Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Stochastic Processes and their Applications 119 (2009) 3549–3582
stochastic processes and their applications
www.elsevier.com/locate/spa

Averaging of stochastic flows: Twist maps and escape from resonance☆

Richard B. Sowers*

Department of Mathematics, University of Illinois at Urbana—Champaign, Urbana, IL 61801, United States

Received 27 February 2008; received in revised form 18 February 2009; accepted 24 April 2009
Available online 26 June 2009

Abstract

Our setup is a classical stochastic averaging one studied by Has'minskiĭ, which is a two-dimensional SDE (on a cylinder) consisting of a fast angular drift and a slow axial diffusion. We seek to understand the asymptotics of the flow generated by this SDE. To do so, we fix n initial points on the cylinder and consider the axial components of the trajectories evolving from these points. We conclude a propagation-of-chaos. There are two components of the limiting n-point motion: a common Brownian motion, and n independent Brownian motions, one for each initial point.

MSC: 60H20; 60F17; 93E15; 34C29

Keywords: Stochastic averaging; Escape from resonance; Stochastic flows of diffeomorphisms

1. Introduction

The goal of this paper is to understand some of the fine structure of stochastic averaging. We will focus on a particularly simple situation. Let $\omega \in C^\infty(\mathbb{R})$ be such that $\dot{\omega} > 0$ at all points of $\mathbb{R}$ and such that for some ω_- and ω_+ in $(0,\infty)$, $\omega_- \le \omega(x) \le \omega_+$ for all $x \in \mathbb{R}$ (think of $\omega(x) = \arctan(x) + \pi$). Secondly, fix σ in

$$C^\infty(\mathbb{S}^1) \overset{\text{def}}{=} \left\{\varphi \in C^\infty(\mathbb{R}) : \varphi(\theta) = \varphi(\theta+1) \text{ for all } \theta \in \mathbb{R}\right\}.$$

☆ This material is based upon work supported by the National Science Foundation under Grant Nos. 0305925 and 0604249.
* Tel.: +217 333 3350; fax: +217 333 9576.
E-mail address: r-sowers@math.uiuc.edu.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2009.04.004

Suppose that $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability triple on which a standard Brownian motion W is defined. Fix $\theta_\circ$ and $x_\circ$ in $\mathbb{R}$. For each $\varepsilon \in (0,1)$, consider the $\mathbb{R}^2$-valued SDE

$$\begin{aligned} d\theta_t^\varepsilon &= \frac{1}{\varepsilon}\omega(X_t^\varepsilon)dt \\ dX_t^\varepsilon &= \sigma(\theta_t^\varepsilon)dW_t \\ (\theta_0^\varepsilon, X_0^\varepsilon) &= (\theta_\circ, x_\circ) \end{aligned} \tag{1}$$

It is well known that X^ε converges in distribution as $\varepsilon \searrow 0$. Define

$$\hat{\sigma} \overset{\text{def}}{=} \left\{ \int_{\theta=0}^1 \sigma^2(\theta)d\theta \right\}^{1/2},$$

and for each $f \in C^2(\mathbb{R})$, define

$$\mathscr{L}_1 f = \frac{1}{2}\hat{\sigma}^2 \ddot{f}.$$

Theorem 1.1 (*Has'minskiĭ [7]*). *We have that* $\lim_{\varepsilon \searrow 0} X^\varepsilon = X$ *(in law; i.e., in* $\mathscr{P}(C([0,\infty);\mathbb{R}))$*) where* X *is Markov with generator* $\mathscr{L}_1$ *and initial distribution* $\delta_{x_\circ}$ *(in other words,* $X = x_\circ + \hat{\sigma} W$*, where* W *is a Brownian motion).*

Our goal is to understand this result from a perspective of *stochastic flows*. Define vector fields $\mathbf{V}_0$ and $\mathbf{V}_1$ on $\mathbb{R}^2$ as

$$(\mathbf{V}_0 f)(\theta, x) = \omega(x)\frac{\partial f}{\partial \theta}(\theta, x) \quad \text{and} \quad (\mathbf{V}_1 f)(\theta, x) = \sigma(\theta)\frac{\partial f}{\partial x}(\theta, x)$$

for all $f \in C^1(\mathbb{R}^2)$ and $(\theta, x) \in \mathbb{R}^2$. For each $\varepsilon \in (0,1)$, let $\{\phi_t^\varepsilon;\ t \geq 0\}$ be the $\text{Diff}(\mathbb{R}^2)$-valued[1] stochastic process such that

$$\begin{aligned} d\phi_t^\varepsilon &= \frac{1}{\varepsilon}\mathbf{V}_0(\phi_t^\varepsilon)dt + \mathbf{V}_1(\phi_t^\varepsilon) \circ dW_t \quad t > 0 \\ \phi_0^\varepsilon &= \text{id} \end{aligned}$$

where id is the identity map (we have used Stratonovich integration here out of respect for the established notation for the theory of stochastic flows on manifolds [5]; it is easy to see that in this case Stratonovich and Ito integrals coincide since the coefficient in $\mathbf{V}_1$ depends only on the angle). Define $\pi_2(p) \overset{\text{def}}{=} x$ for all $p = (\theta, x) \in \mathbb{R}^2$, and define $\Phi_t^\varepsilon \overset{\text{def}}{=} \pi_2 \circ \phi_t^\varepsilon$ for all $\varepsilon \in (0,1)$ and $t \geq 0$. Then the averaging result of Theorem 1.1 is that, defining $p_\circ \overset{\text{def}}{=} (\theta_\circ, x_\circ)$, the law of $\{\Phi_t^\varepsilon(p_\circ);\ t \geq 0\}$ converges, as $\varepsilon \searrow 0$, to an $\mathbb{R}$-valued Markov process with generator $\mathscr{L}_1$ and (naturally) initial distribution $\delta_{\pi_2(p_\circ)}$.

The more general question that we hope to investigate is: what is the limiting distribution, as $\varepsilon \searrow 0$, of $\{\Phi_t^\varepsilon;\ t \geq 0\}$? Namely, for each $\varepsilon \in (0,1)$, $\{\Phi_t^\varepsilon;\ t \geq 0\}$ is a stochastic process in

$$C^\infty(\mathbb{S}^1 \times \mathbb{R}) \overset{\text{def}}{=} \left\{ \varphi \in C^\infty(\mathbb{R} \times \mathbb{R}) : \varphi(\theta, x) = \varphi(\theta + 1, x) \text{ for all } (\theta, x) \in \mathbb{R} \times \mathbb{R} \right\}.$$

[1] For any C^∞ manifold M, $\text{Diff}(M)$ is the group of C^∞ diffeomorphisms of M.

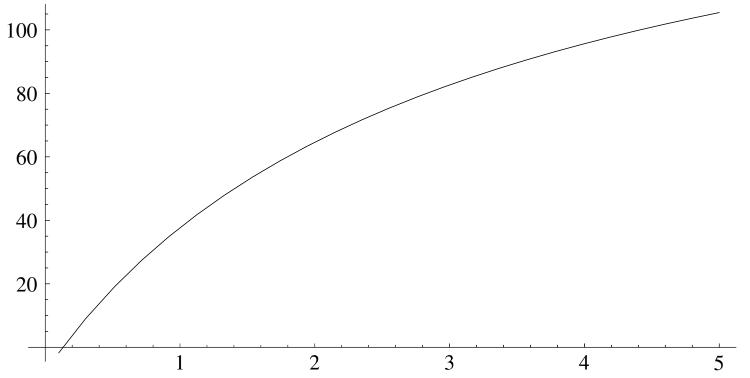

Fig. 1. Graph of $\mathcal{J}$.

Does it converge as $\varepsilon \searrow 0$? Since we can think of Φ^ε as a random map from $\mathbb{R} \times \mathbb{R}$ into $C([0,\infty);\mathbb{R})$ (i.e., map an initial point into an axial trajectory), we should naturally look at the effect of this map on a finite collection of points in $\mathbb{R} \times \mathbb{R}$ (i.e., a finite-dimensional approximation). Our main result is the following. For p and p' in $\mathbb{R}^2$, we say that $p \sim p'$ if $p - p' \in \mathbb{Z} \times \{0\}$. Define also

$$\overline{\sigma} \overset{\text{def}}{=} \int_{\theta=0}^{1} \sigma(\theta)\mathrm{d}\theta.$$

We now need a *escape from resonance* condition. For $\lambda \in \mathbb{R}$, define

$$\mathcal{J}_\circ(\lambda) \overset{\text{def}}{=} \mathrm{e}^{-\lambda^3} \int_{\varsigma=-\infty}^{\lambda} \mathrm{e}^{\varsigma^3}\mathrm{d}\varsigma = \frac{1}{3}\int_{\varsigma=0}^{\infty} \frac{\mathrm{e}^{-\varsigma}}{(\lambda^3-\varsigma)^{2/3}}\mathrm{d}\varsigma. \tag{2}$$

From (16) below, we have that $\mathcal{J}_\circ \in L^1(\mathbb{R})$. Next define

$$\kappa_1 \overset{\text{def}}{=} \left\{ \int_{\vartheta=0}^{1} (\dot{\sigma}(\vartheta))^2 \mathrm{d}\vartheta \right\}^{1/2} \tag{3}$$

and

$$\kappa_2(x) \overset{\text{def}}{=} \left(\frac{2\dot{\omega}(x)}{3\kappa_1^2} \right)^{1/3} \qquad x \in \mathbb{R}. \tag{4}$$

Also, define

$$\mathcal{J}(\varkappa) \overset{\text{def}}{=} \int_{\lambda\in\mathbb{R}} \left\{ \frac{3\varkappa\lambda}{\varkappa+\lambda^2} + \frac{\varkappa-\lambda^2}{(\varkappa+\lambda^2)^2} \right\} \mathcal{J}_\circ(\lambda)\mathrm{d}\lambda$$

for all $\varkappa > 0$. See also [4]. It is easy to see that $\lim_{\varkappa\searrow 0} \mathcal{J}(\varkappa) < 0$. Numerical integration shows that $\lim_{\varkappa\nearrow\infty} \mathcal{J}(\varkappa) > 0$ and a plot of $\mathcal{J}$ is given in Fig. 1. Numerical integration also shows that $\mathcal{J}(\varkappa) > 0$ if $\varkappa > \varkappa_c \approx 0.13064$. Our main result is then the following. Fix $K \subset\subset \mathbb{R}$ such that $\inf_{x\in K} \mathcal{J}(\kappa_2^2(x)) > 0$.

Theorem 1.2. *Fix $\{p_1, p_2 \dots p_n\} \subset \mathbb{R} \times \mathbb{R}$ such that $p_i \not\sim p_j$ if $i \neq j$ and define*

$$\tau_K^\varepsilon \overset{\text{def}}{=} \inf\{t > 0 : \Phi_t^\varepsilon(p_i) \notin K \text{ for some } i \in \{1, 2 \dots n\}\}$$

(in other words, τ_K^ε is the first time that one of the $\Phi^\varepsilon(p_i)$'s leaves K). Then

$$\{(\Phi^\varepsilon_{t\wedge\tau^\varepsilon_K}(p_1), \Phi^\varepsilon_{t\wedge\tau^\varepsilon_K}(p_2) \dots \Phi^\varepsilon_{t\wedge\tau^\varepsilon_K}(p_n));\ t \geq 0\}$$

converges in law (i.e., in the topology of $\mathscr{P}(C([0,\infty]); \mathbb{R}^n)$) as $\varepsilon \searrow 0$ to $\{(X^1_{t\wedge\hat\tau_K}, X^2_{t\wedge\hat\tau_K} \dots X^n_{t\wedge\hat\tau_K});\ t \geq 0\}$, where

$$X_t^i = \pi_2(p_i) + \overline{\sigma} V_t + \sqrt{\hat\sigma^2 - \overline{\sigma}^2}\, W_t^i \tag{5}$$

with V and the W^i's being independent standard Wiener processes and where

$$\hat\tau_K \overset{\text{def}}{=} \inf\{t > 0 : X_t^i \notin K \text{ for some } i \in \{1, 2 \dots n\}\}.$$

By Jensen's inequality, $\hat\sigma^2 \geq \overline{\sigma}^2$. The law of $\{(X_t^1, X_t^2 \dots X_t^n);\ t \geq 0\}$ is that of a d-dimensional Markov process with generator

$$(\mathscr{L}_d f)(\bar p) \overset{\text{def}}{=} \frac{1}{2}\hat\sigma^2 \sum_{1\leq i\leq d} \frac{\partial^2 f}{\partial p_i^2}(\bar p) + \overline{\sigma}^2 \sum_{\substack{1\leq i,j\leq d\\ i\neq j}} \frac{\partial^2 f}{\partial p_i \partial p_j}(\bar p) \tag{6}$$

for all $\bar p = (p_1, p_2 \dots p_d) \in \mathbb{R}^d$ and all $f \in C^2(\mathbb{R}^d)$. We note that thus the limit does *not* correspond to the generator of an evolution in $\text{Diff}^\infty(\mathbb{R})$ (in contrast to the case of [2]). Namely, for distinct p_1 and p_2 in $\mathbb{R} \times \mathbb{R}$, if X^1 and X^2 are given as in (5), then $X_t^1 - X_t^2 = \pi_2(p_1) - \pi_2(p_2) + \sqrt{\hat\sigma^2 - \overline{\sigma}^2}\{W_t^1 - W_t^2\}$; then $X^1 - X^2$ will hit zero at some (random) time t^*. But this is not possible if $X^1_{t^*} = \tilde\varphi(\pi_2(z_1))$ and $X^2_{t^*} = \tilde\varphi(\pi_2(z_2))$ for some (random) diffeomorphism $\tilde\varphi$. In fact, Theorem 1.2 tells us that the limit of Φ^ε should in some sense have an *uncountable* amount of randomness. Namely, fix a standard Brownian motion V and an uncountable collection $\{W^{\mathsf{p}} : \mathsf{p} \in \mathbb{R}^2 \backslash \sim\}$ of independent standard Brownian motions which are also independent of V. Then define $\Phi_t(p) \overset{\text{def}}{=} \pi_2(p) + \overline{\sigma} V_t + \{\hat\sigma^2 - \overline{\sigma}^2\} W_t^{\mathsf{p}}$ for all $p \in \mathsf{p}$ (where, as usual, $\mathsf{p} \in \mathbb{R}^2 \backslash \sim$ is an equivalence class). Then in some sense Φ^ε converges to Φ. We will not make this precise in this paper. We point out that a rigorous proof of this would involve a number of topological complications. For each $t > 0$, Φ_t is a map from $\mathbb{R}^2 \backslash \sim$ to $\mathbb{R}^\infty$; Φ should take values in the space of trajectories whose values are such maps.

The characterization of stochastic flows as diffusions in the diffeomorphism group is due to Baxendale in [1]. In our case that theory takes on the following guise. Define $a : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ as

$$a(p, p') \overset{\text{def}}{=} \begin{cases} \hat\sigma^2 & \text{if } p \sim p' \\ \overline{\sigma}^2 & \text{if } p \not\sim p'. \end{cases}$$

Then the generator of (6) can be written as

$$(\mathscr{L}_d f)(\bar p) \overset{\text{def}}{=} \frac{1}{2} \sum_{1\leq i\leq d} a(p_i, p_i) \frac{\partial^2 f}{\partial p_i^2}(\bar p) + \sum_{\substack{1\leq i,j\leq d\\ i\neq j}} a(p_i, p_j) \frac{\partial^2 f}{\partial p_i \partial p_j}(\bar p)$$

$$\bar p = (p_1, p_2 \dots p_n) \in \mathbb{R}^d.$$

We note that a is discontinuous (at the diagonal); thus the reproducing kernel Hilbert space associated with a is not separable, so we cannot make a decomposition $a(x, y) = \sum_{j=1}^{\infty} e_j(x) e_j(y)$ for some countable collection $\{e_j;\ j \in \mathbb{N}\}$ of functions. This means that the limiting dynamics are not those of a flow of diffeomorphisms. More precisely, let $\pi_2^* \mathrm{Diff}(\mathbb{R})$ be the collection of functions from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$ of the form $f \circ \pi_2$, where $f \in \mathrm{Diff}(\mathbb{R})$. In the case of [2], where ω was constant, $\{\Phi_t^\varepsilon;\ t \geq 0\}$, which is a diffusion in $C^\infty(\mathbb{S}^1 \times \mathbb{R})$, in fact converged (weakly, in the topology of probability measures on $C([0, \infty); C^\infty(\mathbb{S}^1 \times \mathbb{R}))$ to a diffusion in (the smaller space) $\pi_2^* \mathrm{Diff}(\mathbb{R})$. Here $\{\Phi_t^\varepsilon;\ t \geq 0\}$ does not converge (in fact, it does not converge in $C^\infty(\mathbb{S}^1 \times \mathbb{R})$).

The origins of $\hat{\sigma}$ and $\overline{\sigma}$ are natural. Fix $\{p_1, p_2 \ldots p_d\} \subset \mathbb{R} \times \mathbb{R}$ as in the statement of Theorem 1.2. For all $t \geq 0$ and $\varepsilon \in (0, 1)$, let $\vartheta_t^{i,\varepsilon}$ and $Z_t^{i,\varepsilon}$ be the angular and axial components of $\phi_t^\varepsilon(p)$; i.e., $\phi_t^\varepsilon(p) = (\vartheta_t^{i,\varepsilon}, Z_t^{i,\varepsilon})$ (thus $Z_t^{i,\varepsilon} = \Phi_t^\varepsilon(p_i)$). Fix also $f \in C_b^2(\mathbb{R}^d)$. Then for $\varepsilon \in (0, 1)$

$$M_t^\varepsilon \stackrel{\text{def}}{=} f(Z_t^{1,\varepsilon}, Z_t^{2,\varepsilon} \ldots Z_t^{i,\varepsilon}) - \int_{s=0}^t \left\{ \sum_{1 \leq i \leq d} \frac{1}{2} \sigma^2(\vartheta_s^{i,\varepsilon}) \frac{\partial^2 f}{\partial z_i^2}(Z_s^{1,\varepsilon}, Z_s^{2,\varepsilon} \ldots Z_s^{i,\varepsilon}) \right. \\ \left. + \sum_{\substack{1 \leq i,j \leq d \\ i \neq j}} \sigma(\vartheta_s^{i,\varepsilon}) \sigma(\vartheta_s^{j,\varepsilon}) \frac{\partial^2 f}{\partial z_i \partial z_j}(Z_s^{1,\varepsilon}, Z_s^{2,\varepsilon} \ldots Z_s^{i,\varepsilon}) \right\} \mathrm{d}s \tag{7}$$

is a martingale. Recall that the $Z^{i,\varepsilon}$'s are the *slow* variables (the "actions") while the $\vartheta^{i,\varepsilon}$'s are the *fast* variables (the "angles"). We want to average in order to replace $\sigma^2(\vartheta^{i,\varepsilon})$ by $\hat{\sigma}^2$ and $\sigma(\vartheta^{i,\varepsilon})\sigma(\vartheta^{j,\varepsilon})$ by $\overline{\sigma}^2$. If $Z^{i,\varepsilon} \approx z_i$ over a "mesoscopic" time scale, then $\vartheta^{i,\varepsilon}$ roughly evolves like (a speeded up version of) $t \mapsto \theta + \omega(z_i)t$. Since

$$\lim_{T \nearrow \infty} \frac{1}{T} \int_{s=0}^T \sigma^2(\theta + \omega(x_i)s) \mathrm{d}s = \hat{\sigma}^2 \tag{8}$$

we should be able to average the coefficients of the $\frac{\partial^2 f}{\partial z_i^2}$'s. To average the coefficients of the $\frac{\partial^2 f}{\partial z_i z_j}$'s, we move to the torus and observe that for $i \neq j$, the angular coordinates $(\vartheta^{i,\varepsilon}, \vartheta^{j,\varepsilon})$ roughly evolve like (again, a speeded up version of) $t \mapsto (\theta_i + \omega(z_i)t, \theta_j + \omega(z_j)t)$. If $\omega(z_i)$ and $\omega(z_j)$ are not rationally related, i.e.,

$$\{(j, k) \in \mathbb{Z}^2 : j\omega(z_i) + k\omega(z_j) = 0\} = \{(0, 0)\}$$

then

$$\lim_{T \nearrow \infty} \frac{1}{T} \int_{s=0}^T \sigma(\theta_i + \omega(z_i)s) \sigma(\theta_j + \omega(z_j)s) \mathrm{d}s = \overline{\sigma}^2. \tag{9}$$

At this point, the technical challenge of our work becomes apparent: the problem of *resonances*. Resonances are trivial from a probabilistic standpoint, but terrible from a deterministic standpoint. For each nonzero $(j, k) \in \mathbb{Z}^2$, define

$$\mathcal{R}_{k,l} \stackrel{\text{def}}{=} \{(x, y) \in \mathbb{R}^2 : k\omega(x) + l\omega(y) = 0\}.$$

Then each $\mathcal{R}_{k,l}$ is a set of Lebesgue measure zero in the plane (recall that $\dot{\omega} > 0$ on all of $\mathbb{R}$), so

$$\mathcal{R} \stackrel{\text{def}}{=} \bigcup_{\substack{(k,l)\in\mathbb{Z}^2 \\ (k,l)\neq(0,0)}} \mathcal{R}_{k,l}$$

is also a set of measure zero in the plane. If $(\Phi_t^\varepsilon(z_i), \Phi_t^\varepsilon(z_j))$ were to have a sufficiently regular density (as $\varepsilon \searrow 0$) with respect to two-dimensional Lebesgue measure on the plane, then we should be able to ignore $\mathcal{R}$ (some similar calculations appear in [6]). On the other hand, if there were no noise, then one would have to show that the process does not "stick" at the $\mathcal{R}_{k,l}$'s (i.e., one must preclude "capture into resonance"). Often this is a very complicated calculation involving problems of "small divisors".

