

ON THE TANGENT FLOW OF A STOCHASTIC DIFFERENTIAL EQUATION WITH FAST DRIFT

RICHARD B. SOWERS

ABSTRACT. We investigate the behavior of the tangent flow of a stochastic differential equation with a fast drift. The state space of the stochastic differential equation is the two-dimensional cylinder. The fast drift has closed orbits, and we assume that the orbit times vary nontrivially with the axial coordinate. Under a nondegeneracy assumption, we find the rate of growth of the tangent flow. The calculations involve a transformation introduced by Pinsky and Wihstutz.

1. INTRODUCTION

There has recently been a renewed interest in *stochastic averaging*; viz., in the techniques of model-reduction of stochastic differential equations (SDE's) with several scales of speed. Although classical stochastic averaging dates back to the 1960's and the work of (among others) Khasminskii (see [8] and [9]), recent results by Freidlin and Wentzell [6] have expanded the scope of stochastic averaging. This renewed interest, and the development of a number of tools of stochastic analysis since the 1960's suggests that we re-examine stochastic averaging.

Our interest here is in the *stability* of stochastic averaging. In particular, we will investigate the effect of an infinitesimally small perturbation of the initial condition upon the trajectories of a stochastic differential equation (SDE) which is amenable to stochastic averaging; more exactly, we are interested in the tangent flow of an SDE which is amenable to stochastic averaging. In almost any physical model, exact parameters, the initial condition included, are not precisely known. Thus it is of interest to understand how a slight perturbation of the initial condition manifests itself, particularly when a separation of scales occurs; this is a question of stability of the SDE. Secondly, if we know the law of the two-point motion of an SDE, then we can identify the coefficients of the SDE (see [2]); our results are also a small step in the direction of understanding the two-point motion of an SDE which has a fast motion. Baxendale has already investigated related asymptotics of the tangent flow of an SDE with no fast component [3]. Finally, we should mention that our results are a small step towards studying the Fokker-Planck equation of diffusions with fast components. One of the most notable advances in stochastic analysis in the past several decades is the Malliavin calculus, which probabilistically gives analytic information about densities of SDE's (see [4] and [12]). It would be interesting to

1991 *Mathematics Subject Classification.* 34D08, 34C29, 60H10.

Key words and phrases. Floquet, Lyapunov exponent, Pinsky-Wihstutz, stochastic averaging.

This work was supported by NSF DMS 9615877. The author would like to thank Professor Sri Namachchivaya of the Department of Aeronautical and Astronautical Engineering at the University of Illinois at Urbana-Champaign for pointing out the paper by Pinsky and Wihstutz. The author would also like to thank the anonymous referee who insisted upon notational clarity.

see how the techniques of Malliavin calculus could be combined with some of the asymptotics in stochastic averaging. Since the Malliavin covariance matrix crucially depends on the tangent flow, our efforts would hopefully be relevant for such an undertaking.

2. THE PROBLEM AND THE RESULT

Let's state our problem and our results. The state space of our SDE will be the two-dimensional cylinder, which we shall interpret as $(\mathbb{R}/\mathbb{Z}) \times \mathbb{R}$. We shall try to use $x \in \mathbb{R}$ to denote the periodic or angular coordinate (thus all functions of x will be periodic of period 1) and $y \in \mathbb{R}$ to denote the axial coordinate and $z = (x, y) \in \mathbb{R}^2$ to denote the ordered pair of these coordinates. Although we shall always write the components of an element of \mathbb{R}^2 as a row vector, we shall implicitly convert to column vectors in all instances of matrix multiplication. We first define a speed which depends only on the y coordinate.

Definition 1 (Fast Drift). *Fix $\alpha \in C^\infty(\mathbb{R})$ such that*

- $\alpha(y) > 0$ for all $y \in \mathbb{R}$.
- $\alpha'(y) > 0$ for all $y \in \mathbb{R}$.
- For each $n \geq 0$, $\sup_{y \in \mathbb{R}} |\alpha^{(n)}(y)| < \infty$ (where, as usual, $\alpha^{(n)}$ is the n -th derivative of α).

We will let α be the dominant speed in the angular coordinate—see Figure 1. We also need some perturbations.

Definition 2 (Perturbations). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability triple on which are defined two independent standard Wiener processes W^1 and W^2 . For convenience, define the three semimartingales*

$$\gamma_t^0 \stackrel{\text{def}}{=} t, \quad \gamma_t^1 \stackrel{\text{def}}{=} W_t^1, \quad \text{and} \quad \gamma_t^2 \stackrel{\text{def}}{=} W_t^2. \quad t \geq 0$$

For each $i \in \{0, 1, 2\}$ and each $j \in \{1, 2\}$, let $\sigma_i^j \in C^\infty(\mathbb{R}^2)$ be such that

- $\sigma_i^j(x, y) = \sigma_i^j(x + 1, y)$ for all x and y in \mathbb{R}
- For each triplet of nonnegative integers l, m , and n ,

$$\lim_{|y| \rightarrow \infty} \sup_{x \in \mathbb{R}} \left| \frac{\partial^{m+l} \sigma_i^j}{\partial x^m \partial y^l}(x, y) \right| |y|^n = 0.$$

Our main object of study will be the following stochastic differential equation:

$$\begin{aligned} dX_t^\varepsilon(x, y) &= \frac{1}{\varepsilon^2} \alpha(Y_t^\varepsilon(x, y)) dt + \sum_{i=0}^2 \sigma_i^1(X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y)) \circ d\gamma_s^i \\ (1) \quad dY_t^\varepsilon(x, y) &= \sum_{i=0}^2 \sigma_i^2(X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y)) \circ d\gamma_s^i & t \geq 0 \\ X_0^\varepsilon(x, y) &= x \\ Y_0^\varepsilon(x, y) &= y \end{aligned}$$

for all $(x, y) \in \mathbb{R}^2$. Here “ \circ ” corresponds to Stratonovich integration. This SDE is a speeded-up and randomly-perturbed form of the *twist map*

$$\mathfrak{z}_t(x, y) \stackrel{\text{def}}{=} (x + \alpha(y)t, y). \quad t \in \mathbb{R}, (x, y) \in \mathbb{R}^2$$

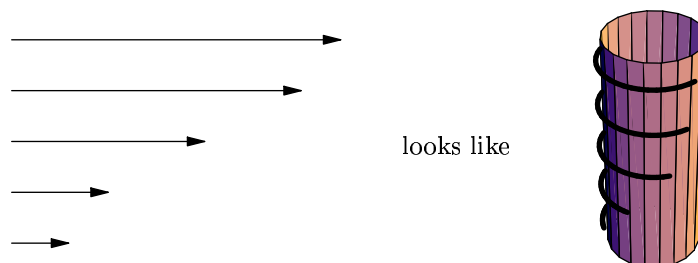
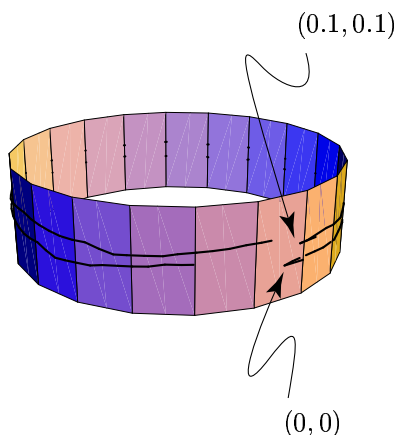


FIGURE 1. Fast Drift

FIGURE 2. Two Solutions (with starting points $(0, 0)$ and $(0.1, 0.1)$)

The orbits of the twist map are periodic (when we consider \mathbb{R}/\mathbb{Z}) and they are of constant speed, with this speed varying from orbit to orbit.

