U. Seifert, and T. Speck, Stochastic thermodynamics of fluctuating density fields: Non-equilibrium free energy differences under coarse-graining, J. Chem. Phys. **139**, 204109 (2013).

[36] S. Pigolotti and A. Vulpiani, Coarse graining of master equations with fast and slow states, J. Chem. Phys. **128**, 154114 (2008).

[37] S. Rahav and C. Jarzynski, Fluctuation relations and coarse-graining, J. Stat. Mech.: Theory Exp. (2007) P09012.

[38] N. G. Van Kampen, Elimination of fast variables, Phys. Rep. **124**, 69 (1985).

[39] P. Hänggi and H. Thomas, Time evolution, correlations, and linear response of non-Markov processes, Z. Phys. B **26**, 85 (1977).

[40] S. C. Kou, X. S. Xie, Generalized Langevin Equation with Fractional Gaussian Noise: Subdiffusion within a Single Protein Molecule, Phys. Rev. Lett. **93**, 180603 (2004).

[41] A. Puglisi *et al.*, Entropy production and coarse graining in Markov processes, J. Stat. Mech.: Theory Exp. (2010) P05015.

[42] P. Bilotto, L. Caprini, and A. Vulpiani, Excess and loss of entropy production for different levels of coarse graining, Phys. Rev. E **104**, 024140 (2021).

[43] J. Mehl *et al.*, Role of Hidden Slow Degrees of Freedom in the Fluctuation Theorem, Phys. Rev. Lett. **108**, 220601 (2012).

[44] H.-M. Chun and J. D. Noh, Hidden entropy production by fast variables, Phys. Rev. E **91**, 052128 (2015).

[45] R. Marino and R. Eichhorn, and E. Aurell, Entropy production of a Brownian ellipsoid in the overdamped limit, Phys. Rev. E **93**, 012132 (2016).

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

[47] L. Cocconi, G. Salbreux, and G. Pruessner, Scaling of entropy production under coarse graining in active disordered media, Phys. Rev. E **105**, L042601 (2022).

[48] A. Ghosal and G. Bisker, Inferring entropy production rate from partially observed Langevin dynamics under coarse-graining, Phys. Chem. Chem. Phys., **24**, 24021 (2022).

[49] D. M. Busiello and A. Maritan, Entropy production in master equations and Fokker-Planck equations: Facing the coarse-graining and recovering the information loss, J. Stat. Mech.: Theory Exp. (2019) 104013.

[50] D. M. Busiello, J. Hidalgo, and A. Maritan, Entropy production for coarse-grained dynamics, New J. Phys. **21**, 073004 (2019).

[51] M. Guenza, Thermodynamic consistency and other challenges in coarse-graining models, Eur. Phys. J.: Spec. Top. **224**, 2177 (2015).

[52] H. Ge, C. Jia, and X. Jin, Martingale structure for general thermodynamic functionals of diffusion processes under second-order averaging, J. Stat. Phys. **184**, 1 (2021).

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

[54] M. Santillán and H. Qian, Irreversible thermodynamics in multiscale stochastic dynamical systems, Phys. Rev. E **83**, 041130 (2011).

[55] P. Strasberg and M. Esposito, Stochastic thermodynamics in the strong coupling regime: An unambiguous approach based on coarse graining, Phys. Rev. E **95**, 062101 (2017).

[56] G. Teza and A. L. Stella, Exact Coarse Graining Preserves Entropy Production out of Equilibrium, Phys. Rev. Lett. **125**, 110601 (2020).

[57] B. Altaner and J. Vollmer, Fluctuation-Preserving Coarse Graining for Biochemical Systems, Phys. Rev. Lett. **108**, 228101 (2012).

[58] K. Kawaguchi and Y. Nakayama, Fluctuation theorem for hidden entropy production, Phys. Rev. E **88**, 022147 (2013).

[59] T. Herpich, T. Cossetto, G. Falasco, and M. Esposito, Stochastic thermodynamics of all-to-all interacting many-body systems, New J. Phys. **22**, 063005 (2020).

[60] T. Herpich, K. Shayanfard, and M. Esposito, Effective thermodynamics of two interacting underdamped Brownian particles, Phys. Rev. E **101**, 022116 (2020).

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

[62] M. Ding and X. Xing, Covariant nonequilibrium thermodynamics from Ito-Langevin dynamics, Phys. Rev. Res. **4**, 033247 (2022).

[63] 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).

[64] M. Ding, F. Liu, and X. Xing, Unified theory of thermodynamics and stochastic thermodynamics for nonlinear Langevin systems driven by non-conservative forces, Phys. Rev. Res. **4**, 043125 (2022).

[65] P. Glansdorff and I. Prigogine, *Thermodynamic Theory of Structure, Stability and Fluctuations* (J. Willey & Sons, Hoboken, NJ, 1971).

[66] S.-i. Sasa and H. Tasaki, Steady state thermodynamics, J. Stat. Phys. **125**, 125 (2006).

[67] M. Ding and X. Xing, Coarse-graining of stochastic thermodynamics (to be published).

[68] L. Bocquet, High friction limit of the Kramers equation: The multiple time-scale approach, Am. J. Phys. **65**, 140 (1997).

[69] 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).

[70] C. M. Bender, S. Orszag, and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*. Vol. 1 (Springer Science & Business Media, New York, 1999).

[71] G. Pavliotis and A. Stuart, *Multiscale Methods: Averaging and Homogenization* (Springer Science & Business Media, New York, 2008).

[72] S. Hottovy, A. McDaniel, G. Volpe, and J. Wehr, The Smoluchowski-Kramers limit of stochastic differential equations with arbitrary state-dependent friction, Commun. Math. Phys. **336**, 1259 (2015)

[73] M. Ottobre and G. A. Pavliotis, Asymptotic analysis for the generalized Langevin equation, Nonlinearity **24**, 1629 (2011).

[74] R. F. Pawula, Approximation of the linear Boltzmann equation by the Fokker-Planck equation, Phys. Rev. **162**, 186 (1967).

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

[76] The formal solution to Eq. (2.2) is $p(t) = e^{t\mathscr{L}} p(0)$. Hence we readily see that dividing $\mathscr{L}_{\mathbf{Y}}$ by ϵ amounts to speeding up the dynamics of $\boldsymbol{y}$ by a factor $1/\epsilon$.

[77] Here the term covariant means that the theory is parameterized by tensor objects with standard transformation properties under nonlinear transformation of variables. The covariance guarantees that theories formulated using different coordinate systems are mathematically equivalent.