Our problem is between the two extremes of full noise and no noise. From a probabilistic perspective the noise is very degenerate. The simplest incarnation of the problem is in the case of $d = 2$ (and indeed this case is definitive); we then want to replace functions of the angle $(\vartheta^{1,\varepsilon}, \vartheta^{2,\varepsilon})$ by effective quantities. It is easy to see that the Lie algebra spanned by $\mathbf{V}_0 \otimes \mathbf{V}_0$ and $\mathbf{V}_1 \otimes \mathbf{V}_1$ is at most two-dimensional; thus the four-dimensional diffusion $\{(\vartheta_t^{1,\varepsilon}, Z_t^{1,\varepsilon}, \vartheta_t^{2,\varepsilon}, Z_t^{2,\varepsilon}); t \geq 0\}$ does not satisfy Hörmander's requirement, and we cannot expect to find a density for the slow variables $(Z_t^{1,\varepsilon}, Z_t^{2,\varepsilon})$ by taking marginals of a four-dimensional density. Moreover, if by luck we could use abstract machinery to show that $(Z_t^{1,\varepsilon}, Z_t^{2,\varepsilon})$ would have a density, it would be ε-dependent and we would need to show that as $\varepsilon \searrow 0$ it would still be regular enough that $(Z_t^{1,\varepsilon}, Z_t^{2,\varepsilon})$ would be unlikely to be near the $\mathcal{R}_{k,l}$'s. Nevertheless, the noise should in *some* way make things simpler than in the deterministic case. In fact, the calculations of [9] show that it is in fact unlikely to be captured into resonance before leaving K; the tangent flow of $\{\phi_t^\varepsilon; t \geq 0\}$ grows in certain directions as $\varepsilon \searrow 0$ (see also [4] and [3]). Thus, even if z_1 and z_2 are close (but do not coincide), $\phi_t^\varepsilon(p_1)$ and $\phi_t^\varepsilon(p_2)$ are repelled, and in fact $Z_t^{1,\varepsilon}$ and $Z_t^{2,\varepsilon}$ are repelled, and hence should move away from at least $\mathcal{R}_{1,-1}$.

Our work thus represents to some extent a contribution to the theory of escape from resonance. In our problem, there is exactly enough noise in the right direction (transversal to the $\mathcal{R}_{j,k}$'s) that we can neglect the resonances. Furthermore, this noise does not vanish too quickly as $\varepsilon \searrow 0$. The work of [9] will help us to formalize this (see Section 4.3).

Notation 1.3. *We will use two cutoff functions. Let* $\varphi_0 \in C^\infty(\mathbb{R}; [0,1])$ *have support in* $[-2, 2]$ *and be such that* $\chi_{[-1,1]} \leq \varphi_0$ *(in other words* φ_0 *is zero except in a small neighborhood of the origin). Similarly, let* $\varphi_1 \in C^\infty(\mathbb{R}; [0,1])$ *be such that* $0 \notin \operatorname{supp} \varphi_1$ *and* $\varphi_1 \geq \chi_{\mathbb{R}\setminus[-1,1]}$ *(i.e.,* φ_1 *is 1 except in a small neighborhood of the origin). Define*

$$\varphi_0^{(-2)}(z) \stackrel{\text{def}}{=} \int_{r=0}^{z} \left\{ \int_{s=0}^{r} \varphi_0(s) \mathrm{d}s \right\} \mathrm{d}r \qquad z \in \mathbb{R}.$$

Note that there is a $\mathsf{K} > 0$ *such that* $|\dot{\varphi}_0^{(-2)}(z)| \leq \mathsf{K}$ *and* $|\varphi_0^{(-2)}(z)| \leq \mathsf{K}|z|$ *for all* $z \in \mathbb{R}$.

For each $j \in \mathbb{Z}$, define

$$\sigma_j \stackrel{\text{def}}{=} \int_{\theta=0}^{1} \exp\left[-2\pi \mathrm{i} j\theta\right] \sigma(\theta) \mathrm{d}\theta$$

so that

$$\sigma(\theta) = \sum_{j \in \mathbb{Z}} \sigma_j \exp[2\pi \mathrm{i} j\theta]$$

for all $\theta \in \mathbb{R}$, and this sum converges pointwise.

Throughout, we let K be a generic constant, which may change from incarnation to incarnation, and which depends only on σ and ω.

2. Nondegeneracy at resonance and proof of main result

Here we organize our calculations and identify the escape from resonance phenomenon which underlies our averaging. This will rely on several calculations which simplify our interests.

Prior to (7), we fixed p_i's and defined $Z^{i,\varepsilon}$ and $\vartheta^{i,\varepsilon}$'s. Define $\bar{Z}^\varepsilon_t \stackrel{\text{def}}{=} (Z^{1,\varepsilon}_t, Z^{1,\varepsilon}_t \dots Z^{d,\varepsilon}_t)$. Also fix i and j in $\{1, 2 \dots n\}$, which we shall use throughout the rest of the paper. The first observation which we wish to exploit is that it is in a sense sufficient to consider the *two-point* motion. Namely, note that our goal is to use approximations like (8) and (9) in (7) to replace the $\sigma(\vartheta^{i,\varepsilon})$'s with constant coefficients. We should be able to separately do this for each i and j in the two sums in (7); since each such term involves at most two angles, the two-point motion should be sufficient. The $Z^{i,\varepsilon}$'s are slow variables, so they should effectively be held constant while we carry out the averaging. The details of these arguments will be in the proof below of Theorem 1.2. Set $\theta^\varepsilon \stackrel{\text{def}}{=} \vartheta^{i,\varepsilon}$, $X^\varepsilon \stackrel{\text{def}}{=} Z^{i,\varepsilon}$, $\psi^\varepsilon \stackrel{\text{def}}{=} \vartheta^{j,\varepsilon}$, and $Y^\varepsilon \stackrel{\text{def}}{=} Z^{j,\varepsilon}$. Set $\theta_\circ \stackrel{\text{def}}{=} \theta^\varepsilon_0$, $x_\circ \stackrel{\text{def}}{=} X^\varepsilon_0$, $\psi_\circ \stackrel{\text{def}}{=} \psi^\varepsilon_0$, and $y_\circ \stackrel{\text{def}}{=} Y^\varepsilon_0$. Then $(\theta^\varepsilon, X^\varepsilon)$ satisfies (1) and $(\psi^\varepsilon, Y^\varepsilon)$ satisfies

$$\begin{aligned} \mathrm{d}\psi^\varepsilon_t &= \frac{1}{\varepsilon}\omega(Y^\varepsilon_t)\mathrm{d}t \\ \mathrm{d}Y^\varepsilon_t &= \sigma(\psi^\varepsilon_t)\mathrm{d}W_t \\ (\psi^\varepsilon_0, Y^\varepsilon_0) &= (\psi_\circ, y_\circ). \end{aligned}$$

Define $\mathsf{C}^\varepsilon_t \stackrel{\text{def}}{=} (X^\varepsilon_t, \theta^\varepsilon_t, Y^\varepsilon_t, \psi^\varepsilon_t)$ for all $\varepsilon \in (0, 1)$ and $t \geq 0$.

Our second simplification follows from a Fourier decomposition, which allows us to efficiently exploit periodicity in the angular variables θ^ε and ψ^ε. Define $C^\infty(\mathbb{T}^2)$ as the collection of $\varphi \in C^\infty(\mathbb{R}^2)$ such that $\varphi(\theta + j, \psi + k) = \varphi(\theta, \psi)$ for all $(\theta, \psi) \in \mathbb{R}^2$ and $(k, l) \in \mathbb{Z}^2$. For $\varphi \in C^\infty(\mathbb{T}^2)$ and $(j, k) \in \mathbb{Z}^2$, define

$$\hat{\varphi}_{j,k} \stackrel{\text{def}}{=} \int_{\theta=0}^1 \int_{\psi=0}^1 \varphi(\theta, \psi) \exp[-2\pi \mathrm{i}(j\theta + k\psi)]\, \mathrm{d}\theta \mathrm{d}\psi;$$

then

$$\lim_{N\to\infty} \sup_{(\theta,\psi)\in\mathbb{R}^2} \left| \varphi(\theta, \psi) - \sum_{\substack{(j,k)\in\mathbb{Z}^2 \\ |j|+|k|\leq N}} \hat{\varphi}_{j,k} \exp[2\pi \mathrm{i}(j\theta + k\psi)] \right| = 0.$$

Lemma 2.1. *Fix* $f^* \in C^2_b(\mathbb{R}^d)$, $0 \leq s_1 \leq s_2 \dots s_n \leq s \leq t$, $\{g_{n'}\}_{n'=1}^n \subset C_b(\mathbb{R}^d)$, *and* $(k, l) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$. *Then*

$$\lim_{\varepsilon \searrow 0} \mathbb{E}\left[\left\{ \int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K} f^*(\bar{Z}^\varepsilon_r) \exp\left[2\pi \mathrm{i}(j\theta^\varepsilon_r + k\psi^\varepsilon_r)\right] \mathrm{d}r \right\} \prod_{j=1}^n g_{n'}(\bar{Z}^\varepsilon_{s_{n'}}) \right] = 0.$$

The proof of our main result follows from this.

Proof of Theorem 1.2. According to the standard theory of convergence of Markov processes, we need to prove three things: tightness, convergence to the limiting martingale characterization, and uniqueness of the solution of the limiting martingale characterization. Tightness of $\{(\Phi^\varepsilon_t(z_1), \Phi^\varepsilon(z_2)\dots\Phi^\varepsilon(z_n));\ t \ge 0\}$ follows from tightness of the one-point motions, which is implied by the convergence result of Theorem 1.1. Uniqueness of the limiting martingale characterization is standard. To show the desired convergence to the limiting martingale problem, we argue as follows. Let f^*, s_i's, s, and t be as in Lemma 2.1. Fix also $\varphi \in C^\infty(\mathbb{T}^2)$ and define $\bar\varphi \stackrel{\text{def}}{=} \int_{[0,1]^2}\varphi(x,y)\mathrm{d}x\mathrm{d}y$. By approximation by a finite Fourier series, we have that

$$\lim_{\varepsilon\searrow 0}\mathbb{E}\left[\left\{\int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K} f^*(\bar Z^\varepsilon_r)\{\varphi(\theta^\varepsilon_r,\psi^\varepsilon_r)-\bar\varphi\}\mathrm{d}r\right\}\prod_{n'=1}^n g_{n'}(\bar Z^\varepsilon_{s_{n'}})\right]=0.$$

Since i and j were arbitrarily chosen at the beginning of the section, we have that

$$\lim_{\varepsilon\searrow 0}\mathbb{E}\left[\left\{\int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K} f^*(\bar Z^\varepsilon_r)\{\varphi(\vartheta^{i,\varepsilon}_r,\vartheta^{j,\varepsilon}_r)-\bar\varphi\}\mathrm{d}r\right\}\prod_{n'=1}^n g_{n'}(\bar Z^\varepsilon_{s_{n'}})\right]=0$$

for all i and j. Letting f^* be of the form $\partial^2 f/\partial z_i^2$ or $\partial^2 f/\partial z_i\partial z_j$ and letting φ be of the form $(\theta,\psi)\mapsto\sigma(\theta)\sigma(\psi)$ or $(\theta,\psi)\mapsto\sigma^2(\theta)$ as indicated by (7), we have that

$$\begin{aligned}\lim_{\varepsilon\searrow 0}\mathbb{E}\Bigg[\Bigg\{\int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K}\Bigg\{&\sum_{1\le i\le d}\frac{1}{2}\sigma^2(\vartheta^{i,\varepsilon}_r)\frac{\partial^2 f}{\partial z_i^2}(Z^{1,\varepsilon}_r, Z^{2,\varepsilon}_r\dots Z^{i,\varepsilon}_r)\\ &+\sum_{\substack{1\le i,j\le d\\ i\ne j}}\sigma(\vartheta^{i,\varepsilon}_r)\sigma(\vartheta^{j,\varepsilon}_r)\frac{\partial^2 f}{\partial z_i\partial z_j}(Z^{1,\varepsilon}_r, Z^{2,\varepsilon}_r\dots Z^{i,\varepsilon}_r)-(\mathscr{L}_d f)(\bar Z^\varepsilon_r)\Bigg\}\mathrm{d}r\Bigg\}\\ &\times\prod_{n'=1}^n g_{n'}(\bar Z^\varepsilon_{s_{n'}})\Bigg]=0.\end{aligned}$$

We finally use the fact that (7) is a martingale to see that in fact

$$\lim_{\varepsilon\searrow 0}\mathbb{E}\left[\left\{f(\bar Z^\varepsilon_{t\wedge\tau^\varepsilon_K})-f(\bar Z^\varepsilon_{s\wedge\tau^\varepsilon_K})-\int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K}(\mathscr{L}_d f)(\bar Z^\varepsilon_r)\mathrm{d}r\right\}\prod_{n'=1}^n g_{n'}(\bar Z^\varepsilon_{s_{n'}})\right]=0,$$

which gives us the desired characterization of the limiting law of the $\bar Z^\varepsilon$'s. □

The task before us is now to prove Lemma 2.1. Fix a nonzero $(k,l)\in\mathbb{Z}^2$ and define

$$\Gamma^\varepsilon_t \stackrel{\text{def}}{=} k\theta^\varepsilon_t + l\psi^\varepsilon_t \quad\text{and}\quad \mathsf{a}^\varepsilon_t \stackrel{\text{def}}{=} k\omega(X^\varepsilon_t)+l\omega(Y^\varepsilon_t).$$

We note that Γ^ε evolves according to $\mathrm{d}\Gamma^\varepsilon_t/\mathrm{d}t=\varepsilon^{-1}\mathsf{a}^\varepsilon_t$. Thus we have a separation of scales in the quantity

$$f(\bar Z^\varepsilon_t)\exp\left[2\pi\mathrm{i}\Gamma^\varepsilon_t\right];$$

Γ^ε moves much faster than Z^ε (and in particular faster than X^ε and Y^ε). Consequently, we should be able to average. To do so, define

$$\Phi(\theta)\stackrel{\text{def}}{=}\frac{\exp[2\pi\mathrm{i}\theta]-1}{2\pi\mathrm{i}}$$

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

$$U_t^\varepsilon \stackrel{\text{def}}{=} \varepsilon f^*(\bar{Z}_t^\varepsilon) \frac{\Phi(\Gamma_t^\varepsilon)}{\mathsf{a}_t^\varepsilon} \chi_{\mathbb{R}^2 \setminus \mathcal{R}_{j,k}}(X_t^\varepsilon, Y_t^\varepsilon)$$

for all $t > 0$. We should then have that $\mathrm{d}U_t^\varepsilon \approx f^*(\bar{Z}_t^\varepsilon) \exp\left[2\pi \mathrm{i} \Gamma_t^\varepsilon\right] \mathrm{d}t$.

To start to make this precise, define

$$F_t^{1,\varepsilon} \stackrel{\text{def}}{=} \frac{1}{2} \sum_{1 \le i,j \le d} \frac{\partial^2 f^*}{\partial z_i \partial z_j}(\bar{Z}_t^\varepsilon) \sigma(\vartheta_t^{i,\varepsilon}) \sigma(\vartheta_t^{j,\varepsilon})$$

$$F_t^{2,\varepsilon} \stackrel{\text{def}}{=} \sum_{1 \le i \le d} \frac{\partial f^*}{\partial z_i}(\bar{Z}_t^\varepsilon) \sigma(\vartheta_t^{i,\varepsilon})$$

$$A_t^{1,\varepsilon} \stackrel{\text{def}}{=} \frac{1}{2}\{k \ddot{\omega}(X_t^\varepsilon) \sigma^2(\theta_t^\varepsilon) + l \ddot{\omega}(Y_t^\varepsilon) \sigma^2(\psi_t^\varepsilon)\}$$

$$A_t^{2,\varepsilon} \stackrel{\text{def}}{=} k \dot{\omega}(X_t^\varepsilon) \sigma(\theta_t^\varepsilon) + l \dot{\omega}(Y_t^\varepsilon) \sigma(\psi_t^\varepsilon)$$

for all $t \ge 0$; then

$$\mathrm{d}\mathsf{a}_t^\varepsilon = A_t^{1,\varepsilon} \mathrm{d}t + A_t^{2,\varepsilon} \mathrm{d}W_t \quad \text{and} \quad \mathrm{d}f^*(\bar{Z}_t^\varepsilon) = F_t^{1,\varepsilon} \mathrm{d}t + F_t^{2,\varepsilon} \mathrm{d}W_t.$$

Applying Ito's formula to U^ε, we then have that on $\left\{\inf_{r \in [0,t]} |\mathsf{a}_r^\varepsilon| > 0\right\}$,

$$U_t^\varepsilon - U_0^\varepsilon = \int_{r=0}^t f^*(\bar{Z}_t^\varepsilon) \exp\left[2\pi \mathrm{i} \Gamma_r^\varepsilon\right] \mathrm{d}r + \int_{r=0}^t \mathscr{E}_{1,\varepsilon}(r) \Phi(\Gamma_r^\varepsilon) \mathrm{d}r + \int_{r=0}^t \mathscr{E}_{2,\varepsilon}(r) \Phi(\Gamma_r^\varepsilon) \mathrm{d}W_r$$

where, for all $r \ge 0$,

$$\mathscr{E}_{1,\varepsilon}(r) \stackrel{\text{def}}{=} \varepsilon \left\{ \frac{F_r^{1,\varepsilon}}{\mathsf{a}_r^\varepsilon} - \frac{f^*(\bar{Z}_r^\varepsilon) A_r^{1,\varepsilon}}{(\mathsf{a}_r^\varepsilon)^2} + \frac{f(\bar{Z}_r^\varepsilon)(A_r^{2,\varepsilon})^2}{(\mathsf{a}_r^\varepsilon)^3} - \frac{F_r^{2,\varepsilon} A_r^{2,\varepsilon}}{(\mathsf{a}_r^\varepsilon)^2} \right\}$$

$$\mathscr{E}_{2,\varepsilon}(r) \stackrel{\text{def}}{=} \varepsilon \left\{ \frac{F_r^{2,\varepsilon}}{\mathsf{a}_r^\varepsilon} - \frac{f^*(\bar{Z}_r^\varepsilon) A_r^{2,\varepsilon}}{(\mathsf{a}_r^\varepsilon)^2} \right\}.$$

Note that $|F_t^{i,\varepsilon}| \le \mathsf{K}$ and $|A_t^{i,\varepsilon}| \le \mathsf{K}$ for $i \in \{1,2\}$, $\varepsilon \in (0,1)$, and $t \ge 0$; thus the only source of the singularity is the a_t^ε in the denominator in U^ε and I^ε. Note also that in fact nothing in our discussion actually precludes *starting* in $\mathcal{R}_{j,k}$.