We are interested in the *stability* of the SDE (1). Standard results (see [5] and [11]) ensure that for each $t \geq 0$, the map $(x, y) \rightarrow (X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y))$ is C^∞ . We can then define the 2×2 matrix $\Xi_t^\varepsilon(x, y)$ by

$$\Xi_t^{\varepsilon, x, y} \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\partial X_t^\varepsilon}{\partial x}(x, y) & \frac{\partial X_t^\varepsilon}{\partial y}(x, y) \\ \frac{\partial Y_t^\varepsilon}{\partial x}(x, y) & \frac{\partial Y_t^\varepsilon}{\partial y}(x, y) \end{pmatrix}$$

for all $t \geq 0$ and $(x, y) \in \mathbb{R}^2$. Heuristically, $\Xi_t^{\varepsilon, x, y}$ measures how a small perturbation in the initial condition (i.e., a small perturbation in the point (x, y)) propagates. We are interested in how this notion of stability interacts with the twist map. If we start the SDE (1) at two nearby points (x, y) and (x', y') , then the two solutions of (1) should quickly separate if the perturbations kick one of the solutions to a higher y (where the angular speed α is larger); see Figure 2. Our goal is to quantify this.

For future reference, let's identify the generator of $(X^\varepsilon, Y^\varepsilon)$. Define the three vector fields (i.e., derivations)

$$(\sigma_i \varphi)(x, y) \stackrel{\text{def}}{=} \sigma_i^1(x, y) \frac{\partial \varphi}{\partial x}(x, y) + \sigma_i^2(x, y) \frac{\partial \varphi}{\partial y}(x, y) \quad \varphi \in C^\infty(\mathbb{R}^2), (x, y) \in \mathbb{R}^2$$

for all $i \in \{0, 1, 2\}$. Also define the *fast* vector field

$$(\mathfrak{F}\varphi)(x, y) \stackrel{\text{def}}{=} \alpha(y) \frac{\partial \varphi}{\partial x}(x, y) \quad \varphi \in C^\infty(\mathbb{R}^2), (x, y) \in \mathbb{R}^2$$

and define the second-order operator Hörmander-form operator

$$\mathcal{L} \stackrel{\text{def}}{=} \sigma_0 + \frac{1}{2} \sum_{i=1}^2 \sigma_i^2.$$

The generator of $(X_t^\varepsilon, Y_t^\varepsilon)$ is the second-order operator $\varepsilon^{-2} \mathfrak{F} + \mathcal{L}$. We also note that for any $\varphi \in C_b^\infty(\mathbb{R}^2)$ and $t \geq 0$,

$$\begin{aligned} & \varphi(X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y)) - \varphi(X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y)) \\ &= \varepsilon^{-2} \int_0^t (\mathfrak{F}\varphi)(X_s^\varepsilon(x, y), Y_s^\varepsilon(x, y)) ds \\ &+ \sum_{i=0}^2 \int_0^t (\sigma_i \varphi)(X_s^\varepsilon(x, y), Y_s^\varepsilon(x, y)) \circ d\gamma_s^i \\ &= \varepsilon^{-2} \int_0^t (\mathfrak{F}\varphi)(X_s^\varepsilon(x, y), Y_s^\varepsilon(x, y)) ds \\ &+ \int_0^t (\mathcal{L}\varphi)(X_s^\varepsilon(x, y), Y_s^\varepsilon(x, y)) ds + \sum_{i=0}^2 \int_0^t (\sigma_i \varphi)(X_s^\varepsilon(x, y), Y_s^\varepsilon(x, y)) dW_s^i. \end{aligned}$$

The larger context of our problem is one of *stochastic averaging*. The point of (1) is that there is an asymptotic *separation of scales*; the angular coordinate $X_t^\varepsilon(x, y)$ varies quickly (with speed ε^{-2}) and $Y_t^\varepsilon(x, y)$ is a slowly-varying quantity. Since the orbits of the angular component are closed (upon mapping to \mathbb{R}/\mathbb{Z}), we should be able to get some closed dynamics for Y_t^ε as ε tends to zero. Note that

$$\inf\{t \geq 0 : \mathfrak{z}_t(x, y) = (x + 1, y)\} = \frac{1}{\alpha(y)}$$

for any $(x, y) \in \mathbb{R}^2$ and

$$\alpha(y) \int_0^{1/\alpha(y)} \varphi(\mathfrak{z}_s(x, y)) ds = \int_0^1 \varphi(x', y) dx'$$

for all $y \in \mathbb{R}$ and any Borel-measurable function $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ which is 1-periodic in the first argument. It is fairly classical that the law of Y^ε converges to that of a Markov process given by an *effective* generator. Define the second-order operator $\bar{\mathcal{L}}$ on $C^\infty(\mathbb{R})$ by

$$(\bar{\mathcal{L}}\varphi)(y) = \int_0^1 (\mathcal{L}\varphi')(x, y) dx \quad y \in \mathbb{R}$$

for all $\varphi \in C^\infty(\mathbb{R})$ where $\varphi'(x, y) \stackrel{\text{def}}{=} \varphi'(x, y)$ for all $(x, y) \in \mathbb{R}^2$. Then

Proposition 1. *The law of $\{Y_t^\varepsilon(x, y); t \geq 0\}$ converges (in the Prohorov topology on $\mathcal{P}(C([0, \infty); \mathbb{R}))$) to the law of a Markov process with generator \mathcal{L} (with domain $C^\infty(\mathbb{R})$) and initial condition δ_y .*

See [7] for a proof.

Let $\|\cdot\|$ be the standard Euclidean distance function on \mathbb{R}^2 . Our main result is:

Main Theorem. *Fix $(x, y) \in \mathbb{R}^2$, $t > 0$, and $V \in \mathbb{R}^2 \setminus \{0\}$. Then under Assumption 1 below,*

$$(2) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \varepsilon^{4/3} \ln \frac{\|\Xi_t^{\varepsilon, x, y} V\|}{\|V\|} - \int_0^t \lambda(Y_s^\varepsilon(x, y)) ds \right|^p \right] = 0,$$

where the map $\lambda : \mathbb{R} \rightarrow \mathbb{R}$ is given in Definition 4 below.

This identifies the rate of growth of $\Xi_t^{\varepsilon, x, y}$ for fixed t as ε tends to zero. Heuristically, we have that

$$\frac{\|\Xi_t^{\varepsilon, x, y} V\|}{\|V\|} \underset{\varepsilon \rightarrow 0}{\asymp} \exp \left[\varepsilon^{-4/3} \int_0^t \lambda(Y_s^\varepsilon(x, y)) ds \right].$$

Thus, if we start the SDE (1) at two nearby but different points, we should expect the two solutions to diverge at an exponential rate of $\varepsilon^{-4/3}$, with the precise rate depending upon the *slowly-varying* coordinate Y^ε (which is in the limit governed by the averaged dynamics of Proposition 1). The novel exponent of $4/3$ comes from a certain nilpotency in the fast dynamics of $\Xi^{\varepsilon, x, y}$.

To define the function λ of (2), we need several things. First, let's compute a certain *invariant distribution*.

Proposition 2. *For each $v_0 \in \mathbb{R}$ and $v_1 > 0$, the PDE*

$$\begin{aligned} v_0 (\sin^2 u_{v_0, v_1})'(\theta) + \frac{v_1}{2} \left(\cos^2 (\cos^2 u_{v_0, v_1})' \right)'(\theta) &= 0 & 0 \leq \theta \leq 2\pi \\ u_{v_0, v_1}(0) &= u_{v_0, v_1}(2\pi) \\ u'_{v_0, v_1}(0) &= u'_{v_0, v_1}(2\pi) \\ \int_0^{2\pi} u_{v_0, v_1}(\theta) d\theta &= 1. \end{aligned}$$

has a unique C^∞ solution.

The proof is deferred until Section 5. The term $(\sin^2 u_{v_0, v_1})'$ means the derivative of the map $\theta \mapsto (\sin^2 \theta) u_{v_0, v_1}(\theta)$ and the term $(\cos^2 (\cos^2 u_{v_0, v_1})')'$ means the derivative of the product of $\theta \mapsto \cos^2 \theta$ and the derivative of the map $\theta \mapsto (\cos^2 \theta) u_{v_0, v_1}(\theta)$. Next, we make a definition

Definition 3. *For each $v_0 \in \mathbb{R}$ and $v_1 > 0$, define*

$$\ell_{v_0, v_1}(\theta) \stackrel{def}{=} v_0 \sin \theta \cos \theta + \frac{v_1}{2} \cos^2 \theta \{1 - 2 \sin^2 \theta\}$$

for all $\theta \in \mathbb{R}$ and define

$$\bar{\ell}(v_0, v_1) \stackrel{def}{=} \int_0^{2\pi} \ell_{v_0, v_1}(\theta) u_{v_0, v_1}(\theta) d\theta.$$

Define

$$\mathbf{v}_0(y) \stackrel{\text{def}}{=} \alpha'(y) \quad \text{and} \quad \mathbf{v}_1(y) \stackrel{\text{def}}{=} \sum_{i=1}^2 \int_0^1 \left(\frac{\partial \sigma_i^2}{\partial x}(x, y) \right)^2 dx$$

for all $y \in \mathbb{R}$. The assumptions of Definition 1 ensure that $\mathbf{v}_0(y)$ is finite for all $y > 0$, and we now impose

Assumption 1 (Nondegeneracy). *For each $y \in \mathbb{R}$, $\mathbf{v}_1(y) > 0$.*

Definition 4. *Define*

$$\lambda(y) = \bar{\ell}(\mathbf{v}_0(y), \mathbf{v}_1(y))$$

for all $y \in \mathbb{R}$.

The following explanation of the parameters \mathbf{v}_1 and \mathbf{v}_2 was suggested by Peter Baxendale. As we discussed, the reason for the growth of the tangent flow is that nearby starting points will get kicked to different orbits, where they will separate due to the difference in angular speed (which is measure by $\mathbf{v}_0 = \alpha'$). The worst case is if the two starting points are on the same orbit; i.e., consider two starting points (x, y) and $(x + \delta, y)$. The primary reason why these two starting points would move to different orbits (remember that the same noise is driving both solutions of the SDE) is if the diffusion coefficients in the y direction are different at (x, y) and $(x + \delta, y)$; i.e., if $\frac{\partial \sigma_1^2}{\partial x}(x, y)$ or $\frac{\partial \sigma_2^2}{\partial x}(x, y)$ is nonzero. The quantity \mathbf{v}_1 measures the effective average of these quantities.

As the basis for all of our analysis, let's write down the SDE for the matrix $\Xi^{\varepsilon, x, y}$. For any $(x, y) \in \mathbb{R}^2$,

$$(3) \quad \begin{aligned} d\Xi_t^{\varepsilon, x, y} &= \varepsilon^{-2} \begin{pmatrix} 0 & \alpha'(Y_t^\varepsilon(x, y)) \\ 0 & 0 \end{pmatrix} \Xi_t^{\varepsilon, x, y} dt \\ &+ \sum_{i=0}^2 \begin{pmatrix} \frac{\partial \sigma_1^1}{\partial x}(X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y)) & \frac{\partial \sigma_1^1}{\partial y}(X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y)) \\ \frac{\partial \sigma_2^1}{\partial x}(X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y)) & \frac{\partial \sigma_2^1}{\partial y}(X_t^\varepsilon(x, y), Y_t^\varepsilon(x, y)) \end{pmatrix} \Xi_t^{\varepsilon, x, y} \circ d\gamma_t^i \\ \Xi_0^{\varepsilon, x, y} &= I. \end{aligned}$$

3. A PINSKY-WIHSTUTZ TRANSFORMATION AND THE KHASHMINSKII-FURSTENBURG FORMULA

We start by using the transformation of Pinsky and Wihstutz [13] to identify the dominant terms of (3). To help clarify things, let's set up some notation. Define the 2×2 transformation matrix

$$W_\varepsilon \stackrel{\text{def}}{=} \begin{pmatrix} \varepsilon^{2/3} & 0 \\ 0 & 1 \end{pmatrix}$$

and note that for any a, b, c , and d in \mathbb{R} ,

$$W_\varepsilon \begin{pmatrix} a & b \\ c & d \end{pmatrix} W_\varepsilon^{-1} = \begin{pmatrix} a & b\varepsilon^{2/3} \\ c\varepsilon^{-2/3} & d \end{pmatrix}.$$

Fix now a point $(x^*, y^*) \in \mathbb{R}^2$. For convenience, we henceforth set $X_t^\varepsilon = X_t^\varepsilon(x^*, y^*)$ and $Y_t^\varepsilon = Y_t^\varepsilon(x^*, y^*)$, and we define

$$\begin{aligned} Z_t^\varepsilon &\stackrel{\text{def}}{=} (X_t^\varepsilon, Y_t^\varepsilon) \\ \hat{\Xi}_t^\varepsilon &\stackrel{\text{def}}{=} W_\varepsilon \Xi_t^{\varepsilon, x^*, y^*} W_\varepsilon^{-1}, \end{aligned}$$