Noting that the highest power of a_t^ε in the denominator is 3, the following lemma thus is a natural goal.

Lemma 2.2 (*Stochastic Nondegeneracy at Resonance*). *Fix* $\nu \in (0, 1/3)$. *For* $t > 0$,

$$\overline{\lim_{\varepsilon \searrow 0}} \mathbb{E}\left[\int_{r=0}^{t \wedge \tau_K^\varepsilon} \chi_{\{|\mathsf{a}_r^\varepsilon| \le \varepsilon^\nu\}} \mathrm{d}r\right] = 0.$$

It turns out that this is exactly what is needed to prove Lemma 2.1.

Proof of Lemma 2.1. To rigorously use Ito's formula, define

$$\tilde{U}^\varepsilon_t \stackrel{\text{def}}{=} \varphi_1\left(\frac{\mathsf{a}^\varepsilon_t}{\varepsilon^\nu}\right) U^\varepsilon_t.$$

By a careful application of Ito's formula, we get that

$$\begin{aligned}\int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K} \varphi_1\left(\frac{\mathsf{a}^\varepsilon_t}{\varepsilon^\nu}\right) f^*(\bar{Z}^\varepsilon_r)\exp\left[2\pi \mathrm{i}\Gamma^\varepsilon_r\right]\mathrm{d}r &= \tilde{U}^\varepsilon_{t\wedge\tau^\varepsilon_K} - \tilde{U}^\varepsilon_{s\wedge\tau^\varepsilon_K} \\ &\quad + \int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K} \bar{\mathscr{E}}_{1,\varepsilon}(r)\exp\left[2\pi\mathrm{i}\Gamma^\varepsilon_r\right]\mathrm{d}r - \sum_{j=2}^4 \int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K} \bar{\mathscr{E}}_{j,\varepsilon}(r)\Phi(\Gamma^\varepsilon_r)\mathrm{d}r \\ &\quad - \sum_{j=5}^6 \int_{r=s\wedge\tau^\varepsilon_K}^{t\wedge\tau^\varepsilon_K} \bar{\mathscr{E}}_{j,\varepsilon}(r)\Phi(\Gamma^\varepsilon_r)\mathrm{d}W_r\end{aligned}$$

where

$$\begin{aligned}&\bar{\mathscr{E}}_{1,\varepsilon}(r) \stackrel{\text{def}}{=} \left\{1-\varphi_1\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)\right\} f^*(\bar{Z}^\varepsilon_r), \qquad \bar{\mathscr{E}}_{2,\varepsilon}(r) \stackrel{\text{def}}{=} \varphi_1\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)\mathscr{E}_{1,\varepsilon}(r) \\ &\bar{\mathscr{E}}_{3,\varepsilon}(r) \stackrel{\text{def}}{=} \dot{\varphi}_1\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)\frac{A^{1,\varepsilon}_r U^\varepsilon_r}{\varepsilon^\nu}, \qquad \bar{\mathscr{E}}_{4,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{1}{2}\ddot{\varphi}_1\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)\frac{(A^{2,\varepsilon}_r)^2 U^\varepsilon_r}{\varepsilon^{2\nu}} \\ &\bar{\mathscr{E}}_{5,\varepsilon}(r) \stackrel{\text{def}}{=} \dot{\varphi}_1\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)\frac{A^{2,\varepsilon}_r \mathscr{E}_{2,\varepsilon}(r)}{\varepsilon^\nu}, \qquad \bar{\mathscr{E}}_{6,\varepsilon}(r) \stackrel{\text{def}}{=} \varphi_1\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)\mathscr{E}_{2,\varepsilon}(r), \\ &\bar{\mathscr{E}}_{7,\varepsilon}(r) \stackrel{\text{def}}{=} \dot{\varphi}_1\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)\frac{U^\varepsilon_r A^{2,\varepsilon}_r}{\varepsilon^\nu}.\end{aligned}$$

Since $\varphi_1 \ge \chi_{\mathbb{R}\setminus[-1,1]}$, $1-\varphi_1 \le \chi_{(-1,1)}$, so

$$|\bar{\mathscr{E}}_{1,\varepsilon}(r)| \le \mathsf{K}\chi_{\{|\mathsf{a}^\varepsilon_r|\le\varepsilon^\nu\}}.$$

Next, since $0 \notin \operatorname{supp}\varphi_1$, there is a $\varrho > 0$ such that $(-\varrho,\varrho)\cap\operatorname{supp}\varphi_1 = \emptyset$. Thus

$$|\tilde{U}^\varepsilon_t| \le \mathsf{K}\frac{\varepsilon}{|\mathsf{a}^\varepsilon_t|}\chi_{\{|\mathsf{a}^\varepsilon_r|\ge\varrho\varepsilon^\nu\}}$$

and

$$\begin{aligned}&|\bar{\mathscr{E}}_{2,\varepsilon}(r)| \le \mathsf{K}\frac{\varepsilon}{|\mathsf{a}^\varepsilon_r|^3}\chi_{\{|\mathsf{a}^\varepsilon_r|\ge\varrho\varepsilon^\nu\}}, \qquad |\bar{\mathscr{E}}_{3,\varepsilon}(r)| \le \mathsf{K}\frac{\varepsilon^{1-\nu}}{|\mathsf{a}^\varepsilon_r|}\chi_{\{|\mathsf{a}^\varepsilon_r|\ge\varrho\varepsilon^\nu\}}, \\ &|\bar{\mathscr{E}}_{4,\varepsilon}(r)| \le \mathsf{K}\frac{\varepsilon^{1-2\nu}}{|\mathsf{a}^\varepsilon_r|}\chi_{\{|\mathsf{a}^\varepsilon_r|\ge\varrho\varepsilon^\nu\}}, \qquad |\bar{\mathscr{E}}_{5,\varepsilon}(r)| \le \mathsf{K}\frac{\varepsilon^{1-\nu}}{|\mathsf{a}^\varepsilon_r|^2}\chi_{\{|\mathsf{a}^\varepsilon_r|\ge\varrho\varepsilon^\nu\}}, \\ &|\bar{\mathscr{E}}_{6,\varepsilon}(r)| \le \mathsf{K}\frac{\varepsilon}{|\mathsf{a}^\varepsilon_r|^2}\chi_{\{|\mathsf{a}^\varepsilon_r|\ge\varrho\varepsilon^\nu\}}, \qquad |\bar{\mathscr{E}}_{7,\varepsilon}(r)| \le \mathsf{K}\frac{\varepsilon^{1-\nu}}{|\mathsf{a}^\varepsilon_r|}\chi_{\{|\mathsf{a}^\varepsilon_r|\ge\varrho\varepsilon^\nu\}}.\end{aligned}$$

Combine things together and use standard calculations to get the desired result. We here use that $1-3\nu > 0$. □

3. Averaging for $k + l \neq 0$

We here prove Lemma 2.2 for $k + l \neq 0$. For $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$, define first

$$\hbar_a(\mathsf{c}) \stackrel{\text{def}}{=} (k\dot\omega(x)\sigma(\theta) + l\dot\omega(y)\sigma(\psi))^2 = (k\dot\omega(x))^2\sigma^2(\theta) + (l\dot\omega(y))^2\sigma^2(\psi)^2 + 2(k\dot\omega(x))(l\dot\omega(y))\sigma(\theta)\sigma(\psi);$$

then $\hbar_a(\mathsf{C}^\varepsilon_t) = (A^{1,\varepsilon}_t)^2$. For $\lambda > 0$, we define a corrector which will allow us to coarse-grain. Set

$$H^\lambda_a(\mathsf{c}) \stackrel{\text{def}}{=} \int_{r=0}^\infty e^{-\lambda r}\, \hbar_a(x, \theta + \omega(x)r, y, \psi + \omega(y)r)\mathrm{d}r$$

for all $\lambda \in (0, 1)$ and $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$.

Lemma 3.1. *We have that $H^\lambda_a \in C^\infty(\mathbb{R}^4)$. Secondly, $\sup_{\lambda\in(0,1)} \lambda\|H^\lambda_a\|_{C^2(\mathbb{R}^4)} < \infty$. Thirdly,*

$$\omega(x)\frac{\partial H^\lambda_a}{\partial\theta}(\mathsf{c}) + \omega(y)\frac{\partial H^\lambda_a}{\partial\psi}(\mathsf{c}) = \lambda H^\lambda_a(\mathsf{c}) - \hbar_a(\mathsf{c})$$

for all $\lambda \in (0, 1)$ and $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$, and finally there is a $\mathsf{K} > 0$ such that

$$\lambda H^\lambda_a(\mathsf{c}) \geq \frac{1}{\mathsf{K}} - \mathsf{K}\lambda - \mathsf{K}|k\omega(x) + l\omega(y)|$$

for all $\lambda \in (0, 1)$ and $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$.

Proof. The regularity and bound on derivatives are easy; one can differentiate under the integral. The formula for the derivative is fairly easy to see; see [10].

Fix $\lambda \in (0, 1)$ and $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$. We can explicitly compute that

$$\begin{aligned}\lambda H^\lambda_a(\mathsf{c}) &= (k\dot\omega(x))^2 \sum_{j,j'\in\mathbb{Z}} \sigma_j\sigma_{j'} \exp\left[2\pi \mathrm{i}(j + j')\theta\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j + j')\omega(x)}\\ &\quad + (l\dot\omega(y))^2 \sum_{j,j'\in\mathbb{Z}} \sigma_j\sigma_{j'} \exp\left[2\pi \mathrm{i}(j + j')\psi\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j + j')\omega(y)}\\ &\quad + 2(k\dot\omega(x))(l\dot\omega(y)) \sum_{j,j'\in\mathbb{Z}} \sigma_j\sigma_{j'} \exp\left[2\pi \mathrm{i}(j\theta + j'\psi)\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j\omega(x) + j'\omega(y))}.\end{aligned}$$

Note that if $j\omega(x) + j'\omega(y) = 0$ and $k\omega(x) + l\omega(y) = 0$, then[2] $jl = j'k$. Let M be the greatest common factor of $|k|$ and $|l|$ and set $k^* \stackrel{\text{def}}{=} k/M$ and $l^* \stackrel{\text{def}}{=} l/M$. Thus k^* and l^* are relatively prime and $jl^* = j'k^*$. Hence $j = mk^*$ and $j' = ml^*$ for some $m \in \mathbb{Z}$; for simplicity, we denote this for future reference as $(j, j') \in (k^*, l^*)\mathbb{Z}$. We hence rewrite λH^λ_a as

$$\lambda H^\lambda_a(\mathsf{c}) = \bar I + (k\dot\omega(x))^2 I_1 + (l\dot\omega(y))^2 I_2 + (k\dot\omega(x))(l\dot\omega(y)) I_3$$

[2] Note that since $\omega > 0$, j, j', k and l must all be nonzero.

where

$$\bar I \stackrel{\text{def}}{=} (k\dot\omega(x))^2 \sum_{j\in\mathbb{Z}} |\sigma_j|^2 + (l\dot\omega(y))^2 \sum_{j\in\mathbb{Z}} |\sigma_j|^2 + 2(k\dot\omega(x))(l\dot\omega(y)) \times \sum_{m\in\mathbb{Z}} \sigma_{mk^*}\sigma_{ml^*} \exp\left[2\pi \mathrm{i} m(k^*\theta + l^*\psi)\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i} m(k^*\omega(x)+l^*\omega(y))}$$

$$I_1 \stackrel{\text{def}}{=} \sum_{\substack{j,j'\in\mathbb{Z}\\ j+j'\neq 0}} \sigma_j\sigma_{j'} \exp\left[2\pi\mathrm{i}(j+j')\theta\right] \frac{\lambda}{\lambda - 2\pi\mathrm{i}(j+j')\omega(x)}$$

$$I_2 \stackrel{\text{def}}{=} \sum_{\substack{j,j'\in\mathbb{Z}\\ j+j'\neq 0}} \sigma_j\sigma_{j'} \exp\left[2\pi\mathrm{i}(j+j')\psi\right] \frac{\lambda}{\lambda - 2\pi\mathrm{i}(j+j')\omega(y)}$$

$$I_3 \stackrel{\text{def}}{=} \sum_{\substack{(j,j')\in\mathbb{Z}^2\setminus(k^*,l^*)\mathbb{Z}\\ |j\omega(x)+j'\omega(y)|\geq\sqrt{\lambda}}} \sigma_j\sigma_{j'} \exp\left[2\pi\mathrm{i}(j\theta + j'\psi)\right] \frac{\lambda}{\lambda - 2\pi\mathrm{i}(j\omega(x)+j'\omega(y))}$$

$$I_4 \stackrel{\text{def}}{=} \sum_{\substack{(j,j')\in\mathbb{Z}^2\setminus(k^*,l^*)\mathbb{Z}\\ |j\omega(x)+j'\omega(y)|<\sqrt{\lambda}}} \sigma_j\sigma_{j'} \exp\left[2\pi\mathrm{i}(j\theta + j'\psi)\right] \frac{\lambda}{\lambda - 2\pi\mathrm{i}(j\omega(x)+j'\omega(y))}.$$

If $j+j'\neq 0$, then $|j+j'|\geq 1$. Thus

$$|I_1| \leq \left(\sum_{j\in\mathbb{Z}}|\sigma_j|\right)^2 \frac{\lambda}{\omega_-}, \qquad |I_2| \leq \left(\sum_{j\in\mathbb{Z}}|\sigma_j|\right)^2 \frac{\lambda}{\omega_-}, \qquad |I_3| \leq \left(\sum_{j\in\mathbb{Z}}|\sigma_j|\right)^2 \sqrt{\lambda}.$$

Next, note that

$$\left|\frac{\lambda}{\lambda - 2\pi\mathrm{i}(j+j')\omega(x)}\right| \leq 1$$

and thus, by Young's inequality,

$$\bar I \geq (k\dot\omega(x))^2 \sum_{j\in\mathbb{Z}\setminus k^*\mathbb{Z}} |\sigma_j|^2 + (l\dot\omega(y))^2 \sum_{j\in\mathbb{Z}\setminus l^*\mathbb{Z}} |\sigma_j|^2$$

and since the period of σ is exactly 1, the right-hand side must be positive.

We finally observe that if $|j\omega(x)+j'\omega(y)|<\sqrt{\lambda}$ and $(j,j')\in\mathbb{Z}^2\setminus(k^*,l^*)\mathbb{Z}$, then

$$\begin{aligned}\left(|j|+|j'|\right)|k^*\omega(x)+l^*\omega(y)| &\geq \left|j(k^*\omega(x)+l^*\omega(y)) - k^*(j\omega(x)+j'\omega(y))\right| \\ &\quad + \left|j'(k^*\omega(x)+l^*\omega(y)) - l^*(j\omega(x)+j'\omega(y))\right| - 2(|k^*|+|l^*|)\sqrt{\lambda} \\ &\geq \left|jl^* - j'k^*\right|(\omega(y)+\omega(x)) - 2(|k^*|+|l^*|)\sqrt{\lambda} \geq 2\omega_- - 2(|k^*|+|l^*|)\sqrt{\lambda}.\end{aligned}$$

Thus if

$$\lambda < \frac{\omega^2}{4(|k^*|+|l^*|)^2}, \tag{10}$$

then $\left(|j|+|j'|\right)|k^*\omega(x)+l^*\omega(y)|\geq\omega_-$, so either

$$|j||k^*\omega(x)+l^*\omega(y)|\geq\omega_-/2 \quad \text{or} \quad |j'||k^*\omega(x)+l^*\omega(y)|\geq\omega_-/2.$$

Hence

$$|I_4| \le 2\left(\sum_{j'\in\mathbb{Z}}|\sigma_{j'}|\right)\left(\sum_{\substack{j\in\mathbb{Z}\\ |j||k^*\omega(x)+l^*\omega(y)|\ge\omega_-/2}}|\sigma_j|\right)$$
$$\le 2|k^*\omega(x)+l^*\omega(y)|\left(\sum_{j'\in\mathbb{Z}}|\sigma_{j'}|\right)\left(\sum_{\substack{j\in\mathbb{Z}\\ |j||k^*\omega(x)+l^*\omega(y)|\ge\omega_-/2}}\frac{|j||\sigma_j|}{|j||k^*\omega(x)+l^*\omega(y)|}\right)$$
$$\le \frac{4}{\omega_-}|k^*\omega(x)+l^*\omega(y)|\left(\sum_{j'\in\mathbb{Z}}|\sigma_{j'}|\right)\left(\sum_{j\in\mathbb{Z}}|j\sigma_j|\right).$$

Note that

$$|k^*\omega(x)+l^*\omega(y)| = \frac{|k\omega(x)+l\omega(y)|}{M} \le |k\omega(x)+l\omega(y)|.$$

This completes the proof of the bound on λH^λ_a if (10) holds. The bound on λH^λ_a is trivial if (10) does not hold. □

We now have

Proof of Lemma 2.2 for $k+l\neq 0$. Set $\lambda_\varepsilon \overset{\text{def}}{=} \varepsilon^\nu$ and define

$$U^\varepsilon_t \overset{\text{def}}{=} 2\varepsilon^{2\nu}\varphi_0^{(-2)}\left(\frac{\mathsf{a}^\varepsilon_t}{\varepsilon^\nu}\right) + \varepsilon H^{\lambda_\varepsilon}_a(\mathsf{C}^\varepsilon_t)\varphi_0\left(\frac{\mathsf{a}^\varepsilon_t}{\varepsilon^\nu}\right).$$

Then

$$\int_{r=0}^{t\wedge\tau^\varepsilon_K}\lambda_\varepsilon H^{\lambda_\varepsilon}_a(\mathsf{C}^\varepsilon_r)\varphi_0\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)\mathrm{d}r = U^\varepsilon_{t\wedge\tau^\varepsilon_K} - U^\varepsilon_0 - \sum_{j=1}^{5}\int_{r=0}^{t\wedge\tau^\varepsilon_K}\mathscr{E}_{j,\varepsilon}(r)\mathrm{d}r$$
$$- \sum_{j=6}^{8}\int_{r=0}^{t\wedge\tau^\varepsilon_K}\mathscr{E}_{j,\varepsilon}(r)\mathrm{d}W_r$$

where

$$\mathscr{E}_{1,\varepsilon}(r) \overset{\text{def}}{=} \varepsilon^\nu\dot{\varphi}_0^{(-2)}\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)A^{2,\varepsilon}_r$$

$$\mathscr{E}_{2,\varepsilon}(r) \overset{\text{def}}{=} \frac{\varepsilon}{2}\left\{\frac{\partial^2 H^{\lambda_\varepsilon}_a}{\partial x^2}(\mathsf{C}^\varepsilon_r)\sigma^2(\theta^\varepsilon_r) + \frac{\partial^2 H^{\lambda_\varepsilon}_a}{\partial y^2}(\mathsf{C}^\varepsilon_r)\sigma^2(\psi^\varepsilon_r)\right\}\varphi_0\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)$$

$$\mathscr{E}_{3,\varepsilon}(r) \overset{\text{def}}{=} \frac{\varepsilon^{1-\nu}}{2}H^{\lambda_\varepsilon}_a(\mathsf{C}^\varepsilon_r)\dot{\varphi}_0\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)A^{2,\varepsilon}_r$$

$$\mathscr{E}_{4,\varepsilon}(r) \overset{\text{def}}{=} \frac{\varepsilon^{1-2\nu}}{2}H^{\lambda_\varepsilon}_a(\mathsf{C}^\varepsilon_r)\ddot{\varphi}_0\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)(A^{1,\varepsilon}_r)^2$$

$$\mathscr{E}_{5,\varepsilon}(r) \overset{\text{def}}{=} \varepsilon^{1-\nu}\left\{\frac{\partial H^{\lambda_\varepsilon}_a}{\partial x}(\mathsf{C}^\varepsilon_r)\sigma(\theta^\varepsilon_r) + \frac{\partial H^{\lambda_\varepsilon}_a}{\partial y}(\mathsf{C}^\varepsilon_r)\sigma(\psi^\varepsilon_r)\right\}\dot{\varphi}_0\left(\frac{\mathsf{a}^\varepsilon_r}{\varepsilon^\nu}\right)A^{1,\varepsilon}_r$$

$$\mathscr{E}_{6,\varepsilon}(r) \stackrel{\text{def}}{=} 2\varepsilon^{\nu}\dot{\varphi}_0^{(-2)}\left(\frac{\mathsf{a}_t^{\varepsilon}}{\varepsilon^{\nu}}\right)A_r^{1,\varepsilon}$$

$$\mathscr{E}_{7,\varepsilon}(r) \stackrel{\text{def}}{=} \varepsilon\left\{\frac{\partial H_a^{\lambda_\varepsilon}}{\partial x}(\mathsf{C}_r^{\varepsilon})\sigma(\theta_r^{\varepsilon}) + \frac{\partial H_a^{\lambda_\varepsilon}}{\partial y}(\mathsf{C}_r^{\varepsilon})\sigma(\psi_r^{\varepsilon})\right\}\varphi_0\left(\frac{\mathsf{a}_r^{\varepsilon}}{\varepsilon^{\nu}}\right)$$

$$\mathscr{E}_{8,\varepsilon}(r) \stackrel{\text{def}}{=} \varepsilon^{1-\nu}H_a^{\lambda_\varepsilon}(\mathsf{C}_r^{\varepsilon})\dot{\varphi}_0\left(\frac{\mathsf{a}_r^{\varepsilon}}{\varepsilon^{\nu}}\right)A_r^{1,\varepsilon}.$$