for all $t \geq 0$ and $\varepsilon > 0$. Define

$$\begin{aligned} S_0^\varepsilon(x, y) &\stackrel{\text{def}}{=} \begin{pmatrix} \varepsilon^{4/3} \frac{\partial \sigma_1^1}{\partial x}(x, y) & \alpha'(y) + \varepsilon^2 \frac{\partial \sigma_1^1}{\partial y}(x, y) \\ \varepsilon^{2/3} \frac{\partial \sigma_2^2}{\partial x}(x, y) & \varepsilon^{4/3} \frac{\partial \sigma_2^2}{\partial y}(x, y) \end{pmatrix} \\ S_i^\varepsilon(x, y) &\stackrel{\text{def}}{=} \begin{pmatrix} \varepsilon^{2/3} \frac{\partial \sigma_1^1}{\partial x}(x, y) & \varepsilon^{4/3} \frac{\partial \sigma_1^1}{\partial y}(x, y) \\ \frac{\partial \sigma_2^2}{\partial x}(x, y) & \varepsilon^{2/3} \frac{\partial \sigma_2^2}{\partial y}(x, y) \end{pmatrix} \quad i \in \{1, 2\} \end{aligned}$$

for all $(x, y) \in \mathbb{R}^2$ and $\varepsilon > 0$. We then have that

$$d\hat{\Xi}_t^\varepsilon = \varepsilon^{-4/3} S_0^\varepsilon(Z_t^\varepsilon) \hat{\Xi}_t^\varepsilon dt + \varepsilon^{-2/3} \sum_{i=1}^2 S_i^\varepsilon(Z_t^\varepsilon) \hat{\Xi}_t^\varepsilon \circ dW_t^i$$

for all $t \geq 0$ and $\varepsilon > 0$. Of course

Lemma 1. *For any nonzero $V \in \mathbb{R}^2 \setminus \{0\}$ and every $t \geq 0$, we have that*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \varepsilon^{4/3} \ln \frac{\|\Xi_t^{\varepsilon, x^*, y^*} V\|}{\|V\|} - \varepsilon^{4/3} \ln \frac{\|\hat{\Xi}_t^\varepsilon V\|}{\|V\|} \right|^p \right] = 0.$$

Proof. The transformation W_ε is polynomial in $\varepsilon^{2/3}$; such transformations are invisible under the desired logarithmic asymptotics. \square

Fix now $V \in \mathbb{R}^2 \setminus \{0\}$. The Khasminskii-Furstenburg formula allows us to study the evolution of $\|\Xi_t^{\varepsilon, x^*, y^*} V\|$. To write all of this out, we adopt most of the notation of [13]. Define $S^1 \stackrel{\text{def}}{=} \{x \in \mathbb{R}^2 : \|x\| = 1\}$ and let $\langle \cdot, \cdot \rangle$ denote the standard inner product in \mathbb{R}^2 . For any 2×2 matrix C , define

$$\begin{aligned} h(C, \theta) &\stackrel{\text{def}}{=} C\theta - \langle C\theta, \theta \rangle \theta \\ q_0(C, \theta) &\stackrel{\text{def}}{=} \langle C\theta, \theta \rangle \quad \theta \in S^1 \\ q_1(C, \theta) &\stackrel{\text{def}}{=} \langle C\theta, C\theta \rangle + \langle C^2\theta, \theta \rangle - 2\langle C\theta, \theta \rangle^2. \end{aligned}$$

Set

$$\theta_t^\varepsilon \stackrel{\text{def}}{=} \frac{\hat{\Xi}_t^\varepsilon V}{\|\hat{\Xi}_t^\varepsilon V\|}$$

for all $t \geq 0$ and $\varepsilon > 0$. Then we have

Proposition 3. *For each $t \geq 0$ and $\varepsilon > 0$,*

$$(4) \quad \ln \frac{\|\hat{\Xi}_t^\varepsilon V\|}{\|V\|} = \varepsilon^{-4/3} \int_0^t q_0(S_0^\varepsilon(Z_s^\varepsilon), \theta_s^\varepsilon) ds + \varepsilon^{-2/3} \sum_{i=1}^2 \int_0^t q_0(S_i^\varepsilon(Z_s^\varepsilon), \theta_s^\varepsilon) \circ dW_s^i$$

and

$$(5) \quad d\theta_t^\varepsilon = \varepsilon^{-4/3} h(S_0^\varepsilon(Z_t^\varepsilon), \theta_t^\varepsilon) dt + \varepsilon^{-2/3} \sum_{i=1}^2 h(S_i^\varepsilon(Z_t^\varepsilon), \theta_t^\varepsilon) \circ dW_t^i.$$

Proof. Straightforward. \square

Let's convert things into a more usable form; namely, let's convert (4) to Ito form and let's identify the generator of the $\mathbb{R}^2 \times S^1$ -valued process $(Z^\varepsilon, \theta^\varepsilon)$. First of all, define

$$S_i^{\varepsilon, \prime}(x, y) \stackrel{\text{def}}{=} \begin{pmatrix} \varepsilon^{2/3} \frac{\partial^2 \sigma_i^1}{\partial x^2} \sigma_i^1 + \varepsilon^{2/3} \frac{\partial^2 \sigma_i^1}{\partial x \partial y} \sigma_i^2 & \varepsilon^{4/3} \frac{\partial^2 \sigma_i^1}{\partial x \partial y} \sigma_i^1 + \varepsilon^{4/3} \frac{\partial^2 \sigma_i^1}{\partial y^2} \sigma_i^2 \\ \frac{\partial^2 \sigma_i^2}{\partial x^2} \sigma_i^1 + \frac{\partial^2 \sigma_i^2}{\partial x \partial y} \sigma_i^2 & \varepsilon^{2/3} \frac{\partial^2 \sigma_i^2}{\partial x \partial y} \sigma_i^1 + \varepsilon^{2/3} \frac{\partial^2 \sigma_i^2}{\partial y^2} \sigma_i^2 \end{pmatrix} (x, y)$$

for $i \in \{1, 2\}$, $(x, y) \in \mathbb{R}^2$, and $\varepsilon > 0$. Then

Corollary 1. *For each $t \geq 0$,*

$$\begin{aligned} \ln \frac{\|\hat{\Xi}_t^\varepsilon V\|}{\|V\|} &= \varepsilon^{-4/3} \int_0^t \left\{ q_0(S_0^\varepsilon(Z_s^\varepsilon), \theta_s^\varepsilon) + \frac{1}{2} \sum_{i=1}^2 q_1(S_i^\varepsilon(Z_s^\varepsilon), \theta_s^\varepsilon) \right\} ds \\ &+ \varepsilon^{-2/3} \sum_{i=1}^2 \int_0^t q_0(S_i^\varepsilon(Z_s^\varepsilon), \theta_s^\varepsilon) dW_s^i + \frac{\varepsilon^{-2/3}}{2} \sum_{i=1}^2 \int_0^t q_0(S_i^{\varepsilon, \prime}(Z_s^\varepsilon), \theta_s^\varepsilon) ds \end{aligned}$$

Let ∇ denote the standard Euclidean gradient operator. For any 2×2 matrix C , define the vector field (derivation) \mathbf{h}_C on S^1 by

$$(\mathbf{h}_C \varphi)(\theta) \stackrel{\text{def}}{=} \langle h(C, \theta), \nabla \tilde{\varphi}(\theta) \rangle$$

for any $\varphi \in C^\infty(S^1)$ and $\theta \in S^1$ and any $\tilde{\varphi} \in C^\infty(\mathbb{R} \setminus \{0\})$ such that $\tilde{\varphi}(\theta) = \varphi(\theta)$ for all $\theta \in S^1$. For each $z \in \mathbb{R}^2$ and $\varepsilon > 0$, define the second-order Hörmander-form differential operator $\mathcal{A}_z^\varepsilon$ on $C^\infty(S^1)$ as

$$\mathcal{A}_z^\varepsilon \stackrel{\text{def}}{=} \mathbf{h}_{S_0^\varepsilon(z)} + \frac{1}{2} \sum_{i=1}^2 \mathbf{h}_{S_i^\varepsilon(z)}^2 + \varepsilon^{2/3} \sum_{i=1}^2 \mathbf{h}_{S_i^{\varepsilon, \prime}(z)}.$$

Heuristically, $\varepsilon^{-4/3} \mathcal{A}_z^\varepsilon$ is the generator of θ^ε (if we fix Z^ε at z). In order to write down the evolution of the pair $(Z^\varepsilon, \theta^\varepsilon)$, we need to extend differential operators from $C^\infty(\mathbb{R}^2)$ (i.e. the σ_i 's and \mathcal{L}) and $C^\infty(S^1)$ (i.e. the \mathbf{h}_C 's and $\mathcal{A}_z^\varepsilon$) to $C^\infty(\mathbb{R}^2 \times S^1)$. We do so in the obvious way; if \mathcal{D} is a differential operator on $C^\infty(\mathbb{R}^2)$ and \mathcal{D}' is a differential operator on $C^\infty(S^1)$, we extend them to $C^\infty(\mathbb{R}^2 \times S^1)$ by the formulæ

$$(6) \quad (\mathcal{D}\varphi)(z, \theta) \stackrel{\text{def}}{=} (\mathcal{D}\varphi(\cdot, \theta))(z) \quad \text{and} \quad (\mathcal{D}'\varphi)(z, \theta) \stackrel{\text{def}}{=} (\mathcal{D}'\varphi(z, \cdot))(\theta)$$

where the “ \cdot ” denotes the variable with respect to which derivatives are being taken. To proceed with our calculations, we now define, for each $\varepsilon > 0$, the second-order differential operator \mathcal{D}^ε on $C^\infty(\mathbb{R}^2 \times S^1)$ as

$$(\mathcal{D}^\varepsilon \varphi)(z, \theta) \stackrel{\text{def}}{=} \sum_{i=1}^2 \{ (\mathbf{h}_{S_i^\varepsilon(z)} \sigma_i \varphi)(z, \theta) + (\sigma_i \mathbf{h}_{S_i^\varepsilon(z)} \varphi)(z, \theta) \}$$

for each $\varphi \in C^\infty(\mathbb{R}^2 \times S^1)$, $z \in \mathbb{R}^2$ and $\theta \in S^1$. In this formula, we interpret $\mathbf{h}_{S_i^\varepsilon(z)} \sigma_i \varphi$ to mean the action (as in (6)) of the vector field $\mathbf{h}_{S_i^\varepsilon(z)}$ on the function $(z, \theta) \mapsto (\sigma_i \varphi)(z, \theta)$ and we interpret $\sigma_i \mathbf{h}_{S_i^\varepsilon(z)} \varphi$ to mean the action (as in (6)) of the vector field σ_i on the function $(z, \theta) \mapsto (\mathbf{h}_{S_i^\varepsilon(z)} \varphi)(z, \theta)$. Combining (1) and (5)

and then converting to Ito form, we see that for any $\Phi \in C_b^\infty(\mathbb{R}^2 \times S^1)$,

$$\begin{aligned}
(7) \quad \Phi(Z_t^\varepsilon, \theta_t^\varepsilon) - \Phi(Z_0^\varepsilon, \theta_0^\varepsilon) &= \varepsilon^{-2} \int_0^t (\mathfrak{F}\Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) ds + \sum_{i=0}^3 \int_0^t (\sigma_i \Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) \circ d\gamma_s^i \\
&+ \varepsilon^{-4/3} \int_0^t (\mathbf{h}_{S_0^\varepsilon(Z_s^\varepsilon)} \Phi)(\theta_s^\varepsilon) ds + \varepsilon^{-2/3} \sum_{i=1}^2 \int_0^t (\mathbf{h}_{S_i^\varepsilon(Z_s^\varepsilon)} \Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) \circ dW_s^i \\
&= \varepsilon^{-2} \int_0^t (\mathfrak{F}\Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) ds + \int_0^t (\mathcal{L}\Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) ds + \sum_{i=1}^2 \int_0^t (\sigma_i \Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) dW_s^i \\
&+ \varepsilon^{-4/3} \int_0^t (\mathcal{A}_{Z_s^\varepsilon}^\varepsilon \Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) ds + \varepsilon^{-2/3} \sum_{i=1}^2 \int_0^t (\mathbf{h}_{S_i^\varepsilon(Z_s^\varepsilon)} \Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) dW_s^i \\
&+ \varepsilon^{-2/3} \sum_{i=1}^2 \int_0^t (\mathcal{D}^\varepsilon \Phi)(Z_s^\varepsilon, \theta_s^\varepsilon) ds
\end{aligned}$$

for all $t \geq 0$ and $\varepsilon > 0$.

4. ASYMPTOTICS

Let's start to identify the dominant parts of things. Define

$$\begin{aligned}
A_0 &\stackrel{\text{def}}{=} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A_1 \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\
\vartheta_0(x, y) &\stackrel{\text{def}}{=} \alpha'(y) \quad \text{and} \quad \vartheta_i(x, y) \stackrel{\text{def}}{=} \frac{\partial \sigma_i^2}{\partial x}(x, y) \quad i \in \{1, 2\}
\end{aligned}$$

for all $(x, y) \in \mathbb{R}^2$. Clearly there is an $M > 0$ such that

$$(8) \quad \begin{aligned}
&\sup_{z \in \mathbb{R}^2} \|S_0^\varepsilon(z) - \vartheta_0(z)A_0\|_{\text{op}} \leq M\varepsilon^{2/3} \\
&\sup_{z \in \mathbb{R}^2} \|S_i^\varepsilon(z) - \vartheta_i(z)A_i\|_{\text{op}} \leq M\varepsilon^{2/3} \quad i \in \{1, 2\}
\end{aligned}$$

for all $\varepsilon > 0$ (where here $\|\cdot\|_{\text{op}}$ is the norm on 2×2 matrices induced by $\|\cdot\|$). Note also that

$$\mathbf{v}_0(y) = \int_0^1 \vartheta_0(x, y) dx \quad \text{and} \quad \mathbf{v}_1(y) \stackrel{\text{def}}{=} \sum_{i=1}^2 \int_0^1 \vartheta_i^2(x, y) dx$$

for all $y \in \mathbb{R}$. Define now

$$\Lambda_1(z, \theta) \stackrel{\text{def}}{=} \vartheta_0(z)q_0(A_0, \theta) + \frac{1}{2} \left\{ \sum_{i=1}^2 \vartheta_i^2(z) \right\} q_1(A_1, \theta)$$

for all $z \in \mathbb{R}^2$ and $\theta \in S^1$; to enforce notational clarity, if $z = (x, y)$, then $\Lambda_1(x, y, \theta) = \Lambda_1(z, \theta)$. Then our first asymptotic simplification is