We then have that $|\mathscr{E}_{j,\varepsilon}(r)| \leq \mathsf{K}\varepsilon^{\nu}$ for $j \in \{1,6\}$ and $|\mathscr{E}_{j,\varepsilon}(r)| \leq \mathsf{K}\varepsilon^{1-2\nu}/\lambda_\varepsilon = \mathsf{K}\varepsilon^{1-3\nu}$ for all other j. We also note that

$$\lambda_\varepsilon H_a^{\lambda_\varepsilon}(\mathsf{C}_r^{\varepsilon})\varphi_0\left(\frac{\mathsf{a}_r^{\varepsilon}}{\varepsilon^{\nu}}\right) \geq \left\{\frac{1}{\mathsf{K}} - \mathsf{K}\varepsilon^{\nu}\right\}\varphi_0\left(\frac{\mathsf{a}_r^{\varepsilon}}{\varepsilon^{\nu}}\right).$$

The claimed result follows.[3] □

4. Averaging for $k + l = 0$

The next task is to average the 1:1 resonance. This is more challenging as it requires a detailed analysis near the diagonal of $(\mathbb{S}^1 \times \mathbb{R}) \times (\mathbb{S}^1 \times \mathbb{R})$. Define $\mathsf{b}_t^{\varepsilon} \stackrel{\text{def}}{=} X_t^{\varepsilon} - Y_t^{\varepsilon}$. The central component of the proof of Lemma 2.2 for $k = -l$ is then

Lemma 4.1. *For each $\varrho > 0$,*

$$\varlimsup_{\varepsilon \searrow 0}\mathbb{E}\left[\int_{r=0}^{t\wedge\tau_K^{\varepsilon}}\chi_{[-\varrho,\varrho]}\left(\frac{\mathsf{b}_r^{\varepsilon}}{\varepsilon^{\nu}}\right)\mathrm{d}r\right] = 0.$$

We will prove this in several stages—the proof will culminate at the end of Section 4.3. First, however, we return to our main goal.

Proof of Lemma 2.2 for $k = -l$**.** Since $\dot{\omega} > 0$ and ω is smooth and invertible and K is compact, there is a $\mathsf{K}_K > 0$ such that $|\omega^{-1}(a) - \omega^{-1}(b)| \leq \mathsf{K}_K|a-b|$ for all a and b in $\omega(K)$. If $k|\omega(X_r^{\varepsilon}) - \omega(Y_r^{\varepsilon})| \leq \varepsilon^{\nu}$, $X_r^{\varepsilon} \in K$ and $Y_r^{\varepsilon} \in K$, then

$$|X_r^{\varepsilon} - Y_r^{\varepsilon}| = \left|\omega^{-1}(\omega(X_r^{\varepsilon})) - \omega^{-1}(\omega(Y_r^{\varepsilon}))\right| \leq \mathsf{K}_K|\omega(X_r^{\varepsilon}) - \omega(Y_r^{\varepsilon})| \leq \frac{\mathsf{K}_K}{k}\varepsilon^{\nu}.$$

Thus

$$\int_{r=0}^{t\wedge\tau_K^{\varepsilon}}\chi_{[-1,1]}\left(\frac{\mathsf{a}_r^{\varepsilon}}{\varepsilon^{\nu}}\right)\mathrm{d}r \leq \int_{r=0}^{t\wedge\tau_K^{\varepsilon}}\chi_{[-\mathsf{K}_K/k,\mathsf{K}_K/k]}\left(\frac{\mathsf{b}_r^{\varepsilon}}{\varepsilon^{\nu}}\right)\mathrm{d}r.$$

The claim then follows from Lemma 4.1. □

To proceed, we further decompose $\chi_{[-\varrho,\varrho]}(\frac{\mathsf{b}_r^{\varepsilon}}{\varepsilon^{\nu}})$. For $\vartheta \in \mathbb{R}$, define the "sawtooth" map

$$\mathbf{s}(\vartheta) \stackrel{\text{def}}{=} \vartheta - \left\lfloor \vartheta + \frac{1}{2}\right\rfloor.$$

[3] We use here the fact that $\nu < 1/3$.

Note that $\mathbf{s}$ is smooth on $\mathbb{R} \setminus (\mathbb{Z} + \frac{1}{2})$, i.e., on the set $\mathbf{s}^{-1}(-\frac{1}{2}, \frac{1}{2})$, and that for $\vartheta \in \mathbf{s}^{-1}(-\frac{1}{2}, \frac{1}{2})$, $\mathbf{s}'(\vartheta) = 1$. Also note that finally, $\sigma(\theta) - \sigma(\psi) = \sigma(\theta) - \sigma(\theta + \mathbf{s}(\psi - \theta))$ for all θ and ψ in $\mathbb{R}$.

We next define

$$\Delta_\varepsilon(r) \stackrel{\text{def}}{=} \mathbf{s}^2(\theta^\varepsilon_r - \psi^\varepsilon_r) + (X^\varepsilon_r - Y^\varepsilon_r)^2 \quad \text{and} \quad \tilde{\Delta}_\varepsilon(r) \stackrel{\text{def}}{=} \mathbf{s}^2(\theta^\varepsilon_r - \psi^\varepsilon_r) + \left(\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3}}\right)^2.$$

Fix $\eta \in (0, 1)$. We then write that

$$\chi_{[-\varrho,\varrho]}\left(\frac{\mathsf{b}^\varepsilon_r}{\varepsilon^\nu}\right) = \sum_{j=1}^3 I^\varepsilon_j(r)$$

where

$$\begin{aligned} I^\varepsilon_1(r) &\stackrel{\text{def}}{=} \chi_{[-\varrho,\varrho]}\left(\frac{\mathsf{b}^\varepsilon_r}{\varepsilon^\nu}\right)\chi_{[1,\infty)}\left(\frac{\Delta_\varepsilon(r)}{\varepsilon^\eta}\right) \\ I^\varepsilon_2(r) &\stackrel{\text{def}}{=} \chi_{[-\varrho,\varrho]}\left(\frac{\mathsf{b}^\varepsilon_r}{\varepsilon^\nu}\right)\chi_{[0,1)}\left(\frac{\Delta_\varepsilon(r)}{\varepsilon^\eta}\right)\chi_{[0,1)}\left(\frac{\tilde{\Delta}_\varepsilon(r)}{\varepsilon^{\eta/4}}\right) \\ I^\varepsilon_3(r) &\stackrel{\text{def}}{=} \chi_{[-\varrho,\varrho]}\left(\frac{\mathsf{b}^\varepsilon_r}{\varepsilon^\nu}\right)\chi_{[0,1)}\left(\frac{\Delta_\varepsilon(r)}{\varepsilon^\eta}\right)\chi_{(1,\infty)}\left(\frac{\tilde{\Delta}_\varepsilon(r)}{\varepsilon^{\eta/4}}\right). \end{aligned} \tag{11}$$

Each of these terms will require some work. We will consider I^ε_1 in Section 4.1, I^ε_2 in Section 4.2, and I^ε_4 in Section 4.3.

4.1. Averaging away from the diagonal

Let's first consider I^ε_1. For $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$, define

$$\hbar_b(\mathsf{c}) \stackrel{\text{def}}{=} (\sigma(\theta) - \sigma(\psi))^2 = \sigma^2(\theta) + \sigma^2(\psi) - 2\sigma(\theta)\sigma(\psi);$$

then

$$\langle \mathsf{b}^\varepsilon \rangle_t = \int_{r=0}^t \hbar_b(\mathsf{C}^\varepsilon_r)\mathrm{d}r.$$

For $\lambda \in (0, 1)$ and $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$, define

$$H^\lambda_b(\mathsf{c}) \stackrel{\text{def}}{=} \int_{r=0}^\infty \mathrm{e}^{-\lambda r}\, \hbar_b(\theta + \omega(x)r, \psi + \omega(y)r)\mathrm{d}r.$$

Lemma 4.2. *We have that* $H^\lambda_b \in C^\infty(\mathbb{R}^4)$. *Secondly,* $\sup_{\lambda\in(0,1)} \lambda \|H^\lambda_b\|_{C^2(\mathbb{R}^4)} < \infty$. *Thirdly,*

$$\omega(x)\frac{\partial H^\lambda_b}{\partial\theta}(\mathsf{c}) + \omega(y)\frac{\partial H^\lambda_b}{\partial\psi}(\mathsf{c}) = \lambda H^\lambda_b(\mathsf{c}) - \hbar_b(\mathsf{c})$$

for all $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$. *Finally, there is a constant* $\mathsf{K} > 0$ *such that*

$$\lambda H^\lambda_b(\mathsf{c}) \geq \frac{1}{\mathsf{K}}\left\{\mathbf{s}^2(\theta - \psi) + |\omega(x) - \omega(y)|^2\right\} - \mathsf{K}\left\{\sqrt{\lambda} + \frac{|x - y|}{\lambda}\right\}$$

for all $\lambda \in (0, 1)$ *and* $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$.

Proof. Again, the regularity and bound on derivatives are easy; one can differentiate under the integral and the formula for the derivative is fairly easy to see.

Fix $\lambda \in (0,1)$ and $\mathsf{c} = (x, \theta, y, \psi) \in \mathbb{R}^4$. We explicitly compute that

$$\begin{aligned}
\lambda H_b^\lambda(\mathsf{c}) &= \sum_{j,j' \in \mathbb{Z}} \sigma_j \sigma_{j'} \exp\left[2\pi \mathrm{i}(j+j')\theta\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j+j')\omega(x)} \\
&\quad + \sum_{j,j' \in \mathbb{Z}} \sigma_j \sigma_{j'} \exp\left[2\pi \mathrm{i}(j+j')\psi\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j+j')\omega(y)} \\
&\quad - 2 \sum_{j,j' \in \mathbb{Z}} \sigma_j \sigma_{j'} \exp\left[2\pi \mathrm{i}(j\theta + j'\psi)\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j\omega(x) + j'\omega(y))} \\
&= \bar{I} + I_1 + I_2 + I_3
\end{aligned}$$

where

$$\begin{aligned}
\bar{I} &\overset{\text{def}}{=} 2 \sum_{j \in \mathbb{Z}\setminus\{0\}} |\sigma_j|^2 \left\{1 - \exp\left[2\pi \mathrm{i} j(\theta - \psi)\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i} j(\omega(x) - \omega(y))}\right\} \\
I_1 &\overset{\text{def}}{=} \sum_{\substack{j,j' \in \mathbb{Z} \\ j+j' \neq 0}} \sigma_j \sigma_{j'} \left\{\exp\left[2\pi \mathrm{i}(j+j')\theta\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j+j')\omega(x)} \right.\\
&\quad \left. + \exp\left[2\pi \mathrm{i}(j+j')\psi\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j+j')\omega(y)}\right\} \\
I_2 &\overset{\text{def}}{=} -2 \sum_{\substack{j,j' \in \mathbb{Z} \\ j+j' \neq 0 \\ |j\omega(x) + j'\omega(y)| < \sqrt{\lambda}}} \sigma_j \sigma_{j'} \exp\left[2\pi \mathrm{i}(j\theta + j'\psi)\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j\omega(x) + j'\omega(y))} \\
I_3 &\overset{\text{def}}{=} -2 \sum_{\substack{j,j' \in \mathbb{Z} \\ j+j' \neq 0 \\ |j\omega(x) + j'\omega(y)| \geq \sqrt{\lambda}}} \sigma_j \sigma_{j'} \exp\left[2\pi \mathrm{i}(j\theta + j'\psi)\right] \frac{\lambda}{\lambda - 2\pi \mathrm{i}(j\omega(x) + j'\omega(y))}.
\end{aligned}$$

We immediately see that

$$|I_1| \leq \left(\sum_{j \in \mathbb{Z}} |\sigma_j|\right)^2 \frac{\lambda}{\omega_-} \quad \text{and} \quad |I_3| \leq 2 \left(\sum_{j \in \mathbb{Z}} |\sigma_j|\right)^2 \sqrt{\lambda}.$$

Next observe that if $j + j' \neq 0$ and $|j\omega(x) + j'\omega(y)| < \sqrt{\lambda}$, then

$$\begin{aligned}
\left(|j| + |j'|\right) |\omega(y) - \omega(x)| &= |j(\omega(y) - \omega(x))| + |j'(\omega(x) - \omega(y))| \\
&\geq |j(\omega(y) - \omega(x)) + j\omega(x) + j'\omega(y)| + |j'(\omega(x) - \omega(y)) \\
&\quad + j\omega(x) + j'\omega(y)| - 2\sqrt{\lambda} \\
&= |j + j'|(\omega(y) + \omega(x)) - 2\sqrt{\lambda} \geq 2\omega_- - 2\sqrt{\lambda},
\end{aligned}$$

so if $\lambda < \omega_-^2/4$, then either $|j||\omega(y)-\omega(x)| > \omega_-/2$ or $|j'||\omega(y)-\omega(x)| > \omega_-/2$. Thus

$$\begin{aligned} |I_2| &\le 2\left\{\sum_{j'\in\mathbb{Z}} |\sigma_{j'}|\right\}\left\{\sum_{\substack{j\in\mathbb{Z}\\ |j||\omega(y)-\omega(x)|>\omega_-/2}} |\sigma_j|\right\} \\ &\le 2|\omega(x)-\omega(y)|\left\{\sum_{j'\in\mathbb{Z}} |\sigma_{j'}|\right\}\left\{\sum_{\substack{j\in\mathbb{Z}\\ |j||\omega(y)-\omega(x)|>\omega_-/2}} \frac{j|\sigma_j|}{|j||\omega(y)-\omega(x)|}\right\} \\ &\le \frac{4}{\omega_-}|\omega(x)-\omega(y)|\left\{\sum_{j'\in\mathbb{Z}} |\sigma_{j'}|\right\}\left\{\sum_{j\in\mathbb{Z}} |j||\sigma_j|\right\}. \end{aligned}$$

We use here the fact that since $\sigma \in C^\infty$, $\sum_{j\in\mathbb{Z}} |j||\sigma_j| < \infty$.

We finally bound $\bar{I}$ from below. We first write that

$$\bar{I} = 4\sum_{j=1}^{\infty} |\sigma_j|^2 \left\{1 - \Re\left(\exp\left[2\pi \mathrm{i} j(\theta-\psi)\right] \frac{\lambda}{\lambda - 2\pi\mathrm{i} j(\omega(x)-\omega(y))}\right)\right\}.$$

We will then optimize over two lower bounds. For convenience, set $\boldsymbol{\Theta} \stackrel{\text{def}}{=} (\theta-\psi)$, $\mathbf{W} \stackrel{\text{def}}{=} 2\pi(\omega(x)-\omega(y))$, and $\bar{w} \stackrel{\text{def}}{=} 2\pi(\omega_+ - \omega_-)$.

First note that for each $j \in \mathbb{Z}$,

$$\left|\exp\left[2\pi\mathrm{i} j(\theta-\psi)\right] \frac{\lambda}{\lambda - 2\pi\mathrm{i} j(\omega(x)-\omega(y))}\right| = \frac{\lambda}{\sqrt{\lambda^2 + j^2\mathbf{W}^2}};$$

thus

$$\bar{I} \ge 4\sum_{j=1}^{\infty} |\sigma_j|^2 \left\{1 - \frac{\lambda}{\sqrt{\lambda^2 + j^2\mathbf{W}^2}}\right\}.$$

Since $q \mapsto \frac{q}{q+j^2\mathbf{W}^2}$ is increasing on $(0,1)$, we get that for $\lambda \in (0,1)$

$$\begin{aligned} 4\left\{1 - \frac{\lambda}{\sqrt{\lambda^2 + j^2\mathbf{W}^2}}\right\} &\ge 4\left\{1 - \frac{1}{\sqrt{1 + j^2\mathbf{W}^2}}\right\} = 4j^2\int_{z=0}^{\mathbf{W}} \frac{z}{(1+j^2z^2)^{3/2}}\mathrm{d}z \\ &\ge \frac{4j^2}{(1+j^2\bar{w}^2)^{3/2}}\int_{z=0}^{\mathbf{W}} z\mathrm{d}z \\ &\ge \frac{2j^2\mathbf{W}^2}{(1+j^2\bar{w}^2)^{3/2}}. \end{aligned}$$

Thus $\bar{I} \ge \varrho_1 \mathbf{W}^2$ where

$$\varrho_1 \stackrel{\text{def}}{=} \|\sigma\|^2_{L^2([0,1))} \left\{\inf_{z\ge 1} \frac{2z^2}{1+z^2\bar{w}^2}\right\}.$$

On the other hand, we can explicitly compute that for each $j \in \{1, 2 \dots\}$,

$$
\begin{aligned}
1 - \Re\left(\exp\left[2\pi \mathrm{i} j \boldsymbol{\Theta}\right] \frac{\lambda}{\lambda - j\mathbf{W}}\right) &= 1 - \cos(2\pi j \boldsymbol{\Theta}) \frac{\lambda^2}{\lambda^2 + j^2\mathbf{W}^2} \\
&\quad + \sin(2\pi j \boldsymbol{\Theta}) \frac{j\lambda\mathbf{W}}{\lambda^2 + j^2\mathbf{W}^2} \\
&= 1 - \cos(2\pi j \boldsymbol{\Theta}) + \cos(2\pi j \boldsymbol{\Theta}) \frac{j^2\mathbf{W}^2}{\lambda^2 + j^2\mathbf{W}^2} + \sin(2\pi \boldsymbol{\Theta}) \frac{j\lambda\mathbf{W}}{\lambda^2 + j^2\mathbf{W}^2} \\
&= 1 - \cos(2\pi j \boldsymbol{\Theta}) + \frac{j\mathbf{W}}{\lambda^2 + j^2\mathbf{W}^2} \Re\left(\mathrm{e}^{2\pi \mathrm{i} j \boldsymbol{\Theta}} (j\mathbf{W} - \sqrt{-1}\lambda)\right) \\
&\geq 1 - \cos(2\pi j \boldsymbol{\Theta}) - \frac{j\mathbf{W}}{\sqrt{\lambda^2 + j^2\mathbf{W}^2}} \\
&\geq 1 - \cos(2\pi j \boldsymbol{\Theta}) - \frac{j\mathbf{W}}{\lambda}.
\end{aligned}
$$