Lemma 2. *For any $p \geq 1$ and any $t \geq 0$,*

$$(9) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \varepsilon^{4/3} \ln \frac{\|\hat{\Xi}_t^\varepsilon V\|}{\|V\|} - \int_0^t \Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon) ds \right|^p \right] = 0.$$

Proof. Use (8). □

We can easily compare various speeds. The speed of Z^ε is ε^{-2} and the speed of θ^ε (using (5)) is $\varepsilon^{-4/3}$. Thus the X^ε component of Z^ε will make many rotations (mapping the X^ε coordinate to \mathbb{R}/\mathbb{Z}) before θ^ε moves any noticeable amount. Thus, we should first average out the X^ε component of Z^ε in (9). Define

$$\Lambda_2(y, \theta) \stackrel{\text{def}}{=} \int_0^1 \Lambda_1(x, y, \theta) dx;$$

we should be able to replace $\Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon)$ in Lemma 2 with $\Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon)$.

Let's now set up some notation to take care of a small difficulty. In order to replace Λ_1 by Λ_2 , we will use an auxiliary function Φ in (7). Unfortunately, we have no control over Φ for $y = \pm\infty$. We need to *localize*. For each $K > 0$ and $\varepsilon > 0$, define the stopping time

$$\tau_K^\varepsilon \stackrel{\text{def}}{=} \inf\{t \geq 0 : |Y_t^\varepsilon| \geq K\}.$$

Then (in light of the SDE for Y^ε in (1))

Lemma 3. *We have that*

$$\lim_{K \rightarrow \infty} \sup_{\varepsilon > 0} \mathbb{P}\{\tau_K^\varepsilon < t\} = 0.$$

We now can make the stated replacement.

Lemma 4. *For any $p \geq 1$ and any $t \geq 0$,*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \int_0^t \{\Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon) - \Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon)\} ds \right|^p \right] = 0.$$

Proof. First of all, note that for any $K > 0$ and $\varepsilon > 0$,

$$\begin{aligned} (10) \quad \mathbb{E} \left[\left| \int_0^t \{\Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon) - \Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon)\} ds \right|^p \right]^{1/p} \\ \leq \mathbb{E} \left[\left| \int_0^{t \wedge \tau_K^\varepsilon} \{\Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon) - \Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon)\} ds \right|^p \right]^{1/p} \\ + \mathbb{E} \left[\left| \int_{t \wedge \tau_K^\varepsilon}^t \{\Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon) - \Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon)\} ds \right|^p \right]^{1/p}. \end{aligned}$$

By Definitions 1 and 2 there is an $M > 0$ such that $|\Lambda_1(z, \theta)| \leq M$ and $|\Lambda_2(y, \theta)| \leq M$ for all $z \in \mathbb{R}^2$ and $y \in \mathbb{R}$. Thus,

$$(11) \quad \mathbb{E} \left[\left| \int_{t \wedge \tau_K^\varepsilon}^t \{\Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon) - \Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon)\} ds \right|^p \right]^{1/p} \leq 2Mt \mathbb{P}\{\tau_K^\varepsilon < t\}^{1/p}.$$

Next, define

$$\Psi(x, y, \theta) \stackrel{\text{def}}{=} \frac{1}{\alpha(y)} \left\{ \int_0^x \Lambda_1(x', y, \theta) dx' - x \Lambda_2(y, \theta) \right\}$$

for all $(x, y) \in \mathbb{R}^2$ and $\theta \in S^1$ (here is the place where we need to localize; as α tends to zero, Ψ tends to infinity). Note that $\Psi \in C^\infty(\mathbb{R}^2 \times S^1)$ and that

$$(\mathfrak{F}\Psi)(x, y, \theta) = \Lambda_1(x, y, \theta) - \Lambda_2(y, \theta)$$

for all $(x, y) \in \mathbb{R}^2$ and $\theta \in S^1$. Taking $\Phi = \varepsilon^2 \Psi$ in (7), we have that

$$\begin{aligned} \int_0^{t \wedge \tau_K^\varepsilon} \{\Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon) - \Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon)\} ds &= \varepsilon^2 \left\{ \Psi(Z_{t \wedge \tau_K^\varepsilon}^\varepsilon, \theta_{t \wedge \tau_K^\varepsilon}^\varepsilon) - \Psi(Z_0^\varepsilon, \theta_0^\varepsilon) \right\} \\ &\quad - \varepsilon^2 \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{L}\Psi)(Z_s^\varepsilon, \theta_s^\varepsilon) ds - \varepsilon^2 \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\sigma_i \Psi)(Z_s^\varepsilon, \theta_s^\varepsilon) dW_s^i \\ &\quad - \varepsilon^{2/3} \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{A}_{Z_s^\varepsilon}^\varepsilon \Psi)(Z_s^\varepsilon, \theta_s^\varepsilon) ds - \varepsilon^{4/3} \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\mathbf{h}_{S_i^\varepsilon}(Z_s^\varepsilon) \Psi)(Z_s^\varepsilon, \theta_s^\varepsilon) dW_s^i \\ &\quad - \varepsilon^{4/3} \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{D}^\varepsilon \Psi)(Z_s^\varepsilon, \theta_s^\varepsilon) ds. \end{aligned}$$

Letting ε tend to zero, we get that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \int_0^{t \wedge \tau_K^\varepsilon} \{\Lambda_1(Z_s^\varepsilon, \theta_s^\varepsilon) - \Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon)\} ds \right|^p \right] = 0$$

for each $K > 0$. Use this and (11) to first let ε tend to zero and then let K tend to infinity in (10). This gives the stated result. \square

Next, we average over θ^ε . Of course to do so, we should identify the relevant reference measure for θ^ε . Define first the second-order Hörmander-form operator

$$\mathcal{A}_z^0 \stackrel{\text{def}}{=} \vartheta_0(z) \mathbf{h}_{A_0} + \frac{1}{2} \left\{ \sum_{i=1}^2 \vartheta_i^2(z) \right\} \mathbf{h}_{A_1}^2$$

for each $z \in \mathbb{R}^2$. Note that for any $\varphi \in C^\infty(S^1)$, there is an $M > 0$ such that

$$\sup_{\substack{z \in \mathbb{R}^2 \\ \theta \in S^1}} |(\mathcal{A}_z^0 \varphi)(\theta) - (\mathcal{A}_z^\varepsilon \varphi)(\theta)| \leq M \varepsilon^{2/3}.$$