Hence

$$
\bar{I} \geq \bar{I}^\circ(\boldsymbol{\Theta}) - \mathsf{K}_2 \frac{\mathbf{W}}{\lambda}
$$

where $\mathsf{K}_2 \stackrel{\text{def}}{=} 4 \sum_{j=1}^\infty j|\sigma_j|^2$ and

$$
\bar{I}^\circ(\vartheta) \stackrel{\text{def}}{=} 4 \sum_{j=1}^\infty |\sigma_j|^2 \{1 - \cos(2\pi j \vartheta)\}
$$

for all $\vartheta \in \mathbb{R}$. We claim that

$$
\varrho_2 \stackrel{\text{def}}{=} \inf_{\vartheta \in \mathbb{R}} \frac{\bar{I}^\circ(\vartheta)}{\mathbf{s}^2(\vartheta)} > 0.
$$

First note that $\bar{I}^\circ$ is right-continuous. Fix $\vartheta \in \mathbb{R} \setminus \mathbb{Z}$. Then $|\mathbf{s}(\vartheta)| > 0$ and

$$
\bar{I}^\circ(\vartheta) = 4 \sum_{\substack{1 \leq j \leq \infty \\ j \notin \mathbb{Z}/\theta}} |\sigma_j|^2 \{1 - \cos(2\pi j \vartheta)\}.
$$

Since $1 - \cos(2\pi j \vartheta) > 0$ for all positive integers $j \notin \mathbb{Z}/\vartheta$, if $\bar{I}^\circ(\theta) = 0$ then $\sigma_j = 0$ for all positive $j \notin \mathbb{Z}/\vartheta$. This violates the assumption that σ is exactly 1-periodic; hence $\bar{I}^\circ > 0$ on $\mathbb{R} \setminus \mathbb{Z}$. Next fix $\{\vartheta_n\} \in \mathbb{R}$ such that $\lim_{n\to\infty} \vartheta_n = 0$. Fix $j_* \in \{1, 2 \dots \infty\}$ such that $\sigma_{j_*} \neq 0$ (which is possible since σ has period exactly 1). Then

$$
\varliminf_{n\to\infty} \frac{\bar{I}^\circ(\vartheta_n)}{\mathbf{s}^2(\vartheta_n)} \geq 4|\sigma_{j_*}|^2 \varliminf_{n\to\infty} \frac{1 - \cos(2\pi j \vartheta_n)}{\vartheta_n^2} = \frac{1}{2}(2\pi j^*)^2.
$$

Collecting our thoughts together, we have that $\varrho_2 > 0$. Thus

$$
\bar{I} \geq \varrho_2 d^2(\boldsymbol{\Theta}) - \mathsf{K}_2 \frac{\mathbf{W}}{\lambda}.
$$

Combining together our two bounds, we have that

$$\bar{I} \geq \max\left\{\varrho_1 \mathbf{W}^2, \varrho_2 d^2(\boldsymbol{\Theta}) - \mathsf{K}_2 \frac{\mathbf{W}}{\lambda}\right\} \geq \max\left\{\varrho_1 \mathbf{W}^2, \varrho_2 d^2(\boldsymbol{\Theta})\right\} - \mathsf{K}_2 \frac{\mathbf{W}}{\lambda}$$
$$\geq \frac{1}{2}\left\{\varrho_1 \mathbf{W}^2 + \varrho_2 d^2(\boldsymbol{\Theta})\right\} - \mathsf{K}_2 \frac{\mathbf{W}}{\lambda}.$$

Combine our estimates together and use the fact that $\lambda < \sqrt{\lambda}$ for $\lambda \in (0,1)$. □

We can now complete this subsection and bound I_1^ε of (11).

Lemma 4.3. *We have that*

$$\mathbb{E}\left[\int_{r=0}^{t\wedge\tau_K^\varepsilon} I_1^\varepsilon(r)\mathrm{d}r\right] \leq \mathsf{K}\left\{\varepsilon^{\nu-\eta} + \varepsilon^{1-\eta-5\nu/2}\right\}.$$

Proof. Set $\lambda_\varepsilon \stackrel{\text{def}}{=} \varepsilon^{\nu/2}$ and define

$$U_t^\varepsilon \stackrel{\text{def}}{=} 2\varrho^2 \varepsilon^{2\nu-\eta} \varphi_0^{(-2)}\left(\frac{1}{\varrho}\frac{\mathsf{b}_t^\varepsilon}{\varepsilon^\nu}\right) + \varepsilon^{1-\eta} H_a^{\lambda_\varepsilon}(\mathsf{C}_t^\varepsilon)\varphi_0\left(\frac{1}{\varrho}\frac{\mathsf{b}_t^\varepsilon}{\varepsilon^\nu}\right).$$

Then

$$\int_{r=0}^{t\wedge\tau_K^\varepsilon} \lambda_\varepsilon H_a^{\lambda_\varepsilon}(\mathsf{C}_r^\varepsilon)\varphi_0\left(\frac{1}{\varrho}\frac{\mathsf{b}_r^\varepsilon}{\varepsilon^\nu}\right)\mathrm{d}r = U_{t\wedge\tau_K^\varepsilon}^\varepsilon - U_0^\varepsilon - \sum_{j=1}^{3}\int_{r=0}^{t\wedge\tau_K^\varepsilon} \mathscr{E}_{j,\varepsilon}(r)\mathrm{d}r$$
$$- \sum_{j=4}^{6}\int_{r=s\wedge\tau_K^\varepsilon}^{t\wedge\tau_K^\varepsilon} \mathscr{E}_{j,\varepsilon}(r)\mathrm{d}W_r$$

where

$$\mathscr{E}_{1,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{1-\eta}}{2}\left\{\frac{\partial^2 H_a^{\lambda_\varepsilon}}{\partial x^2}(\mathsf{C}_r^\varepsilon)\sigma^2(\theta_r^\varepsilon) + \frac{\partial^2 H_a^{\lambda_\varepsilon}}{\partial y^2}(\mathsf{C}_r^\varepsilon)\sigma^2(\psi_r^\varepsilon)\right\}\varphi_0\left(\frac{1}{\varrho}\frac{\mathsf{b}_r^\varepsilon}{\varepsilon^\nu}\right)$$

$$\mathscr{E}_{2,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{1-\eta-2\nu}}{2\varrho^2} H_a^{\lambda_\varepsilon}(\mathsf{C}_r^\varepsilon)\ddot{\varphi}_0\left(\frac{1}{\varrho}\frac{\mathsf{b}_r^\varepsilon}{\varepsilon^\nu}\right)\hbar_b(\mathsf{C}_r^\varepsilon)$$

$$\mathscr{E}_{3,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{1-\eta-\nu}}{\varrho}\left\{\frac{\partial H_a^{\lambda_\varepsilon}}{\partial x}(\mathsf{C}_r^\varepsilon)\sigma(\theta_r^\varepsilon) + \frac{\partial H_a^{\lambda_\varepsilon}}{\partial y}(\mathsf{C}_r^\varepsilon)\sigma(\psi_r^\varepsilon)\right\}\dot{\varphi}_0\left(\frac{1}{\varrho}\frac{\mathsf{b}_r^\varepsilon}{\varepsilon^\nu}\right)(\sigma(\theta_r^\varepsilon) - \sigma(\psi_r^\varepsilon))$$

$$\mathscr{E}_{4,\varepsilon}(r) \stackrel{\text{def}}{=} 2\varrho\varepsilon^{\nu-\eta}\dot{\varphi}_0\left(\frac{1}{\varrho}\frac{\mathsf{b}_r^\varepsilon}{\varepsilon^\nu}\right)(\sigma(\theta_r^\varepsilon) - \sigma(\psi_r^\varepsilon))$$

$$\mathscr{E}_{5,\varepsilon}(r) \stackrel{\text{def}}{=} \varepsilon^{1-\eta}\left\{\frac{\partial H_a^{\lambda_\varepsilon}}{\partial x}(\mathsf{C}_r^\varepsilon)\sigma(\theta_r^\varepsilon) + \frac{\partial H_a^{\lambda_\varepsilon}}{\partial y}(\mathsf{C}_r^\varepsilon)\sigma(\psi_r^\varepsilon)\right\}\varphi_0\left(\frac{1}{\varrho}\frac{\mathsf{b}_r^\varepsilon}{\varepsilon^\nu}\right)$$

$$\mathscr{E}_{6,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{1-\eta-\nu}}{\varrho} H_a^{\lambda_\varepsilon}(\mathsf{C}_r^\varepsilon)\dot{\varphi}_0\left(\frac{1}{\varrho}\frac{\mathsf{b}_r^\varepsilon}{\varepsilon^\nu}\right)(\sigma(\theta_r^\varepsilon) - \sigma(\psi_r^\varepsilon)).$$

We then have that $|\mathscr{E}_{4,\varepsilon}(r)| \leq \mathsf{K}\varepsilon^{\nu-\eta}$ and $|\mathscr{E}_{j,\varepsilon}(r)| \leq \mathsf{K}\varepsilon^{1-\eta-2\nu}/\lambda_\varepsilon = \mathsf{K}\varepsilon^{1-\eta-5\nu/2}$ for all other j. We also note that

$$\begin{aligned}
&\varepsilon^{-\eta}\lambda_\varepsilon H_b^{\lambda_\varepsilon}(\mathsf{c})\varphi_0\left(\frac{x-y}{\varepsilon^\nu}\right)\\
&\quad\geq \frac{\varepsilon^{-\eta}}{\mathsf{K}}\left\{\mathbf{s}^2(\theta-\psi)+|\omega(x)-\omega(y)|^2\right\}\varphi_0\left(\frac{x-y}{\varepsilon^\nu}\right)-\mathsf{K}\frac{\varepsilon^\nu}{\lambda_\varepsilon}-\mathsf{K}\sqrt{\lambda_\varepsilon}\\
&\quad\geq \frac{1}{\mathsf{K}}\frac{\mathbf{s}^2(\theta-\psi)+|\omega(x)-\omega(y)|^2}{\varepsilon^\eta}\varphi_0\left(\frac{x-y}{\varepsilon^\nu}\right)\chi_{[1,\infty)}\left(\frac{\Delta_\varepsilon(r)}{\varepsilon^\eta}\right)-\mathsf{K}\varepsilon^{\nu/2}.
\end{aligned}$$

We next note that

$$\begin{aligned}
&\varliminf_{\varepsilon\searrow 0}\inf\left\{\frac{\mathbf{s}^2(\theta-\psi)+|\omega(x)-\omega(y)|^2}{\varepsilon^\eta}:(x,\theta,y,\psi)\in\mathbb{R}^2,|x|<L,|y|<L\right.\\
&\qquad\left.\text{and } \mathbf{s}^2(\theta-\psi)+|x-y|^2>\varepsilon^\eta\right\}>0.
\end{aligned}$$

The claimed result follows.[4] □

4.2. *Intermediate averaging*

Our next task in the analysis of the errors of (11) is I_3^ε. We start by simplifying the problem.

Lemma 4.4. *For $\varepsilon \in (0, \bar{\varepsilon})$, we have that*

$$|I_3^\varepsilon(r)| \leq \varphi_0\left(\frac{\mathbf{s}(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\varepsilon^{\eta/2}}\right)\varphi_1\left(\sqrt{2}\frac{X_r^\varepsilon-Y_r^\varepsilon}{\varepsilon^{1/3+\eta/8}}\right).$$

Proof. If $\Delta_\varepsilon(r) < \varepsilon^\eta$, then $|\mathbf{s}(\theta_r^\varepsilon-\psi_r^\varepsilon)| \leq \varepsilon^{\eta/2}$. If in addition $\tilde{\Delta}_\varepsilon(r) > \varepsilon^{\eta/4}$, then we have that

$$\varepsilon^{\eta/4} \leq \mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)+\left(\frac{X_r^\varepsilon-Y_r^\varepsilon}{\varepsilon^{1/3}}\right)^2 \leq \varepsilon^\eta+\left(\frac{X_r^\varepsilon-Y_r^\varepsilon}{\varepsilon^{1/3}}\right)^2.$$

Let $\bar{\varepsilon} \in (0,1)$ be such that $\varepsilon^\eta < \frac{1}{2}\varepsilon^{\eta/4}$ for $\varepsilon \in (0,\bar{\varepsilon})$. Thus

$$I_3^\varepsilon(r) \leq \chi_{[0,1)}\left(\frac{\mathbf{s}(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\varepsilon^{\eta/2}}\right)\chi_{[1/\sqrt{2},\infty)}\left(\frac{X_r^\varepsilon-Y_r^\varepsilon}{\varepsilon^{1/3+\eta/8}}\right).$$

The claim readily follows. □

We now claim that as long as $2\left|\frac{X_r^\varepsilon-Y_r^\varepsilon}{\varepsilon^{1/3+\eta/8}}\right| < \varrho$ is uniformly bounded away from zero, then $\theta^\varepsilon-\psi^\varepsilon$ is quickly varying. Indeed, in this case we have that

$$\mathrm{d}(\theta_r^\varepsilon-\psi_r^\varepsilon)=\frac{1}{\varepsilon}\left\{\omega(X_r^\varepsilon)-\omega(Y_r^\varepsilon)\right\}\mathrm{d}r\approx\dot{\omega}(X_r^\varepsilon)\left\{\frac{X_r^\varepsilon-Y_r^\varepsilon}{\varepsilon}\right\}\mathrm{d}r.$$

The other fact we have at our disposal is the fact that the support of $\varphi_0\left(\frac{\mathbf{s}(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\varepsilon^{\eta/2}}\right)$ is "thin".

[4] We use here the fact that $\nu > \eta$ and $\eta + 5\nu/2 < 7\nu/2 < 1$.

To quantify all of this, define

$$H_\varepsilon(\vartheta) \stackrel{\text{def}}{=} \int_{r=0}^{\infty} \mathrm{e}^{-r}\varphi_0\left(\frac{\mathbf{s}(\vartheta+r)}{\varepsilon^{\eta/2}}\right)\mathrm{d}r$$

for all $\varepsilon \in (0,1)$ and $\vartheta \in \mathbb{R}$.

Lemma 4.5. *For each* $\varepsilon \in (0,1)$, $H_\varepsilon \in C(\mathbb{R}) \cap C^\infty(\mathbf{s}^{-1}(-1/2,1/2))$. *Secondly,*

$$H'_\varepsilon(\vartheta) = H_\varepsilon(\vartheta) - \varphi_0\left(\frac{\mathbf{s}(\vartheta)}{\varepsilon^{\eta/2}}\right)$$

for all $\vartheta \in \mathbb{R}$. *Finally,* $\sup_{\varepsilon\in(0,1)} \varepsilon^{-\eta/2}\|H_\varepsilon\|_{C(\mathbb{R})} < \infty$.

Proof. We have that

$$H_\varepsilon(\vartheta) = \mathrm{e}^{\vartheta}\int_{r=\vartheta}^{\infty} \mathrm{e}^{-r}\varphi_0\left(\frac{\mathbf{s}(r)}{\varepsilon^{\eta/2}}\right)\mathrm{d}r$$

for all $\vartheta \in \mathbb{R}\setminus\mathbb{Z}$. This easily implies the stated ODE. Since H_ε is 1-periodic, a bound on $\|H_\varepsilon\|_{C(\mathbb{R})}$ will follow from a bound on $\|H_\varepsilon\|_{C([0,1])}$. For $\vartheta \in [0,1)$,

$$H_\varepsilon(\vartheta) = -\mathrm{e}^{\vartheta}\left\{\int_{r=\vartheta}^{1} \mathrm{e}^{-r}\varphi_0\left(\frac{\mathbf{s}(r)}{\varepsilon^{\eta/2}}\right)\mathrm{d}r + \left\{\sum_{k=1}^{\infty}\mathrm{e}^{-k}\right\}\int_{r=0}^{1} \mathrm{e}^{-r}\varphi_0\left(\frac{\mathbf{s}(r)}{\varepsilon^{\eta/2}}\right)\mathrm{d}r\right\}.$$

This gives us the remainder of the claim. □

We can now finish off our analysis of I_3^ε of (11).

Lemma 4.6. *We have that*

$$\varlimsup_{\varepsilon\searrow 0}\mathbb{E}\left[\int_{r=0}^{t\wedge\tau_K^\varepsilon} I_3^\varepsilon(r)\mathrm{d}r\right] \le \mathsf{K}\varepsilon^{\eta/8}.$$

Proof. We first write that

$$\int_{r=0}^{t\wedge\tau_K^\varepsilon} I_3^\varepsilon(r)\mathrm{d}r \le \int_{r=0}^{t\wedge\tau_K^\varepsilon}\varphi_0\left(\frac{\mathbf{s}(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\varepsilon^{\eta/2}}\right)\varphi_1\left(\sqrt{2}\frac{X_r^\varepsilon-Y_r^\varepsilon}{\varepsilon^{1/3+\eta/8}}\right)\mathrm{d}r.$$

Set

$$U_r^\varepsilon \stackrel{\text{def}}{=} \varepsilon^{2/3-\eta/8}\,\mathrm{sgn}(X_r^\varepsilon-Y_r^\varepsilon)H_\varepsilon(\theta_r^\varepsilon-\psi_r^\varepsilon)\varphi_1\left(\sqrt{2}\frac{X_r^\varepsilon-Y_r^\varepsilon}{\varepsilon^{1/3+\eta/8}}\right).$$

Thus

$$U_{t\wedge\tau_K^\varepsilon}^\varepsilon - U_0^\varepsilon = \int_{r=0}^{t\wedge\tau_K^\varepsilon} A_\varepsilon(r)\mathrm{d}r + \sum_{j=1}^{2}\int_{r=0}^{t\wedge\tau_K^\varepsilon}\mathscr{E}_{j,\varepsilon}(r)\mathrm{d}r - \int_{r=0}^{t\wedge\tau_K^\varepsilon}\mathscr{E}_{3,\varepsilon}(r)\mathrm{d}W_r$$

where

$$A_\varepsilon(r) \stackrel{\text{def}}{=} \varepsilon^{-1/3-\eta/8}\operatorname{sgn}(X^\varepsilon_r - Y^\varepsilon_r)\{\omega(X^\varepsilon_r) - \omega(Y^\varepsilon_r)\}\varphi_0\left(\frac{\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)}{\varepsilon^{\eta/2}}\right)\varphi_1\left(\sqrt{2}\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3+\eta/8}}\right)$$

$$\mathscr{E}_{1,\varepsilon}(r) \stackrel{\text{def}}{=} -\varepsilon^{-1/3-\eta/8}\operatorname{sgn}(X^\varepsilon_r - Y^\varepsilon_r)\{\omega(X^\varepsilon_r) - \omega(Y^\varepsilon_r)\}H_\varepsilon(\theta^\varepsilon_r - \psi^\varepsilon_r)\varphi_1\left(\sqrt{2}\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3+\eta/8}}\right)$$

$$\mathscr{E}_{2,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{2/3-\eta/8}}{\varepsilon^{2/3+\eta/4}}\operatorname{sgn}(X^\varepsilon_r - Y^\varepsilon_r)H_\varepsilon(\theta^\varepsilon_r - \psi^\varepsilon_r)\ddot{\varphi}_1\left(\sqrt{2}\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3+\eta/8}}\right)(\sigma(\theta^\varepsilon_r) - \sigma(\psi^\varepsilon_r))^2$$

$$\mathscr{E}_{3,\varepsilon}(r) \stackrel{\text{def}}{=} \sqrt{2}\frac{\varepsilon^{2/3-\eta/8}}{\varepsilon^{1/3+\eta/8}}\operatorname{sgn}(X^\varepsilon_r - Y^\varepsilon_r)H_\varepsilon(\theta^\varepsilon_r - \psi^\varepsilon_r)\dot{\varphi}_1\left(\sqrt{2}\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3+\eta/8}}\right)(\sigma(\theta^\varepsilon_r) - \sigma(\psi^\varepsilon_r)).$$

We then compute that

$$|U^\varepsilon_r| \le \mathsf{K}\varepsilon^{2/3-\eta/8+\eta/2} = \mathsf{K}\varepsilon^{2/3+3\eta/8}$$
$$|\mathscr{E}_{1,\varepsilon}(r)| \le \mathsf{K}\varepsilon^{\eta/2} \qquad |\mathscr{E}_{2,\varepsilon}(r)| \le \mathsf{K}\varepsilon^{\eta/2-\eta/8-\eta/4} = \mathsf{K}\varepsilon^{\eta/8}$$
$$|\mathscr{E}_{3,\varepsilon}(r)| \le \mathsf{K}\varepsilon^{1/3-\eta/4+\eta/2} = \mathsf{K}\varepsilon^{1/3+\eta/4}.$$

We finally observe that

$$\varliminf_{\varepsilon\searrow 0}\inf\left\{\frac{(\omega(x)-\omega(y))\operatorname{sgn}(x-y)}{\varepsilon^{1/3+\eta/8}} : |x| < L, |y| < L \text{ and } \varphi_1\left(\sqrt{2}\frac{x-y}{\varepsilon^{1/3+\eta/8}}\right) \neq 0\right\} > 0.$$

This completes the proof. □

4.3. *Averaging near the diagonal*

We finally consider I^ε_2 of (11), which measures the amount of time that $\tilde{\Delta}_\varepsilon(r)$ is small, i.e., the amount of time that $(X^\varepsilon, \theta^\varepsilon)$ and $(Y^\varepsilon, \psi^\varepsilon)$ are close (in an appropriate sense). The essence of our argument is that there are three time scales when $(X^\varepsilon, \theta^\varepsilon)$ and $(Y^\varepsilon, \psi^\varepsilon)$ are close (this was developed in [9]; see also [4,3]). The axial coordinates X^ε and Y^ε vary the most slowly (the *macroscopic* scale). The angles θ^ε and ψ^ε vary the most quickly (the *microscopic* scale). In between, we have a *mesoscopic* scale, on which the angle of the vector $(\varepsilon^{-1/3}(X^\varepsilon - Y^\varepsilon), \mathbf{s}(\theta^\varepsilon - \psi^\varepsilon))$ fluctuates. Note that by uniqueness of solutions to SDE's,

$$\inf\{r \ge 0 : \tilde{\Delta}_\varepsilon(r) = 0\} = \infty$$

$\mathbb{P}$-a.s. (if $\tilde{\Delta}_\varepsilon(r) = 0$, then $(X^\varepsilon_r, \theta^\varepsilon_r) \sim (Y^\varepsilon_r, \psi^\varepsilon_r)$). Fix a symbol $\star$ and define $\mathbb{R}^\star \stackrel{\text{def}}{=} \mathbb{R}\cup\{\star\}$; we give $\mathbb{R}^\star$ the usual topology of one-point compactification [8]. For $\varepsilon \in (0,1)$ and $r \ge 0$, define

$$\Lambda^\varepsilon_r \stackrel{\text{def}}{=} \begin{cases} \dfrac{\varepsilon^{-1/3}(X^\varepsilon_r - Y^\varepsilon_r)}{\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)} & \text{if } \mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r) \neq 0 \\ \star & \text{if } \mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r) = 0. \end{cases}$$

Note that if $\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r) \neq 0$, then

$$\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3}} = \frac{\Lambda^\varepsilon_r}{\sqrt{1+(\Lambda^\varepsilon_r)^2}}\sqrt{\tilde{\Delta}_\varepsilon(r)}$$
$$|\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)| = \frac{1}{\sqrt{1+(\Lambda^\varepsilon_r)^2}}\sqrt{\tilde{\Delta}_\varepsilon(r)}.$$

Thus we can make a coordinate change from $(X^\varepsilon, \theta^\varepsilon, Y^\varepsilon, \psi^\varepsilon)$ to $(X^\varepsilon, \theta^\varepsilon, \Lambda^\varepsilon, \tilde{\Delta}_\varepsilon)$. For any $r > 0$ such that $\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r) \neq 0$, we informally have that

$$\begin{aligned} \mathrm{d}\Lambda^\varepsilon_r &= \frac{\varepsilon^{-1/3}(\sigma(\theta^\varepsilon_r) - \sigma(\psi^\varepsilon_r))}{\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)}\mathrm{d}W_r - \varepsilon^{-4/3}\frac{(X^\varepsilon_r - Y^\varepsilon_r)(\omega(X^\varepsilon_r) - \omega(Y^\varepsilon_r))}{\mathbf{s}^2(\theta^\varepsilon_r - \psi^\varepsilon_r)}\mathrm{d}r \\ &\approx \varepsilon^{-1/3}\dot{\sigma}(\theta^\varepsilon_r)\mathrm{d}W_r - \varepsilon^{-2/3}\dot{\omega}(X^\varepsilon_r)(\Lambda^\varepsilon_r)^2\mathrm{d}r. \end{aligned}$$

The import of this is that we expect to see fluctuations of Λ^ε_r on time scales of order $\varepsilon^{2/3}$. Since the fluctuations of θ^ε are on the shorter time scale ε, we should further be able to simplify the effective dynamics of Λ^ε. Recalling (3), we have that

$$\mathrm{d}\Lambda^\varepsilon_r \approx \varepsilon^{-1/3}\kappa_1\mathrm{d}W_r - \varepsilon^{-2/3}\dot{\omega}(X^\varepsilon_r)(\Lambda^\varepsilon_r)^2\mathrm{d}r. \tag{12}$$

We finally note that this SDE implies that Λ^ε should reach an invariant measure on a time interval of order $\varepsilon^{2/3}$. This invariant measure can in fact be explicitly described; see (17).

Note that

$$\begin{aligned} \mathrm{d}\tilde{\Delta}_\varepsilon(r) &= \frac{2}{\varepsilon}\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)\left\{\omega(X^\varepsilon_r) - \omega(Y^\varepsilon_r)\right\}\mathrm{d}r + \frac{2}{\varepsilon^{2/3}}(X^\varepsilon_r - Y^\varepsilon_r)(\sigma(\theta^\varepsilon_r) - \sigma(\psi^\varepsilon_r))\mathrm{d}W_r \\ &\quad + \frac{1}{\varepsilon^{2/3}}(\sigma(\theta^\varepsilon_r) - \sigma(\psi^\varepsilon))^2\mathrm{d}r = \varepsilon^{-2/3}A_\varepsilon(r)\mathrm{d}r + \varepsilon^{-1/3}B_\varepsilon(r)\mathrm{d}W_r \end{aligned}$$

for $r > 0$ such that $|\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)| \neq \frac{1}{2}$, where

$$\begin{aligned} A_\varepsilon(r) &= 2\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)\left(\frac{\omega(X^\varepsilon_r) - \omega(Y^\varepsilon_r)}{\varepsilon^{1/3}}\right) + (\sigma(\theta^\varepsilon_r) - \sigma(\psi^\varepsilon_r))^2 \\ B_\varepsilon(r) &= 2\left(\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3}}\right)(\sigma(\theta^\varepsilon_r) - \sigma(\psi^\varepsilon_r)). \end{aligned} \tag{13}$$

Also, $|A_\varepsilon(r)| \leq \mathsf{K}\tilde{\Delta}_\varepsilon(r)$ and $|B_\varepsilon(r)| \leq \mathsf{K}\tilde{\Delta}_\varepsilon(r)$.

We will proceed in a number of steps. For convenience, set

$$\tilde{\Delta}'_\varepsilon(r) \stackrel{\text{def}}{=} \frac{\tilde{\Delta}_\varepsilon(r)}{\varepsilon^{\eta/4}}.$$

Then set

$$u_{1,\varepsilon}(r) \stackrel{\text{def}}{=} \left\{\frac{A_\varepsilon(r)}{\tilde{\Delta}_\varepsilon(r)} - \frac{1}{2}\frac{B^2_\varepsilon(r)}{\tilde{\Delta}^2_\varepsilon(r)}\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r)) \tag{14}$$

for all $\varepsilon \in (0, 1)$ and $r \geq 0$.

Define $\ell(z) \stackrel{\text{def}}{=} \ln(e + 1/z)$ for all $z > 0$; then ℓ is bounded from above for z large and behaves like $z \mapsto \ln z^{-1}$ for z small. We also note that we will localize all of our steps by τ^ε_K.

Our first result is essentially to show that $\tilde{\Delta}'_\varepsilon$ doesn't spend too much time near the origin. We should almost be able to do this by bounding $u_{1,\varepsilon}$. It would be an easy task if the term in braces in (14) were bounded from below away from zero. In fact, this will only be true once we do a bit more averaging. We also note that the bound on $u_{1,\varepsilon}$ will use the bounds of Sections 4.2 and 4.3 to control some of the errors.

Lemma 4.7. *For each $L > 0$,*

$$\mathbb{E}\left[\int_{r=0}^{t\wedge\tau_K^\varepsilon} u_{1,\varepsilon}(r)\mathrm{d}r\right] \le \mathsf{K}\ell(\varepsilon)\mathbb{E}\left[\int_{r=0}^{t\wedge\tau_K^\varepsilon} \left\{I_1^\varepsilon(r) + I_3^\varepsilon(r)\right\}\mathrm{d}r\right] + \mathsf{K}\varepsilon^{2/3}\ell(\varepsilon).$$

Proof. Fix $\delta > 0$ and set

$$U_r^{\delta,\varepsilon} \stackrel{\text{def}}{=} \varepsilon^{2/3}\ln(\tilde{\Delta}_\varepsilon(r) + \delta)\varphi_0(\tilde{\Delta}'_\varepsilon(r)).$$

By Ito's rule,

$$\int_{r=0}^{t\wedge\tau_K^\varepsilon} u_{1,\varepsilon}^\delta(r)\mathrm{d}r = U_{t\wedge\tau_K^\varepsilon}^{\delta,\varepsilon} - U_0^{\delta,\varepsilon} - \int_{r=0}^{t\wedge\tau_K^\varepsilon} \mathscr{E}_{\delta,\varepsilon}(r)\mathrm{d}r - \{M_{t\wedge\tau_K^\varepsilon} - M_0\},$$

where

$$u_{1,\varepsilon}^\delta(r) \stackrel{\text{def}}{=} \left\{\frac{A_\varepsilon(r)}{\tilde{\Delta}_\varepsilon(r)+\delta} - \frac{1}{2}\frac{B_\varepsilon^2(r)}{(\tilde{\Delta}_\varepsilon(r)+\delta)^2}\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

$$\mathscr{E}_{\delta,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{A_{1,\varepsilon}(r)}{\varepsilon^{\eta/4}}\ln(\tilde{\Delta}_\varepsilon(r)+\delta)\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r)) + \frac{B_\varepsilon(r)}{\tilde{\Delta}_\varepsilon(r)+\delta}\frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r)) + \frac{1}{2}\frac{B_\varepsilon^2(r)}{\varepsilon^{\eta/2}}\ln(\tilde{\Delta}_\varepsilon(r)+\delta)\ddot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))$$

and where M is a martingale. By dominated convergence,

$$\mathbb{E}\left[\int_{r=0}^{t\wedge\tau_K^\varepsilon} u_{1,\varepsilon}(r)\mathrm{d}r\right] = \lim_{\delta\searrow 0}\mathbb{E}\left[\int_{r=0}^{t\wedge\tau_K^\varepsilon} u_{1,\varepsilon}^\delta(r)\mathrm{d}r\right].$$

Keeping in mind that $(x_\circ, \theta_\circ, y_0, \psi_0) \in \mathbb{R}^4$ is fixed, there is a $\mathsf{K} > 0$ such that

$$|U_0^{\delta,\varepsilon}| \le \mathsf{K}\varepsilon^{2/3}\ell(\varepsilon) \quad \text{and} \quad U_{t\wedge\tau_K^\varepsilon}^{\delta,\varepsilon} \le \mathsf{K}\varepsilon^{2/3}\ell(\varepsilon)$$

for all δ and ε in $(0, 1)$. Then $|\ln(\tilde{\Delta}_\varepsilon(r)+\delta)| \le \mathsf{K}\ell(\varepsilon)$ for all ε and δ in $(0, 1)$ and all $r \in [0, \tau_K^\varepsilon]$ such that $\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))$ or $\ddot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))$ is nonzero. Thus

$$|\mathscr{E}_{\delta,\varepsilon}(r)| \le \mathsf{K}\ell(\varepsilon)\chi_{[1,2]}(|\tilde{\Delta}'_\varepsilon(r)|)$$

for all ε and δ in $(0, 1)$ and $r \in [0, \tau_K^\varepsilon]$. If $|\tilde{\Delta}'_\varepsilon(r)| \in [1, 2]$, then[5] $|X_r^\varepsilon - Y_r^\varepsilon| \le \sqrt{2}\varepsilon^{1/3+\eta/8} \le \varrho\varepsilon^\nu$. Hence

$$\chi_{[1,2]}(|\tilde{\Delta}'_\varepsilon(r)|) \le \chi_{[-\varrho,\varrho]}\left(\frac{\mathsf{b}_r^\varepsilon}{\varepsilon^\nu}\right)\chi_{[1,\infty)}\left(\frac{\tilde{\Delta}_\varepsilon(r)}{\varepsilon^{\eta/4}}\right) \le I_1^\varepsilon(r) + I_3^\varepsilon(r)$$

for all $\varepsilon \in (0, 1)$ and $r \in [0, \tau_K^\varepsilon]$. Combining our calculations, we get the desired result. Note that we do not need a lower bound on $U_{t\wedge\tau_K^\varepsilon}^{\delta,\varepsilon}$. □

We now want to start replacing the term in braces in (14) by simpler expressions. First, we replace differences in (13) by (first-order) Taylor approximations; this basically allows us to

[5] We use here that $1/3 + \eta/8 > \nu$.

consider the tangent flow, which was treated in [9]. Define

$$\check{A}_{2,\varepsilon}(r) \stackrel{\text{def}}{=} 2\dot{\omega}(X^\varepsilon_r)\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)\left(\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3}}\right) + (\dot{\sigma}(\theta^\varepsilon_r))^2\mathbf{s}^2(\theta^\varepsilon_r - \psi^\varepsilon_r)$$

$$\check{B}_{2,\varepsilon}(r) \stackrel{\text{def}}{=} 2\dot{\sigma}(\theta^\varepsilon_r)\left(\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3}}\right)\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)$$

$$u_{2,\varepsilon}(r) \stackrel{\text{def}}{=} \left\{\frac{\check{A}_{2,\varepsilon}(r)}{\tilde{\Delta}_\varepsilon(r)} - \frac{1}{2}\frac{\check{B}^2_{2,\varepsilon}(r)}{\tilde{\Delta}^2_\varepsilon(r)}\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

for all $\varepsilon \in (0, 1)$ and $r \geq 0$.

Lemma 4.8. *For each $L > 0$, we have that*

$$\varlimsup_{\varepsilon \searrow 0} \mathbb{E}\left[\int_{r=0}^{t\wedge\tau^\varepsilon_K} |u_{1,\varepsilon}(r) - u_{2,\varepsilon}(r)|\mathrm{d}r\right] \leq 0.$$

Proof. Note that

$$|A_\varepsilon(r) - \check{A}_{2,\varepsilon}(r)| \leq \mathsf{K}\tilde{\Delta}^{3/2}_\varepsilon(r) \quad \text{and} \quad |B_\varepsilon(r) - \check{B}_{2,\varepsilon}(r)| \leq \mathsf{K}\tilde{\Delta}^{3/2}_\varepsilon(r).$$

Thus

$$|u_{1,\varepsilon}(r) - u_{2,\varepsilon}(r)| \leq \mathsf{K}\left\{\frac{\tilde{\Delta}^{3/2}_\varepsilon(r)}{\tilde{\Delta}_\varepsilon(r)} + \frac{\tilde{\Delta}^{5/2}_\varepsilon(r)}{\tilde{\Delta}^2_\varepsilon(r)}\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r)) \leq \mathsf{K}\varepsilon^{\eta/8}. \quad \square$$

We can now sequentially treat the different time scales. We first average the fast variable, i.e., θ^ε (the microscopic time scale); this simplifies the $\dot{\sigma}(\theta^\varepsilon)$ term in $u_{2,\varepsilon}$. Set

$$\check{A}_{3,\varepsilon}(r) \stackrel{\text{def}}{=} 2\dot{\omega}(X^\varepsilon_r)\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)\left(\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3}}\right) + \kappa_1^2\mathbf{s}^2(\theta^\varepsilon_r - \psi^\varepsilon_r)$$

$$\check{B}_{3,\varepsilon}(r) \stackrel{\text{def}}{=} 2\kappa_1\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)\left(\frac{X^\varepsilon_r - Y^\varepsilon_r}{\varepsilon^{1/3}}\right)$$

$$u_{3,\varepsilon}(r) \stackrel{\text{def}}{=} \left\{\frac{\check{A}_{3,\varepsilon}(r)}{\tilde{\Delta}_\varepsilon(r)} - \frac{1}{2}\frac{\check{B}^2_{3,\varepsilon}(r)}{\tilde{\Delta}^2_\varepsilon(r)}\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r)).$$

We will need the following averaging corrector. For $\lambda \in (0, 1)$ and $(x, \theta) \in \mathbb{R}^2$, set

$$H^\lambda_{\kappa_1}(x, \theta) \stackrel{\text{def}}{=} \int_{s=0}^\infty \mathrm{e}^{-\lambda s}(\dot{\sigma}(\theta + \omega(x)s))^2\mathrm{d}s.$$

Lemma 4.9. *For each $\lambda \in (0, 1)$, $H^\lambda_{\kappa_1} \in C^\infty(\mathbb{S}^1 \times \mathbb{R})$. Secondly, $\sup_{\lambda\in(0,1)} \lambda\|H^\lambda_{\kappa_1}\|_{C^2(\mathbb{R}^2)} < \infty$. Thirdly,*

$$\omega(x)\frac{\partial H^\lambda_{\kappa_1}}{\partial\theta}(x, \theta) = \lambda H^\lambda_{\kappa_1}(x, \theta) - (\dot{\sigma}(\theta))^2$$

for all $(x, \theta) \in \mathbb{R}^2$. Finally, there is a constant $\mathsf{K} > 0$ such that

$$\left|\lambda H^\lambda(x, \theta) - {\kappa_1}^2\right| \leq \mathsf{K}\lambda$$

for all $(x, \theta) \in \mathbb{R}^2$ and $\lambda \in (0, 1)$.

Proof. The regularity, bounds on derivatives, and PDE follow from direct calculations [10]. To prove the final bound, define

$$\Phi(\theta) \overset{\text{def}}{=} \int_{r=0}^{\theta} \left\{ (\dot\sigma(r))^2 - \kappa_1{}^2 \right\} \mathrm{d}r;$$

for all $\theta \in \mathbb{R}$. Then we can integrate by parts to see that

$$\begin{aligned}
\lambda H^{\lambda}_{\kappa_1}(x,\theta) - \kappa_1{}^2 &= \lambda \int_{r=0}^{\infty} \mathrm{e}^{-\lambda r}\, \dot\Phi(\theta + \omega(x) r)\mathrm{d}r \\
&= -\lambda \Phi(\theta) + \lambda^2 \int_{r=0}^{\infty} \mathrm{e}^{-\lambda r}\, \Phi(\theta + \omega(x) r)\mathrm{d}r \\
&= -\lambda \Phi(\theta) + \lambda \int_{r=0}^{\infty} \mathrm{e}^{-r}\, \Phi\left(\theta + \omega(x)\frac{r}{\lambda}\right)\mathrm{d}r.
\end{aligned}$$

This implies the stated result after observing that Φ is bounded. □

We can now show that $u_{2,\varepsilon}$ and $u_{3,\varepsilon}$ are close.