Thus, θ^ε roughly evolves according to $\varepsilon^{-4/3} \mathcal{A}_{Z_s^\varepsilon}^0$. Again, we are faced with the fact that Z^ε moves much faster than θ^ε . Thus, we should define the averaged second-order Hörmander-form operator

$$(12) \quad \vec{\mathcal{A}}_y \stackrel{\text{def}}{=} \mathbf{v}_0(y) \mathbf{h}_{A_0} + \frac{1}{2} \mathbf{v}_1(y) \mathbf{h}_{A_1}^2$$

for all $y \in \mathbb{R}$; of course

$$(\vec{\mathcal{A}}_y \varphi)(\theta) = \int_{x=0}^1 (\mathcal{A}_{x,y}^0 \varphi)(\theta) dx$$

for all $y \in \mathbb{R}$ and $\varphi \in C^\infty(S^1)$. Then $\vec{\mathcal{A}}_y$ is a hypoelliptic operator on S^1 for each $y \in \mathbb{R}$ (see [1] or [13] or the calculations of Section 5). Thus, for each $y \in \mathbb{R}$ there is a unique $\bar{\mu}_y \in \mathcal{P}(S^1)$ such that

$$\int_{z \in S^1} (\vec{\mathcal{A}}_y \varphi)(z) \bar{\mu}_y(dz) = 0$$

for all $\varphi \in C^\infty(S^1)$. Finally, define

$$(13) \quad \lambda(y) \stackrel{\text{def}}{=} \int_{\theta \in S^1} \Lambda_2(y, \theta) \bar{\mu}_y(d\theta).$$

(we shall see in Section 5 that this definition coincides with that of Definition 4). We then have

Lemma 5. *For any $p \geq 1$ and any $t \geq 0$,*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \int_0^t \{\Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon) - \lambda(Y_s^\varepsilon)\} ds \right|^p \right] = 0.$$

Proof. Of course we need only prove that for any $K > 0$

$$(14) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \int_0^{t \wedge \tau_K^\varepsilon} \{\Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon) - \lambda(Y_s^\varepsilon)\} ds \right|^p \right] = 0.$$

By the Fredholm alternative and standard smoothness results, for each $y \in \mathbb{R}$ there is a $\hat{\Psi}_y \in C^\infty(S^1)$ such that

$$(\mathcal{A}_y \hat{\Psi}_y)(y, \theta) = \Lambda_2(y, \theta) - \lambda(y)$$

for all $y \in \mathbb{R}$ (again, here is the need for localization; if $\mathbf{v}_1(y)$ is close to zero, the PDE for $\hat{\Psi}_y$ becomes a singular perturbation problem). Define $\Psi_1(x, y, \theta) \stackrel{\text{def}}{=} \hat{\Psi}_y(\theta)$. Note that since Ψ_1 does not depend on the first coordinate x , $\mathfrak{F}\Psi_1 \equiv 0$ and furthermore

$$\int_{x=0}^1 (\mathcal{A}_{x,y}^0 \Psi_1)(x, y, \theta) dx = (\mathcal{A}_y \hat{\Psi}_y)(\theta) = \Lambda_2(y, \theta) - \lambda(y)$$

for all $y \in \mathbb{R}$ and $\theta \in S^1$. Setting $\Phi = \varepsilon^{4/3} \Psi_1$ in (7), we can calculate that

$$\begin{aligned} & \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{A}_{Z_s^\varepsilon}^\varepsilon \Psi_1)(Z_s^\varepsilon, \theta_s^\varepsilon) ds \\ &= \varepsilon^{4/3} \left\{ \Psi_1(Z_{t \wedge \tau_K^\varepsilon}^\varepsilon, \theta_{t \wedge \tau_K^\varepsilon}^\varepsilon) - \Psi_1(Z_0^\varepsilon, \theta_0^\varepsilon) \right\} - \varepsilon^{4/3} \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{L} \Psi_1)(Z_s^\varepsilon, \theta_s^\varepsilon) ds \\ & - \varepsilon^{4/3} \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\sigma_i \Psi_1)(Z_s^\varepsilon, \theta_s^\varepsilon) dW_s^i - \varepsilon^{2/3} \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\mathbf{h}_{S_i^\varepsilon}(Z_s^\varepsilon) \Psi_1)(Z_s^\varepsilon, \theta_s^\varepsilon) dW_s^i \\ & \quad - \varepsilon^{2/3} \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{D}^\varepsilon \Psi_1)(Z_s^\varepsilon, \theta_s^\varepsilon) ds. \end{aligned}$$

Thus

$$(15) \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{A}_{Z_s^\varepsilon}^0 \Psi_1)(Z_s^\varepsilon, \theta_s^\varepsilon) ds \right|^p \right] = 0.$$

Now define

$$\Psi_2(x, y, \theta) \stackrel{\text{def}}{=} \frac{1}{\alpha(y)} \left\{ \int_0^x (\mathcal{A}_{x',y}^0 \Psi_1)(x', y, \theta) dx' - x \{ \Lambda_2(y, \theta) - \lambda(y) \} \right\}$$

for all $(x, y) \in \mathbb{R}^2$ and $\theta \in S^1$. Then

$$(\mathfrak{F} \Psi_2)(x, y, \theta) = (\mathcal{A}_{x,y}^0 \Psi_1)(x, y, \theta) - \{ \Lambda_2(y, \theta) - \lambda(y) \}$$

for all $(x, y) \in \mathbb{R}^2$ and $\theta \in S^1$. Again using (7) (with $\Phi = \varepsilon^2 \Psi_2$), we get that

$$\begin{aligned} & \int_0^{t \wedge \tau_K^\varepsilon} \left\{ (\mathcal{A}_{Z_s^\varepsilon}^0 \Psi_1)(Z_s^\varepsilon, \theta_s^\varepsilon) - \{\Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon) - \lambda(Y_s^\varepsilon)\} \right\} ds \\ &= \varepsilon^2 \left\{ \Psi_2(Z_{t \wedge \tau_K^\varepsilon}^\varepsilon, \theta_{t \wedge \tau_K^\varepsilon}^\varepsilon) - \Psi_2(Z_0^\varepsilon, \theta_0^\varepsilon) \right\} - \varepsilon^2 \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{L} \Psi_2)(Z_s^\varepsilon, \theta_s^\varepsilon) ds \\ & \quad - \varepsilon^2 \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\sigma_i \Psi_2)(Z_s^\varepsilon, \theta_s^\varepsilon) dW_s^i - \varepsilon^{2/3} \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{A}_{Z_s^\varepsilon}^\varepsilon \Psi_2)(Z_s^\varepsilon, \theta_s^\varepsilon) ds \\ & \quad - \varepsilon^{4/3} \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\mathbf{h}_{S_i^\varepsilon(Z_s^\varepsilon)} \Psi_2)(Z_s^\varepsilon, \theta_s^\varepsilon) dW_s^i - \varepsilon^{4/3} \sum_{i=1}^2 \int_0^{t \wedge \tau_K^\varepsilon} (\mathcal{D}^\varepsilon \Psi_2)(Z_s^\varepsilon, \theta_s^\varepsilon) ds. \end{aligned}$$

This implies that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\left| \int_0^{t \wedge \tau_K^\varepsilon} \left\{ (\mathcal{A}_{Z_s^\varepsilon}^0 \Psi_1)(Z_s^\varepsilon, \theta_s^\varepsilon) - \{\Lambda_2(Y_s^\varepsilon, \theta_s^\varepsilon) - \lambda(Y_s^\varepsilon)\} \right\} ds \right|^p \right] = 0.$$

Combine this with (15) to get (14). \square

This completes the proof of the main result.