Lemma 4.10. *For each $L > 0$,*

$$\lim_{\varepsilon \searrow 0} \mathbb{E}\left[\left| \int_{r=0}^{t\wedge\tau^{\varepsilon}_K} \left\{ u_{2,\varepsilon}(r) - u_{3,\varepsilon}(r) \right\} \mathrm{d}r \right| \right] = 0.$$

Proof. Set $\lambda_\varepsilon \overset{\text{def}}{=} \varepsilon^{1/3}$. Define

$$q^{\varepsilon}_r \overset{\text{def}}{=} \left(\frac{X^{\varepsilon}_r - Y^{\varepsilon}_r}{\varepsilon^{1/3}} \right)^2, \qquad Q^{\varepsilon}_r \overset{\text{def}}{=} \frac{1}{\tilde\Delta_\varepsilon(r)} - 2\frac{q^{\varepsilon}_r}{\tilde\Delta^2_\varepsilon(r)}$$

$$\begin{aligned}
U^{\varepsilon}_r &\overset{\text{def}}{=} \varepsilon H^{\lambda_\varepsilon}_{\kappa_1}(X^{\varepsilon}_r, \theta^{\varepsilon}_r)\mathbf{s}^2(\theta^{\varepsilon}_r - \psi^{\varepsilon}_r) Q^{\varepsilon}_r \varphi_0(\tilde\Delta'_\varepsilon(r)) \\
&= \varepsilon H^{\lambda_\varepsilon}_{\kappa_1}(X^{\varepsilon}_r, \theta^{\varepsilon}_r)\frac{\mathbf{s}^2(\theta^{\varepsilon}_r - \psi^{\varepsilon}_r)}{\tilde\Delta_\varepsilon(r)} \left\{ 1 - 2\frac{q^{\varepsilon}_r}{\tilde\Delta_\varepsilon(r)} \right\} \varphi_0(\tilde\Delta'_\varepsilon(r)).
\end{aligned}$$

We note that informally

$$\mathrm{d}q^{\varepsilon}_r = \varepsilon^{-1/3} B_\varepsilon(r)\mathrm{d}W_r + \varepsilon^{-2/3}\, \hbar_b(\mathsf{C}^{\varepsilon}_r)\mathrm{d}r$$

$$\begin{aligned}
\mathrm{d}Q^{\varepsilon}_r &= \varepsilon^{-2/3} \left\{ -\frac{A_\varepsilon(r) + 2\,\hbar_b(\mathsf{C}^{\varepsilon}_r)}{\tilde\Delta^2_\varepsilon(r)} + \frac{4q^{\varepsilon}_r A_\varepsilon(r) + 5B^2_\varepsilon(r)}{\tilde\Delta^3_\varepsilon(r)} - \frac{6q^{\varepsilon}_r B^2_\varepsilon(r)}{\tilde\Delta^4_\varepsilon(r)} \right\} \mathrm{d}r \\
&\quad + \varepsilon^{-1/3} \left\{ \frac{-3B_\varepsilon(r)}{\tilde\Delta^2_\varepsilon(r)} + \frac{4q^{\varepsilon}_r B_\varepsilon(r)}{\tilde\Delta^3_\varepsilon(r)} \right\} \mathrm{d}W_r.
\end{aligned}$$

By Ito's rule,

$$\begin{aligned}\int_{r=0}^{t\wedge\tau_K^\varepsilon}\{u_{3,\varepsilon}(r)-u_{2,\varepsilon}(r)\}\mathrm{d}r &= U^\varepsilon_{t\wedge\tau_K^\varepsilon}-U_0^\varepsilon-\int_{r=0}^{t\wedge\tau_K^\varepsilon}\mathscr{E}_{1,\varepsilon}(r)\mathrm{d}r\\
&\quad-\frac{\varepsilon}{2}\int_{r=0}^{t\wedge\tau_K^\varepsilon}\frac{\partial^2 H^{\lambda_\varepsilon}_{\kappa_1}}{\partial x^2}(X_r^\varepsilon,\theta_r^\varepsilon)\sigma^2(\theta_r^\varepsilon)\mathscr{E}_{2,\varepsilon}(r)\mathrm{d}r\\
&\quad-\varepsilon^{2/3}\sum_{j=3}^{4}\int_{r=0}^{t\wedge\tau_K^\varepsilon}\frac{\partial H^{\lambda_\varepsilon}_{\kappa_1}}{\partial x}(X_r^\varepsilon,\theta_r^\varepsilon)\sigma(\theta_r^\varepsilon)\mathscr{E}_{2,\varepsilon}(r)\mathrm{d}r\\
&\quad-\varepsilon^{1/3}\sum_{j=5}^{8}\int_{r=0}^{t\wedge\tau_K^\varepsilon}H^{\lambda_\varepsilon}_{\kappa_1}(X_r^\varepsilon,\theta_r^\varepsilon)\mathscr{E}_{j,\varepsilon}(r)\mathrm{d}r\\
&\quad-\varepsilon\int_{r=0}^{t\wedge\tau_K^\varepsilon}\frac{\partial H^{\lambda_\varepsilon}_{\kappa_1}}{\partial x}(X_r^\varepsilon,\theta_r^\varepsilon)\sigma(\theta_r^\varepsilon)\mathscr{E}_{2,\varepsilon}(r)\mathrm{d}W_r\\
&\quad-\varepsilon^{2/3}\sum_{j=9}^{10}\int_{r=0}^{t\wedge\tau_K^\varepsilon}H^{\lambda_\varepsilon}_{\kappa_1}(X_r^\varepsilon,\theta_r^\varepsilon)\mathscr{E}_{j,\varepsilon}(r)\mathrm{d}W_r\end{aligned}$$

where

$$\begin{aligned}
\mathscr{E}_{1,\varepsilon}(r)&\overset{\text{def}}{=}\left\{\lambda_\varepsilon H^{\lambda_\varepsilon}_{\kappa_1}(X_r^\varepsilon,\theta_r^\varepsilon)-\kappa_1^2\right\}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\left\{1-2\frac{q_r^\varepsilon}{\tilde\Delta_\varepsilon(r)}\right\}\varphi_0(\tilde\Delta'_\varepsilon(r))\\
\mathscr{E}_{2,\varepsilon}(r)&\overset{\text{def}}{=}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\left\{1-2\frac{q_r^\varepsilon}{\tilde\Delta_\varepsilon(r)}\right\}\varphi_0(\tilde\Delta'_\varepsilon(r))\\
\mathscr{E}_{3,\varepsilon}(r)&\overset{\text{def}}{=}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\left\{\frac{-3B_\varepsilon(r)}{\tilde\Delta_\varepsilon(r)}+\frac{4q_r^\varepsilon B_\varepsilon(r)}{\tilde\Delta^2_\varepsilon(r)}\right\}\varphi_0(\tilde\Delta'_\varepsilon(r))\\
\mathscr{E}_{4,\varepsilon}(r)&\overset{\text{def}}{=}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\left\{1-2\frac{q_r^\varepsilon}{\tilde\Delta_\varepsilon(r)}\right\}\dot\varphi_0(\tilde\Delta'_\varepsilon(r))\frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}\\
\mathscr{E}_{5,\varepsilon}(r)&\overset{\text{def}}{=}2\frac{1}{\tilde\Delta_\varepsilon(r)}\left(\mathbf{s}(\theta_r^\varepsilon-\psi_r^\varepsilon)\left\{\frac{\omega(X_r^\varepsilon)-\omega(Y_r^\varepsilon)}{\varepsilon^{1/3}}\right\}\right)\left\{1-2\frac{q_r^\varepsilon}{\tilde\Delta_\varepsilon(r)}\right\}\varphi_0(\tilde\Delta'_\varepsilon(r))\\
\mathscr{E}_{6,\varepsilon}(r)&\overset{\text{def}}{=}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\\
&\quad\times\left\{-\frac{A_\varepsilon(r)+2\,\hbar_b(\mathsf{C}_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}+\frac{4q_r^\varepsilon A_\varepsilon(r)+5B^2_\varepsilon(r)}{\tilde\Delta^2_\varepsilon(r)}-\frac{6q_r^\varepsilon B^2_\varepsilon(r)}{\tilde\Delta^3_\varepsilon(r)}\right\}\varphi_0(\tilde\Delta'_\varepsilon(r))\\
\mathscr{E}_{7,\varepsilon}(r)&\overset{\text{def}}{=}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\left\{1-2\frac{q_r^\varepsilon}{\tilde\Delta^2_\varepsilon(r)}\right\}\left\{\dot\varphi_0(\tilde\Delta'_\varepsilon(r))\frac{A_{1,\varepsilon}(r)}{\varepsilon^{\eta/4}}+\frac{1}{2}\ddot\varphi_0(\tilde\Delta'_\varepsilon(r))\frac{B^2_\varepsilon(r)}{\varepsilon^{\eta/2}}\right\}\\
\mathscr{E}_{8,\varepsilon}(r)&\overset{\text{def}}{=}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\left\{\frac{-3B_\varepsilon(r)}{\tilde\Delta_\varepsilon(r)}+\frac{4q_r^\varepsilon B_\varepsilon(r)}{\tilde\Delta^2_\varepsilon(r)}\right\}\dot\varphi_0(\tilde\Delta'_\varepsilon(r))\frac{B_{1,\varepsilon}(r)}{\varepsilon^{\eta/4}}\\
\mathscr{E}_{9,\varepsilon}(r)&\overset{\text{def}}{=}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\left\{\frac{-3B_\varepsilon(r)}{\tilde\Delta_\varepsilon(r)}+\frac{4q_r^\varepsilon B_\varepsilon(r)}{\tilde\Delta^2_\varepsilon(r)}\right\}\varphi_0(\tilde\Delta'_\varepsilon(r))\\
\mathscr{E}_{10,\varepsilon}(r)&\overset{\text{def}}{=}\frac{\mathbf{s}^2(\theta_r^\varepsilon-\psi_r^\varepsilon)}{\tilde\Delta_\varepsilon(r)}\left\{1-2\frac{q_r^\varepsilon}{\tilde\Delta_\varepsilon(r)}\right\}\dot\varphi_0(\tilde\Delta'_\varepsilon(r))\frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}.
\end{aligned}$$

Using Lemma 4.9, we have that $|\mathscr{E}_{1,\varepsilon}(r)| \leq \mathsf{K}\lambda_\varepsilon = \mathsf{K}\varepsilon^{1/6}$ and $\varepsilon^{1/3}|\mathscr{E}_{j,\varepsilon}(r)| \leq \mathsf{K}\varepsilon^{1/3}/\lambda_\varepsilon = \mathsf{K}\varepsilon^{1/6}$ for $j \in \{2, 3 \ldots 10\}$. We similarly have that $|U^\varepsilon_r| \leq \mathsf{K}\varepsilon/\lambda_\varepsilon = \mathsf{K}\varepsilon^{5/6}$. $\square$

We now want to use the mesoscopic time scale suggested by (12) and average out Λ^ε. This will take some work as (12) is only an approximate description of the dynamics of Λ^ε. We first rewrite some things. Define $G \in C(\mathbb{R}^\star \times \mathbb{R})$ as

$$G(\lambda, x) = \begin{cases} \dfrac{2\dot\omega(x)\lambda + \kappa_1^2}{1+\lambda^2} - \dfrac{2\kappa_1^2\lambda^2}{(1+\lambda^2)^2} & \text{if } (\lambda, x) \in \mathbb{R}\times\mathbb{R} \\ 0 & \text{if } (\lambda, x) \in \{\star\}\times\mathbb{R}. \end{cases}$$

Then

$$u_{3,\varepsilon}(r) = G(\Lambda^\varepsilon_r, X^\varepsilon_r)\varphi_0(\tilde\Delta'_\varepsilon(r)). \tag{15}$$

We want to replace the dependence on Λ^ε (the mesoscopic variable) with an effective constant. We first define an averaging measure. We recall $\mathcal{J}_\circ$ of (2) and κ_2 of (4). Clearly

$$\lim_{|\lambda|\to\infty} \lambda^2 \mathcal{J}_\circ(\lambda) = \frac{1}{3}; \tag{16}$$

thus $\mathcal{J}_\circ \in L^1(\mathbb{R})$. Next define

$$(\mathbf{A}G)(x) \overset{\text{def}}{=} \frac{\int_{\lambda\in\mathbb{R}} G(\lambda, x)\mathcal{J}_\circ(\kappa_2(x)\lambda)\mathrm{d}\lambda}{\int_{\lambda\in\mathbb{R}} \mathcal{J}_\circ(\kappa_2(x)\lambda)\mathrm{d}\lambda} = \frac{\int_{\lambda\in\mathbb{R}} g\left(\frac{\lambda}{\kappa_2(x)}, x\right)\mathcal{J}_\circ(\lambda)\mathrm{d}\lambda}{\int_{\lambda\in\mathbb{R}} \mathcal{J}_\circ(\lambda)\mathrm{d}\lambda}. \tag{17}$$

Let's now construct the corrector needed to replace G by $\mathbf{A}G$.

Lemma 4.11. *There is a* $\Upsilon \in C(\mathbb{R}^\star \times \mathbb{R}) \cap C^2(\mathbb{R}\times\mathbb{R})$ *such that*

$$\begin{aligned} &\frac{\kappa_1^2}{2}\frac{\partial^2 \Upsilon}{\partial\lambda^2}(\lambda, x) - \dot\omega(x)\lambda^2\frac{\partial\Upsilon}{\partial\lambda}(\lambda, x) = G(\lambda, x) - (\mathbf{A}G)(x) \quad (\lambda, x) \in \mathbb{R}\times\mathbb{R} \\ &\lim_{|\lambda|\to\infty} \Upsilon(\lambda, x) = 0 \quad x \in \mathbb{R}. \end{aligned}$$

We also have that

$$\begin{aligned} &\lim_{|\lambda|\to\infty} \lambda^2 \frac{\partial\Upsilon}{\partial\lambda}(\lambda, x) = \frac{2}{3}\frac{(\mathbf{A}G)(x)}{\kappa_1^2\kappa_2^3(x)} \\ &\lim_{|\lambda|\to\infty} \frac{\partial^2\Upsilon}{\partial\lambda^2}(\lambda, x) = 0 \end{aligned}$$

for all $x \in \mathbb{R}$. *Finally,*

$$\sup_{\substack{i,j\in\{0,1,2\}\\ i+j\leq 2\\ \lambda\in\mathbb{R}\\ x\in K}} (1+|\lambda|)^{i+1}\left|\frac{\partial^{i+j}\Upsilon}{\partial\lambda^i\partial x^j}(\lambda, x)\right| < \infty \tag{18}$$

for each $K \subset\subset \mathbb{R}$.

Proof. Define

$$T(\lambda, x) \stackrel{\text{def}}{=} \frac{2}{\kappa_1^2 \kappa_2^2(x)} \left\{ g\left(\frac{\lambda}{\kappa_2(x)}, x \right) - (\mathbf{A}g)(x) \right\}$$
$$\Upsilon(\lambda, x) \stackrel{\text{def}}{=} - \int_{\varsigma=-\infty}^{\kappa_2(x)\lambda} e^{\varsigma^3} \left\{ \int_{\nu=\varsigma}^{\infty} T(\nu, x) e^{-\nu^3} d\nu \right\} d\varsigma$$

for all $(\lambda, x) \in \mathbb{R} \times \mathbb{R}$. For all $x \in \mathbb{R}$, $g(\cdot, x) \in C(\mathbb{R}^\star)$ so $\bar{g}(x) \stackrel{\text{def}}{=} \sup_{\lambda \in \mathbb{R}} |T(\lambda, x)|$ is finite. Hence

$$\left| e^{\lambda^3} \int_{\varsigma=\lambda}^{\infty} T(\varsigma, x) e^{-\varsigma^3} d\varsigma \right| \leq \bar{g}(x) \mathcal{J}_\circ(-\lambda)$$

for all $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}$, so by (16), we know that Υ is well-defined and also that $\lim_{\lambda \to -\infty} \Upsilon(\lambda, x) = 0$ for each $x \in \mathbb{R}$. We also note that

$$\begin{aligned}
\lim_{\lambda \to \infty} \Upsilon(\lambda, x) &= \int_{\varsigma=-\infty}^{\infty} e^{\varsigma^3} \left\{ \int_{\nu=\varsigma}^{\infty} T(\nu, x) e^{-\nu^3} d\nu \right\} d\varsigma \\
&= \int_{\nu=-\infty}^{\infty} T(\nu, x) e^{-\nu^3} \left\{ \int_{\varsigma=-\infty}^{\nu} e^{\varsigma^3} d\varsigma \right\} dr \\
&= \int_{r=-\infty}^{\infty} T(r, x) \mathcal{J}_\circ(r) dr = \frac{2}{\kappa_1^2 \kappa_2^2(x)} \\
&\quad \times \left\{ \int_{r=-\infty}^{\infty} g\left(\frac{r}{\kappa_2(x)}, x \right) \mathcal{J}_\circ(r) dr - (\mathbf{A}g)(x) \int_{r=-\infty}^{\infty} \mathcal{J}_\circ(r) dr \right\} = 0.
\end{aligned}$$

Differentiating the formula for Υ, we get that

$$\frac{\partial \Upsilon}{\partial \lambda}(\lambda, x) = -\kappa_2(x) \exp\left[\kappa_2^3(x) \lambda^3 \right] \int_{\varsigma=\kappa_2(x)\lambda}^{\infty} T(\varsigma, x) e^{-\varsigma^3} d\varsigma$$
$$\frac{\partial^2 \Upsilon}{\partial \lambda^2}(\lambda, x) = 3\kappa_2^3(x) \lambda^2 \frac{\partial \Upsilon}{\partial \lambda}(\lambda, x) + \kappa_2^2(x) T(\kappa_2(x)\lambda, x)$$

for all $(\lambda, x) \in \mathbb{R} \times \mathbb{R}$. The second formula is equivalent to the differential equation for Υ. We next compute that

We can rewrite the first formula as

$$\begin{aligned}
\frac{\partial \Upsilon}{\partial \lambda}(\lambda, x) &= -\frac{\kappa_2(x)}{3} \int_{\varsigma=0}^{\infty} \frac{T\left((\varsigma + (\kappa_2(x)\lambda)^3)^{1/3}, x \right)}{\left\{ \varsigma + (\kappa_2(x)\lambda)^3 \right\}^{2/3}} e^{-\varsigma} d\varsigma \\
&= -\frac{2}{3\kappa_1^2 \kappa_2(x)} \int_{\varsigma=0}^{\infty} \frac{g\left(\frac{(\varsigma + (\kappa_2(x)\lambda)^3)^{1/3}}{\kappa_2(x)}, x \right) - (\mathbf{A}g)(x)}{\left\{ \varsigma + (\kappa_2(x)\lambda)^3 \right\}^{2/3}} e^{-\varsigma} d\varsigma.
\end{aligned}$$

The claimed limit of $\lambda^2 \frac{\partial \Upsilon}{\partial \lambda}(\lambda, x)$ follows, as do the claimed bounds on $\partial^{i+j} \Upsilon / \partial \lambda^i \partial x^j$ for $i \geq 1$. The bound for $i = 0$ follows by integrating $\partial^{1+j} \Upsilon / \partial \lambda \partial x^j$. $\square$

This should help us to average Λ^ε out of (15); of course in doing so, we still have the fact that (13) is only approximate. This will force us to take several extra steps. Set

$$\alpha_\varepsilon(r) \stackrel{\text{def}}{=} \int_{\gamma=0}^1 \dot{\omega}(Y_r^\varepsilon + \gamma(X_r^\varepsilon - Y_r^\varepsilon))\mathrm{d}\gamma$$
$$\beta_\varepsilon(r) \stackrel{\text{def}}{=} \int_{\gamma=0}^1 \dot{\sigma}(Y_r^\varepsilon + \gamma(X_r^\varepsilon - Y_r^\varepsilon))\mathrm{d}\gamma.$$

Define now

$$v_{1,\varepsilon}(r) \stackrel{\text{def}}{=} \left\{-\alpha_\varepsilon(r)(\Lambda_r^\varepsilon)^2 \frac{\partial \Upsilon}{\partial \lambda}(\Lambda_r^\varepsilon, X_r^\varepsilon) + \frac{1}{2}\beta_\varepsilon^2(r)\frac{\partial^2 \Upsilon}{\partial \lambda^2}(\Lambda_r^\varepsilon, X_r^\varepsilon)\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

if $\Lambda_r^\varepsilon \in \mathbb{R}$, and set

$$v_{1,\varepsilon}(r) \stackrel{\text{def}}{=} -\alpha_\varepsilon(r)\frac{2}{3}\frac{(\mathbf{A}G)(X_r^\varepsilon)}{\kappa_1^2\kappa_2^3(X_r^\varepsilon)}\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

if $\Lambda_r^\varepsilon = \star$ (this case can also be defined by continuity). The point of $v_{1,\varepsilon}$ is that it is approximately $\{G - (\mathbf{A}G)\}\varphi_\circ(\tilde{\Delta}'_\varepsilon)$, where $\mathbf{A}G$ has averaged out the Λ^ε term. This allows us to proceed to the next step, but we use the estimates of Sections 4.2 and 4.3 to bound some of the errors.