Proof of Main Theorem. Combine together Lemmas 1, 2, 4, and 5. \square

5. EXPLICIT FORMULÆ

Let's now extract the formulæ of Proposition 2 and Definitions 3 and 4. To do so, we want to map various objects (viz. differential operators and measures) from S^1 to \mathbb{R} . To start, define $\Theta : \mathbb{R} \rightarrow S^1$ by

$$\Theta(\varsigma) \stackrel{\text{def}}{=} (\cos \varsigma, \sin \varsigma). \quad \varsigma \in \mathbb{R}$$

Fix next a $\varphi \in C^\infty(\mathbb{R}^2 \setminus \{0\})$ and define

$$(16) \quad \tilde{\varphi}(\varsigma) = \varphi(\Theta(\varsigma)).$$

An easy computation shows that

$$\tilde{\varphi}'(\varsigma) = -\sin \varsigma \frac{\partial \varphi}{\partial x}(\Theta(\varsigma)) + \cos \varsigma \frac{\partial \varphi}{\partial y}(\Theta(\varsigma)).$$

It is fairly easy to see that

$$\begin{aligned} h(A_0, \Theta(\varsigma)) &= (\sin^3 \varsigma, -(\sin^2 \varsigma)(\cos \varsigma)) \\ h(A_1, \Theta(\varsigma)) &= (-(\cos^2 \varsigma)(\sin \varsigma), \cos^3 \varsigma) \end{aligned}$$

for all $\varsigma \in \mathbb{R}$. This implies that

$$(\mathbf{h}_{A_0} \varphi)(\Theta(\varsigma)) = -(\sin^2 \varsigma) \tilde{\varphi}'(\varsigma) \quad (\mathbf{h}_{A_1} \varphi)(\Theta(\varsigma)) = (\cos^2 \varsigma) \tilde{\varphi}'(\varsigma)$$

for all $\varsigma \in \mathbb{R}$. For each $v_1 \in \mathbb{R}$ and $v_2 > 0$, define now the second-order operator $\mathcal{A}_{v_0, v_1}^{\mathbb{R}}$ on $C^\infty(\mathbb{R})$ by

$$(\mathcal{A}_{v_0, v_1}^{\mathbb{R}} \hat{\varphi})(\varsigma) = -v_0 (\sin^2 \varsigma) \hat{\varphi}'(\varsigma) + \frac{1}{2} v_1 (\cos^2 \varsigma) (\cos^2 \hat{\varphi}')'(\varsigma)$$

for all $\varsigma \in \mathbb{R}$ and $\hat{\varphi} \in C^\infty(\mathbb{R})$. Here the term $(\cos^2 \hat{\varphi}')'$ means the derivative of the map $\varsigma \mapsto (\cos^2 \varsigma) \hat{\varphi}'(\varsigma)$. Thus we have

Lemma 6. *For each $v_0 \in \mathbb{R}$ and $v_1 > 0$, the PDE of Proposition 2 has a unique C^∞ solution. Secondly, the map $(v_0, v_1, \varsigma) \mapsto u_{v_0, v_1}(\varsigma)$ is C^∞ . Finally,*

$$\mu_y(A) = \int_{\substack{\varsigma \in [0, 2\pi] \\ \Theta(\varsigma) \in A}} u_{v_0(y), v_1(y)}(\varsigma) d\varsigma$$

for all $A \in \mathcal{B}(S^1)$ and all $y \in \mathbb{R}$.

Proof. From the above calculations and (12), we have that

$$(\mathcal{A}_y \bar{\varphi})(\Theta(\varsigma)) = \left(\mathcal{A}_{v_0(y), v_1(y)}^{\mathbb{R}} \bar{\varphi} \right)(\varsigma)$$

if we have (16). It is fairly easy to see that $\mathcal{A}_{v_0, v_1}^{\mathbb{R}}$ is hypoelliptic for each $v_0 \in \mathbb{R}$ and $v_1 > 0$ ($(\mathbf{h}_{A_1} \varphi)(\Theta(\varsigma)) = 0$ for all φ if and only if $\varsigma = \pi(k + 1/2)$ for some integer k , but $([\mathbf{h}_{A_0}, [\mathbf{h}_{A_0}, \mathbf{h}_{A_1}]] \varphi)(\Theta((k + 1/2)\pi)) = 0$ only if $\hat{\varphi}'((k + 1/2)\pi) = 0$). Standard estimates imply that u_{v_0, v_1} is well-defined and C^∞ (see, for example [4] and the references therein) and $(v_0, v_1, \varsigma) \mapsto u_{v_0, v_1}(\varsigma)$ is indeed C^∞ (see, for example, [10]). \square

To get to the explicit formula of Definition 4, we calculate that

$$q_0(A_0, \Theta(\varsigma)) = (\sin \varsigma)(\cos \varsigma) \quad \text{and} \quad q_1(A_1, \Theta(\varsigma)) = \cos^2 \varsigma - 2(\cos^2 \varsigma)(\sin^2 \varsigma)$$

for all $\varsigma \in \mathbb{R}$ and use (13).

REFERENCES

1. L. Arnold, E. Oeljeklaus, and É. Pardoux, *Almost sure and moment stability for linear Ito equations*, Lyapunov exponents (Bremen, 1984), Springer, Berlin, 1986, pp. 129–159.
2. Peter Baxendale, *Brownian motions in the diffeomorphism group. I*, Compositio Math. **53** (1984), no. 1, 19–50.
3. ———, *Stability along trajectories at a stochastic bifurcation point*, Stochastic dynamics (Bremen, 1997), Springer, New York, 1999, pp. 1–25.
4. Denis R. Bell, *The Malliavin calculus*, Longman Scientific & Technical, Harlow, 1987.
5. K. D. Elworthy, *Stochastic differential equations on manifolds*, Cambridge University Press, Cambridge, 1982.
6. Mark I. Freidlin and Alexander D. Wentzell, *Random perturbations of Hamiltonian systems*, Mem. Amer. Math. Soc. **109** (1994), no. 523, viii+82.
7. R. Z. Has'minskiĭ, *Diffusion processes with a small parameter*, Izv. Akad. Nauk SSSR Ser. Mat. **27** (1963), 1281–1300.
8. ———, *A limit theorem for solutions of differential equations with a random right hand part*, Teor. Veroyatnost. i Primenen **11** (1966), 444–462.
9. ———, *On the principle of averaging the Ito's stochastic differential equations*, Kybernetika (Prague) **4** (1968), 260–279.
10. Tosio Kato, *Perturbation theory for linear operators*, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition.
11. Hiroshi Kunita, *Stochastic flows and stochastic differential equations*, Cambridge University Press, Cambridge, 1990.
12. David Nualart, *Analysis on Wiener space and anticipating stochastic calculus*, Lectures on probability theory and statistics (Saint-Flour, 1995), Springer, Berlin, 1998, pp. 123–227.
13. M. A. Pinsky and V. Wihstutz, *Lyapunov exponents of nilpotent Ito systems*, Stochastics **25** (1988), no. 1, 43–57.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, IL 61801

E-mail address: r-sowers@math.uiuc.edu