Lemma 4.12. *For each $L > 0$,*

$$\lim_{\varepsilon\searrow 0} \mathbb{E}\left[\left|\int_{r=0}^{t\wedge\tau_K^\varepsilon} v_{1,\varepsilon}(r)\mathrm{d}r\right|\right] \le \mathsf{K}\mathbb{E}\left[\int_{r=0}^{t\wedge\tau_K^\varepsilon}\left\{I_1^\varepsilon(r) + I_3^\varepsilon(r)\right\}\mathrm{d}r\right] + \mathsf{K}\varepsilon^{1/3}.$$

Proof. Set

$$U_r^\varepsilon \stackrel{\text{def}}{=} \varepsilon^{2/3}\Upsilon(\Lambda_r^\varepsilon, X_r^\varepsilon)\varphi_0(\tilde{\Delta}'_\varepsilon(r)).$$

Then

$$\int_{r=0}^{t\wedge\tau_K^\varepsilon} v_{1,\varepsilon}(r)\mathrm{d}r = U^\varepsilon_{t\wedge\tau_K^\varepsilon} - U_0^\varepsilon - \sum_{j=1}^5\int_{r=0}^{t\wedge\tau_K^\varepsilon}\mathscr{E}_{j,\varepsilon}(r)\mathrm{d}r - \int_{r=0}^{t\wedge\tau_K^\varepsilon}\mathscr{E}_{6,\varepsilon}(r)\mathrm{d}W_r$$

where

$$\mathscr{E}_{1,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{1}{2}\varepsilon^{2/3}\frac{\partial^2\Upsilon}{\partial x^2}(\Lambda_r^\varepsilon, X_r^\varepsilon)\sigma^2(\theta_r^\varepsilon)\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$
$$\mathscr{E}_{2,\varepsilon}(r) \stackrel{\text{def}}{=} \varepsilon^{1/3}\frac{\partial^2\Upsilon}{\partial\lambda\partial x}(\Lambda_r^\varepsilon, X_r^\varepsilon)\sigma(\theta_r^\varepsilon)\beta_\varepsilon(r)\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$
$$\mathscr{E}_{3,\varepsilon}(r) \stackrel{\text{def}}{=} \Upsilon(\Lambda_r^\varepsilon, X_r^\varepsilon)\left\{\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))\frac{A_\varepsilon(r)}{\varepsilon^{\eta/4}} + \frac{1}{2}\ddot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))\frac{B_\varepsilon(r)}{\varepsilon^{\eta/2}}\right\}$$
$$\mathscr{E}_{4,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\partial\Upsilon}{\partial\lambda}(\Lambda_r^\varepsilon, X_r^\varepsilon)\beta_\varepsilon(r)\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))\frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}$$
$$\mathscr{E}_{5,\varepsilon}(r) \stackrel{\text{def}}{=} \varepsilon^{1/3}\frac{\partial\Upsilon}{\partial x}(\Lambda_r^\varepsilon, X_r^\varepsilon)\sigma(\theta_r^\varepsilon)\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))\frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}$$

$$\begin{aligned}\mathscr{E}_{6,\varepsilon}(r) &\overset{\text{def}}{=} \varepsilon^{1/3}\frac{\partial \Upsilon}{\partial \lambda}(\Lambda^\varepsilon_r, X^\varepsilon_r)\beta_\varepsilon(r)\varphi_0(\tilde{\Delta}'_\varepsilon(r)) + \varepsilon^{2/3}\frac{\partial \Upsilon}{\partial x}(\Lambda^\varepsilon_r, X^\varepsilon_r)\sigma(\theta^\varepsilon_r)\varphi_0(\tilde{\Delta}'_\varepsilon(r))\\ &\quad + \varepsilon^{1/3}\Upsilon(\Lambda^\varepsilon_r, X^\varepsilon_r)\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))\frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}.\end{aligned}$$

Here we have defined, by continuity,

$$\frac{\partial^{i+j}\Upsilon}{\partial\lambda^i\partial x^j}(\star, x) \overset{\text{def}}{=} 0$$

for all relevant nonnegative integers i and j and all $x \in \mathbb{R}$. We easily see that $\mathscr{E}_{1,\varepsilon}$, $\mathscr{E}_{2,\varepsilon}$, $\mathscr{E}_{5,\varepsilon}$ and $\mathscr{E}_{6,\varepsilon}$ are all bounded by $\mathsf{K}\varepsilon^{1/3}$. As in the proof of Lemma 4.7, both $\mathscr{E}_{3,\varepsilon}(r)$ and $\mathscr{E}_{4,\varepsilon}(r)$ are bounded by

$$\mathsf{K}\chi_{[1,2]}(|\tilde{\Delta}'_\varepsilon(r)|) \le \mathsf{K}\chi_{[-\varrho,\varrho]}\left(\frac{\mathsf{b}^\varepsilon_r}{\varepsilon^\nu}\right)\chi_{[1,\infty)}\left(\frac{\tilde{\Delta}_\varepsilon(r)}{\varepsilon^{\eta/4}}\right) \le \mathsf{K}\{I^\varepsilon_1(r) + I^\varepsilon_3(r)\}.$$

The claimed result easily follows.[6] □

Next define

$$v_{2,\varepsilon}(r) \overset{\text{def}}{=} \left\{-\dot{\omega}(X^\varepsilon_r)(\Lambda^\varepsilon_r)^2\frac{\partial\Upsilon}{\partial\lambda}(\Lambda^\varepsilon_r, X^\varepsilon_r) + \frac{1}{2}(\dot{\sigma}(\theta^\varepsilon_r))^2\frac{\partial^2\Upsilon}{\partial\lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r)\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

if $\Lambda^\varepsilon_r \in \mathbb{R}$, and set

$$v_{2,\varepsilon}(r) \overset{\text{def}}{=} -\dot{\omega}(X^\varepsilon_r)\frac{2}{3}\frac{(\mathbf{A}G)(X^\varepsilon_r)}{\kappa_1^2\kappa_2^3(X^\varepsilon_r)}\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

if $\Lambda^\varepsilon_r = \star$. This is even closer to $\{G - (\mathbf{A}G)\}\varphi_\circ(\tilde{\Delta}'_\varepsilon)$; we have replaced α_ε and β_ε by $\dot{\omega}(X^\varepsilon_r)$ and $\dot{\sigma}(\theta^\varepsilon_r)$.

Lemma 4.13. *We have that*

$$\lim_{\varepsilon\searrow 0}\mathbb{E}\left[\left|\int_{r=0}^{t\wedge\tau^\varepsilon_K} v_{1,\varepsilon}(r) - v_{2,\varepsilon}(r)\mathrm{d}r\right|\right] = 0.$$

Proof. If $\varphi_0(\tilde{\Delta}'_\varepsilon(r)) > 0$, then

$$|\alpha_\varepsilon(r) - \dot{\omega}(X^\varepsilon_r)| \le \mathsf{K}|X^\varepsilon_r - Y^\varepsilon_r| \le \mathsf{K}\varepsilon^{1/3}\sqrt{\tilde{\Delta}_\varepsilon(r)} \le \mathsf{K}\varepsilon^{1/3+\eta/8}$$

$$|\beta_\varepsilon(r) - (\dot{\sigma}(\theta^\varepsilon_r))^2| \le \mathsf{K}|\mathbf{s}(\theta^\varepsilon_r - \psi^\varepsilon_r)| \le \mathsf{K}\sqrt{\tilde{\Delta}_\varepsilon(r)} \le \mathsf{K}\varepsilon^{\eta/8}.$$

This gives us the desired result. □

We have almost finished. We next replace $(\dot{\sigma}(\theta^\varepsilon_r))^2$ by its average; recall that θ^ε is the microscopic (fastest) variable. Define

$$\begin{aligned}v_{3,\varepsilon}(r) &\overset{\text{def}}{=} \left\{-\dot{\omega}(X^\varepsilon_r)(\Lambda^\varepsilon_r)^2\frac{\partial\Upsilon}{\partial\lambda}(\Lambda^\varepsilon_r, X^\varepsilon_r) + \frac{\kappa_1^2}{2}\frac{\partial^2\Upsilon}{\partial\lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r)\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r))\\ &= u_{3,\varepsilon}(r) - (\mathbf{A}G)(X^\varepsilon_r)\varphi_0(\tilde{\Delta}'_\varepsilon(r))\end{aligned}$$

[6] Here is where we most need the localization by τ^ε_K, so that we can take advantage of the uniform bounds of (18).

if $\Lambda^\varepsilon_r \in \mathbb{R}$, and set

$$v_{3,\varepsilon}(r) \stackrel{\text{def}}{=} -\dot{\omega}(X^\varepsilon_r)\frac{2}{3}\frac{(\mathbf{A}G)(X^\varepsilon_r)}{\kappa_1^2\kappa_2^3(X^\varepsilon_r)}\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

if $\Lambda^\varepsilon_r = \star$.

Lemma 4.14. *For each $L > 0$,*

$$\lim_{\varepsilon\searrow 0}\mathbb{E}\left[\left|\int_{r=0}^{t\wedge\tau^\varepsilon_K}\{v_{2,\varepsilon}(r) - v_{3,\varepsilon}(r)\}\mathrm{d}r\right|\right] = 0.$$

Proof. Set $\lambda_\varepsilon \stackrel{\text{def}}{=} \varepsilon^{1/6}$ and then set

$$U^\varepsilon_r \stackrel{\text{def}}{=} \frac{\varepsilon}{2}H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r)\frac{\partial^2\Upsilon}{\partial\lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r)\varphi_0(\tilde{\Delta}'_\varepsilon(r)).$$

Then

$$\int_{r=0}^{t\wedge\tau^\varepsilon_K}\{v_{3,\varepsilon}(r) - v_{2,\varepsilon}(r)\}\mathrm{d}r = U^\varepsilon_{t\wedge\tau^\varepsilon_K} - U^\varepsilon_0 - \sum_{j=1}^{11}\int_{r=0}^{t\wedge\tau^\varepsilon_K}\mathscr{E}_{j,\varepsilon}(r)\mathrm{d}r$$
$$- \sum_{j=12}^{15}\int_{r=0}^{t\wedge\tau^\varepsilon_K}\mathscr{E}_{j,\varepsilon}(r)\mathrm{d}W_r$$

where

$$\mathscr{E}_{1,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{1}{2}\left\{\lambda_\varepsilon H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r) - \kappa_1^2\right\}\frac{\partial^2\Upsilon}{\partial\lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r)\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

$$\mathscr{E}_{2,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon}{2}\frac{\partial^2 H^{\lambda_\varepsilon}_{\kappa_1}}{\partial x^2}(\theta^\varepsilon_r, X^\varepsilon_r)\sigma^2(\theta^\varepsilon_r)\frac{\partial^2\Upsilon}{\partial\lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r)\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

$$\mathscr{E}_{3,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{1/3}}{2}H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r)$$
$$\times\left\{-\alpha_\varepsilon(r)(\Lambda^\varepsilon_r)^2\frac{\partial^3\Upsilon}{\partial\lambda^3}(\Lambda^\varepsilon_r, X^\varepsilon_r) + \frac{1}{2}\beta^2_\varepsilon(r)\frac{\partial^4\Upsilon}{\partial\lambda^4}(\Lambda^\varepsilon_r, X^\varepsilon_r)\right\}\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

$$\mathscr{E}_{4,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon}{2}H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r)\frac{\partial^4\Upsilon}{\partial\lambda^2\partial x^2}(\Lambda^\varepsilon_r, X^\varepsilon_r)\sigma^2(\theta^\varepsilon_r)\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

$$\mathscr{E}_{5,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{1/3}}{2}H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r)\frac{\partial^2\Upsilon}{\partial\lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r)\left\{\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))\frac{A_\varepsilon(r)}{\varepsilon^{\eta/4}} + \frac{1}{2}\ddot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))\frac{B^2_\varepsilon(r)}{\varepsilon^{\eta/2}}\right\}$$

$$\mathscr{E}_{6,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon}{2}\frac{\partial H^{\lambda_\varepsilon}_{\kappa_1}}{\partial x}(\theta^\varepsilon_r, X^\varepsilon_r)\sigma^2(\theta^\varepsilon_r)\frac{\partial^3\Upsilon}{\partial\lambda^2\partial x}(\Lambda^\varepsilon_r, X^\varepsilon_r)\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

$$\mathscr{E}_{7,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{2/3}}{2}\frac{\partial H^{\lambda_\varepsilon}_{\kappa_1}}{\partial x}(\theta^\varepsilon_r, X^\varepsilon_r)\sigma(\theta^\varepsilon_r)\frac{\partial^3\Upsilon}{\partial\lambda^3}(\Lambda^\varepsilon_r, X^\varepsilon_r)\beta_\varepsilon(r)\varphi_0(\tilde{\Delta}'_\varepsilon(r))$$

$$\mathscr{E}_{8,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{2/3}}{2}\frac{\partial H^{\lambda_\varepsilon}_{\kappa_1}}{\partial x}(\theta^\varepsilon_r, X^\varepsilon_r)\sigma(\theta^\varepsilon_r)\frac{\partial^2\Upsilon}{\partial\lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r)\dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r))\frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}$$

$$\mathscr{E}_{9,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon}{2} H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r) \frac{\partial^4 \Upsilon}{\partial \lambda^3 \partial x}(\Lambda^\varepsilon_r, X^\varepsilon_r) \beta^\varepsilon(r) \sigma(\theta^\varepsilon_r) \varphi_0(\tilde{\Delta}'_\varepsilon(r))$$
$$\mathscr{E}_{10,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{1/3}}{2} H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r) \frac{\partial^3 \Upsilon}{\partial \lambda^3}(\Lambda^\varepsilon_r, X^\varepsilon_r) \beta_\varepsilon(r) \dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r)) \frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}$$
$$\mathscr{E}_{11,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{1/3}}{2} H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r) \frac{\partial^3 \Upsilon}{\partial \lambda^2 \partial x}(\Lambda^\varepsilon_r, X^\varepsilon_r) \sigma(\theta^\varepsilon_r) \dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r)) \frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}$$
$$\mathscr{E}_{12,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon}{2} \frac{\partial H^{\lambda_\varepsilon}_{\kappa_1}}{\partial x}(\theta^\varepsilon_r, X^\varepsilon_r) \sigma(\theta^\varepsilon_r) \frac{\partial^2 \Upsilon}{\partial \lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r) \varphi_0(\tilde{\Delta}'_\varepsilon(r))$$
$$\mathscr{E}_{13,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{2/3}}{2} H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r) \frac{\partial^3 \Upsilon}{\partial \lambda^3}(\Lambda^\varepsilon_r, X^\varepsilon_r) \beta_\varepsilon(r) \varphi_0(\tilde{\Delta}'_\varepsilon(r))$$
$$\mathscr{E}_{14,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon}{2} H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r) \frac{\partial^3 \Upsilon}{\partial \lambda^2 \partial x}(\Lambda^\varepsilon_r, X^\varepsilon_r) \sigma(\theta^\varepsilon_r) \varphi_0(\tilde{\Delta}'_\varepsilon(r))$$
$$\mathscr{E}_{15,\varepsilon}(r) \stackrel{\text{def}}{=} \frac{\varepsilon^{2/3}}{2} H^{\lambda_\varepsilon}_{\kappa_1}(\theta^\varepsilon_r, X^\varepsilon_r) \frac{\partial^2 \Upsilon}{\partial \lambda^2}(\Lambda^\varepsilon_r, X^\varepsilon_r) \dot{\varphi}_0(\tilde{\Delta}'_\varepsilon(r)) \frac{B_\varepsilon(r)}{\varepsilon^{\eta/4}}.$$

We have that $|\mathscr{E}_{1,\varepsilon}| \le \mathsf{K}\lambda_\varepsilon = \mathsf{K}\varepsilon^{1/6}$. All of the remaining $\mathscr{E}_{j,\varepsilon}$'s can be bounded by $\mathsf{K}\varepsilon^{1/3}/\lambda_\varepsilon = \mathsf{K}\varepsilon^{1/6}$. The claim then follows. □

We can now pull everything together and rigorously complete the mesoscopic averaging that we need to do. Recall that the goal of this subsection is finding a bound on I^ε_2 of (11). This allows us to finish the proof of Lemma 4.1.

Lemma 4.15. *We have that*

$$\lim_{\varepsilon \searrow 0} \mathbb{E}\left[\int_{r=0}^{t\wedge\tau^\varepsilon_K} I^\varepsilon_2(r)\mathrm{d}r\right] = 0.$$

Proof. We first write that

$$\int_{r=0}^{t\wedge\tau^\varepsilon_K} I^\varepsilon_2(r)\mathrm{d}r \le \int_{r=0}^{t\wedge\tau^\varepsilon_K} \varphi_0(\tilde{\Delta}'_\varepsilon(r))\mathrm{d}r.$$

We proceed by writing that

$$(\mathbf{A}G)(x) = \frac{\kappa_1^2 \kappa_2^2(x)}{\int_{\lambda\in\mathbb{R}} \mathcal{J}_\circ(\lambda)\mathrm{d}\lambda} \mathcal{J}(\kappa_2^2(x)).$$

Since $\mathcal{J}(\kappa_2^2(x)) > 0$ for all $x \in K$, we must have that $\upsilon_K \stackrel{\text{def}}{=} \inf_{x\in K}(\mathbf{A}G)(x) > 0$. We consequently can write that

$$\begin{aligned}\upsilon_K \varphi_0(\tilde{\Delta}'_\varepsilon(r)) &\le (\mathbf{A}G)(X^\varepsilon_r)\varphi_0(\tilde{\Delta}'_\varepsilon(r)) \le u_{3,\varepsilon}(r) - v_{3,\varepsilon}(r) \\ &= \{u_{3,\varepsilon}(r) - u_{2,\varepsilon}(r)\} + \{u_{2,\varepsilon}(r) - u_{1,\varepsilon}(r)\} + u_{1,\varepsilon}(r) \\ &\quad - \{v_{3,\varepsilon}(r) - v_{2,\varepsilon}(r)\} - \{v_{2,\varepsilon}(r) - v_{1,\varepsilon}(r)\} - v_{1,\varepsilon}(r).\end{aligned}$$

We combine our estimates and Lemmas 4.3 and 4.6 together to get that

$$\lim_{\varepsilon \searrow 0} \mathbb{E}\left[\int_{r=0}^{t\wedge\tau^\varepsilon_K} \varphi_0(\tilde{\Delta}'_\varepsilon(r))\mathrm{d}r\right] = 0. \quad \square$$

We can finally finish our proof of Lemma 4.1.

Proof of Lemma 4.1. Combine Lemmas 4.3, 4.6 and 4.15. □

References

[1] Peter Baxendale, Brownian motions in the diffeomorphism group. I, Compos. Math. 53 (1) (1984) 19–50.
[2] Peter H. Baxendale, Stochastic averaging and asymptotic behavior of the stochastic Duffing–van der Pol equation, Stochastic Process. Appl. 113 (2) (2004) 235–272.
[3] Peter H. Baxendale, Levon Goukasian, Lyapunov exponents of nilpotent Itô systems with random coefficients, Stochastic Process. Appl. 95 (2) (2001) 219–233.
[4] Peter H. Baxendale, Levon Goukasian, Lyapunov exponents for small random perturbations of Hamiltonian systems, Ann. Probab. 30 (1) (2002) 101–134.
[5] K.D. Elworthy, Stochastic Differential Equations on Manifolds, Cambridge University Press, Cambridge, 1982.
[6] M.I. Freidlin, A.D. Wentzell, Averaging principle for stochastic perturbations of multifrequency systems, Stoch. Dyn. 3 (3) (2003) 393–408.
[7] R.Z. Has'minskiĭ, Diffusion processes with a small parameter, Izv. Akad. Nauk SSSR Ser. Mat. 27 (1963) 1281–1300.
[8] H.L. Royden, Real Analysis, third ed., Macmillan Publishing Company, New York, 1988.
[9] Richard B. Sowers, On the tangent flow of a stochastic differential equation with fast drift, Trans. Amer. Math. Soc. 353 (4) (2001) 1321–1334 (electronic).
[10] Richard B. Sowers, Stochastic averaging near long heteroclinic orbits, Stoch. Dyn. 7 (2) (2007) 187–228.